

Substituting the End for the Whole: Why Voters Respond Primarily to the Election-Year Economy

Andrew Healy Loyola Marymount University
Gabriel S. Lenz University of California, Berkeley

According to numerous studies, the election-year economy influences presidential election results far more than cumulative growth throughout the term. Here we describe a series of surveys and experiments that point to an intriguing explanation for this pattern that runs contrary to standard political science explanations, but one that accords with a large psychological literature. Voters, we find, actually intend to judge presidents on cumulative growth. However, since that characteristic is not readily available to them, voters inadvertently substitute election-year performance because it is more easily accessible. This “end-heuristic” explanation for voters’ election-year emphasis reflects a general tendency for people to simplify retrospective assessments by substituting conditions at the end for the whole. The end-heuristic explanation also suggests a remedy, a way to align voters’ actions with their intentions. Providing people with the attribute they are seeking—cumulative growth—eliminates the election-year emphasis.

Voters, the data suggest, reward incumbents not broadly for economic growth throughout incumbents’ terms, but narrowly for conditions in the six months or year before Election Day (e.g., Achen and Bartels 2004; Alesina, Londregan, and Rosenthal 1993; Fair 1978; Kiewiet 1983; Kramer 1971). Surveying the evidence, Achen and Bartels conclude that “long-term economic growth, whether in real income or GDP, contributes little or nothing to the incumbent party’s electoral prospects” (2004, 19). In this article, we seek to understand this behavior, which creates at least three major problems for democratic accountability.

First, voter behavior incentivizes politicians to focus on election-year growth, with substantial consequences for public welfare. As Tufte wrote, it creates “a bias toward policies with immediate, highly visible benefits and deferred, hidden costs—myopic policies for myopic voters” (1978, 143). In perhaps the clearest U.S. example, the Nixon administration pressured Federal Reserve Chair-

man Arthur Burns to pursue expansionary monetary policy in late 1971 and early 1972, which helped precipitate high inflation in the years that followed (Abrams 2006; Greider 1989). Second, the election-year focus can lead to adverse selection. Instead of selecting the best economic leaders, democracies may select the best economic manipulators. Finally, the election-year focus may make voters less likely to select the best leaders simply by introducing noise, turning democracy into a game of “musical chairs” where elections are determined not so much by the quality of the candidates, but by the vagaries of the business cycle (Achen and Bartels 2004). Bartels (2008, 110) finds that if voters focused more on the overall economy during presidents’ terms, they would have elected a different president in three postwar elections (1952, 1968, and 2000).¹

So why do voters, who presumably would like to maximize economic welfare and elect capable leaders, respond primarily to the election-year economy? More generally,

Andrew Healy is Associate Professor, Loyola Marymount University, Department of Economics, 1 LMU Drive, University Hall 4229, Los Angeles, CA 90045 (ahealy@lmu.edu). Gabriel S. Lenz is Associate Professor, University of California, Berkeley, Charles and Louise Travers Department of Political Science, 210 Barrows #1950, Berkeley, CA 94720 (glenz@berkeley.edu).

We thank Larry Bartels, Adam Berinsky, Eitan Hersh, Justin Fox, Jens Hainmueller, Paco Flores-Macias, Aaron Kaufman, George Krause, Neil Malhotra, Mike Myers, Leif Nelson, Michael Peress, Kai Quek, Ken Shepsle, Joe Simmons, and Gabor Simonovits for helpful comments and suggestions, as well as seminar participants at Brown University, Colorado State University, Harvard University, MIT, Princeton University, UC Berkeley, UC Merced, University of Rochester, University of Wyoming, and Yale University. Replication data are available from the AJPS Data Archive on Dataverse (<http://dvn.iq.harvard.edu/dvn/dv/ajps>).

¹Voters’ election-year emphasis can also affect the entry decisions of challengers, who face greater uncertainty due to the unknown election-year economy. The election-year emphasis may also limit incentives to pursue strong growth earlier in a term if mean reversion in growth makes robust growth in the election year easier to achieve after relatively weak growth early in the term.

American Journal of Political Science, Vol. 58, No. 1, January 2014, Pp. 31–47

©2013, Midwest Political Science Association

DOI: 10.1111/ajps.12053

why do recency effects emerge not just with the economy but also in domains such as legislative appropriations, crime, and terrorist attacks?² To test the potential explanations, we consider evidence from a series of surveys and experiments.

We first attempt to answer a fundamental but unaddressed question relevant to all explanations: what are voters' intentions? Across a wide variety of elicitation methods, we find that voters intend to weigh the years of a president's term similarly. To determine why voters place much greater weight on the election-year economy despite their intentions to do otherwise, we conduct studies where we show participants economic conditions during presidents' terms. After viewing simple plots of either historical or hypothetical yearly personal-income growth during each term, participants evaluate overall economic conditions. We then examine how much influence (weight) each year has on their evaluations. Even though participants observe growth for all four years, their behavior closely resembles actual voter behavior. That is, despite intending to weigh income growth in all years roughly equally, they nevertheless overweigh the last year.

We first use these studies to test two standard explanations for why voters overweigh the end: that voters either simply fail to remember earlier conditions (e.g., Sarafidis 2007) or that they perceive the election year as more informative (e.g., MacKuen, Erikson, and Stimson 1992). Instead of supporting these explanations, however, our findings point to an intriguing alternative, which we call the "end-heuristic" explanation. While this explanation deviates from the ones offered by political science for voter behavior, it accords with a large body of evidence from psychology.³

According to this explanation, voters intend to judge presidents on overall performance rather than election-year conditions, but overall performance during the term is not readily available to them. Even if some candidates, such as Obama in the 2012 election, attempt to focus on the long term, the media focus mainly on recent economic

conditions.⁴ More generally, it takes cognitive effort to discern total performance over the term. With total performance not readily available, voters do what decision makers have been shown to do in many other environments where retrospective assessments are involved: they simplify a complicated evaluation problem by substituting end conditions for overall conditions (Ariely and Carmon 2000; Redelmeier and Kahneman 1996; Varey and Kahneman 1992). As Kahneman (2011) recently put it, "people who face a difficult question often answer an easier one instead, without realizing it" (Kahneman 2003; Kahneman and Frederick 2002). We argue that voters are doing something similar when confronted with the cognitively difficult task of summing conditions over an extended period, precisely the kind of situation where psychologists have found that people tend to rely on the end to represent the whole. Despite trying to evaluate overall performance over a term, our subjects appear to focus on the election-year economy because that attribute is an easily available substitute for the overall growth for which they are searching.

We then use a variation on our experimental design to test directly the end-heuristic explanation. If voters are indeed accidentally substituting election-year growth for cumulative growth, then curing this tendency—that is, helping them apply the weights they intend to apply—should be straightforward. As suggested by research in psychology (e.g., Liersch and McKenzie 2009), we should only have to make cumulative growth equally available. We find precisely this pattern. When we provide participants with information on cumulative performance (e.g., total income growth during incumbents' terms or income in levels), they no longer focus on the election year. Instead, they weigh all years roughly equally. Thus, while our results provide evidence of voters failing to act the way they intend with important implications for democratic accountability, they also point to a potential remedy for recency effects that government statistical agencies, the news media, and candidates could adopt.

Why the Election-Year Economy? Three Explanations

Political scientists have considered a variety of explanations for the weight voters place on the election-year

²For example, see Shepsle et al. (2009, 357), Muthoo and Shepsle (2010), and Levitt (1997). Several studies also find evidence of recency effects and the perverse incentives they create with natural disaster policy (Healy and Malhotra 2009) and terrorist attacks (Aksoy 2011; Bali 2007; Montalvo 2011).

³While the end-heuristic explanation has not been tested as an explanation for myopic economic voting, Bechtel and Hainmueller (2011) argue that usage of a "peak-and-end" heuristic can help to explain why German voters rewarded robust disaster response. In addition, a growing political science literature has considered how other heuristics affect voters (e.g., Berry and Howell 2007; Duch and Stevenson 2008; Duch and Tyran 2012; Kuklinski and Quirk 2000; Oliver and Ha 2007).

⁴We classified a random sample of broadcast media transcripts covering the U.S. economy in the months preceding the 2000, 2004, and 2008 presidential elections according to the time period they covered. Ninety-six percent of those articles covered economic conditions in the year preceding the election, at least in part, with much less coverage of earlier conditions. Please see the supporting information for complete details.

economy. An obvious explanation pertains to memory. Compared to their memories of a year ago, voters' memories of economic conditions in earlier years of the term may be faint or inaccurate (e.g., Sarafidis 2007). Even if they intend to answer Reagan's question ("Are you better off than you were four years ago?"), they may not be able to simply because they lack the necessary information about earlier economic conditions.

A second explanation is that voters may see the election-year economy as more informative about the quality of the president's economic stewardship. They may view the president's policies as taking months or years to influence the economy. Therefore, they may perceive economic performance at the end as particularly revealing about the effectiveness of those policies and so more predictive of future growth under that incumbent (e.g., Erikson, MacKuen, and Stimson 2000, 2002; MacKuen, Erikson, and Stimson 1992).

A third explanation, which has received considerably less attention from political scientists, stems from research in psychology. Psychologists have documented a pervasive human tendency to substitute the end for the whole when retrospectively assessing experiences. More recent work concludes that this behavior reflects an even more general tendency to substitute one attribute for another without realizing it, a tendency that Kahneman calls "attribute substitution" (Kahneman 2003; Kahneman and Frederick 2002). When people attempt to judge a target attribute, such as the change in their welfare under an incumbent president, they search for a reasonable value. For some judgments, this search finishes almost immediately because the required value is readily accessible (e.g., the question "How old are you?"). For many judgments, however, the target attribute does not readily come to mind, but the search for it evokes related attributes. People then substitute the related attribute for the target attribute, without realizing it. Regarding the economy, even if voters have a general idea of economic conditions in previous years (i.e., memory is not a problem), putting all the information together requires cognitive effort. As a result, voters may simplify the problem they intended to address ("How well did the economy perform during the president's term?") by substituting with an answer to a related question ("How has the economy been recently?").

Not only is attribute substitution pervasive, but it also appears to be particularly common with retrospective assessments like those involved with evaluating economic performance during a presidential term. These assessments are examples of what Kahneman (2003) calls "extensional attributes," defined as aggregated properties of the evaluation targets such as those that involve

summing over a period of time. Since these attributes are generally not readily available to individuals—they require calculations—people are especially prone to substituting other attributes such as the end. Indeed, people seem to evaluate entire experiences from colonoscopies to shopping according to the conditions at the peak and at the end, a phenomenon sometimes referred to as a "Peak-End Rule" (e.g., Fredrickson and Kahneman 1993; Redelmeier and Kahneman 1996). The peak part of the rule has figured less prominently in subsequent research than in the early studies of colonoscopies and aversive sounds—possibly because peak pain or discomfort is more salient—suggesting that the end part of the rule is more robust to other environments (e.g., Ariely 1998).

A series of experiments conducted by Kahneman et al. (1993) provides an example of this research. Participants began the studies by immersing one hand in very cold water for 10 seconds. The researchers then told participants to expect three more trials of this kind, but actually conducted only two. In the short trial, participants kept one hand in water at 14°C for 60 seconds. In the long trial, the immersion lasted 90 seconds. The experimenters kept the water at 14°C for the first 60 seconds, at which point they gradually raised the temperature from 14°C to 15°C over the next 30 seconds (unbeknownst to the participants). After a seven-minute delay, the researchers called participants for a third trial, informed them that they would repeat one of the two previous procedures, gave them a choice of whether the first or the second trial should be repeated, and asked them to answer several questions about the first two trials.

The results were striking. Even though participants experienced more total pain in the long trial, a large majority preferred to repeat it. The slight decrease in pain in the last 30 seconds of the long trial led participants to remember the entire trial as less painful than it actually was—they judged the experience on the end rather than the whole. A similar phenomenon occurs in many kinds of retrospective evaluations, including vacations (Loewenstein and Prelec 1993), emotional episodes (Fredrickson and Kahneman 1993; Varey and Kahneman 1992), TV advertisements (Baumgartner, Suja, and Padgett 1997), gambling (Ross and Simonson 1991), and academic performance (Zauberman, Diehl, and Ariely 2006).

These studies are also consistent with the finding from the retrospective voting literature that voters appear to vote as pocketbook voters on their own economic experiences (Kramer 1983; Lewis-Beck 1985) and as sociotropic observers of the national economy (Kinder and Kiewiet 1981). In the psychological studies, individuals

TABLE 1 Descriptions of Surveys and Experiments

Motivation	Description	N
A. Elicit intended weights	1. Survey about how much weight voters intend to place on the economy in each year of presidential terms	1602
	2. Experiment where we elicited weights without directly asking participants about weights	572
B. Test explanations based on memory and salience	1. Experiment showing participants income growth from the historical record for presidents	232
	2. Experiment showing participants hypothetical (randomly drawn) income growth for presidents	209
C. Test explanations based on the end potentially being more informative	1. Survey about how predictive each year of the term is for future income growth (same sample as B1)	232
	2. Experiment showing participants hypothetical income growth for four-year periods, but with no mention of politics, presidents, terms, etc.	81
	3. Experiment where participants saw either four years or four terms of income growth under governors (growth randomly drawn)	139
D. Test end-heuristic explanation by looking for ways to eliminate the election-year emphasis	1. Experiment where participants saw either yearly income growth or both cumulative and yearly growth for the same underlying data	116
	2. Experiment where participants saw either yearly growth or both income in levels and yearly growth for the same underlying data	187

Note: We conducted these studies between March 2010 and July 2012. Study A1 was a nationally representative sample (on demographics) collected by Survey Sampling International in May 2010. The other studies listed here use Mechanical Turk samples. We also conducted numerous replication studies and several pilot studies, some of which used student subjects. We report these in the supporting information.

appear to apply the end heuristic not only when they evaluate their own experiences, but also when they evaluate the experiences of others. In their study of colonoscopy patients, for example, Redelmeier and Kahneman (1996) found an end heuristic explained not only patients' evaluations of their colonoscopy experiences, but also the administering physicians' estimates of the pain that their patients experienced. Moreover, when asked which patients should have received more anesthetic, the doctors answered not according to total pain, but on pain at the peak and the end. Studies have found similar end bias when participants evaluated data on people's discomfort with unpleasant experiences such as listening to loud drilling noises (Varey and Kahneman 1992), when participants evaluated data on factories' production of defective products (reject rate), and when they evaluated tests of other individuals' intelligence (Zauberman, Diehl, and Ariely 2006). In short, observers appear to apply (or misapply) the same end heuristic in retrospective evaluations as those undergoing the experiences. They both substitute the end when evaluating the whole.

Summarizing the Surveys and Experiments

As described earlier, we use a series of surveys and experiments to disentangle competing explanations for why sociotropic voters put so much weight on the election-year economy. Table 1 presents a list of the studies that we show in the article. When presenting regression results and summary statistics below, we include the study identifier from Table 1 so that readers can easily trace results back to the sample that generated them by referring to the table. We also conducted more than a dozen additional studies that replicate and extend these findings; we present these in the supporting information.

Voters' Actual and Intended Behavior

Before examining the underlying reasons, we describe the extent to which voters assign extra weight to the election-year economy. Many studies have documented

that voters' decisions depend on election-year growth rather than cumulative growth during the entire term (e.g., Achen and Bartels 2004; Fair 1978; Kiewiet 1983; Kramer 1971). Figure 1 illustrates this pattern, showing the relationship between the incumbent party's vote margin and election-year income growth. The relationship is strikingly strong—when incumbents preside over robust election-year growth, they win in landslides, such as in 1964 and 1984. Similar graphs for growth in earlier years of a president's term evince no such relationship.

In Table 2, we show the estimated weights that actual voters assigned to economic growth in each year of presidential terms, using data from 1944 to 2008. We obtain these estimates by regressing the incumbent party's vote share margin on real disposable income (RDI) growth in each of the previous four years (see Column 1). Consistent with previous work (e.g., Bartels 2008), each percentage point of election-year income growth increases the incumbent party's expected margin by just over four percentage points ($b = 4.37$, $p < 0.01$). In contrast, third-year growth has a much smaller effect ($b = 1.48$), which is not statistically significant at conventional levels. First- and second-year growth have point estimates that are close to zero. Converting those coefficients to relative weights suggests that voters put 75% of the weight on election-year growth, 25% of the weight on third-year growth, and no weight on first-year or second-year growth (see Column 2).

FIGURE 1 Election-Year Growth and Incumbent Party Vote Share

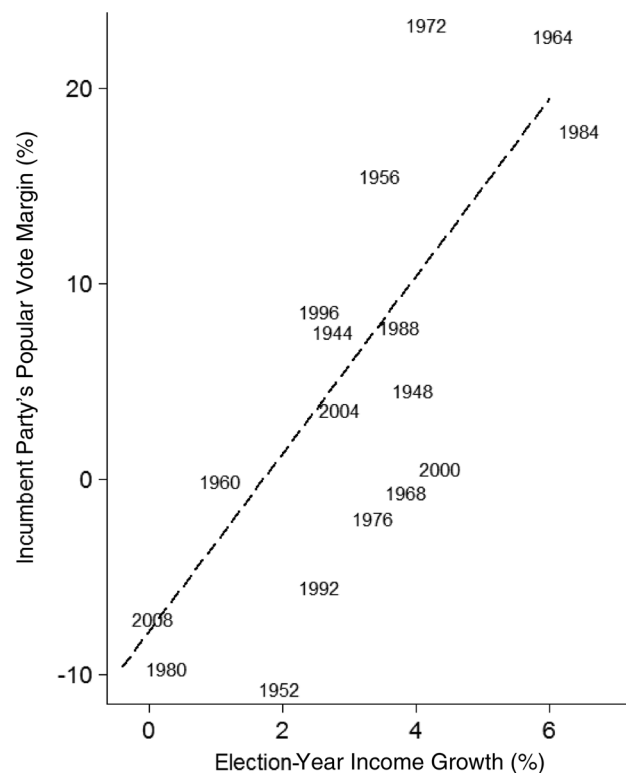


TABLE 2 Actual and Intended Weights for Income Growth

	Actual Weights from Presidential Elections		Intended Weights
	Regression Estimates for Actual Vote Share (1944–2008)	Regression Estimates Converted to % Weights	Large-Scale Internet Sample
	(1)	(2)	(3)
Year 1 Growth	−0.032 (0.70)	0%	20.6% (0.3%)
Year 2 Growth	−0.43 (0.70)	0%	22.9% (0.2%)
Year 3 Growth	1.48 (0.94)	25%	27.0% (0.3%)
Year 4 Growth	4.37 (1.07)	75%	29.6% (0.4%)
R-squared	0.64		
Number of terms	17		
Number of participants			1602
Study			A1

Note: The question asked in Column 3 was “When evaluating the economy during presidents’ terms, how much weight do you typically give to each year? (Please enter a percent for each year, with the total equaling 100%).” Respondents’ answers were constrained to sum to 100%. Standard errors are in parentheses (unstandardized coefficients). The constant is not shown.

To understand this election-year focus, a fundamental first question is whether this behavior is intentional. Do voters mean to put so much weight on the election year? Several explanations suggest that it is intentional. For example, voters may consciously choose to assign extra weight to the election-year economy if they believe it provides more information than earlier conditions about incumbent performance and potential future growth under that incumbent (Erikson, MacKuen, and Stimson 2002).

Our survey responses suggest, however, that voters' behavior may not be intentional. In a national Internet survey fielded by Survey Sampling International (SSI) in May 2010 (Study A1),⁵ 1,602 survey respondents answered the following question: "When evaluating the economy during presidents' terms, how much weight do you typically give to each year? (Please enter a percent for each year, with the total equaling 100%.)" The answers were constrained to sum to 100%. We summarize respondents' intended weights in the third column of Table 2. Although voters actually place weights of 0%, 0%, 25%, and 75% on Years 1 to 4, respectively, the average intended weights in our survey display a different pattern. On average, respondents reported that they intended to use weights of 20.6%, 22.9%, 27.0%, and 29.6% for Years 1 through 4 of a president's term.⁶

Of course, these results are only suggestive.⁷ Not only are they from a single question from a single sample, but aspects of the question may also have led participants to weigh the years equally. In particular, since participants' weights must add up to 100, a quick solution is 25% for each year. At the same time, answering 0%, 0%, 0%, 100% is also an easy solution—one that much more closely matches voters' actual behavior—but one only 20 of 1,602 individuals provided.

To investigate whether the structure of the question could have led to our findings, we conducted several additional studies, finding similar results (see the supporting information for complete details). For example, we tried eliminating the easy answer by requiring the weights to

sum to 70% instead of 100% (see pp. 5–6). We also tried providing participants with predefined sets of weights to choose between, e.g., 0%, 0%, 25%, 75%, or 20%, 23%, 27%, 30% (see pp. 7–9).

In case participants simply failed to understand the questions involving weights, we also tried eliciting their intentions without referring to weights. In Study A2, we asked 572 participants how much their overall evaluation of a president's term would go up if growth improved in a given year (randomized to be either Year 1, 2, 3, or 4). Participants' response options ranged from "Does Not Change" to "Goes Up Substantially." If individuals intended to place more weight on later years, their evaluations should increase least if presented with a Year 1 growth increase and most when presented with a Year 4 growth increase. Using a between-subjects design, we asked participants this hypothetical for only one of the four years: e.g., "by how much would your rating change if growth in Year 4 was three percentage points higher?" Since we only asked participants about a single year, they seem unlikely to be motivated by possible experimenter demand for equal weights. As with the other approaches, participants focused more or less equally on all four years (see the supporting information, pp. 10–11). In sum, no matter how we elicited people's intentions, we arrived at the same conclusion: people intend to weigh all years similarly, with only slightly more weight on later years.

Memory-Based Explanations

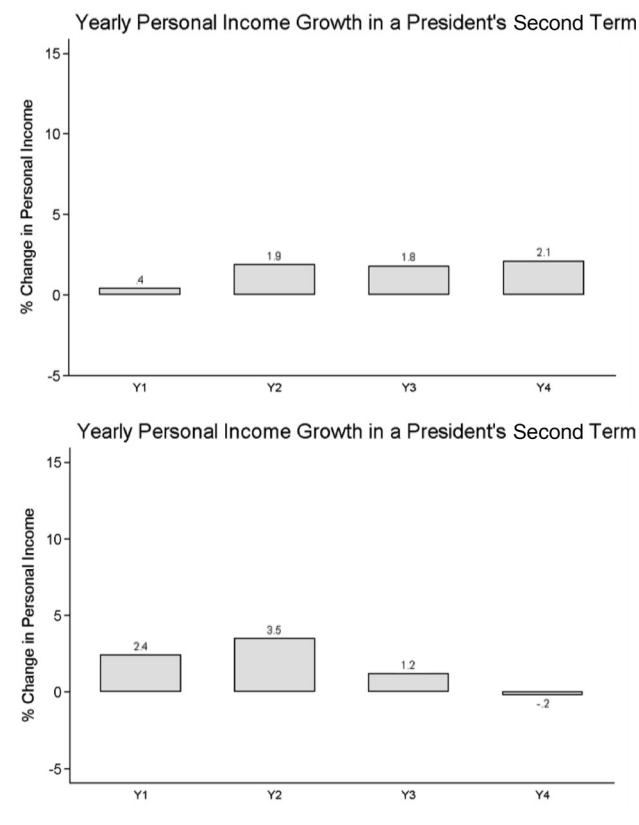
To investigate why voters disproportionately focus on the election-year economy despite no apparent intention to do so, we start by considering memory. Simply put, voters may lack memory of economic conditions in earlier years. Even if voters do not entirely forget earlier conditions, they may primarily put weight on the more recent conditions that they remember better.

To investigate this explanation, we ran experiments where we showed participants information about the economy during presidents' terms. Our motivation was to remove memory constraints about economic performance. In the first experiment (Study B1), participants saw bar plots of income growth showing all four years of historical terms (1941–2008). We presented these plots without referring to the presidents who presided over each term. Figure 2 presents an example of these bar plots (see the supporting information for complete instructions). These 17 terms are the same ones we used in the regression to estimate the actual weights that real-world voters

⁵SSI administered the survey between May 17 and May 19, 2010. We did not employ quotas but asked SSI to construct a target population that matched the (18 and over) census population on education, gender, age, geography, and income. Although the resulting sample is not a probability sample, it is a representative national sample on these demographics.

⁶We also elicited intended weights from our participants after they completed the experiments we describe below, and those weights display a very similar pattern (see the supporting information).

⁷In particular, studies have shown that people have difficulties self-reporting on their own cognitive processes (Nisbett and Wilson 1977).

FIGURE 2 Examples of Two Plots Shown to Participants in Study B1

used (Table 2).⁸ Below each plot, we asked, "How would you rate the condition of the national economy during this period? Is it very good, fairly good, fairly bad, or very bad?" To put income growth and these evaluations on a similar scale, we recoded responses to vary from 0 to 10, with 10 corresponding to "very good." We take the average of the responses for each of the 17 terms and call this variable *Economic Evaluations*.

For this study, we recruited 232 participants through Amazon's Mechanical Turk web service, accepting only those with a 90%-or-better acceptance rate on previous work and paying them \$0.25.⁹ As we note in a later section, the main findings replicate in a national sample. Before analyzing the data, we dropped participants who failed to evaluate every term or who failed a test of attention to instructions (about 10% of the sample). We follow similar procedures for all the Mechanical Turk studies.

⁸We labeled those terms as referring to second-term presidents. We explain the motivation in the next section.

⁹Please see the supporting information for additional details on our sample and Mechanical Turk, as well as Berinsky, Huber, and Lenz (2012).

Even when presented with data from all years in a simple format, respondents put substantially more weight on the economy in the final year of presidents' terms. To illustrate this behavior, consider the two examples in Figure 2. In the top panel, which shows Bill Clinton's first term, growth was only moderate on average, but ends on a relatively strong note. In the bottom panel, which shows Jimmy Carter's only term, cumulative growth was actually stronger than in Clinton's first term, but it slows substantially by the election year. Even though cumulative income growth was higher during Carter's term (6.9% versus 6.2% in Clinton's), participants rated the economy during Clinton's first term as much stronger. On a 10-point scale, they rated Clinton's as 6.2 on average and Carter's as only 3.9 ($p < 0.001$ for the two-tailed test of equality). Participants therefore appeared to focus more on the end, even though they could see all four years and even though we only asked them about the economy's condition, not about the president's performance.

In Table 3, we show that this pattern holds more generally. To do so, we estimate a regression similar to the retrospective voting model in Table 2, but with our participants' ratings of the 17 historical economies as the dependent variable (instead of incumbent party vote margin). That is, to determine the influence that each year had on participants' economic evaluations, we regress participants' average ratings of these economies on the percentage change in income growth in each of the four years (so the N is 17, not 232).¹⁰ As shown in the first column, participants did indeed overweigh the last year, putting an estimated weight of 0.81 on Year 4 growth, compared to -0.13 , 0.16 , and 0.49 , for Years 1, 2, and 3, respectively.

Even though we presented the data from the historical record without presidents' names, participants' evaluations responded to these economies much as real-world election results did. To illustrate this similarity, we recoded both the incumbent party's actual vote margin in these 17 elections (the dependent variable [DV] in Column 1 of Table 2) and our economic evaluations variable (the DV in Column 1 of Table 3) to vary from 0 to 1. Then we regressed each of these rescaled variables on income growth in the four years of presidents' terms. The results indicate that real-world voters and our experimental participants weighed each year of economic growth similarly, even though voters experienced actual

¹⁰We also estimated specifications in which our dependent variable was the individual evaluation, rather than the average across all of the subjects. With or without subject fixed effects, we obtain similar results for the coefficients for each of the years. We focus on the results obtained by averaging across subjects because it leads to the most conservative statistical inference.

TABLE 3 Responses to Historical Income Growth and to Hypothetical Income Growth

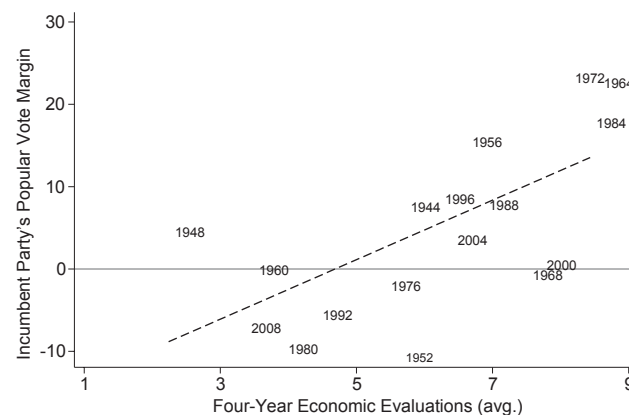
DV:	Historical Data			Hypothetical Data
	Rescaled Dependent Variable			
	Participants' Economic Evaluations (1)	Actual Election Results (1944–2008) (2)	Participants' Economic Evaluations (3)	Participants' Economic Evaluations (4)
Year 1 Growth	−0.13 (0.061)	−0.0009 (0.021)	−0.02 (0.010)	0.27 (0.050)
Year 2 Growth	0.16 (0.061)	−0.01 (0.021)	0.03 (0.010)	0.44 (0.055)
Year 3 Growth	0.49 (0.081)	0.04 (0.028)	0.08 (0.013)	0.50 (0.045)
Year 4 Growth	0.81 (0.093)	0.13 (0.032)	0.13 (0.015)	0.79 (0.060)
R-squared	0.92	0.64	0.92	0.94
Number of terms	17	17	17	25
Number of participants	232		232	209
p-values: Comparing Year 4 weight to				
Year 1 weight	<0.001			<0.001
Year 2 weight	<0.001			<0.001
Year 3 weight	0.030			<0.001
Study	B1		B1	B2

Note: Regression standard errors are in parentheses (unstandardized coefficients). The constants are not shown.

income growth, whereas our respondents only saw it on their computer screens. The point estimates suggest that our experimental participants may put somewhat more weight on Years 2 and 3 than did real-world voters, but the point estimates for Years 1 and 4 are nearly the same. To illustrate this correspondence visually, Figure 3 plots the incumbent party's actual vote share margin from 1944 to 2008 against the average rating of the economy by our experimental participants, showing that our participants' naïve ratings predict the actual election results with remarkable accuracy. This similarity in responses supports the external validity of the experimental results.

Since we use historical data for Study B1, a potential concern is that fourth-year income growth may follow a unique pattern that could drive these findings.¹¹ To address this concern, we ran a replication study (Study B2) where 209 participants observed growth rates that, rather than being drawn from the historical record, were drawn at random from a normal distribution that reflects the income growth rates prevailing in the postwar period (mean

growth equal to two with a standard deviation also equal to two). We randomly selected 25 four-year terms in this way. Again, we found that participants put extra weight

FIGURE 3 Participants' Evaluations of Actual Growth Predict Incumbent Party Vote Share

Note: This figure plots the incumbent party's presidential vote share margin from 1944 to 2008 against the average rating of the economy by our experimental participants in Study B1.

¹¹ Fourth-year income growth does appear somewhat unusual, correlating less strongly with other years in the term than those years correlate with each other (see p. 27 of the supporting information).

on Year 4. The last column of Table 3 presents the results. Participants put roughly 60% more weight on Year 4 than Year 3 ($p < 0.001$) and 80% more weight on Year 4 than Year 2 ($p < 0.001$). In the supporting information, we report additional replications, including a replication of the historical data experiment (see p. 13 of the supporting information), a replication with income-growth bars shown horizontally instead of vertically with the most recent year on top (to rule out a simple visual explanation, see pp. 14–15), several replications with hypothetical data (see p. 16), and finally one where we relabeled the plots as describing the murder rate under governors (see p. 16). In all the studies, we find the same pattern of overweighing later years, especially the last year.

Altogether, the results indicate that, even when people observe information from all four years of presidents' terms displayed in an equally salient way, they continue to assign extra weight to conditions in the most recent year—they appear to behave similarly to real-world voters. The results therefore imply that memory limitations cannot explain most of voters' focus on economic conditions at the end of presidents' terms, even though it may still play some role for real-world voters who lack the information that our participants observed. Our findings suggest that voters would continue to focus on the end even if they knew income growth in all four years.

Do Voters Think the Election-Year Economy Is More Informative?

If it is not memory, do voters put extra weight on the election-year economy because they see it as being more informative about incumbent performance or future growth? As discussed earlier, voters may perceive election-year economic conditions as more informative about incumbents' ability to manage the economy, and so more predictive about future growth under that incumbent (Erikson, MacKuen, and Stimson 2000, 2002; MacKuen, Erikson, and Stimson 1992).¹² Voters may think along these lines especially with first-term pres-

idents, whose policies may take time to influence the economy.

Several findings we have already presented are inconsistent with this explanation. First, citizens' stated intentions to put similar weights on all four years suggest that they do not perceive the election-year economy to be more informative than conditions in earlier years. Second, the experiments described above asked not for prospective judgments or evaluations of incumbents' performance, but only for evaluations of economic conditions during the terms (the murder rate study asked about performance, with the results essentially unchanged). Third, in the experiments described in the previous section (Studies B1 and B2), we told participants that the income-growth data were from the incumbents' second terms.¹³ We did so precisely to address the concern about first-term presidents' policies taking time to influence the economy. As a result, seeing later years in these experiments as more informative about incumbent quality seems especially unlikely. Participants would have to believe that Year 8 of an incumbent's term is considerably more informative about the incumbent's quality than, say, Year 6.

To further assess whether voters perceive the election-year economy to be particularly informative, we conducted additional surveys and experiments. As a starting place, we first investigated whether participants saw income growth from particular years as more informative about income growth in future years (Study C1). After 232 participants rated 17 real-world economies for the experiment we presented earlier (Study B1), we asked, "Say you were trying to forecast the average economy in the four years following these periods. Would all years during these terms be equally predictive of the future? Would later years matter a little more or a little less than earlier years? Tell us what you think by assigning percentage weights to each year." Based on their responses, participants did not appear to see later years as considerably more informative: the average weights they reported are 19.8%, 22.8%, 26.5%, and 30.9%, respectively (see p. 17 of the supporting information for details). Although this study's sample is not a random sample, the answers that respondents gave are similar to the overall weights for related questions in both the national sample and the other online samples reported above.

¹²While voters could perceive the election-year economy to be a better predictor of future growth than conditions in earlier years, actual election-year RDI growth does not predict future RDI growth particularly well. We present evidence on this point in the supporting information, but stronger evidence comes from Achen and Bartels' (2004) finding that those incumbents whom voters reelect, who generally presided over strong election-year economies, did not produce higher levels of post-election GDP growth than would have been produced by those incumbents whom voters did not reelect. Using a different specification, Alesina, Londregan, and Rosenthal (1993, 20) reach a similar conclusion.

¹³To check whether participants read and remembered these instructions, we asked them at the end of a replication study (see p. 13 of the supporting information) whether these hypothetical presidents were in their first terms, second terms, or third terms. Of the 64 respondents in the replication study, only one answered this incorrectly. The results are identical when we exclude this individual.

As with our studies on voters' intended weights, we conducted several studies to address potential concerns and alternative explanations for this finding, but we always found a similar result. The supporting information presents the details. First, we replicated this study with a different sample (see p. 18). Second, we asked a sample of 257 participants how a hypothetical income-growth change in a given year would affect their projections of future growth, randomizing whether respondents observed a Year 1, 2, 3, or 4 growth increase (see p. 19). Third, in a nationally representative sample with 987 participants, we asked participants during the 2012 election year a straightforward question about how important economic conditions were for predicting growth over the next four years, randomly assigning respondents to respond to increased growth either "this year," "last year," "two years ago," or "three years ago" (see p. 20). In a fourth study, we elicited the valuation that subjects put on the growth rate for a given year by asking them to predict future growth after being given the opportunity to obtain information (see p. 21). These latter three studies used between-subject designs where participants evaluated a single year, so participants were unaware of our interest in how people weigh years relative to each other. Across the range of question wordings and samples, we find that people do not see the election year as being much more predictive than earlier years in the term.

As we noted above, the studies where we show participants information on all four years of presidents' terms also seem inconsistent with people seeing later years as more informative. They seem inconsistent because we asked participants to rate the economies themselves, not the president's handling of the economy, so participants should not have been considering policy lags. A concern with those studies, however, is that the language likely raised the salience of politics, potentially leading participants to answer the economic conditions question as if it were about a president's handling of the economy. If people perceive the election-year economy to be more informative about presidential performance, they then could have conceivably put extra weight on the election-year economy even though we asked them to evaluate the economy's performance. To address this concern, we conducted another experiment where we again showed participants four-year economies with randomly drawn income-growth data, but this time we eliminated all mention of presidents or politics from the recruitment language and from the experiment itself (Study C2). (See p. 22 of the supporting information for an example of the plots that participants observed.) The findings in this experiment did not change. Column 1 of Table 4 presents

the regression weights, showing that participants continued to overweigh later years.

We also conducted another experiment designed to make it particularly implausible that the final period could be more informative about incumbent performance (Study C3). We randomly assigned participants to see two versions of the 25 hypothetical income-growth plots. In the "four-year" condition, respondents saw plots just like those shown above, except we told them that the plots showed yearly income growth in a state for a hypothetical governor's term, instead of a hypothetical president's term. In the "four-term" condition, we told participants that the plots showed the average income growth in each *term* for governors with especially long tenures: four terms, with each term consisting of four years. We kept the plots identical except we retitled them and relabeled the four income-growth bars, not as Years 1–4, but as Terms 1–4. Figure 4 shows an example. If voters simply viewed the last year of a term as more informative, the tendency to give extra weight to later data points should diminish in this condition due to the long period covered by four gubernatorial terms. In fact, however, participants continued to put substantially more weight on growth in the last period. Table 4 presents the regression coefficients for these two conditions. The pattern of weights is largely the same whether participants are evaluating four years (Column 2) or four terms (Column 3).

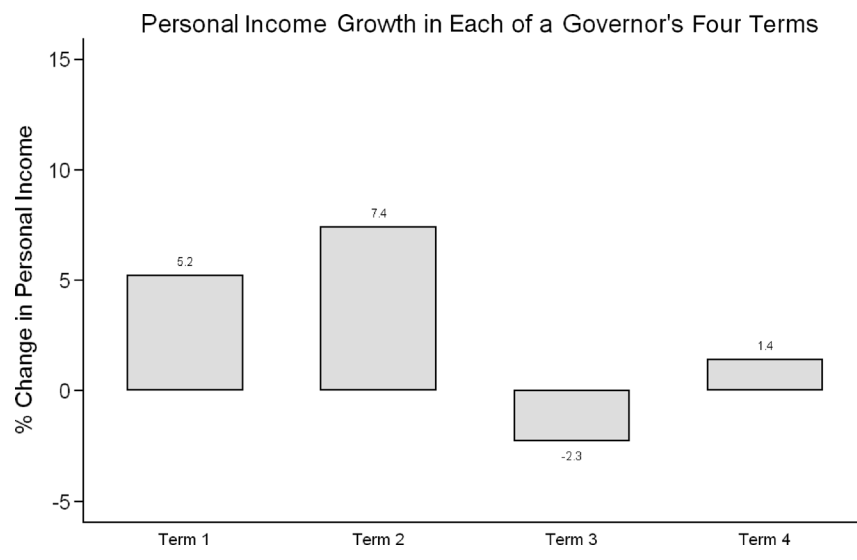
Finally, we considered the possibility that participants' behavior could arise due to a perception that peaks or trends are particularly informative about economic performance. As we noted earlier, psychology studies have found that peaks (or troughs) and trends can influence retrospective judgments, though they appear to do so less often than do conditions at the end.¹⁴ To investigate whether peaks or troughs could account for the weight participants put on the election year, we pooled data from five hypothetical yearly income-growth studies (e.g., B2 and replication studies). We found little sign that years with the most or the least income growth (peaks or troughs) influenced evaluations. On the other hand, we did find some evidence that trends matter, especially downward trends, but controlling for trends left our main findings unchanged. That is, trends do not account for participants' tendency to focus on Year 4 growth. Please see the supporting information for the analysis (pp. 23–24).

¹⁴For example, Carmon and Kahneman (1996) found that respondents applied only an end rule when retrospectively evaluating experiences of waiting in line (Kahneman 2000).

TABLE 4 Robustness Checks: Weights for Hypothetical Growth Data

	President	Governor	
	DV: Participants' Economic Evaluations		
	Political Context Removed	Evaluated over Four Years	Evaluated over Four Terms
	(1)	(2)	(3)
Year 1 Growth	0.22 (0.062)	0.11 (0.081)	0.18 (0.080)
Year 2 Growth	0.43 (0.055)	0.26 (0.074)	0.33 (0.073)
Year 3 Growth	0.49 (0.064)	0.40 (0.083)	0.47 (0.082)
Year 4 Growth	0.72 (0.058)	0.69 (0.069)	0.64 (0.068)
R-squared	0.95	0.89	0.89
Number of terms	25	25	25
Number of participants	81	75	65
p-values: Comparing Year 4 weight to			
Year 1 weight	<0.001	<0.001	<0.001
Year 2 weight	0.003	<0.001	<0.001
Year 3 weight	0.010	0.021	0.145
Study	C2		C3

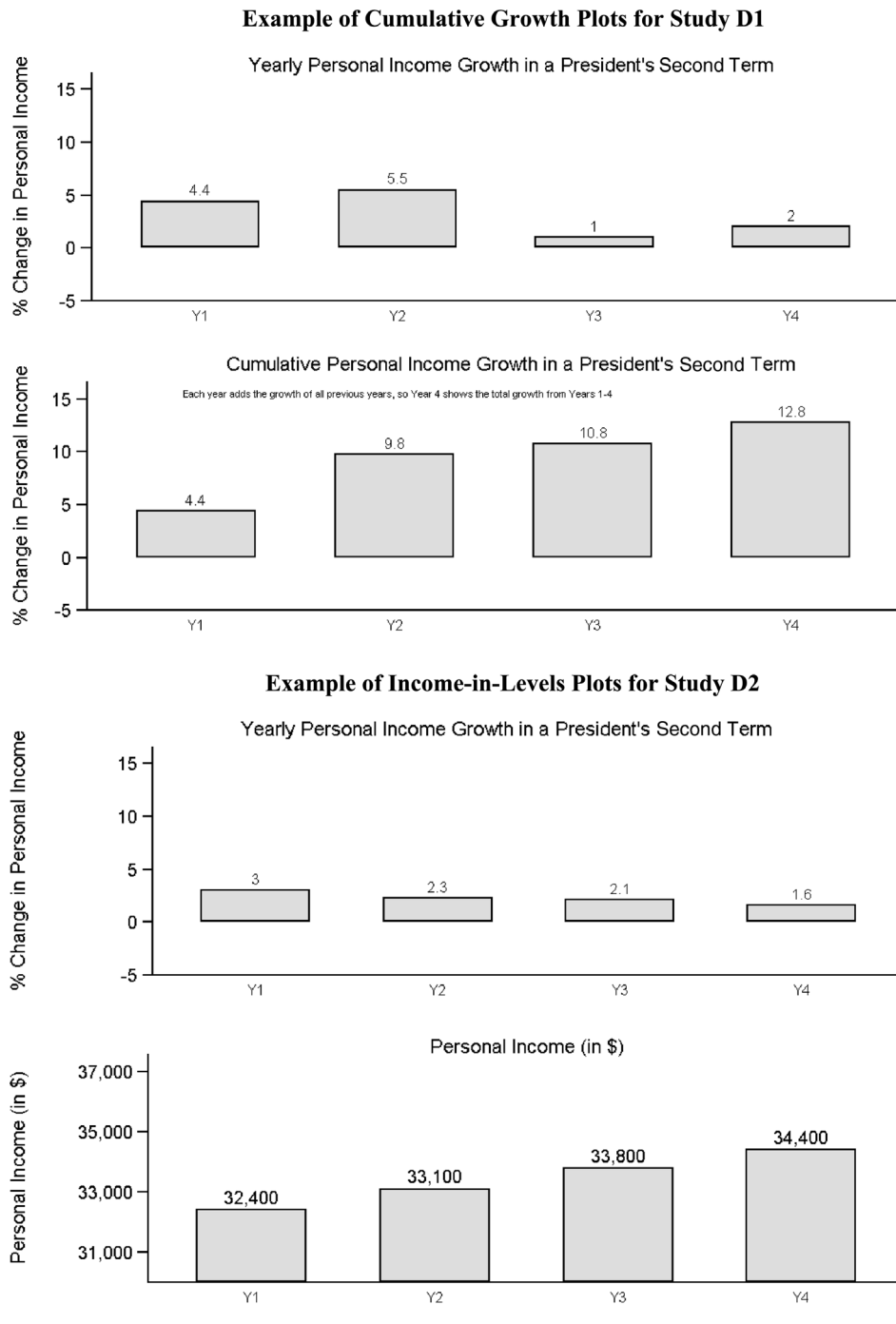
Note: Regression standard errors are in parentheses (unstandardized coefficients). The constants are not shown.

FIGURE 4 Example of Four-Term Growth Plot for Governors

Aligning Voters' Actions with Their Intentions

The previous three sections presented an array of evidence inconsistent with the memory and informative-

ness explanations for voters' emphasis on the election-year economy. In contrast, the end-heuristic explanation can account for the findings. In this section, we directly test a key end-heuristic prediction—one that the other explanations do not predict. If voters are inadvertently

FIGURE 5 Cumulative and Level Growth Plots

applying an end heuristic—substituting economic growth at the end for economic conditions throughout the term—then curing this tendency should be straightforward. We should only have to make readily available the attribute people say they are searching for, something approaching total growth, and people will apply the decision rule they say they intend to implement.

To test this prediction, we conducted several experiments. The first had two conditions (Study D1). In the control condition, participants saw yearly growth plots just like those described in the earlier experiments. In the treatment condition, participants saw both yearly growth and cumulative growth (one on top of the other). Figure 5 (top) presents an example. The cumulative growth plot

TABLE 5 Eliminating the Election-Year Emphasis: Cumulative Growth and Level Plots

	Cumulative Growth Experiment		Income-in-Levels Experiment	
	DV: Participants' Economic Evaluations			
	Yearly Growth (1)	Cumulative and Yearly Growth (2)	Yearly Growth (3)	Income in Levels and Yearly Growth (4)
Year 1 Growth	0.12 (0.060)	0.44 (0.063)	0.19 (0.065)	
Year 2 Growth	0.37 (0.053)	0.50 (0.056)	0.31 (0.059)	0.39 (0.053)
Year 3 Growth	0.46 (0.058)	0.54 (0.061)	0.40 (0.066)	0.47 (0.065)
Year 4 Growth	0.59 (0.040)	0.50 (0.042)	0.63 (0.055)	0.52 (0.054)
R-squared	0.94	0.94	0.92	0.90
Number of terms	25	25	25	25
Number of participants	54	62	97	90
p-values: Comparing Year 4 weight to				
Year 1 weight	<0.001	0.412	<0.001	
Year 2 weight	0.004	0.968	<0.001	0.049
Year 3 weight	0.051	0.504	0.026	0.591
Study		D1		D2

Note: Regression standard errors are in parentheses (unstandardized coefficients). The constants are not shown.

simply adds up the current- and prior-year growth.¹⁵ The figure makes cumulative growth readily accessible but also allows participants to choose between yearly growth and cumulative growth.¹⁶ In the instructions for the study, we provided a brief explanation of how cumulative growth relates to yearly growth, an explanation that accompanied each cumulative plot.¹⁷ The study had 25 terms and used simulated data as described for the earlier studies.

When yearly growth and cumulative growth are equally available, do people focus on cumulative growth?

¹⁵We calculate yearly growth and cumulative growth with natural logs, i.e., yearly percent change = $\ln(RDI_t) - \ln(RDI_{t-1})$. This method converts gains and losses to an equivalent scale and means that cumulative change can be calculated by adding up the yearly percent change from each year.

¹⁶In a pilot study, we showed one group just the cumulative plots, and the results were essentially the same. However, we were concerned about participants potentially misinterpreting the cumulative growth bars as referring to yearly growth. By showing participants both, the experiment forced them to choose, making misunderstanding less of a concern.

¹⁷After rating all the terms, we tested participants' understanding of cumulative growth by asking them a simple question that involves calculating cumulative growth from yearly growth. Ninety percent of the participants answered correctly.

In other words, do they no longer put extra weight on the change in income occurring in Year 4? Table 5 shows estimates of the weights given to each year's income growth for both conditions. When participants saw the standard plot, the end mattered more, as in the studies above. When participants saw both the yearly and cumulative plots, however, they gave roughly equal weight to each year. The weights change from 0.12, 0.37, 0.46, and 0.59, for Years 1–4 respectively, to 0.44, 0.50, 0.54, and 0.50. The ratio of the weight participants place on the fourth year relative to the first decreases substantially, from 4.92 to 1.11 ($p < 0.001$).¹⁸ When we give participants equal access both to cumulative growth and to yearly growth, they therefore rely on the attribute—cumulative growth—that is more informative about overall performance.¹⁹

One potential concern with presenting data in this format is that people may not think in cumulative

¹⁸This test is based on results from individual-level regressions.

¹⁹Moreover, we can reject the hypothesis that people weigh economic conditions in line with their intentions when they only see growth information ($p < 0.001$), but not when they see both growth and cumulative information ($p = 0.156$). See page 28 of the supporting information for details on these tests.

percentage terms. Another, more intuitive, way to make overall economic performance easily available is simply to show the level of per capita income in each of the four years. For example, participants could see that per capita income was \$32,500 in the first year, \$34,000 in the second, and so on. By comparing the fourth year to the first year, participants can easily see the cumulative growth that occurred.

To determine whether presenting income in levels would also eliminate end bias, we conducted another experiment (Study D2). We randomly assigned participants to see yearly growth plots or plots that showed both yearly growth and income in levels. Figure 5 (bottom) presents an example of these plots.²⁰ We drew the initial value (the Year 1 per capita income level) from a normal distribution with a mean of \$32,000 (the average per capita RDI since 2000) and a standard deviation of 100. One issue with the level plots is that respondents cannot observe growth for the first year. For them to see first-year growth, we would have to show them per capita income from the last year of the previous administration. As a result, we estimate the weights that participants assign to the growth rates that occur in each of the preceding three years, as opposed to four years.²¹

As with the cumulative plots, the level plots also appear to reduce the extra weight voters assign to the election-year economy. Column 3 of Table 5 shows the weights for the control condition (growth only) and Column 4 does so for the level condition, using the identical underlying data. Although we cannot compare the first year, the ratio of the weight participants place on the fourth year relative to the second decreases substantially, from 2.04 to 1.40 ($p = 0.030$).²² These results suggest that simply showing voters income in levels, in addition to growth, can reduce the extra weight that people put on the election-year economy.

To address potential concerns with these studies, we conducted six replications. One concern is that the sentence explaining cumulative growth that we added to the standard instructions from Study B1 could have led participants to privilege cumulative growth over yearly growth—though the condition with income in levels lacked any additional instructions and produced similar weights for all years. Another concern is that we

asked participants not about presidential performance, but about the economy during the term, which could push people toward weighing all years equally in the cumulative and in the level conditions—though it obviously did not do so in the yearly growth conditions from the previous section. To address these concerns, we replicated the cumulative study twice. In the first, we explained yearly growth with equal detail as cumulative growth (see p. 29 in the supporting information). In the second, we did not explain cumulative growth in any detail (i.e., identical instructions as in the conditions with only yearly growth; see p. 30). In both replications, we also asked about the president's performance on the economy rather than the economy itself.²³ We likewise replicated the level experiment with the president's performance question (see pp. 31–32). We also worried that the steadier upward trend often seen in level or cumulative growth plots could explain people's responses, so we replicated the results showing participants growth and level information in text form rather than in graph form (see p. 33). Finally, we replicated these findings in a national Internet sample that observed only a single term, a between-subject design (see pp. 34–35). Across all these studies, the primary results remain the same.

Our potential cures for myopic voting—presenting cumulative growth or income in levels—are in keeping with other research on reducing an end heuristic's impact on retrospective evaluations. For example, Liersch and McKenzie (2009) showed participants hypothetical sequences of unpleasant episodes and found that respondents substantially overweighted discomfort at the end (using the same sequences as in Varey and Kahneman's [1992] original demonstration). By varying the graphical presentation of the sequences to make cumulative discomfort more easily accessible, Liersch and McKenzie (2009) found that they could eliminate end bias.²⁴

Taken together, the findings in this section provide further evidence for the end-heuristic explanation. Voters appear to want to evaluate the economy as a whole, not just the end. However, since the whole is not readily available to them—it requires adding up growth across years—they substitute the end. When people observe cumulative growth or income in levels so that the whole is equally available, they no longer make this substitution. Instead, they evaluate the economy as they intend to, focusing more on overall growth. In contrast to the

²⁰Consistent with the cumulative condition, the results refer to subjects who observed growth data on top and level data on the bottom. We also ran a condition with the level data on top. In this condition, subjects put slightly, but not significantly, more weight on later years.

²¹If we include Year 1 growth in the regression, it enters with a near-zero coefficient, and the other coefficients remain nearly unchanged.

²²This test is based on results from individual-level regressions.

²³We thank an anonymous reviewer for bringing these concerns to our attention.

²⁴In Liersch and McKenzie's (2009) article, graphs that make cumulative discomfort easily accessible (e.g., their Figure 1b) reduce end bias. On the other hand, other graphs that do not make cumulative discomfort easily accessible (e.g., Figures 1d and 1e) fail to reduce end bias.

alternative hypotheses described earlier, the end-heuristic hypothesis predicts this pattern.

Conclusions

Voters' tendency to focus on election-year performance has substantial consequences for policy and election outcomes. It may lead elections to favor the best manipulators as opposed to the best stewards of overall economic prosperity. It may incentivize incumbents to take action to inflate election-year growth even at the cost of larger long-term economic damage. Such election-year machinations have led to post-election inflation and recession in countries ranging from the United States and Russia to Mexico and Turkey (Drazen 2001; Gonzalez 2002; Greider 1989; Rogoff 2004). Also, since even the best manipulators have limited control over short-term economic conditions, voters' extra weight on the election-year economy may turn elections into a game of "musical chairs," with voters judging incumbents on whether the business cycle happens to be up or down in the election year (Achen and Bartels 2004).

The studies presented in this article help us understand the underlying causes and potential solutions to this phenomenon. Instead of limited voter memory or voters perceiving the election-year economy as particularly informative, our findings support a simple explanation based on research in psychology. Voters appear to judge incumbents on the election-year economy because conditions at the end are an easily available attribute that they can substitute for the entire experience—in lieu of engaging in the more complicated task of evaluating the whole period. Previous research suggests that people inadvertently engage in this kind of attribute substitution for precisely these kinds of retrospective assessments, whether for colonoscopies or economic evaluations, that require summing up a series of previous experiences. Our results are thus consistent with the finding that, when making quick intuitive judgments, "people take their heuristics off-the-shelf, use them unknowingly and automatically, and rarely worry about their accuracy" (Kuklinski and Quirk 2000, 156), a pattern that holds not only for amateurs, but also for experts such as statisticians making quick judgments (Tversky and Kahneman 1971).

On one hand, our results are discouraging. They imply that the high weight voters assign to the election-year economy does not reflect an intentional decision, such as one based on the belief that election-year growth is more informative or does a better job of forecasting the future. They also imply that the end heuristic should influence decisions across other domains of incumbent

performance, not just the economy, and so may pose a more general problem for democratic accountability.²⁵ In addition to holding for different offices, term lengths, visual presentations, and question wording, our findings also remained essentially the same when participants observed plots referring to crime rather than the economy.

On the other hand, the end-heuristic explanation implies a simple fix. Since voters do not intend to overweigh the election year, they should correct the bias when overall growth is easily available. This is precisely what we found. While voters are not likely to observe the plots that our respondents saw, our results indicate that relatively simple changes in the information context may enable voters to hold their leaders accountable more effectively. Even without the graphs, simply focusing on cumulative growth or income in levels in economic news and discussions may be enough to change voter behavior. Government departments, the news media, or even candidates thus may be able to reduce voters' unintentional shortsightedness by changing how they frame economic data.

Of course, curing this tendency in the real world may be much more difficult than in the lab. For one, representing the complicated process of retrospective voting is difficult with any lab studies, even with the variety of conditions we presented in the article. In addition, the immediacy of the current economy and voters' limited incentives to learn about politics (Downs 1957) are formidable barriers. Nevertheless, understanding that voters' focus on the election-year economy reflects a cognitive bias not only provides a fuller picture of the retrospective voter, but also suggests solutions that could improve democratic accountability.

References

- Abrams, Burton A. 2006. "How Richard Nixon Pressured Arthur Burns: Evidence from the Nixon Tapes." *Journal of Economic Perspectives* 20(4): 177–88.
- Achen, Christopher H., and Larry M. Bartels. 2004. "Musical Chairs: Pocketbook Voting and the Limits of Democratic Accountability." Princeton University. Unpublished manuscript.
- Aksoy, Deniz. 2011. "Elections and the Timing of Terrorist Attacks in Democracies." Pennsylvania State University. Unpublished manuscript.
- Alesina, Alberto, John Londregan, and Howard Rosenthal. 1993. "A Model of the Political Economy of the United States." *American Political Science Review* 87(1): 12–33.

²⁵For example, Bechtel and Hainmueller (2011) reference the "peak-and-end" heuristic to explain voters' reactions to government disaster response before elections, an interpretation that our results support.

- Ariely, Dan. 1998. "Combining Experiences over Time: The Effects of Duration, Intensity Changes and On-Line Measurements on Retrospective Pain Evaluations." *Journal of Behavioral Decision Making* 11(1): 19–45.
- Ariely, Dan, and Ziv Carmon. 2000. "Gestalt Characteristics of Experienced Profiles: The Defining Features of Summarized Events." *Journal of Behavioral Decision Making* 13(2): 191–201.
- Bali, Valentina A. 2007. "Terror and Elections: Lessons from Spain." *Electoral Studies* 26(3): 669–87.
- Bartels, Larry M. 2008. *Unequal Democracy: The Political Economy of the New Gilded Age*. Princeton, NJ: Princeton University Press.
- Baumgartner, Hans, Mita Sujan, and Dan Padgett. 1997. "Patterns of Affective Reactions to Advertisements: The Integration of Moment-to-Moment Responses into Overall Judgments." *Journal of Marketing Research* 34(2): 219–32.
- Bechtel, Michael M., and Jens Hainmueller. 2011. "How Lasting Is Voter Gratitude? An Analysis of the Short- and Long-Term Electoral Returns to Beneficial Policy." *American Journal of Political Science* 55(4): 851–67.
- Berinsky, Adam J., Gregory A. Huber, and Gabriel S. Lenz. 2012. "Evaluating Online Labor Markets for Experimental Research: Amazon.com's Mechanical Turk." *Political Analysis* 20(3): 351–68.
- Berry, Christopher, and William Howell. 2007. "Accountability and Local Elections: Rethinking Retrospective Voting." *Journal of Politics* 69(3): 844–58.
- Carmon, Ziv, and Daniel Kahneman. 1996. "The Experienced Utility of Queuing." Duke University. Unpublished manuscript.
- Downs, Anthony. 1957. *An Economic Theory of Democracy*. New York: Harper and Row.
- Drazen, Allan. 2001. "The Political Business Cycle after 25 Years." *NBER Macroeconomics Annual* 15: 75–138.
- Duch, Raymond M., and Randolph T. Stevenson. 2008. *The Economic Vote: How Political and Economic Institutions Condition Election Results*. New York: Cambridge University Press.
- Duch, Raymond, and Jean-Robert Tyran. 2012. "Coalition Context, Voter Heuristics, and the Coalition-Directed Vote." Oxford University. Unpublished manuscript.
- Erikson, Robert S., Michael B. MacKuen, and James A. Stimson. 2000. "Bankers or Peasants Revisited: Economic Expectations and Presidential Approval." *Electoral Studies* 19(2–3): 295–312.
- Erikson, Robert S., Michael B. MacKuen, and James A. Stimson. 2002. *The Macro Polity*. New York: Cambridge University Press.
- Fair, Ray C. 1978. "The Effect of Economic Events on Votes for President." *The Review of Economics and Statistics* 60(2): 159–73.
- Fredrickson, Barbara L., and Daniel Kahneman. 1993. "Duration Neglect in Retrospective Evaluations of Affective Episodes." *Journal of Personality and Social Psychology* 65(1): 45–55.
- Gonzalez, Maria de los Angeles. 2002. "Do Changes in Democracy Affect the Political Budget Cycle? Evidence from Mexico." *Review of Development Economics* 6(2): 204–24.
- Greider, William. 1989. *Secrets of the Temple: How the Federal Reserve Runs the Country*. New York: Simon & Schuster.
- Healy, Andrew, and Neil Malhotra. 2009. "Myopic Voters and Natural Disaster Policy." *American Political Science Review* 103(3): 387–406.
- Kahneman, Daniel. 2000. "Evaluation by Moments: Past and Future." In *Choices, Values, and Frames*, ed. Daniel Kahneman and Amos Tversky. New York: Cambridge University Press, 693–708.
- Kahneman, Daniel. 2003. "Maps of Bounded Rationality: Psychology for Behavioral Economics." *American Economic Review* 93(5): 1449–75.
- Kahneman, Daniel. 2011. "The Surety of Fools." *New York Times Sunday Magazine*, October 19, 30–33, 62.
- Kahneman, Daniel, and Shane Frederick. 2002. "Representativeness Revisited: Attribute Substitution in Intuitive Judgment." In *Heuristics of Intuitive Judgment*, ed. T. Gilovich, D. Griffin, and D. Kahneman. New York: Cambridge University Press, 49–81.
- Kahneman, Daniel, Barbara L. Fredrickson, Charles A. Schreiber, and Donald A. Redelmeier. 1993. "When More Pain Is Preferred to Less: Adding a Better End." *Psychological Science* 4(6): 401–405.
- Kiewiet, D. Roderick. 1983. *Macroeconomics and Micropolitics: The Electoral Effects of Economic Issues*. Chicago: University of Chicago Press.
- Kinder, Donald R., and D. Roderick Kiewiet. 1981. "Sociotropic Politics: The American Case." *British Journal of Political Science* 11(2): 129–61.
- Kramer, Gerald H. 1971. "Short-Term Fluctuations in U.S. Voting Behavior, 1896–1964." *American Political Science Review* 65(1): 131–43.
- Kramer, Gerald H. 1983. "The Ecological Fallacy Revisited: Aggregate- versus Individual-Level Findings on Economic and Elections, and Sociotropic Voting." *American Political Science Review* 77(1): 92–111.
- Kuklinski, James H., and Paul J. Quirk. 2000. "Reconsidering the Rational Public." In *Elements of Reason: Cognition, Choice, and the Bounds of Rationality*, ed. Arthur Lupia and Mathew Daniel McCubbins. New York: Cambridge University Press, 153–82.
- Levitt, Steven. 1997. "Using Electoral Cycles in Police Hiring to Estimate the Effect of Police on Crime." *American Economic Review* 87(3): 270–90.
- Lewis-Beck, Michael S. 1985. "Pocketbook Voting in U.S. National Election Studies: Fact or Artifact?" *American Journal of Political Science* 29(2): 348–56.
- Liersch, Michael J., and Craig R. M. McKenzie. 2009. "Duration Neglect by Numbers: And Its Elimination by Graphs." *Organizational Behavior and Human Decision Processes* 108(2): 303–14.
- Loewenstein, Georgen F., and Drazen Prelec. 1993. "Preferences for Sequences of Outcomes." *Psychological Review* 100(1): 91–108.
- MacKuen, Michael B., Robert S. Erikson, and James A. Stimson. 1992. "Peasants or Bankers? The American Electorate and the U.S. Economy." *American Political Science Review* 86(3): 597–611.
- Montalvo, José G. 2011. "Voting after the Bombings: A Natural Experiment on the Effect of Terrorist Attacks on Democratic Elections." *Review of Economics and Statistics* 93(4): 1146–54.

- Muthoo, Abhinay, and Kenneth A. Shepsle. 2010. "Information, Institutions and Constitutional Arrangements." *Public Choice* 144(1): 1–36.
- Nisbett, Richard E., and Timothy Decamp Wilson. 1977. "Telling More Than We Can Know: Verbal Reports on Mental Processes." *Psychological Review* 84(3): 231–59.
- Oliver, J. Eric, and Shang E. Ha. 2007. "Vote Choice in Suburban Elections." *American Political Science Review* 101(3): 393–408.
- Redelmeier, Donald A., and Daniel Kahneman. 1996. "Patients' Memories of Painful Medical Treatments." *Pain* 66(1): 3–8.
- Rogoff, Kenneth. 2004. "Bush Throws a Party." *Foreign Policy* 141: 80–88.
- Ross, William T., and Itamar Simonson. 1991. "Evaluations of Pairs of Experiences: A Preference for Happy Endings." *Journal of Behavioral Decision Making* 4(4): 273–82.
- Sarafidis, Yianis. 2007. "What Have You Done for Me Lately? Release of Information and Strategic Manipulation of Memories." *Economic Journal* 117(518): 307–26.
- Shepsle, Kenneth A., Robert P. Van Houweling, Samuel J. Abrams, and Peter C. Hanson. 2009. "The Senate Electoral Cycle and Bicameral Appropriations Politics." *American Journal of Political Science* 53(2): 343–59.
- Tufte, Edward R. 1978. *Political Control of the Economy*. Princeton, NJ: Princeton University Press.
- Tversky, Amos, and Daniel Kahneman. 1971. "Belief in the Law of Small Numbers." *Psychological Bulletin* 76(2): 105–10.
- Varey, Carol, and Daniel Kahneman. 1992. "Experiences Extended across Time: Evaluation of Moments and Episodes." *Journal of Behavioral Decision Making* 5(3): 169–85.
- Zauberman, Gal, Kristin Diehl, and Dan Ariely. 2006. "Hedonic versus Informational Evaluations: Task Dependent Preferences for Sequences of Outcomes." *Journal of Behavioral Decision Making* 19(3): 191–211.

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's website:

- I. Voters' Intended Behavior: Additional Studies
 - A. Histograms for respondents' intended weights
 - B. Eliciting intended weights from the main experimental subjects
 - C. Eliminating the easy $4 \times 25\%$ answer
 - D. Question with predefined weights — Internet sample
 - E. Question with predefined weights — student sample
 - F. Question without referring to weights
 - G. How far do you look back with George W. Bush?
- II. Memory-based Explanations for End Bias
 - A. Replication of Table 3's historical-data study with a separate sample of 64 participants
 - B. Experiments changing the order of the bars and the orientation of the plots
 - C. Replication of Table 3: Additional hypothetical income growth studies
- III. How Informative Is the Election-Year Economy?
 - A. Additional information for simple survey question
 - B. Replication of Study C1
 - C. Predicting future growth when seeing the change for a single year
 - D. Importance of each year when predicting the future economy (YouGov/Polimetrics)
 - E. Hypothetically paying for growth information to predict the future
 - F. Example of growth plot without political context (Study C2)
 - G. Trends and Peaks
 1. Do trends matter? Pooled analysis of randomly drawn experiments on yearly income growth
 2. Do peaks matter? Pooled analysis of randomly drawn experiments on yearly growth
 - H. Is the election-year economy actually more informative about future growth?
 - I. Is election-year growth more informative about earlier growth?
- IV. Aligning Voters' Actions with Their Intentions
 - A. Testing whether voters act in accordance with their stated intentions
 - B. Addressing concerns about demand with the cumulative growth experiments
 - C. Addressing concerns about demand with cumulative experiments (minimal instructions)
 - D. Addressing concerns about demand with the level experiment
 - E. Level experiment using text instead of plots
 - F. Income-in-levels experiment where participants saw only one term (SSI sample)
 - G. Simple income-in-levels experiment
 - H. Average-in-plot experiment
 - I. Twenty-year study
- V. Additional Analysis and Robustness Checks
 - A. Media coverage of the economy: Content analysis
 - B. Replication with Indian subjects
 - C. Accounting for potential repeat takers on Mechanical Turk
 1. Excluding repeat IP addresses across studies
 2. Excluding repeat IP addresses within study
- VI. Experimental Instructions and Related Information
 - A. Recruitment language for Mechanical Turk studies
 - B. Description of task for Mechanical Turk studies
 - C. Mechanical Turk validation
 1. Published validation studies
 2. Citations for MTurk validation studies and selected MTurk publications
 3. Mechanical Turk demographics
 4. Attention to task and screener questions
- VII. Individual-level regressions referenced in footnotes 18 and 22
- VIII. When can people add up the bars in plots?