Economic Evaluations and Partisan Faultfinding: When are Americans Most Likely to Answer Survey Questions Honestly?

This paper introduces a simple framework for understanding which survey questions are more likely elicit political identity-influenced responses. We use daily data from Gallup to test which survey self-reports exhibit more or less susceptibility to politicization, finding the highest likelihood of politicization for macro-social questions. Conversely, we show that self-reported assessments of personal finances are less sensitive to partisan motivated responding. Taken together, our results uncover scope conditions for how to interpret self-reported views of the economy, and argue that measures of public opinion which have not yet been strongly politicized are better proxies for capturing the underlying welfare of the public.

In considering the outcome of a national election, the usual first question, as far as domestic politics goes, is "How's the economy doing?"

Michael S. Lewis-Beck and Mary Stegmaier (2019)

Introduction

There is a positive relationship between the economy and incumbent support [CITES]. And the evidence is robust, initially documented by linking objective measures of economic health with realized incumbent support, and more recently by turning to individual-level survey data in which self-reported assessments of the economy are powerful predictors of self-reported approval of the current government or self-reported vote intentions. It is the latter body of evidence that is the focus of this paper, in which we provide a cautionary brake on how to interpret the patterns found in individual survey data by highlighting the potential for reverse causality, and describing two previously neglected scope conditions: the phrasing of how economic evaluations are elicited, and the placement of these questions within the broader survey. In so doing, we raise questions about the empirical support for one of the most debated theories in political science: retrospective voting.

The positive association between evaluations of the economy and of the current leadership have been used to defend a theory of retrospective voting, in which voters evaluate how the economy has been recently and then either support or punish the incumbent accordingly. Such an interpretation must assume that the self-reported assessment of the economy is exogenous to subsequent evaluations of the current government. Put differently, scholars must assume that the objective state of the economy influences an individuals perception of it, and this perception than conditions their approval / vote intentions. However, we show that economic evaluations are themselves endogenous to co-partisanship with the president, joining a long line of research that highlights the degree to which partisanship colors one's perception of the economy [CITES]. In addition, we identify two additional dimensions by which the magnitude of this reversed relationship is stronger: the extent to which the question about the economy is phrased sociotropically, and the extent to which the rest of the survey includes political questions. We argue that

both of these dimensions prompt respondents to behave more like cheerleading partisans who condition their self-reported assessment of the economy on whether a co-partisan president sits in the White House.

In the following sections, we first define the concepts of socio- and egotropism, and link them with a Zallerian model of considerations (Zaller 1992), which forms the foundation of our intuition. This theoretical intuition generates two hypotheses that we take to the data, both of which expect that survey design can, unintentionally, elicit stronger or weaker partisan responses to questions that are nominally apolitical (such as the state of the economy). We employ a wide-range of methods to evaluate our daily Gallup survey data, including variable importance tests, regression discontinuity design, and Bayesian changepoint analysis. Our results confirm both hypotheses, highlighting the vulnerability of nominally apolitical questions to partisan bias due to two dimensions of survey design: question wording and survey composition.

By providing descriptive evidence of (1) the possibility of reverse causality between economic evaluations and political approval, and (2) the extent to which minor differences in question wording and survey design can inflate these associations, we sound a note of concern about the empirical evidence supporting a retrospective voting model. However, our contribution only speaks to the individual-level survey data that has constituted the majority of recent work on this topic. As such, we are silent on the aggregate patterns that link objective economic conditions with voting behavior [CITES]. Furthermore, our contribution doesn't mean that some of the causally identified experimental work on this topic is similarly vulnerable to reverse causality concerns [CITES]. Nevertheless, we provide an additional dimension on which survey design can inflate or depress associations in observational data that are the bedrock on which theoretical consensus are drawn, and connects with the rich literature on survey design [CITES].

Related Literature

Scholars have long understood the sensitivity of survey self-reports to question wording and survey design. Wording effects such as whether to "allow" or "forbid" a

certain behavior can produce dramatically divergent conclusions about public sentiment [CITE SCHUMAN AND PRESSER 1981]. Multi-year debates over the sensitivity of conclusions to survey design have played out in the pages of the discipline's flagship journal [CITE THE MCKUEN VS OTHERS FIGHTS OF THE 90s]. And the arrangement and sequencing of questions within a survey has well-documented downstream effects on subsequent questions to the extent that experimentally manipulating these factors is a common approach to causal identification in survey experimental research [CITE A BILLION THINGS].

In tandem with the growing appreciation of these sources of fragility in empirical political science is the recognition that Democrats and Republicans have substantially different views of reality. Since Bartels' observation in 2002, myriad papers have been written about how Democrats and Republicans see the world in fundamentally different ways, ranging from whether Iraq had weapons of mass destruction prior to the U.S. invasion in 2003 [CITES]; George W. Bush's responsibility for allowing 9/11 [CITES]; the magnitude of and threat posed by climate change [CITES]; the health risks associated with Covid-19 [CITES]; to a parade of claims made by Former President Trump on everything from the size of his inauguration crowd to the ... [CITES].

Implicit in these findings is a shared assumption that surveys are measuring genuine views of the people who respond. But as Bullock and Lenz (2019) [CITE] summarize in their comprehensive review of the literature, it is not clear how much of these differences are attributable to expressive responding motives such as partisan cheerleading, versus what they term "congenial inference". The former suggests that evidence of partisan differences in perceptions of objective reality are driven primarily by dishonest responses on surveys. The latter provides an explanation for why these partisan gaps in perception might be genuine, and carries much graver implications for democratic deliberation. In their review, they highlight a swath of recent research that manipulates the incentives for accurate answers via monetary incentives, interpreting the evidence in light of this distinction, and calling for additional research on the topic [CITE SOOD AND PRIOR AND ALL THE REST].

One of the favored outcomes used in these studies are questions about economic conditions, which have been shown to be partisan in nature, despite ostensibly concerning

an objective state of the world. This is also the outcome we focus on in this paper, and demonstrate how variation in the way in which these evaluations are elicited, and the design of the surveys in which they are asked, can augment or suppress the strength of the partisan bias in the responses. In so doing, we rely exclusively on observational data, and are therefore unable to disentangle cheerleading from congenial inference. As we discuss below, it might be that more sociotropic questions prime respondents to consider their co-partisan status with the president, and then express their support for (or disdain of) the president by inflating (or reducing) their evaluations of the economy. Conversely, more sociotropic questions are inherently more vague in the sense they are about conditions that individual respondents might be less aware of, prompting them to rely on heuristics like co-partisanship with the president in a boundedly rational manner.

Theoretical Intuition

Before describing the intuition by which question wording might elicit more politicized responses, we review the concepts of socio- and egotropism, which are terms that can help us distinguish between two types of considerations voters may have when making decisions, and the concept of a "consideration" itself, which we summarize from Zaller 1992.

A "consideration" is an attribute or aspect of a political topic that influences an individual's summary judgement on that issue by combining cognition and affect. As in Zaller's original description, a relevant consideration to the political topic of defense spending is that the Pentagon spends too much money – a consideration that combines a cognitive piece (e.g. thinking about Pentagon spending) with the affect of believing this amount is excessive. In our setting, a relevant consideration to the topic of the economy is the responsibility attributed to one party or another for overseeing the economy: a consideration that combines the cognition of thinking about how much control the president or partisans in Congress might actually have, with the affect of whether the respondent views a given political party in a positive or a negative light. The consideration of partisan responsibility is thus the core explanation for the degree to which a survey response exhibits some type of partisan bias, either due to cheerleading or bounded rational

expectations.[1]

Turning to the terms "egotropism" and "sociotropism", both adjectives incorporate the root word "tropic," derived from the Greek *tropos* meaning a turning or an orientation. In a broader scientific context, tropic tends to describe an orientation toward a particular stimulus or objective. In biology, for instance, plants exhibit "phototropic" behavior when they orient themselves and grow toward a light source. Within social sciences, an orientation or inclination in decision-making can also be conceptualized in such tropic terms: thus, sociotropic denotes an orientation towards societal concerns, egotropic indicates an orientation towards personal concerns, and a middle ground has been labelled "communotropic" (Rogers and Tyszler 2018) or also "mecro-economic voting" (Ansolabehere, Meredith, and Snowberg 2014), and the relationships between community-level outcomes and economic evaluations have also recently been examined (Bisbee and Zilinsky 2023).

Combining these perspectives, our core claim is that the consideration of partisan responsibility is more likely to be activated when a survey question is worded in a more sociotropic way, or is surrounded by many other sociotropic questions. For example, asking about the "state of the national economy" is often used as proxy for how the public perceives economic conditions, but it may also activate the consideration of partisan responsibility due to its sociotropic framing, which then injects partisan bias into responses.[2] Conversely, asking about an individual's satisfaction with their finances is an alternative measure of the public's perception of economic conditions, which is less likely to activate the partisan

---

[1]We are agnostic on which of these explanations of partisan bias holds. It might be purely expressive cheerleading in which the survey respondent experiences some cognitive benefit from signaling group affiliation by either inflating their approval of an economy shepherded by a co-partisan president, or depressing their approval of an economy when an out-partisan holds office. Conversely, it might be boundedly rational in the sense that partisans might rationally expect better economic conditions when their party holds power. In either case, the causal arrow between leadership approval and economic evaluations is reversed.

[2]But even if respondents are more likely view such a question as an opportunity to give voice to their partisan opinions (relative to questions which elicit egotropic evaluations), we do not mean to suggest that response substitution is inevitable. Rather, our aim is to quantify the possibility that a response is influenced by respondents' partisan identities.

responsibility consideration. Similarly, asking a question about personal finances directly after a battery of questions about one's health and health-related behaviors is likely to be more insulated from a prior question about presidential approval compared to the same question asked shortly after the same presidential approval question.

The preceding argument motivates the following empirical expectations:

H1: Sociotropically-phrased questions should exhibit greater partisan bias than ecotropically-phrased questions.

H2: Removing egotropically-phrased questions from a survey will increase the overall partisan bias found in subsequent questions.

We evaluate these two hypotheses in the following sections. Importantly, we do not claim to provide a set of guidelines about how best to phrase survey questions, nor how to design a survey that minimizes partisan-biased responses. We only highlight a previously neglected dimension of survey design that can contribute to the types of biased responses which might inflate theoretically-relevant associations between self-reported economic evaluations and political attitudes (approval for the current leadership) or behaviors (vote intention). Our hope is that applied researchers might consider these dimensions when designing and piloting their surveys. Subsequent work that experimentally manipulates both dimensions would help put more precision on our results, although we are skeptical that a comprehensive list of question wordings and survey designs is possible, given the myriad purposes for which applied researchers ask these questions.

Data and Methods

Data

We describe empirical patterns that are consistent with our framework using daily Gallup survey data from 2008 to 2018. These data ask a variety of questions that, we

argue, provide variation in the spectrum of interest: egotropism on the one hand and sociotropism on the other. This variation manifests in two ways. The first dimension of variation is across questions pertaining to economic and financial conditions, which varied in the degree to which they present egotropic versus sociotropic frames. The list of questions we examine is presented below, ranging from the most egotropic to the most sociotropic, along with reference questions which we would expect to be highly politicized. We bold parts of each question which pertain to the dimension of tropism.[3]

- Economic situation (personal and national):

  - "Are **you** satisfied or dissatisfied with **your** standard of living, all the things **you** can buy and do?"

  - "[Agree/Dis] Compared to the people **you** spend time with, **you** are satisfied with your standard of living."

  - "Have there been times in the past twelve months when **you** did not have enough money: To buy food that **you** or **your** family needed?"

  - "Would **you** be able right now to make a major purchase, such as a car, appliance, or furniture, or pay for a significant home repair if **you** needed to?"

  - "At this time, are **you** cutting back on how much money **you** spend each week, or not?"

  - "Did **you** worry yesterday that **you** spent too much money, or not?"

  - "[Agree/Dis] In the last seven days, **you** have worried about money."

[3]Note that our operationalization of the concept of tropism is necessarily coarse. We do not mean to suggest that second-person pronouns guarantee insulating from partisan motivated responding, nor that a mention of a country writ large invites this type of bias. For example, the life ladder questions – despite not explicitly referencing the country writ large – nevertheless evoke comparisons to a hypothetical life that might prompt considerations of a broader reference group. Similarly, the standard of living implicitly suggests a comparison group against which one's standard is measured. Conversely, precise questions about instances over the past 12 months in which the respondent did not have enough money, are less likely to stimulate sociotropic considerations, and are therefore less amenable to partisan motivated responding.

- "[Agree/Dis] **You** have more than enough money to do what **you** want to do."

- "[Agree/Dis] **You** are watching **your** spending very closely."

- "Do **you** have enough money to buy the things **you** need, or not?"

- "Are **you** feeling better about your financial situation these days, or not?"

- "Are **you** feeling pretty good these days about the amount of money **you** have to spend, or not?"

- "How would you rate economic conditions **in this country** today – as excellent, good, only fair, or poor?"

- "Right now, do you think that economic conditions **in this country, as a whole**, are getting better or getting worse?"

- Life Ladder

  - "Please imagine a ladder with steps numbered from zero at the bottom to ten at the top. The top of the ladder represents the best possible life for **you** and the bottom of the ladder represents the worst possible life for **you**. On which step of the ladder would you say **you personally feel you** stand at this time?"

  - "Please imagine a ladder with steps numbered from zero at the bottom to ten at the top. The top of the ladder represents the best possible life for **you** and the bottom of the ladder represents the worst possible life for **you**. On which step of the ladder would you say **you** will stand about five years from now?"

  - "**You** are proud of **your** community or the area where **you** live."

- Pure politics:

  - "Do you approve or disapprove of the way Donald Trump is handling his job as president?"

  - "Do you have a favorable or unfavorable opinion of Hillary Clinton, or haven't you heard of them?"

  - "Do you approve or disapprove of the way Barack Obama is handling his job as president?"

The second dimension of variation is found *within* survey questions thanks to a design change Gallup implemented on January 1st, 2013. Prior to this date, Gallup included the above-listed questions of interest along with a battery of questions about the respondents' physical and emotional health, all of which were asked to a daily sample of 1,000 respondents. After January 1st, 2013, Gallup divided its daily sample of 1,000 respondents into two groups of 500, one of which was given only the strictly egotropic questions about their mental and emotional well-being, while the other answered only the battery of economic and financial questions. This decision meant that respondents who had previously been asked about their financial situation as part of a broader survey asking about their weight, stress levels, and other health issues, began answering the same question in a smaller survey that only asked about a range of economic and financial questions, along with explicitly political questions regarding presidential approval, trust in government and policy preferences. We argue that this shift in the design made even the most egotropically-phrased questions about economics more sociotropic by association, thereby prompting more partisan considerations and, by extension, more partisan responses.

The individual-level responses from Gallup are geocoded, allowing us to merge in an extensive set of contextual variables with each respondent. This additional information will allow us to account for the state of the community where respondents live. This is important because, if we observe that Democrats and Republicans systematically report, for instance, diverging economic expectations, we need to account for the possibility that one of these groups may reside in areas which are economically declining/improving at a particular point in time, and differentially from the other group(s).

Methods

To characterize the degree to which these questions evoke partisan bias in responses, we employ two complementary approaches. Our first approach measures how *important* different variables are for predicting the outcomes of interest. To do so, we combine a random forest model with permutation analysis. The random forest algorithm relieves us from having to either assume that every predictor enters linearly, or to determine the

functional form *a priori*. Furthermore, random forests rely on searching across random selections of the covariate space for break points that best divide observations by the outcome of interest. Using this non-parametric machine learning algorithm, we use permutation methods wherein we randomly reshuffle predictors one at a time, and compare the model's predictive accuracy under the permuted setting to that obtained with the original data. Predictors whose permutations produce more substantial declines in the random forest algorithm's predictive accuracy are interpreted to be more important to characterizing an outcome.[4]

Our second approach is more descriptive in nature and simply documents the rapid pivots in evaluations of the national economy that correspond to changes in the office of the president. Specifically we focus on two changes in the partisanship of the president: the transition from George W. Bush to Barack Obama in January of 2009, and the transition from Barack Obama to Donald Trump in January of 2017. We examine the degree to which co-partisans of the preceding president shifted their responses to a variety of nominally non-partisan survey questions. We interpret questions which exhibit dramatic shifts in opinions associated with the change in the presidency as more susceptible to partisan motivated responding. Furthermore, we argue that the speed of these shifts is unlikely to reflect rational updating about future conditions, and is instead more plausibly connected to expressive responding.[5]

Finally, we combine these tools to document the over-time trends in the variable importance measure for partisanship. If the explanatory power of partisanship spikes prior to presidential elections — especially for the more sociotropically-framed questions — such behaviors of survey respondents would be consistent with our theoretical framework. In addition, we compare the predictive power of partisanship just before, and just following, the change in Gallup's survey design on January 1st, 2013.

---

[4]We demonstrate that our conclusions are robust to alternative methods – including LASSO regressions (Figure S2) as well as calculating the expected percentage reduction in error (ePRE, Herron 1999) from a logistic regression model (Figure S3) – in the Supporting Information.

[5]We argue that, while there might be (boundedly) rational explanations for updating one's beliefs about the *future* health of the economy, this seems less plausible for contemporary questions about how the economy is doing today.

Results

We divide the presentation of our results according to the methods described above. First, we examine measures of variable importance across outcomes. We expect to find that the relative importance of partisanship in predicting outcomes varies as a function of how egotropic the questions are, with the most egotropic questions being least associated with partisanship and the most sociotropic questions exhibiting evidence consistent with partisan motivated responding. Second, we exploit the rich time series cross sectional nature of our data to test whether partisan motivated responding is stable over time. We expect the importance of party affiliation (for sociotropic questions) will increase in the weeks and months prior to national elections, and that there should be evidence of a rapid flip among partisans following a change in the partisanship of the elected president.

Variable Importance

We begin with estimating a measure of variable importance which calculates the reduction in a model's predictive accuracy when one of the variables is randomly reshuffled (permuted) – breaking the empirical relationship between it and the outcome – while holding all other variables constant. Here, the baseline performance is the fully specified model with all variables unpermuted. Importantly, this comparison captures the relative decline in model performance *when all other predictors are included*. Specifically, we estimate a set of random forests predicting responses to 17 survey questions, and predictor of interest is partisanship. Other potential predictors include respondents' income, gender, age, education, and information about data the economy (county labor force measures including the unemployment rate, the labor force participation rate, average wages, and employment by industry), crime (county arrests per capita by offense), demographic composition (county population by race, age, gender, and rural) and public health outcomes in the respondents' communities (male and female death rates and life expectancies by age cohort). We find that when a question is either political or sociotropic nature, these question characteristics are strongly associated with the degree to which partisanship – and

by extension, partisan motivated responding – influences survey responses (Figure 1).[6] Consistent with our theoretical argument, partisanship is approximately twice as predictive in evaluating the national economy as it is in forecasting respondents' (self-reported) feelings about their personal financial situation.

How influential is party ID relative to other predictors? Figure 2 compares the variable importance for a selection of individual and contextual-level measures across two outcome measures: whether the respondent has enough money (egotropic) and whether the respondent thinks that the national economy is improving (sociotropic). Consistent with expectations, non-political predictors are substantially less prognostic of the more sociotropic question, and substantially more prognostic of the more egotropic question. For example, average local wages are more prognostic of responses An important extension to our work that is beyond the scope of this paper would involve a systematic analysis of the contextual measures of local economic, health, and crime-related factors to understand the degree to which these presumably important phenomena are reflected in survey self-reports.[7]

Flip-flopping with the President

To motivate the subsequent analysis, consider this purely descriptive plot of daily views of the economy, broken out by party ID (Figure 3). There is striking evidence that Democrats and Republicans effectively traded places in the weeks following the inauguration of new presidents (Obama in January of 2009, Trump in January of 2017). While this dataset is not a genuine panel (same individuals are not observed over time), these data are based on nationally representative samples, making it quite implausible that these patterns could be explained by wholly different partisans responding after each

---

[6]This approach to variable importance allows us to compare different types of outcome variables, including those with ordinal responses.

[7]Regressions of this nature lie at the heart of a rich literature in IPE (among other subfields) that seeks to link the local consequences of import competition and offshoring with political beliefs. See Mansfield and Mutz 2009; Ballard-Rosa, Jensen, and Scheve 2022 for some examples.
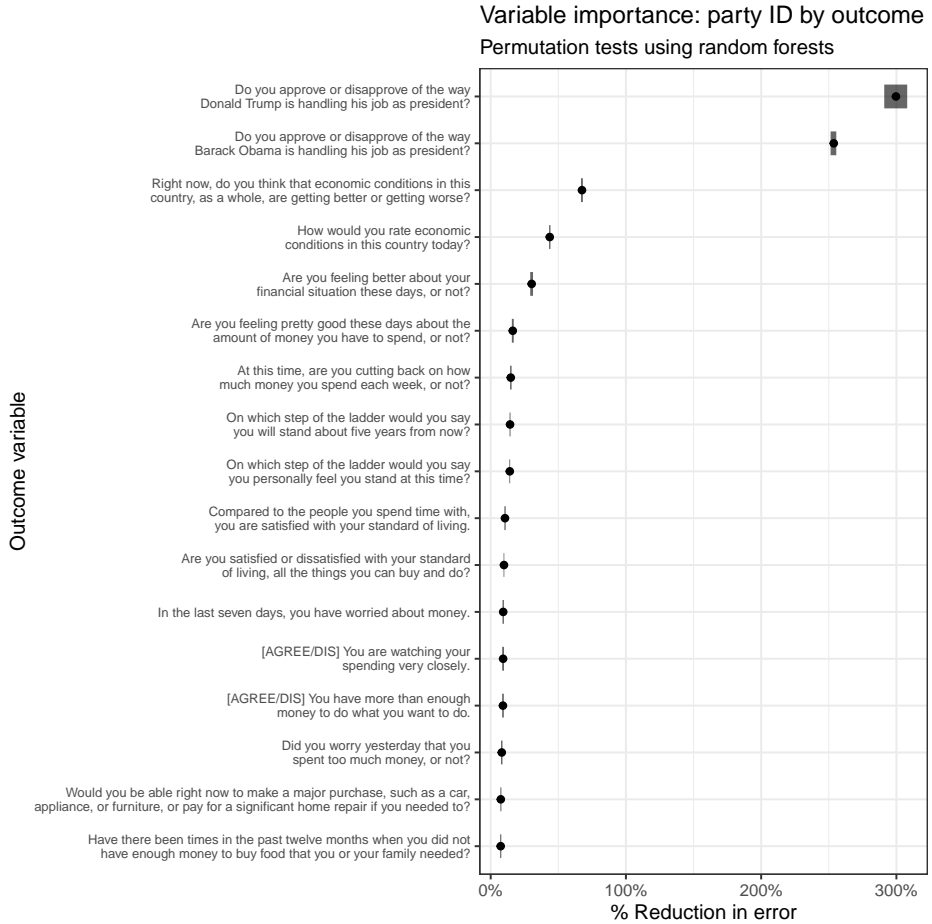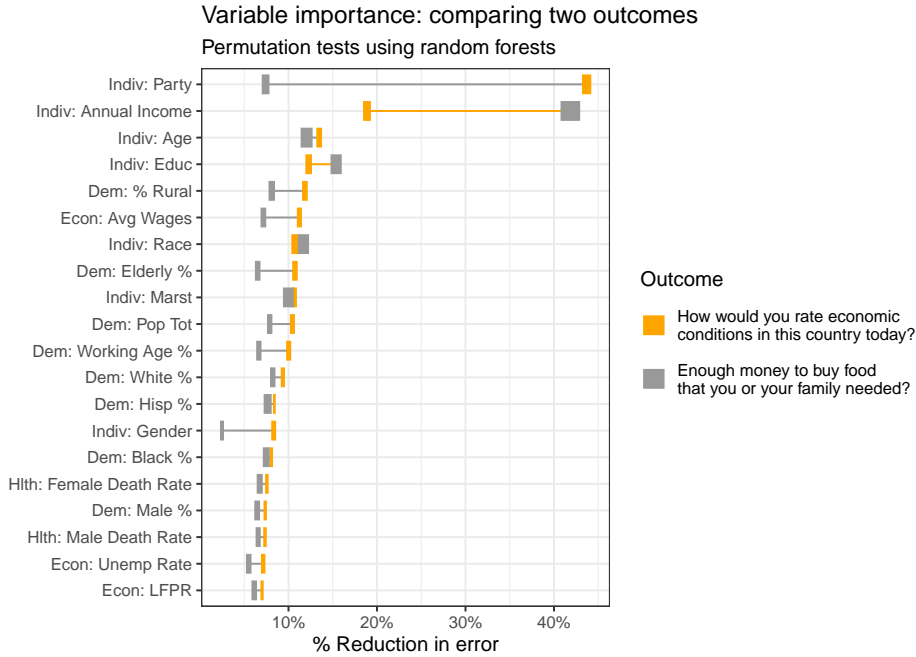
Figure 1. *Percent reduction in model accuracy (x-axis) associated with permuting the respondent's partisanship, relative to a random forest with an exhaustive set of individual and contextual predictors, across survey questions (y-axis). Upper and lower bounds based on 10-fold cross validation indicated with the widths of the boxes.*

inauguration from those that responded prior. Instead, this can be viewed as clear evidence of motivated views and specifically – unless partisans truly modified their beliefs in a short span – motivated *responding*.

Nevertheless, there might be confounders which explain these transitions. For

Figure 2. *Percent reduction in model accuracy (x-axis) associated with permuting predictors (y-axis), relative to a random forest with an exhaustive set of individual and contextual predictors, comparing a sociotropic question on the national economy (in grey) to an egotropic question on whether the respondent has enough money (yellow color). Upper and lower bounds based on 10-fold cross validation indicated with the widths of the boxes. Each row indicates whether a variable was measured by Gallup and pertains to respondents ("Indiv.") or whether it is a contextual variable which was merged with the observations in the individual-level Gallup data (e.g. "Dem."=demographics; "Hlth"=local health outcomes).*

example, dissatisfaction with the 2008 election outcome might correspond to negative outcomes in Republican areas, producing bleaker views as a function of genuine beliefs about the economy instead of purely expressive responding. To account for similar possibilities, we run a flexible random forest model that predicts different outcomes as a function of all individual and contextual predictors for each money in the data, and then use these models to predict opinions over time by party ID. Figure 4 visualizes the results from a selection of outcomes, ranging from the most egotropic (whether a respondent experienced stress during the preceding day) on the top to the most sociotropic (whether

Daily average views by party ID
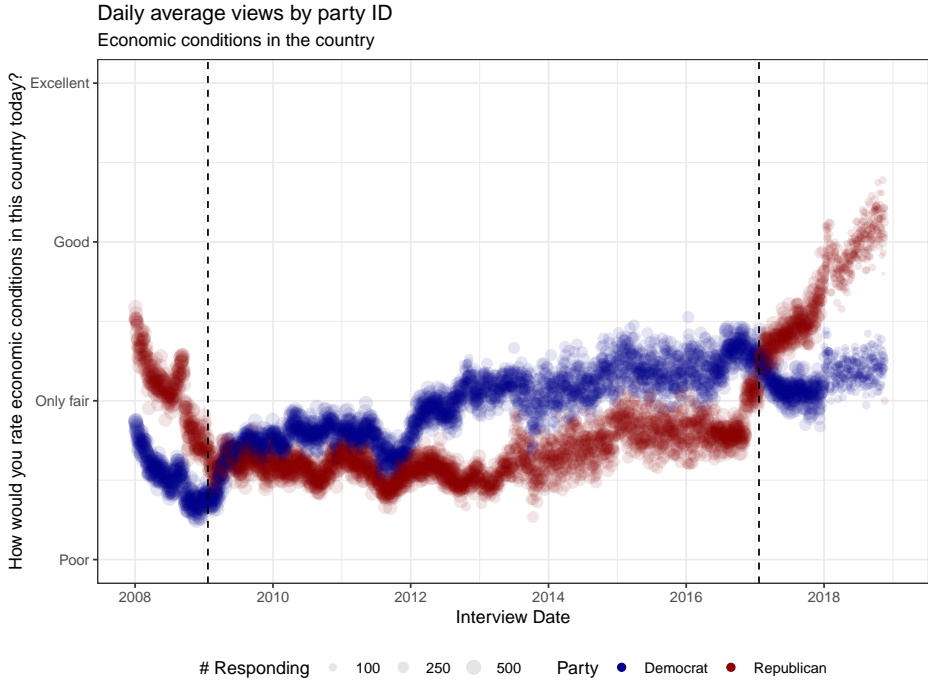Economic conditions in the country

Figure 3.   Daily average views of the economy on a 4-point scale by party ID of the respondent. Partisans include leaners (Independents and non-responders are dropped). Dashed vertical lines denote inauguration dates.

the national economy is improving) in the bottom panel. As illustrated, the evidence of flips in responses by partisans around the inaugurations of Obama and Trump is striking for the most sociotropic measures, and less so as we move up the panels of the figure.

The Dynamics of Variable Importance

In our next analysis, we combine the two preceding approaches and calculate variable importance as a function of time. Specifically, we show in Figure 5 that the predictive power of party ID for sociotropic measures tracks election cycles, with increasing relevance in the final months of a presidential campaign, and diminishing strikingly during
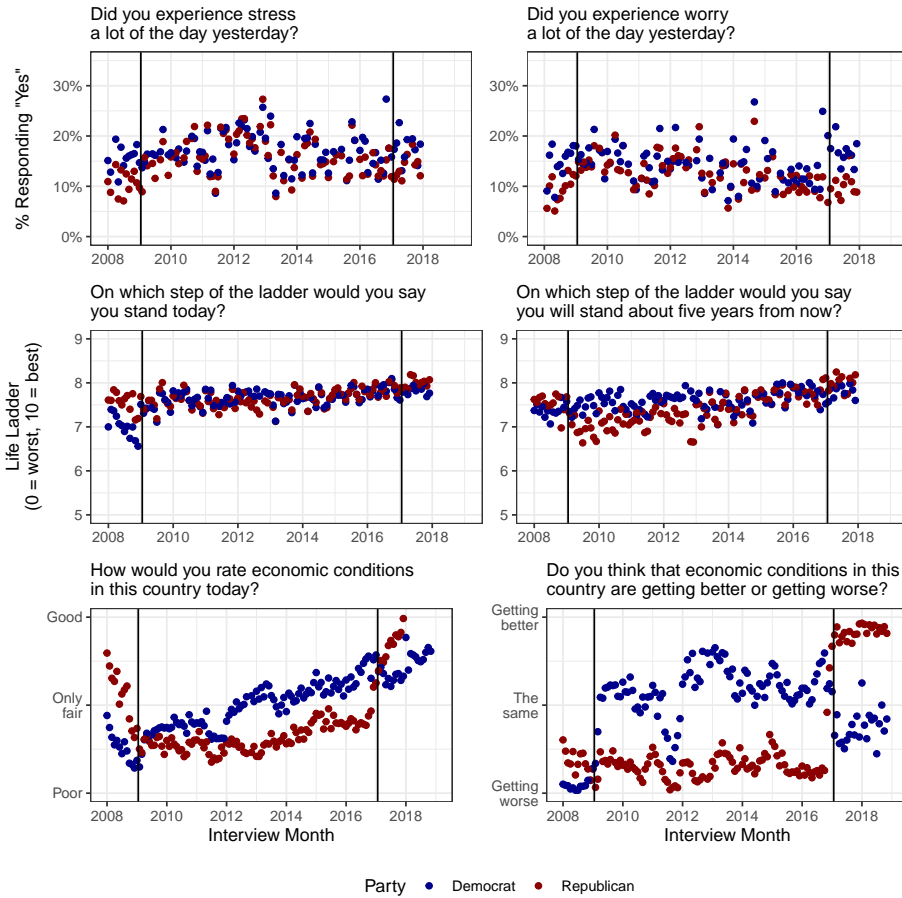
Figure 4. *Predicted attitudes from flexible random forest model.*

the period between the election and the inauguration (indicated by grey rectangles in Figure 5. This over-time evidence is not found for the egotropic question about whether the respondent has enough money, although there is a small increase in party ID after Trump's inauguration.

The monthly evidence reported in Figure 5 is consistent with campaign effects and with the conjecture that subjective reports of (national) economic conditions are less

informative (or at least answered 'less truthfully') during intense electoral campaigns. Conversely, responses to a question about personal finances are less tainted by partisan motivated responding, and the (limited) influence of partisanship is stable over entire (4-year) electoral terms of U.S. presidents.
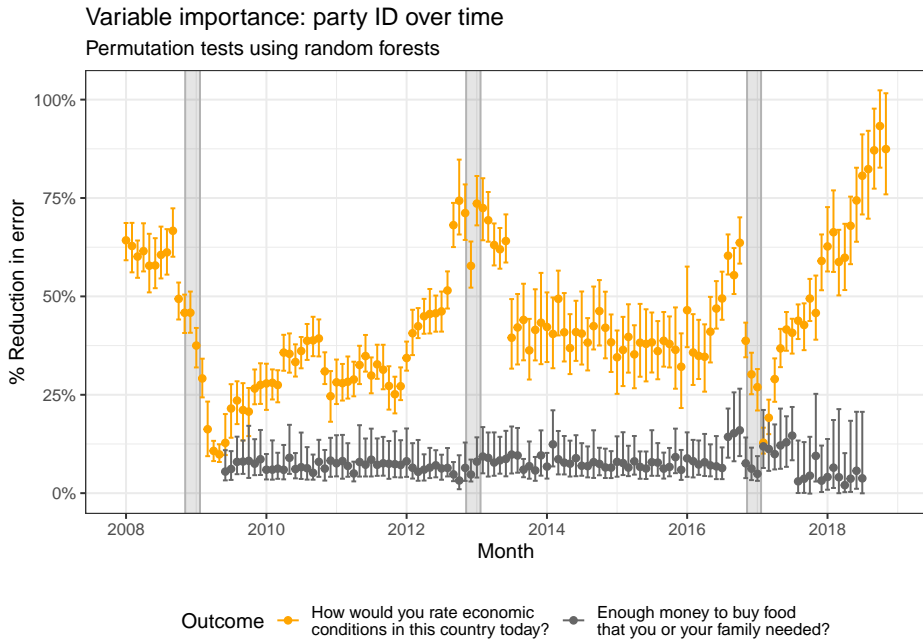
Variable importance: party ID over time
Permutation tests using random forests

Figure 5. Variable importance of party affiliation (from models estimated on a monthly basis), for two outcomes: whether the national economy is improving (grey) and whether the respondent has enough money (orange).

Survey Design

The preceding analyses all rest on our interpretation of some questions as more sociotropic than others, which we posit can elicit more partisan considerations and, thus, more expressive responding. While we feel comfortable in our claim that the phrasing of questions about the national economy are more sociotropic in nature than questions about whether the respondent has enough money, we recognize that these are subjective claims.

To further support our argument that the sociotropic nature of survey design can elicit more expressive responding along partisan lines, we hold constant the survey questions and document how the predictive power of partisanship shifts along with the overarching design of the survey.

Specifically, we take advantage of a change in the design of the daily Gallup tracking polls that occurred on January 1st, 2013. Prior to this date, all questions were asked of 1,000 respondents daily, including both questions on politics and the economy, as well as questions on personal health. After January 1st, 2013, Gallup divided these into two separate surveys, each fielded to 500 respondents. Both surveys started with a question about presidential approval, priming respondents to consider political responsibility for the economy. However, in the pre-2013 design, many of the more egotropic questions pertaining to the respondent's economic situation were located after the battery of well-being questions that asked about deeply personal issues involving physical and mental health. But when these surveys were split apart, the same egotropic questions regarding the respondent's economic situation no longer came after the broader set of strictly egotropic questions. Instead, the rest of the survey is much more explicitly political in nature, asking only questions about the respondent's employment status and conditions at their job, before turning to the same egotropic questions asked prior to 2013. We visualize these differences in Figure **??**, and include two examples of the pre- and post-change survey in the Supporting Information for reference.

We argue that these changes in survey design removed a buffer between the explicitly political question about presidential approval, and the egotropic questions of interest to our analysis. In particular, we believe that the questions about physical health and well-being are sufficiently non-political to reduce whatever downstream influence the initial presidential approval question may have primed. But without them acting as a buffer, the political considerations elicited by the initial presidential approval question manifest more strongly, rendering even the egotropic questions about one's individual economic conditions endogenous to their partisanship and that of the president. If the relevant considerations that prompt expressive responding are sensitive not just to a specific question, but also to the broader organization of a survey, we would expect to find greater evidence of expressive responding on the *same egotropic questions* after January 1st, 2013 when they are no longer associated with a broader battery of strictly personal questions

Pre-Change **Post-Change**

| Presidential Approval | | Presidential Approval |

| Health & Well-being | |

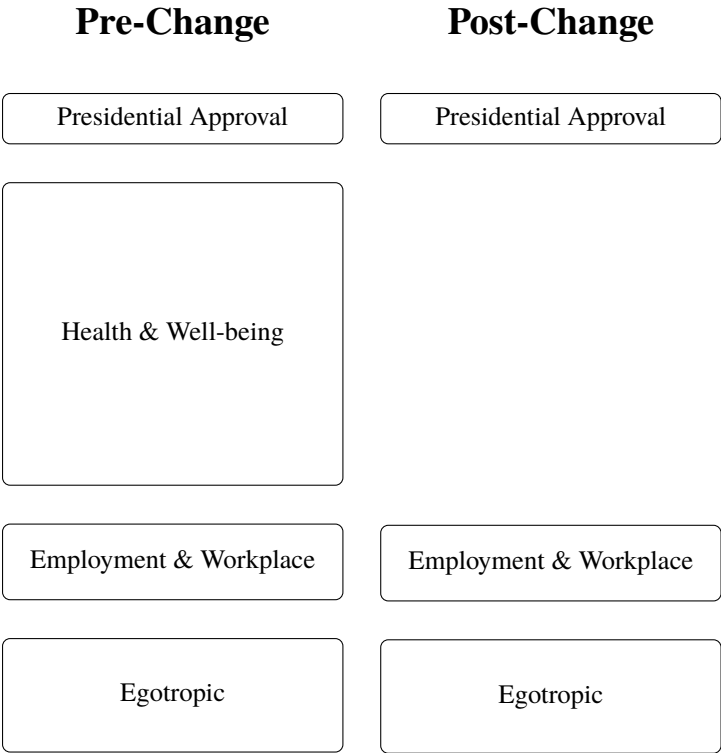| Employment & Workplace | | Employment & Workplace |

| Egotropic | | Egotropic |

Figure 6. Design change between 2012 and 2013. Health & well-being battery of questions was long, specific to the respondent, and wholly unrelated to politics. Without this buffer, we expect that the political considerations primed by the initial question on presidential approval more heavily influence the egotropic questions about individual economic conditions.

about physical and mental health (in line with H2).

Indeed, this appears to be the case, as illustrated in Figure 7 which plots the variable importance of partisanship on a number of egotropic questions pertaining to the respondent's economic situation. Vertical dashed lines indicate the January 1st, 2013 date when the sample was divided, and participants responding to these egotropic questions were no longer doing so as part of a larger battery of questions about their mental and physical well-being. As illustrated, across all egotropic questions there is clear evidence of a discontinuous increase in the predictive power of the respondent's partisanship after

January 1st, 2013. Indeed, Bayesian Change Point tests confirm the visual intuition, identifying this date as when there is a structural break in the time series measure of the importance of partisanship (see the Supporting Information).
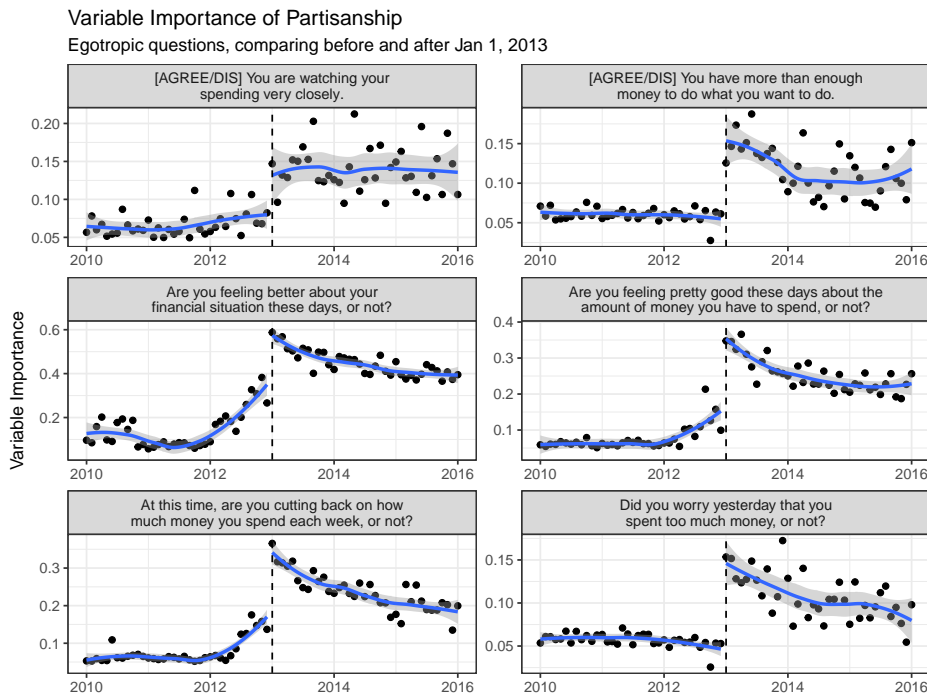

Figure 7.   Variable importance of party affiliation (from models estimated on a monthly basis) on a battery of egotropically-phrased questions about the respondent's economic condition. Loess smoothers fit separately prior to, and following, January 1st 2013 when Gallup split the survey into two samples.

Note, however, that (for a subset of items) there is also evidence of "anticipation", indicating that there might be more at play here than only the change in the survey design. One potential explanation might be the expiration of the Bush-era tax cuts which the Obama administration allowed to lapse at the beginning of 2013. Another might be the simpler argument that all public opinion becomes more partisan during presidential elections, such as the one that occurred in the fall of 2012. Indeed, descriptive evidence of the partisan gap suggests that the increase in variable importance is driven more

by Republican respondents, who grow significantly less optimistic about their personal finances after January 1st.

Nevertheless, we argue that the sharp discontinuity around this date is unlikely to solely reflect the expiration of the tax cuts, nor the increased salience of partisanship around the 2012 presidential election. While these events may have helped increase the predictive power of partisanship in the preceding months, across all facets the discontinuity between December and January is stark. We argue that these patterns are consistent with our theory about how the design of surveys, both in terms of the wording of specific questions as well as in the overall emphasis of a survey, can exaggerate partisan differences.

Conclusion

Implicit in the expansive body of empirical work that documents evidence of economic voting is that self-reported evaluations of the economy are causally prior to attitudes toward and support for the government. Yet in individual-level survey data, the sensitivity of these self-reported evaluations is well-known [CITES]. In this letter, we proposed two previously neglected dimensions of survey design that, we argue, exacerbate politically-biased responses to nominally apolitical questions. We argue that the "macro" nature of the question can facilitate partisan behavior (e.g. cheerleading or, conversely, faultfinding on surveys) by making it easier for respondents to engage in response substitution (linking their answers to a question about the economy to considerations of co- or out-partisan leadership). This extends prior work on the sensitivity of survey responses to question ordering, among other factors; see, e.g. Palmer and Duch (2001).[8] Similarly, we argue that embedding economic evaluations in surveys with more sociotropic questions,

---

[8]Separately, no evidence of partisan priming effects is found in Tyler and Iyengar (2023) in the context of thermometer questions for estimating affective polarization. In the latter experiment, participants who completed a survey largely devoid of political content before giving their thermometer scores showed similar levels of polarization to those who responded to an extensive set of politically charged evaluations with obvious partisan valence. Their outcome measure, however, was purely political, whereas our interest lies in the politicization of economic evaluations.

or questions that ask respondents to think like a partisan, also can inflate the association between partisanship and their appraisal of the economy.

By highlighting these two previously neglected dimensions, we hope to help future researchers design and field surveys that avoid introducing unwanted political bias. However, we do not provide an exhaustive list of question types that elicit stronger or weaker responses, nor do we provide a detailed how-to on what questions should and should not be mixed together in a survey. We hope applied researchers can use our findings to pilot surveys that test different question wordings and survey designs which fit their specific research question.

References

Ansolabehere, Stephen, M Meredith, and Erik Snowberg. 2014. "Mecro-Economic Voting: Local Information and Micro-Perceptions of the Macro-Economy." *Economics & Politics* 26 (3): 380–410.

Ballard-Rosa, Cameron, Amalie Jensen, and Kenneth Scheve. 2022. "Economic decline, social identity, and authoritarian values in the United States." *International Studies Quarterly* 66 (1): sqab027.

Bisbee, James, and Jan Zilinsky. 2023. "Geographic Boundaries and Local Economic Conditions Matter for Views of the Economy." *Political Analysis* 31 (2): 288–294.

Herron, Michael C. 1999. "Postestimation Uncertainty in Limited Dependent Variable Models." *Political Analysis* 8, no. 1 (January): 83–98.

Lewis-Beck, Michael S., and Mary Stegmaier. 2019. "247Economic Voting." In *The Oxford Handbook of Public Choice, Volume 1*. Oxford University Press, February. ISBN: 9780190469733. doi:10.1093/oxfordhb/9780190469733.013.12. eprint: https://academic.oup.com/book/0/chapter/347392345/chapter-ag-pdf/44434063/book_40414_section_347392345.ag.pdf. https://doi.org/10.1093/oxfordhb/9780190469733.013.12.

Mansfield, Edward D, and Diana C Mutz. 2009. "Support for Free Trade: Self-Interest, Sociotropic Politics, and Out-Group Anxiety." *Int. Org.* 63, no. 03 (July): 425.

Palmer, Harvey D, and Raymond M Duch. 2001. "Do surveys provide representative or whimsical assessments of the economy?" *Political Analysis* 9 (1): 58–77.

Rogers, Jonathan, and Marcelo Tyszler. 2018. "Information and economic voting." *Political science research and methods* 6 (4): 753–770.

Tyler, Matthew, and Shanto Iyengar. 2023. "Testing the Robustness of the ANES Feeling Thermometer Indicators of Affective Polarization." *American Political Science Review:* 1–7. doi:10.1017/S0003055423001302.

Zaller, John R. 1992. *The Nature and Origins of Mass Opinion.* Cambridge University Press.

Supporting information

Data Sources

We list the sources of our data (which we merged with the Gallup individual-level data) in Table S1 below. All sources are public with the exception of the Gallup data itself, which we obtained via our institution with geolocated respondents.

Variable	Source
Unemployment rate	BLS
Labor force participation rate	BLS
Total employment	QCEW
Wages	QCEW
Death rates	CDC
Life expectancy	IHME*
Crime / arrests	UCR (FBI)

*Note: *Institute for Health Metrics and Evaluation*

TABLE S1 *Contextual variables used as (potential) predictors of answers to questions about life, emotions, and the economy.*

Alternative Measures of Variable Importance

Our main results rely on permutation tests of variable importance using a random forest. Below, we confirm our findings are robust to alternative approaches to evaluating which measures are most prognostic, including LASSO regressions and expected percentage reduction in error (ePRE, Herron 1999), the latter of which is very similar in spirit to the permutation tests of our main results.

Starting with the LASSO approach, we visualize two questions from 2016 in Figure S1. In the left panel, we demonstrate that the most prognostic covariates of an individual's satisfaction with their standard of living include their income, marital status, educational attainment, and age. In the right panel, we find that their views of the national economy are far more strongly associated with their partisanship.

How do these results generalize across all outcomes and all periods? Figure S2 summarizes the LASSO results by indicating the λ penalty at which each predictor is included, and highlighting the highest predictor with solid black borders. As illustrated, partisanship is the most important predictor for questions about the trend of the national economy, the current state of the national economy, and then approval for presidents Obama and Trump, as well as Hillary Clinton's favorability. As we move down the y-axis toward more egotropically-framed outcomes, we find weaker predictive power of partisanship, replaced by age, income, and marital status.

An alternative approach to characterizing the prognostic power of a variable is to evaluate how much better we are at predicting an outcome when we add partisanship to a regression model. The simplest version of this is to compare a naive model that simply predicts the modal outcome category to a logistic regression of the outcome on an indicator for whether the respondent is a co-partisan of the president. We summarize the improvement in predictive accuracy using the expected percentage reduction in error (ePRE, Herron 1999) and visualize the results in Figure S3, which support our substantive argument that sociotropically framed questions are more sensitive to partisan motivated responding.
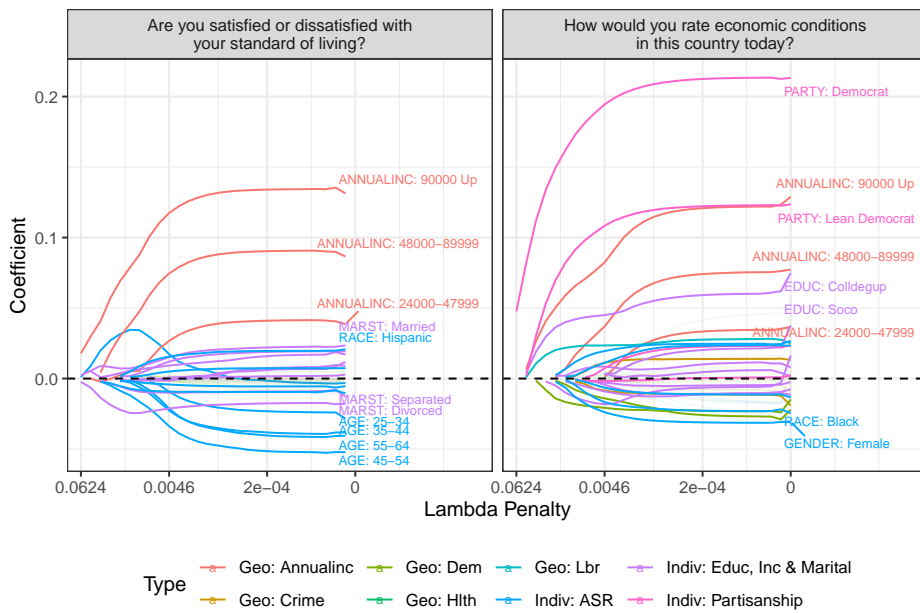
LASSO Results

Figure S1.   Estimated coefficients (y-axes) across different levels of L1 norm penalties (x-axes) for views on the economy (right panel) and self-reported satisfaction with one's standard of living (left panel). Labeled opaque values are consistently included in optimal model across 100 bootstrapped simulations. Data: 2016 Gallup responses augmented with contextual data (see Table S1).
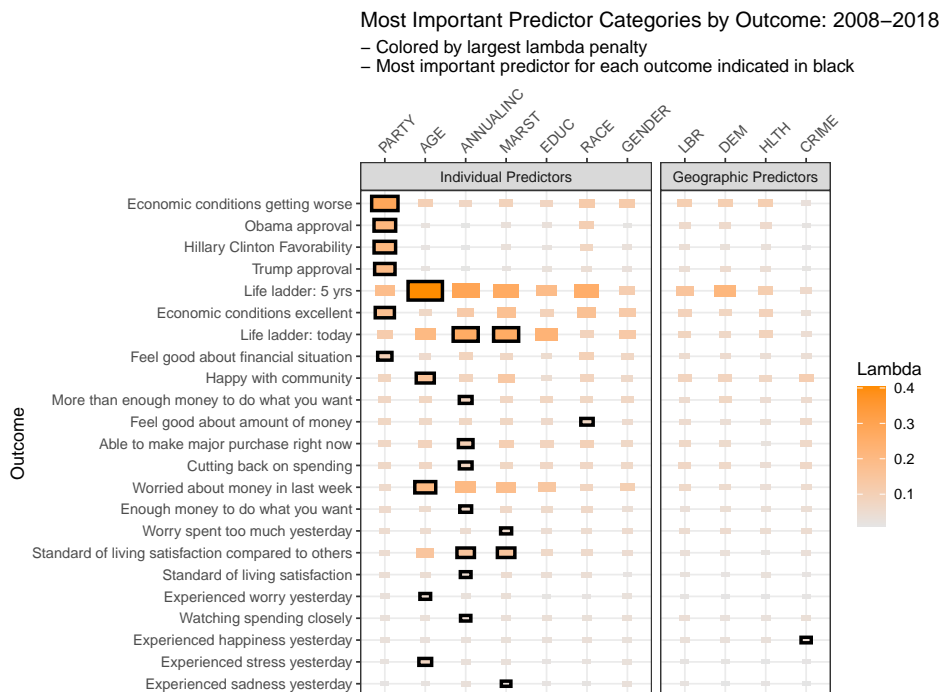
Most Important Predictor Categories by Outcome: 2008–2018
– Colored by largest lambda penalty
– Most important predictor for each outcome indicated in black

Figure S2. λ *penalties at which different predictors (x-axis, top) are included in the LASSO regression for different outcomes (y-axis) across all years of the data. The largest* λ *penalty for each outcome is outlined in black (note that the coarseness of the grid for possible* λ *penalties means that some predictors are tied for importance for some outcomes).*
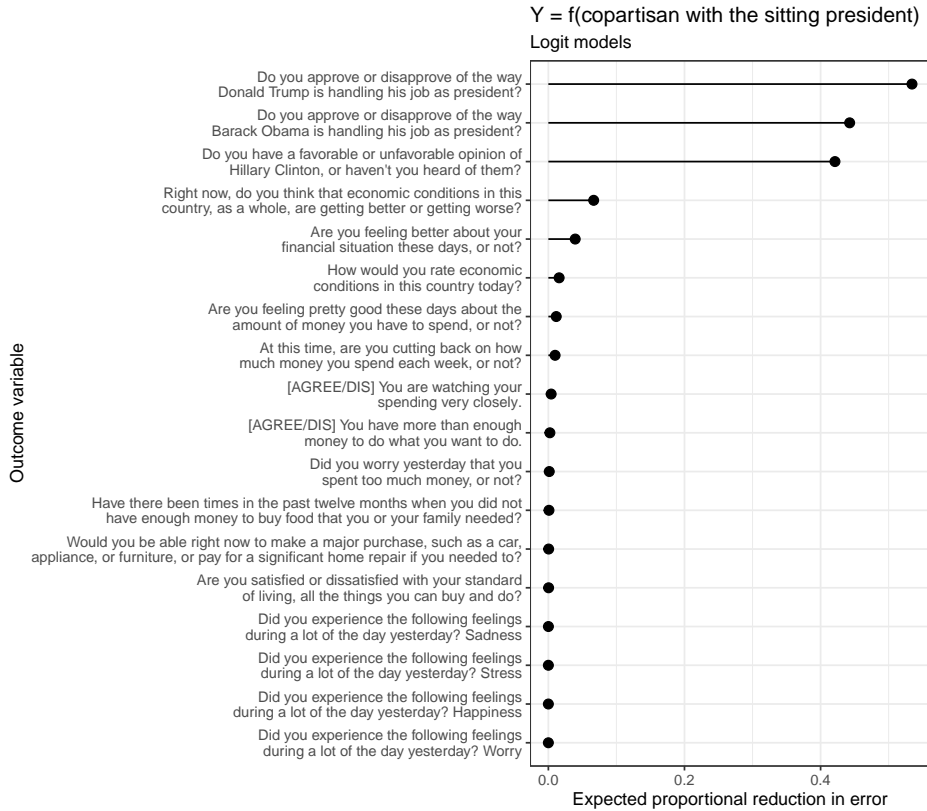
Figure S3.   Expected percentage reduction in error (ePRE, Herron 1999) (x-axis) associated with the inclusion of an indicator for co-partisanship with the president, across binary survey questions (y-axis).

Extended Analysis of Survey Design Change

Our main analyses focused on the change in the variable importance measure of partisanship just prior to, and just following, the change in Gallup's survey design on January 1st, 2013. We present several extensions here.

First, we evaluate these results subject to finer-grained temporal units, including weeks and days. As illustrated in Figures S4 and S5, the substantive conclusions hold, albeit with increasing noise as we disaggregate to smaller temporal units.

Variable Importance of Partisanship
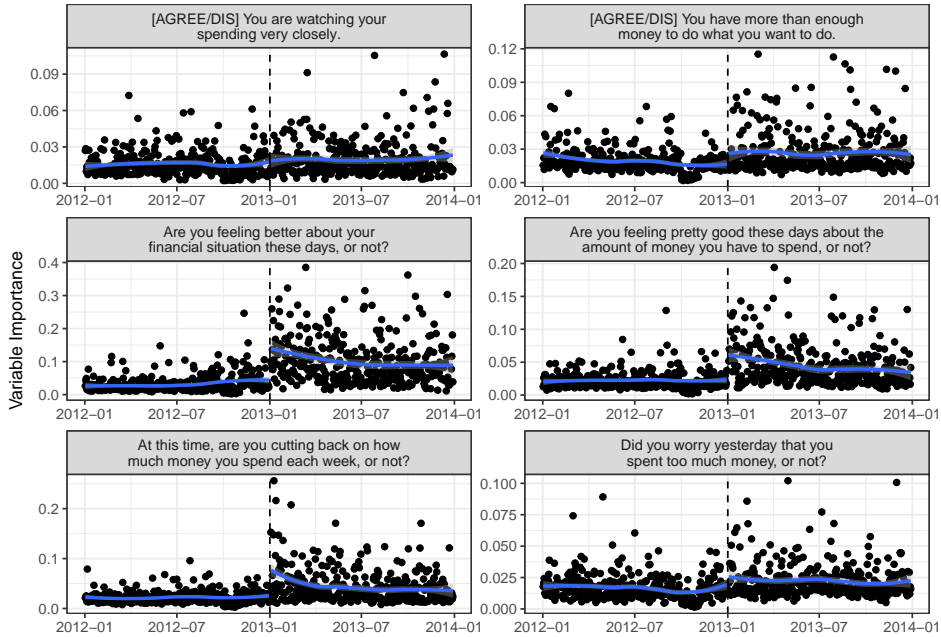Egotropic questions, comparing before and after Jan 1, 2013
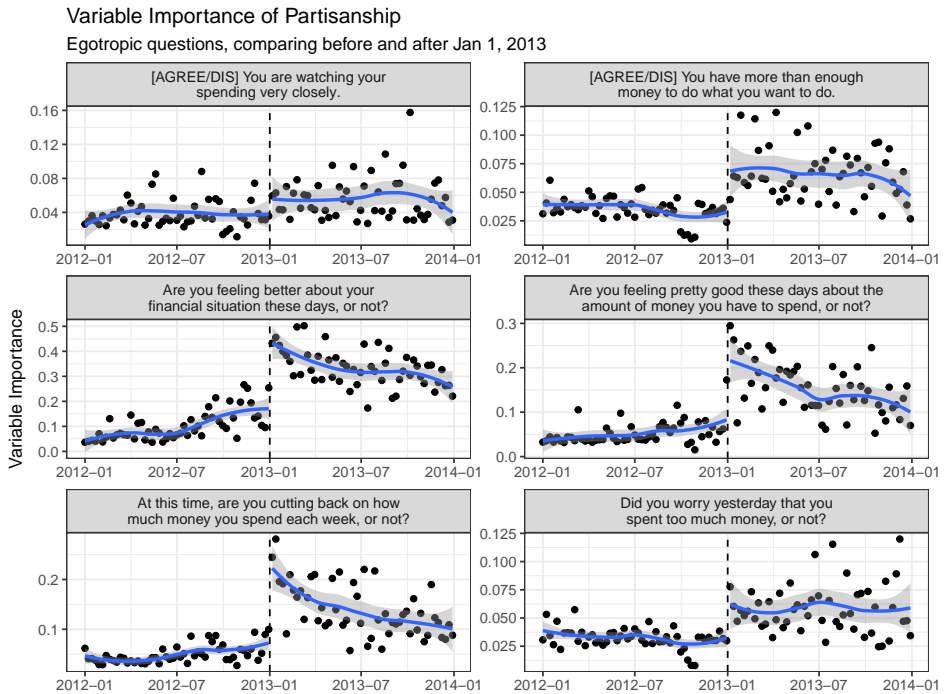
Figure S4. Variable importance of party affiliation (from models estimated on a monthly basis) on a battery of egotropically-phrased questions about the respondent's economic condition. Loess smoothers fit separately prior to, and following, January 1st 2013 when Gallup split the survey into two samples. Data aggregated to week level.

Second, we apply Bayesian Change Point Analysis (BCP) to let the data inform if

Figure S5. *Variable importance of party affiliation (from models estimated on a monthly basis) on a battery of egotropically-phrased questions about the respondent's economic condition. Loess smoothers fit separately prior to, and following, January 1st 2013 when Gallup split the survey into two samples. Data aggregated to day level.*

and when a discontinuous break appears in these data. As illustrated in Figures S6, S7, and S8, the analysis consistently chooses the period on or around January 1st, 2013 across different outcomes, regardless of whether we aggregate to the month, week, or day.
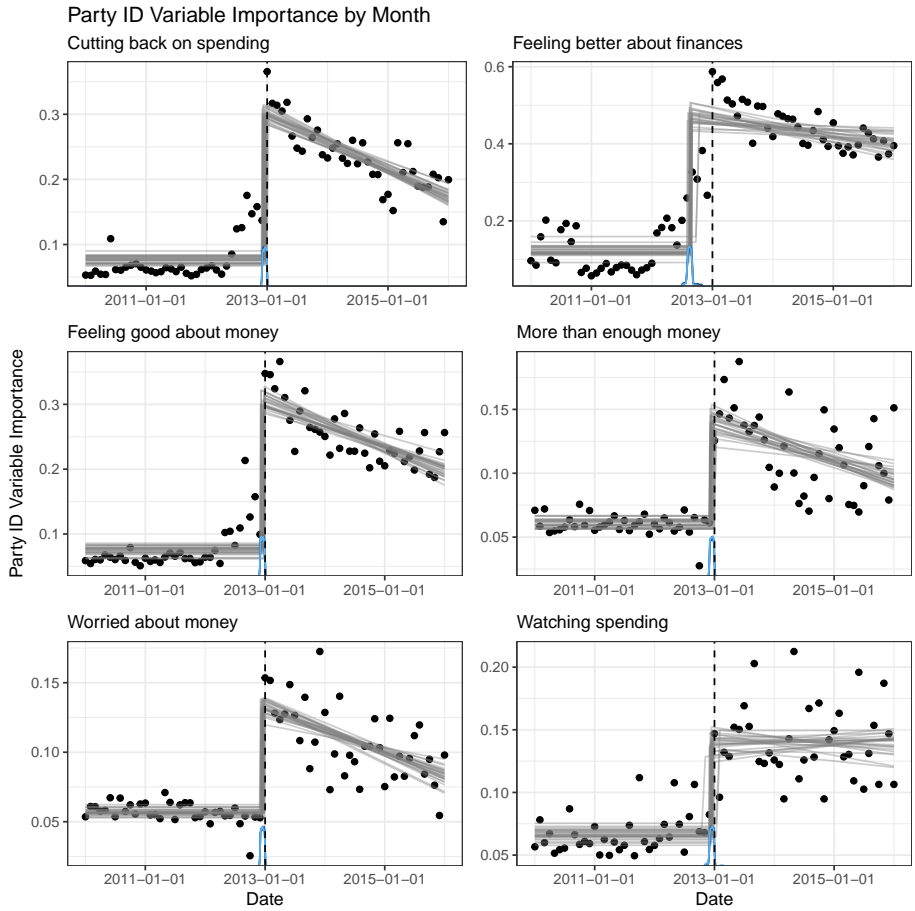
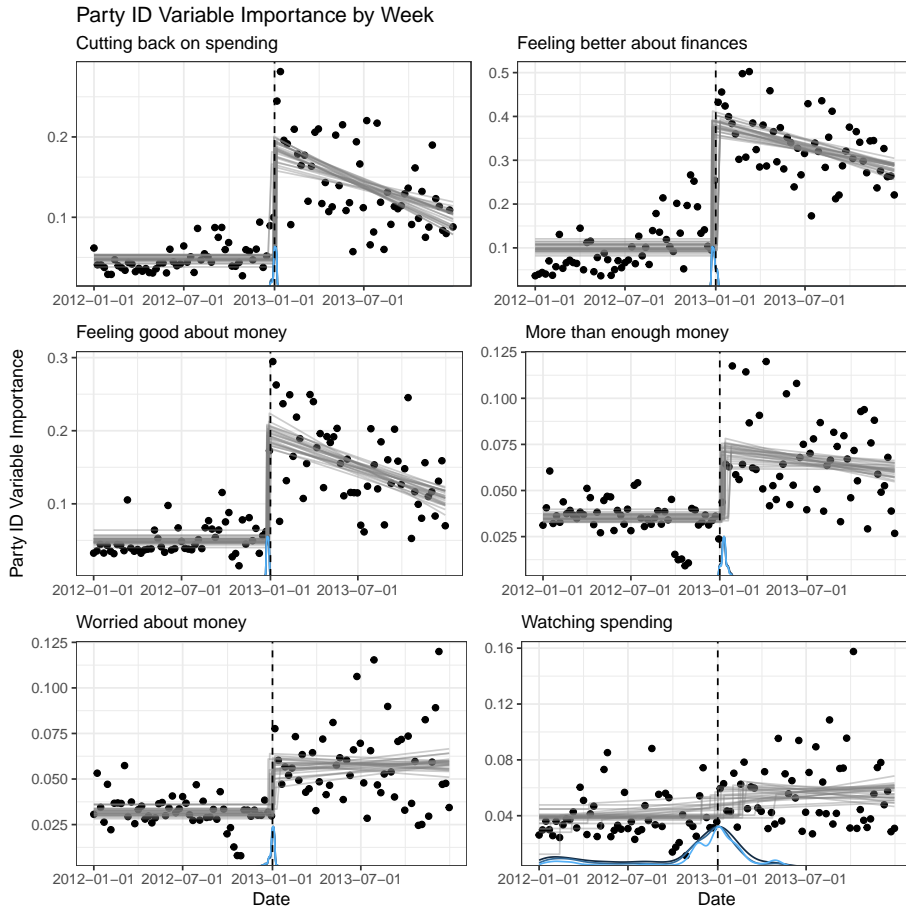Figure S6. *Bayesian change point detection, data aggregated to months.*

Figure S7.   Bayesian change point detection, data aggregated to weeks.
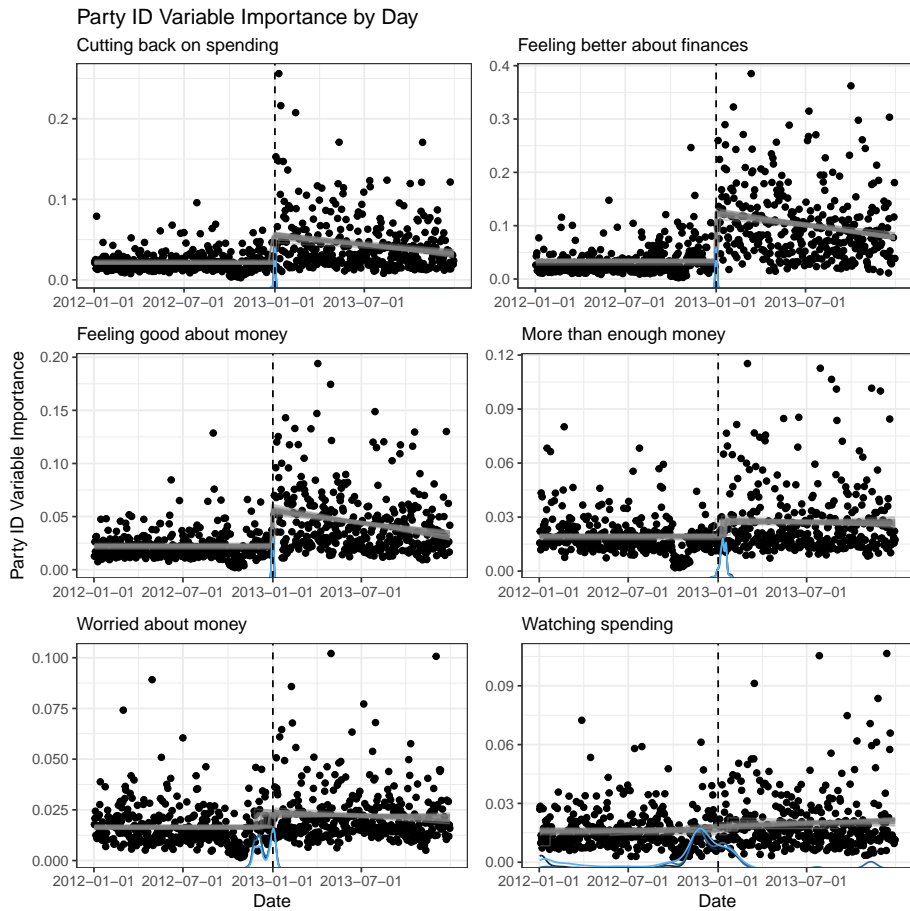
Party ID Variable Importance by Day

Figure S8.   Bayesian change point detection, data aggregated to days.