Political Surveys Bias Self-Reported Economic Perceptions*

Jack Bailey†

Abstract

If voters are to hold governments to account for the state of the economy, they must know how it has changed. Indeed, this is a prerequisite for democratic accountability. Yet the perceptions that voters report often show signs of clear partisan bias. At present, we do not know if this bias is real or instead due to priming in political surveys. To test this, I assign subjects at random to either a political or non-political survey. I then record their economic perceptions and compare the results for each group. I show that political surveys do worsen partisan bias, though only among supporters of the incumbent party. Still, much partisan bias remains unexplained, even in the non-political condition. So, while economic perception items remain biased, we can at least be sure that most people respond to them in a similar way no matter the survey context.

Introduction

To hold governments to account for the state of the economy, voters must first know how it has changed. Indeed, this is an essential prerequisite for democratic accountability (Healy and Malhotra 2013; Ashworth 2012). Thus, voters should notice that conditions improve when the economy grows and worsen when it shrinks. Just as this variation in perception is important, so too are its consequences. If voters are to reward and punish appropriately, then they should be more likely to support the incumbent where they also think that the economy has improved. This two-step process — first, of economic updating; second, of electoral sanctioning — is crucial for good governance. Rather than force voters to suffer fools, it lets them "kick the rascals out" if they fail to live up to expectations (Stegmaier, Lewis-Beck, and Brown 2019).

Though this idea has great normative appeal, reality often falls short. This is because voters are not so dispassionate when it comes to judging economic management. Instead, all manner of considerations influence the decisions they make. For instance, voters tend to rely on pre-existing beliefs when processing new information. Evidence of this behavior is rife in political surveys, where respondents often report economic perceptions that show clear signs of partisan bias: those who support the incumbent tend to be more positive, and those who support the opposition more negative, than those who support no party at all (for recent evidence of this phenomenon, see Bailey 2019; Bisgaard 2019; De Vries, Hobolt, and Tilley 2018).

Given the potential ramifications, much work now focuses on mitigating this bias. Yet we still do not know if it is meaningful or, instead, a result of partisan priming in political surveys. In test this possibility in this short research note. To do so, I use new survey experimental data from the 2019 UK General Election campaign.

I find that political surveys worsen partisan bias in voters' self-reported economic perceptions. But this is true only for those who voted for the incumbent at the last election. What's more, much partisan bias remains unexplained. Thus, while economic perception items are far from perfect, survey researchers and economic voting scholars can at least be sure that most people respond to them in a similar way no matter the survey context.

Economic Perceptions, Partisan Bias, and Political Surveys

Party identification biases the economic perceptions that voters report (see, for example, De Vries, Hobolt, and Tilley 2018; Bartels 2002; Conover, Feldman, and Knight 1987). What's more, this bias serves to undermine accountability mechanisms that make democracy possible (Healy and Malhotra 2013; Anderson 2007). The reason for this is simple. Partisan voters report economic perceptions that paint their party in a positive light. For example, incumbent supporters tend to report more positive economic perceptions. Opposition supporters, instead, tend to report more negative ones. Thus, we cannot be sure that voters will hold their party to account on the basis of economic management when it gains power.

While we know that party identification affects reported economic perceptions, we are less certain why this is the case. Clearly, this is an important gap in our understanding. If we do not know what causes partisan bias, we cannot hope to mitigate its worst effects. To this end, the literature on economic perceptions offers four competing hypotheses. Yet, though they are competing, they are not exclusive. Rather, all four likely influence the economic perceptions that voters report to some extent or another. I discuss each below in turn.

The first potential cause of partisan bias is consistency-motivated reasoning (Kunda 1990). This holds that voters weight new information based on how congruent it is with their existing beliefs (Hill 2017). This behavior weakens voters' ability to hold parties to account. Still, it is easy to see why it might play a role in their economic calculus. Existing research shows that the economy shapes wider perceptions of party competence (Green and Jennings 2017). And this can lead to much psychological discomfort for partisan voters (Groenendyk 2013). Consider an incumbent supporter during the middle of an economic downturn. Such a voter must contend with two conflicting beliefs: that the economy is doing badly and that their own party is managing the economy well. Consistency-motivated reasoning offers them a way out of this predicament. By down-weighting incongruent information, they can either ignore evidence that the economy is doing badly or, instead, update their perceptions based only on information that pins the blame on someone else (Bisgaard 2019, 2015; Tilley and Hobolt 2011).

The second potential cause is partisan cuing (Brady and Sniderman 1985). Like consistency-motivated reasoning, this too is psychological in nature. It suggests that voters make use of cognitive shortcuts. This is necessary because, as have long known, many voters pay little attention to politics (Campbell et al. 1960). As such, they have a hard time when it comes to making decisions about political matters. The partisan cuing literature argues that they resolve this problem by making a simple substitution. Rather than derive their own belief, they rely on their favorite party's position on the issue instead. They may do this either out of party loyalty or the belief that they would have come to the same conclusion were they fully-informed (Ramirez and Erickson 2014; Brader, Tucker, and Duell 2013). Evidence in favor of partisan cuing is most striking where it concerns party elites. Bisgaard and Slothuus (2018), for instance, shows that when the Danish government began to consider the budget deficit in a negative light, its supporters came to do so too despite not having done so just a short time before.

The third potential cause is expressive responding (Schaffner and Luks 2018; Bullock et al. 2015). Unlike the preceding two explanations, it does not rely on voter psychology to explain partisan bias. Instead, it contends that survey respondents use survey items to signal their support for a particular party. For example, a respondent might report that the economy has gotten better not because they believe it to be so, but because they support the incumbent party. Recent survey experimental evidence shows that expressive responding almost certainly occurs. Though concerned with factual questions, Bullock et al. (2015) and Prior, Sood, and Khanna (2015) run similar experiments where they manipulate the incentive to engage in expressive responding. Respondents in the treatment groups received a small cash reward where they admitted either to not knowing the answer or happened to give the correct answer to a series of factual questions about the economy and other policy-related topics. Respondents in the control groups received no such reward. In both cases, the authors find that partisan disagreement was lower under the treatment than under the control, implying that some responses serve only to signal respondents' party preferences.

The fourth and final potential cause is item-order effects. These occur where the order in which survey questions are asked affects the answers that respondents give. If non-political

items precede political ones, they may *personalize* respondents' answers. Likewise, where political items precede non-political ones, they may *politicize* them instead (Sears and Lau 1983). Such item-order effects can be both large and long-lasting. Indeed, even where several buffer items separate them, political questions still come to bias the economic perceptions that respondents report (Wilcox and Wlezien 1993). Further, as many electoral surveys begin by asking their respondents how they voted or how they intend to vote, this politicization of economic perception items is probably relatively common.

Though distinct, all four causes share a common catalyst: the political survey context. That is to say, partisan priming in political surveys might worsen their effect. For the sake of illustration, consider expressive responding. If the survey context implies that the survey administrator does not care about politics, then respondents also face fewer incentives to engage in partisan cheer-leading. Likewise, consider motivated reasoning and partisan cuing. If the survey context does not encourage respondents to consider the economy through a partisan lens, then it seems reasonable to expect them to be less likely to rely on partisanship to determine what they think about the state of the economy. It is for this reason that consumer most confidence surveys rarely ask about respondents' party affiliation (Curtin 2019).

As a result, the political survey context *itself* might moderate how party identification affects the economic perceptions that voters report. And, given that partisan bias varies direction based on party identification, so too should political survey effects. Thus, we should expect incumbent supporters to be *more* likely to report positive and *less* likely to report negative economic perceptions in political compared to non-political surveys. We should expect opposition supporters, instead, to do the opposite. This implies the two following hypotheses:

Hypothesis 1: Incumbent partisans will be more likely to choose *positive* answers to economic perception items in a political survey than similar incumbent partisans in a non-political survey.

Hypothesis 2: Opposition partisans will be more likely to choose *negative* answers to economic perception items in a political survey than similar opposition partisans in a non-political survey.

Experimental Design

I use a simple survey experiment to test my hypotheses. The market research and polling company YouGov collected the corresponding data from its panel of eligible British voters[1]. Further, data collection occurred between the 6th and the 8th November 2019. This was a deliberate choice and ensured that data collection coincided with the beginning of the campaign for the 2019 UK General Election.

The British case is especially useful and provides a strong test of my argument for two reasons. First, the data collection process coincided with the start of the 2019 UK General Election campaign. This implies that my subjects were exposed to a general politicization of the information environment. As a result, we might expect participants to bias their responses in *non-political* surveys too. Thus, any differences between my treatment group and my control group will likely be conservative. Second, data collection occurred at a time of economic uncertainty. Though the economy was not in recession, it was not growing much either. The most recent GDP data at the time showed that the UK economy had contracted by 0.2% in the previous quarter. This is important as new evidence shows that even strong partisans "get it" when the going gets tough (Bisgaard 2019; De Vries, Hobolt, and Tilley 2018) and that this leads partisan bias to diminish (Bailey 2019; Stanig 2013). As such, it seems reasonable to expect the economic circumstances at the time to provide less partisan bias for my treatment to manipulate.

In the first stage of the experiment, I drew a blocked sample from YouGov's online panel[2]. The first blocked contained only those panelists who had voted for the incumbent Conservative Party at the last election in 2017, the second only those who had voted for an opposition party, and the third only those who had not voted at all[3]. To determine my sample size, I conducted

---

[1]Note that while YouGov uses non-probability sampling, its data are not from a convenience sample. Rather the company ensures that its data are nationally-representative using a method it calls "active sampling" (Twyman 2008). This approach has proven robust and the company's surveys often yield results substantively similar to those collected using random probability sampling (Sanders et al. 2007).

[2]This design is deliberately non-representative as it ensures that I am just as able to detect a treatment effect for incumbent voters as I am for opposition voters or non-voters. As such, I do not weight my data. Regardless, this is likely to make little difference. As Miratrix et al. (2018) show, also using YouGov data, for "high-quality, broadly representative samples recruited by top online survey firms, sample quantities, which do not rely on weights, are often sufficient" (p. 275).

[3]Retention was high. Just 3.5% (91) of respondents failed to finish the survey. Of these, 48 left before being assigned to a condition, 18 left after being assigned to the treatment, and 25 left after being assigned to the control.

a simulation-based power analysis. The results from 6,000 simulated experiments showed that I would need a sample of around 2,500 respondents to reach a power level of 80%[4].

I do not have my participants report their voting behavior during the experiment, but rely instead on contemporaneous data that YouGov collected after the 2017 election. As such, misreporting bias or other related issues should be low. Some might argue that it would be better to use participants' *current* party identification and not how they voted in the past. After all, attitudes and choices change over time. While this is a reasonable objection, it is not possible to include such an item without undermining the non-political survey context. Further, using past voting behavior has one particular advantage: voters cannot undo it. This may explain why it appears to exert such a considerable effect on the economic perceptions that voters report in political surveys (Anderson, Mendes, and Tverdova 2004).

In the second stage of the experiment, I exploited YouGov's day-to-day operations to administer my treatment. As a large commercial polling company, YouGov runs many simultaneous political and non-political surveys. It also runs them in tandem. As a result, panelists are used to surveys that concern one topic then switch to another. My treatment group first completed a version of YouGov's standard voting intention poll. This includes five questions that concern voting behavior and the perception of party leaders. The control group, instead, completed a survey on dental hygiene. This had an almost identical structure to the political survey. For example, it asked the same number of questions, the same type of questions, and included the same number of response options in all cases. Further, it also used only questions that YouGov had fielded in the past to ensure that it was believable[5]. In all cases, participants had an equal chance of being assigned to the treatment or to the control.

In the third and final stage of the experiment, I again exploited YouGov's day-to-day operations, this time to measure my participants' economic perceptions. After receiving their treatment, both groups saw the topic of the survey switch from politics or dental hygiene to the economy. I then asked them to report their own retrospective economic and financial perceptions. As I used a sample of eligible British voters, I followed the lead of the British

---

[4]For more information, see appendix A.
[5]A full questionnaire is available in appendix B.

Election Study Internet Panel (Fieldhouse, Green, Evans, Mellon, Prosser, et al. 2020) and had my participants answer the two following questions:

- Now, a few questions about economic conditions. How does the *financial situation of your household* now compare with what it was 12 months ago?

- How do you think the *general economic situation in this country* has changed over the *last 12 months*?

These items have their origins in consumer confidence surveys (Katona 1951), entered political science via *The American Voter* (Campbell et al. 1960), and are now ubiquitous in economic voting research (Lewis-Beck and Stegmaier 2007)[6]. By and large, the literature on economic perceptions and partisan bias focuses only on national-level items (for recent examples, see Dassonneville and Lewis-Beck 2019; Anson 2017; Hansford and Gomez 2015). While I do include this item, I also asked my subjects to report their personal financial perceptions too. Doing so serves a useful purpose: it provides both a benchmark for any national-level effects and helps to prevent an unusual one-question-long topic.

In both cases, my subjects faced the exact same response options. They could answer each question on a five-point Likert-type scale that ranged from *"1 – Got a lot worse"* to *"5 – Got a lot better"*. They could also report that they did not know how either the national or their own personal economic situation compared to what it was 12 months ago. Where this was the case, I removed these participants using list-wise deletion[7].

Figure 1 shows the raw percentages for each response option stratified by party and treatment status. As we would expect, these show that incumbent partisans are more positive than do nonvoters. Further, this is true under both the treatment and the control. For example, 17.6% of incumbent partisans in the treatment condition (a political survey) said that the

---

[6]Note that the consumer confidence surveys from which these items originate rarely field questions of partisanship as they are known to engender emotional states that bias how survey respondents answer economic perception questions (Curtin 2019)

[7]List-wise deletion can produce biased estimates if data are not missing completely at random. As I estimate treatment effects on each specific response category, this should not be an issue. Still, simulation studies show that list-wise deletion yields less biased estimates than multiple imputation where data are missing not at random (Pepinsky 2018). Even so, I include these data as a robustness check (see appendix D). This does not change my results. Further, participants were no more likely to answer "Don't know" under the treatment than the control.
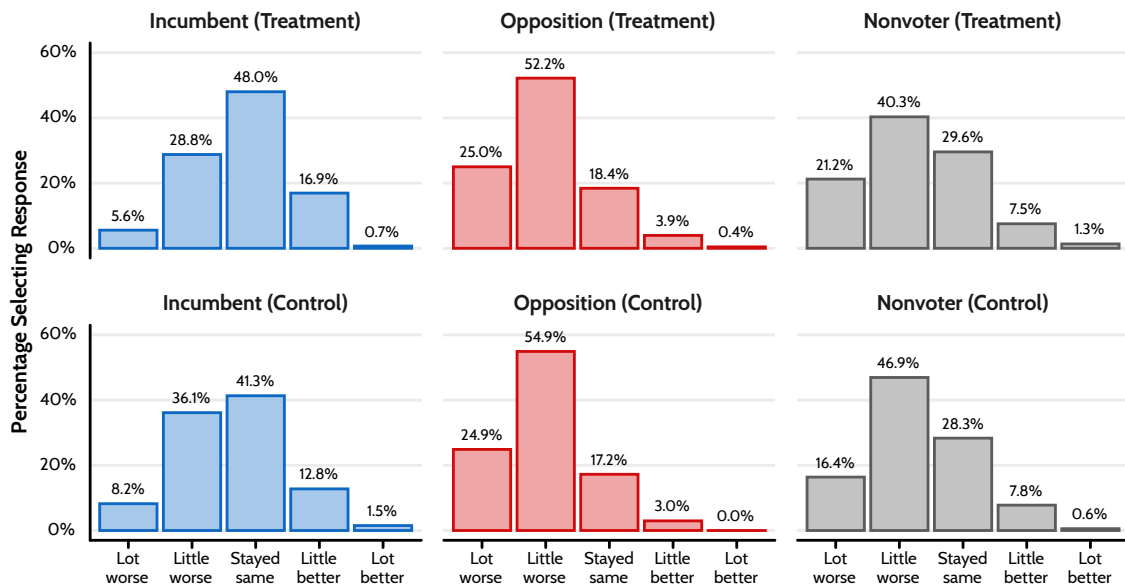
Figure 1: Descriptive figures show that incumbent partisans (left column) tend to be more positive than nonvoters (right column) under both the treatment and control. Likewise, opposition partisans (middle column) tend to be more negative than nonvoters. Further, these figures also suggest evidence in favour of my first hypothesis that incumbent partisans in the treatment would be more positive than incumbent partisans in the control, though not my second hypothesis that opposition partisans in the treatment would be more negative than opposition partisans in the control.

economy had gotten a lot or a little better while only 8.9% of nonvoters in the treatment condition said the same. Likewise, opposition partisans are more negative than nonvoters. Among opposition partisans in the treatment condition, 77.2% said that the economy had got a lot or a little worse whereas 61.6% of similar nonvoters made the same judgment.

Though still descriptive, figure 1 also suggests evidence in favor of my first hypothesis. Incumbent partisans in the political survey treatment condition were 3.3% more likely to say that things had gotten better than similar incumbent partisans in the control condition. They were also -10.0% less likely to say that the economy had gotten worse. Yet these data suggest little evidence in favor of my second hypothesis. Indeed, opposition partisans in the political survey treatment condition were 1.4% more likely to say that things had gotten better and 2.6% less likely to say that things had gotten worse than opposition partisans in the non-political control condition. While informative, any inferences that we make from these descriptive statistics do not account for the uncertainty inherent in the sample. To do so requires a more rigorous approach, which I describe in greater detail below.

Estimating Response-Specific Treatment Effects

Economic perception items yield ordinal data. Yet many researchers treat them as continuous. This is convenient, as it allows them to estimate treatment effects using only a simple comparison of means. But this simplicity belies drawbacks that include false positives, false negatives, and even estimates with incorrect signs (Liddell and Kruschke 2018).

One argument for treating these items as continuous is that while the outcome is ordinal, subgroup means and differences are continuous. This is true. But it is not clear what such treatment effects even imply. Indeed, when ordinal variables have three or more response options, there are an *infinite* combination of response distributions that could produce any given difference in means.

Better then to estimate treatment effects for the choices that survey respondents really face: the ordinal variable's response options. To do so, one might expect to use ordered regression. But these models face similar problems. Figure 2 shows why. Ordered regression treats the ordinal distribution that we observe (bottom row) as a function of a continuous one that we do not (top row). It then uses a set of threshold parameters (gray dotted lines) to convert between the two. These divide the latent continuous distribution into as many segments as there are response options. The area between two thresholds then gives the probability of each response occurring.

The first column shows the baseline case. Here, each response has an equal probability. To change this, we can adjust either the latent distribution's mean or variance. This has three consequences. When we adjust the mean, the latent distribution *shifts* up or down the scale (second column). This alters the area between the thresholds and moves the ordinal distribution in the same direction. When we instead adjust the variance, the latent distribution either *compresses*, squeezing the ordinal distribution's probability mass (third column), or *disperses*, piling up probability mass at the extremes (fourth panel).

As figure 2 shows, compression and dispersion can have large effects on the ordinal distribution. Yet conventional ordered regression accounts only for shift. This is a problem, as treatments may affect the outcome without shifting the probability mass to one end or
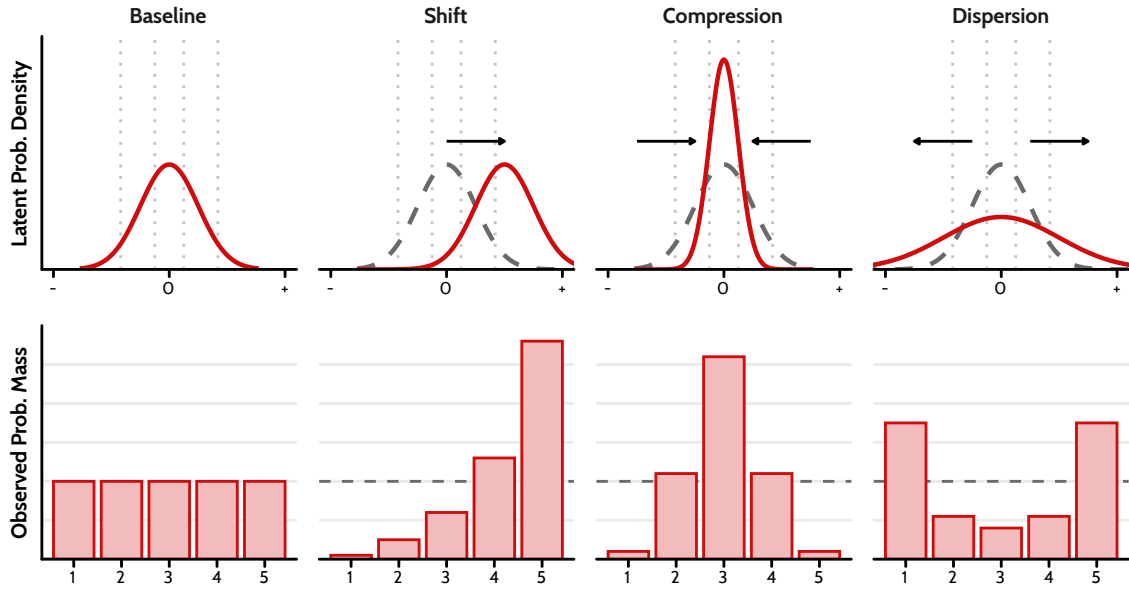
Figure 2: Ordered regression assumes that the observed ordinal scale is a function of a latent continuous one. Imagine that each column shows a different treatment group and that the baseline group is the control. Participants have an equal probability of selecting each response. When the the latent scale *shifts*, so too does the probability of selecting a higher value on the observed scale. Likewise, when it *compresses* or *disperses*, the observed scale follows suit. As we can see, each may have a large effect. Yet most ordered regression models account only for shift.

the other. Dealing with this is difficult using Frequentist methods. Thus, in line with recent recommendations (Liddell and Kruschke 2018), I use Bayesian methods instead[8].

My model is as follows. Let E_i be person i's reported retrospective economic perceptions. In line with existing economic voting research, this item is measured on a five-point Likert-type scale as described above and which takes a value that varies between 1 = "Got a lot worse" and 5 = "Got a lot better". In order to model the data as ordinal, we assume that the observed ordered variable, E_i, is a function of some latent continuous variable, E_i^*. We then assume that this latent continuous variable follows a normal distribution with mean, μ_i, and standard deviation, σ_i:

$$E_i^* \sim \text{Normal}(\mu_i, \sigma_i)$$

Likewise, the observed ordinal outcome variable, E_i, takes a particular value as follows:

---

[8]Note that Bayesian models require prior distributions. In this case, I specify a set of conservative and weakly informative priors for each parameter. I discuss my choices in greater detail in appendix C.

$$E_i = k \text{ if } \tau_{k-1} \leq E_i^* \leq \tau_k \text{ for } k = 1, ..., K$$

Here, τ_k for $k = 0, ..., K$ represent threshold parameters which segment the latent continuous distribution. We fix the 0th and k^{th} thresholds equal to $-\infty$ and $+\infty$, such that $-\infty = \tau_0 < \tau_1 < ... < \tau_{k-1} < \tau_k = \infty$. As such, the probability that $E_i = k$ is:

$$\Pr(E_i = k) = \Phi(\frac{\tau_k - \mu_i}{\sigma_i}) - \Phi(\frac{\tau_{k-1} - \mu_i}{\sigma_i})$$

Where Φ is the cumulative distribution function of the normal distribution with mean μ_i and standard deviation σ_i. As I discuss above, both influence the ordinal distribution that we observe. Likewise, both may also vary according either to party preference or treatment status:

$$\mu_i = \beta_1 T_i + \beta_2 I_i + \beta_3 O_i + \beta_4(T_i \times I_i) + \beta_5(T_i \times O_i)$$

$$log(^1\!/\sigma_i) = \delta_1 T_i + \delta_2 I_i + \delta_3 O_i + \delta_4(T_i \times I_i) + \delta_5(T_i \times O_i)$$

Here, T_i takes the value 1 where person i is in the treatment group. Likewise, I_i and O_i take the value 1 where person i voted for the incumbent or an opposition party at the last election, respectively. Rather than model σ_i, I instead model $log(^1\!/\sigma_i)$. Further, to identify the model, I fix σ_i to 1 for the baseline category (non-voters).

Though somewhat complex, this method is robust to the problems I discuss above. Still, like any ordered regression model, it produces parameters that are hard to interpret. Fortunately, as Bayesian models are generative (Lambert 2018) we can have them estimate treatment effects on the more intuitive probability scale while also incorporating any inherent uncertainty. I do this below, and compute my treatment effects for each response category as follows[9]:

$$\text{ATE}_k = \Pr(E_i = k | T_i = 1) - \Pr(E_i = k \mid T_i = 0)$$

---

[9]As it is much easier to interpret treatment effects measured in percentage point terms than as log-odds, and because space is limited, I save any regression tables for the appendix (see tables A2–A5).
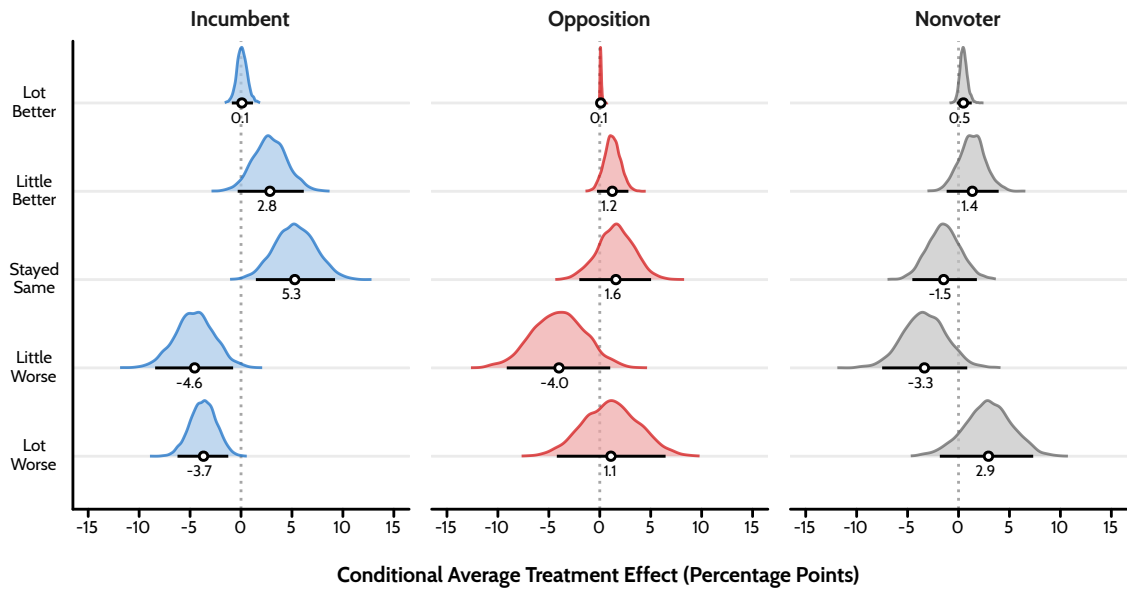
Results

Figure 3: Political surveys cause incumbent voters to report different economic perceptions (left panel). To this end, they were less likely to say that the economy had either "got a little worse" or "got a lot worse". They were also more likely to say that the economy had "stayed the same" or that it had "got a little better". Note that the treatment had little effect on reporting that the economy had "got a lot better" for any group and also that these estimates were very precise. This is because almost no one said that the economy had "got a lot better" in either group. Here, density plots show the posterior distribution of conditional average treatment effects. Further, black bars show their 95% credible intervals and point estimates their medians.

In line with my expectations, my results show that political surveys *do* affect the economic perceptions that respondents report in political surveys. Yet, these political survey treatment effects are limited only to those respondents who voted for the incumbent Conservative Party at the last general election in 2017[10]. Figure 3 shows these estimates in graphical form. Here, density plots show the posterior distribution of each treatment effect, black bars their 95% credible intervals, and point estimates their median. Each density curve reflects the difference in the probability of reporting a given response under the treatment versus the control. Thus, a positive value implies that the political survey treatment increased the probability of a respondent picking a given response by a given number of percentage points compared to similar respondents in the non-political survey control.

---

[10]This finding is robust to a range of tests. See appendix D.

The left-most panel shows how the treatment affected those who voted for the incumbent party in 2017. In general, these effects are in line with my expectation that incumbent-supporting subjects would be more positive in the political survey. In the treatment group, these respondents were 3.7 percentage points (95% CI: 1.3 to 6.2) less likely to say that the economy "got a lot worse" over the past twelve months and 4.6 percentage points (95% CI: 0.8 to 8.4) less likely to say that it "got a little worse". In comparison, they were 2.8 percentage points (95% CI: -0.3 to 6.2) *more* likely to say that the economy "got a little better".

Interestingly, these subjects were no more likely to say that the economy "got a lot better" (0.1, 95% CI: -0.9 to 1.2). This effect was also much more precise than for other responses. Though this may seem unusual, it arises only because almost no one reported that the economy "got a lot better". This is not uncommon, at least in the British case, even when the economy is booming (see Bailey 2019). Finally, those reporting that the economy "stayed the same" made up the difference. These participants were 5.3 percentage points (95% CI: 1.5 to 9.2) more likely to pick this option under the treatment compared to the control.

Contrary to my expectations, the effect of taking a political survey was less clear where participants voted for an opposition party at the last election. These subjects were not much more likely to say that the economy "got a lot better" (0.1, 95% CI: 0.0 to 0.3), "got a little better" (1.2, 95% CI: -0.3 to 2.8), or "stayed the same" (1.6, 95% CI: -2.0 to 5.0) where they took the political survey treatment. And, while they were 1.1 percentage points (95% CI: -4.2 to 6.5) more likely to say that the economy "got a lot worse", they were in fact 4.0 percentage points (95% CI: 0.3 to 9.1) *less* likely to say that it "got a little worse". Evidence of any treatment effects here was only very weak. In both cases, the range of plausible values was large and included many values practically-equivalent to zero. Interestingly, non-voters showed a similar pattern of treatment effects to opposition voters, though were even more muted. This is perhaps unsurprising, given that the participants who comprised this group presumably had little sense of party identification.

Political Surveys and Partisan Bias

One question remains unanswered: how much partisan bias do political surveys account for? With only a single experiment to draw upon, this is difficult to know. Yet we can approximate this proportion by assuming that my treatment effects represent upper-bounds on the true effect. As I discuss above, my estimates are likely conservative. As such, treating them as an *upper-* and not *lower-*bounds is also conservative as the true value may be larger.

Computing the proportion of bias *within* the experiment is simple if we rely on the log-odds changes shown in table A2. One need only divide the treatment's main effect and its interaction with partisanship by its main effect, its interaction, and the main effect of partisanship. In the present case, this suggests that around 30.3 percent (95% CI: 11.4 to 47.3) of the partisan bias present in incumbent supporters self-reported economic perceptions is due to the political survey context itself.

While informative, this estimate is limited only to a single case. It would be better to compute a *distribution* of proportions using data from many points in time. The British Election Study Internet Panel, 2014–2023 (Fieldhouse, Green, Evans, Mellon, Prosser, et al. 2020), provides one such source of data. The BESIP includes the national economic perceptions item in fifteen separate waves. These cover the period between April 2014 and November 2019. I fit a similar ordered regression model to each wave of the data then, as the data do not vary the survey context, use the treatment effect from my survey experiment to approximate the proportion of partisan bias due to the political survey context under the assumption that it remains constant.

Figure 4 shows the resulting estimates. For incumbent supporters, these range from a low of 20.1 percent (95% CI: 6.7 to 34.0) in wave 6 to a high of 31.7 percent (95% CI: 10.8 to 53.5) in wave 13. The average across all waves is 27.4 percent (95% CI: 9.4 to 45.8), suggesting that around one-quarter of all partisan bias in the economic perceptions that incumbent partisans report is due to political survey effects. The equivalent effects for opposition supporters are much weaker and much more uncertain.
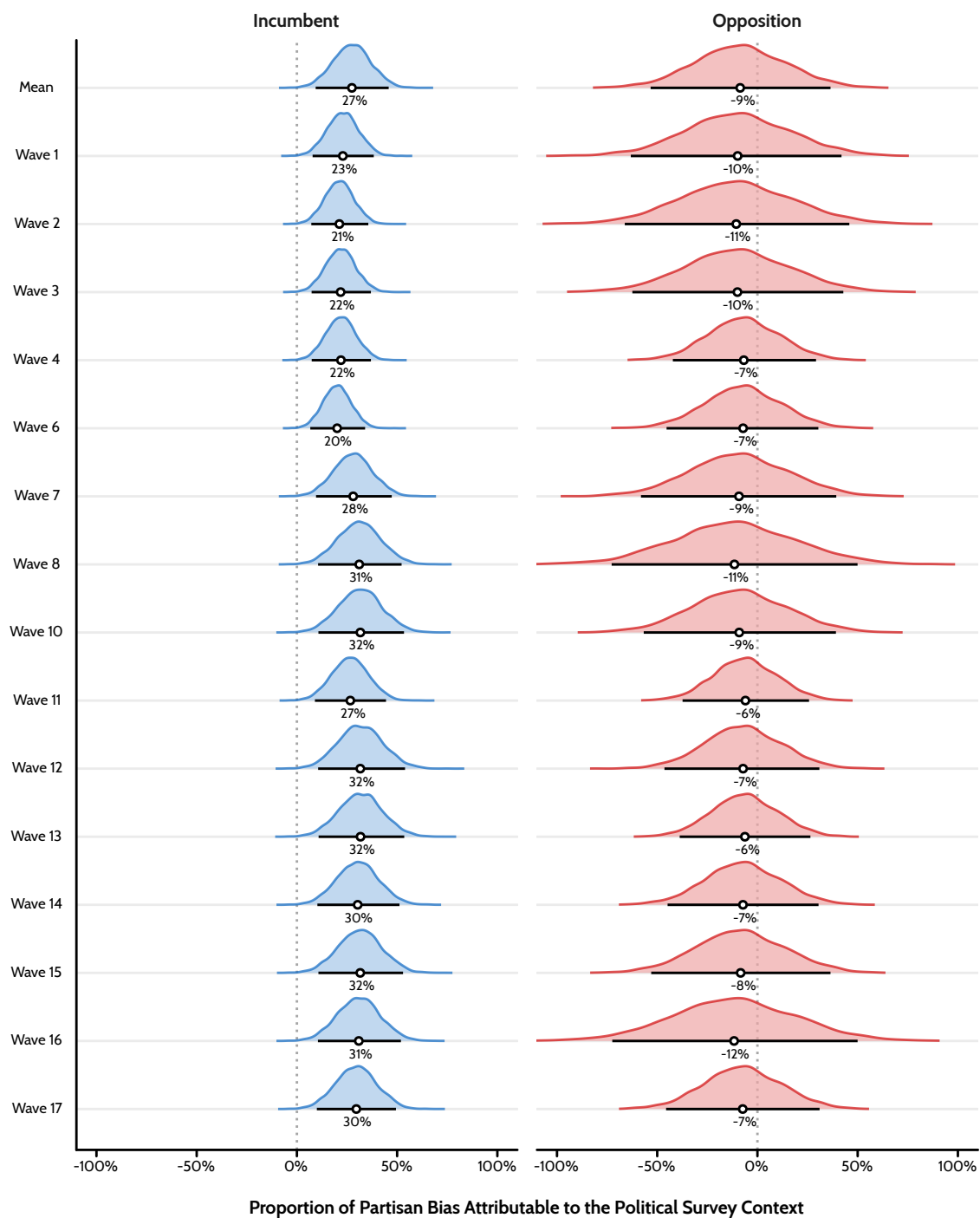
Figure 4: Assuming that the estimates I present here represent upper-bounds, political surveys account for around one-quarter of all partisan bias in the economic perceptions that voters report in fifteen recent waves of the British Election Study Internet Panel (BESIP). In reality, these estimates are likely conservative. As such, the true amount of partisan bias that political surveys account for may be larger outside of election periods. Note that density plots here show the posterior distribution of the proportion of partisan bias attributable to the political survey context. Further, black bars show their 95% credible interval and point estimates their median.

Discussion and Conclusion

Survey research often proceeds as though survey respondents say what they mean. This is especially true when it comes to studying both the economic vote and voters' economic perceptions. Most often, this research assumes that differences between groups that exist *within the survey* reflect real differences that exist *outside of the survey* (Bullock and Lenz 2019). My results show that this is not always the case. Some partisan bias arises simply due to the political survey context itself. In particular, I show that incumbent partisans report more positive economic perceptions in political compared to non-political surveys.

Why might this affect incumbent partisans but not opposition partisans? One explanation is that different partisans face different pressures when the economy is middling or poor. First, let us consider opposition partisans. When things are bad, these voters' primed and unprimed responses should coincide. Thus, they should show little difference in partisan bias. Now consider incumbent partisans, who face the opposite pressure. For this group, political priming leads them to report that things have gotten better. As a result, the economic climate causes a gap to open between the perceptions that they report in political and non-political survey contexts. If this is correct, and the state of the economy moderates political survey treatment effects, then future research might find that its effect reverses when the economy is doing well.

As the political survey context worsens partisan bias in self-reported economic perceptions, the most pressing issue is to work out how this affects applied research. For the economic vote, the outcome is mixed. Measuring economic perceptions in political surveys almost certainly worsens economic voting's endogeneity problem (Visconti 2017; Evans and Pickup 2010; Evans and Andersen 2006). Yet any effects appear limited only to incumbent partisans and much partisan bias remains constant. For political science in general, my results raise questions of measurement and validity. After all, these are likely not the only items sensitive to changes in survey context. Take items that measure attitudes towards immigration. It seems reasonable to expect voters to report different attitudes when primed to consider party politics or, say, changes in the labor market or even unrelated topics like dental hygiene.

16

The consequences are most serious for macro research. This is because small differences at the individual-level can yield large differences at the aggregate-level. Most often, this research aggregates these data in one of two ways. First, they compute the proportion of respondents who say that the economy has gotten better. However, my results suggest that doing so would over-estimate how rosy incumbent supporters think things really are where these figures are split by partisanship (see, for example, Enns, Kellstedt, and McAvoy 2012). Second, they compute net economic perceptions. That is, the proportion of respondents who report that the economy has gotten better minus those who report that it has gotten worse. Yet using net figures both over-estimates incumbent positivity and under-estimates incumbent negativity. For example, the results in figure 2 suggest a difference in net economic perceptions of almost 12 percentage points depending on the survey context. Some might argue that this is no issue for economic voting research that most often relies on real economic statistics (though see Lewis-Beck, Martini, and Kiewiet 2013). This might be true. Even so, it would still be a problem for the broader analysis of attitudes and opinions in mass publics.

Future research should consider if political survey effects are constant or, instead, if they vary by external context. For example, I have alluded to the possibility that my results might change in good economic times. Ultimately, this remains to be tested. Nevertheless, my results suggest a fruitful avenue for future research: we might opt not to adjust our *models*, but rather to adjust our *designs*. Research in this vein has already begun (Visconti 2017). One obvious suggestion would be to field separate surveys to measure respondents' party political and non-party political attitudes and beliefs. Though this might be more costly, the gains could be considerable if it reduced nuisance variation that muddies our inferences. As a result, students of the economic vote might gain both a better understanding of how the economy affects voters' behavior and how voters come to update their economic perceptions.

References

Agresti, Alan. 2010. *Analysis of Ordinal Categorical Data*. 2nd ed. Wiley Series in Probability and Statistics. Hoboken, N.J: Wiley.

Anderson, Christopher J. 2007. "The End of Economic Voting? Contingency Dilemmas and the Limits of Democratic Accountability." *Annual Review of Political Science* 10 (1): 271–96. https://doi.org/10.1146/annurev.polisci.10.050806.155344.

Anderson, Christopher J., Silvia M. Mendes, and Yuliya V. Tverdova. 2004. "Endogenous Economic Voting: Evidence from the 1997 British Election." *Electoral Studies* 23 (4): 683–708.

Anson, Ian G. 2017. "'That's Not How It Works': Economic Indicators and the Construction of Partisan Economic Narratives." *Journal of Elections, Public Opinion and Parties* 27 (2): 213–34. https://doi.org/10.1080/17457289.2016.1215319.

Ashworth, Scott. 2012. "Electoral Accountability: Recent Theoretical and Empirical Work." *Annual Review of Political Science* 15 (1): 183–201. https://doi.org/10.1146/annurev-polisci-031710-103823.

Bailey, Jack. 2019. "The Fact Remains: Party ID Moderates How Voters Respond to Economic Change." *Electoral Studies* 61. https://doi.org/10.1016/j.electstud.2019.102071.

Bartels, Larry M. 2002. "Beyond the Running Tally: Partisan Bias in Political Perceptions." *Political Behavior* 24 (2): 117–50.

Bisgaard, Martin. 2015. "Bias Will Find a Way: Economic Perceptions, Attributions of Blame, and Partisan-Motivated Reasoning During Crisis." *The Journal of Politics* 77 (3): 849–60. https://doi.org/10.1086/681591.

———. 2019. "How Getting the Facts Right Can Fuel Partisan-Motivated Reasoning." *American Journal of Political Science* 63 (4): 824–39. https://doi.org/10.1111/ajps.12432.

Bisgaard, Martin, and Rune Slothuus. 2018. "Partisan Elites as Culprits? How Party Cues Shape Partisan Perceptual Gaps." *American Journal of Political Science* 62 (2): 456–69. https://doi.org/10.1111/ajps.12349.

Brader, Ted, Joshua A. Tucker, and Dominik Duell. 2013. "Which Parties Can Lead

Opinion? Experimental Evidence on Partisan Cue Taking in Multiparty Democracies." *Comparative Political Studies* 46 (11): 1485–1517. https://doi.org/10.1177/0010414012453452.

Brady, Henry E., and Paul M. Sniderman. 1985. "Attitude Attribution: A Group Basis for Political Reasoning." *American Political Science Review* 79 (4): 1061–78. https://doi.org/10.2307/1956248.

Bullock, John G., Alan S. Gerber, Seth J. Hill, and Gregory A. Huber. 2015. "Partisan Bias in Factual Beliefs About Politics." *Quarterly Journal of Political Science* 10 (4): 519–78. https://doi.org/10.1561/100.00014074.

Bullock, John G., and Gabriel Lenz. 2019. "Partisan Bias in Surveys." *Annual Review of Political Science* 22 (1): 325–42. https://doi.org/10.1146/annurev-polisci-051117-050904.

Campbell, Angus, Philip E. Converse, Warren E. Miller, and Donald E. Stokes. 1960. *The American Voter*. Chicago, IL: The University of Chicago Press.

Conover, Pamela Johnston, Stanley Feldman, and Kathleen Knight. 1987. "The Personal and Political Underpinnings of Economic Forecasts." *American Journal of Political Science* 31 (3): 559–83. https://doi.org/10.2307/2111283.

Curtin, Richard. 2019. "Consumer Expectations: A New Paradigm." *Business Economics* 54 (4): 199–210. https://doi.org/10.1057/s11369-019-00148-1.

Dassonneville, Ruth, and Michael S. Lewis-Beck. 2019. "A Changing Economic Vote in Western Europe? Long-Term Vs. Short-Term Forces." *European Political Science Review* 11 (1): 91–108. https://doi.org/10.1017/S1755773918000231.

De Vries, Catherine E., Sara B. Hobolt, and James Tilley. 2018. "Facing up to the Facts: What Causes Economic Perceptions?" *Electoral Studies* 51 (February): 115–22. https://doi.org/10.1016/j.electstud.2017.09.006.

Enns, Peter K., Paul M. Kellstedt, and Gregory E. McAvoy. 2012. "The Consequences of Partisanship in Economic Perceptions." *Public Opinion Quarterly* 76 (2): 287–310. https://doi.org/10.1093/poq/nfs016.

Evans, Geoffrey, and Robert Andersen. 2006. "The Political Conditioning of Economic

Perceptions." *The Journal of Politics* 68 (1): 194–207. https://doi.org/10.1111/j.1468-2508.2006.00380.x.

Evans, Geoffrey, and Mark Pickup. 2010. "Reversing the Causal Arrow: The Political Conditioning of Economic Perceptions in the 2000-2004 U.S. Presidential Election Cycle." *The Journal of Politics* 72 (4): 1236–51. https://doi.org/10.1017/S0022381610000654.

Fieldhouse, Edward, Jane Green, Geoffrey Evans, Jonathan Mellon, and Christopher Prosser. 2020. "British Election Study Internet Panel, 2014-2023."

Fieldhouse, Edward, Jane Green, Geoffrey Evans, Jonathan Mellon, Christopher Prosser, Hermann Schmitt, and Cees van der Eijk. 2020. *Electoral Shocks: The Volatile Voter in a Turbulent World*. New York, NY: Oxford University Press.

Gelman, Andrew, and John Carlin. 2014. "Beyond Power Calculations: Assessing Type S (Sign) and Type M (Magnitude) Errors." *Perspectives on Psychological Science* 9 (6): 641–51. https://doi.org/10.1177/1745691614551642.

Green, Jane, and Will Jennings. 2017. *The Politics of Competence: Parties, Public Opinion and Voters*. Cambridge: Cambridge University Press.

Groenendyk, Eric W. 2013. *Competing Motives in the Partisan Mind: How Loyalty and Responsiveness Shape Party Identification and Democracy*. Series in Political Psychology. New York: Oxford University Press.

Hansford, Thomas G., and Brad T. Gomez. 2015. "Reevaluating the Sociotropic Economic Voting Hypothesis." *Electoral Studies* 39: 15–25. https://doi.org/10.1016/j.electstud.2015.03.005.

Healy, Andrew, and Neil Malhotra. 2013. "Retrospective Voting Reconsidered." *Annual Review of Political Science* 16 (1): 285–306. https://doi.org/10.1146/annurev-polisci-032211-212920.

Hill, Seth J. 2017. "Learning Together Slowly: Bayesian Learning About Political Facts." *The Journal of Politics* 79 (4): 1403–18. https://doi.org/10.1086/692739.

Katona, George. 1951. *Psychological Analysis of Economic Behavior*. New York, NY: McGraw-Hill.

Kinder, Donald R., and D. Roderick Kiewiet. 1979. "Economic Discontent and Political Behavior: The Role of Personal Grievances and Collective Economic Judgments in Congressional Voting." *American Journal of Political Science* 23 (3): 495–527. https://doi.org/10.2307/2111027.

———. 1981. "Sociotropic Politics: The American Case." *British Journal of Political Science* 11 (2): 129–61. https://doi.org/10.2307/193580.

Kruschke, John. 2015. *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan*. Second. Cambridge, MA: Academic Press.

Kruschke, John K., and Torrin M. Liddell. 2017. "Bayesian Data Analysis for Newcomers." *Psychonomic Bulletin & Review*, April, 1–23. https://doi.org/10.3758/s13423-017-1272-1.

Kunda, Ziva. 1990. "The Case for Motivated Reasoning." *Psychological Bulletin* 108 (3): 480–98. https://doi.org/10.1037/0033-2909.108.3.480.

Lambert, B. 2018. *A Student's Guide to Bayesian Statistics*. London, UK: Sage.

Lewis-Beck, Michael S., and Marina Costa Lobo. 2017. "The Economic Vote: Ordinary Vs. Extraordinary Times." In *The Sage Handbook of Electoral Behaviour*, edited by Kai Arzheimer, Jocelyn Evans, and Michael S. Lewis-Beck, 2:606–30. Thousand Oaks, CA: Sage.

Lewis-Beck, Michael S., Nicholas F. Martini, and D. Roderick Kiewiet. 2013. "The Nature of Economic Perceptions in Mass Publics." *Electoral Studies* 32 (3): 524–28. https://doi.org/10.1016/j.electstud.2013.05.026.

Lewis-Beck, Michael S, and Martin Paldam. 2000. "Economic Voting: An Introduction." *Electoral Studies* 19 (2): 113–21. https://doi.org/10.1016/S0261-3794(99)00042-6.

Lewis-Beck, Michael S., and Mary Stegmaier. 2007. "Economic Models of Voting." In *The Oxford Handbook of Political Behavior*, edited by Russell J. Dalton and Hans-Dieter Klingemann, 518–37. Oxford: Oxford University Press.

Liddell, Torrin M., and John K. Kruschke. 2018. "Analyzing Ordinal Data with Metric Models: What Could Possibly Go Wrong?" *Journal of Experimental Social Psychology* 79 (November): 328–48. https://doi.org/10.1016/j.jesp.2018.08.009.

Mccullagh, Peter. 1980. "Regression Models for Ordinal Data." *Journal of the Royal*

Statistical Society. Series B (Methodological) 42 (2): 109–42.

McElreath, Richard. 2020. *Statistical Rethinking: A Bayesian Course with Examples in R and Stan*. Second. CRC Texts in Statistical Science. Boca Raton: CRC Press.

Miratrix, Luke W., Jasjeet S. Sekhon, Alexander G. Theodoridis, and Luis F. Campos. 2018. "Worth Weighting? How to Think About and Use Weights in Survey Experiments." *Political Analysis* 26 (3): 275–91. https://doi.org/10.1017/pan.2018.1.

Paldam, Martin. 1981. "A Preliminary Survey of the Theories and Findings on Vote and Popularity Functions." *European Journal of Political Research* 9 (2): 181–99. https://doi.org/10.1111/j.1475-6765.1981.tb00598.x.

Pepinsky, Thomas B. 2018. "A Note on Listwise Deletion Versus Multiple Imputation." *Political Analysis* 26 (4): 480–88. https://doi.org/10.1017/pan.2018.18.

Prior, Markus, Gaurav Sood, and Kabir Khanna. 2015. "You Cannot Be Serious: The Impact of Accuracy Incentives on Partisan Bias in Reports of Economic Perceptions." *Quarterly Journal of Political Science* 10 (4): 489–518. https://doi.org/10.1561/100.00014127.

Ramirez, Mark D., and Nathan Erickson. 2014. "Partisan Bias and Information Discounting in Economic Judgments: Partisan Bias and Information Discounting." *Political Psychology* 35 (3): 401–15. https://doi.org/10.1111/pops.12064.

Sanders, David, Harold D. Clarke, Marianne C. Stewart, and Paul Whiteley. 2007. "Does Mode Matter for Modeling Political Choice? Evidence from the 2005 British Election Study." *Political Analysis* 15 (3): 257–85. https://doi.org/10.1093/pan/mpl010.

Schaffner, Brian F., and Samantha Luks. 2018. "Misinformation or Expressive Responding? What an Inauguration Crowd Can Tell Us About the Source of Political Misinformation in Surveys." *Public Opinion Quarterly* 82 (1): 135–47. https://doi.org/10.1093/poq/nfx042.

Sears, David O., and Richard R. Lau. 1983. "Inducing Apparently Self-Interested Political Preferences." *American Journal of Political Science* 27 (2): 223–52. https://doi.org/10.2307/2111016.

Sorace, Miriam, and Sara B. Hobolt. 2018. "Distorted Perceptions: How Leavers and

Remainers View the Economy and with What Consequences." *LSE British Politics and Policy*. https://blogs.lse.ac.uk/politicsandpolicy/selective-perception-brexit-economy/.

Stanig, Piero. 2013. "Political Polarization in Retrospective Economic Evaluations During Recessions and Recoveries." *Electoral Studies*, Special Symposium: The new research agenda on electoral integrity, 32 (4): 729–45. https://doi.org/10.1016/j.electstud.2013.05.029.

Stegmaier, Mary, Michael S. Lewis-Beck, and Lincoln Brown. 2019. "The Economic Voter Decides." *Oxford Research Encyclopedia of Politics*, May. https://doi.org/10.1093/acrefore/978 0190228637.013.931.

Tilley, James, and Sara B. Hobolt. 2011. "Is the Government to Blame? An Experimental Test of How Partisanship Shapes Perceptions of Performance and Responsibility." *The Journal of Politics* 73 (2): 316–30. https://doi.org/10.1017/S0022381611000168.

Tilley, James, Anja Neundorf, and Sara B. Hobolt. 2018. "When the Pound in People's Pocket Matters: How Changes to Personal Financial Circumstances Affect Party Choice." *The Journal of Politics* 80 (2): 555–69. https://doi.org/10.1086/694549.

Twyman, Joe. 2008. "Getting It Right: YouGov and Online Survey Research in Britain." *Journal of Elections, Public Opinion and Parties* 18 (4): 343–54. https://doi.org/10.1080/1745 7280802305169.

Visconti, Giancarlo. 2017. "Economic Perceptions and Electoral Choices: A Design-Based Approach." *Political Science Research and Methods*, September, 1–19. https://doi.org/10.101 7/psrm.2017.26.

Wilcox, Nathaniel, and Christopher Wlezien. 1993. "The Contamination of Responses to Survey Items: Economic Perceptions and Political Judgments." *Political Analysis* 5: 181–213.
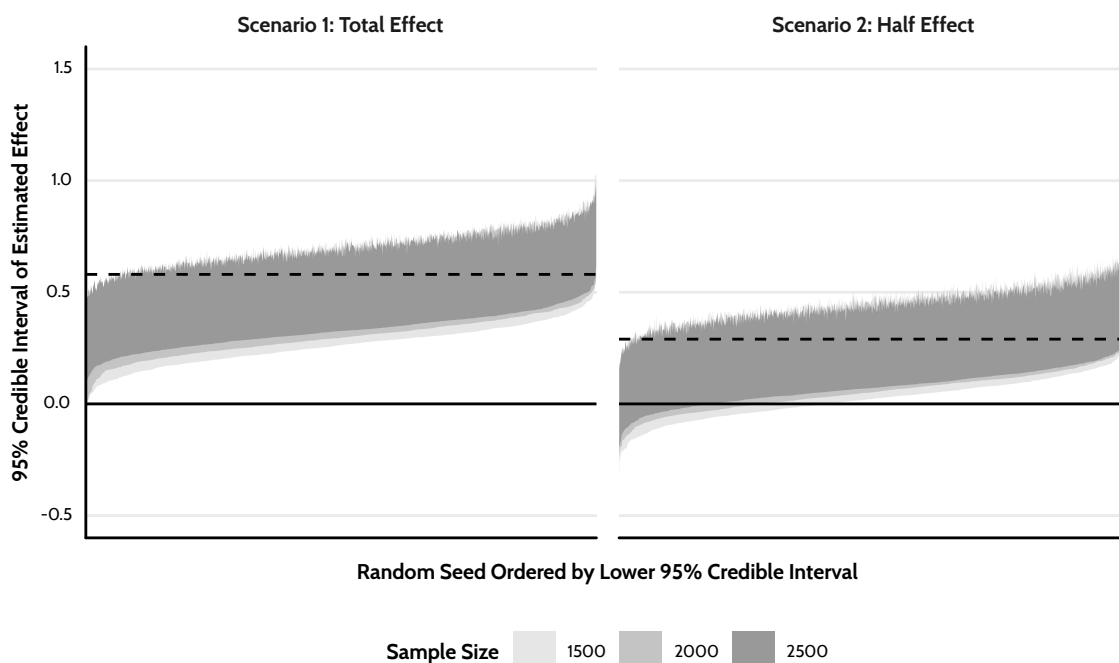
Appendix A: Power Analysis

Figure A1: Outcomes from 6,000 simulation-based power analyses, ordered by lower 95% credible interval. In scenario 1, I assume that political survey effects account for the total effect I find in the observational data. In scenario 2, I assume instead that political survey effects account for half of this effect. All samples achieve 80% in scenario 1. Only a sample of 2,500 achieved 80% power in scenario 2.

Before fielding my experiment, it was essential that I determine an appropriate sample size. To do so, I conducted a simulation-based power analysis. This approach was necessary as my dependent variable was ordinal. Unlike continuous data, it is not possible to conduct power analyses for ordinal data by hand. It was important that the effect sizes I used in my power analysis be of a realistic and reasonable size. To do so, I fit a similar ordered-probit model to the one I discussed above to data from wave 16 of the British Election Study Internet Panel (Fieldhouse, Green, Evans, Mellon, Prosser, et al. 2020). I then used the regression parameters to establish informative prior expectations for what effect sizes I might expect.

The BESIP data are observational. As such, the resulting estimates tell us only the net effect that respondents' past voting behavior and the political survey context have on the economic perceptions that they report. We do not, however, know what proportion of these effects each

24

accounts for. I proposed two hypothetical scenarios as the basis of my power analysis. In the first, I assumed that political survey effects accounted for the total effect I observed in the observational data. In the second scenario, I instead assumed that they accounted for only half of it. I then simulated 1,000 experiments for each across three sample sizes ($n = 1,500$, $n = 2,000$, and $n = 2,500$). These matched the blocked structure of my experiment. To make sure that my results did not depend on the random seed I used to simulate my data, I incremented it by one for each simulation. This gave 6,000 simulated experiments in total. In simulating my data, I focused on the treatment's effect on incumbent partisans. Further, I set my desired level of power at 80%. As I expect this effect to be positive, this implies that 80% of the effects from my simulation should have a lower 95% credible interval exceeds zero.

Figure A1 shows the outcomes of all 6,000 simulated experiments, ordered by their lower 95% credible interval. For scenario 1, all sample sizes achieved the desired level of power. Indeed, every simulated experiment yielded estimates that were greater than zero matter the sample size. This was not the case for scenario 2. Instead, every sample size included at least some simulations with lower 95% credible intervals that did not exceed zero. Here, a sample size of 1,500 corresponded with a power level of 60%; 2,000 with a power level of 74%; and 2,500 with a power level of 84%. Thus, I opted for the latter to exceed 80% power.

Appendix B: Questionnaire

<table>
<tr><td align="center">Treatment</td><td align="center">Control</td></tr>
</table>

Q1. If there were a general election held tomorrow, which party would you vote for?

Q1. Imagine that you need to buy toothpaste in the near future, which brand would you choose?

Treatment	Control
1. *Conservative*	1. *Colgate*
2. *Labour*	2. *Sensodyne*
3. *Liberal Democrat*	3. *Aquafresh*
4. *Scottish National Party (SNP)*	4. *Oral-B*
5. *Plaid Cymru*	5. *Macleans*
6. *Brexit Party*	6. *Arm & Hammer*
7. *Green*	7. *Crest*
8. *Some other party*	8. *Some other brand*
9. *Would not vote*	9. *I would not buy toothpaste*
10. *Don't know*	10. *Don't know*

Q2. On a scale of 0 (certain NOT to vote) to 10 (absolutely certain to vote), how likely would you be to vote in a general election tomorrow?

1. *0 – Certain NOT to vote*
2. *1*
3. *2*
4. *3*
5. *4*
6. *5*
7. *6*
8. *7*
9. *8*
10. *9*
11. *10 – Absolutely certain to vote*
12. *Don't know*

Q3. Who do you think would make the best Prime Minister?

1. *Boris Johnson*
2. *Jeremy Corbyn*
3. *Jo Swinson*
4. *Don't know*

Q2. On a scale of 0 (not at all important) to 10 (very important), how important do you think dental hygiene is in everyday life?

1. *0 – Not at all important*
2. *1*
3. *2*
4. *3*
5. *4*
6. *5*
7. *6*
8. *7*
9. *8*
10. *9*
11. *10 – Very important*
12. *Don't know*

Q3. Generally speaking, what type of toothbrush do you use?

1. *Manual*
2. *Electric*
3. *I do not have a toothbrush*
4. *Don't know*

Q4. In hindsight, do you think Britain was right or wrong to vote to leave the European Union?

1. *Right to leave*
2. *Wrong to leave*
3. *Neither right nor wrong*
4. *Don't know*

Q5. How well or badly do you think the government are doing at handling Britain's exit from the European Union?

1. *Very well*
2. *Fairly well*
3. *Neither well nor badly*
4. *Fairly badly*
5. *Very badly*
6. *Don't know*

Q4. When brushing your teeth, do you...

1. *Wet your toothbrush, then apply toothpaste?*
2. *Apply toothpaste, then wet your toothbrush?*
3. *Not wet your toothbrush at all*
4. *Don't know*

Q5. Generally speaking, on average how many times do you brush your teeth every day?

1. *Never*
2. *Once*
3. *Twice*
4. *Three times*
5. *More than three times*
6. *Don't know*

Treatment ends. Subsequent questions are identical for each group.

Q6. Now, a few questions about economic conditions. How does the *financial situation of your household* now compare with what it was 12 months ago?

1. *Got a lot worse*

2. *Got a little worse*

3. *Stayed the same*

4. *Got a little better*

5. *Got a lot better*

6. *Don't know*

Q7. How do you think the *general economic situation* in this country has changed over the last 12 months?

1. *Got a lot worse*

2. *Got a little worse*

3. *Stayed the same*

4. *Got a little better*

5. *Got a lot better*

6. *Don't know*

Appendix C: Prior Distributions for Ordered Regression Models

As I discuss in my methods section above, my experiment includes an outcome variables that is ordinal rather than continuous or binary. Though others often treat these data as though they are continuous for the sake of convenience, this practice is prone to a whole host of serious inferential pitfalls. Further, though more robust, almost all conventional ordered regression models face similar issues. As such, I use Bayesian methods to implement an extended ordered regression model that overcomes these problems, thereby allowing me to estimate any treatment effects in a principled manner that respects the nature of the data.

Though similar, the Bayesian approach to statistical analysis does introduce some points of difference compared to the classical statistics that dominates much political science research. Most notably, it requires that one specify a prior distribution over each parameter in one's model before fitting it to the data[11]. As well as allowing us to shift focus from the likelihood (*"what is the probability of the data given the hypothesis?"*) to the posterior distribution (*"what is the probability of the hypothesis given the data?"*), these "priors" also serve two useful purposes. First, they allow us to incorporate any pre-existing knowledge that we might be privy to into our models. This might include the results of a previous analysis (thereby having our model expect results similar to the previous case before it sees the data) or simply our understanding of the nature of the model and what values it is reasonable for certain parameters to take (for example, we know that it is not possible for probabilities to be negative). Second, they make our models skeptical by nature and shrink any parameter estimates towards the prior. This "regularization" protects against over-fitting common to maximum likelihood-based approaches, especially where sample sizes are small (McElreath 2020; Lambert 2018; Gelman and Carlin 2014).

In this case, there is little existing evidence on which to draw. Though others have fielded the same economic perception items in the past, they have tended to do so in an observational, rather than an experimental, context. As a result, I do not have a good idea about what values my parameters might take before I fit the model to the data. To complicate matters further,

---

[11] For an introduction to Bayesian statistics and Bayesian methods, see McElreath (2020); Lambert (2018); Kruschke and Liddell (2017); Kruschke (2015)

ordered regression models have many moving parts, which interact with one another to produce the implied ordinal outcome. To overcome this complication, I take a principled approach below and decide my priors based on a series of prior predictive simulations. In doing so, I seek to ensure that all my decisions be conservative so that any treatment effects I estimate will be robust. This has two implications. First, that I should assume all effects to be zero before fitting my model to the data. Second, that I should also assume all possible combination of response probabilities to be equally likely. The various ordered regression models that I present in this paper rely on the same three sets of parameters, each of which serves a distinct function. I discuss each specific parameter in turn, below.

Threshold Parameters

Ordered regression models work by translating between an ordinal variable that we observe and a continuous variable that we do not. To perform this feat, they rely on a series of threshold parameters that split the latent continuous distribution into as many segments as their are response options. The are of each segment then corresponds to the probability that the response option that it represents will occur. Absent any knowledge about the nature of the data, the most conservative assumption that we can make is that any possible combination of responses is as likely as any other.

As we measure each response option in terms of its probability of occurring, it is worth also thinking on the probability scale when setting our priors. Thus, given this assumption, we should expect the probability of a threshold parameter landing on any point on the probability scale to be constant. The issue then is to find a prior on the latent probit scale on which the model operates, which ranges from $-\infty$ to $+\infty$, that gives a flat prior on the probability scale that we really care about, which is bounded by 0 and 1. Fortunately, this is not difficult and the answer is relatively well-known: a Normal(0, 1) prior on the former gives a uniform prior on the latter.

The four left-most panels of figure A2 show the resulting prior distributions that this collective prior over all thresholds implies for each specific threshold. The right-most panel

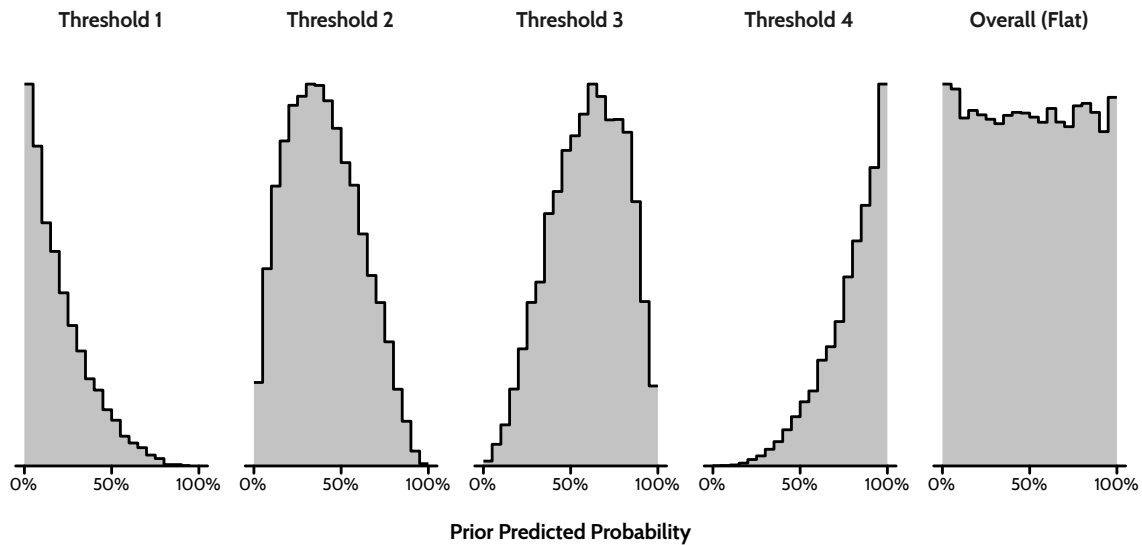## Normal(0, 1) Prior Over Threshold Distributions

Figure A2: Before seeing the data, the most conservative assumption that we can make about the distribution of response options is that each combination is as likely as any other. This implies that prior thresholds must be able to take any possible value on the probability scale, conditional on the constraints of the model itself. A Normal(0, 1) prior on the latent probit scale yields fulfills this requirement and implies a flat prior on the probability scale. The figure above shows the resulting prior threshold parameters. Note also that the probability of a threshold occurring at any point on the scale is constant over the entire range of the scale (with any deviations arising only due to random noise in the simulation process).

instead shows the implied prior over the whole probability scale (i.e. the distribution that we would find were we to stack each threshold distribution on top of each other). As the histograms in the figure make clear, the implied priors for each threshold are non-informative and, in all cases, take a wide range of possible values. For example, the priors shown here allow for a non-zero probability that the first threshold occurs as high as the 80% mark and the fourth threshold as low as 20%. That the first response option corresponds to respondents reporting that the economy "got a lot worse" and the last that it has "got a lot better" reaffirms just how non-informative these priors really are.

Though we do not specify priors for each specific threshold, these prior predictive simulations suggest that they take their own distinctive shapes nonetheless. This phenomenon arises due to the constraints that both the prior and the model impose on the values that these parameters can take. For example, each threshold is constrained to take only values smaller than

those of the threshold that follow it. As a result, it is not possible for any threshold to cover the entire space as this would leave the others with nowhere to go. Likewise, the collective nature of the prior means that the priors for each threshold *must* result in a flat prior overall. The result is the set of symmetrical distributions that we see above.

Beta Parameters

Threshold parameters segment the latent outcome distribution, though do not move. Beta parameters, instead, shift the latent distribution up and down its scale. This movement then serves to shift the probability mass of each observed response option in turn. As such, we can interpret beta parameters much like regression coefficients in linear and logistic regression models, which perform a similar role.

Before seeing the data, the most conservative assumption that one can make about these effects is that they are equal to zero (i.e. that they are null). Doing so is simple and, in the absence of any better information, uncontroversial. More difficult however is determining how uncertain these priors should be. One the one hand, a tight prior around zero will be very conservative, but perhaps to the extent that it ignores perfectly informative data. On the other, a loose prior will pay closer attention to the data, but perhaps to the extent that it will result in over-fitting.

For models with continuous outcomes, things are straightforward. If the prior is very wide, then it is also likely to cover the full range of plausible values that its respective parameter might take. But ordinal variables are not continuous and, as a result, this common practice can lead to perverse implications. Unlike models with continuous outcomes, wide priors on the latent probit scale *do not* give wide priors on the outcome scale. This is because the outcome scale takes only a finite set of discrete values. As a result, wide priors on the latent probit scale instead imply U-shaped priors due to probability mass piling up at the extremes. This problem then multiplies — quite literally — where the data include variables that exhibit a high degree of variation (for example age, which in the study of voting behavior might take any value between 18 and 100) or where the sum of all variables is large (such as when a model contains many parameters).

Figure A3 displays this phenomenon across different priors and different values of beta.
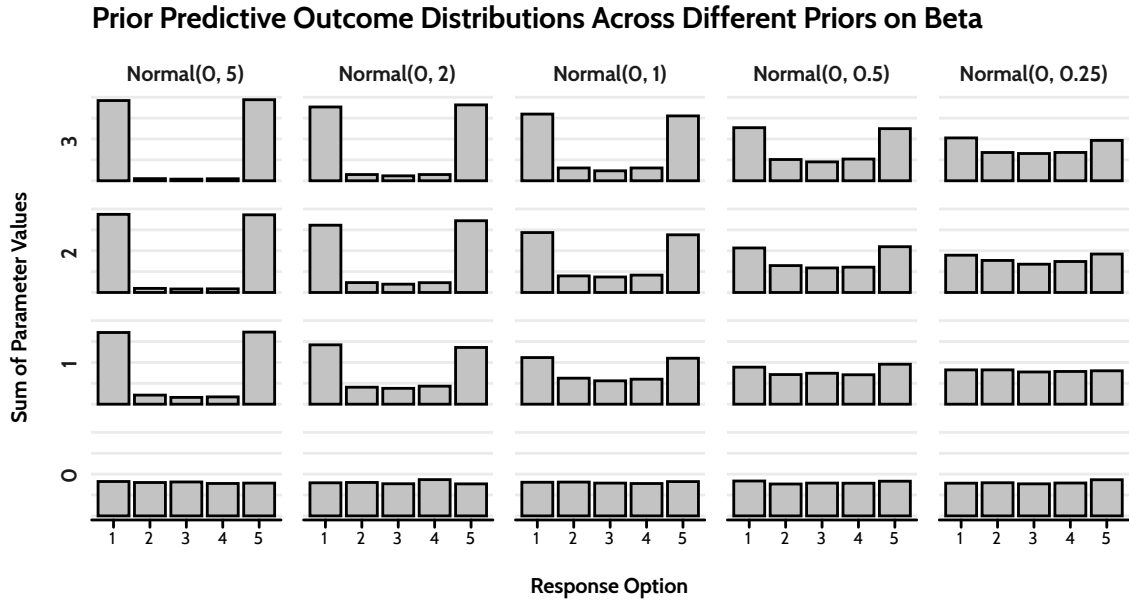
Prior Predictive Outcome Distributions Across Different Priors on Beta

Figure A3: While it is common to set wide priors on beta values where the outcome is continuous, such "non-informative" priors have perverse consequences when the outcome is ordered. This is because they make the latent scale too diffuse, thereby concentrating almost all of the prior probability mass at the two extremes of the observed ordinal outcome scale. Further, models that include many independent variables or independent variables that take extreme values worsen this problem further.

Where these sum to zero, only the thresholds determine the response distribution, which I fix to ensure that each response has a prior probability of 20% where betas sum to zero. As the figure shows, when this sum exceeds zero the prior probability of responding with either a 1 or a 5 increases. This is true for all priors, though the effect is most pronounced where the prior standard deviations are large. In each model in this paper, the sum of parameters increases where respondents voted at the last election or are in the treatment group. In light of this, using a prior on beta with a large standard deviation is akin to assuming that these participants are more likely to say either that the economy has "got a lot worse" or "got a lot better". Perhaps counter-intuitively, smaller standard deviations are, thus, less informative. Thus, I use the least informative prior — Normal(0, 0.25) — for all beta values in my models.

Delta Parameters

Whereas beta parameters shift the latent outcome distribution, delta parameters instead compress or disperse it at a given point. This serves to redistribute the observe outcome's

34

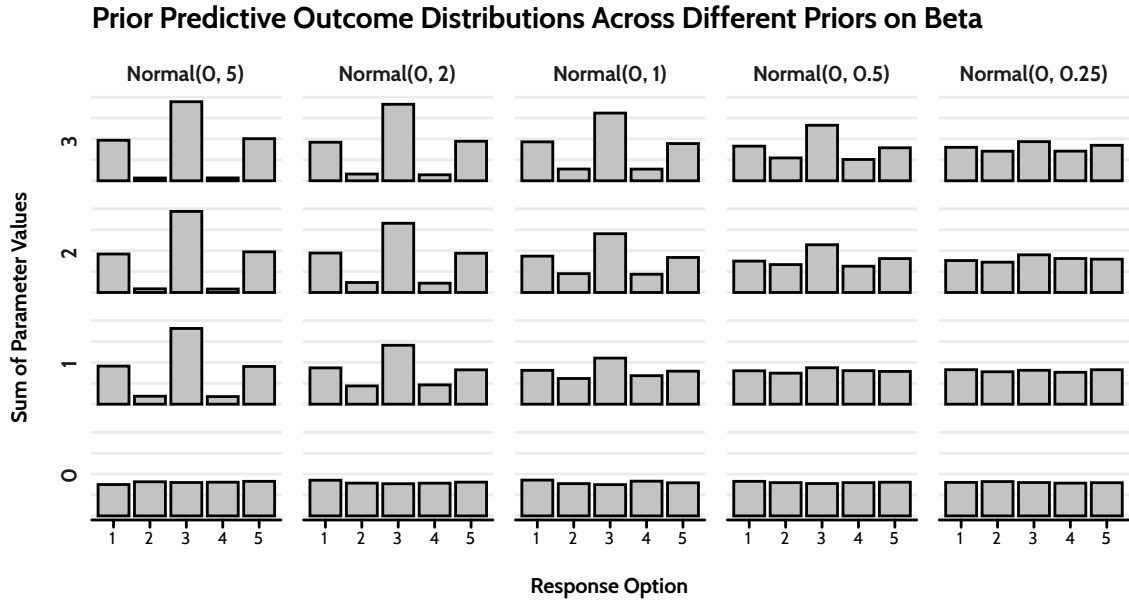**Prior Predictive Outcome Distributions Across Different Priors on Beta**

Figure A4: Setting diffuse priors on delta parameters can also have perverse consequences. In this case, they instead concentrate the prior probability mass in the middle and at the two extremes of the observed ordinal variable. Again, this is likely to be worse where models also include many independent variables or independent variables that take extreme values.

probability mass towards central or extreme responses, conditional on its place on the scale. As in the previous case, the most conservative assumption that we can make before seeing the data is to expect these parameters to be equal to zero. Where this is true, the standard deviation of the latent outcome distribution does not vary across participants. Again, this is simple to achieve and, again, things become more complicated when it comes to setting the standard deviation. The problem is the same as before: large standard deviations imply more, not less, informative outcomes.

Figure A4 is similar to figure A4 and shows the implication that different priors and different values of delta have on the implied prior outcome distribution. As before, I fix all thresholds in this case to imply an equal chance of any response option being selected and also fix all beta parameters to 0. While wide priors on the beta parameters produced U-shaped distributions, wide priors on the delta parameters do not. Instead, they increase the prior probability of central and extreme responses, resulting in a crown-like distribution. Note, however, that this pattern is conditional on the choice of thresholds and that U-shaped distributions may arise here too under different circumstances. As before, each response has an

equal probability where the sum of delta parameters is zero. As this sum increases, the central response option becomes much more likely and extreme responses somewhat more likely. This implies that tighter standard deviations are also less informative in this case too. Given this, I opt to use

While the beta parameters produced U-shaped distributions, these do not. Instead, they increase the prior probability of central and extreme responses. Note, however, that this is conditional on the choice of thresholds. Where these do not equalize the probability of each response, a U-shape can still emerge. As before, each response has an equal probability when the sum of parameter values is zero. As this sum increases, the central response becomes much more likely. Likewise, extreme responses become somewhat more likely too. Given this, I take the right-most distribution — a Normal(0, 0.25) prior — as the prior for each of my delta values as it remains most constant across all cases.

Appendix D: Robustness Checks

There are three plausible objections to the results I report above. First, that the treatment effects occur due to some mechanism other than partisan bias. Second, that the theory does not generalize to other types of electoral identification. And, third, that the results are sensitive to my model specification. I test each below. The first tests if the treatment mechanism relies on partisan bias. To do so, I apply the same test to participants' reported *personal* economic perceptions. Past research finds that these show little sensitivity to party identification. The second tests if the theory generalizes to other types of identification. In particular, voting behavior at the 2016 referendum on European Union membership. The third tests if the findings are robust to different methods. In this case, by substituting ordered regression for multinomial regression instead.

Personal Economic Perceptions and Partisan Bias as a Potential Mechanism

Above, I assume that my findings result from partisan bias. This seems reasonable given existing research (De Vries, Hobolt, and Tilley 2018; Bartels 2002; Conover, Feldman, and Knight 1987). Even so, a skeptic might argue that I have not yet provided good evidence that this is indeed the case. Instead, they might argue that some other mechanism is reasonable for my findings. As a result, the pattern that I observe might also apply to any other dependent variable. This is a reasonable objection, as my design does not allow me to tease apart any intermediary steps in the causal chain between survey context and reported economic perceptions. Fortunately, there are ways to reduce this uncertainty. One is to test how the treatment affects a similar item that we know suffers from little partisan bias. Voters' perceptions of their own personal finances are on such possibility. Like national-level items, these too have their origin in consumer confidence surveys (Katona 1951). But, unlike national-level items, they are much less sensitive to partisan bias. This makes sense. After all, many would argue that the government is less accountable for any one person's well-being than it is for the well-being of the nation as a whole (Lewis-Beck and Costa Lobo 2017; Lewis-Beck and Paldam 2000; Paldam 1981; Kinder and Kiewiet 1981, 1979; though see Tilley, Neundorf, and Hobolt 2018).
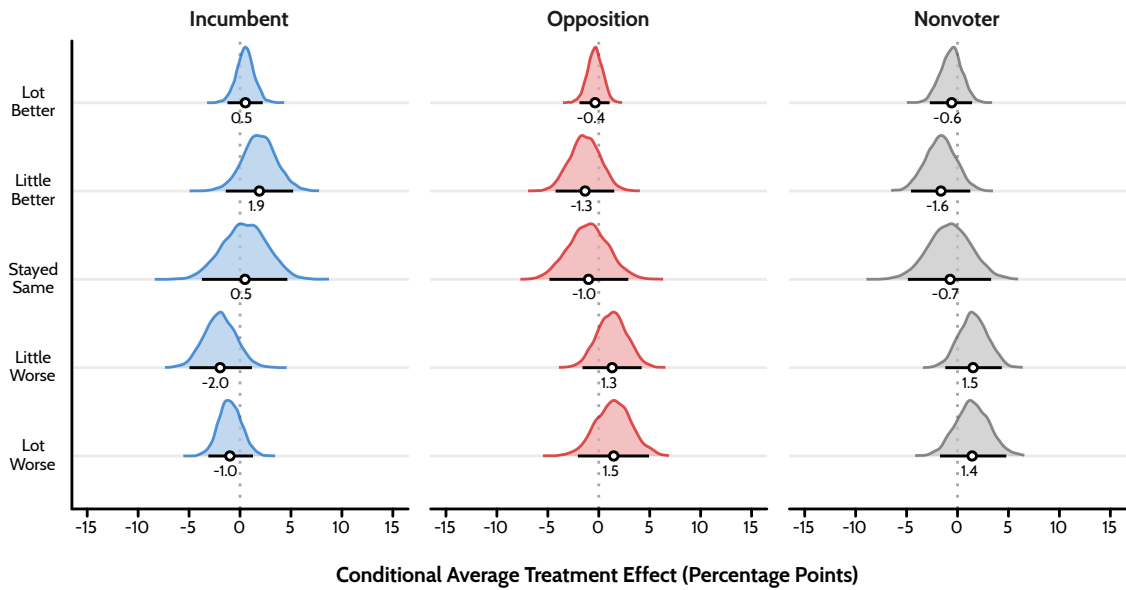
Figure A5: Political surveys do not cause voters to report different perceptions of their own personal finances (left panel). This is unsurprising, since prior research shows that they are much less sensitive to party identification. Positive values imply that those in the treatment group were more likely to report a given response. Negative values imply the opposite. In general, treatment effects showed the expected signs. Incumbent voters were more positive. Likewise, opposition voters were more negative. Even so, in all cases, the distribution of treatment estimates were centered on small values and had a plausible chance of being practically-equivalent to zero. Here, density plots show the posterior distribution of conditional average treatment effects. Further, black bars show their 95% credible intervals and point estimates their medians.

Figure A5 shows how the treatment affected the personal economic perceptions that my participants reported. As before, I condition these estimates on prior voting behavior for the same reasons as above. In this case, all treatment effects have the expected signs. That is, incumbent supporters are more positive and opposition supporters more negative under the treatment. This might, then, suggest the presence of at least some partisan bias. Yet, in all cases, point estimates are small. These range in size from only -0.4 (95% CI: -1.9 to 1.1) to 1.9 percentage points (95% CI: -1.4 to 5.2). Further, these effects have 95% credible intervals that, in all cases, are very uncertain.

Taken together, these results suggest little evidence that political surveys affect the personal economic perceptions that respondents report. Were some other mechanism responsible for the treatment effects I find, this might not be the case. Instead, I find that the treatment might have a similar effect for both items. Instead, both sets of results are consistent with

Table A1: There is a strong relationship between voting behaviour at the 2016 referendum on European Union membership and the 2017 general election. Even so, this relationship is not absolute. For example, some who voted for the incumbent Conservative Party in 2017 voted to remain in the EU in 2016. Likewise, some who voted for an opposition party voted to leave the EU.

	Nonvoter	Incumbent	Opposition	Total
Nonvoter	60.7% (495)	2.2% (20)	6.2% (56)	21.6% (571)
Leave	20.4% (166)	70.5% (648)	26.4% (239)	39.9% (1053)
Remain	18.9% (154)	27.3% (251)	67.4% (611)	38.5% (1016)
Total	100.0% (815)	100.0% (919)	100.0% (906)	100.0% (2640)

existing theory and the argument that I present above. That is, respondents must have a reason to assign responsibility to the government if political surveys are to prime respondents to respond in a different way. This does not seem to be the case for perceptions of one's personal finances. Instead, they appear to exhibit little partisan bias, leaving the treatment with nothing to manipulate. Of course, it is never possible to rule out any other mechanism with absolute certainty. Still, these results do at least make such a possibility seem much less likely.

Generalization of Treatment Effects Across Different Types of Electoral Identification

If the theory that underpins my analysis is robust, it should generalize to other types of political identification. The British case is useful in this respect. Due to the 2016 referendum on EU membership. the country now has *two* forms of electoral identification[12]. The first is conventional party identification. The second is identification with either the Leave or Remain side at the EU referendum. Further, recent evidence shows that the latter also affects how voters report to perceive the economy (Fieldhouse, Green, Evans, Mellon, and Prosser 2020; Sorace and Hobolt 2018). As the Leave side won, supporting it is now, for all intents and purposes, akin to supporting the incumbent party. By the same logic, supporting Remain is now akin to supporting an opposition party. Accordingly, we should expect any treatment effects to generalize to EU referendum identification in the same way as does party identification.

EU and party identification are not unrelated. But, the former does still cut across the latter to a meaningful extent. Table A1 makes this clear. It shows the proportion of participants who

---

[12]And yet more still in Scotland, where unionist versus nationalist identities rose to prominence as a result of the 2014 referendum on Scottish independence.
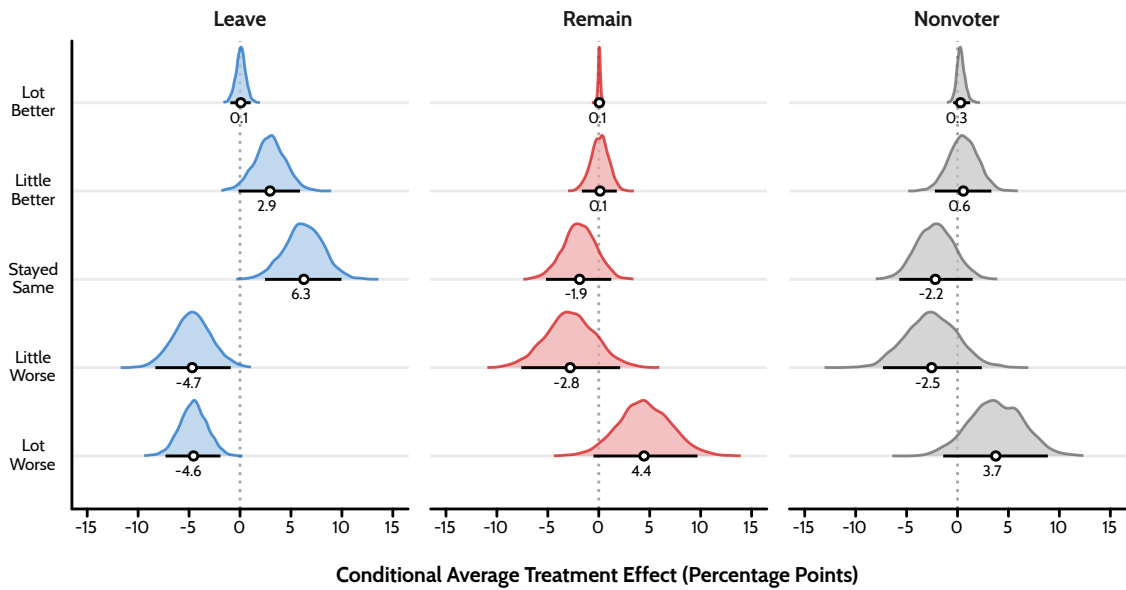
Figure A6: Political surveys cause participants to report different perceptions of the national economy, conditional on their voting behavior at the 2016 referendum on European Union membership. Like with party identification, these effects are most pronounced where they voted for the winning side (Leave). But, in this case, there is also good evidence of a treatment effect on Remain voters too. Positive values imply that those in the treatment group were more likely to report a given response. Negative values imply the opposite. Density plots show the posterior distribution of conditional average treatment effects. Black bars show their 95% credible intervals and point estimates their medians.

voted for each combination of options at the 2016 referendum on EU membership and the 2017 general election. As we can see, participants who voted for the incumbent Conservative Party in 2017 most often voted to leave in 2016. Likewise, those who voted for an opposition party most often voted to remain. But this is not true in all cases. For example, 27.3% of participants who voted for the incumbent Conservative Party also voted to remain in the EU. Similarly, 26.4% of participants who voted for an opposition party also voted to leave. Further, 12.1% of participants voted only in 2016. Thus, we should not expect treatment effects for EU identification to be mere reflections of those across party identification.

Fortunately, the fourth question on the political survey primed voters to consider how they voted at the 2016 referendum (see appendix B). Figure A6 shows the corresponding treatment effects. In this case, Leave supporters in the treatment group were -4.7 percentage points (95% CI: -8.3 to -0.9) less likely to report either that the economy "got a little worse" and -4.6 percentage

points (95% CI: -7.3 to -1.9) less likely to say that it "got a lot worse". They were also 6.3 percentage points (95% CI: 2.5 to 10.0) more likely to report that it had "stayed the same" and 2.9 percentage points (95% CI: -0.1 to 5.9) more likely to report that it "got a little better". Again, almost no one said that the economy "got a lot better" and there was no meaningful treatment effect (0.1, 95% CI: -0.9 to 1.0).

Those who voted Remain also showed similar effects to opposition voters. Yet they were much more likely to say that the economy "got a lot worse" in the last twelve months. This effect was large (4.4, 95% CI: -0.5 to 9.7). Further, though its 95% credible interval crossed zero, 96% of the posterior distribution was greater than zero. Thus, we can be reasonably confident that the true effect is, in fact, greater than zero. Likewise, given these results, we can also be confident that the treatment generalizes to other types of electoral identification too.

Sensitivity of Treatment Effects to Modeling Assumptions

Ordered regression models estimate effects that are consistent across threshold parameters and, thus, across responses. This is known as the proportional odds assumption (Agresti 2010; Mccullagh 1980). Consider the present case. The treatment has a positive effect on the national economic perceptions that incumbents report when measured on the probit scale (see Table A2 in the appendix). This is why they are more likely to say that the economy "got a little better" or "stayed the same" and less likely to say that the economy "got a little worse" or "got a lot worse". But, of course, this assumption may not hold. Instead, the treatment might have a unique effect on each response option.

To relax this assumption, we can use multinomial regression instead. Figure A7 shows the resulting estimates from such a model. Note that the multinomial model is less efficient and, thus, estimates tend to be less precise. Even so, they still lead to the same conclusion: that political survey effects are most clear where participants voted for the incumbent at the last election. Here, incumbent voters were -7.3 percentage points (95% CI: -13.4 to -1.6) less likely to say that the economy "got a little worse" and -2.5 percentage points (95% CI: -5.5 to 0.7) that it "got a lot worse". They were also 5.8 percentage points (95% CI: 0.0 to 11.6) more likely to say that the
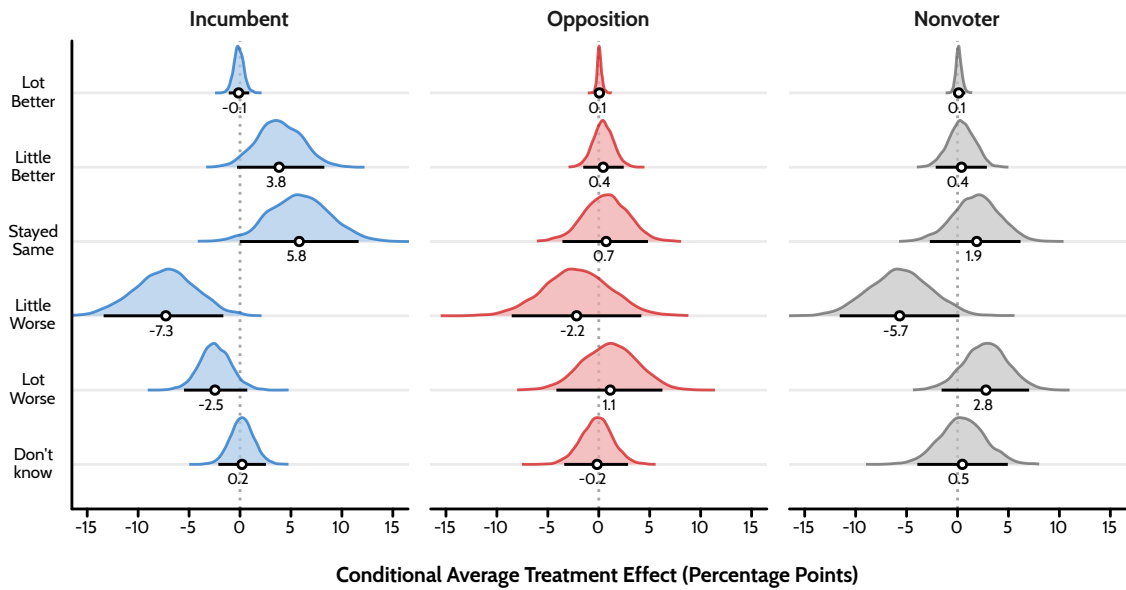
Figure A7: Using a multinomial rather than an ordinal model does little to change the results. We still find that political surveys cause incumbent voters to report more positive economic perceptions (left panel). Positive values imply that those in the treatment group were more likely to report a given response. Negative values imply the opposite. Density plots show the posterior distribution of conditional average treatment effects. Further, black bars show their 95% credible intervals and point estimates their medians.

economy had "stayed the same" and 3.8 percentage points (95% CI: -0.3 to 8.3) more likely to say that it "got a little better". Results for opposition supporters and non-voters differ little to the results in figure 3. Further, there appear to be no difference in the propensity of respondents to answer "Don't know" under the political survey treatment compared to the non-political survey control. As such, my conclusions appear robust to both model specification and missing data.

Table A2: Incumbent supporters report more positive economic perceptions when they answer in a political versus a non-political survey. Data come from a survey experiment conducted by YouGov between 6th and 8th November 2019, the start of the campaign for the 2019 UK General Election.

	Median	Error	2.5%	97.5%
Mean				
Treatment	−0.02	0.07	−0.16	0.12
Incumbent	0.41	0.07	0.29	0.54
Opposition	−0.35	0.07	−0.49	−0.22
Treatment × Inc.	0.20	0.09	0.02	0.38
Treatment × Opp.	0.05	0.09	−0.14	0.22
Discrimination				
Treatment	−0.10	0.06	−0.21	0.02
Incumbent	0.05	0.06	−0.06	0.16
Opposition	0.14	0.06	0.02	0.25
Treatment × Inc.	0.18	0.08	0.03	0.33
Treatment × Opp.	0.01	0.08	−0.16	0.16
Threshold				
Threshold 1	−0.94	0.07	−1.07	−0.81
Threshold 2	0.31	0.05	0.21	0.42
Threshold 3	1.40	0.08	1.25	1.56
Threshold 4	2.52	0.13	2.26	2.79
N (Individuals)				2,487

Table A3: Respondents do not differ in how they report their personal economic perceptions when they answer in a political versus a non-political survey. Data come from a survey experiment conducted by YouGov between 6th and 8th November 2019, the start of the campaign for the 2019 UK General Election.

	Median	Error	2.5%	97.5%
Mean				
Treatment	−0.08	0.07	−0.21	0.05
Incumbent	0.04	0.06	−0.09	0.16
Opposition	−0.17	0.07	−0.29	−0.04
Treatment × Inc.	0.15	0.09	−0.02	0.32
Treatment × Opp.	0.01	0.09	−0.17	0.19
Discrimination				
Treatment	−0.01	0.06	−0.12	0.10
Incumbent	0.13	0.05	0.03	0.24
Opposition	0.04	0.05	−0.07	0.14
Treatment × Inc.	0.02	0.07	−0.13	0.16
Treatment × Opp.	0.00	0.07	−0.14	0.15
Threshold				
Threshold 1	−1.37	0.08	−1.52	−1.23
Threshold 2	−0.44	0.05	−0.55	−0.34
Threshold 3	0.69	0.06	0.58	0.81
Threshold 4	1.68	0.09	1.52	1.85
N (Individuals)				2, 591

Table A4: Like party identification, those who voted for the winning side at the 2016 referendum on European Union membership report more positive economic perceptions when they answer in a political versus a non-political survey. Data come from a survey experiment conducted by YouGov between 6th and 8th November 2019, the start of the campaign for the 2019 UK General Election.

	Median	Error	2.5%	97.5%
Mean				
Treatment	−0.07	0.08	−0.22	0.09
Leave	0.35	0.07	0.22	0.50
Remain	−0.28	0.07	−0.42	−0.14
Treatment × Leave	0.27	0.10	0.08	0.46
Treatment × Remain	−0.03	0.10	−0.23	0.16
Discrimination				
Treatment	−0.09	0.07	−0.22	0.04
Leave	0.02	0.06	−0.09	0.14
Remain	0.05	0.06	−0.07	0.17
Treatment × Leave	0.18	0.08	0.02	0.35
Treatment × Remain	0.02	0.09	−0.14	0.19
Threshold				
Threshold 1	−0.96	0.08	−1.12	−0.81
Threshold 2	0.32	0.06	0.20	0.44
Threshold 3	1.42	0.09	1.25	1.59
Threshold 4	2.51	0.15	2.23	2.80
N (Individuals)				2, 459

Table A5: That incumbent supporters report more positive economic perceptions when they answer in a political versus a non-political survey is robust to model specification. Data come from a survey experiment conducted by YouGov between 6th and 8th November 2019, the start of the campaign for the 2019 UK General Election.

	Median	Error	2.5%	97.5%
Got a lot worse				
Intercept	−0.50	0.13	−0.76	−0.24
Treatment	0.10	0.17	−0.24	0.43
Incumbent	−1.06	0.19	−1.44	−0.70
Opposition	0.80	0.17	0.46	1.14
Treatment × Inc.	−0.58	0.27	−1.12	−0.06
Treatment × Opp.	−0.09	0.23	−0.52	0.36
Got a little worse				
Intercept	0.51	0.11	0.30	0.73
Treatment	−0.22	0.14	−0.50	0.06
Incumbent	−0.62	0.14	−0.90	−0.35
Opposition	0.62	0.15	0.32	0.90
Treatment × Inc.	−0.15	0.19	−0.54	0.23
Treatment × Opp.	0.14	0.20	−0.25	0.53
Got a little better				
Intercept	−1.34	0.17	−1.68	−1.03
Treatment	−0.02	0.21	−0.42	0.38
Incumbent	0.18	0.20	−0.21	0.57
Opposition	−0.32	0.25	−0.83	0.17
Treatment × Inc.	0.16	0.25	−0.35	0.64
Treatment × Opp.	0.10	0.31	−0.52	0.71
Got a lot better				
Intercept	−3.59	0.35	−4.32	−2.94
Treatment	0.10	0.36	−0.60	0.82
Incumbent	0.04	0.36	−0.64	0.73
Opposition	−0.29	0.41	−1.10	0.49
Treatment × Inc.	−0.36	0.40	−1.18	0.41
Treatment × Opp.	0.03	0.43	−0.81	0.88
Don't know				
Intercept	−0.75	0.17	−1.09	−0.43
Treatment	−0.04	0.23	−0.48	0.43
Incumbent	−1.80	0.32	−2.46	−1.20
Opposition	−0.28	0.27	−0.84	0.24
Treatment × Inc.	−0.04	0.45	−0.93	0.81
Treatment × Opp.	−0.04	0.38	−0.78	0.72
N (Individuals)				2,671