Erroneous Beliefs and Political Approval: Evidence from the Coronavirus Pandemic*

Matthew Lilley[†] and Brian Wheaton[‡]

June 23, 2021

Abstract

Are politicians rewarded for good performance? This requires that public perceptions of performance are accurate. Examining the case of the coronavirus pandemic, we conduct an incentivized survey asking respondents how pairs of states have performed relative to one another in terms of deaths per-capita. We find that beliefs are only modestly more accurate than random guessing, and it is not the actual death rate - but rather *beliefs* about the death rate – that drive governor approval. We replicate these findings in both an identical follow-up survey later in the pandemic and in an additional survey with experimental variation.

*We are grateful to Harvard Business School for grant funding and to Robert Barro, Ben Enke, Ed Glaeser, and Gautam Rao for excellent comments and advice. We also thank the participants of the Behavioral Economics workshop at Harvard.

[†]Harvard University Department of Economics. Contact: matthewlilley@fas.harvard.edu

[‡]Harvard University Department of Economics. Contact: bwheaton@g.harvard.edu

The question of whether politicians are rewarded for good performance and penalized for bad performance is a matter of paramount political-economic importance. This question – central to models of retrospective voting – is crucial because the existence of such rewards/penalties may incentivize elected leaders to pursue socially beneficial outcomes, helping ensure the accountability of elected government to its constituents and the healthy functioning of democracy. A government that is able to generate perceptions of good performance despite poor actual performance may be able to evade responsibility for its actions.

In order for politicians to be rewarded or penalized in this way, however, it is first necessary that public perceptions of performance be at least somewhat accurate. A crucial challenge is that it is often difficult to objectively measure performance. First, there are a multitude of dimensions of both the policies pursued by politicians and the outcomes over which they preside – many of which may be difficult to measure in any objective sense. Second, it can be unclear what role politicians have on each of these dimensions. For example, a growing literature studies the extent to which leaders have actual effects on economic growth, and its findings have been mixed. All of these factors may lead to imperfectly-accurate perceptions of performance.

The precise questions that emerge from these observations are (i) whether voters do actually have accurate beliefs about performance, (ii) whether politicians are rewarded for having good outcomes or merely for being perceived as having good outcomes, and (iii) whether inaccurate beliefs yield any cost to society. To answer these questions, we study the Coronavirus Pandemic of 2020-21, which we regard as a setting highly amenable to the investigation of our research questions. During the pandemic, the entire apparatus of state government shifted its priorities toward managing and mitigating coronavirus. Plentiful data on coronavirus cases, testing, and deaths was available at the state level (and finer geographies) on a daily basis. Governors possessed an extraordinarily wide degree of latitude to implement policy responses of their choosing, with comparatively little encumbrance from legislatures. Meanwhile, they also became the highly-visible public faces of their states' efforts, with some – such as New York's Andrew Cuomo and California's Gavin Newsom – holding daily or weekly coronavirus briefings. Furthermore, many opinion polls throughout the period focused specifically on public approval of their governor's handling of the coronavirus pandemic. All this renders the pandemic an ideal setting for studying the accuracy of public perceptions about the performance of their leaders – and the implications of that accuracy.

We conducted an incentivized mTurk survey at the end of July 2020 (during the pandemic's "summer wave"), primarily asking respondents to provide their best guess of how pairs of states performed relative to one another in terms of deaths per-capita. We additionally asked a variety of demographic questions, questions about political identification, and benchmarking questions designed to gauge respondents' perceptions of how well the states should have performed, given preexisting characteristics such as their population density and setting aside factors of leadership/political competence. The survey consisted of approximately 400 mTurk Masters located in the United States, each of whom was compensated a base rate of $1.50 along with a potential incentive bonus for answering the primary questions correctly. We subsequently ran an identical survey three months later, at the end of October 2020 (during the beginning of the pandemic's fall/winter wave).

We find that individuals perform better than random guessing in their pairwise comparisons of state performance – but not substantially better. Respondents only correctly guessed which state performed worse 63.4% of the time. Their performance was an identical 63.4% when comparing their home state to another state. Respondents tended to think that states like Florida and Texas – which received substantial critical media coverage – performed substantially worse than they actually did. We investigated whether there existed any in-group bias in beliefs, finding at most weak evidence of Republican (Democratic) respondents holding biased beliefs about how positively Republican (Democratic) states performed, in relative terms. These results were fairly stable across both the July and October waves of the survey.

Using data from SafeGraph and the Understanding America Study on social distancing behavior, we show that these beliefs about state's performance – and not actual state performance – have bearing on social distancing behavior. Individuals engage in less social distancing when their state is erroneously perceived to have performed better in terms of coronavirus deaths. We take this both as evidence that the measured beliefs are real and that erroneous beliefs may distort behavior in a way potentially harmful to society.

Next, we turn to the question of whether politicians are rewarded for good outcomes or merely perceptions of good outcomes. To do this, we regress respondents' guesses about death rates on state fixed-effects in order to provide a measure of how badly people think each state is doing. Next, using opinion-polling data from The COVID States Project on state-level approval of governor handling of the pandemic, we regress these measures of approval on the actual state death rate and these aforementioned fixed effects that capture beliefs. We find that it is not the actual death rate – but rather beliefs about the death rate – which drive governor approval. Controlling for beliefs about the death rate, the effect of a higher actual death rate on approval is actually positive, consistent with a potential role for governor media visibility (which tended to be higher in harder-hit states) in boosting approval. As an alternative approach, instead of using the COVID States opinion-polling data on approval of governor coronavirus handling, we use an identical question internal to the survey and run individual-level regressions of governor approval on respondents' beliefs about deaths in their home state. This yields the same result – incorrect beliefs strongly affect political approval. All these results, too, are highly stable across both the July and October waves of the survey.

We argue that, in these regressions, making a causal interpretation is reasonable, as reverse causality would entail disliking the politician in question and consequently having negatively-biased beliefs about the state's coronavirus performance. If this was widely the case, we should expect to observe substantial partisan in-group bias in beliefs about states' performance, but as previously noted, we find minimal such bias in the data. Furthermore, the result is robust to the addition of a broad variety of demographic and political control variables, which should net out effects due to pre-existing attitudes to politicians. However, to gain further evidence on causality, we ran an additional survey in December 2020 – this one leveraging experimental variation. Firstly, given that respondents are imperfectly informed about state performance, we elicit governor approval conditional on different counterfactual levels of performance in terms of coronavirus deaths. We find that conditional governor approval is falling sharply in the hypothetical death rate. Second, we shock respondent beliefs about their state's performance (by eliciting their priors and providing them with the true information), and elicit their ex post governor approval. Exogenously inducing higher beliefs about the number of deaths corresponds to lower governor approval. That is, in both experiments, respondents' approval of their governor moves in the expected direction. Taken as a whole, these results suggest complications for retrospective models of voting relying on accurate perceptions of performance.

Prior Literature

Our work relates most directly to the broad literature on retrospective voting, which originated over a half-century ago. Key (1966) seminally argued that "voters are not fools" – that is, that they update their beliefs and actions based on government performance, rewarding or punishing politicians accordingly. Key's informal intuition was subsequently formalized in models by Barro (1973) and Ferejohn (1986). In these models, by re-electing high-performing politicians and voting out poorly-performing ones, voters incentivize good performance by politicians (and thus good outcomes). These theories represented an important divergence from the theretofore standard conception of the voter as mostly lacking in information and voting entirely on the basis of promised future political outcomes rather than past performance. On the empirical front, a large subset of this literature, beginning with Kramer (1971), Fair (1978), and Fiorina (1981), has studied whether voters reward or penalize politicians for economic outcomes, which are taken as objective performance indicators.

Later theoretical frameworks enriched the mechanisms underlying retrospective voting. Persson & Tabellini (2002) and Duch & Stevenson (2008) view retrospective voters as learning about incumbent quality through incumbent performance during his/her period in office. Voters then choose between re-electing an incumbent leader of known quality or voting the incumbent out of office and taking a new draw from the quality distribution. Ashworth (2005) models the effects that such a mechanism have on politician decision-making and effort allocation over the course of a career. Recent empirical papers have exploited a variety of natural experiments (e.g., Alt et al. (2011), Gasper & Reeves (2011), Bechtel & Hainmueller (2011), Reeves & Gimpel (2012), Stokes (2016)) and controlled experiments (e.g., Malhotra & Kuo (2008), Malhotra & Margalit (2014)).

An important subset of this literature has focused on how behavioral biases and cognitive limitations might interact with the concept of retrospective voting. In a complex world, voters may choose to rely on heuristics rather than evaluate all information carefully. This strand began with the observation (initially made by Kramer (1971), Fair (1978), and Tufte (1978)) that the election-year economy appeared to have larger impacts on voting behavior than conditions in other years of the incumbent's tenure, suggesting a form of availability bias. Huber et al. (2012) and Healy & Lenz (2014) examine this phenomenon in more detail. More generally, it has been argued that voters reward or punish politicians because they are happy or sad for reasons that have nothing to do with incumbent performance, such as foreign economic conditions or football games (Schwarz & Clore (1983), Achen & Bartels (2004), Wolfers (2007), Healy & Malhotra (2010), Campello & Zucco (2016), Busby et al. (2017)). These findings are often attributed to a combination of behavioral biases and difficulties in attributing responsibility for outcomes.

Our work relates most closely to this strand of the literature, which probes the behavioral contours of retrospective voting. One notable contrast is that whereas many of these prior studies focus on the effects of irrelevant outcomes (e.g. football games) on voting, our setting enables us to analyse how voters respond to politician actions that they may legitimately (and plausibly) believe are able to substantially impact outcomes. Further, we are able to precisely distinguish between the true and false components of beliefs held by the public about an important outcome (coronavirus mortalities in their home state), thereby enabling us to directly assess the importance of behavioral considerations relative to rational considerations underlying the mechanism of retrospective voting and reward/punishment of politicians. We also contribute by examining a domain (mortalities from the coronavirus pandemic) distinct from the standard economic outcomes most typically studied in the retrospective voting literature.

More broadly, this paper relates to some of the key insights of public choice theory (Olson (1971), Downs (1957)). For many political matters, there is little private benefit for holding correct beliefs: the effects of a politician's performance, good or bad, cannot be avoided or amplified merely by knowing about them, and the chance of a voter being pivotal in an election are infinitesimally small. Accordingly, people often have limited incentive to obtain correct beliefs about political matters, and instead choose to be *rationally ignorant* (Downs (1957)). Further, in many domains, most people have limited ability to obtain correct beliefs - the world is full of rich, often seemingly conflicting, information, from which discerning the truth can be difficult. Given these frictions, if politicians think they are able to manipulate beliefs, it should be unsurprising that they would expend effort to do so.

Data

Our primary data obtained by conducting two separate surveys on Amazon Mechanical Turk. The first elicited beliefs about the pandemic, whereas the second sought to exogenously shock those beliefs.

In the first survey, participants were asked, for 10 randomly-drawn pairs of states, to guess which state in the pair had fared worse up to that point in terms of coronavirus mortalities per capita.[1] The question about each pair was immediately followed-up by a more precise question asking how much worse, as a percent-

[1]The first 5 of the 10 pairs were constrained to include the respondent's home state as one of the states in the pair, since individuals may plausibly have a more accurate picture of the pandemic situation in their home state.

age, they believe their chosen state had fared. Next, for 5 randomly-drawn pairs of states, participants were asked their beliefs about which state would have performed worse (and how much worse) due to pre-existing non-political factors such as population density, population age, presence of international travelers, and anything else they deemed relevant. Also included in the survey were demographic questions on sex, age, race, education, income, and state of residence. Political questions – respondents' Presidential election vote in 2016, their party identification, and the level of their approval for their governor's handling of the pandemic – were also asked. The timing of the initial wave of our survey roughly corresponded to the summer peak of coronavirus cases and deaths. Later, to validate our results and check for consistency, we subsequently ran an identical second wave of the survey in the fall. This timing corresponded to the beginning of the fall/winter increase in coronavirus cases and deaths.

In the second survey, we conducted an information-revelation experiment. In the experiment, we randomly assigned participants to either a treatment group, a control group, or a hypothetical group. We asked the control group of participants for their guess of coronavirus deaths and pandemic employment declines within their home state, followed by a question on the level of their approval of their governor's handling of the pandemic. The treatment group was provided with information on the true figures before being asked about their approval of their governor's handling of the pandemic. The disparity between their priors and the true information induces a shock to the beliefs of respondents in the treatment group, allowing us to discern the effect (if any) of beliefs on governor approval. The hypothetical group was asked a series of hypothetical questions: whether they would approve of their governor's handling of the pandemic if they learned that the true coronavirus death rate (or the true decline in employment since the start of the pandemic) was X, for a variety of values of X (at least one of which is true).

Additional technical details about the surveys are available in Appendix B.

Social Distancing Data

We primarily draw on social distancing data from two sources: SafeGraph and the USC Understanding America Study. SafeGraph compiles measures of mobility based on cellphone usage and location patterns. Since the early stages of the coronavirus pandemic, SafeGraph has made a Social Distancing Metrics dataset freely available to academic researchers for the purpose of studying behavior during the pandemic. This dataset contains measures such as time spent at home, time spent outside the home, and distance traveled. It has been widely used in the economics literature on the coronavirus pandemic. However, the SafeGraph social distancing data is noisy. For example, the median number of total minutes logged per day (at home plus outside the home) varies substantially across places due to factors such as variation in the amount of time people keep location services activated on their phones and variation in restrictions by service providers with regard to how frequently SafeGraph can ping cellphones. To combat these challenges, we leverage the fact that the data is available at the daily level from January 2019 onward; we generate normalized versions of the metrics - computing the change in behavior relative to one year prior.

We additionally use social distancing data from the USC Understanding America Study. Since March 10, 2020, USC has asked their panel of Understanding America Study respondents a series of questions about the coronavirus pandemic, including a few about social distancing behavior (such as whether the respondent has gone outside in the past 7 days or had any close contact with non-household members over the past 7 days). This individual-level data is freely available to academic researchers.

Governor Approval and Other Outcomes

Since March 2020, The Covid States Project,[2] a multi-university group of multidisciplinary researchers has released a variety of periodic reports on the status of the pandemic and related indicators at the state level. Amongst these reports have been state opinion polling data on approval of governor handling of the pandemic, termed Executive Approval reports by the Project. This data is publicly-available online through the Project's website and we utilize it as our key outcome, using opinion-polling data from their July Wave with our July mTurk survey and opinion-polling data from their October Wave with our October mTurk survey.

As noted, we collect data on approval of governor handling of the coronavirus pandemic in our mTurk survey as well, to obtain an additional source of this data.

We obtain data on each state's 2016 Presidential Election victor and margin of victory from Dave Leip's Election Atlas.[3]

Empirical Strategy

In order to identify the effects of both actual coronavirus deaths and beliefs about coronavirus deaths, we apply the following procedure. We first note that our data on our main survey question is at the state-pair level. Respondents are asked which of State A and State B they believe has experienced a higher coronavirus death rate and, subsequently, the factor by which they think deaths are higher in their chosen

[2] covidstates.org

[3] https://uselectionatlas.org/

state. From this, we construct the logarithm of the ratio of the coronavirus death rates in the two states. The logarithmic transformation ensures the data is coherently normalised and allows for ease of interpretation of subsequent regression coefficients. Each observation can be represented as either the log of the factor, X, by which State A's coronavirus death rate exceeds State B's coronavirus death rate or the log of the factor $1/X$ by which State B's coronavirus death rate exceeds State A's coronavirus death rate.

From this data, we wish to extract an estimate of average beliefs about each state's death rate relative to other states. Since each observation pertains to the relative level of deaths in a state pair, (separately) for each survey wave we regress respondents' guesses about relative death rates on state indicator variables as follows, in order to estimate state fixed-effects:

$$\log X_{isr} = \gamma_s + \delta_r + u_{isr}, \tag{1}$$

where X_{isr} denotes the guess of respondent i about the factor by which the death rate of state s exceeds the death rate of state r, δ_r and γ_s denote state fixed-effects for state r and state s respectively, and u_{isr} is the error term. The estimated fixed effect for each state can be extracted as an estimate of beliefs regarding a state's death rate, as desired.

However, an immediate challenge arises. There is no convincing theoretical reason for any particular rotation of any particular observation, namely which state should be considered s and which should be considered r. Further, with a set of fixed effects for s and another for r, this regression generates two separate estimates of beliefs about each state. The two sets of point estimates will not in general be equal, will vary based on arbitrary rotation of datapoints, and it is unclear which (or what combination of them) should be interpreted as beliefs.

Fortunately, there exists a simple and elegant fix that works by negating the arbitrary nature of rotation decisions. That is, we duplicate each observation in the dataset, representing each observation with both rotations and weighting each by half in our analysis. By construction, this yields $\gamma_s = -\delta_s$.[4] The estimated γ_s vector thus provides a measure of how badly, on average, people think state s is doing in terms of coronavirus death rates, with a higher value corresponding to perceptions of higher deaths.

We exploit this same procedure to extract several other pertinent measures from the survey data. First, we again take respondents' guesses about death rates and construct the extent of erroneous beliefs held by individual i about how much higher death rates are in state s relative to r. This involves dividing the guessed ratio by the true ratio, or in logarithms,

$$\log \tilde{X}_{isr} = \log X_{isr} - \log \frac{d_s}{d_r},$$

where d_s and d_r are the respective per capita death rates. Second, we use our survey question on what respondents expected the relative death rate B_{isr} should be in a given state pair, taking into account factors like population density and age, while putting aside factors of political competence. In these two cases cases, we re-estimate Equation (1) replacing X_{isr} with $\tilde{X}_{isr}$ and B_{isr} in turn, and extract the resulting state fixed effect estimates. This yields state-level measures of average er-

[4]To see this, suppose the OLS estimates are $\gamma_s \neq -\delta_s$, and note that $y_{ijk} = -y_{ikj}$. This yields fitted values $\hat{y}_{ijk}$ and $\hat{y}_{ikj}$, and residuals $\hat{u}_{ijk}$ and $\hat{u}_{ikj}$. Consider alternate candidate solution vectors $\tilde{\gamma_s} = (\gamma_s - \delta_s)/2, \tilde{\delta_s} = (\delta_s - \gamma_s)/2$ (such that $\tilde{\gamma_s} = -\tilde{\delta_s}$). This yields fitted values $\tilde{\hat{y}}_{ijk} = (\hat{y}_{ijk} - \hat{y}_{ikj})/2 = -\tilde{\hat{y}}_{ikj}$ and analogously residuals $\tilde{\hat{u}}_{ijk} = (\hat{u}_{ijk} - \hat{u}_{ikj})/2 = -\tilde{\hat{u}}_{ikj}$. For any real scalars $a \neq b, a^2 + b^2 > 2 * [(a+b)/2]^2$, so this constitutes an improvement under the OLS objective function, a contradiction. Note that an analogous argument holds when $y_{ijk} = y_{ikj}$ (in which case, $\gamma_s = \delta_s$) or with an additive constant in either case ($y_{ijk} = a \pm y_{ikj}$).

roneous beliefs about deaths rates, and benchmark expected death rates abstracting from political competence, for each state.

Next, with these state level measures in hand, we regress our outcomes of interest – most notably, political approval – on these estimated fixed-effects and on the natural logarithm of the actual state death rate from coronavirus. That is,

$$Y_s = \alpha + \beta_1 \log DeathsPerMil_s + \beta_2 \cdot \hat{\gamma}_s + \varepsilon_s, \tag{2}$$

where Y_s is a state-level outcome of interest (such as governor approval rate for handling of the pandemic), $DeathsPerMil_s$ is the actual coronavirus death rate per million population, $\hat{\gamma}_s$ are the fixed-effects estimated in the preceding regression, and ε_s is the error term. Thus the effect of a 1% increase in actual deaths on the outcome Y is given by $\beta_1/100$. The effect of a 1% increase in believed deaths on the outcome Y is given by $\beta_2/100$. As noted, we weight each observation by one-half, and we also use robust standard errors. In specifications where we pool the first and second wave of our survey, we add a fixed-effect for the wave and cluster by state. For robustness, we run additional specifications with an assortment of demographic and political control variables added to the above regression equation. This is done to partial out any correlation of beliefs with these controls, which themselves may plausibly drive political approval.

We also run a specification where we focus on the effect of erroneous beliefs by replacing $\hat{\gamma}_s$ with our measure of erroneous beliefs about state j. This yields a regression with a slightly modified interpretation of the coefficients: β_1 now corresponds to the effect of the actual deaths, holding the error in beliefs constant (rather than holding beliefs themselves constant), while β_2 corresponds to the effect of the erroneous component of beliefs.[5] Since political approval may effectively handicap each governor based on how exposed people deem their state as being due to pre-existing factors (largely) outside of government control, we also add the measure of expected death rates abstracting from political competence as a control variable to Equation (2). We then run an alternative version of all of the above specifications wherein we use respondents' binary guess about which state has a higher death rate (as opposed to the precise continuous factor by which the state has more deaths per capita).

Finally, for the governor approval outcome, we run an alternative one-step regression leveraging our internal mTurk survey data on respondent approval of governor handling of the coronavirus pandemic. Because we have individual-level data on this outcome, it is not necessary to generate state fixed-effects for use in a second-stage regression. We can instead retain the observations involving the respondent's home state and run a version of regression equation (2) which, in place of $\hat{\gamma}_s$, directly includes the mean log of the respondent's guesses X_{isr} of how much worse his home state, s, has performed relative to some randomly-selected states r. That is,

$$Y_{is} = \alpha + \beta_1 \log DeathsPerMil_s + \beta_2 \overline{\log X_{is}} + \varepsilon_{is}. \tag{3}$$

[5]Since the true log death rate ratio $\log {}^{d_s}/_{d_r}$ is a state-pair (or state-pair-wave) specific constant, the calculated erroneous beliefs are identical (up to an additive constant) to $\hat{\gamma}_s - log(d_s)$. As a result, in these subsequent regressions, the coefficients on this measure of erroneous beliefs, controlling for the log of actual death rates, are identical to the coefficients on beliefs controlling for the log of actual death rates.

Results

In Table 1, we first report some simple descriptive statistics about the characteristics of our sample. Average age, share male, and median household income of the sample are consistent with the U.S. general population. The sample, however, has a somewhat higher education, share of non-Hispanic whites, and share of liberals/Democrats than the U.S. general population. In certain specifications, we control for these variables in order to ensure that the deviations from representativeness have no effect on our results.

Accuracy and Bias in Beliefs

Survey respondents correctly guessed which state had performed worse (through the date of the survey) in terms of coronavirus death rates only 63.4% of the time.[6] Restricting only to state pairs involving the respondent's home leaves this figure almost exactly unchanged (also 63.4%). Respondents beliefs about their home state are no more accurate than their beliefs about randomly-selected other states. Respondents also performed poorly when estimating states' relative per capita death rates, with a mean (absolute) error of 101 log points. For each log point higher that a state's death rate actually is, respondents estimate deaths to be only 0.28 log points higher. While this weak relationship may merely indicate imperfect knowledge, respondent beliefs are also poorly calibrated: for every log point higher respondents predict a state's death rate to be, deaths are only actually 0.38 log points higher.[7]

[6] Since completely uninformed random guessing would yield a 50% correct rate, this is consistent, for example, with respondents only truly knowing the answer in 26.8% of cases.

[7] Rational expectations imposes no restriction on the former statistic, but requires a unit change in beliefs to predict a unit difference of the actual outcome. Under perfect knowledge, both coefficients are one.

Figure 1 displays a scatterplot of the relative frequency with which survey respondents guess each state had a higher death rate (than states with which it is being compared in the pairwise questions) against the relative frequency it actually had a higher death rate. This reveals which states actually performed better than respondents believe (those above the 45-degree reference line) and which states performed worse. As can be seen from the scatterplot, the states with the largest positive gap between actual and perceived performance (i.e., those most erroneously perceived as performing poorly) are Texas and Florida – two states which received particularly negative media coverage despite having moderate contemporaneous death rates. States with the largest negative gap between actual and perceived performance (i.e., those most erroneously perceived as performing well) are Colorado, Rhode Island, Connecticut, and Massachusetts – the latter three being New England states which were quite intensely impacted by the first wave of the pandemic but received limited attention in the media relative to New York, which only performed slightly worse but was front-and-center in terms of media coverage in early months of the pandemic.

Given the politically-charged nature of discussions surrounding state performance during the pandemic, one might wonder whether there exists any partisan bias in perceptions of death rates. That is, do Democrats have unjustifiably positive views of the performance of Democratic states while Republicans have unjustifiably positive views of the performance of Republican states? To study this question, we regress the natural logarithm of individual respondents' excess believed deaths (believed minus actual) on an indicator variable for the governor's partisan alignment and a "cross-party" indicator variable for whether the governor is of the opposite political party to the respondent. We also include an indicator variable for whether the respondent is from the state in question. The regression analysis in Table 2 follows this approach. As seen in column (1), there is strong evidence that beliefs about Republican states' death rates are excessively pessimistic relative to beliefs about Democratic states' death rates. Turning to the coefficient on the cross-party indicator variable, there is at most weak evidence of modest partisan in-group bias. Respondents, when considering a state whose governor is of the opposite party affiliation, believe that the state's deaths per-capita are 5.8% higher relative to respondents who share a party affiliation with the governor. This coefficient is small in magnitude and, furthermore, it is only statistically-significant at the 10% level. Columns (2) adds actual log deaths per capita as a control variable, yielding no meaningful change in the estimates of partisan in-group bias. However, adding the control for log deaths causes the coefficient on governor party to become a tightly estimated zero. In other words the negative bias in beliefs about states with Republican governors in column (1) is an artefact of respondents being largely unaware which states had done better, combined with the average Republican-led state then having fewer per capita deaths. Column (3) adds state-by-wave fixed-effects, such that in-group bias is identified only from within-state variation in beliefs; namely, for each state, the partisan difference in beliefs about that state held by respondents. Neither the magnitude or significance of the estimated partisan in-group bias is meaningfully altered. Across specifications, the home state coefficient is close to zero; beliefs about a state differ little irrespective of where the respondent lives.

Effects on Social Distancing Behavior

We next examine whether these erroneous beliefs translate into behavioral differences. During the coronavirus pandemic, efforts to "flatten the curve" of coronavirus cases by encouraging individuals to spend as much time as possible quarantining at home – as opposed to outside – were central to the public health response. Erroneous beliefs about the intensity of the pandemic might lead to distortions in behavior, potentially inflicting costs upon society. To test this, we run versions of the regression specification described in Equation (2) – in this case, with measures of social distancing as the outcome variables. We use SafeGraph's measure of percentage of time spent at home per day along with measures from the Understanding America Study (UAS) of the share of people in each state who went outside or who had close contact with a non-household member at any time in the past 7 days.[8]

Table 3 cycles through the same regression specifications used in the preceding section, now with the median percentage of time at home outcome. The specification in column (1) corresponds to Equation (2). The results show that it is not deaths, but rather beliefs about deaths, which affect the percentage of time individuals spend at home. In particular, a one standard deviation (38 log point) increase in believed deaths translates into a (roughly) 2.28 percentage-point greater increase in the percentage of time individuals spent at home (relative to the analogous period in 2019). Consequently, erroneously believing that a state has had fewer deaths than it actually has may lead to a sub-optimal amount of social distancing behavior. In column (2), we add as controls a variety of state demographics, political factors such as 2016 Trump vote share interacted with governor party, a rolling 7-day average of coronavirus cases and deaths in the state, and beliefs about benchmark death rates (how high a death rate respondents expected given the state's pre-existing non-political characteristics). Statistical significance of the

[8]We use SafeGraph/UAS data for the 30 days following the occurrence of our mTurk survey. That is, we merge observations from the July wave of our survey with corresponding SafeGraph/UAS data on social distancing behavior from the subsequent 30 days after the end of that wave; we merge observations from the October wave of our survey with SafeGraph/UAS data from the subsequent 30 days after the end of that wave.

coefficient on believed deaths is retained, albeit at the 10% level. Columns (3) and (4) repeat this analysis with the share of individuals who went outside in the past 7 days (from the UAS data) as their outcome. Here, too, beliefs of a higher death rate (unlike a higher *actual* death rate) translate into less time spent outside. In terms of magnitudes, column (3) suggests that a one standard deviation increase in believed deaths translates into a 4.94 percentage-point decrease in the share of individuals who went outside in the past 7 days. Columns (5) and (6) find an analogous result for the share of individuals who had close contact with someone outside their household.[9]

Effects on Political Approval (Observational)

We next turn to the key question of how beliefs about state performance affect political approval. Table 4 displays versions of the regression specification described in Equation (2), with state-level average approval ratings of governor handling of the pandemic as the outcome variable. Column (1), however, begins with a univariate regression of governor approval on the log of the death rate. There is evidence of a positive (albeit slightly weak) association between the death rate and governor coronavirus approval. A one standard deviation (90 log point) increase in deaths roughly translates into a 2.6 percentage-point increase in governor approval. This regression, however, masks a more complex relationship. Column (2) is the specification directly corresponding to Equation (2). It reveals that higher believed deaths are associated with lower approval, whereas higher actual deaths are associated with higher approval. A one standard deviation increase in actual deaths translates into a 5.7 percentage-point increase in governor approval; a one

[9] The Covidstates survey data includes some measures of social distancing as well – whether individuals had recently visited with friends, avoided contact with others, avoided crowded places, or limited contact outside their household. Running analogous regressions with these outcomes again yields very similar results.

standard deviation (38 log point) increase in believed deaths translates into a 4.3 percentage-point decrease in governor approval. In other words, the intuitive relationship whereby voters punish politicians for bad outcomes (here, deaths) is entirely driven by perceptions of the outcome, not the outcome itself. This suggests potential challenges to ensuring politicians are properly incentivized through public opinion and voting.

Column (3) contains the results of the analogous specification in which we transform beliefs into "excess beliefs" by subtracting the truth from beliefs. This re-frames the regression specification to yield a slightly different interpretation. Holding the error in beliefs constant, a one standard deviation increase in the death rate is associated with a (non-significant) 4.5 percentage-point decrease in governor approval. Holding actual deaths constant, a one standard-deviation increase in the false component of beliefs about the death rate is associated with a (significant) 7.8 percentage-point decrease in governor approval.

Columns (4) and (5) repeat the exercises of columns (2) through (3), with an added control capturing beliefs about benchmark death rates. These benchmarks measure how high a death rate respondents would have expected in each state given its pre-existing characteristics (e.g., population density, population age, exposure to international travelers, etc.), putting aside factors of political competence. The addition of this benchmark deaths control increases the coefficient on beliefs approximately two-fold. The positive coefficient on benchmark deaths is consistent with governors being graded on a curve based on their state's perceived inherent exposure to the pandemic. The overall conclusion is substantively unchanged. Finally, column (6) adds a variety of state-level control variables to the specification in column (5), including governor pre-pandemic approval rating, governor party indicator variables, 2016 Presidential election margin (Trump minus Clinton) interacted with governor party, and the natural logarithm of the state's past-seven-day average of new coronavirus cases and deaths (in case recent outcomes correlate with beliefs and are also responsible for driving governor approval). The key results are robust to adding these controls, with the estimated effect of beliefs on approval becoming modestly smaller, but more tightly estimated.

In Table 5, this same analysis is repeated, except with the binary version of the variable containing individuals' guesses about which states experienced higher death rates from coronavirus. We provide these results to generate general robustness to functional form changes, but we emphasize that this discrete measure throws away variation regarding the relative extent to which respondents believe states have performed differently. While the magnitude of the coefficients changes slightly, the overall conclusions are largely unchanged.

An additional analysis is conducted in Table 6, now using the specification in Equation (3) along with the individual-level data on governor approval that we collected in our mTurk survey. Since we run these regressions at the individual level, they focus on beliefs about one's home state. As noted previously, the accuracy of respondents' beliefs about death rates in their home states is identical to their accuracy about other states, so there is little reason to expect different results. Indeed, the magnitudes of the coefficients differ slightly from the main specifications – notably, the positive coefficients on the actual deaths variable appear larger. The conclusions, however, are unchanged and the effects remain strongly significant under this alternative approach.

Of course, it must be noted that our measures of individuals' beliefs about state performance are observational rather than the product of experimental variation. Nonetheless, we believe that making a causal interpretation regarding the impact of beliefs on political approval is sound for two reasons. The most obvious threat to identification is some form of endogeneity or reverse causality. For example, individuals who happen to approve of a state's governor may be more likely to hold optimistic beliefs about the performance of that state. However, if this was the case, then we should expect to observe substantial partisanship in beliefs about states' performance, but as shown above, we find minimal such bias in the data. Furthermore, any endogeneity associated with political affiliations of states and their governors should largely be captured by controlling for political and demographic factors about the state (e.g., the partisan lean of the state and the governor's pre-existing approval rating). As we show, our results are robust to adding these controls, and it is plausible that, conditional on these observables, beliefs are exogenous of unobserved determinants of governor approval.

Effects on Political Approval (Experimental)

To provide further evidence on this front, we turn to the results of our survey experiment. Two approaches in our survey experiment yield information about the responsiveness of governor approval to beliefs. The first approach entails simply asking respondents whether they would approve of their governor if, hypothetically, they learned that deaths per capita were X, for a variety of values of X. Specifically, we elicit conditional approval for each of three different deaths rates, one of which is the true death rate for the state and the other two of which are randomly drawn from the set of other states' death rates. To the extent that approval is varying in X, individuals are admitting that their approval of their governor's handling of the pandemic is indeed responsive to their beliefs about how well the pandemic was handled in their state. Columns (1) through (3) of Table 7 regress an indicator for individual-level approval on the natural logarithm of the hypothetical value of deaths per capita the individual is presented with. Column (1) is a simple univariate regression, whereas column (2) adds controls for individual demographic characteristics. Column (3) adds person fixed-effects, identifying solely off within-person variation. In each of these specifications, the result that approval is indeed responsive to beliefs emerges – with very strong levels of statistical significance. A hypothetical one standard deviation increase in deaths per-capita is associated with an approximately 19 percentage-point decline in approval of governor coronavirus handling in each of these specifications. In columns (4) through (6), we provide additional evidence using hypothetical questions pertaining to a different domain – the percent decline in employment since the start of the pandemic. In this domain, too, approval is responsive to beliefs. A hypothetical one standard deviation increase in employment loss is associated with an approximately 9 percentage-point decline in approval.

The second approach to identifying effects of beliefs on governor approval is a more standard information-revelation experiment. We randomize respondents into a control group or a treatment group. In the control group, they are asked to guess their state's performance in terms of deaths per capita and then about the extent to which they approve of their governor's handling of the coronavirus pandemic. In the treatment group, they are asked to guess their state's performance – and then told their state's true performance – before being asked about approval of their governor's coronavirus handling. The treatment group thus receives a shock to their beliefs, allowing us to measure the effect of a shift in these beliefs on governor approval. We do precisely this in Table 8. Column (1) displays the results of the simplest version of such a specification, regressing approval on a treatment group indicator, state (log) deaths per-capita, and the interaction term thereof. The interaction term isolates how the information revealed affects governor approval for treatment group members (relative to control group individuals). This is accordingly the key variable of interest. A one standard deviation shock to believed deaths per capita is estimated to lead to approximately a 14 percentage-point decline in approval of governor coronavirus handling. Column (2) adds controls for individual demographics, while column (3) utilizes only within-state variation. It also controls for prior beliefs, namely believed deaths and benchmark deaths per-capita, both of which are interacted with treatment status. The results are broadly stable across specifications. Columns (4) through (6) repeat the exercise for employment instead of deaths, again finding an analogous effect.

Conclusion

In order to shed light on whether the public rewards (or penalizes) politicians for their performance in office and thereby contribute further to the literature on retrospective voting, we study public perceptions of coronavirus death rates and governor approval ratings during the 2020-21 coronavirus (COVID-19) pandemic. We note that, in order for the public to reward or penalize politicians for their performance, it is necessary for the public to have an accurate understanding of that performance. Errors or biases may ameliorate this ability – and thus undermine the incentive structure for politicians to continue performing well. We ran an incentivized survey on Amazon Mechanical Turk in July (Wave 1) and October (Wave 2) of 2020 asking respondents to provide their best guesses, for 10 randomly-drawn pairs of states, which state had the higher death rate (and by how much). We find that respondents choose the correct state 63.4% of the time. We find little to no evidence of partisan in-group bias, though respondents systematically overestimate death rates in Texas and Florida, states which received substantial media attention despite moderate death rates. Using data on social distancing behavior, we also show that these erroneous beliefs about state performance translate into altered social-distancing behavior.

Turning to the question of how these partially-erroneous beliefs translate into governor approval, we find that governor approval is driven by beliefs about death rates, not actual death rates. This remains true if one controls for individuals' perceptions of how well the states should have performed, setting aside factors of leadership/political competence. We thus conclude that considerations related to imperfect information on the part of the public may generate frictions in the operation of retrospective voting models and ability of voters to reward (penalize) good (bad) performance on the part of politicians.

References

Achen, C. H., & Bartels, L. M. (2004). Blind retrospection: Electoral responses to drought, flu and shark attacks.

Alt, J., Bueno de Mesquita, E., & Rose, S. (2011). Disentangling Accountability and Competence in Elections: Evidence from U.S. Term Limits. *The Journal of Politics*, *73*(1), 171–186.

Ashworth, S. (2005). Reputational Dynamics and Political Careers. *Journal of Law, Economics, & Organization*, *21*(2), 441–466. Publisher: Oxford University Press.

Bechtel, M. M., & Hainmueller, J. (2011). How Lasting Is Voter Gratitude? An Analysis of the Short- and Long-Term Electoral Returns to Beneficial Policy. *American Journal of Political Science*, *55*(4), 852–868. Publisher: Midwest Political Science Association, Wiley.

Busby, E. C., Druckman, J. N., & Fredendall, A. (2017). The Political Relevance of Irrelevant Events. *The Journal of Politics*, *79*(1), 346–350. Publisher: The University of Chicago Press.

Campello, D., & Zucco, C. (2016). Presidential Success and the World Economy. *The Journal of Politics*, *78*(2), 589–602.

Downs, A. (1957). *An Economic Theory of Democracy*. New York, NY: Harper and Row.

Duch, R. M., & Stevenson, R. T. (2008). *The Economic Vote: How Political and Economic Institutions Condition Election Results*. New York: Cambridge University Press.

Fair, R. C. (1978). The Effect of Economic Events on Votes for President. *The Review of Economics and Statistics*, *60*(2), 159–173. Publisher: The MIT Press.

Fiorina, M. P. (1981). *Retrospective Voting in American National Elections*. New Haven, CT: Yale University Press.

Gasper, J. T., & Reeves, A. (2011). Make It Rain? Retrospection and the Attentive Electorate in the Context of Natural Disasters. *American Journal of Political Science*, *55*(2), 340–355.

Healy, A., & Lenz, G. S. (2014). Substituting the End for the Whole: Why Voters Respond Primarily to the Election-Year Economy. *American Journal of Political Science*, *58*(1), 31–47. Publisher: Midwest Political Science Association, Wiley.

Healy, A., & Malhotra, N. (2010). Random Events, Economic Losses, and Retrospective Voting: Implications for Democratic Competence. *Quarterly Journal of Political Science*, *5*(2), 193–208. Publisher: Now Publishers, Inc.

Huber, G. A., Hill, S. J., & Lenz, G. S. (2012). Sources of Bias in Retrospective Decision Making: Experimental Evidence on Voters' Limitations in Controlling Incumbents. *American Political Science Review*, *106*(4), 720–741. Publisher: Cambridge University Press.

Kramer, G. H. (1971). Short-Term Fluctuations in U.S. Voting Behavior, 1896-1964. *The American Political Science Review*, *65*(1), 131–143. Publisher: American Political Science Association, Cambridge University Press.

Malhotra, N., & Kuo, A. G. (2008). Attributing Blame: The Public's Response to Hurricane Katrina. *The Journal of Politics*, *70*(1), 120–135. Publisher: The University of Chicago Press, Southern Political Science Association.

Malhotra, N., & Margalit, Y. (2014). Expectation Setting and Retrospective Voting. *The Journal of Politics*, *76*(4), 1000–1016. Publisher: The University of Chicago Press.

Olson, M. (1971). *The Logic of Collective Action: Public Goods and the Theory of Groups*. Cambridge, MA: Harvard University Press.

Persson, T., & Tabellini, G. (2002). *Political Economics: Explaining Economic Policy*. Cambridge, MA: The MIT Press.

Reeves, A., & Gimpel, J. G. (2012). Ecologies of Unease: Geographic Context and National Economic Evaluations. *Political Behavior*, *34*(3), 507–534.

Schwarz, N., & Clore, G. (1983). Mood, Misattribution, and Judgments of Well-Being: Informative and Directive Functions of Affective States. *Journal of Personality and Social Psychology*, *45*(3), 513–523.

Stokes, L. C. (2016). Electoral Backlash against Climate Policy: A Natural Experiment on Retrospective Voting and Local Resistance to Public Policy. *American Journal of Political Science*, *60*(4), 958–974.

Tufte, E. R. (1978). *Political Control of the Economy*. Princeton, NJ: Princeton University Press.

Wolfers, J. (2007). Are Voters Rational? Evidence from Gubernatorial Elections.

Appendices

A Tables and Figures

Table 1 – Descriptive Statistics of mTurk Sample

	mTurk Sample	U.S. Population (18+)
Share Male	0.485	0.487
Median Age	40	47
Share White, non-Hisp.	0.726	0.628
Share w/ BA or Greater	0.491	0.306
Median HH Income	65,885	65,000
Share Clinton Voters	0.423	0.264
Share Trump Voters	0.274	0.252
Share Liberals	0.483	0.279
Share Democrats	0.426	0.354
Observations	613	-

U.S. Population data is from the 2019 American Community Survey (demographic variables), 2016 Presidential election returns data (voting variables), and 2020 American National Election Study (ideology variables).

Table 2 – Partisan In-Group Bias in Beliefs

	(1)	(2)	(3)
	Excess Believed Deaths (Log)		
	Baseline	+ Actual Deaths	Within State
Republican Governor	0.251***	0.014	
	(0.028)	(0.020)	
Cross Party Governor	0.058*	0.040	0.044**
	(0.032)	(0.025)	(0.022)
Home State	0.011	0.080***	-0.041*
	(0.036)	(0.027)	(0.024)
Deaths Per Capita (Log)		-0.717***	
		(0.015)	
State x Wave FE	No	No	Yes
R-squared	0.0210	0.4434	0.5456
Observations	12024	12024	12024

Robust standard errors clustered by individual respondent. * $p<0.10$, ** $p<0.05$, *** $p<0.01$.

Table 3 – Effects of Actual and Believed COVID Deaths on Social Distancing Behavior

	% Time at Home		Outside		Non-HH Contact	
	(1) Beliefs	(2) Controls	(3) Beliefs	(4) Controls	(5) Beliefs	(6) Controls
Deaths Per Capita (Log)	0.00	-0.00	0.06***	0.06***	0.04	0.06***
	(0.01)	(0.01)	(0.02)	(0.02)	(0.03)	(0.02)
Believed Relative Deaths (Log)	0.06***	0.03*	-0.13***	-0.11*	-0.19***	-0.15**
	(0.02)	(0.02)	(0.03)	(0.06)	(0.05)	(0.06)
Benchmark Relative Deaths (Log)		0.00		-0.04		-0.08
		(0.02)		(0.05)		(0.06)
State Controls	No	Yes	No	Yes	No	Yes
Wave FE	Yes	Yes	Yes	Yes	Yes	Yes
Outcome Mean	-0.0142	-0.0142	0.7645	0.7645	0.6881	0.6881
R-squared	0.3943	0.7181	0.2095	0.3127	0.2570	0.4614
Observations	102	96	102	96	102	96

State controls include political controls and disease controls. Political controls are pre-pandemic Governor approval rating, fixed effects for party of the Governor, and 2016 Trump-Clinton net margin interacted with Governor party. Disease controls are 7 day moving averages of cases and deaths in the state. Robust standard errors clustered by state. * $p<0.10$, ** $p<0.05$, *** $p<0.01$.

Table 4 – Effects of Actual and Believed COVID Deaths (Continuous) on Governor Approval

	(1)	(2)	(3)	(4)	(5)	(6)
	Governor COVID Handling Approval					
	Deaths	(1) + Beliefs	(1) + Excess Beliefs	(2) + Benchmark	(3) + Benchmark	(4) + Controls
Deaths Per Capita (Log)	0.029**	0.063***	-0.050	0.059***	-0.152**	0.054***
	(0.014)	(0.013)	(0.032)	(0.012)	(0.075)	(0.014)
Believed Relative Deaths (Log)		-0.113***		-0.212***		-0.131**
		(0.034)		(0.072)		(0.049)
Excess Believed Relative Deaths (Log)			-0.113***		-0.212***	
			(0.034)		(0.072)	
Benchmark Relative Deaths (Log)				0.118*	0.118*	0.086*
				(0.067)	(0.067)	(0.045)
State Controls	No	No	No	No	No	Yes
Wave FE	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.1165	0.2141	0.2141	0.2691	0.2691	0.5676
Observations	100	100	100	100	100	96

Outcome variable mean is 0.503. State controls include political controls and disease controls. Political controls are pre-pandemic Governor approval rating, fixed effects for party of the Governor, and 2016 Trump-Clinton net margin interacted with Governor party. Disease controls are 7 day moving averages of cases and deaths in the state. Robust standard errors clustered by state. * $p<0.10$, ** $p<0.05$, *** $p<0.01$

Table 5 – Effects of Actual and Believed COVID Deaths (Binary) on Governor Approval

	(1)	(2)	(3)	(4)	(5)	(6)
	Governor COVID Handling Approval					
	Deaths	(1) + Beliefs	(1) + Excess Beliefs	(2) + Benchmark	(3) + Benchmark	(4) + Controls
Deaths Per Capita (Log)	0.029**	0.056***	0.007	0.055***	0.012	0.046***
	(0.014)	(0.015)	(0.019)	(0.015)	(0.025)	(0.014)
Probability Believed Worse		-0.204**		-0.334**		-0.112
		(0.084)		(0.127)		(0.091)
Net Probability Erroneously Believed Worse			-0.121**		-0.108*	
			(0.058)		(0.063)	
Probability Benchmark Worse				0.140	-0.025	0.041
				(0.112)	(0.085)	(0.081)
State Controls	No	No	No	No	No	Yes
Wave FE	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.1165	0.1901	0.1595	0.2067	0.1605	0.5329
Observations	100	100	100	100	100	96

Outcome variable mean is 0.503. State controls include political controls and disease controls. Political controls are pre-pandemic Governor approval rating, fixed effects for party of the Governor, and 2016 Trump-Clinton net margin interacted with Governor party. Disease controls are 7 day moving averages of cases and deaths in the state. Robust standard errors clustered by state. * $p<0.10$, ** $p<0.05$, *** $p<0.01$.

Table 6 – Individual-Level Specifications: Effects of Actual and Believed COVID Deaths on Governor Approval

	(1)	(2)	(3)	(4)	(5)	(6)
	Governor COVID Handling Approval					
	Deaths	(1) + Beliefs	(1) + Excess Beliefs	(2) + Benchmark	(4) + Controls	(5) + Demographics
Deaths Per Capita (Log)	0.098***	0.156***	-0.007	0.160***	0.094***	0.087***
	(0.034)	(0.039)	(0.032)	(0.040)	(0.027)	(0.026)
Believed Relative Deaths (Log)		-0.181***		-0.160***	-0.141***	-0.132***
		(0.042)		(0.044)	(0.032)	(0.034)
Excess Believed Relative Deaths (Log)			-0.147***			
			(0.036)			
Benchmark Relative Deaths (Log)				-0.038	-0.044*	-0.038
				(0.023)	(0.026)	(0.023)
State Controls	No	No	No	No	Yes	Yes
Demographic Controls	No	No	No	No	No	Yes
Wave FE	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.0216	0.0821	0.0649	0.0852	0.1776	0.2993
Observations	612	612	612	612	589	589

Outcome variable mean is 0.489. Because these are individual-level specifications, both the left-hand-side approval rating and the right-hand-side variables contain only observations pertaining to the respondents' home state. State controls include political controls and disease controls pertaining to the state. Political controls are pre-pandemic Governor approval rating, fixed effects for party of the Governor, and 2016 Trump-Clinton net margin interacted with Governor party. Disease controls are 7 day moving averages of cases and deaths in the state. Demographic controls are fixed effects for individual sex, race, household income, education, a quadratic in age, and party ID interacted with the party of an individual's state governor. Robust standard errors clustered by state. * p<0.10, ** p<0.05, *** p<0.01.

Table 7 – Experimental Effects of Beliefs on Governor Approval (Hypothetical Approach)

	Deaths Hypothetical			Employment Hypothetical		
	(1)	(2)	(3)	(4)	(5)	(6)
Deaths Per Capita (Log)	-0.311***	-0.332***	-0.339***			
	(0.034)	(0.031)	(0.038)			
Employment Decline (%)				-0.030***	-0.034***	-0.038***
				(0.008)	(0.007)	(0.009)
Demographics	No	Yes	No	No	Yes	No
Outcome Mean	0.4585	0.4585	0.4585	0.5520	0.5520	0.5520
Person FE	No	No	Yes	No	No	Yes
R-squared	0.1235	0.2516	0.7769	0.0327	0.1758	0.7247
Observations	615	615	615	615	615	615

Demographic controls are dummies for individual sex, race, household income, education, a quadratic in age, and party ID interacted with the party of an individual's state governor. Robust standard errors clustered by individual respondent. * $p<0.10$, ** $p<0.05$, *** $p<0.01$

Table 8 – Experimental Effects of Beliefs on Governor Approval (Information Approach)

	Deaths Information			Employment Information		
	(1)	(2)	(3)	(4)	(5)	(6)
Deaths Information	-0.250**	-0.178	-0.280**			
× Deaths Per Capita (Log)	(0.117)	(0.124)	(0.121)			
Employment Information				-0.044**	-0.047**	-0.060**
× Employment Decline (%)				(0.022)	(0.021)	(0.023)
Information Treatment	Yes	Yes	Yes	Yes	Yes	Yes
State Outcomes	Yes	Yes	No	Yes	Yes	No
State FE	No	No	Yes	No	No	Yes
Initial Beliefs	No	No	Yes	No	No	Yes
Initial Beliefs x Treatment	No	No	Yes	No	No	Yes
Demographics	No	Yes	Yes	No	Yes	Yes
R-squared	0.0191	0.1814	0.3234	0.0420	0.1735	0.3297
Observations	346	344	344	351	350	350

Outcome variable mean is 0.580. Information treatment controls are fixed effects for treatment group. State outcome controls are the log of deaths per capita (columns 1-3) and employment decline (%) (columns 4-6). Initial beliefs are actual and benchmark beliefs of the respondent for the log of deaths per capita (columns 1-3) and employment decline (%) (columns 4-6). Demographic controls are fixed effects for individual sex, race, household income, education, a quadratic in age, and party ID interacted with the party of an individual's state governor. Robust standard errors. * $p<0.10$, ** $p<0.05$, *** $p<0.01$

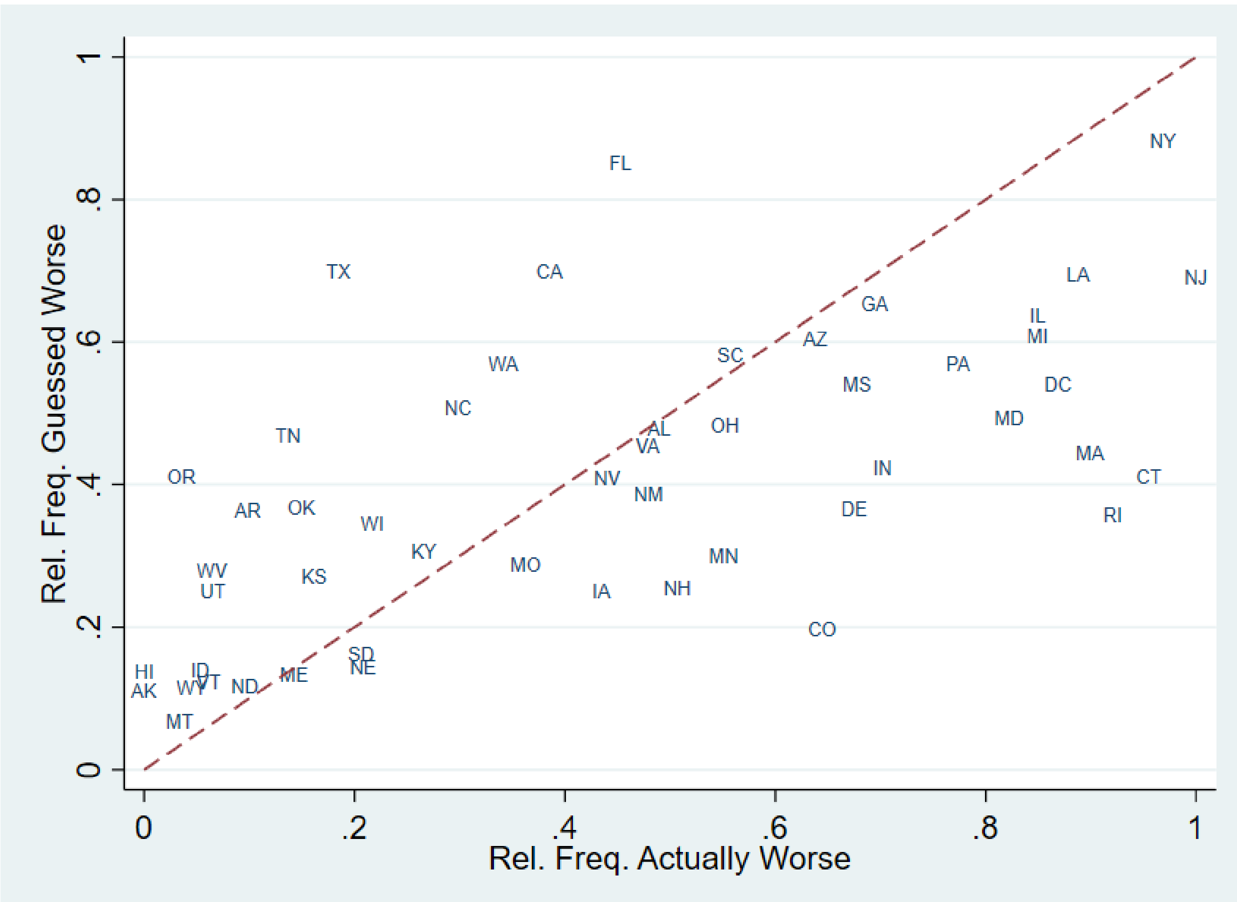

Figure 1 – Beliefs Versus Actual Relative State Performance

B Survey Details

Amazon Mechanical Turk is an online platform on which users can opt-in to completing various tasks in exchange for monetary compensation.

Our initial survey was conducted in two waves. The first wave occurred between July 22nd and August 10th, 2020 and involved approximately 400 respondents. Our sample was limited to "mTurk Masters," mTurk workers specifically designated by Amazon as top performers due to consistent high-quality answers. Respondents were also required to be US residents. Generally speaking, mTurk workers skew younger than the general population, but this is somewhat less true of mTurk Masters. We compensated respondents with a base rate of $1.50, topped up with an incentive bonus of up to $0.50 for accuracy.[10]

The second wave of our initial survey was conducted on October 14th and October 15th, 2020. Since this largely conducted for validation purposes, we recruited a smaller sample of approximately 200 respondents. For this survey, instead of restricting participation to U.S. mTurk Masters, we restricted to U.S. mTurk workers who had completed at least 500 tasks with a success rate of at least 99%.[11] Compensation was again $1.50, with an incentive bonus of up to $0.50. The questionnaire for our survey can be found in full in the Online Appendix.

Finally, our information-revelation experiment was conducted on December 21st and December 22nd, 2020 with a sample size of approximately 600 respondents. We again restricted our sample to U.S. mTurk workers who had completed at least 500 tasks with a success rate of at least 99%. We compensated respondents with a base rate of $0.80, topped up with an incentive bonus of up to $0.30. The questionnaire for our survey experiment can be found in the Online Appendix.

[10]This compensation was later increased – ultimately to a base rate of $2.50 and an incentive bonus of up to $0.75 – in order to attract additional respondents.

[11]We had exhausted the supply of U.S. mTurk Masters who were willing to take our survey at the compensation we offered.