Identifying and Measuring Conditional Policy Preferences: The Case of Opening Schools During a Pandemic

Jonathan Green

Postdoctoral Research Fellow, Network Science Institute, Northeastern University

Matthew Baum

Marvin Kalb Professor of Global Communications, Harvard Kennedy School

James Druckman

Payson S. Wild Professor of Political Science and IPR Fellow, Northwestern University

David Lazer

Distinguished Professor of Political Science and Computer and Information Science
Northeastern University

Katherine Ognyanova

Assistant Professor of Communication, Rutgers University

Matthew Simonson

Network Science Institute, Northeastern University

Roy Perlis

Professor of Psychiatry, Harvard Medical School

Mauricio Santillana

Assistant Professor of Pediatrics and Epidemiology, Harvard University

Version: December 13, 2020

DRAFT
Please do not quote or distribute without permission.

ABSTRACT

An individual's issue preferences are non-separable when they depend on other issue outcomes (Lacy 2001a), presenting measurement challenges for traditional survey research. The researchers extend this logic to the broader case of conditional preferences, in which policy preferences depend on the status of conditions with inherent levels of uncertainty—and are not necessarily policies themselves. They demonstrate new approaches for measuring conditional preferences in two large-scale survey experiments regarding the conditions under which citizens would support reopening schools in their communities during the COVID-19 pandemic. By drawing on recently developed methods at the intersection of machine learning and causal inference, the authors identify which citizens are most likely to have school reopening preferences that depend on additional considerations. The results highlight the advantages of using such approaches to measure conditional preferences, which represent an underappreciated and general phenomenon in public opinion

Click here for supplementary information for this paper.

Introduction

Policy complexity and uncertainty present challenges for public opinion research. Elected officials' policy proposals frequently contain multiple components, implicating multiple considerations. A politician may propose increasing spending on social welfare programs *on the condition* that beneficiaries provide evidence they are seeking employment. Or, they may propose increasing taxes to fund a *specific* public service. Asking citizens for their opinions on these packaged proposals often requires violating standard survey research best practices by asking respondents to summarize preferences on multiple issues in one response. However, isolating each element of a multi-faceted proposal in independent survey items may miss the essence of policy opinions.

This presents a critical problem when interlocking (i.e., conditional) policy proposals represent attempts at faithful delegate-based representation, reflecting perceived complexity in constituents' preferences that would entail relationships between considerations. Policymakers may in fact expect their constituents to only support expanding social welfare programs if individuals perceived to be undeserving are excluded from said benefits. Or they may expect them to oppose raising taxes without assurances that the revenue would not be wasted. Failure to reflect these relationships between considerations, when they exist, undermines surveys' abilities to communicate what constituents think about actual policies.

In such cases of *non-separable preferences* (Lacy 2001a), citizens' preferences regarding a policy depend on the status of at least one additional policy. When many citizens hold non-separable preferences on a given policy, standard survey items ask too much of respondents while saying too little. Such items force respondents to make unstated inferences regarding the status of additional policies relevant to their responses (Neblo 2015). For example, respondents may be asked to report a preference on a social welfare proposal without knowing its eligibility requirements. Whether they infer that certain eligibility requirements would be in place will affect their response, even if they have no way of communicating this inference to the researcher. As this issue does not arise from explicit violations of survey research best-practices, it is likely to go unnoticed.

Non-separable preferences can be generalized to a broader phenomenon, *conditional preferences*, which are familiar to economic theory (Skiadas 1997; Bicchieri 2010).[1] For our purposes, a respondent's preference on a given policy is conditional if it depends on additional considerations (e.g., support for new social welfare spending if and only if it does not increase inflation). A conditional preference then is non-separable if the consideration on which their preference depends is an additional policy (e.g., support for new social welfare spending if and only if taxes are increased).

Conditional preferences are likely to present particular measurement challenges for public opinion researchers when there is uncertainty regarding important additional considerations,

---

[1] Lacy (2001a) also notes the similarity between non-separable preferences and "dependent preferences" (Keeney and Raiffa 1993).

such as uncertainty regarding consequences of the policy itself. This sort of uncertainty is ubiquitous in public policymaking (Manski 2019). Indeed, debates regarding public policy are often debates over uncertain outcomes as much as they are debates over competing goals -- debates in which the relative merits of a proposal turn on whether it will (or will not) do the good (bad) things its proponents (opponents) claim it will. Public support for such proposals may depend on who is able to reduce uncertainty about these additional outcomes in which directions, persuading citizens that conditions relevant to their preferences have or have not been met. In short, public policy debates often invoke uncertain conditional outcomes, but survey researchers have few tools at their disposal to capture this.

Put another way, even though policy support regularly depends on additional considerations, researchers routinely ignore those dependencies in their measures. This is likely because, due to their complexity, such preferences are difficult to measure and recover in a systematic fashion. However, this should not mean that they are unavailable for public opinion research. Nor does it mean that researchers should be satisfied with unreliable, imprecise measurement. Instead, such issues call for measurement approaches aimed at recovering both *how many* and *which* citizens hold preferences that depend on additional considerations. We use recently-developed methods at the intersection of causal inference and machine learning (Athey and Imbens 2016; Wager and Athey 2018; Athey, Tibshirani, and Wager 2019) to develop such an approach, systematically identifying conditional preferences when and for whom they exist.

To illustrate our approach, we examine attitudes toward reopening schools during the COVID-19 pandemic, a case in which many citizens' preferences are likely to depend on additional considerations left unstated in traditional opinion measures. We find that most respondents hold some form of conditional preference regarding reopening schools. Specifically, most respondents are predicted to be more supportive of school reopening when given reason to believe that doing so could be done safely. We find that uncertainty regarding the safety of reopening can be reduced by credible source cues and changes in local case rates, while school-level safety measures have smaller effects. Further, we find that different respondents are sensitive to different conditions; in particular, respondents who support Joe Biden for president and those who are more conscientious about following public health recommendations tend to exhibit greater conditionality in their reopening preferences. Not only does this suggest a meaningful structure underlying attitudes regarding reopening policy, it highlights possibilities for representatives to both learn in more detail what their constituents want and to identify merits on which they may be persuaded. In our case, survey researchers regularly asked for citizens' preferences on various school reopening plans (Jones 2020; Horowitz 2020), and those survey results received widespread media attention. Yet, our findings suggest that these topline results did not reflect the complexity of many citizens' opinions, based on uncertainty regarding the safety of reopening.

Our results demonstrate the importance of taking conditional preferences seriously in public opinion research, and strategies for measuring them in a systematic manner. While we study one specific issue here, our analytical framework could be extended to other instances of conditional preferences, a common phenomenon in public opinion that has thus far received

insufficient attention. Ignoring conditionality in issue preferences there are theoretical reasons to expect it can provide a misleading account of both what the public wants and whether their preferences reflect meaningful structure. When respondents have multiple possible preferences on a given policy depending on the status of unstated conditions -- a dynamic that frequently occurs across a wide range of policy proposals -- what are in fact complex attitudes can present the appearance of non-attitudes (Lacy 2001a). As such, our results highlight the need for renewed interest in the study of conditional policy preferences, and demonstrate how recently developed methodological tools make such research easier to carry out.

Measuring Conditional Preferences

Lacy (2001a) draws on prior theories of non-separable preferences in formal models of political choice (e.g., Hinich and Munger 1997) to develop a theory of non-separable preferences for the survey response. Per Lacy, a survey respondent's preferences are non-separable if their position on Proposal A depends on the status of Issue B. For instance, nearly half of respondents in Lacy's survey indicated they would want higher (lower) taxes *if* anti-crime spending increased (decreased), demonstrating that their preferences on tax rates depended on the provision of particular public services. Some of the policies Lacy asked about exhibited high rates of dependency on other issues -- tax rates and spending levels were often dependent on what the tax revenue would be used for and what other money was being spent -- while other policies, such as abortion rights, were relatively separable from other issues.[2] Substantively, Lacy's findings challenged prior interpretations of question order effects as being attributable to non-attitudes among respondents low in political sophistication (Zaller 1992) -- instead showing that they were attributable to non-separable preferences, which were present among more than a quarter of respondents in seven of the ten issue pairs tested in his survey.

For our purposes, we consider non-separable preferences to be a specific case of the more general *conditional* preferences, in which preferences on Proposal A depend on Consideration B regardless of whether that consideration concerns an additional policy issue. Despite their relative prevalence in the mass public, conditional preferences are rarely studied as an outcome of interest in political science research. Indeed, since its publication, Lacy's theory of non-separable preferences has been cited just over 100 times,[3] typically in the context of voting behavior (Stoetzer and Zittlau 2015; Franchino and Zucchini 2015) or for its optimistic interpretation of framing and question order effects (Druckman 2004; Gaines, Kuklinski, and Quirk 2007) rather than for measuring non-separable issue preferences directly. Moreover, when such preferences are the direct subject of empirical inquiry, they are typically measured using single-shot survey questions (Tate 2003; Neblo et al. 2010), with the goal of estimating their prevalence in the aggregate. They are rarely, if ever, elicited in survey experiments as the

---

[2] For example, support for spending on the Aid to Families with Dependent Children (AFDC) program was higher conditional on abortion rights being restricted, but support for abortion rights did not appreciably change conditional on funding levels for AFDC.

[3] As of this writing, per Google Scholar. A companion piece in *Political Analysis* (Lacy 2001b) has been cited less than thirty times.

primary dependent variable; and they are rarely, if ever, the subject of common-content items on publicly-available surveys such as the ANES, CCES, or GSS.

This is likely *not* because conditional preferences are rare or inconsequential, but because they are difficult (and expensive) to measure. As Lacy acknowledges, it is exceedingly difficult to explore the full range of relevant conditions in conventional public opinion surveys. This requires asking respondents to report preferences on (or more complex still, rank) all possible alternatives with respect to the relevant conditions, inflating the time and attention required to complete the survey. However, Lacy offers a handful of practical suggestions for measuring conditional preferences. Relevant to our approach, when there are theoretical reasons to expect specific conditions on which attitudes depend, Lacy recommends making those relevant conditions explicit.[4] However, as he notes, time and resource constraints limit the extent to which researchers are able to ask respondents about the range of conditional outcomes.

We therefore consider it advantageous, resources permitting, to build on this recommendation using large-scale survey experiments, which Lacy did not explicitly recommend but have become much more practical in the intervening years (e.g. Mutz 2011). In an experimental context, researchers can manipulate both the presence of information about relevant conditions (allowing for comparisons to a control group) and its nature (allowing for comparisons between treatment conditions). The control group allows for comparisons to the results researchers would get absent any concern for conditional preferences (as is typically done). The varied conditions in the treatment groups -- with each respondent seeing only one version of the question -- allow for inferences regarding said conditionality without inflating survey length for any given respondent. To our knowledge this is the first such extension of Lacy's work to identify conditional preferences regarding public policy at the individual level in an experimental setting.

Eliciting Conditional Preferences Regarding School Reopening

The policy we consider here is reopening schools in the midst of the COVID-19 pandemic, for which we expect many respondents to hold conditional preferences. During the summer and fall of 2020 it was unclear how safe it would be to open schools months later, but government officials needed to make decisions well in advance of scheduled opening dates. There were compelling cases to be made both for and against opening schools: On the one hand, keeping schools closed likely carried long-term negative consequences for students missing valuable in-person instruction (Levinson, Cevik, and Lipsitch 2020), threatening to exacerbate existing inequalities in student outcomes while also being incredibly taxing for parents and slowing communities' abilities to resume other economic activities. At the same time, schools held the potential to be hotbeds of disease transmission (Guthrie et al. 2020). Without adequate resources and planning for in-person instruction during a pandemic, and without successfully

---

[4] Additional recommendations include providing respondents with the full questionnaire in advance of having them answer any items, to guard against question order effects, as well as allowing respondents to rank sets of possible alternatives.

reducing community spread of the disease, opening schools presented health risks for students, teachers, parents, administrators, and, by extension, entire communities.

Many citizens were thus likely to prefer that schools open, contingent on the condition that doing so would not exacerbate community disease transmission. Whether they thought this condition would be met would depend on whether they had been given a good reason to believe that opening schools would be done safely -- resolving their uncertainty regarding safety. However, this presents an additional layer of complexity: the vast majority of citizens could not know for themselves whether it was safe to open schools, and would instead be relying on trusted sources for relevant signals. Different citizens may be responsive to different messengers (Nelson and Garst 2005), or have different criteria they consider necessary to support reopening. Asking respondents for a simple thumbs up or down -- as typical surveys did on the topic -- without providing such information would force them to make inferences regarding these additional factors for themselves. By extension, in order to consider the resulting aggregate measures to be valid representations of the public's preferences, researchers would be forced to make the implicit, strong (and surely inaccurate) assumption that respondents resolved uncertainty regarding these conditions in essentially random fashion.

We investigate this complexity with two large-scale survey experiments conducted in July and October of 2020 (n = 19,057 and 16,128, respectively), with non-probability samples recruited through the survey vendor PureSpectrum (see further details in the Appendix). Prior research has found that experimental results in non-probability samples of this nature are generally similar to results that would be obtained using probability samples representative of the population (Mullinix et al. 2015)[5]. The first experiment asked respondents whether schools should reopen in the fall if a randomly assigned messenger indicated it was safe to do so, with a control group of respondents who were asked for their preference on reopening absent any signal. The second experiment asked respondents whether schools should be open under varying conditions regarding local prevalence of COVID-19 and schools' testing policies. This experiment included both a control group, which reported their preferences on reopening absent any information regarding the status of these conditions, and a "ceiling" condition that asked respondents to report their preferences regarding reopening if COVID-19 completely disappeared and it was definitely safe to do so. The latter contrasts with the former in that rather than leaving conditions relevant to the safety of reopening unstated, it renders them irrelevant.

Study 1

For the first experiment, we asked the control group, "Do you support reopening schools in your community for face-to-face classes this fall?" This item is based on standard approaches to measuring policy preferences, and is similar to items fielded by Gallup and Pew (Jones 2020; Horowitz 2020). We asked the treatment group, "If [messenger] said that it was safe for schools in your community to reopen for in-person classes this fall, would you support school reopening?" The response options were on a five-point scale ranging from strong opposition to

---

[5] Recent work also finds that experimental results obtained using such samples prior to the COVID-19 pandemic are robust to replication during the COVID-19 pandemic (Peyton, Huber, and Coppock 2020).

strong support, with a separate not sure option that is recoded to the middle of the scale. The messenger was randomized to be one of:
- The Centers for Disease Control
- The White House
- Donald Trump
- Your state's governor
- Your district's School Superintendent
- Leading scientists from the National Academy of Science.
- [if the respondent had school-age children] Your children's school principal

It is important to vary the source cues in this manner because, in the real world, citizens will hear that reopening schools is safe from specific institutions and individuals, and will rely on such cues to resolve their information uncertainty (Baum and Groeling 2009). These institutions and individuals will be differentially credible. While prior research finds that citizens tend to seek out and trust information from relevant institutions during times of crisis (Albertson and Gadarian 2015), emerging research indicates that partisanship weighs heavily on pandemic-related attitudes and behaviors (Allcott et al., forthcoming.; Bhanot and Hopkins 2020; Druckman et al. 2020; Gadarian, Goodman, and Pepinsky 2020; Gollwitzer et al. 2020).This means including messengers with varying degrees of political valence. It is also important to include state and local officials -- namely, governors, superintendents, and principals -- because the COVID-19 pandemic did not affect geographic areas equally or at the same time and local officials in those areas were responsible for significant portion of the policy response (Holman, Farris, and Sumner 2020). As such, citizens who are situated in communities with more severe exposure to the virus -- either in terms of cumulative case counts or recent case trends -- may be more sensitive to sources who are more attuned to local conditions than national officials or agencies (Malhotra and Kuo 2008).

Study 2

The second experiment is similar to the first, but differs in a handful of key respects. The control condition asked respondents a standard policy preference item: "Do you support or oppose having schools in your community open for in-person classes full-time?" The primary treatment conditions varied whether rapid testing was or was not mandatory, and whether more or fewer cases per day, worded as "If rapid testing for COVID-19 [was/was not] mandatory and there were [more/fewer] cases per day than there are now, would you support or oppose having schools in your community open for in-person classes full-time?" Finally, we included a "ceiling" version of the item, which asked "If COVID-19 ceased to exist and it was definitely safe to do so, would you support or oppose having schools in your community open for in-person classes full time?" This ceiling condition was designed to identify the maximum extent to which COVID-19 might affect attitudes regarding reopening by explicitly telling respondents to ignore it.

This second design addresses two potential limitations in Study 1. First, the ceiling condition provides a standard for conditionality to which responses in each treatment arm can be compared. Mandatory testing and lower case rates lower the risk posed by COVID-19

associated with reopening schools, but the hypothetical state of the world where both are the case is still riskier than the hypothetical state of the world where the virus has disappeared entirely. Second, by varying specific conditions in this design, rather than varying elite messengers sending the same cue, we are able to speak directly to the relevant conditions underpinning reopening preferences. As in, sources cues regarding safety in Study 1 encourage respondents to infer particular states of the world relevant to their preferences -- such as local case rates declining or school-level safety measures being put in place -- while Study 2 explicitly tells respondents what the statuses of those potential states of the world are.

Identifying Conditional Preferences with the Causal Random Forest

As discussed above, preferences regarding reopening schools during the COVID-19 pandemic have the potential to be highly complex. There are three key sources of such complexity. The first is basic conditionality. Some respondents would support reopening schools absent any additional information, but many would only support reopening schools if relevant conditions would be met. The second is which conditions are considered more important -- which source cues are considered more credible (Study 1) and which tangible conditions are considered more directly related to the safety of reopening (Study 2).

The final source of complexity is the potential for interactions between the treatments and respondent characteristics such as demographics, public health behaviors, and geographic context. Different citizens are likely to hold different levels of concern regarding COVID-19 based on their direct experiences with the virus at home and in their communities, as well as their personal and political identities, and will by extension be differentially responsive to attempts to reduce their level of uncertainty regarding the safety of school reopening. This leads us to consider a variety of covariates on which there may be heterogeneous treatment effects within and between experimental conditions based on respondent demographics, health behaviors, and geographic context (see Appendix for descriptions of all independent variables).

As we have only weak theoretical expectations regarding the effects these relationships between covariates should have on our outcome, we apply recently-developed machine learning methods to let such interactions emerge from the data itself -- as opposed to manually searching for them. Specifically, we analyze our survey experiments using the causal random forest (Athey and Imbens 2016; Wager and Athey 2018; Athey, Tibshirani, and Wager 2019; Ratkovic 2021). The causal random forest is an extension of the commonly-used random forest algorithm (Breiman 2001) that alters the target of the estimator, optimizing on difference in outcome between treatment conditions at each split as opposed to differences in the overall outcome. It is implemented using the open source grf package in R (Tibshirani et al. 2020).

The causal random forest is most useful when there are theoretical reasons to expect heterogeneous treatment, but the specific nature of these effects is not known in advance. In such cases, it can be tempting for researchers to manually specify a variety of models to explore the data, preserving the one that generates the most or most interesting statistically significant coefficients. Results generated using this approach are likely to be artifacts of model

specification, and are unlikely to replicate (Gelman and Loken 2013). The causal random forest avoids these pitfalls by applying principles from machine learning that are useful for generating good out-of-sample predictions (Cranmer and Desmarais 2017) to Rubin's (1974) potential outcomes framework for causal inference.

Specifically, it does so by randomly partitioning the data into a splitting subsample and an estimating subsample in each iteration of the algorithm. The splitting subsample is used to identify which splits, in which variables, in what order, maximize conditional average treatment effects in that subset of the data. However, the eventual estimation of conditional average treatment effects is determined by applying those splits to the held-out data. Put another way, trees in a causal forest satisfy what is referred to as the "honesty condition" when each observation is used to either identify splits or to estimate effects, but not both (Wager and Athey 2018).

As with other bootstrap-aggregated, or "bagged," learners, this process is repeated over a large number of iterations, with random partitioning of observations and random subsets of variables considered for splitting in each iteration. Individual level treatment effects are then estimated for each observation by passing them through the trees for which they were not used in the splitting subsample and averaging the predicted differences in outcomes. Wager and Athey (2018) show that these individual level effects are asymptotically normal, allowing for the construction of uncertainty intervals around the point estimates. Taken together, the process allows for the recovery of heterogeneous treatment effects that may not be expected *a priori* but are robust to out-of-sample prediction.

The causal random forest presents a handful of important advantages over prior applications of machine learning-based methods for detecting heterogeneous treatment effects in political science research (Imai and Strauss 2011; Green and Kern 2012; Imai and Ratkovic 2013; Grimmer, Messing, and Westwood 2017). Many of these methods were not designed for the express purpose of identifying heterogeneous treatment effects, and instead optimize on predicting the outcome generally, including the treatment condition as a covariate. As Ratkovic (2021) points out, in most prediction settings many confounding variables are likely to explain much more of the variation in the outcome than treatment assignment, and approaches that are not explicitly optimized for subgroup estimation will place more emphasis on these confounding relationships. While useful for predicting the outcome itself, these confounders may not be useful for recovering *differences* in the outcome conditional on treatment assignment (we elaborate on this in the Appendix). Conversely, previous machine learning methods developed for the purpose of detecting subgroup effects have not provided for the estimation of effects at the individual level. The ability to generate conditional average treatment effects with individual-level specificity, which can then be aggregated back up to examine group-level differences in response to different treatments, make the causal random forest ideal for addressing our research question.

Specifically, the causal random forest allows us to account for each of the three aforementioned layers of complexity regarding conditional preferences, in descending order of granularity. First,

it allows us to estimate baseline conditionality: the prevalence of conditional preferences regarding school reopening as measured by the average treatment effects between each treatment condition and the control condition. The larger these effects, the more respondents hold conditional preferences. Second, it allows us to compare conditions, identifying messengers are perceived to be more credible signals (in resolving uncertainty) that reopening is safe (Study 1) and tangible conditions respondents consider more important (Study 2). Finally, it allows us to recover within-condition differences in which respondents are more or less sensitive to the conditions made explicit. We examine each of these outcomes in the next section.

Results: Study 1

Overall Prevalence of Conditional Preferences

Figure 1 shows the overall distributions of responses within each experimental condition in Study 1. A plurality of respondents in the control condition oppose reopening schools in their community for in-person classes in the fall. The modal respondent strongly opposes the policy and just over 20% of respondents indicate any level of support. Each treatment condition has a distribution of responses differing from the control in some way. These differences typically take the form of more support and less opposition, indicating that respondents are more inclined to support reopening when the source says it is safe. The exceptions to this general trend are for the treatment conditions associated with the Trump administration -- the White House and Donald Trump himself -- which generate higher shares of strong opposition to reopening. However, between roughly 20 and 30% of respondents neither support nor oppose reopening across all conditions, indicating that no treatment is sufficient to eliminate uncertainty regarding reopening.

Figure 1: Topline Distributions (Study 1)

Distributions of School Reopening Responses by Experimental Condition

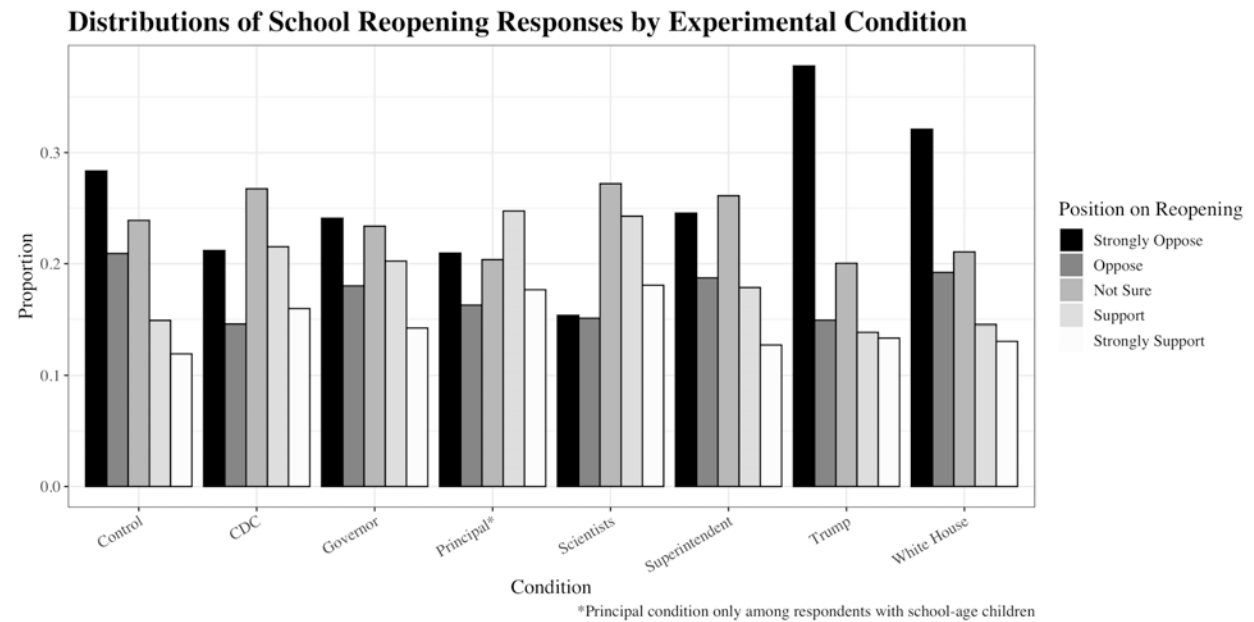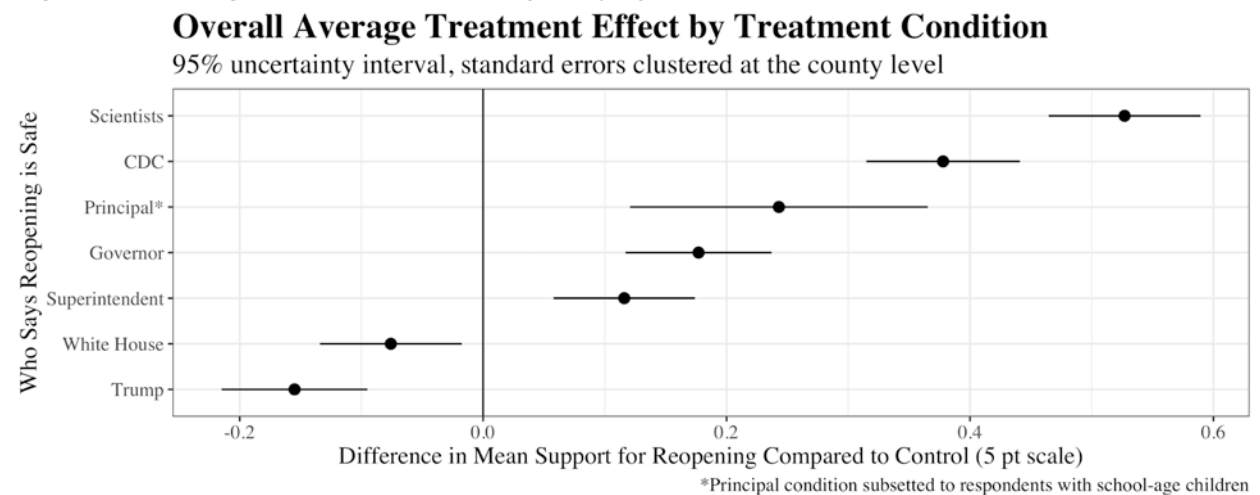*Principal condition only among respondents with school-age children

Figure 2 shows overall average treatment effects in each arm of the experiment, in comparison to the control group. In line with previous scholarship (Albertson and Gadarian 2015), we find that institutions with relevant expertise -- leading scientists from the National Academy of Science and the Centers for Disease Control -- have the strongest positive effects on support for reopening, with leading scientists increasing average support by over a half point on a five-point scale relative to control. By contrast, the White House and Donald Trump himself have *negative* and significant treatment effects, indicating that, on average, respondents become *less* supportive of reopening schools in the fall when the Trump administration is the source of the signal that reopening is safe -- consistent with prior findings of anti-cue effects (Goren, Federico, and Kittilson 2009; Nicholson 2012).

Figure 2: Average Treatment Effects (Study 1)

Overall Average Treatment Effect by Treatment Condition

95% uncertainty interval, standard errors clustered at the county level

*Principal condition subsetted to respondents with school-age children

These overall treatment effects establish the presence, but not prevalence, of conditional preferences in the aggregate. In order to estimate how many individuals hold conditional preferences, we turn to the individual-level treatment effects generated by the causal forest. Table 1 sorts the treatment conditions by the share of respondents whose individual-level treatment effects were statistically distinguishable from zero at the $p < .05$ level with cluster-robust standard errors. The table also shows the share of respondents with significant effects in each direction (Positive and Negative, with Null accounting for the remainder), as well as the share of respondents whose effects were significantly higher or lower than the *average* treatment effect (Above Average and Below Average). For example, roughly 9% of respondents are expected to be significantly more supportive of reopening schools in the Donald Trump condition relative to the control condition, compared to 46% of respondents for whom the expected effect is negative and significant (the effects for the remaining 45% are not statistically distinguishable from zero). This imbalance is partially attributable to ceiling effects among respondents who support Donald Trump's re-election, who are likelier than other respondents to support reopening schools in the control condition and therefore have less room for their support to increase (this is explored further in the Appendix). In addition, 28% of respondents in this condition have an expected treatment effect that is significantly above the overall average, while 32% have an expected effect significantly below the average, further illustrating the polarizing nature of this particular cue.

Table 1: Proportion of statistically significant treatment effect types by treatment condition (Study 1)

Comparison to Control	Positive	Negative	Null	Above Average	Below Average
Leading Scientists	0.979	0	0.021	0.156	0.154
Centers for Disease Control	0.852	0	0.148	0.093	0.101
Donald Trump	0.088	0.46	0.452	0.28	0.321
Your Children's School Principal*	0.534	0	0.466	0.001	0.012
Your State's Governor	0.444	0	0.556	0.034	0.034
The White House	0.004	0.198	0.798	0.045	0.055
You District's School Superintendent	0.187	0	0.813	0.005	0.003

*Among respondents with school-age children

A majority of respondents in a majority of conditions exhibit some form of conditional preferences, in that they hold a different view regarding reopening schools for in-person instruction in the fall given a signal that doing so would be safe compared to when this condition is left unstated. Moreover, even in most conditions with overall positive effects on average, there are non-negligible shares of respondents for whom the expected treatment effect is not

statistically distinguishable from zero, suggesting that different respondents place sufficiently distinct levels of trust in the source cue to produce heterogeneous effects. We also find that, in conditions featuring either Donald Trump or the scientific community, between 20 and 60 percent of respondents demonstrate predicted treatment effects that are themselves significantly above or below the average effect, further indicating treatment effect heterogeneity. These results highlight the importance of taking conditional preferences into account when there are theoretical reasons to expect them. Taken in isolation, the result in the control condition could be interpreted as overwhelming opposition to reopening schools in the fall. However, the treatment conditions show that a large share of citizens are willing to be convinced, and are more willing to be convinced by some messengers than others. Put another way, the standard survey measure would not be a validity indicator of individuals' preferences since their preferences are conditional on safety -- something about which there was in fact information available to reduce the uncertainty. Indeed, the information environment was saturated with elite messages at the time and citizens clearly conditioned their preferences on exposure to these messages.

Figure 3 visualizes these patterns, plotting individual-level treatment effects, relative to the control, in each condition. The y-axes sort respondents by their predicted treatment effects; the x-axes reflect the magnitude and direction of those effects, and have equal ranges to aid comparisons. Black dots represent point estimates and gray lines represent 95% uncertainty intervals with standard errors clustered at the county level. Average treatment effects across all respondents are shown with dashed vertical lines, with solid vertical lines at zero. So a condition with a relatively constant average treatment effect would show a relatively vertical line, with the vast majority of respondents predicted to have similar treatment effects. A condition with heterogeneous effects would show a steeper slope with more uncertainty intervals that do not touch the dashed horizontal line.

The figure shows that overall average treatment effects can mask marked heterogeneities. A roughly half-point effect on average in the Leading Scientists condition includes some respondents for whom the predicted effect is slight and some for whom it approaches a full point on a five-point scale. The slight negative effect in the Donald Trump condition includes some respondents who respond to the treatment positively and many who are predicted to support reopening at even lower rates than the overall average effect would suggest. Other conditions, such as the superintendent and principal conditions, show relatively constant effects.

Figure 3: Ranked individual-level treatment effects (Study 1)

Individual-level treatment effects in all conditions

Overall average treatment effects shown with dashed lines
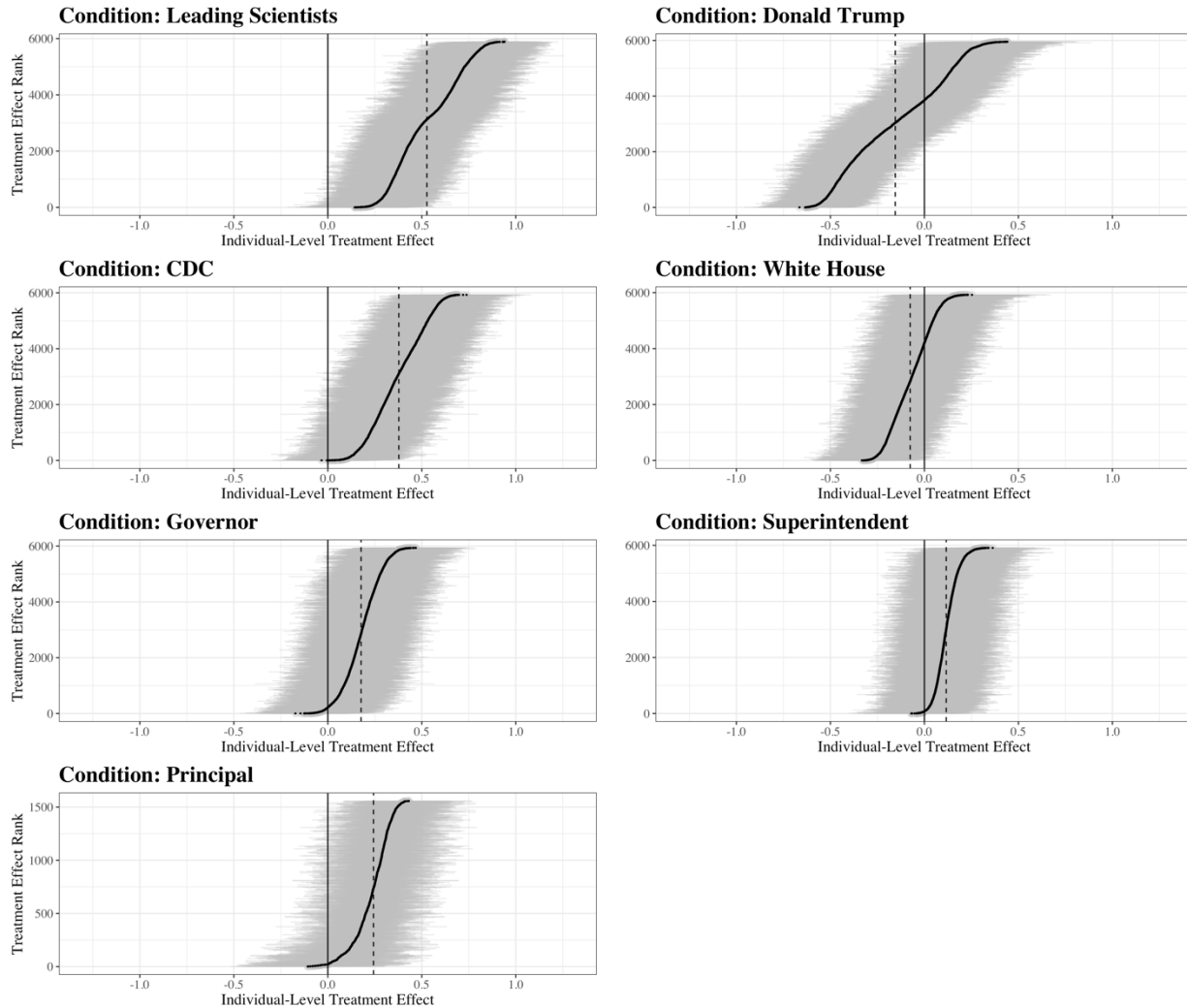
Cross-trained predictions for respondents in treatment condition or control; standard errors clustered by county

It is important to cautiously interpret the presence of negative treatment effects of the Trump administration (both the White House and Donald Trump specifically) signaling that reopening schools is safe, relative to there being no signal at all, on support for reopening. As indicated by the raw distributions in Figure 1 and elaborated on in the Appendix, this is primarily due to Democratic respondents (and Strong Democrats in particular) intensifying in what was already opposition to re-opening in the control condition, rather than respondents moving from support to opposition generally. In this context, such expression is suggestive of low levels of trust in the Trump administration among Democrats to handle the COVID-19 pandemic responsibly.
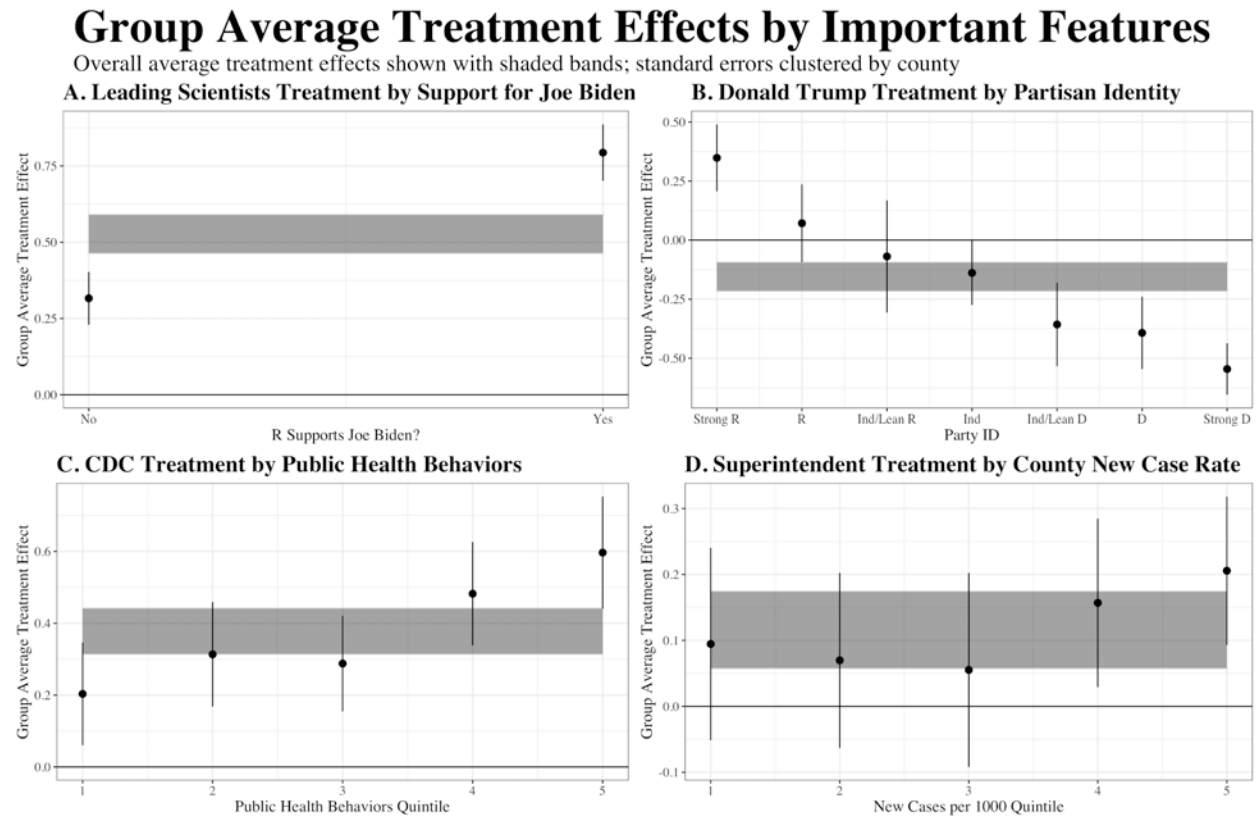
Who Holds Conditional Preferences?

While the above demonstrates the prevalence of conditional preferences regarding reopening schools in the fall of 2020, it does not speak to the question of which citizens hold which conditional preferences.

We begin to address this question in Figures 4a-4d, plotting group average treatment effects by changes in variables that emerged as most important for predicting treatment effects in the given experimental condition (we discuss variable importances across all conditions in more detail in the Appendix). We show this in the three conditions with the highest degree of heterogeneity and the one with the lowest, as measured by the share of respondents whose individual effects were above or below the overall average treatment effect. Specifically, we show group average treatment effects by support for Joe Biden in the Leading Scientists condition, partisan identity in the Donald Trump condition, public health behaviors in the CDC condition, and new cases per 1000 people in the respondent's county in the superintendent condition, respectively.

Sensitivity to cues from the scientific community varies by both political and public health-related characteristics, shown in Figures 4a and 4c. The overall positive average treatment effect in the Leading Scientists condition masks significant degrees of heterogeneity by respondents' political identities. While the average treatment effect in the Leading Scientists condition is positive among all subgroups, it is nevertheless polarized by political characteristics such that the effect among Joe Biden supporters is well above the average and the effect among respondents who do not affirmatively support Joe Biden is well below the average. However, we also find that respondents who report being more conscientious[6] in following public health recommendations report a greater degree of sensitivity to cues from the Centers for Disease Control that reopening schools is safe -- particularly at the high end of the behavior scale.

---

[6] Drawing on the method used in (Abramson et al. 2020), we bin this continuous scale of public health behaviors by quintile and calculate average treatment effects within each bin.

Figure 4: Group Average Treatment Effects by Important Features

Group Average Treatment Effects by Important Features
Overall average treatment effects shown with shaded bands; standard errors clustered by county

A. Leading Scientists Treatment by Support for Joe Biden

B. Donald Trump Treatment by Partisan Identity

C. CDC Treatment by Public Health Behaviors

**D. Superintendent Treatment by County New Case Rate**

Unlike the combination of political and public health factors that emerge as important in the experimental conditions featuring members of the scientific community, sensitivity to Donald Trump signaling that school reopening is safe is dominated by political factors. The most prominent of these factors is partisan identity, visualized in Figure 4b. Despite a negative overall average treatment effect, Strong Republicans are expected to be more supportive of reopening schools when the president signals it is safe compared to the control condition, while all non-Republicans (including independents) are expected to be less supportive and Strong Democrats especially so. Conversely, we do not find strong evidence of any particular factor driving treatment effect heterogeneity in the conditions that do not directly or indirectly implicate respondents' political identities -- those featuring local school administrators (district superintendents and, for respondents with school-age children, their principals). This is shown in Figure 4d, which shows the marginal effects of moving between respondents with increasingly acute new case trends (based on the seven-day moving average of new cases in their county). While there is weak evidence that respondents in counties where cases are increasing at higher rates per capita are more sensitive to signals from their districts' superintendents that reopening is safe, there do not appear to be any subgroups for whom treatment effects are meaningfully different from the overall, small but positive, average effect.

We further explore heterogeneities in treatment effects by plotting the relationships between two of the covariates that emerged as important features for prediction across multiple treatment conditions -- candidate support and public health behaviors -- and predicted treatment effects

relative to the control condition in Figure 5. As one might expect, candidate support is strongly associated with predicted treatment effects in the conditions mentioning Donald Trump and the White House, as well as the conditions mentioning the scientific community -- reflecting prior findings that scientific expertise is politicized in the United States (Gauchat 2012). In these conditions, predicted treatment effects do not vary to a great extent by public health behaviors -- save for slight increases in predicted effects at the very high end of the scale in the conditions featuring cues from the scientific community. However, in conditions that do not implicate national political dynamics to the same extent -- namely the principal and state governor conditions -- the extent to which respondents are predicted to increase their support for reopening schools varies more by the degree to which they report engaging in recommended public health behaviors, and less by which candidate they plan on supporting for president.

Figure 5: Relationships Between Politics and Public Health (Study 1)

Predicted Treatment Effects by Public Health Behaviors and Candidate Support
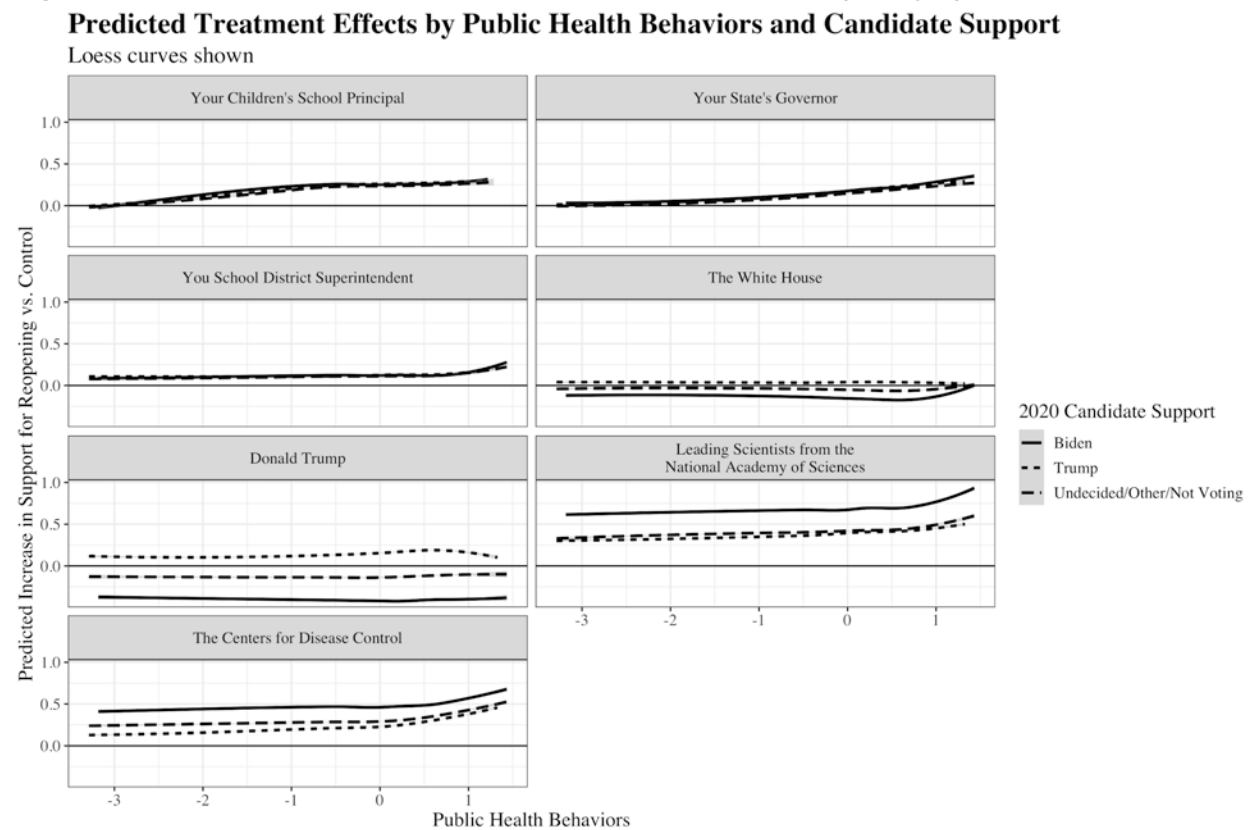
The degree to which the slopes and intercepts vary by condition in Figure 5 corresponds with overall levels of heterogeneity in treatment effects shown previously in Table 1. The Leading Scientists and Donald Trump conditions, for example, have sizable shares of respondents exhibiting individual level effects significantly different than the overall average, and see greater differences in predicted effects. The reverse holds for conditions such as Superintendent, in which nearly all individual-level effects are not statistically distinguishable from the average treatment effects and no variables emerge as being especially important for predicting treatment effects within that condition. This is especially apparent in Figure 5, where predicted effects are essentially constant across all levels of public health behaviors and candidate support.

Not only does this suggest that the ability to affect preferences regarding policies that involve uncertain conditions depends in part on who attempts to reduce that uncertainty, it also suggests that different audiences will be differentially receptive to different messengers. Moreover, it suggests that citizens' receptivities to different messengers exhibit a considerable degree of structure. People who are more conscientious about following public health recommendations change their preferences regarding school reopening when members of the scientific community say it is safe. Partisans exhibit polarized reactions to Donald Trump indicating that reopening schools is safe. These results are interesting precisely because they can be interpreted as unsurprising: what could otherwise appear to be cold opposition to reopening schools on the surface masks a considerable degree of structure and nuance in the public's attitudes regarding the issue.

Furthermore, that different covariates emerge as being differentially important in different treatment conditions underscores the difficulty of fully accounting for the complexities in conditional preferences using traditional techniques, and highlights the value the methodological approach we use here adds.[7] Overall, the substantive findings presented here regarding which citizens were most receptive to which messengers may not be surprising, but we could not have systematically produced them using traditional parametric approaches. Not only are the underlying mechanisms too complex to have theorized prior to analysis, attempting to recover them using ordinary least squares would almost certainly have produced overfit, unreliable, misleading results. Applying the tools developed for making good out-of-sample predictions broadens the scope of questions on which we can make systematic causal inferences.

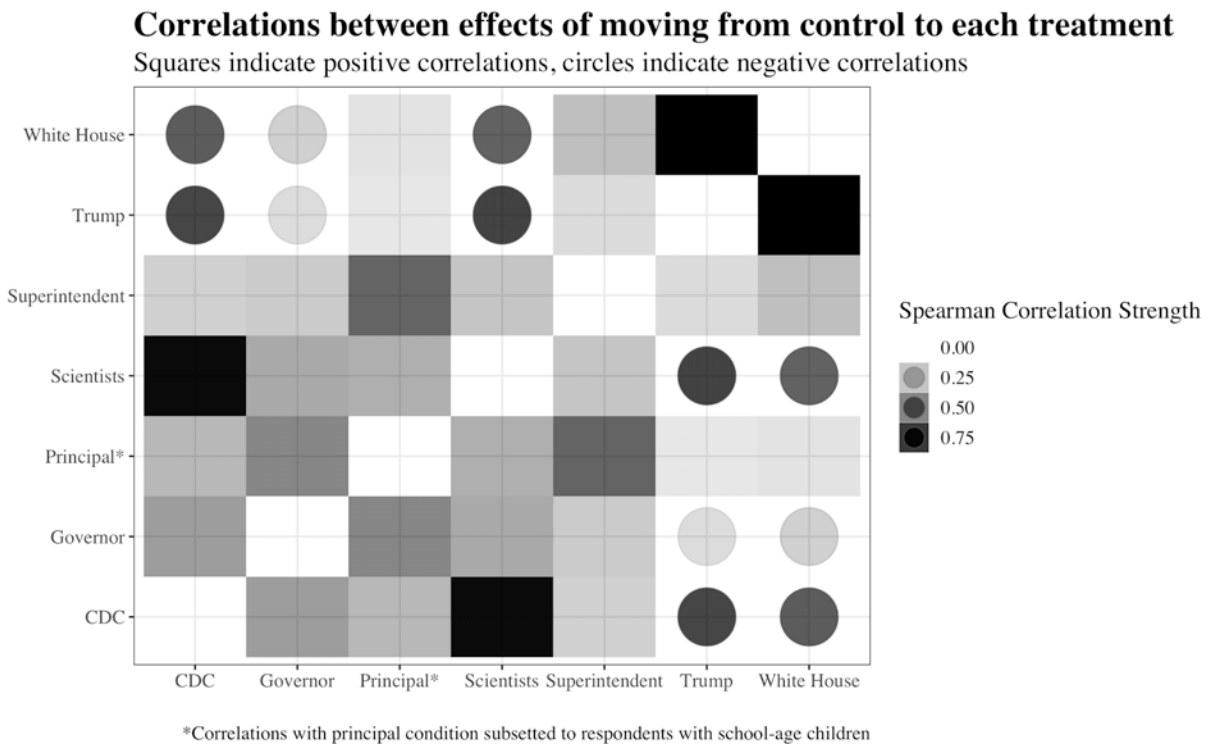*Relationships Between Conditional Preferences*

While the above illustrates the effects of individual treatments in comparison to the control group, it does not speak directly to how different treatments relate to one another. We address this question two ways: first, by examining correlations between treatment effects in comparison to the control group; and second, by changing the comparison to be between different treatment arms directly.

In order to estimate correlations between all treatment conditions and the control, we generate individual-level treatment effects for all respondents for every treatment condition to which respondents had equal propensity to be assigned, regardless of whether they were assigned to it. Which is to say, we do not generate out-of-sample predictions for the effect of moving from the control condition to the "your children's school principal" condition for respondents who do not have children. While we are not able to estimate cluster-robust standard errors for these individual-level out-of-sample predictions, this limitation does not affect correlations between their point estimates.

---

[7] Our findings also qualify recent work suggesting that treatment effects in survey experiments tend to be homogeneous (Coppock, Leeper, and Mullinix 2018), suggesting that conditional preferences present an underexplored area of survey research in which heterogeneous treatment effects may be more likely to manifest.

The correlations between treatment effects is visualized in Figure 6. Respondents are differentially sensitive to different messengers in comparison to the control condition, and these patterns generalize to different types of messengers rather than specific treatment conditions. Specifically, the more sensitive a respondent is to members of the scientific community saying that reopening is safe, the less sensitive that respondent is likely to be to the Trump administration saying so, and vice versa. Sensitivity to local officials such as governors and superintendents is positively, but not as strongly, correlated with sensitivity to the scientific community.

Figure 6: Correlations between treatment effects (Study 1)

Correlations between effects of moving from control to each treatment

Squares indicate positive correlations, circles indicate negative correlations

*Correlations with principal condition subsetted to respondents with school-age children

Taken together, these findings illustrate the significant degree to which respondents' preferences on reopening public schools depend on their perception that doing so would not exacerbate the COVID-19 pandemic, and which sorts of respondents are open to being persuaded by which sorts of messengers regarding the status of that relevant condition. Taken only in the aggregate, our findings would be consistent with pre-pandemic research indicating that citizens tend to place their trust in relevant institutions, such as the scientific community, during times of crisis (Albertson and Gadarian 2015). However, this general trust in relevant expertise is not consistent across respondents. The heterogeneities in sensitivity to the scientific community that emerge in our data reflect ongoing patterns in U.S. politics of partisan identities playing an important role in how public health information is sent and received (Baum 2011; Gadarian, Goodman, and Pepinsky 2020). While these results speak to the basic conditionality of reopening preferences, and which messengers different respondents are most sensitive to,

they do not speak to which tangible states of the world respondents consider relevant when evaluating whether reopening will be safe. We turn to this question with Study 2.

Results: Study 2

Figure 8 shows the overall distributions of responses in each experimental condition from Study 2. There is indeed a strong preference that schools be open for in-person classes full time in the hypothetical scenario where COVID-19 suddenly ceases to exist, compared to a plurality that remains opposed to schools being open under the status quo control condition with no additional information provided. We also observe differences in distributions between treatment conditions. It is apparent that support for having schools open is higher given lower case rates, and higher still when combined with mandatory testing, though neither of these conditions are fully sufficient to bring support up to the levels observed in the ceiling condition in which respondents were asked to consider the hypothetical scenario in which COVID-19 disappeared overnight. In short, these descriptive results clearly show conditionality in preferences regarding reopening schools.

**Figure 8: Topline Distributions (Study 2)**
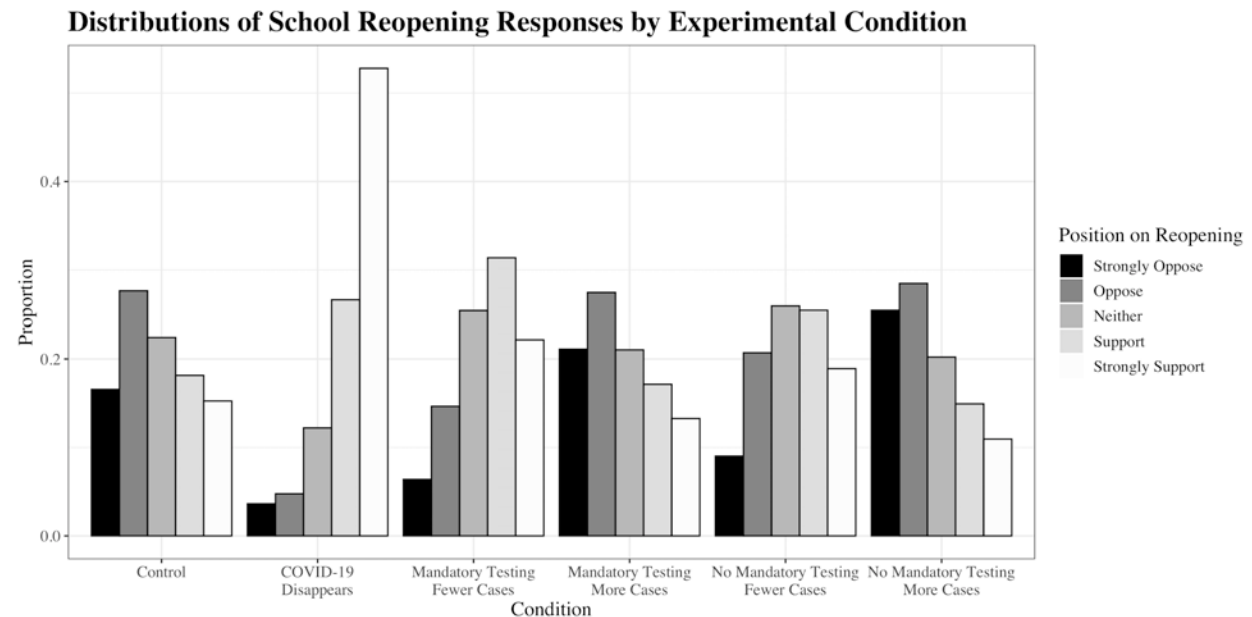
We turn to the overall average treatment effects from Study 2 in Figure 9, which reflect the differences in distributions shown in Figure 8. As the rank-ordering of the average treatment effects indicates, respondents tend to be more sensitive to case rates than to testing. The treatment effects are positive in the conditions where case rates are lower and negative when they are higher, with testing policies affecting the magnitude but not the sign of these effects.

Figure 9: Average Treatment Effects (Study 2)

Overall Average Treatment Effect by Treatment Condition
95% uncertainty interval, standard errors clustered at the county level
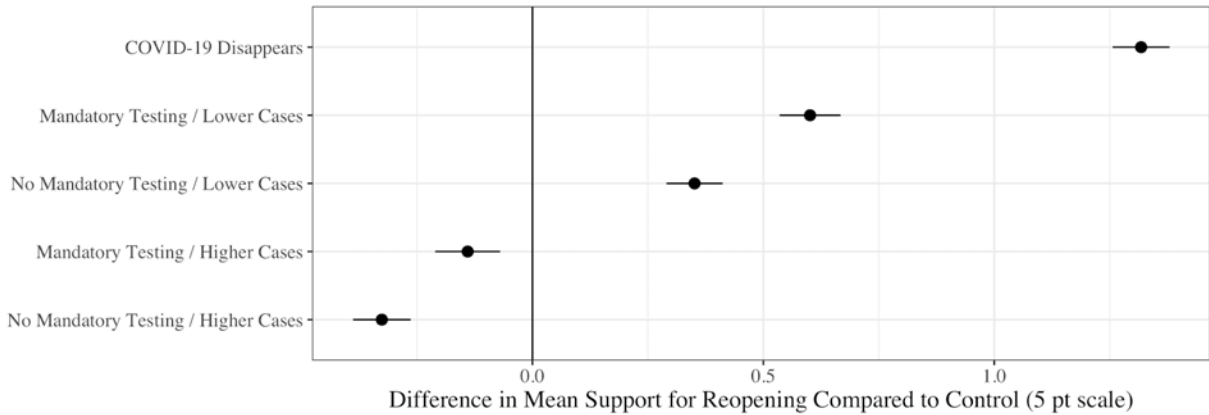
We next show the proportions of respondents with various types of predicted treatment effects in Table 2. Relative to Study 1, we observe less heterogeneity in most conditions and there are no conditions with polarizing effects. That is, within a given condition, all statistically significant predicted effects are either positive *or* negative. However, it is important to note that despite all respondents having positive and significant predicted treatment effects in the ceiling condition, where COVID-19 disappears and it is definitely safe, more than half of respondents have predicted effects that are statistically distinguishable from the overall average in either direction. This may be due to ceiling effects, or it may constitute evidence that preferences are less sensitive to conditionality, among respondents who would support reopening schools under the status quo.

Table 2: Proportion of statistically significant treatment effect types by treatment condition (Study 2)

| Comparison to Control	Positive	Negative	Null	Above Average	Below Average
COVID-19 Disappears	1	0	0	0.36	0.268
Mandatory Testing / Lower Cases	0.945	0	0.055	0.177	0.159
No Mandatory Testing / Lower Cases	0.933	0	0.067	0.014	0.049
No Mandatory Testing / Higher Cases	0	0.886	0.114	0.024	0.037
Mandatory Testing / Higher Cases	0	0.294	0.706	0.031	0.062

These trends are visualized in Figure 10, which plots predicted individual-level effects by treatment condition. As the figure shows, three of the treatment conditions (both where local case rates increase, and when case rates decrease without mandatory testing) show relatively homogeneous effects, while two (lower case rates with mandatory testing, and the ceiling condition where COVID-19 disappears) have notable levels of heterogeneity. This heterogeneity

is especially apparent in the ceiling condition, where predicted effects range from slightly less than half a point on a five point scale to nearly two points.

Figure 10: Ranked individual-level treatment effects (Study 2)

Predicted individual-level treatment effects in all conditions

Overall average treatment effects shown with dashed lines
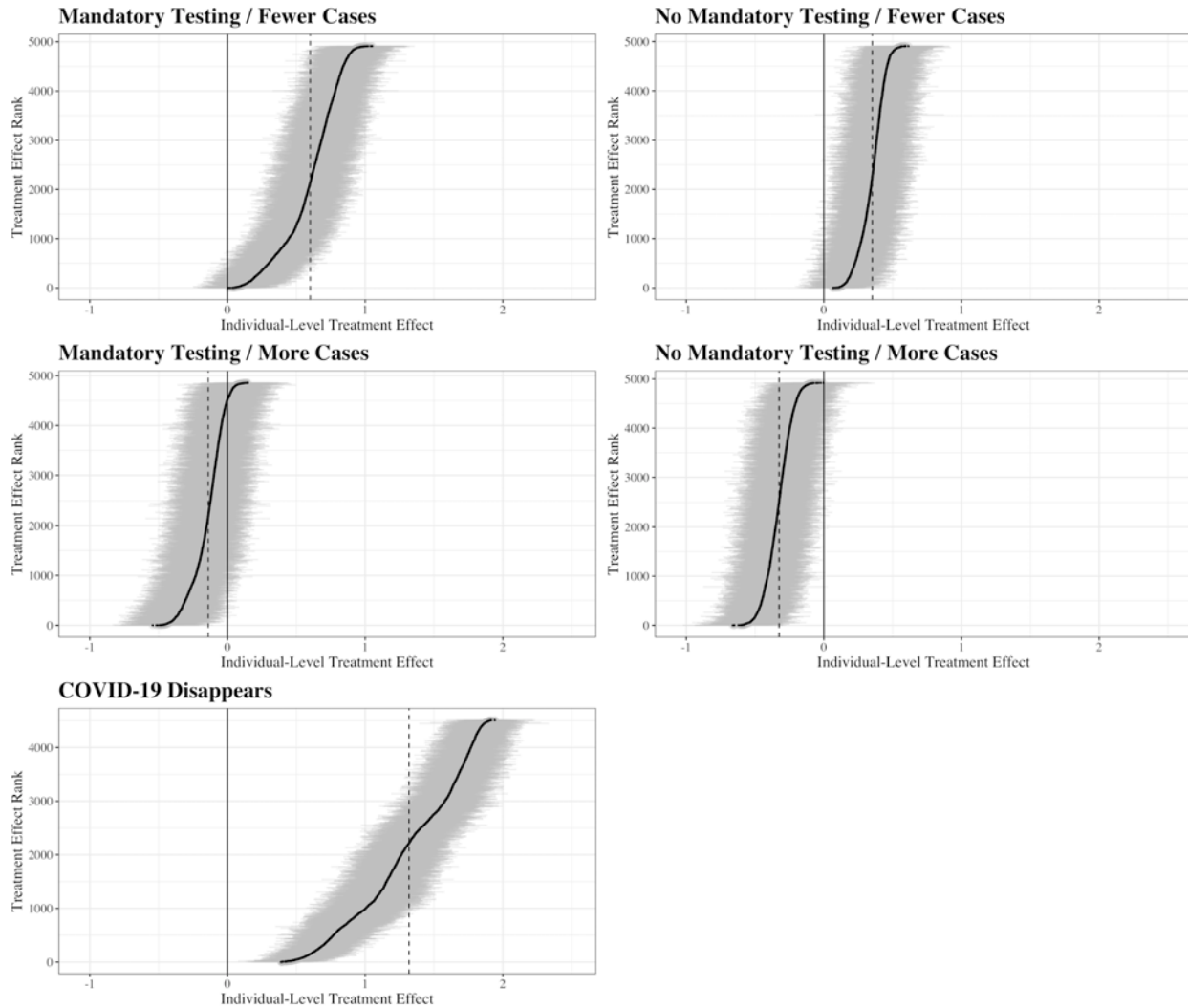
Cross-trained predictions for respondents in treatment condition or control; standard errors clustered by county

We next examine the specific interaction between political preferences and public health behaviors -- the two factors identified as most important for predicting heterogeneities in treatment effects (see Appendix for further details) -- by plotting the loess curves derived from observed characteristics and predicted treatment effects in Figure 11. There is a non-linear interaction between political preferences and public health behaviors that is particularly apparent in the two conditions with the greatest heterogeneity in effects: the mandatory testing / lower case rates condition and the ceiling condition where COVID-19 disappears completely.

While the positive effects in these conditions are larger in general for respondents who say they support Joe Biden than for respondents who are undecided and who say they support Trump, respectively, these differences become much smaller among those who report above-average adherence to public health recommendations regarding avoiding contact with other people, avoiding crowds, and wearing a mask when outside the home. One element of this finding is a ceiling effect -- Trump supporters who report lower adherence to public health guidelines are more likely to *already* support having public schools open. Hence, there is less room for their support to increase. Yet, this also highlights the conditionality of preferences regarding reopening policy. There are large numbers of respondents -- including Trump supporters -- who do not support having public schools open for in-person classes under the status quo but are open to the idea under the right conditions.

Figure 11: Relationships Between Politics and Public Health (Study 2)

Predicted Treatment Effects by Public Health Behaviors and Candidate Preferences
Loess curves shown
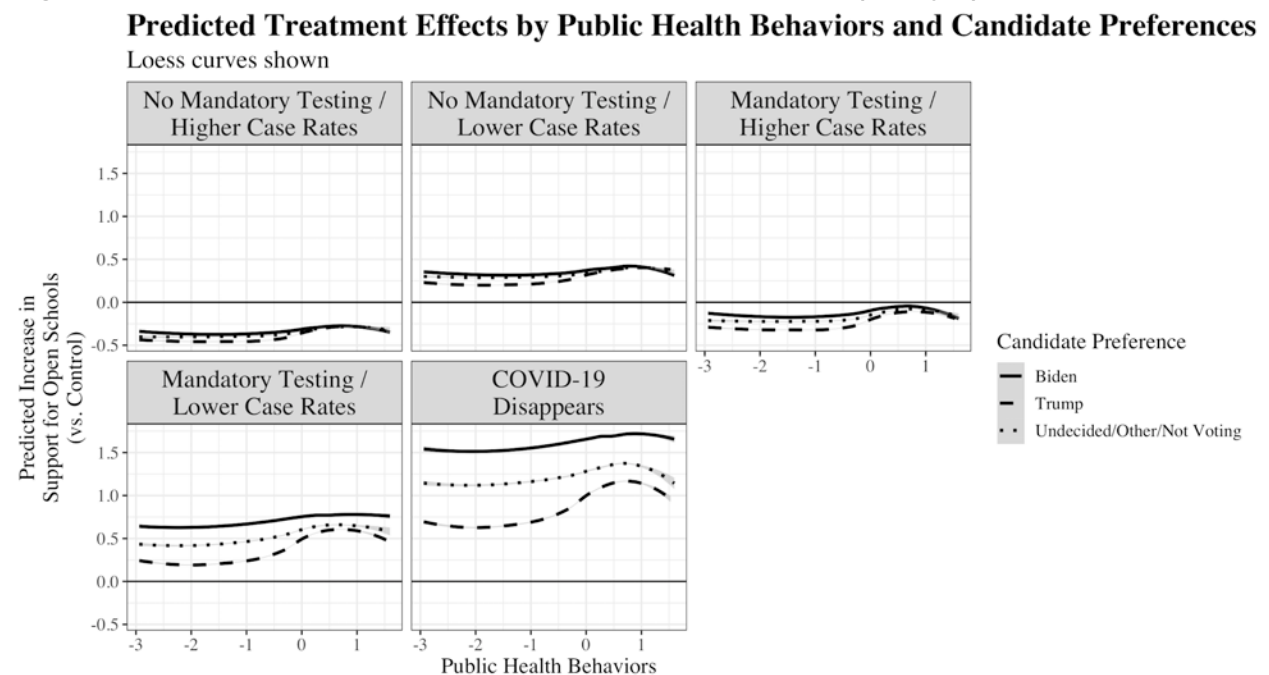
Discussion

Representative democracy is premised on the idea that public policy ought to bear some relationship to public opinion (Pitkin 1967; Mansbridge 2011). This makes the task of identifying what people think their political communities should do crucial for representation. However, complexity in policy design and uncertainty with respect to policies' potential consequences pose marked challenges for traditional public opinion research. Scholars often attribute measurement issues arising from these challenges to errors on the part of the respondents, who cannot be trusted to reliably report their preferences -- if they have preferences at all. It should come as no surprise, then, that policymakers frequently discount or ignore survey-based estimates of constituents' policy preferences when such information is made readily available to them (Herbst 1998; Kalla and Porter 2019). Building on prior work in survey theory and practice, we argue that measurement issues are likely to arise when survey respondents are forced to

grapple with complexity or uncertainty that makes policy preferences difficult to reduce to a single survey response. In such cases, the problem is not that citizens lack meaningful preferences to report to public opinion researchers and, by extension, policymakers. The problem is that standard measurement approaches do not reliably convey them.

This is apparent in the specific case examined here: opening schools during a pandemic. Preferences regarding this policy are difficult to measure from the perspective of traditional public opinion research for two related reasons. The first is that preferences on this policy proposal are highly conditional -- there are many citizens who want public schools to reopen on the condition that it would be done safely, but have not been persuaded that this is the case. This means that simply asking respondents whether they support reopening public schools forces those respondents to package multiple and potentially conflicting considerations into a single survey response. The second reason is that it is not enough to simply provide citizens with a cue that reopening schools is safe to estimate the extent to which they can be persuaded to support reopening. Outside of the survey context, that cue will come from some source who will reduce citizens' uncertainty regarding the policy at different rates -- based on both the degree to which citizens trust that source and the actual states of the world in their communities.

We apply recently developed methodological techniques to two large survey experiments that elicit preferences under a variety of states of relevant conditions, in comparison to control groups in which the statuses of all such conditions are left unstated. This allows us to show both the conditional nature of preferences regarding the relevant policy and differences in the degrees to which different citizens are sensitive to different conditions. Importantly, these methods further allow us to estimate individual-level treatment effects for respondents between any two experimental conditions. While this cannot fully resolve the fundamental problem of causal inference insofar as we cannot know *exactly* how each respondent would have reported their preferences if assigned to a different treatment condition (Rubin 1974), the principles from machine learning that researchers have previously used to generate robust out-of-sample predictions allow us to generate good estimates of a variety of potential outcomes for all of our respondents. Findings such as these can be used to inform representatives as they navigate the legitimately complex set of tradeoffs associated with policies such as reopening schools during pandemic conditions.

Our application of the causal random forest to large-scale survey experiments can be extended across a broad range of issue dimensions on which preferences are likely to contain similar levels of complexity, making what constituents want more intelligible for their representatives. While it does carry some limitations -- most notably, by requiring a larger survey in order to have the necessary power to estimate (potentially heterogeneous) effects while holding half the data out from estimation -- these are far outweighed by the increased breadth and depth it affords public opinion researchers interested in studying conditional preferences. Importantly, all analyses shown here rely on open-source, off-the-shelf statistical software that can easily be adapted to new data.

Indeed, we suspect that many if not most issue preferences are somewhat conditional insofar as policies are proposed with particular goals, which can be understood as conditions they aim to bring about. The degree to which this requires augmenting traditional survey methodology will depend on the degree to which preferences are conditional, the level of uncertainty associated with the relevant condition(s), and the degree to which citizens are likely to differ in their sensitivities to various attempts at reducing this uncertainty. In cases where there are reasons to suspect significant levels of conditionality, single-shot survey items or even traditional survey experiments are likely to be limited in ways that frustrate their use for informing policy action. However, as we show here, this is not a reason to avoid seeking the public's input on questions of complex, uncertain policy, assuming that they will not be capable of signaling their preferences to their elected representatives. Instead, it is a reason to invest in identifying the structure underlying their considerations. Failure to do so will further frustrate policy representation.

References

Abramson, Scott, Korhan Kocak, Asya Magazinnik, and Anton Strezhnev. 2020. "Improving Preference Elicitation in Conjoint Designs Using Machine Learning for Heterogeneous Effects." https://polmeth.theopenscholar.com/files/polmeth2020/files/polmeth_magazinnik.pdf.

Albertson, Bethany, and Shana Kushner Gadarian. 2015. *Anxious Politics: Democratic Citizenship in a Threatening World*. Cambridge University Press.

Allcott, Hunt, Levi Boxell, Jacob Conaway, Matthew Gentzkow, Michael Thaler, and David Yang. n.d. "Polarization and Public Health: Partisan Differences in Social Distancing During the Coronavirus Pandemic." *Journal of Public Economics*. http://web.stanford.edu/~gentzkow/research/social_distancing.pdf.

Athey, Susan, and Guido Imbens. 2016. "Recursive Partitioning for Heterogeneous Causal Effects." *Proceedings of the National Academy of Sciences* 113 (27): 7353–7360.

Athey, Susan, Julie Tibshirani, and Stefan Wager. 2019. "Generalized Random Forests." *Annals of Statistics* 47 (2): 1148–78.

Baum, Matthew. 2011. "Red State, Blue State, Flu State: Media Self-Selection and Partisan Gaps in Swine Flu Vaccinations." *Journal of Health Politics, Policy and Law* 36 (6): 1021–1059.

Baum, Matthew, and Tim Groeling. 2009. "Shot by the Messenger: Partisan Cues and Public Opinion Regarding National Security and War." *Political Behavior* 31 (2): 157–86.

Bhanot, Syon, and Daniel Hopkins. 2020. "Partisan Polarization and Resistance to Elite Messages: Results from Survey Experiments on Social Distancing." *Journal of Behavioral Public Administration*, July. https://cpb-us-w2.wpmucdn.com/web.sas.upenn.edu/dist/f/49/files/2020/07/Forthcoming-Partisan-Polarization-and-Resistance-to-Elite-Messages-JBPA.pdf.

Bicchieri, Cristina. 2010. "Norms, Preferences, and Conditional Behavior." *Politics, Philosophy, & Economics* 9 (3): 297–313.

Breiman, Leo. 2001. "Random Forests." *Machine Learning* 45: 5–32.

Coppock, Alexander, Thomas Leeper, and Kevin Mullinix. 2018. "Generalizability of Heterogeneous Treatment Effect Estimates across Samples." *Proceedings of the National Academy of Sciences* 115 (49): 12441–46.

Cranmer, Skyler, and Bruce Desmarais. 2017. "What Can We Learn from Predictive Modeling?" *Political Analysis* 25 (2): 145–66.

Druckman, James. 2004. "Political Preference Formation: Competition, Deliberation, and the (Ir)Relevance of Framing Effects." *American Political Science Review* 98 (4): 671–86.

Druckman, James, Samara Klar, Yanna Krupnikov, Matthew Levendusky, and John Barry Ryan. 2020. "Affective Polarization, Local Contexts and Public Opinion in America." *Nature Human Behavior*. https://doi.org/10.1038/s41562-020-01012-5.

Franchino, Fabio, and Francesco Zucchini. 2015. "Voting in Multi-Dimensional Space: A Conjoint Analysis Employing Valence and Ideology Attributes of Candidates." *Political Science Research and Methods* 3 (2): 221–41.

Gadarian, Shana Kushner, Sara Wallace Goodman, and Thomas Pepinsky. 2020. "Partisanship, Health Behavior, and Policy Attitudes in the Early Stages of the COVID-19 Pandemic." https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3562796.

Gaines, Brian, James Kuklinski, and Paul Quirk. 2007. "The Logic of the Survey Experiment Reexamined." *Political Analysis* 15 (1): 1–20.

Gauchat, Gordon. 2012. "Politicization of Science in the Public Sphere: A Study of Public Trust in the United States." *American Sociological Review* 77 (2): 167–87.

Gelman, Andrew, and Eric Loken. 2013. "The Garden of Forking Paths: Why Multiple

Comparisons Can Be a Problem, Even When There Is No 'fishing Expedition' or 'p-Hacking.'" http://www.stat.columbia.edu/~gelman/research/unpublished/p_hacking.pdf.

Gollwitzer, Anton, Cameron Martel, William Brady, Philip Pärnamets, Isaac Freedman, Eric Knowles, and Jay Van Bavel. 2020. "Partisan Differences in Physical Distancing Are Linked to Health Outcomes During the COVID-19 Pandemic." *Nature Human Behavior*.

Goren, Paul, Christopher Federico, and Miki Kittilson. 2009. "Source Cues, Partisan Identities, and Political Value Expression." *American Journal of Political Science* 53 (4): 805–20.

Green, Donald, and Holger Kern. 2012. "Modeling Heterogeneous Treatment Effects in Survey Experiments with Bayesian Additive Regression Trees." *Public Opinion Quarterly* 76 (3): 491–511.

Grimmer, Justin, Solomon Messing, and Sean Westwood. 2017. "Estimating Heterogeneous Treatment Effects and the Effects of Heterogeneous Treatments with Ensemble Methods." *Political Analysis* 25 (4): 413–34.

Guthrie, Brandon, Diana Tordoff, Julianne Meisner, Lorenzo Tolentino, Wenwen Jiang, Sherrilynne Fuller, Dylan Green, Diana Louden, and Jennifer Ross. 2020. "Summary of School Re-Opening Models and Implementation Approaches During the COVID 19 Pandemic." University of Washington Department of Global Health. https://globalhealth.washington.edu/sites/default/files/COVID-19%20Schools%20Summary%20%28updated%29.pdf.

Herbst, Susan. 1998. *Reading Public Opinion: How Political Actors View the Democratic Process*. University of Chicago Press.

Hinich, Melvin, and Michael Munger. 1997. *Analytical Politics*. Cambridge University Press.

Holman, Mirya, Emily Farris, and Jane Lawrence Sumner. 2020. "Local Political Institutions and First-Mover Policy Responses to COVID-19." *Journal of Political Institutions and Political Economy* 1 (4): 523–41.

Horowitz, Juliana Menasce. 2020. "Republicans, Democrats Differ over Factors K-12 Schools Should Consider in Deciding Whether to Reopen." *Fact Tank: News in the Numbers* (blog). August 5, 2020. https://www.pewresearch.org/fact-tank/2020/08/05/republicans-democrats-differ-over-factors-k-12-schools-should-consider-in-deciding-whether-to-reopen/.

Imai, Kosuke, and Marc Ratkovic. 2013. "Estimating Treatment Effect Heterogeneity in Randomized Program Evaluation." *Annals of Applied Statistics* 7 (1): 443–70.

Imai, Kosuke, and Aaron Strauss. 2011. "Estimation of Heterogeneous Treatment Effects from Randomized Experiments, with Application to the Optimal Planning of the Get-Out-the-Vote Campaign." *Political Analysis* 19: 1–19.

Jones, Jeffrey. 2020. "Fewer U.S. Parents Want Full-Time In-Person Fall Schooling." *Gallup* (blog). August 3, 2020. https://news.gallup.com/poll/316412/fewer-parents-full-time-person-fall-schooling.aspx.

Kalla, Joshua, and Ethan Porter. 2019. "Correcting Bias in Perceptions of Public Opinion Among American Elected Officials: Results from Two Field Experiments." https://osf.io/c2sp6/.

Keeney, Ralph, and Howard Raiffa. 1993. *Decisions with Multiple Objectives: Preferences and Value Tradeoffs*. Cambridge University Press.

Lacy, Dean. 2001a. "A Theory of Nonseparable Preferences in Survey Responses." *American Journal of Political Science* 45 (2): 239–58.

———. 2001b. "Nonseparable Preferences, Measurement Error, and Unstable Survey Responses." *Political Analysis* 9 (2): 95–115.

Levinson, Meira, Muge Cevik, and Marc Lipsitch. 2020. "Reopening Primary Schools during the Pandemic." *New England Journal of Medicine*. https://doi.org/10.1056/NEJMms2024920.

Malhotra, Neil, and Alexander Kuo. 2008. "Attributing Blame: The Public's Response to Hurricane Katrina." *Journal of Politics* 70 (1): 120–35.

Mansbridge, Jane. 2011. "Clarifying the Concept of Representation." *American Political Science*

Review 105 (3): 621–30.

Manski, Charles. 2019. "Communicating Uncertainty in Policy Analysis." *Proceedings of the National Academy of Sciences* 116 (16): 7634–41.

Mullinix, Kevin, Thomas Leeper, James Druckman, and Jeremy Freese. 2015. "The Generalizability of Survey Experiments." *Journal of Experimental Political Science* 2 (2): 109–38.

Mutz, Diana. 2011. *Population-Based Survey Experiments*. Princeton University Press.

Neblo, Michael. 2015. *Deliberative Democracy between Theory and Practice*. Cambridge University Press.

Neblo, Michael, Kevin Esterling, Ryan Kennedy, David Lazer, and Anand Sokhey. 2010. "Who Wants to Deliberate--And Why?" *American Political Science Review* 104 (3): 566–83.

Nelson, Thomas, and Jennifer Garst. 2005. "Values-Based Political Messages and Persuasion: Relationships among Speaker, Recipient, and Evoked Values." *Political Psychology* 26 (4): 489–516.

Nicholson, Stephen. 2012. "Polarizing Cues." *American Journal of Political Science* 56 (1): 52–66.

Peyton, Kyle, Gregory Huber, and Alexander Coppock. 2020. "The Generalizability of Online Experiments Conducted During the COVID-19 Pandemic." https://osf.io/preprints/socarxiv/s45yg/.

Pitkin, Hanna. 1967. *The Concept of Representation*. University of California Press.

Ratkovic, Marc. 2021. "Subgroup Analysis: Pitfalls, Promise, and Honesty." In *Advances in Experimental Political Science*. Cambridge University Press.

Rubin, Donald. 1974. "Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies." *Journal of Educational Psychology* 66 (5): 688–701.

Skiadas, Costis. 1997. "Conditioning and Aggregation of Preferences." *Econometrica* 65 (2): 347–67.

Stoetzer, Lukas, and Steffen Zittlau. 2015. "Multidimensional Spatial Voting with Non-Separable Preferences." *Political Analysis* 23 (3): 415–28.

Tate, Katherine. 2003. "Black Opinion on the Legitimacy of Racial Redistricting and Minority-Majority Districts." *American Political Science Review* 97 (1): 45–56.

Tibshirani, Julie, Susan Athey, Rina Friedberg, Vitor Hadad, David Hirshberg, Luke Miner, Erik Sverdrup, Stefan Wager, and Martin Wright. 2020. *Grf: Generalized Random Forests* (version 1.2.0). R. https://github.com/grf-labs/grf.

Wager, Stefan, and Susan Athey. 2018. "Estimation and Inference of Heterogeneous Treatment Effects Using Random Forests." *Journal of the American Statistical Association* 113 (523): 1228–42.

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