

Research Article

Evaluating Excuses: How the Public Judges Noncompliance

Amanda Driscoll*, Jay N. Krehbiel[†], Michael J. Nelson[‡], Taran Samarth^{**}

Abstract: Public officials often make policy but delegate its implementation. Yet, for reasons ranging from intransigence to incompetence, those tasked with implementation may not faithfully carry out policies. If implementors can frame noncompliance in a way that engenders sympathy, they may be able to disrupt the policy-making process with limited public backlash. We examine if the public's willingness to excuse noncompliance varies with the implementing actor's stated rationale for its failing to carry out the policy. Drawing on a survey experiment fielded in Germany, we find that the public is more sympathetic to resource-based, rather than principled, justifications for noncompliance, though the size of the effect is small. Further, contrary to fears that the pandemic would decay democratic functioning by leading citizens to be more forgiving of emergency-based inaction, we find no evidence that the public is more accepting of noncompliance justified on the base of the pandemic.

Keywords: Noncompliance, policy implementation, Covid-19, federalism

Supplements: Open data, [Open materials](#), [Preregistration](#)

The faithful implementation of policy poses a fundamental challenge for policymakers. While those tasked with implementation may not always like the policies they are asked to put into practice, it is widely assumed that overt noncompliance carries with it costs large enough to keep recalcitrant implementors in line (Carlin et al. 2022). However, for reasons ranging from intransigence to incompetence, implementors may not faithfully put policies into practice (Huber and McCarty 2004; Börzel et al. 2010; Fjølseth and Carrubba 2018). Studies typically assume that implementors face constant costs for noncompliance irrespective of the reason for not following through (e.g., Carrubba 2005; Vanberg 2001). Focusing on one potential cost of noncompliance—public reaction—we expect that the costs of noncompliance vary with the implementor's stated rationale for failing to execute the policy.

The potential for implementors to limit costs for noncompliance by tempering public constraints poses a normative concern for policymakers. If the public is swayed by implementors' excuses, clever public officials can undermine the democratic process by framing their noncompliance to engender public sympathy and limit fallout. But, if the public's reaction is stable regardless of the excuse, fear of a negative public reaction can provide a guardrail to foster faithful compliance.

Drawing on a survey experiment fielded on a nationally-representative sample of Germans at the height of the Covid-19 pandemic, we examine the public's reaction to a state government's failure to implement a hypothetical national education policy. We vary the rationale for noncompliance and evaluate whether the Covid-19-based justifications sway the public's reaction. We find that the public is more sympathetic to resource-based, rather than principled, justifications for noncompliance, though the effect size is small. Further, contrary to fears that citizens are more forgiving of pandemic-justified inaction, we find no evidence that the public is more accepting of noncompliance justified on this basis.

*Florida State University, [†]West Virginia University, [‡]Pennsylvania State University, ^{**}Yale University

Address correspondence to Nelson at mjn15@psu.edu.

Copyright: © 2023. The authors license this article under the terms of the Creative Commons Attribution 4.0 International License.

Implementation and Noncompliance

Governments formulate and pass policies with the intent of carrying out the letter of the law, but their agents sometimes fail to do so due to lack of capacity, statutory complexity, or unforeseen obstacles to perfect implementation (Mbaye 2001; Toshkov 2008). Research also documents that lack of faithful implementation might be strategically deployed to shirk responsibilities or forestall electoral retribution for unpopular policies (König and Mäder 2014; Fjelstul and Carrubba 2018). Most accounts that include some public costs to noncompliance assume they are large enough to be important and binding, but generally do not consider how the public's evaluation of noncompliance might vary in response to the rationale for nonimplementation (Vanberg 2001; Carrubba 2005). Although we know that, on balance, citizens do not like noncompliance (Carlin et al. 2022), there is little systematic evidence about how the public evaluates governmental lack of follow through, despite an abundance of compliance-related survey items on public opinion surveys (e.g., Levi, Sacks and Tyler 2009; Krönke 2018).

One possibility is that the rationales used by implementors to justify noncompliance change the public's response to their inaction. Indeed, particularly since citizens do not homogeneously follow politics closely, they are susceptible to accepting noncompliance and the rationale for doing so on its face, especially when citizens are largely unable to independently review the authenticity of justifications for noncompliance. Importantly, implementors—particularly elected officials—have incentives to leverage these dynamics and thereby reduce the penalties they face for noncompliance as they seek re-election and support to pursue other initiatives.

These rationales are essentially frames; by shaping the considerations that come to mind as individuals consider an issue, especially one—like governmental noncompliance—that has not occupied much of their thought, frames have the power to shape sentiment (Chong and Druckman 2007). The types of frames that can shape opinion are vast; perhaps most relevant to our study are those about a policy's stated rationale: Americans are famously more likely to approve of “assistance to the poor” than “welfare” (Smith 1987; but see Huber and Paris 2013), and Nelson and Driscoll (2023) demonstrate that people are more supportive of court packing efforts that purport to improve case processing rather than lend one party an ideological judicial advantage. Recent research (e.g., Meza and Zizumb-Colunga 2021) links noncompliance and narrative frames directly, suggesting that frames change the size of the accepted zone of noncompliance, thereby affecting the likelihood public officials engage in noncompliance. Our study, by contrast, is more in line with the literature on psychological framing effects, examining how these frames affect the public's *response* to noncompliance.

Types of Noncompliance

Research on policy implementation points to three primary motives for government non-compliance (Coombs 1980). Governments may not have the *resources* to comply—a state government may not be able to implement an unfunded mandate, for example, without federal support. Compliance may be precluded by *emergencies* that require officials to prioritize more urgent matters. Or governments may simply believe, on *principle*, a policy ought not to be implemented and choose to not comply even if they have the capacity to do so.

These justifications for noncompliance can reasonably mitigate political penalties from the public, especially when they suggest that noncompliance is not willful or intentional (regardless of the genuine cause for noncompliance). And these rationales can be deployed as frames to mask principled or policy-based noncompliance behind resource-based justifications that may be viewed more sympathetically by the public. The public may defer to governments that provide resource-based frames, since they cannot authenticate or comprehensively survey the government's available resources themselves. Rather, assuming citizen trust in government, we expect that the public will be more likely to accept policy noncompliance when justified due to resource constraints, compared to a policy-based frame. Thus:

H_1 : Noncompliant actors receive greater support from the public when using resource-based justifications for noncompliance, compared to policy-based justifications.

Pressing emergencies that divert governmental attention are likely to be viewed as similarly urgent by the public, particularly given that the public tends to extend greater approval and policy latitude to leaders during crisis and war (Davis and Silver 2004; Oneal and Bryan 1995; Huddy, Feldman and Weber 2007). Considering greater deference to governing authorities during crises, we expect that noncompliance is more likely to be tolerated by the public when justified as a result of constraints imposed by an ongoing crisis. Moreover, we expect that resource-based justifications are especially persuasive during a crisis that is legible to the public (for example, a pandemic or public safety threat), as the public has good reason to believe the government's claim is authentic. Therefore:

H₂: Noncompliant actors receive greater support from the public when invoking an ongoing crisis to justify noncompliance.

H₃: Noncompliant actors who invoke resource-based justifications for noncompliance receive greater support from the public during a crisis than outside of crisis.

Research Design

Our data come from a survey of 3,697 Germans fielded in partnership with YouGov from June 30 to July 13, 2020.¹ Germany presents an ideal case for our theory, as Germans' strong commitment to norms like the rule of law suggest they should be less susceptible to accepting excuses for noncompliance (Eurobarometer 2019). Moreover, the country's federal structure and accompanying division of policymaking across jurisdictions provides a natural context in which to evaluate our expectations.

Respondents read that the national government had announced a new policy in an area traditionally handled by the German *Länder*, public school curriculum. Respondents then learned that a state government refused to comply with the national curriculum policy. Here, respondents were randomly assigned into a 2x2 design: half of the respondents read that the nonimplementation was due to a lack of resources ("it does not have the resources to implement the new curriculum in their schools"), and the other half of the respondents read a policy-based explanation: that the state "knows best" about education. In the second treatment, half of the respondents were provided additional information that the state used the Covid-19 pandemic to justify their noncompliance.² We are cognizant that some respondents may have viewed this wording as off-putting, perhaps decreasing their support for the state government's proposal due to the brashness of the rationale. All respondents read the following text: "Suppose that a state announced that it would ignore the federal government's law mandating a new uniform curriculum in public schools. The state government stated that it would ignore the federal government's new law because..." The wording of the four treatments were as follows:

- **No Covid-Lacks Resources:** "it does not have the resources to implement the new curriculum in their schools."
- **Covid-Lacks Resources:** "it does not have the resources to implement the new curriculum in their schools in light of the crisis caused by the COVID-19 pandemic."
- **No Covid-Knows Best:** "it knows best how to educate its children."
- **Covid-Knows Best:** "it knows best how to educate its children, especially during the COVID-19 crisis."

Our outcome variable is the scaled response to four items answered by respondents after the vignette: support for the state government's response; belief that the state government's response was a legitimate exercise of power; support for the federal government to reduce state government funding in response; and support in an upcoming election for a political party that endorsed the state government's action. Full question wording is available in Appendix A. We scaled the resulting measure from 0-1 with higher values indicating more support for the noncompliance; we report in Appendix C the results using the individual indicator variables as outcome variables as well as multivariate models that include demographic and political control variables. Finally, balance tests (discussed in Appendix D) provide no evidence to suggest that the randomization was not successful.

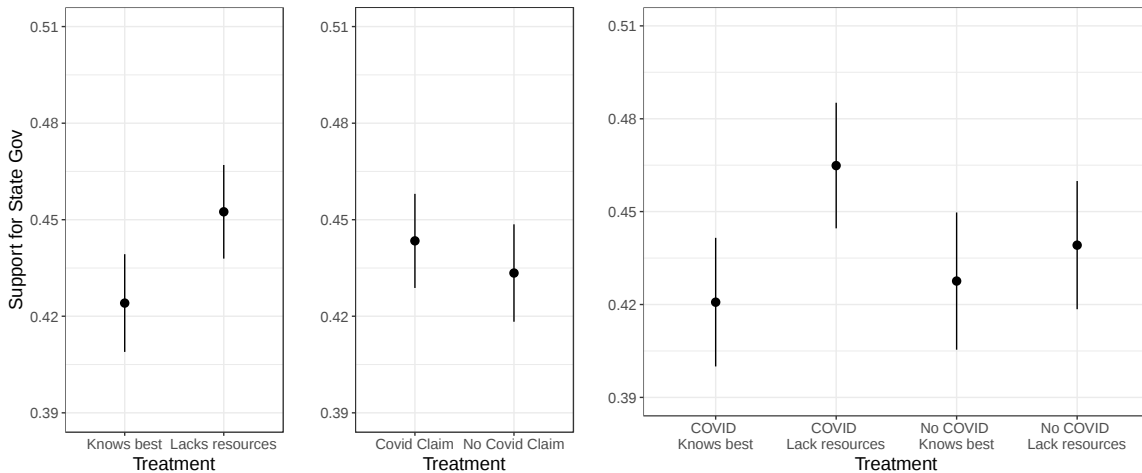
Results

Our first hypothesis was that respondents would be more sympathetic toward noncompliance justified based on resource limitations than principles. As shown in the left-hand panel of Figure 1, the data support this hypothesis. The average level of support for noncompliance for those assigned to the resource limitations treatment was 0.45; for those in the principled noncompliance group, the mean is 0.42. The difference in means is statistically significant: $p < 0.01$. Importantly, however, the substantive size is small: the difference in means corresponds to about a 10% change in a standard deviation of the outcome.

Our second hypothesis concerned the effect of the pandemic as a justification. We expected that respondents would be more tolerant of noncompliance justified based on the pandemic than excuses that did not cite the public health emergency. The results, shown in the middle panel of Figure 1, do not support our expectation: the average level of support for respondents assigned to read about the pandemic-based justification had an average level of support of 0.44; for those that did not read about an emergency-based justification, the average was 0.43. This difference in means is not statistically significant: $p = 0.43$.

Our final hypothesis concerned the conditional effect of an emergency justification. We find no evidence of a statistically significant interaction between the two treatments in a linear regression (see Table C4). However, as the right-hand panel of Figure 1 makes clear, this combination of treatments is associated with the highest level of support among the respondents. Respondents who read a justification that mentioned COVID were more supportive ($p < .01$) when the state claimed a lack of resources (versus knowing best), whereas respondents who did not read a COVID justification expressed no different levels of support ($p = .45$) when the state claimed a lack of resources or to know best.

Figure 1: Experimental Results



Notes: The dots plot the average value of Support for State Government and the whiskers provide 95% confidence intervals. Higher values of the y-axis indicate more support for the state government's noncompliance with the federal government.

Discussion

Can public officials frame their noncompliance with policy implementation to limit public backlash? Drawing on a survey experiment fielded during the Covid-19 pandemic in Germany, we assessed the effects of three types of justification—policy-based, resource-based, and emergency-based—on the public's response to noncompliance. Our results suggest that the public's reaction is relatively invariant. While we observe some statistically significant differences across our treatments, none of the effect sizes are large, and at no point do we observe a treatment associated with approval of the noncompliance. While resource-based justifications (especially those that also cite an emergency) are slightly more effective at evoking public sympathy, they are

not sufficient to achieve public approval of the government's failure to follow through on policy implementation.

The advantage of our survey experiment is that it exposed all respondents equally to noncompliance, limiting differences in citizen monitoring of noncompliance that might weaken these findings. Still, we are cognizant of several limitations of our analysis. We studied only one type of policy noncompliance in one country during a historic event, and we did so using a survey experiment which abstracts away from the heterogeneous attention to politics and access to information that plague real-world politics, perhaps inflating the size of treatment effects relative to real-world (Barabas and Jerit 2010) or giving rise to demand effects that inflate the size of our findings (but see Mummolo and Peterson 2019). And, the effect of both concerns would be to exaggerate the effects we observe relative to the real world; given the small size of our effects in the experimental setting, our results suggest that the ability of public officials to use frames to shield themselves from public backlash is minimal.

We close by highlighting two potential implications of our findings. First, the results suggest that delegation may not result in citizens giving a great deal of deference to those implementing the policy, even when those doing so claim to be better informed about what policy is most appropriate (e.g. Huber and McCarty 2004). One potential interpretation of this is that citizens, at least in the case of Germany, prioritize compliance with the proper policymaking process over asserting greater expertise or citing resource limitations. Second, the paper speaks to the literature exploring the contexts under which officials might be more or less willing to engage in noncompliance (e.g. Meza and Zizumbo-Colunga 2021). Whereas extant studies have emphasized factors such as the presence of NGOs and the need for bureaucrats to alter policies to improve their delivery (Braithwaite 2006, Gofen 2015), our study suggests that public reactions may be of less importance since citizens appear to censure noncompliance largely irrespective of its rationale. While this consistency on the public's part may be positive in the sense that it suggests citizens are not easily convinced of the propriety of acting outside the conventional legal policymaking structures, it at the same time reveals the potential for political constraints on implementors seeking to adapt policies to fit the circumstances.

Acknowledgment

This material is based upon work supported by the National Science Foundation under Grant Nos. SES-2027653, SES-2027671, and SES-2027664. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation. The experiment was cleared by the West Virginia University Institutional Review Board for compliance with Human Subjects protocol.

Notes

1. For technical details about the survey, see Appendix A.
2. We randomized whether the federal government cited the pandemic as its rationale for the policy. That treatment had no effect on reactions to the subnational government's noncompliance. See Appendix B.

References

- Barabas, J. and Jerit, J., 2010. "Are Survey Experiments Externally Valid?", *American Political Science Review*, 104(2): 226–242.
- Braithwaite, J. 2006. "Responsive Regulation and Developing Economies." *World Development* 34(5):884–98. doi: 10.1016/j.worlddev.2005.04.021.
- Borzel, T. A., Hofmann T., Panke, D. & Sprungk, C. 2010. "Obstinate and Inefficient: Why Member States Do Not Comply with European Law." *Comparative Political Studies* 43(11), 1363–1390.
- Carlin, R. E., Castrellón, M., Gauri, V., Jaramillo Sierra, I. C., & Staton, J. K. 2022. "Public Reactions to Non-compliance with Judicial Orders." *American Political Science Review* 116(1), 265–282.
- Carrubba, C. J. 2005. "Courts and Compliance in International Regulatory Regimes." *Journal of Politics* 67(3), 669–689.
- Chong, D. and Druckman, J.N., 2007. "Framing theory," *Annual Review of Political Science*, 10(2007): 103–126.
- Coombs, F. S. 1980. "The Bases of Noncompliance with a Policy." *Policy Studies Journal* 8(6), 885–892.

- Davis, D. W. & Silver, B. D. 2004. "Civil Liberties vs. Security: Public Opinion in the Context of the Terrorist Attacks on America." *American Journal of Political Science* 48(1), 28–46.
- Eurobarometer. 2019. *Special Eurobarometer 489 Report: Rule of Law*. European Commission, Directorate-General for Communication.
- Fjelstul, J. C. & Carrubba, C. J. 2018. "The Politics of International Oversight: Strategic Monitoring and Legal Compliance in the European Union." *American Political Science Review* 112(3), 429–445.
- Gofen, A. 2015. "Reconciling Policy Dissonance: patterns of Governmental Response to Policy Non-compliance." *Policy Sciences* 48(1):3–24. doi: 10.1007/s11077-014-9202-9.
- Huber, J. D. & McCarty, N. 2004. "Bureaucratic Capacity, Delegation, and Political Reform." *American Political Science Review* 98(3), 481–494.
- Huber, G.A. and Paris, C., 2013. "Assessing the programmatic equivalence assumption in question wording experiments: Understanding Why Americans like Assistance to the Poor More than Welfare," *Public Opinion Quarterly*, 77(1): 385-397.
- Huddy, L, Feldman, S. & Weber, C. 2007. "The Political Consequences of Perceived Threat and Felt Insecurity." *The Annals of the American Academy of Political and Social Science* 614, 131–153.
- König, T. & Mäder, L. 2014. "The Strategic Nature of Compliance: An Empirical Evaluation of Law Implementation in the Central Monitoring System of the European Union." *American Journal of Political Science* 58(1), 246–263.
- Krönke, M. 2018. *Bounded Autonomy: What limits Zimbabweans' Trust in their Courts and Electoral Commission*. Afrobarometer Policy Paper No. 52.
- Levi, M, Sacks, M. & Tyler, T. R. 2009. "Conceptualizing Legitimacy, Measuring Legitimizing Beliefs." *American Behavioral Scientist* 53(3), 354–375.
- Mbaye, H. A. D. 2001. "Why National States Comply with Supranational Law." *European Union Politics* 2(3), 259–281.
- Meza, O. & Zizumbo-Colunga, D. 2021. "Constructions of Noncompliance: Narratives and Contexts in the Case of Administrative Corruption." *International Public Management Journal* 24(5), 623-645.
- Mummolo, J. and Peterson, E., 2019. "Demand Effects in Survey Experiments: An Empirical Assessment," *American Political Science Review*, 113(2):517-529.
- Nelson, M.J. and Driscoll, A. 2023. Accountability for Court Packing," *Journal of Law and Courts*, 1-22.
- Oneal, J. R. & Bryan, A. L. 1995. "The rally 'round the flag effect in U.S. foreign policy crises, 1950–1985." *Political Behavior* 17(4), 379–401.
- Smith, T.W., 1987. "That which we call welfare by any other name would smell sweeter an analysis of the impact of question wording on response patterns," *Public Opinion Quarterly*, 51(1): 75-83.
- Toshkov, D. 2008. "Embracing European Law: Compliance with EU Directives in Central and Eastern Europe." *European Union Politics* 9(3), 379–402.
- Vanberg, G. 2001. "Legislative-Judicial Relations: A Game-Theoretic Approach to Constitutional Review." *American Journal of Political Science* 45, 346-361.

Appendix

Appendix A. The YouGov Survey

Sampling Details

YouGov interviewed 4,729 respondents in Wave 1 who were then matched down to a sample of 4,400 to produce the final Wave 1 dataset. The respondents were matched to a sampling frame on gender, age, and education. This frame was constructed by stratified sampling from the 2018 Eurobarometer with selection within strata by weighted sampling (using the person weights on the public use file). The matched cases were weighted to the sampling frame using propensity scores. The matched cases and the frame were combined and a logistic regression was estimated for inclusion in the frame. The propensity score function included age, gender, years of education, and state. The propensity scores were grouped into deciles of the estimated propensity score in the frame and post-stratified according to these deciles. The weights were then post-stratified on 2017 General Election vote choice, and a stratification of gender, state, age (4-categories), and education (4-categories) to produce the final weight. In Wave 2, YouGov re-contacted all 4,400 wave 1 respondents and completed 3,697 completed wave 2 interviews. In Wave 2, YouGov re-contacted the Wave 2 respondents and completed 3,198 Wave 3 interviews. YouGov prepared a wave 3 weight following the same procedures as Wave 1.

Questionnaire Details

The purpose of the survey was to understand respondents' support for the rule of law in light of the Covid-19 pandemic. The survey was approximately 12-15 minutes in length. The survey opened by asking respondents about their experience with the Covid-19 pandemic, including their concern about contracting the virus, their support for masking policies, their attitudes toward vaccination, and their satisfaction with the government's response to the pandemic. Then, respondents were asked about their support for democracy, the rule of law, and the German judicial system. The experiment came at the end of the surveys so as to ensure that responses to the earlier items were not somehow contaminated by the experimental treatment.

Outcome Variable Details

After reading about the state government's noncompliance, we asked the respondents four questions:

- Would you support or oppose the state government's refusal to implement the federal law? (38% supported the state government at least "somewhat")
- Do you believe this action by the state government is a legitimate exercise of power? (52% believed this was at least "somewhat" a legitimate exercise of power)
- To what extent would you support a decision by the Federal government to reduce funding for major infrastructure projects in this state because of the state government's refusal to implement the uniform national curriculum law? (50% would at least "somewhat" support)
- If the next state elections were held next Sunday, how likely would you be to support a party that supported the state government's refusal to implement the uniform national curriculum law? (41% would be at least "somewhat likely")

The items load onto a single dimension with $\alpha = 0.61$, though the third item—which was reverse coded—loads poorly with a loading of only -0.11. We show in Appendix C that we reach the same conclusions when the outcome variable excludes the reverse-coded item.

Appendix B. First stage of experiment

All respondents read about a hypothetical new German law that would claw back control over local education to the national government. Half of the respondents read that this action was justified by the federal government as part of an exercise of emergency powers due to the Covid-19 pandemic. Note that respondents were assigned to each of these conditions such that half received the Covid-19 justification in the first stage of the experiment. In other words, we randomized the respondents into eight conditions: the 2x2 implementation experiment for those respondents who did (and did not) read that the federal government justified its initial action due to the pandemic.

The first portion of the vignette read:

Some policies in Germany are conducted by the Federal government while others are conducted by state governments. As you may know, most decisions about education policy are made by state governments. Imagine that the Federal government passed a new law establishing a uniform national curriculum for students. The law significantly limits the ability of state officials to determine what students learn. The Federal government justified the measure as [an emergency exercise of its authority] necessary to guarantee a quality education for all students [in light of the Coronavirus (COVID-19) crisis].

We then asked respondents four questions:

- Would you support or oppose the Federal government's action? (78% at least "some- what support")
- Do you believe this action by the federal government is a legitimate exercise of power? (76% at least "somewhat legitimate")
- Do you believe that the state governments should have to follow this policy? (79% at least "somewhat" believe it should have to be followed)
- If the next federal election were held next Sunday, how likely would you be to support a party that supported the creation of a uniform national curriculum? (74% would be at least "somewhat likely")

These four items form a reliable scale ($\alpha = 0.89$) that loads onto a single dimension through factor analysis. Policy support loads at .84; perceived legitimacy and support for state compliance at .87; and party support at .67. We use as our outcome variable the scores from the factor analysis, rescaled from 0-1 such that higher values indicate more support for the federal government's action. For reference, both the four- and three-item scales from the second stage of the experiment (state noncompliance) are negatively correlated ($r = -.38$) with the federal support scale from the first stage (federal government's implementation of the curriculum policy).

First Stage Experiment Results

Respondents express more support ($p < .01$) for the national curriculum policy when the federal government relies strictly on the desire for "quality education," versus including a reference to the COVID-19 pandemic. The treatment effect is 11% of a standard deviation in the scaled federal support outcome variable. Respondents view the policy as more legitimate ($p < .01$) and expect more state compliance ($p < .01$) when the federal government uses the "quality education" justification without a reference to the pandemic than with, but respondents express no greater intent to vote for a party proposing a national curriculum ($p = .80$).

Accounting for First Stage Treatment Assignment in Second Stage Results

While the results from the first stage of the experiment are interesting in their own right, readers may be concerned about whether the results we present in the paper are contingent upon first-stage treatment assignment. We report the results of a linear regression controlling for the first-stage treatment in Table B1 and replicate Figure 1 from the main results with the predicted values from this regression in Figure B1. We find that the results are unaffected by the first stage treatment.

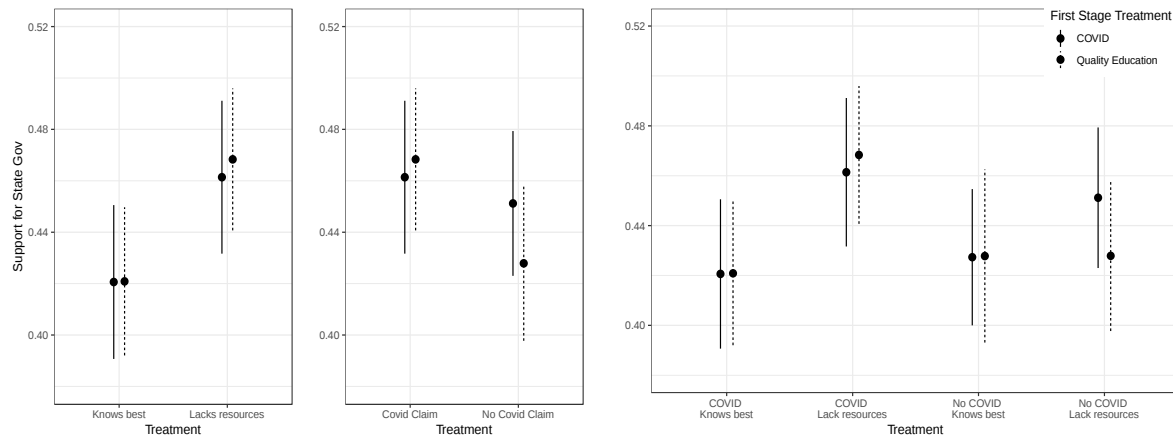


Figure B1: The dots represent the predicted values of Support for State Government, using the regression from Table B1 and the whiskers provide 95% confidence intervals. Higher values of the y-axis indicate more support for the state government's noncompliance with the federal government.

Table B1: Second Stage Results Accounting for First Stage Treatment

	Support for State Gov.
State: No COVID Claim	0.007 (0.021)
State: Lack resources	0.041* (0.022)
Federal: Quality education	0.0002 (0.021)
No COVID × Lack resources	−0.017 (0.029)
No COVID × Quality education	0.0002 (0.031)
Lack resources × Quality education	0.007 (0.030)
No Covid × Lack Resources × Quality education	−0.030 (0.043)
Constant	0.421*** (0.015)
N	3,677
Log Likelihood	−708.457
AIC	1,432.914
*p < .05; **p < .01; ***p < .001	

Appendix C. Additional Results

This appendix presents additional regression results to supplement the results presented in the main paper. Table C1 presents the baseline regression results, using both the four-item outcome variable discussed in the paper and the three-item version described in Appendix B. The results are consistent across outcome variables.

Readers may be interested in how our results vary across the individual items that comprise our outcome variable. Respondents who read a justification about the state lacking resources expressed more support for noncompliance ($p < .01$) than those who read that the state knew best, but respondents reported no different levels of legitimacy ($p = .051$), support for federal punishment ($p = .18$), or electoral support for anti-compliance parties ($p = .54$). Respondents who read assertions that COVID influenced the decision to not comply did not report different levels in any outcome variable compared to those who did not read a COVID-related noncompliance justification. In the 2x2 analysis, among respondents who read a COVID noncompliance justification, those who read that the state lacked resources expressed higher support for noncompliance ($p < .01$) and perceived legitimacy of noncompliance ($p = .04$) than those who read that the state knew best. Levels of support for federal punishment for noncompliance and electoral support for anti-compliance parties are not significantly different across the four cells. In other words, it appears that noncompliance support and noncompliance legitimacy drive differences in the scaled support outcome, while support for a federal response and support for anti-compliance parties are unaffected. Models mirroring the specifications from Table C1 on the individual items as outcome variables are provided in Table C2.

Though, as we show in Appendix D, we have no reason to suspect that randomization was not successful, readers may be concerned that some confounder affected the results presented in the paper. For robustness against key demographic and political variables, we regress the scaled four-item state support outcome on demographics, COVID concern, rule of law, and partisanship and report the results in Table C3. We also re-estimate the model with the 3-item scale, dropping the reverse-coded item about support for the federal government withdrawing funding for the state since it loads poorly onto the scale. Our results from the bivariate analysis are robust to the inclusion of covariates and the smaller composite scale for state government support: the “lack resources” treatment effect holds across both models, while we still do not identify an effect from using a COVID justification. Unsurprisingly, approval of the state government is positively associated with scaled state support, while approval of the federal government is negatively associated. Rule of law is also negatively associated with support for the state.

In Table C4, we provide results using the individual items as outcome variables but include control variables. We find that the “lack resources” justification is only associated with greater support for noncompliance, but has no effect on the legitimacy, punishment, or electoral outcomes. Consistent with our other results, we fail to identify an effect from COVID-related justifications on any outcome.

Table C1: Composite Support Regression, State Stage (No Control Variables)

	Support for State Gov. (4-item)		Support for State Gov. (3-item)	
	(1)	(2)	(3)	(4)
No COVID Claim	-0.010 (0.011)	0.007 (0.015)	-0.010 (0.011)	0.006 (0.015)
Lack resources	0.028** (0.011)	0.044** (0.015)	0.028** (0.011)	0.043** (0.015)
No COVID X Lack resources		-0.033 (0.021)		-0.031 (0.021)
Constant	0.429*** (0.009)	0.421*** (0.011)	0.430*** (0.009)	0.422*** (0.010)
N	3,677	3,677	3,680	3,680
Log Likelihood	-711.100	-709.400	-679.207	-677.596
AIC	1,428.200	1,426.800	1,364.414	1,363.193

*p < .05; **p < .01; ***p < .001

Table C2: Outcome Variable Regressions, State Stage (No Control Variables)

	Noncompliance Support	Perceived Legitimacy	Support for Federal Punishment	Support for Anti- compliance Party
	(1)	(2)	(3)	(4)
No COVID Claim	0.045 (0.056)	-0.016 (0.054)	0.026 (0.054)	-0.068 (0.058)
Lack resources	0.167** (0.053)	0.106* (0.051)	0.037 (0.053)	0.009 (0.054)
No COVID X Lack resources	-0.125 (0.076)	-0.071 (0.073)	0.028 (0.075)	0.029 (0.078)
Constant	2.185*** (0.038)	2.451*** (0.038)	2.417*** (0.038)	2.309*** (0.040)
N	3,694	3,687	3,692	3,690
Log Likelihood	-5,313.907	-5,219.479	-5,343.485	-5,318.322
AIC	10,635.810	10,446.960	10,694.970	10,644.640

*p < .05; **p < .01; ***p < .001

Table C3: Composite Support Regression, State Stage

	Support for State Gov. (4-item) (1)	Support for State Gov. (3-item) (2)
No COVID Claim	0.005 (0.015)	0.004 (0.015)
Lack resources	0.031* (0.015)	0.030* (0.015)
No COVID × Lack resources	−0.028 (0.022)	−0.027 (0.021)
Fed. Gov. Approval	−0.118*** (0.023)	−0.116*** (0.023)
State Gov. Approval	0.124*** (0.022)	0.124*** (0.022)
COVID Concern	0.022 (0.018)	0.024 (0.018)
1+ Children in HH	0.029* (0.014)	0.029* (0.014)
Rule of Law	−0.199*** (0.027)	−0.198*** (0.027)
Conservatism	0.066* (0.033)	0.065* (0.033)
Education	−0.024 (0.024)	−0.023 (0.024)
Age	−0.001 (0.0004)	−0.001 (0.0004)
Woman	0.009 (0.011)	0.009 (0.011)
Lived in DDR	−0.023 (0.014)	−0.022 (0.014)
Rurality	−0.013 (0.019)	−0.013 (0.018)
Consider AFD	0.019 (0.020)	0.019 (0.019)
Constant	0.588*** (0.040)	0.584*** (0.039)
N	3,384	3,387
Log Likelihood	−555.081	−525.167
AIC	1,142.162	1,082.334

* p < .05; ** p < .01; *** p < .001

Table C4: Outcome Variable Regressions, State Stage

	Noncompliance Support (1)	Perceived Legitimacy (2)	Support for Federal Punishment (3)	Support for Anti- compliance Party (4)
No COVID Claim	0.028 (0.055)	−0.004 (0.054)	0.061 (0.054)	−0.052 (0.060)
Lack resources	0.115* (0.054)	0.089 (0.052)	0.046 (0.053)	−0.005 (0.056)
No COVID × Lack resources	−0.102 (0.076)	−0.075 (0.074)	−0.006 (0.075)	0.027 (0.080)
Fed. Gov. Approval	−0.438*** (0.083)	−0.210* (0.082)	0.435*** (0.082)	−0.193* (0.094)
State Gov. Approval	0.355*** (0.083)	0.445*** (0.078)	−0.215** (0.081)	0.302*** (0.088)
COVID Concern	0.022 (0.065)	0.145* (0.066)	0.216** (0.068)	0.141* (0.071)
1+ Children in HH	0.102* (0.050)	0.052 (0.048)	0.021 (0.048)	0.082 (0.051)
Rule of Law	−0.688*** (0.095)	−0.332*** (0.093)	−0.812*** (0.087)	−0.604*** (0.096)
Conservatism	0.211 (0.117)	0.282* (0.113)	0.130 (0.106)	0.001 (0.124)
Education	−0.103 (0.090)	−0.060 (0.099)	0.233* (0.097)	0.065 (0.092)
Age	−0.003* (0.001)	−0.001 (0.001)	0.0003 (0.001)	−0.0001 (0.001)
Woman	0.025 (0.039)	0.054 (0.038)	−0.109** (0.038)	−0.021 (0.041)
Lived in DDR	−0.095 (0.050)	−0.001 (0.047)	0.098* (0.046)	−0.057 (0.052)
Rurality	−0.031 (0.068)	−0.046 (0.064)	0.075 (0.064)	−0.073 (0.068)
Consider AFD	0.079 (0.072)	−0.019 (0.069)	−0.012 (0.068)	0.098 (0.074)
Constant	2.927*** (0.149)	2.409*** (0.148)	2.327*** (0.150)	2.605*** (0.156)
N	3,399	3,394	3,397	3,395
Log Likelihood	−4,785.605	−4,740.136	−4,809.949	−4,835.183

AIC	9,603.209	9,512.272	9,651.898	9,702.366
-----	-----------	-----------	-----------	-----------

* $p < .05$; ** $p < .01$; *** $p < .001$

Appendix D. Balance Tests

We use YouGov’s built-in randomization software to randomly assign the experimental treatments. Although treatments should be randomly assigned by design and unassociated with any covariates, we test for association between key demographics and treatment assignment to confirm covariate balance. As shown in Table D1, we find no evidence for imbalance in the assignment of the COVID claim treatment by respondent age ($p = .11$), previous residence in the DDR ($p = .43$), left-right political ideology ($p = .53$), rural residence ($p = .12$), or gender identity ($p = .88$). We find no evidence for imbalance in the assignment of the justification treatment by respondent age ($p = .24$), previous residence in the DDR ($p = .63$), left-right political ideology ($p = .30$), rural residence ($p = .59$), or gender identity ($p = .68$). Our core results are robust to the inclusion of demographic and political covariates, as shown in Tables C3 and C4.

Table D1: Balance Tests

Covariate	P-value from χ^2 test for association across COVID claim treatment assignment	P-value from χ^2 test for association across justification treatment assignment
Age (binned into ten-year intervals)	.11	.24
Previous residence in the DDR	.43	.63
Left-right political ideology	.53	.30
Rural residence	.12	.59
Gender identity	.88	.68