Do COVID-19 Lockdown Policies Weaken Civic Attitudes? Survey Evidence from the United States and Europe*

Moritz Marbach^{a,c,Ⓕ}, Dalston Ward^{c,Ⓕ}, and Dominik Hangartner^{b,c,d}

^aThe Bush School of Government & Public Service, Texas A&M University, College Station, TX 77843

^bCenter for Comparative and International Studies, ETH Zurich, 8092 Zurich, Switzerland

^cImmigration Policy Lab, Stanford University, Stanford, CA 94305, and ETH Zurich, 8092 Zurich, Switzerland

^dDepartment of Government, London School of Economics and Political Science, London WC2A 2AE, UK

[Ⓕ]M.M. and D.W. contributed equally and the ordering of their names is randomized

Abstract

The spread of COVID-19 has prompted governments to implement restrictions to public life and the economy (broadly, “lockdown policies”). We evaluate the short- and medium-term impacts of lockdown policies on civic attitudes relevant for the health of democracy. Using survey data collected daily between March and May 2020 in the United States and four European countries from 27,317 respondents and a difference-in-differences design, we document that lockdown policies increase authoritarian values and, to a lesser degree, support for autocracy. We find no effects on satisfaction with democracy and the government, out-group hostility, and generalized trust. Additional analyses reveal that the effects persist for at least seven weeks, that these effects persist after the lifting of lockdowns, and that economic support packages have limited ability to alleviate these effects. We discuss the implications of our findings for how governments might need to accompany lockdowns with measures that strengthen civic culture.

*We thank Scherwin Bajka and Miriam Hofstetter for outstanding research assistance. AER [Ⓕ]-certification code for random author ordering of M.M. and D.W.: QxfVNndapvn.

The COVID-19 pandemic has quickly become the largest public health crisis in living memory for countries across the globe. Governments' priority has been to contain and slow the spread of COVID-19 through the implementation of a range of lockdown measures that are extraordinary, at least for democratic countries. While several studies point to the effectiveness of lockdown measures in decreasing COVID-19-related infections and deaths ([Acemoglu et al., 2020](#); [Chernozhukov et al., 2020](#)), these restrictive policies might also exacerbate the economic ([Alvarez et al., 2020](#)), social ([Bonaccorsi et al., 2020](#)), and psychological ([Holman et al., 2020](#)) ramifications of the pandemic.

These developments—the public health crisis, the economic recession, and the lockdown measures—have the potential for major and lasting political repercussions. Already during the first wave, newspapers, public intellectuals, and governments have voiced concerns that the COVID-19 outbreak could trigger a democratic backslide, fuel the rise of populist leaders, and foster authoritarian values and out-group hostility ([Baker, 2020](#); [Mudde, 2020](#)).

Evidence from past health crises suggest that these concerns are not unfounded. Research documents that the Black Death pandemic (1346–1353) in Europe fueled antisemitism and the persecution of Jewish minorities ([Jedwab et al., 2019](#)), that the 1918 influenza pandemic decreased generalized trust in the US ([Aassve et al., 2020](#)), that the 2014 Ebola outbreak triggered anti-immigrant attitudes ([Adida et al., 2018](#)) and an analysis of several past epidemics finds negative effects on confidence and trust in the government ([Aksoy et al., 2020](#)).

This evidence from historical health crises is joined by a rapidly growing literature on the political repercussions of the ongoing COVID-19 pandemic. A sizeable number of surveys document that support for democracy has been eroding ([Alsan et al., 2020](#); [Amat et al., 2020](#)), while support for autocratic rule, illiberal policies, and outgroup hostility have been increasing during the first wave of the COVID-10 pandemic ([Arceneaux et al., 2020](#); [Bartos et al., 2020](#)). Note, however, that several studies suggest that this skeptical view of democracy is often joined by “rally 'round the flag effects” leading to higher support for the incumbent government ([Bol et al., 2020](#); [De Vries et al., 2020](#); [Reeskens et al., 2020](#); [Schraff, 2020](#); [Yam et al., 2020](#)).

We contribute to this research on the general political consequences of COVID-19 by providing causal evidence on the specific impacts of lockdown policies. Understanding the repercussions of these policies is important because, for those citizens not directly infected with COVID-19, lockdown policies are some of the most visible consequences of the pandemic. Furthermore, understanding the effects of lockdown policies matters because governments have control over their introduction and removal; in contrast, policymakers can only indirectly influ-

ence COVID-19's spread. Our design, described in more detail below, not only estimates the effects of introducing lockdown policies, but allows us to examine whether their effects dissipate after their removal, at least in the short- and medium-term. The effects of removals are relevant for both theoretical and policy reasons, yet have not been well covered in the existing literature on COVID-19's political repercussions.

Two inferential challenges stand in the way of estimating effects of lockdown policies. First, to disentangle the effects of lockdown policies from those of COVID-19's spread, researchers need a credible control group of similar areas, ideally within the same country, with less stringent lockdown policies (or no lockdown at all). The few existing studies of the effects of lockdowns address this challenge by resorting to survey experiments that ask respondents to react to hypothetical policy scenarios ([Arceneaux et al., 2020](#)) or general-purpose surveys that were already in the field at the onset of the pandemic, and which allow for before-and-after comparisons of real-world, national-level policies using interrupted time-series model in a single country ([Reeskens et al., 2020](#); [Sibley et al., 2020](#)) or panel data models covering several countries ([Bol et al., 2020](#)).

The second inferential challenge to comprehensively assessing the political impact of lockdown policies is accessing a broad range of attitudinal measures—particularly those that past research has identified as relevant for the health of democracy ([Almond and Verba, 2015](#)). Ideally, these measures not only cover satisfaction with and support for democracy as a form of government (two questions covered by some general-purpose surveys) but also attitudinal foundations of civic culture, including authoritarian values, generalized trust, and out-group hostility, which might dynamically respond to lockdown policies.

Our tailored survey allows us to overcome both inferential challenges. To leverage subnational variation in lockdown policies, we fielded cross-sectional surveys in France, Germany, Switzerland, the United Kingdom, and the United States on a daily basis between March, May and April 2020. These surveys with a total sample size of 27,317 respondents allow us to compare the attitudes of respondents sampled from the same population before and after the introduction (and removal) of lockdown measures while using similar respondents in other regions of the same country that introduced (removed) lockdown policies later on as a control group. A large battery of survey items enables us to trace the effect of lockdown policies along three dimensions of support for democracy (diffuse and specific support as well as satisfaction with democracy) and three dimensions of civic attitudes that are foundational for democracy (right-wing authoritarianism, out-group hostility and generalized trust). We discuss the motivation for

and measurement of these dimensions in the next section.

1 Measuring Civic Attitudes and Lockdown Policy

To measure how public attitudes shift in response to governmental policies, we fielded daily surveys in France, Germany, Switzerland, the United Kingdom and the United States. In Europe, we use the Respondi online panel while we recruited respondents from Amazon MTurk in the US. During the field period which began near the pandemic's onset in Europe and the US on March 19, 2020 and ended 2 months later on May 19, 2020, we surveyed 27,317 adult respondents with about 90-100 respondents daily per country. Median completion time for our survey was about 5 minutes. In addition to socio-demographic characteristics including age, sex and place of residence, we collected information on the outcome measures described below.

We collected information on the following six dimensions of civic attitudes. First, we measure two dimensions of support for democracy ([Linde and Ekman, 2003](#)): support for democracy as a principle, i.e., diffuse support, and satisfaction with how democracy works in practice and with the government, i.e., specific support. We build our measure of diffuse support for democracy on [Norris \(1999\)](#) and ask respondents to indicate their preferences on a 4-point scale over four political regimes: autocracy, technocracy, military rule, and democracy. We aggregate these survey questions into a single scale using polychoric principal component analysis (PCA); we also apply PCA to all other outcomes discussed below. To measure respondents' satisfaction with how democracy works, we use a direct, standard 4-point question as well as respondents' intentions to turnout in the next general election, treating the intention to engage with democracy as implicit approval of it as a system of government ([Foa and Mounk, 2016](#)). Attitudes towards the incumbent government are elicited using a direct question about respondents' satisfaction with the government as well as respondents' intention to support the incumbent parties or candidate in the next general election.

We supplement these measures of support for democracy with three civic attitudes that extant research has identified as relevant foundations ([Almond and Verba, 2015](#)). First, we use the scale developed by [Bizumic and Duckitt \(2018\)](#) to measure two facets of authoritarianism: authoritarian aggression (i.e., aggressiveness directed against perceived outgroups) and authoritarian submission (i.e., submissiveness to the authorities perceived to be legitimate). Scholars of authoritarianism have long contended that (right-wing) authoritarians' support for subordination of individual freedom and autonomy to the collective and its authority conflicts with a

democratic culture (e.g., [Altemeyer and Altemeyer, 1981](#)). Since [Adorno et al.](#) study on ‘The authoritarian personality’ (1950), most research in sociology and social psychology has treated authoritarian attitudes as a rather fixed personal characteristics. Yet, there is a similarly long stream of research that has argued, and empirically documented, that authoritarian values can be, in part, consequences of economic, political or societal change ([Rokeach, 1960](#); [Lipset, 1960](#); [Ballard-Rosa et al., 2021](#), see, e.g.,). With our authoritarianism scale, we build on this second stream to capture if and how authoritarian attitudes respond to lockdown policies.

Second, to measure out-group hostility, we adapt standard feeling thermometer questions ([Hangartner et al., 2019](#)), and take the difference between the feeling thermometer toward co-citizen (in-group) and the average feeling thermometer toward three immigrant out-groups differing in their origin and cultural distance. Among these out-groups, we deliberately choose origins that were not directly connected to the COVID-19’s origin, as our intent was to not to measure scapegoating of a particular group that politicians or the media may tie to the pandemic, but rather, more generalized hostility towards outsiders that was unrelated to COVID-19’s spread. The third element of civic culture that we capture is generalized trust. We measure this concept with standard generalized-trust survey item and feeling thermometer ratings of co-citizens. For more details, see SM 2.

We combine these outcome data with information on the extent and timing of lockdown policies. For each of the studied countries, we collected information on national and regional government policies restricting citizens’ behavior from official government websites covering the two-month study period in the five surveyed countries. Based on this information, we construct a six-point Lockdown Index (LI). For each of the following three policies in place, we added one to the score: closures of non-essential shops and non-take-out service at bars and restaurants; closures of schools and/or daycares; forbidding of public and private events of 50 people or more. If there is a recommendation to stay at home and/or shelter in place, we also added one. If there is an order to stay at home and/or shelter in place but if that order has several exceptions for essential activities, we added two. Finally, if there is a stay at home and/or shelter in place order with only very limited exceptions and/or strong enforcement, we added three. A value of six corresponds to the strict lockdown enforced in France, while a value of three would correspond to closed schools, restrictions on events, and the ”safer-at-home” orders that some US states enacted. We use the same coding when it comes to government policies reversing previous restrictions but instead of adding the value, we subtract it from the index. The final index encodes the strength of national and regional government restrictions in each region and

each day during our study period. For more details, see SM 2.

Figure 1 shows the developments of our outcome measures and the Lockdown Index over the study period. Over these two months, support for democratic principles among respondents in the United States declined and support for authoritarianism increased. Among the European respondents, support for democracy increased slightly while authoritarian values declined. Respondents' satisfaction with how democracy functions in the United States increased between March and April while it declined to some degree in Europe; government satisfaction followed a similar pattern in both regions. In contrast, in neither Europe nor the US do we observe a discernible trend among respondents' generalized social trust or relative out-group hostility.

Figure 1: Estimates for survey items and scales between March 19 and May 19, 2020 for respondents located in the United States and Europe. The lines describe the trend based on a generalized additive regression with shrinkage cubic regression splines fitted on daily survey estimates displayed as dots. To ease interpretation, estimates and trend lines are rescaled to zero for March 19.

Estimating The Effect of Lockdown Policies

Are these trends shaped by how governments on the local and federal level respond to the pandemic? While at the beginning of our field period, about an equal share of respondents in

the United States and Europe reported being affected by government restrictions, the shares diverged considerably throughout the field period. The share of respondents in the US reporting to be affected by restrictions grew substantially while it remained generally steady early on and before dropping in Europe. The co-occurrence in trends in attitude and the exposure to government tougher COVID-19 response policies suggests that the two are connected. We next present evidence that there is, indeed, a causal relationship.

To assess the effects of COVID-19-related lockdown policies, we take advantage of the staggered imposing and lifting of restrictions in the regions of the surveyed countries and compare responses before and after changes in restrictions. This approach follows a difference-in-differences (DID) design [Angrist and Pischke \(2009\)](#), estimating differences in outcomes between regions with varying levels of restrictions net of any pre-existing differences between regions. We estimate effects with the standard difference-in-difference model specification of linear regression with fixed effects for regions and days.

To obtain a balanced panel of average survey outcomes across regions and days, we use Bayesian multilevel modeling with post-stratification [Gelman and Little \(1997\)](#); [Park et al. \(2004\)](#). Post-stratifying the data allows us to address any imbalance in the panel resulting from the fact that due to survey sampling, we are not surveying respondents from all (small) regions every survey day, and to simultaneously adjust for demographic differences between our sample and the population. Our Bayesian multilevel model predicts a survey response as a function of respondents' place of residence, survey day, their age (in 10 age groups) and sex. We estimate separate models for each survey outcome with a linear predictor taking the following form:

$$\mu_{d,r,a} = \alpha_0 + \alpha_{d,r}^{\text{day} \times \text{reg}} + \alpha_{a,r}^{\text{age-sex} \times \text{reg}} + \alpha_{a,c[r],w[d]}^{\text{age-sex} \times \text{cntry} \times \text{week}}$$

The parameter $\alpha_{\text{day,reg}}$ captures the mean on each survey day in each region net of all regional differences in the age-sex distribution (captured by the parameter $\alpha_{\text{age-sex,reg}}$), shifts in the country-specific age-sex distribution across the survey weeks ($\alpha_{\text{age-sex, cntry, week}}$) and the grand mean (α_0). We use r to index a region, d a survey day, c a country and w a survey week. We model all the three parameter batches hierarchically:

$$\begin{aligned}
\alpha_{d,r}^{\text{day x region}} &\sim N(\alpha_d^{\text{day}} + \alpha_r^{\text{reg}} + \alpha_{d,c[r]}^{\text{day x ctry}}, \sigma_{\text{day x reg}}^2) \\
\alpha_{d,c}^{\text{day x ctry}} &\sim N(\alpha_c^{\text{ctry}}, \sigma_{\text{day x ctry}}^2) \\
\alpha_{a,r}^{\text{age-sex x reg}} &\sim N(\alpha_a^{\text{age-sex}}, \sigma_{\text{age-sex x reg}}^2) \\
\alpha_{a,c,w}^{\text{age-sex x ctry x week}} &\sim N(\alpha_w^{\text{week}}, \sigma_{\text{age-sex x ctry x week}}^2).
\end{aligned}$$

The most important feature of this hierarchical structure is that it shrinks the daily regional mean toward the sum of the daily mean, the regional mean and the country-specific, daily deviation. This shrinkage allows to reduce noise in our post-stratification estimates originating in regions with few daily observations. To complete the model we follow standard practice and assign weakly informative prior densities. For more details, see SM 2.

Our post-stratification estimate $\theta_{d,r}$ —the average survey outcome in region r on survey day d —is constructed by summing over a draw from the posterior predictive density for each age-sex group in each region and on each survey day ($\mu_{d,r,a}^{\text{pred}}$), weighted by the relative size of the age-sex group in a particular region, i.e.:

$$\theta_{d,r} = \sum_a \mu_{d,r,a}^{\text{pred}} \frac{N_{r,a}}{N_r}.$$

To construct the weights, we rely on data from Eurostat (European Statistical Office) and the US Census Bureau. We construct 1,000 balanced panel datasets with these post-stratification estimates by repeatedly drawing from the posterior predictive density.

To estimate the effect of the introducing and reversing restrictions, we combine our post-stratification estimates with the Lockdown Index (LI). Our baseline specification, used for the results in Figure 2, is a two-way fixed-effect estimator:

$$\theta_{d,r} = \gamma_0 + \gamma_d + \gamma_r + \delta \text{LI}_{d,r} + \epsilon_{d,r}$$

where γ_d and γ_r denote the day and regional fixed-effect. We estimate the model on each of the 1,000 balanced panel datasets and summarize the posterior distribution of the δ using the 95% quantile and the mean.

In additional specifications we study of the effect of introducing and reversing severe restrictions ($LI > 3$). These specifications follow a difference-in-difference design with lead and lags for the treatment to study the evolution of the treatment effect (Figure 3). Our specification for introducing severe restrictions takes the following form:

$$\theta_{d,r} = \gamma_0 + \gamma_d + \gamma_r + \sum_{m=-6}^6 \beta_m D(m)_{d,r} + \epsilon_{d,r},$$

where m is the day before/after a policy change and $D(m)_{d,r}$ it is the treatment indicator that takes on the value one on day m for those regions that introduce restrictions (except for $D(6)_{d,r}$ which takes the value one on day 6 and every day thereafter). In this way, the coefficients on the indicators $D(1)_{d,r} - D(6)_{d,r}$ measure the effect of restrictions on day one to six (and thereafter) while the coefficients on the remaining indicators ($D(-6)_{d,r} - D(0)_{d,r}$) are placebo checks to test if there are non-parallel pre-treatment trends. The specification for the reversal of restrictions is the same except that $D(m)_{d,r}$ encodes a reversal.

To estimate the effect of introducing (and reversing) severe restrictions, we fit the DID model to each of the 1,000 balanced panel datasets but we subset to regions that introduced or reversed restrictions during the study period (switcher regions). We remove all regions that either always had restrictions or that never introduced them. When estimating the effect of a reversal, we also remove all days prior to the introduction of the restriction, so as to not violate the parallel trends assumption by construction.

Small Effects of Lockdown Policies on Civic Attitudes

Figure 2 reports our main results. Beginning with panel (A), we find mixed evidence for the effects of lockdown policies on the three scales measuring support for democracy. On the political system scale, which captures preferences over regime types, we see that increasingly severe lockdown policies raise support for autocracy (here and below, we report the posterior poster probability p that the effect is smaller than zero: $p < 0.02$). The effect, however is small: each additional point up the Lockdown Index (e.g., going from no restriction on event size to limiting events to 10 or fewer people) leads to an increase in support for autocracy by 1% of a standard deviation. For both satisfaction with democracy and the government, we find, on average, small positive effects—roughly 0.5% of standard deviation—but we cannot reliably

conclude that these effects are positive ($p = 0.15$ and 0.11 , respectively). Together, these results suggest that the restrictive measures introduced at the height of the first wave of the COVID-19 pandemic had only marginal effects on support for democracy.

We now turn to three measures of attitudes reflecting the health of a civic culture: authoritarianism, generalized trust, and hostility toward outgroups. The analysis reveals that increasingly strict lockdown policies fuelled authoritarian attitudes. The estimates allow us to confidently conclude that this effect is positive ($p < 0.01$) and substantively meaningful: the estimates suggest that going from no restrictions to a full lockdown, e.g., as occurred in France, increases authoritarian values by 10.2% of a standard deviation. We do not find any reliable effects on the other two measures of the attitudinal foundations of democracy. For both generalized trust and relative outgroup hostility, the estimated effects of restrictive policies are centered at zero.

The final panel of Figure 2 presents the results of a manipulation check that supports the validity of our difference-in-differences design. We estimate the effect of lockdown policies on the rate at which citizens report that current government policies restrict their movement. If our design works as intended, we would expect restrictive policies to have a strong positive effect on this outcome. Such a pattern is already visible in the bottom row of Figure 1; our formal estimate confirms this relationship. The analysis indicates that an additional restrictive policy increases the share of citizens reporting a restriction by about 3.5 percentage points ($p < 0.01$).

Next, we consider the dynamics of lockdown policies' impact on civic attitudes: How soon after these policies were introduced did attitudes harden? When governments removed restrictions, did citizens subsequently decrease their support for autocratic government and their endorsement of authoritarian values? The analyses reported in Figure 3 shed light on these questions. These analyses recode our measure of lockdown policies into a binary variable and then estimate the effects of restrictive policies in the days immediately before and after their introduction or removal.

For the political system scale (panel A), we see that 5 to 6 days prior to the introduction of restrictions, there are no differences between regions that will and will not implement restrictions. Once the restrictions are introduced, attitudes only gradually shift towards autocracy, peaking five days after policies are implemented and then staying at roughly that level afterwards. Shifting the focus to the effects of lifting lockdown policies, the analysis uncovers no evidence that removing these restrictions led citizens to move back towards favoring democracy.

The results for authoritarianism tell a largely similar story, although there are two relevant differences. First, we identify an anticipation effect of lockdown policies on the day before they

Figure 2: Bayesian posterior estimates (dots) and 95% credible estimates for the effect of the six-point Lockdown Index on citizens' attitudes and preferences. Estimates are based on linear regression model with region and day fixed-effects fitted on post-stratified survey estimates ($N = 7,254$).

are announced, which likely result from reports of impending restrictions in the press. Second, the effect on authoritarianism more resembles a one-time shift rather than a gradual change. Turning to the effects of lifting lockdown policies, we find no evidence that removals led to a decline in authoritarianism.

The SM provide several additional analyses and robustness tests. First, we present effects separately for all of the survey items comprising our scales (Figure S.11). For these disaggregated measures, we find some of the strongest lockdown effects on (i) support for stratocracy (i.e., military rule), agreement with the (authoritarian) statements that (ii) crime and public disorder necessitates a harder crackdown on troublemakers, and that (iii) the country needs more

discipline, with everyone following the leaders. One potential concern of our main estimates is that rather than changing citizens' attitudes, the introduction of lockdown policies shape how respondents interpret our survey questions. While this type of differential item functioning might affect how respondents answer question (iii) about the need to follow the leaders, this is less likely to affect questions on the preferences over stratocracy and crackdowns on crime and public disorder. Furthermore, if differential item functioning were solely responsible for driving our effects, we might expect that they would dissipate after lockdown policies are lifted.

Second, we corroborate our main findings by fitting regression models to the raw survey data, i.e., without post-stratification (Figure S.12). We additionally show that the effects are robust to controlling for the daily number of COVID-19-related deaths per region (Figures S.5 and S.6) or for allowing for more flexible temporal fixed effects (Figure S.7). Complementing Figure 3, we further present these specifications for our other outcome variables (Figure S.1). Lastly, we show overall effects of introductions (Figure S.3) and reversals (Figure S.4) using the binary version of our restriction measure.

(A) Political system scale

(B) Authoritarian attitudes scale

Figure 3: Bayesian posterior estimates (dots) and 95% credible estimates for the effect of introducing and removing restrictions (six-point Lockdown Index above 3) on the political system and authoritarianism scale. Estimates are based on linear regression model with region and day fixed-effects fitted on post-stratified survey estimates. We remove all regions without policy variation and all days prior to the introduction of the restrictions when estimating the effect of removals. $N = 3,844$ (left panels) and $N = 3,806$ (right panels).

Discussion

Our results show that the introduction of restrictive policies increased authoritarianism and support for non-democratic regimes, but had few consequences for other civic attitudes. Together, they suggest a concentrated impact of restrictive policies during the first wave of COVID-19 on attitudes underpinning pluralistic and democratic societies: while we uncover a turn toward authoritarianism, this is not accompanied by scapegoating of out-groups, a loss of generalized trust, or diminished satisfaction with government performance.

Our dynamic analysis in Figure 3 reveals important asymmetries in the effects of the introduction and lifting of lockdowns: governments’ decisions to implement lockdown measures have effects for non-democratic and authoritarian values that are not simply remedied by their removal. To explore whether these effects decay over longer periods of time, with or without removals, we refit the “restriction introduction” models from Figure 3 to estimate the effects of policies on a weekly basis for up to seven weeks post-lockdown. The results, shown in Figure S.10, indicate that even seven weeks after the introduction of lockdowns, the effects on regime preferences and authoritarianism persist. These persistent effects, combined with the fact that subsequent waves of the COVID-19 pandemic have led to the reimposing of lockdowns, raise the specter of compounding effects that may begin to chip away at an important pillar of democratic civic culture.

In their efforts to fight the pandemic, governments often combined lockdown policies with supportive measures to strengthen the economy and aid particularly affected business and workers. Can these supportive policies dampen the harmful effects of lockdown policies on civic attitudes? To investigate this possibility, we augment our baseline DID models by interacting lockdown stringency with a measure of the number of economic support packages announced by the government (see SM section 1.2 for details on this measure). The results of this analysis, shown in Figure S.8, provide little evidence that economic support packages dampen the effects of lockdown policies on either the political system scale (one-sided interaction $p = 0.31$) or the authoritarianism scale (one-sided interaction $p = 0.34$). Though targeted support packages are important for lightening the economic burden of the COVID-19 pandemic, these results point to the need for governments to consider other approaches to soften the repercussions of lockdown policies on civic attitudes. We discuss this point in the concluding paragraph.

Conclusion

Our findings of small but, at least over the short- and medium-term, persistent effects of lockdown policies on authoritarian values and support for non-democratic forms of government—two domains of civic attitudes essential for the functioning of democracy—establish empirical foundations for some of the concerns voiced by commentators and governments about the possible political repercussions of the pandemic (Baker, 2020; Mudde, 2020). Yet, our findings do not paint an entirely bleak picture, since several other important correlates of a healthy civic culture—from evaluations of government performance to generalized trust and out-group hostility—were not affected by lockdown policies during the first wave of the COVID-19 pandemic. Furthermore, the effect we find on preferences for non-democratic regimes is small in magnitude. While these mixed findings require a nuanced assessment of the political repercussions of lockdown policies so far, they serve as a reminder that the pandemic and the policies implemented to mitigate its spread have the potential to catalyze receding public support for one of the pillars of liberal democracy (Foa and Mounk, 2016; Runciman, 2018).

In the public and scholarly debate surrounding the effectiveness and drawbacks of lockdown policies, the trade-off between the economy and public health has been extensively discussed (Acemoglu et al., 2020; Alvarez et al., 2020). Our study points to separate trade-off dimension that to date, has received less attention: the small but persistent effects of lockdown policies on support for non-democratic regimes and authoritarian values. While its harder to quantify the costs of an increase in authoritarian attitudes vis-a-vis, say, GDP output loss, the long-term consequences might nevertheless be concerning. Compounding these concerns is the finding that government aid distributed to particularly affected industries and workers seems ineffective in reducing the political repercussions of lockdown policies documented here.

Given the strong evidence on the effectiveness of lockdown measures in containing the spread of COVID-19 (see, e.g., Chernozhukov et al., 2020), the lesson for policy-makers from our study is, of course, not to prematurely lift measures for concerns about eroding certain civic attitudes. Instead, our findings point to the importance of coupling restrictive policies with measures that counteract the rise in authoritarian values and support for autocracy in liberal democracies. In addition to more long-term measures shown to strengthen civic culture such as civic education programs for students (Langton and Jennings, 1968; Levine, 2007) and adults (Blair, 2003; Finkel, 2003), accompanying lockdown measures with public communication about the evidence base, purpose and timing of the various restrictions (Bellantoni et al.,

2020) might be particularly effective in mitigating some of the negative repercussions documented here. We hope that our study can sound a clarion call for more research into actionable and scalable interventions that effectively strengthen civic attitudes in the post-COVID world.

References

- Aassve, A., G. Alfani, F. Gandolfi, M. Le Moglie, et al. (2020). Epidemics and Trust: The Case of the Spanish Flu. Working Paper 661, IGIER.
- Acemoglu, D., V. Chernozhukov, I. Werning, and M. D. Whinston (2020). A Multi-Risk SIR Model With Optimally Targeted Lockdown. Working Paper 27102, National Bureau of Economic Research.
- Adida, C. L., K. Y. Dionne, and M. R. Platas (2018). Ebola, Elections, and Immigration: How Politicizing an Epidemic can Shape Public Attitudes. *Politics, Groups, and Identities*, 1–27.
- Adorno, T., E. Frenkel-Brenswik, D. J. Levinson, and R. N. Sanford (1950). *The Authoritarian Personality*. Verso Books.
- Aksoy, C. G., B. Eichengreen, and O. Saka (2020). The Political Scar of Epidemics. Working Paper 27401, National Bureau of Economic Research.
- Almond, G. A. and S. Verba (2015). *The Civic Culture: Political Attitudes and Democracy in Five Nations*. Princeton: Princeton University Press.
- Alsan, M., L. Braghieri, S. Eichmeyer, M. J. Kim, S. Stantcheva, and D. Y. Yang (2020). Civil Liberties in Times of Crisis. Working Paper 27972, National Bureau of Economic Research.
- Altemeyer, R. A. and B. Altemeyer (1981). *Right-wing Authoritarianism*. University of Manitoba press.
- Alvarez, F. E., D. Argente, and F. Lippi (2020). A Simple Planning Problem for Covid-19 Lockdown. Working Paper 26981, National Bureau of Economic Research.
- Amat, F., A. Arenas, A. Falcó-Gimeno, and J. Muñoz (2020). Pandemics Meet Democracy. Experimental Evidence From the COVID-19 Crisis in Spain. Available at SocArXiv: <https://doi.org/10.31235/osf.io/dkusw>.
- Angrist, J. D. and J.-S. Pischke (2009). *Mostly Harmless Econometrics. An Empiricist's Companion*. Princeton: Princeton University Press.
- Arceneaux, K., B. N. Bakker, S. Hobolt, and C. E. De Vries (2020). Is COVID-19 a Threat to Liberal Democracy? Available at PsyArxiv: <https://psyarxiv.com/8e4pa/>.

- Baker, P. C. (2020). How Will the World Emerge from the Coronavirus Crisis. *The Guardian*.
- Ballard-Rosa, C., M. Malik, S. Rickard, and K. Scheve (2021). The Economic Origins of Authoritarian Values: Evidence from Local Trade Shocks in the United Kingdom. *Comparative Political Studies* 0(107), 1–33.
- Bartos, V., M. Bauer, J. Cahlíková, and J. Chytilová (2020). Covid-19 Crisis Fuels Hostility against Foreigners. Working Paper 13250, IZA.
- Bellantoni, A., K. Badr, and C. Alfonsi (2020). Transparency, Communication and Trust: The Role of Public Communication in Responding to the Wave of Disinformation About the New Coronavirus. OECD Policy Responses to Coronavirus (COVID-19), OECD.
- Bizumic, B. and J. Duckitt (2018). Investigating Right Wing Authoritarianism With a Very Short Authoritarianism Scale. *Journal of Social and Political Psychology* 6(1), 129–150.
- Blair, H. (2003). Jump-Starting Democracy: Adult Civic Education and Democratic Participation in Three Countries. *Democratization* 10(1), 53–76.
- Bol, D., M. Giani, A. Blais, and P. J. Loewen (2020). The Effect of COVID-19 Lockdowns on Political Support: Some Good News for Democracy? *European Journal of Political Research* (forthcoming).
- Bonaccorsi, G., F. Pierri, M. Cinelli, A. Flori, A. Galeazzi, F. Porcelli, A. L. Schmidt, C. M. Valensise, A. Scala, and W. Quattrocioni (2020). Economic and Social Consequences of Human Mobility Restrictions Under COVID-19. *Proceedings of the National Academy of Sciences* 117(27), 15530–15535.
- Chernozhukov, V., H. Kasahara, and P. Schrimpf (2020). Causal Impact of Masks, Policies, Behavior on Early COVID-19 Pandemic in the US. *Journal of Econometrics* 220(1), 23–62.
- De Vries, C. E., B. N. Bakker, S. Hobolt, and K. Arceneaux (2020). Crisis Signaling: How Italy’s Coronavirus Lockdown Affected Incumbent Support in Other European Countries. Available at SSRN: <https://ssrn.com/abstract=3606149>.
- Finkel, S. E. (2003). Can Democracy be Taught? *Journal of Democracy* 14(4), 137–151.
- Foa, R. S. and Y. Mounk (2016). The Danger of Deconsolidation: The Democratic Disconnect. *Journal of democracy* 27(3), 5–17.

- Gelman, A. and T. C. Little (1997). Postratification into Many Categories Using Hierarchical Logistic Regression. *Survey Methodology* 23(2), 127–135.
- Hangartner, D., E. Dinas, M. Marbach, K. Matakos, and D. Xefteris (2019). Does Exposure to the Refugee Crisis Make Natives More Hostile? *American Political Science Review* 113(2), 442–455.
- Holman, E. A., R. R. Thompson, D. R. Garfin, and R. C. Silver (2020). The Unfolding COVID-19 Pandemic: A Probability-Based, Nationally Representative Study of Mental Health in the United States. *Science Advances* 6(42), eabd5390.
- Jedwab, R., N. D. Johnson, and M. Koyama (2019). Negative Shocks and Mass Persecutions: Evidence From the Black Death. *Journal of Economic Growth* 24(4), 345–395.
- Langton, K. P. and M. K. Jennings (1968). Political Socialization and the High School Civics Curriculum in the United States. *American Political Science Review* 62(3), 852–867.
- Levine, P. (2007). *The Future of Democracy: Developing the Next Generation of American Citizens*. Medford: Tufts University Press.
- Linde, J. and J. Ekman (2003). Satisfaction With Democracy: A Note on a Frequently Used Indicator in Comparative Politics. *European Journal of Political Research* 42(3), 391–408.
- Lipset, S. M. (1960). *Political Man: The Social Bases of Politics*. Anchor Books.
- Mudde, C. (2020). 'Wartime' Coronavirus Powers Could Hurt Our Democracy – Without Keeping Us Safe. *The Guardian*.
- Norris, P. (1999). *Critical Citizens: Global Support for Democratic Government*. Oxford: Oxford.
- Park, D. K., A. Gelman, and J. Bafumi (2004). Bayesian Multilevel Estimation with Poststratification: State-Level Estimates from National Polls. *Political Analysis* 12(4), 375–385.
- Reeskens, T., Q. Muis, I. Sieben, L. Vandecasteele, R. Luijckx, and L. Halman (2020). Stability or Change of Public Opinion and Values During the Coronavirus Crisis? Exploring Dutch Longitudinal Panel Data. *European Societies* (forthcoming).

- Rokeach, M. (1960). *The Open and Closed Mind: Investigations into the Nature of Belief Systems and Personality Systems*. Basic books.
- Runciman, D. (2018). *How Democracy Ends*. New York: Basic Books.
- Schraff, D. (2020). Political Trust During the Covid-19 Pandemic: Rally Around the Flag or Lockdown Effects? *European Journal of Political Research* (forthcoming).
- Sibley, C. G., L. M. Greaves, N. Satherley, M. S. Wilson, N. C. Overall, C. H. Lee, P. Milojev, J. Bulbulia, D. Osborne, and T. L. Milfont (2020). Effects of the COVID-19 Pandemic and Nationwide Lockdown on Trust, Attitudes Toward Government, and Well-Being. *American Psychologist* 75(5), 618–630.
- Yam, K. C., J. C. Jackson, C. M. Barnes, J. Lau, X. Qin, and H. Y. Lee (2020). The Rise of COVID-19 Cases Is Associated With Support for World Leaders. *Proceedings of the National Academy of Sciences* 117(41), 25429–25433.

Supplementary materials

“Do COVID-19 Lockdown Policies Weaken Civic Attitudes? Survey Evidence from the United States and Europe”

Contents

1	Data Collection and Preparation	2
1.1	Survey	2
1.2	Government Policies	4
1.3	COVID-19 Death Counts	6
1.4	Data Preparation	7
1.5	Outcome Coding	7
2	Methodology	8
2.1	Multilevel regression with poststratification	8
2.1.1	Bayesian Multilevel Model	8
2.1.2	Model Selection	11
2.1.3	Poststratification	11
2.2	Difference-in-Differences	12
3	Additional Results	13
	References	29

1 Data Collection and Preparation

1.1 Survey

The target population for our daily, cross-sectional surveys included all adult residents that are at least 18 years old and were living in France (excluding overseas territories and Corsica), the United Kingdom (UK; excluding Northern Ireland and British overseas territories), Switzerland, Germany and the United States (US; excluding Alaska and Hawaii).

The daily target sample was $N = 75$ per country for the four European countries and $N = 100$ for the US. We realized, on average, 85 survey completions (per country) in Europe and 104 in the US. Field work began on March 16, 2020 in the US and on March 19, 2020 in Europe. We ended field work on May 19 in Europe and continued surveying in the US until July 9. Our analysis sample is the common field period from March 19 to May 19.

We conducted our surveys online using existing respondent pools. We used online surveys as they pose no health risks to enumerators or respondents. For the four European countries, we contracted with the survey company Respondi, and for the US, we recruited respondents from Amazon’s Mechanical Turk (MTurk). Extant research shows that the online survey providers with which we contract, provide samples that allow for inferences (see Berinsky et al., 2012 for MTurk and Bansak et al., 2016 for Respondi).

All adults (of least 18 years of age) that lived in the five study countries were eligible to participate. Participants were informed by our online survey providers that they had been invited to participate in a social science study. Participants had the opportunity to decide to click or not on an invitation link. Participants that wished to participate clicked on the invitation link and were directed to the survey landing page provided by the survey software Qualtrics.

On average, respondents received 0.70 Euros for completing our survey. The survey data collection was approved by the Ethics Commission of ETH Zurich (EK 2020-N-43).

Our survey included a series of survey items to measure the concepts described in the main text. We describe these survey items below. We additionally collected some basic demographic information about respondents, including their region of residence.

The first dimension of civic attitudes that our survey captured is regime support. We elicited respondents’ support for different regimes using the following survey prompt: “Below, we describe various types of political systems. For each one, would you say it is a very good, fairly good, fairly bad or very bad way of governing?” The associated items read as follows:

- Having a strong leader who does not have to bother with parliament
- Having experts, not government, make decisions according to what they think is best for the country

- Having the army rule the country
- Having a democratic political system

We next asked respondents about their satisfaction with the performance of democracy in their country and the current government. We asked “How satisfied are you with the way democracy works in [Respondent’s country]?” and “How satisfied are you with the way the people now in national office are handling [Respondent’s country’s] affairs?” Respondents could answer with a four-point scale ranging from very satisfied to very dissatisfied.

In addition, we asked respondents about whom they intended to vote for in the next general election. In the US, the survey question read as follows: “If there were a presidential election tomorrow, would you vote for the Democrat, the Republican, or another candidate?” (followed by the items “Democrat”, “Republican”, “Another candidate” and “I wouldn’t vote”). In Europe the survey question read: “If there were a national election tomorrow, for which party would you vote?” We list all major parties as well as the items “Other (please specify)” and “I wouldn’t vote” as response options.

The second dimension of civic attitudes our survey aimed to capture gauge the attitudinal foundations of a civic culture. To measure authoritarian attitudes, we asked respondents to indicate agreement or disagreement with the following statements using a seven-point scale ranging from very strongly agree to very strongly disagree:

- It’s great that many young people today are prepared to defy authority.
- What our country needs most is discipline, with everyone following our leaders in unity.
- Our society does NOT need tougher government and stricter laws.
- The facts on crime and the recent public disorders show we have to crack down harder on troublemakers, if we are going preserve law and order.

To measure generalized trust we used the survey question: “Generally speaking, would you say that most people can be trusted or that you can’t be too careful in dealing with people?” Respondents were presented an 11-point scale with poles labeled “0 - Can’t be too careful” and “10 - Most people can be trusted.”

We used feeling thermometer questions to measure outgroup hostility. The survey prompt read as follows: “We would like to get your feelings toward a number of groups on a feeling thermometer. A rating of 0 degrees means you feel as cold and negative as possible. A rating of 100 degrees means you feel as warm and positive as possible. You would rate the group at 50 degrees if you don’t feel particularly positive or negative toward the group.”

- How do you feel toward [demonym for people from respondent's country]?
- How do you feel toward immigrants from Europe?
- How do you feel toward immigrants from Syria?
- How do you feel toward immigrants from Australia?

1.2 Government Policies

Our data on governments' policy responses come from official government websites and media outlets. We collected this information at the national level for all five countries, and across various regional level: states (Germany and the US), cantons (Switzerland), regions (France), and constituent countries (UK). Our data collection covered all policies issued by governments in responses to the pandemic up to May 19. We focused on measures, guidelines, laws, and ordinances that aim to contain or slow the spread of COVID-19 and/or compensate citizens and business for the hardship due to COVID-19. In addition, we also collected information on the lifting of restrictions.

In total, we collected information on 2,961 policies across all survey countries (US: 1,364; UK: 123; France: 57; Germany: 874 and Switzerland 543).

Based on information about the date of announcement and content of each policy, we constructed a six-point *Lockdown Index*. For each of the following three policies in place, we added one to the score: closures of non-essential shops and non-take-out service at bars and restaurants; closures of schools and/or daycares; forbidding of public and private events of 50 people or more. If there is a recommendation to stay at home and/or shelter in place, we also added one. If this is an order to stay at home and/or shelter in place but if that order has several exceptions for essential activities, we added two. Finally, if there is a stay at home and/or shelter in place order with only very limited exceptions and/or strong enforcement, we added three. We used the same coding when it comes to government policies reversing previous restrictions but instead of adding the value, we subtract it from the index. The final index encodes the strength of national and regional government restrictions in each region and each day during our study period. Figure S.1 shows the average *Lockdown Index* across five countries.

For additional analysis reported below, we also collected data on the number of economic support packages announced by the government. Our *Economic support index (ESI)* is constructed by summing the number of support packages enacted by the government along three dimensions: rent and/or mortgage forbearance; small-business loans; direct monetary transfers to citizens (e.g., unemployment and direct cash transfers). For each new package announced, we added one to the index on the day it is announced. For example, the government announces a stimulus package that

Figure S.1: Average *Lockdown Index* across five countries during the analysis period (March 19 to May 19).

increases unemployment benefits and offers loans to small business. The index goes from zero to two. One week later, the government announces support for rent payments and additional small business loans—the index then rises from two to four. We collected this data at the national and regional levels.

Our data collection and measurement strategy for the *Lockdown Index* is distinct from the Oxford COVID-19 Government Response Tracker (OxCGRT) and the CoronaNet Research Project, which focus primarily on measuring the stringency of COVID-19 response policies on the national level (Hale et al., 2020; Cheng et al., 2020). The OxCGRT recently began to also release data on policies at the subnational level in the US and construct a subnational stringency index (Hale et al., 2020). Their stringency index is made up of eight indicators: School closures; closings of workplaces; cancellations of public events; limits on private gatherings; closing of public transport and orders to “shelter-in-place” and otherwise confine to the home; restrictions on internal movement between cities/regions and restrictions on international travel. Each of these indicators is ordinal

	M1	M2
Lockdown Index=2	18.30*** (1.52)	14.59*** (0.86)
Lockdown Index=3	20.81*** (1.43)	18.46*** (0.89)
Lockdown Index=4	28.07*** (1.40)	21.97*** (0.94)
Lockdown Index=5	29.95*** (1.41)	25.59*** (1.01)
R ²	0.26	0.81
Num. obs.	3,100	3,100
State FE	No	Yes
Day FE	No	Yes

*** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$

Table S.1: OLS regression estimates and standard error. The dependent variable is the daily OxCGRT Stringency Index. Data include all days between March 19 and May 19 across 50 US States (incl. Washington D.C.).

and they are summed to construct the additive Stringency Index.

We here compare our *Lockdown Index* with the OxCGRT stringency index, by regressing the OxCGRT stringency index on a series of indicators for the *Lockdown Index* levels. Table S.1 demonstrates that there is a robust positive correlation despite only partial overlap between the constituting indicators. The six-point *Lockdown Index* explains about a quarter of the variation in the near-continuous OxCGRT stringency index. When we add fixed effects for each state and each day, this correlation persists and the R^2 increases to 0.81. This means that state-specific changes of the *Lockdown Index*—the variation used by the difference-in-differences (DID) models in our analysis—are strongly associated with the state-specific changes in the OxCGRT stringency index.

1.3 COVID-19 Death Counts

Data on the number of COVID-19 related deaths come from public health officials (e.g., the Swiss Federal Health Office or the Robert Koch Institute) whenever possible. We supplement the public health data with information from public health researchers when necessary. We focus on deaths due to COVID-19 in favor of cases for two reasons: first, the number of deaths is less likely to also depend on a country's testing regime, and second, France did not report case totals at the subnational level for the period covered by our data.

1.4 Data Preparation

We realized a gross sample of 30,200 survey responses. From those, we remove everyone that did not consent to the survey terms or that was a minor at the time of survey ($N = 274$ survey responses). In Europe, some respondents started the survey, closed it and restarted a few hours later ($N = 70$ individuals). We deal with these as follows: We first remove all survey participations with a completion rate of less than 10%, then chronologically order each respondents' attempts, and keep the first response. This removes $N = 73$ survey responses from the gross sample. Lastly, we remove all survey responses outside the common field work period of March 19 to May 19 ($N = 2,457$ survey responses). The final net sample includes 27,317 survey responses. Table S.2 shows the descriptive statistics for our survey sample.

Variable	Mean	SD	Q(0.05)	Q(0.25)	Q(0.75)	Q(0.95)
Age	47.71	15.55	24.00	34.00	60.00	73.00
Female	0.48	0.50	0.00	0.00	1.00	1.00
Political System Scale	-0.00	1.00	-1.23	-0.80	0.71	1.81
Authoritarian Attitudes Scale	-0.00	1.00	-1.80	-0.55	0.67	1.72
Democratic Satisfaction Scale	0.00	1.00	-2.48	-0.41	0.42	1.25
Government Satisfaction Scale	-0.00	1.00	-1.49	-0.83	1.04	1.71
Trust Scale	0.00	1.00	-1.83	-0.64	0.76	1.48
Relative Outgroup Hostility	21.25	26.56	-13.67	0.00	37.67	70.00
Restricted Movement	0.60	0.49	0.00	0.00	1.00	1.00

Table S.2: Descriptive statistics for the survey sample used in the multilevel models.

1.5 Outcome Coding

We next describe the outcome measures that we construct from the survey items. In the list below, outcome measures used in the main analysis are in *italic* and survey items are in **bold** (see above for descriptions of our survey items). With the exception of the outcome *Relative outgroup hostility*, we construct outcome scores by fitting principal component analyses (PCA) with polychoric correlations on the pooled data. We standardize the PCA scores to mean of zero and standard deviation of one.

- *Political system scale*: Scores from the first component of a PCA of **Democracy best**, **Technocracy best**, **Stratocracy best**, **Autocracy best**.
- *Democratic satisfaction scale*: Scores from the first component of a PCA of **Democratic satisfaction** and responses of “I wouldn’t vote” to **Vote choice**.

- *Government satisfaction scale*: Scores from the first component of a PCA of **Government satisfaction** and reported intention to support the incumbent parties/candidate **Vote choice**.
- *Authoritarian attitudes scale*: Scores from the first component of a PCA of **Defy authority**, **Follow leaders**, **Strict laws**, and **Crackdown**.
- *Trust scale*: Scores from the first component of a PCA of **Interpersonal Trust** and **Feeling Thermometer: Natives**.
- *Relative outgroup hostility*: The difference between **Feeling thermometer: Natives** and the respondent-level average of **Feeling thermometer: Europeans**, **Feeling thermometer: Australians**, and **Feeling thermometer: Syrians**.
- *Restricted movement*: Indicator for a response of “Yes” to the item “Due to Coron-avirus/COVID-19, do you face any restrictions in leaving your home today?”

2 Methodology

In this section, we describe how we construct the daily, regional estimates using multilevel regression with poststratification (MRP; [Gelman and Little, 1997](#); [Park et al., 2004](#)). We use these estimates as inputs for the difference-in-differences (DID) models that allow us to estimate the effect of government restrictions.

2.1 Multilevel regression with poststratification

In a first step we construct a Bayesian multilevel model that predicts a survey outcome as a function of respondents’ place of residence, survey day, their age and sex. We then use the posterior predictive density of this model to form an estimate for all age-sex groups in all regions and across all survey days. By summing over all age-sex groups within a region and per survey day, we obtain an estimate of our target quantity.

Using multilevel regression with poststratification (MRP) allows us to obtain daily estimates across all regions that are more efficient as compared to raw survey means. Simultaneously, MRP allows us to account for i) potential differences between the age-sex distribution in the sample and the population as well as ii) potential shifts in the age-sex distribution over time.

2.1.1 Bayesian Multilevel Model

For each survey outcome, we estimate a separate multilevel model. The linear predictor of all models takes the following form:

$$\mu_{d,r,a} = \alpha_0 + \alpha_{d,r}^{\text{day} \times \text{reg}} + \alpha_{a,r}^{\text{age-sex} \times \text{reg}} + \alpha_{a,c[r],w[d]}^{\text{age-sex} \times \text{cntry} \times \text{week}}$$

The parameter $\alpha_{\text{day,reg}}$ captures the mean on each survey day in each region net of all regional differences in the age-sex distribution (captured by the parameter $\alpha_{\text{age-sex,reg}}$), shifts in the country-specific age-sex distribution across the survey weeks ($\alpha_{\text{age-sex, cntry, week}}$) and the grand mean (α_0). We use the Gelman-Hill notation (Gelman and Hill, 2006) to let r (region) or d (survey day) select the corresponding index c (country) and w (survey week).

We model all the three parameter batches hierarchically:

$$\begin{aligned}\alpha_{d,r}^{\text{day} \times \text{region}} &\sim N(\alpha_d^{\text{day}} + \alpha_r^{\text{reg}} + \alpha_{d,c[r]}^{\text{day} \times \text{ctry}}, \sigma_{\text{day} \times \text{reg}}^2) \\ \alpha_{d,c}^{\text{day} \times \text{ctry}} &\sim N(\alpha_c^{\text{ctry}}, \sigma_{\text{day} \times \text{ctry}}^2) \\ \alpha_{a,r}^{\text{age-sex} \times \text{reg}} &\sim N(\alpha_a^{\text{age-sex}}, \sigma_{\text{age-sex} \times \text{reg}}^2) \\ \alpha_{a,c,w}^{\text{age-sex} \times \text{ctry} \times \text{week}} &\sim N(\alpha_w^{\text{week}}, \sigma_{\text{age-sex} \times \text{ctry} \times \text{week}}^2).\end{aligned}$$

The most important feature of this hierarchical structure is that it shrinks the daily regional mean toward the sum of the daily mean, the regional mean and the country-specific, daily deviation. This shrinkage reduces noise in our poststratification estimates originating from regions with few daily observations.

To complete the model we assign prior densities to all parameters. We follow standard practice and assign weakly informative prior densities (Gelman et al., 2014). For all group-specific parameters we use normal densities:

$$\begin{aligned}\alpha_r^{\text{reg}} &\sim N(0, \sigma_{\text{reg}}^2) \\ \alpha_c^{\text{ctry}} &\sim N(0, \sigma_{\text{ctry}}^2) \\ \alpha_d^{\text{day}} &\sim N(0, \sigma_{\text{day}}^2) \\ \alpha_w^{\text{week}} &\sim N(0, \sigma_{\text{week}}^2) \\ \alpha_a^{\text{age-sex}} &\sim N(0, \sigma_{\text{age-sex}}^2).\end{aligned}$$

We use a unit-exponential density as prior for all five scale parameters of these normal densities and another a zero-centered normal density for the grand mean:

$$\sigma_i^2 \sim \text{Exp}(1)$$

$$\alpha_0 \sim N(0, 10).$$

We implement this model using the R package `rstanarm` (Goodrich et al., 2020) which is specifically tailored to estimate Bayesian multilevel models using Hamiltonian Monte Carlo (HMC) and the No U-Turn Sampler (Hoffman and Gelman, 2014). The generic syntax for our model with a continuous outcome takes the following form:

```
mod <- stan_glmer(formula=
  outcome ~ (1|country) + (1|region) +
    (1|svyweek) + (1|svydate) +
    (1|ageg_sex) +
    (1|svydate:country) +
    (1|svydate:region) +
    x(1|ageg_sex:region) +
    x(1|svyweek:ageg_sex:country),
  family = gaussian(link = "identity"),
  na.action = "na.omit",
  prior = default_prior_coef(family),
  prior_intercept = default_prior_intercept(family),
  prior_aux = exponential(autoscale = TRUE),
  prior_covariance = decov(),
  prior_PD = FALSE,
  chains = 8,
  cores = 8,
  iter = 2000,
  algorithm = "sampling")
```

We use 8 chains and iterate each chain 2,000 times. Half of the samples for each chain are used as warm-up to tune the sampler and the other half are taken as samples from the posterior density. Table S.1 reports a series of standard HMC diagnostics (Betancourt, 2017). All of those suggest that the estimates are reliable.

Outcome	Diverg. (in %)	Max. treedepth exceeded (in %)	Pathological chains?	ESS < 800 (in %)	$\hat{R} > 1.1$ (in %)
Political System Scale	0.04	0	0	0.01	0
Authoritarian Attitudes Scale	0.04	0	0	0.01	0
Democratic Satisfaction Scale	0.09	0	0	0.01	0
Government Satisfaction Scale	0.18	0	0	0.03	0
Trust Scale	0.06	0	0	0.01	0
Relative Outgroup Hostility	0.33	0	0	0.00	0
Restricted Movement	0.00	0	0	0.01	0

Table S.1: HMC diagnostics. In the order of columns: Percent divergent iterations, percent of iterations saturating the maximum tree depth, number of chains for which E-BFMI indicates pathological behavior, percent of parameters with an effective sample size less than 800 (10% of the total sample size), and percent of parameters with a potential scale reduction factor of more than 1.1.

2.1.2 Model Selection

We explored two competing specifications for the multilevel model. One featured a less complex hierarchical parameter structure that does not include the term $\alpha_{d,c}^{\text{day} \times \text{ctry}}$. This simplified parameter structure implies that the daily regional mean is not shrunk toward the daily country mean but only toward the sum of the regional mean and the daily mean. Another more complex specification included replacing the term $\alpha_{a,c,w}^{\text{age-sex} \times \text{cntry} \times \text{week}}$ with $\alpha_{a,c,d}^{\text{age-sex} \times \text{cntry} \times \text{day}}$ which allows for daily shifts in the country-specific age-sex distribution. This additional flexibility could accommodate potential differences, for example, in the age-sex distribution among participants at weekends and weekdays.

We compared all three specifications using Bayesian leave-one-out cross-validation as described in [Vehtari et al. \(2017\)](#) and implemented in the R package `loo`. Table S.2 lists the expected log pointwise predictive density (ELPD) for each model, the difference between the two largest ELPD and the associated standard error of this difference. The ELPD is a measure of a model’s out-of-sample predictive performance. Models that have a better out-of-sample predictive performance (a higher ELPD) are generally preferred. The ELPD is estimated based on Pareto smoothed importance sampling. All estimation diagnostics suggest that the reported ELPD estimates are reliable.

Table S.2 demonstrates that for 5 out of 7 survey outcomes, our preferred specification has the highest ELPD. Further, it shows that the difference in ELPD between our preferred specification and the alternative with highest ELPD for these three outcomes is always less than two standard errors. These results give us confidence in our specification selection.

2.1.3 Poststratification

Our poststratification estimate measures the average survey outcome in region r on survey day d , and is denoted by $\theta_{d,r}$. We construct it by summing over the posterior predictive values for each age-sex group in each region and on each survey day ($\mu_{d,r,a}^{\text{pred}}$), weighted by the relative size of the

Outcome	ELPD				Diff.	SE(Diff.)
	M1	M2	M3	max		
Political System Scale	-36430.79	-36465.65	-36436.28	M1	-5.48	3.77
Authoritarian Attitudes Scale	-36121.82	-36124.59	-36129.88	M1	-2.76	3.49
Relative Outgroup Hostility	-123146.62	-123146.75	-123145.79	M3	-0.84	0.48
Democratic Satisfaction Scale	-33901.11	-33900.89	-33902.31	M2	-0.22	3.12
Government Satisfaction Scale	-33033.99	-33035.86	-33050.65	M1	-1.87	3.38
Trust Scale	-37205.69	-37210.25	-37209.93	M1	-4.24	2.75
Restricted Movement	-15891.76	-15897.41	-16005.07	M1	-5.65	5.11

Table S.2: Estimated expected log pointwise predictive density (ELPD) for all survey outcomes and three model specifications: The authors’ preferred specification (M1), a less complex specification (M2) and a more complex specification (M3). The column “max” reports which specification has the highest ELPD, column “Diff.” reports the estimated difference between the two largest ELPD values and “SE(Diff.)” the associated standard error for this difference.

the age group in a particular region:

$$\theta_{d,r} = \sum_a \mu_{d,r,a}^{pred} \frac{N_{r,a}}{N_r}.$$

We construct 1,000 poststratification datasets by drawing parameters tuples from the posterior predictive density.

The age-sex distribution for each region ($N_{r,a}$) comes from Eurostat and US Census Bureau. More specifically, we construct a population frame by age group, sex and sub-national region that includes all adults (18+ years) living in France (excluding overseas territories and Corsica), the UK (excluding Northern Ireland and British overseas territories), Switzerland, Germany and the US (excluding Alaska and Hawaii).

2.2 Difference-in-Differences

To estimate the effect of the introducing and reversing restrictions, we combine our poststratification estimates with the *Lockdown Index* (LI). Our baseline specification, use for the estimates in Figure 2, is a two-way fixed-effect estimator:

$$\theta_{d,r} = \gamma_0 + \gamma_d + \gamma_r + \delta \text{LI}_{d,r} + \varepsilon_{d,r}$$

where γ_d and γ_r denote the day and regional fixed-effect. We estimate the model on each of the 1,000 poststratification datasets constructed above. We summarize the posterior distribution of the δ using the 95% quantile and the mean.

In additional specifications we study of the effect of introducing and reversing severe restrictions ($\text{LI} > 3$). These specifications mimic a DID design (Angrist and Pischke, 2009) with lead

and lags for the treatment to study the evolution of the treatment effect. Our specification for introducing severe restrictions takes the following form:

$$\theta_{d,r} = \gamma_0 + \gamma_d + \gamma_r + \sum_{m=-6}^6 \beta_m D(m)_{d,r} + \varepsilon_{d,r},$$

where m is the day before/after a policy change and $D(m)_{d,r}$ it is the treatment indicator that takes on the value one on day m for those regions that introduce restrictions (except for $D(6)_{d,r}$ which takes the value one on day 6 and every day thereafter). In this way, the coefficients on the indicators $D(1)_{d,r} - D(6)_{d,r}$ measure the effect of restrictions on day one to six (and thereafter) while the coefficients on the remaining indicators ($D(-6)_{d,r} - D(0)_{d,r}$) are placebo checks to test if there are non-parallel pre-treatment trends. The specification for the reversal of restrictions is the same except that $D(m)_{d,r}$ encodes a reversal of a restriction.

We estimate these DID specifications on each of the balanced panel datasets constructed above, but we subset to regions that introduced or reversed restrictions during the study period (switcher regions). We remove all regions that either always had restrictions or that never introduced them. When estimating the effect of a reversal, we also remove all days prior to the introduction of the restriction, so as to not violate the parallel trends assumption by construction.

3 Additional Results

Table S.1 gives the estimates from our baseline specifications as displayed in Figure 2 in the main text. Figure S.2 replicates this baseline specification using the binary version of *Lockdown Index* (i.e., the version of the index used in Figure 3 of the main text). Figures S.3 and S.4 show the estimates separately for the introduction of restrictions and their reversals again using the binary version of *Lockdown Index* and on the subsets we use for Figure 3 of the main text. Figure S.1 replicates Figure 3 from the main text for all other outcomes. Figure S.10 uses weekly instead of daily lag indicators.

Figure S.5 shows the estimates from an extended baseline specification in which we add the number of COVID-19 related deaths at the regional level to the model (Figure S.6 adds the square root of the death counts as well). This helps us ensure that our comparisons focus on respondents facing similar conditions but with divergent policy responses. However, the spread of COVID-19 may be influenced by local policies—making it a post-treatment variable that could introduce bias. For this reason, we prefer to fit models both with and without this variable.

We extend our specification to include more flexible forms of the fixed effects for days in Figure S.7. Instead of including a single time fixed effect for all regions, these specifications estimate a

separate time fixed effect for groups of regions with similar values on a selected covariate. We do this for three covariates: log(population density), support for the incumbent in the most recent national election, and total deaths from COVID-19 (per capita) at the start of our survey period. For the first two, we create four groups, one for all the regions in the bottom quartile of their within-country distribution of the given covariate, another for the second quartile, and so on for each of the four quartiles. We cannot use this approach for the COVID-19 spread measure, as a large number of regions had zero deaths at the start of our survey period. Instead, we create two groups, one for regions with more COVID-19 deaths than the within-country median and another for regions with fewer deaths than the within-country median. This approach weakens our identification assumption, as it means we only need to assume parallel trends within the groups defined by the covariate to identify the effect of lockdowns. As Figure S.7 shows, this alternative time-fixed effects approach makes little difference in our estimates.

We next report the results of an extended baseline specification in which we add the running daily count of economic support packages announced by the government and interact this measure with the binary version of the *Lockdown Index*. Figure S.8 displays marginal effect estimates from these models, i.e., the effect of lockdown policies conditional on a given amount of economic support packages. This simple specification allows us to explore if there is an indication that economic support package moderate the effect of the *Lockdown Index*. That does not seem to be the case. For only one of our measures—the *Government satisfaction scale*—is the interaction statistically reliable. For *Political system scale* and *Authoritarian attitudes scale*, the two measures where we find reliable effects in our baseline specification, the effects of lockdowns vary little across amounts of economic support.

In Figure S.9 we report the estimates from a specification in which we interact the binary version of the *Lockdown Index* with an indicator measuring if a region was above/below the median outcome on March 19/20. We find little evidence for substantial heterogeneity. If anything, the estimates suggest that regions with a higher prevalence of authoritarian attitudes are, on average, more affected by lockdowns.

Figure S.11 shows the estimates when we fit our baseline specification for each survey item separately rather than combining them into a single score. To obtain these estimates we proceeded in the same fashion as with our survey outcomes: We estimated a series of Bayesian multilevel models, then post-stratified the survey responses and finally fitted the DID models. The only difference is that we used a standard ordinal regression model (also sometimes called a cumulative model) instead of a linear regression model to predict the survey responses. Figure S.11 demonstrates that the effect on our *Authoritarian attitudes scale* is primarily driven by respondents being more likely to agree with the statements that it is important to follow leaders and to crackdown on troublemakers. The effect on the *Political system scale* seems to be mostly driven by an increase

in support for military rule—potentially pointing to respondents’ desire for swift actions to combat the pandemic—although there are also small (but unreliable) negative effects on support for democracy and small (unreliable) positive effects on support for autocracy.

Figure S.12 replicates the pattern of our baseline estimates using an alternative estimation methodology. For reasons described above, we prefer to first post-stratify the data via Bayesian multilevel modeling, but we here show the results of estimating the DID models directly on the “raw” survey data. We use the same baseline specification as in section 2.2 but also show models with age, gender, and the interaction between the two as covariates. To account for potentially correlated errors within regions, we cluster the standard errors of the coefficients by region. To account for missing values in the outcomes, we imputed the data via multiple imputation using the R package *mice* and pool the regression estimates according to Rubin’s rules (Rubin, 1987).

	(1) Political system scale	(2) Democratic satisfaction scale	(3) Government satisfaction scale	(4) Authoritarian attitudes scale	(5) Trust scale	(6) Relative outgroup hostility	(7) Restricted movement
Lockdown Index	0.011 (0.006)	0.004 (0.004)	0.007 (0.005)	0.017 (0.005)	0.002 (0.004)	-0.008 (0.054)	0.036 (0.004)
Region FE	✓	✓	✓	✓	✓	✓	✓
Day FE	✓	✓	✓	✓	✓	✓	✓
<i>N</i>	7,254	7,254	7,254	7,254	7,254	7,254	7,254

Table S.1: Bayesian posterior estimates and standard deviation for the effect of government restrictions on citizens' attitudes and preferences. Estimates are based on linear regression models with region and day fixed-effects fitted on post-stratified survey estimates.

Figure S.1: (Figure 3: Additional outcomes) Bayesian posterior estimates (dots) and 95% credible estimates for the effect of introducing and removing restrictions. Estimates are based on linear regression model with region and day fixed-effects fitted on post-stratified survey estimates for those regions with policy variance.

Figure S.2: (Figure 2: Binary restriction measure) Bayesian posterior estimates (dots) and 95% credible estimates for the effect of restrictions on citizens' attitudes and preferences.

Figure S.3: (Figure 2: Binary measure of restriction introductions). Bayesian posterior estimates (dots) and 95% credible estimates for the effect of restrictions on citizens' attitudes and preferences.

Figure S.4: (Figure 2: Binary measure of restriction reversals). Bayesian posterior estimates (dots) and 95% credible estimates for the effect of restrictions on citizens' attitudes and preferences.

Figure S.5: (Figure 2: Adjusting for COVID-19 death counts) Bayesian posterior estimates (dots) and 95% credible estimates for the effect of restrictions on citizens' attitudes and preferences.

Figure S.6: (Figure 2: Adjusting for COVID-19 death counts and counts²) Bayesian posterior estimates (dots) and 95% credible estimates for the effect of restrictions on citizens' attitudes and preferences.

Figure S.7: Figure 2: Separate time-FE for groups based on select covariates Bayesian posterior estimates (dots) and 95% credible estimates for the effect of restrictions on citizens' attitudes and preferences.

Figure S.8: (Figure 2: Interacting lockdown policies (binary measure) with economic support package counts) Bayesian posterior estimates (dots) and 95% credible estimates for the marginal effect of restrictions on citizens' attitudes and preferences conditional on the amount of economic support packages.

Figure S.9: (Figure 2: Heterogeneity by high/low baseline). Bayesian posterior estimates (dots) and 95% credible estimates for the effect of restrictions on citizens' attitudes and preferences.

Figure S.10: (Figure 3: Weekly lag indicators). Bayesian posterior estimates (dots) and 95% credible estimates for the effect of restrictions on citizens' attitudes and preferences.

Figure S.11: (Figure 2: Score components). Bayesian posterior estimates (dots) and 95% credible estimates for the effect of restrictions on citizens' attitudes and preferences.

Figure S.12: Estimates from two-way fixed-effects regressions (dots) and 95% confidence intervals (after pooling across $m = 10$ multiply imputed datasets) for the effects of restrictions on citizens' attitudes and preferences. The unit of analysis is the survey respondent. $N = 25,778$.

References

- Angrist, J. D. and J.-S. Pischke (2009). *Mostly Harmless Econometrics. An Empiricist's Companion*. Princeton: Princeton University Press.
- Bansak, K., J. Hainmueller, and D. Hangartner (2016). How economic, humanitarian, and religious concerns shape european attitudes toward asylum seekers. *Science* 354(6309), 217–222.
- Berinsky, A. J., G. A. Huber, and G. S. Lenz (2012). Evaluating online labor markets for experimental research: Amazon. com's mechanical turk. *Political analysis* 20(3), 351–368.
- Betancourt, M. (2017). A conceptual introduction to hamiltonian monte carlo. *arXiv preprint arXiv:1701.02434*.
- Cheng, C., J. Barceló, A. S. Hartnett, R. Kubinec, and L. Messerschmidt (2020). Covid-19 government response event dataset (corononet v. 1.0). *Nature human behaviour* 4(7), 756–768.
- Gelman, A., J. B. Carlin, H. S. Stern, D. B. D. Dunson, A. Vehtari, and D. B. Rubin (2014). *Bayesian Data Analysis* (3rd ed.). Boca Raton: Chapman & Hall/CRC.
- Gelman, A. and J. Hill (2006). *Data Analysis using Regression and Multilevel/Hierarchical Models*. Cambridge: Cambridge University Press.
- Gelman, A. and T. C. Little (1997). Postratification into many categories using hierarchical logistic regression. *Survey Methodology* 23(2), 127–135.
- Goodrich, B., J. Gabry, I. Ali, and S. Brilleman (2020). rstanarm: Bayesian applied regression modeling via Stan. R package version 2.21.1.
- Hale, T., T. Atav, L. Hallas, B. Kira, T. Phillips, A. Petherick, and A. Pott (2020). Variation in us states' responses to covid-19. Technical report, Blavatnik School of Government.
- Hale, T., S. Webster, A. Petherick, T. Phillips, and B. Kira (2020). Oxford covid-19 government response tracker. Technical report, Blavatnik School of Government.
- Hoffman, M. D. and A. Gelman (2014). The no-u-turn sampler: adaptively setting path lengths in hamiltonian monte carlo. *J. Mach. Learn. Res.* 15(1), 1593–1623.
- Park, D. K., A. Gelman, and J. Bafumi (2004). Bayesian multilevel estimation with poststratification: State-level estimates from national polls. *Political Analysis* 12(4), 375–385.
- Rubin, D. B. (1987). *Multiple Imputation for Nonresponse in Surveys*. New York: J. Wiley & Sons.

Vehtari, A., A. Gelman, and J. Gabry (2017). Practical bayesian model evaluation using leave-one-out cross-validation and waic. *Statistics and Computing* 27(5), 1413–1432.