Cross-National Measures of the Intensity of COVID-19 Public Health Policies

ROBERT KUBINEC^{*1}, JOAN BARCELÓ¹, RAFAEL GOLDSZMIDT⁵, VANJA GRUJIC²,
TIMOTHY MODEL³, CARESS SCHENK⁴, CINDY CHENG⁶, THOMAS HALE⁷, ALLISON
SPENCER HARTNETT⁸, LUCA MESSERSCHMIDT⁶, ANNA PETHERICK⁷, AND SVANHILDUR
THORVALDSDOTTIR⁶

¹*Division of Social Sciences, New York University Abu Dhabi, Abu Dhabi, United Arab Emirates*

²*Faculty of Law, University of Brasilia, Brasília, Brazil*

³*Fors Marsh Group, Arlington, VA, USA*

⁴*School of Sciences and Humanities, Political Science and International Relations, Nazarbayev University, Astana, Kazakhstan*

⁵*Brazilian School of Public and Business Administration, Fundação Getulio Vargas, Rio de Janeiro, Brazil*

⁶*Hochschule für Politik at the Technical University of Munich (TUM) and the TUM School of Governance, Munich, Germany*

⁷*Blavatnik School of Government, Oxford University, Oxford, United Kingdom*

⁸*Department of Political Science and International Relations, University of Southern California, Los Angeles, CA, USA*

November 7, 2021

Abstract

We present six new indices that measure the intensity of government responses to COVID-19 within distinct policy domains: social distancing, schools, businesses, health monitoring, health resources and mask wearing. We create these measures by combining two of the most comprehensive COVID-19 datasets, the CoronaNet COVID-19 Government Response Event Dataset and the Oxford COVID-19 Government Response Tracker, using a Bayesian time-varying measurement model. Our daily indices track policymakers' policy intensity to each of these policy domains from 1 January, 2020 to 14 January, 2021, for over 180 countries. In creating these measures, we are able to provide aggregate summaries of policy-making activities which captures the intensity of a country's government response to COVID-19 in a given policy domain. Using these measures, we are able to clarify findings on the role of lockdowns in preventing infections in the early pandemic period. In terms of policy adoption, we show that more democratic countries and countries with greater levels of economic interdependence tend to have less intense policy responses to the pandemic.

^{*}Corresponding author: rmk7@nyu.edu

Since the outbreak of the COVID-19 pandemic, research groups all around the world have invested considerable work to collect data on government policies aimed at preventing transmission of the SARS CoV-2 virus. However, the scale of the pandemic and the diversity of government responses makes it difficult to aggregate the increasingly complex available datasets into higher level measures which can be more accurate summaries of policy-making activity. Creating aggregated measures of the intensity of police responses to distinct policy domains is necessary to be able to answer questions about the effect of disparate policies, not only in terms of combating COVID-19 infections,¹⁻⁴ but also on a diverse array of socio-economic outcomes, including mental health,⁵⁻⁸ economic hardship,⁹⁻¹⁵ political partisanship,¹⁶⁻²³ and the reasons why some countries were better positioned to implement restrictions to control the pandemic.²⁴⁻²⁷

Despite their differences, existing pandemic policy tracking projects²⁸⁻³⁸ seek to measure de jure policies, recording, specifically, on day t in country c , the set of policies governments put in place to combat COVID-19. Each of these de jure policies represents a discrete commitment by a policymaker within a given policy domain, such as encouraging mask wearing or promoting social distancing. Our objective in studying COVID-19 restrictions is, ultimately, to know what policies should be prioritized to stop the COVID-19 pandemic as well as future ones. That is, we want to know whether the devotion of effort and resources by governments to a given policy domain results in a strong suppressive effect on the spread of the virus. Each de jure policy can be understood as an increase in the strictness or breadth of a country’s policy within a given policy domain, though of course some policies represent much greater intensity than others. To be able to understand how government policies compare, we need an aggregate measure of the overall level of intensity across indicators. We employ the term policy intensity to describe the level of investment in a policy domain; hence, our term of policy intensity is conceptually similar to that of policy commitment adopted in other applications.³⁹

We introduce here model-based aggregated measures built on the theoretical prior that policymakers face a cost-benefit decision problem regarding COVID-19 suppression with a fixed budget constraint^{40,41} (see Methods section for details). Policymakers can choose how much of their resources to dedicate to COVID-19 suppression policies, such as general social distancing, masks, or business restrictions, by trading these off against non-pandemic policy goals. The relative costs and benefits of allocation decisions are broadly related to the potential financial, political or administrative ramifications of policies.

As a result, we define policy intensity as the level of policy investment in a given policy domain. As such, it is the result two parameters: (i) policy stringency: the level of policy strictness within a given policy domain (e.g., a country that implements a policy that bans mass gatherings of more than 10 people is stricter than a country that implements a policy that bans mass gatherings of more than 1,000 people within the policy domain of social distancing); and, (ii) policy breadth: the breadth of governmental interventions

within a given policy domain (e.g., a country that implements a policy that bans mass gatherings of more than 10 people and a dawn-to-dusk curfew has more policy breadth than a country that implements a policy that bans mass gatherings of more than 10 people alone). Therefore, each additional policy, policy dismantling, policy strengthening, or policy relaxation that a policymaker imposes in a given domain will increase or decrease policy intensity, though by varying amounts depending on the importance and cost of that policy. Obtaining measures of the overall level of policy intensity would allow us to more accurately answer more specific research questions such as the following about mask policies:

Does devoting more policymaker time, resources and effort to increase mask-wearing reduce the incidence of COVID-19?

We note that this is a different question from what we can ask with *de jure* indicators, namely,

Which types of mask policies are the most effective at reducing the incidence of COVID-19?

To help answer research questions requiring aggregate measures of policy activity, we have chosen to create indices for six different policy domains which are central to COVID-19 public health policy and are actively employed in research: 1) general social distancing,⁴² 2) masks policies,^{43,44} 3) business restrictions,⁴⁵ 4) schools restrictions,⁴⁶ 5) health monitoring,⁴⁷ and 6) health resources.⁴⁸ We elaborate the assumptions underlying the definition of these policy categories, and their respective costs, under the subheading 'Conceptualization of Indices' in the Results section.

Instead of relying on a single existing dataset to estimate policy intensity, we pool indicators from two widely-used data sources, the CoronaNet Research Project⁴⁹ and the Oxford COVID-19 Government Response Tracker (OxCGRT).⁵⁰ To produce estimates that combine information from both sources, we employ time-varying statistical measurement models that allow us to use 164 *de jure* policy indicators to derive our six time-varying policy intensity scores for over 180 countries from 1 January, 2020 to 14 January, 2021 along with estimates of potential measurement error in the scores.^{49,51}

With these new measures of policy intensity, we explore how public health, economic, political and social factors may affect policy adoption. Our results, which are also robust to estimates of measurement uncertainty in the indices, suggest that the intensity of social distancing and business restrictions are the most strongly associated with reducing mobility in workplaces and retail establishments. However, only social distancing policy intensity is also associated with reducing personal contact, i.e., personal contact with people outside the home appears to rise when business restrictions and school restrictions become more intense.

We also find that the quality and nature of state institutions are strongly related to the level of policy intensity. Democracies and countries with a greater respect for human rights are less likely to increase

the intensity of business restrictions and social distancing policies. In addition, we show that economic interdependence, as measured by the proportion of trade and foreign direct investment (FDI), is strongly associated with reduced business, school, and general social distancing policy intensity.

To show the importance of these indices for understanding COVID-19 suppression, we replicate and extend an influential study by Flaxman et al.⁴² of lockdowns in the early pandemic period in Europe. Their analysis employed dichotomous measures of policies adopted by countries, including what they describe as lockdowns, public events restrictions, school closures, self-isolation policies and public encouragements of social distancing. We resolve a limitation in this study in that their COVID-19 measures are binary indicators and highly correlated, making it difficult to obtain separate estimates for different policy domains.^{52,53} Using our measures of policy intensity, which provide a much more granular assessment of policymaker activity, we find that the intensity of general social distancing appears to have a more limited role in reducing the spread of COVID-19 compared to the intensity of business restrictions (including restaurants and bar closures). We also note that there is some evidence that mask policy intensity, which was not included in the original analysis, is associated with reduced COVID-19 spread in Europe despite having only been implemented towards the end of the first wave.

Data Availability

For the data to reproduce all results, we refer the reader to our Github repository: https://github.com/saudiwin/corona_index. We refer the reader as well to a blog post describing the data and how to incorporate measurement uncertainty into analyses using the R software package, which can be accessed from http://www.robertkubinec.com/post/err_in_vars/.

To download the underlying data and estimated indexes directly, the following two links contain CSV files:

1. Estimated indices in day-country-index-index type form: https://drive.google.com/uc?export=download&id=1dMCTVPrf-tJyhv_uxr0yAQO-Elx0QOCG
2. Underlying policy data indicators from CoronaNet and OxCGRT: <https://drive.google.com/uc?id=1lorcowHnF0Vl6pxBjMdjTC4yPhHBLJI&export=download>

We refer the reader to our supplementary information for a full code book describing each of the underlying policy indicators.

Code Availability

All of the estimations were performed in the R software package, and R scripts to reproduce all results are available via a public Github repository: https://github.com/saudiwin/corona_index.

Background

We describe in this section assumptions underpinning the interpretation we assign to the latent variable we estimate, which necessarily combines heterogeneous policies from both the CoronaNet and OxCGRT trackers. For example, the policy evaluation literature has so far focused on lockdowns as monolithic policies.^{1, 4, 42, 43, 54–57} However, lockdowns can also be understood a bundle of different policy options that happen concurrently, such as curfews, restrictions on mass gatherings, and stay-at-home orders, to name only a few.⁵⁸ While the former conceptualization implicitly assumes that the manner in which a lockdown is implemented does not matter, only the outcome of limited mobility does, the latter conceptualization implicitly assumes that both the outcome of limited mobility and the mechanism through which it is delivered (e.g. via curfew or restrictions of mass gatherings) are important for a given research question.

In developing these indices, we imagined different policy domain as representing different paradigms of how the virus is spread. Our model assumes that policymakers are rational actors who are trading off between costs and benefits of devoting resources to a given policy domain, where each policy represents an additional investment in the policy domain (see Methods section in the supplementary information for details). We define more clearly how we delimited the index policy domain below:

- General social distancing: captures restrictions on individual mobility. These can include lockdowns, curfews, travel bans and other policies which restrict mobility. Under this paradigm, governments seek to restrict the spread of the virus by conceptualising people as the vector of the disease; the less people move, in theory, the less the disease can spread. The direct cost of this policy is the physical autonomy of the people under a government’s jurisdiction, which may induce political or mental-well being costs, among others.
- Masks: captures the wearing of face-masks. This can include wearing masks outdoors or indoors as well as the type of mask that must be worn. In this case, government conceptualizes the aerosols individuals produce as the vector of the disease and seek to restrict the spread of the virus by containing the spread of these aerosols. The direct cost of this policy is an individual’s bodily autonomy which may induce induce various political, social or mental costs, among others.

- **Business restrictions:** captures restrictions on private commerce and industry. This can include policies that close all businesses except for grocery stores, policies which allow for takeaway in restaurants and the like. Governments using this paradigm conceptualize the economic activities and corresponding physical circulation of people needed to sustain people’s livelihoods as a potential vector of a disease. The direct cost of these policies is mainly economic, including risks of unemployment and recessions.
- **Schools restrictions:** captures restriction on education provision. These can include the closure of secondary schools, transitions to online teaching, or postponement of school exams and the like. Educational activities are conceptualized by governments as being a potential vector for disease under this paradigm insofar as the provision of normal education requires a high degree of circulation between different groups of people (e.g. children interacting with teachers; children interacting with caregivers; teachers interacting with children). Direct costs include a child’s educational and social development and which may also have knock-on effects in terms of childcare costs, effects on caregiver work performance and the like.
- **Health Monitoring:** captures monitoring of individual health statuses. These can include measures like contact tracing, temperature checks at stores or health questionnaires. When implementing health monitoring policies, governments seek to restrict the spread of the virus by being able to accurately and swiftly track who has been affected by the disease. Correspondingly, the implication is that the government then only limits the mobility of the affected people and redirects their behaviour to limit the spread of the virus (e.g. individuals with high temperatures are not permitted to enter stores). These policies have direct costs to individual privacy which may induce costs on mental well-being as well as political costs.
- **Health Resources:** captures medical resources employed to treat the disease. Examples may include increased funding for personal protective gear, postponement of elective surgeries or the availability of tele-health medicine. Under this paradigm, governments respond to the pandemic by adjusting their ability to medically treat the actual disease itself in the form of hospitals, staff, and materials. Costs of these policies may include the direct financial cost of funding this policies as well as the health of non-COVID patients who experience delays in accessing medical resources.

A full description of the underlying policy indicators from both trackers is available in the supplementary information.

Methods

Employing statistical measurement models from different data sources provides at least two unique benefits. First, it allows us to create indices which contains more information than either data source could provide on its own. The value of combining the CoronaNet and OxCGRt datasets is particularly high given that they rely on disparate taxonomies which otherwise could not be easily combined. Second, and most crucially, by employing measurement methods, our indices are able to move beyond de jure policy indicators to a measure of the unobserved effort and resources that policymakers have invested in a given policy domain. That is, by designing our model to measure a single-dimensional latent trait of policy intensity per policy domain, we can model their causes and consequences at a more conceptual level and we gain uncertainty intervals that describe how certain we can be about the true level of policy intensity in a given country.

In creating the indices, we combine indicators from the CoronaNet taxonomy, which itself relies on 20 policy categories, with OxCGRt indicators that cover the same aspect of a policy category. To summarize the composition of our indices, see Table 5 in the supplementary information. It shows how each of the six indices is composed through the policy classification in the CoronaNet and OxCGRt taxonomy by top-level and sub-level variables. We include 154 indicators from the CoronaNet dataset and 10 from OxCGRt. We use more indicators from CoronaNet because of differences in how these two projects code public health policies, which we further explore in the supplemental information. In general, the CoronaNet project takes the approach of separately coding all aspects of COVID-19 related policies, while the OxCGRt produces ordinal scores for each policy type.

One complication is that the CoronaNet dataset contains provincial and even municipal information about policies for all countries in the dataset, though coverage varies. As such, we needed a way to include this information to produce national-level estimates. We employed a simple algorithm in which we assigned a value of +1 for each policy in force at the national level and 0 otherwise. For over-lapping policy coverage at lower levels of government we then add to this baseline score. For each province and city with the same policy in force, we add the proportion of the country’s population in the province or city to this baseline score. The resulting score could reach a maximum of 3 if a binary-coded policy was implemented at the national level, among all policies and among all cities in a country. Of course, we only rarely observe this maximum level of policy enforcement in our indices, but we believe the additive nature of combining lower levels of policy enforcement helps communicate the multiple levels of policy while still maintaining parsimony in how we aggregate the information. The resulting score is then a continuous variable and is modeled as such.

In total we have 189 countries in both datasets with coverage from January 1st, 2020 to January 14th,

2021, resulting in 76,380 country-day observations with all of the 164 pooled indicators. As policy coding remains an ongoing task, we intend to release further periodic updates to our indices as more data is available. A complete codebook of all of the indicators is available in the supplemental information.

Combining this data into relevant latent constructs is a difficult challenge given the scope of the data and the diversity in variable types, including binary, continuous, ordinal variables, and the need to incorporate over-time variation. To accomplish this, we employ a Bayesian version of item-response models known as ideal point models^{49,51} that are estimated with the Hamiltonian Markov Chain Monte Carlo (MCMC) sampler Stan.⁵⁹ The ideal point approach can be understood as a simple formal model in which an actor i must decide between two policy alternatives j and j' .⁶⁰ If we assume that there is a one-dimensional utility score $U(\cdot)$ assigned by actor i to j and j' , then actor i will always choose j over j' if $U(j) > U(j')$. In our context, this implies that countries will choose those policies j which have the smallest distance to their ideal point score. As we mentioned, we can then interpret the ideal point score as a measure of the relative level of commitment or intensity in the policy domain, while the position of the policy item corresponds to its relative cost and presumed efficacy to the policy actor. Only those actors who commit a lot of resources are able to implement high-cost policies.

We note that this approach differs from exploratory methods such as principal component analysis and factor analysis that are commonly used for dimension reduction. In comparison to these methods, we begin with a formal model of human behavior and impose more structure into the estimation. As a result, we can assign a more meaningful interpretation to the latent construct than simply the eigenvector which best spans the multidimensional space of the predictors. As we show later, this allows us to focus the model on higher-cost policies to be able to discriminate between high and low-cost policies in the indices.

If we index countries α by i and our policy factor scores δ by j , we can estimate country ideal point scores for each day t using the following parametric equation for each indicator value Y_{ijt} :

$$Pr(Y_{ijt}) \sim g(\alpha_{it}\delta_j - \beta_j) \quad (1)$$

where the function $g(\cdot)$ represents a statistical distribution appropriate for the particular indicator j , such as an ordered logit for an ordered indicator, a logistic regression for a binary indicator, and a Gaussian distribution for a continuous indicator. In this formula, the country index scores are represented by the α_{it} and the policy discrimination (factor) scores by the δ_j . Conditional on the model's assumptions, we can then interpret our uncertainty in the index scores α_{it} as uncertainty derived from measurement error while still being able to estimate the most plausible value for α_{it} .

The sign of the policy discrimination scores δ_j reflects which way the policy loads on the latent scale,

i.e., whether it positively or negatively predicts higher or lower values of the indices. The absolute value of the discrimination score indicates how much information about the respective index is contained within this particular indicator. In other words, policies that have high discrimination are much more likely to be adopted by countries with high levels of intensity than low-intensity countries. We can for this reason infer that these policies must be both relatively costly for these countries to implement and strongly related to the core policy goal. Importantly, there is still uncertainty relating to how countries evaluate the costs and benefits of different de jure policies, which is why we necessarily have measurement error. In addition, we note that a policy could potentially be costly yet have low discrimination if it did not contribute to governments’ investments in a policy domain, such as if we inadvertently added policies in an unrelated domain like power generation.

We refer interested readers to prior work⁵¹ for a more thorough presentation of the joint Bayesian posterior that we estimate to obtain these parameter values. In essence, statistical models appropriate for different variable types are jointly estimated with shared parameters for the country index scores and the policy discriminations. We allow the country index scores to vary via a simple random-walk process in which each country’s score in time t is equal to its score in $t - 1$ plus Gaussian-distributed noise σ_i . This time process permits some smoothing of the resulting scores. We put a boundary-avoiding prior on the σ_i parameters, $\sigma_i \sim \text{InvGamma}(2, 5)$, to improve sampling as policy indicators for some indices do not change very often over time.

One important area we have to consider when employing this model is that the likelihood is not globally identified without further information added to the model. In particular, we have to prevent scale rotations of the indices, so we assign a very strong prior of **Normal**(1, 0.001) on the value of one policy discrimination for each index to ensure a single rotation. We choose policies that are relatively high-cost to assign unit discrimination to ensure that we estimate policy intensity as opposed to another latent dimension (we refer the reader to the Results section for the discrimination parameter values). However, we let all the other policy discrimination parameters float so that we can see if they all map on to the same latent trait, which is a strong test of the validity of the underlying indices.

To fit the models, we run four independent MCMC chains with Stan for 250 warmup iterations and 300 sampling iterations. To test convergence, we report split- $\hat{R}$ diagnostic plots in the supplementary information for each index model, which show clear convergence of the chains given these sampling parameters.

To convert the latent scores to the same scale, we first standardize each score and then employ the inverse logit function to convert the score to a bounded number between 0 and 1. We then multiply by 100 to produce scales with a maximum and minimum observed range of 0 and 100 respectively.

We note that there are three potential ways that the measurement error, or uncertainty in the posterior

estimates, of our indices can be incorporated in applied analyses. These methods all involve treating the indices as Normally-distributed random variables rather than observed values, where the standard deviation of the Normal distribution for each observation is taken from the standard deviation of the indices' posterior draws. We briefly discuss three methods.

The simplest approach is known as errors-in-variables models and have existed for decades.⁶¹ In these models, predictors can be included with a covariance matrix specifying Normally-distributed measurement error. This covariance matrix then inflates the standard errors in ordinary least squares (OLS) regressions. The covariance matrix, however, only specifies a single value (standard deviation) of measurement error for each index variable.

By contrast, two approaches exist that allow measurement error to be included at the observation level, which is potentially important as uncertainty can increase or decrease over time as our previous figures show. The first and most direct method is to include the indices in a Bayesian regression model as Normally-distributed random variables with a mean specified by the index value and a standard deviation from the estimated posterior deviation.⁶² This estimation is quite powerful in that it completely replaces the predictors with random variables with uncertainty that can vary over time.

An analogue to this approach can be achieved via frequentist methods via bootstrapping.⁶³ For each draw of the posterior values of the indices, a model can be fit with the indices as data. The estimates can then be calculated as the average coefficient estimate over these draws, and the confidence interval as the 5% to 95% empirical interval of these estimates.

In the analyses in which we model the adoption of social distancing, business and school policy intensity, we employ Bayesian specifications to directly incorporate measurement error at the observation level (i.e., including the estimated posterior standard deviation for each country-day index estimate). To look at associations between time-varying and cross-sectional predictors and these indices, we predict the posterior median values of each index assuming Gaussian measurement error with estimated standard deviation $\hat{\sigma}_{it}$ from the posterior standard deviation of each index estimate α_{it} for country i and day t . To do so, we employ the R package `brms`⁶⁴ with default weakly informative priors.

To incorporate the fact that policymakers are simultaneously implementing the three different types of policy indices we predict, we allow for the errors to be correlated across specifications. To do so, we estimate a multivariate Normal distribution for a set of countries I , days T and predictors P ,

$$A \sim \mathbf{MVN}(M\beta, \Sigma) \quad (2)$$

where A is a matrix with IT rows and 3 columns containing the $\hat{\alpha}_{it}$ index estimates, M is a array of

data of dimension $ITxPx3$, β is a $3xP$ matrix of coefficients and Σ is an estimated IxI covariance matrix of unmodeled (residual) correlation between the indices. This correlation matrix can help provide an indication of strategic behavior when policymakers change the policy intensity level, such as whether the different policy types are substitutes (inversely correlated) or complements (positively correlated).

Because there is missing data in the predictor variables, we create five imputed datasets non-parametrically with random forests.^{65,66} We estimate independent Markov chains for each imputed dataset and then combine the resulting posterior estimates to increase the variance of the posterior distribution to reasonable levels.⁶⁷

In our replication and extension of Flaxman et al.,⁴² we employed the same Bayesian susceptible-infected-recovered (SIR) model specification both with and without measurement error in our indices. To do so, we modified their code to allow us to include predictors of R_t , the parameter in their model which measures the reproduction rate of SARS-CoV-2, as Gaussian-distributed variables with pre-specified standard deviations for our model with measurement error. For both models with and without measurement error, we constrain the predictor coefficients to be positive to achieve identification, as the authors did in their original paper. The main difference with the original specification is that we did not use varying effects for the lockdown policy category given that lockdowns do not feature as prominently in our results.

Results

In what follows, we present our approach to calculating the indices and explain how they relate to underlying indicators, as well as discuss the validity of the indices and the underlying data. We further discuss our results of models which predict the level of policy intensity across countries and our re-analysis of an influential study of the effect of public health policies at preventing the disease.

Description of Indices

We report the full distributions of each index for all countries in Figure 1. As the number of countries in the indices are too many to be able to distinguish them in a single plot, we select ten countries from diverse areas of the world and show indices scores for each in Figure 2. As each of these figures shows, there are broad similarities across the indices with policy intensity increasing in the early months of the pandemic, followed by more country-specific trends as the pandemic ebbed and flowed within borders. In general, the indices with the highest amount of over-time variation include our school index, business index, and general social-distancing index. The latter is particularly bolstered by the fact that OxCGRT has several indicators for this index, which increases significantly the amount of information about general social distancing policies.

Plot shows the median posterior estimates for country policy intensity scores with the 5% to 95% posterior uncertainty intervals in gray.

Figure 1: Policy Intensity Index Scores for All Countries

To better understand the underlying components of the indices, Figures 3 and 4 show the distribution of discrimination parameters for each de jure policy indicator in each index separately. These parameters represent the degree and direction of association between the observed policies and our latent policy intensity construct. As can be seen, the discrimination parameters almost exclusively line up on the right hand-side of the plot, with some parameters estimated at close to zero. This behavior indicates that the model only found policy indicators positively associated with the latent variable. In other words, there is reason to believe that our policy indicators are all measuring a process of varying degrees of policy intensity rather than quite distinct policy domains that have no real relationship to each other. As such, it represents an important validity test for the approach to creating our indices. Furthermore, we find that the OxCGRT indicators are always in the same direction as the CoronaNet indicators, which shows that the indicators are tracking the same latent dimension despite different coders and coding schema.

The discrimination parameters that are close to zero indicate policies that contribute little information about the overall intensity of that policy domain. As we used relatively high-cost policies to identify the model (see Methods section in the supplementary information), policies with low discrimination are likely either policies that are relatively easy to implement or are ones that are substantively unrelated to the larger policy domain.

Plot shows the median posterior estimates for country policy intensity scores with the 5% to 95% posterior uncertainty intervals in gray.

Figure 2: Policy Intensity Index Scores for 10 Countries

Plot shows the median posterior estimates for policy indicator discrimination parameters. The 5% to 95% posterior uncertainty intervals are also plotted but do not appear because they are too small to be visible.

Figure 3: Discrimination Parameters I

Plot shows the median posterior estimates for policy indicator discrimination parameters. The 5% to 95% posterior uncertainty intervals are also plotted but do not appear because they are too small to be visible.

Figure 4: Discrimination Parameters II

Understanding the Indices

In order to analyze the relation between our indices and their two component dataset sources, we calculated simple summary measures of the indicators for each dataset by summing over all the indicators for a given day and standardizing them. We evaluated correlations between the estimated indices and the aggregated raw data a) over time, considering country-mean centered values, and b) across countries, considering cross-sectional correlations for each day of 2020. The correlations over time between the indices and the aggregated raw data were, in general, of moderate size for CoronaNet and OxCGRT, respectively: Business (.52 and .58), Schools (.32 and .57), General social distancing (.54 and .91), Health Monitoring (.51 and -.01), Health Resources (.66 and .62) and Masks (.27 and .59). Figure 5 shows that the average correlations in the cross-sections are strongly reduced after the initial policy ramp-up period, and were on average lower than the correlations over time, particularly for the OxCGRT indicators: Business (.42 and .26), Schools (.35 and .18), Social Distancing (.40 and .79), Health Monitoring (.51 and .04), Health Resources (.61 and .20) and Masks (.27 and .26)—again, for CoronaNet and OxCGRT, respectively. The higher correlations observed in the earlier months of the pandemic may be associated with some type of herd behaviour in which, in the face of extreme uncertainty about pandemic evolution and the efficiency of different policies, most countries adopted a wide variety of measures in a short time span. After these initial isomorphic movements, with additional information about the pandemic progression, countries adjusted their strategies and policy mix, potentially giving more weight to the perceptions of costs and benefits of distinct policies in their specific national contexts.^{68–70}

While the limited magnitude of most correlations between the raw data of Coronanet and OxCGRT

Figure 5: Daily Cross-sectional Correlations between the Aggregated Raw Data and Estimated Indices

and the indices may be explained by the reduction in measurement error of the combined indices, it also suggests that these two sources provide unique and non-redundant information in terms of de-jure policies. The latter perspective is supported when we compare the associations of the underlying indicators and the indices with mobility measures. As an illustration, we estimate fixed-effects models with two different measures of mobility as the dependent variables (retail and recreation visits and workplace mobility) and two different policy dimensions, which are expected to be directly associated with mobility (social distancing and business). The magnitude of the R^2 varies from 1% to 44% across models, and it is not always the case that the R^2 is higher for the indices (see Supplementary Tables 1 and 2). As such, it is clear that the measurement model which we define in the Methods section is parsing the underlying data in a way that selects a latent trait which is distinct from the indicators themselves, as we would expect given the behavioral model we employ.

Model of Policy Adoption

Our list of covariates for which we look at associations with the estimated indices includes nine time-varying and ten cross-sectional predictors. Tables 3 and 4 in the supplementary information provide summary

statistics for the time-varying and cross-sectional predictors, respectively. In our set of time-varying predictors, we include counts of COVID-19 cases and deaths to account for how governments enact policies in response to rising case and death numbers, though we note that this is for the purposes of statistical adjustment, not to conclusively determine the effectiveness of the policies.⁶⁸ We also include aggregated response data from Facebook surveys (using the percent who answer yes) about personal contact, financial anxiety, and general anxiety to explore how behavior and attitudes factor into policymaking.⁷¹ Likewise, we include Google data on retail, workplace, grocery, and parks mobility to investigate how governments use COVID-19 policies in response to changes in mobility.¹⁰ Google mobility data are measured as the percent change in mobility from baseline, where the baseline was captured during the pre-pandemic weeks of 2020.

In line with existing research on correlates of COVID-19 policy diffusion, we also include a series of cross-sectional social, economic, and political predictors. These cross-sectional predictors include the following economic covariates: World Bank measures of foreign direct investment (FDI, net inflows in USD), trade (as a share of GDP), and GDP per capita. We also include a Gini index from the Standardized World Income Inequality Database, measured on a scale of 0 - 1 where values closer to 1 indicates greater income inequality. Our political covariates include the Varieties of Democracy measure of Bureaucratic Corruption, the Fund for Peace’s measure of State Fragility, the Global Health Security Index’s measure of Pandemic Preparedness, population density, the Polity Score to measure democracy, and a Female Leadership indicator.⁷² We note that the gender of a country’s leader can vary over time during the sample period. We also include a count of days until the next national elections in each country.

In the SI we present in more detail results of regression models predicting our social distancing, school closures, and business restrictions intensity indices in Table 1 in the SI. In SI Figures 5 and 6 we also show predicted values of the indices given values of covariates to show the strength of the associations we uncover. We note here some of the most intriguing patterns. In terms of mental health, the intensity of business restrictions, social distancing, and to an even greater extent school restrictions, is associated with fewer people reporting feeling anxious within the past seven days per our Facebook survey data. However, we note that financial anxiety increased when business and social distancing measures became more intense, though not to as great a degree as the reduction in general anxiety. Importantly, only social distancing intensity is associated with a decrease in the reporting of personal contacts outside the home, as such contacts rose slightly under business restrictions and quite strongly with school restrictions.

In terms of political and social covariates, we find that countries with more democratic institutions and a greater respect for human rights had less intense policies of all three types, though the association is stronger for our measure of respect for human rights than democratic institutions in general. This quite strong association aligns with recent research suggesting that COVID-19 restrictions can be a way for governments

to aggrandize power and abuse citizens' rights.^{73,74} Furthermore, we show that our measures of economic interdependence, particularly FDI and trade as a percentage of GDP, are strongly associated with reduced policy intensity for businesses, schools and social distancing policies. This level of interdependence contrasts strongly with the level of wealth as countries with a larger GDP were in fact more likely to have higher rates of policy intensity.

Finally, we note that we can show with our estimates that the different policy intensity scores tend to be complements, rather than substitutes, as can be seen in Figure 7 in the SI. In other words, policymakers tend to increase policy intensity in tandem across business, school and social distancing measures, rather than trading off intensity across policy types.

Application to Estimating COVID-19 Suppression

We next turn to illustrating the use of our indices as predictors of COVID-19 mortality outcomes with a replication and extension of Flaxman et al.'s study of government policies' suppression effects.⁴² In their analysis of the early pandemic in Europe, each of the variables they measured (lockdowns, public events restrictions, school closures, self-isolation policies and public encouragements of social distancing) was coded as either 0 before policy adoption or 1 for every day following policy adoption to the end of the sample period (31 December, 2019 to 4 May, 2020). While the study was remarkably successful at estimating the R_t (time-varying reproduction number) of SARS-CoV-2 spread from limited and biased data, the authors noted that a limitation of their findings came from coding issues with their policies:

“Most interventions were implemented in rapid succession in many countries, and as such it is difficult to disentangle the individual effect sizes of each intervention (p. 260).”

This difficulty in disentangling effect sizes stems partially from the fact that their five policy variables are highly correlated. In their response⁵³ to a critique⁵² of their estimates, the authors again noted that it was impossible to fully resolve these issues without “more granular information on NPIs [that] will in future hopefully give a more nuanced understanding of which measures—whether mandatory or voluntary—contributed most to reductions in transmission.”

As can be seen in the bottom half of Figure 5, our policy intensity scores for the same sample period never have a correlation value higher than 0.74 and are often as low as 0.21, indicating that the information in the indices still varies even with the similar country responses during the early period of the pandemic. For this reason, our policy intensity scores help resolve the data reduction problem by eliminating highly correlated variables and substituting a much more informative continuous measure with a clear theoretical basis.

Figure 6: Comparison of Correlation of Policy Measures in Flaxman et al. and the Policy Intensity Scores

Figure 7: Flaxman et al. Replication Using Policy Index Scores both with and without Measurement Error

Moreover, the advantage for applied modeling is not only theoretical. Due to these highly correlated measures, Flaxman et al. could not obtain estimates when including all policy indicators as predictors of R_t . SI Figure 1 shows their multivariate results, which show that only lockdowns are able to reduce R_t . As the authors note in their supplemental information (Figure 19), when each of their policy indicators are entered separately, they all are associated with very large reductions in R_t . In other words, the very large effect on R_t attributed to lockdowns in SI Figure 1 is partly an artifact of the high correlation of the policy indicators.

To show how our indices provide more information, we report our re-analysis of the Flaxman et al. model with all six of our policy scores in Figure 7. We estimated two specifications; the first with our scores as covariates without measurement error and the second with normally-distributed measurement error as described in the Methods section in the SI.

The low correlation between covariates is a result of the fact that our estimation approach of our indices employed a diverse variety of data in each policy type. This means that the coefficients in Figure 7 can take on a variety of values. We also note that uncertainty in general is much higher than the original results. In addition, the model with measurement error shows wider uncertainty intervals than the model without error despite the artificial lower bound. On the whole, both models show that the intensity of mask policies actually had the strongest association with reduced deaths, followed by business restrictions, general social distancing (our closest concept to lockdown policies), and school closures. These findings are an interesting comparison to the original paper, which did not have a measure of mask policies and instead inferred that

the steep drop in cases in Europe came about solely via social distancing measures.

There are significant differences between the model with and without measurement error beyond the inflation of uncertainty. The model with measurement error tends to put more weight on the intensity of mask policies while deflating the role of business restrictions and general social distancing. While it might appear that the measurement model’s mask coefficient does not have an uncertainty interval, that is solely due to the non-linear nature of R_t . The coefficient on mask policies in the measurement error model (6.9, **UI**: [6.18, 7.78]) is double that of the model without measurement error (2.77, **UI**: [1.5, 4.07]), and the exponentiation to convert to R_t means that effect is concentrated at very large reductions in R_t .

SI Figure 2 shows a time series plot of our indices for the sample period of the original study with mask scores highlighted. The figure reveals that mask policies did not come into effect until close to end of the sample period in late April, which is why they were not included in the original analysis. Furthermore, this fact indicates that the size of the masks coefficient in Figure 7 is likely affected by the limitations of the sample period. With more data, it is likely that the masks’ estimated effect on R_t would decrease as more waves occurred even with mask policies in place. However, we think that the model’s results do capture a potentially important association between mask policies and the end of the first wave of the pandemic in Europe that the original study did not identify.

To examine the robustness of the result to individual countries, in SI Figures 3 and 4 we estimated models in which we left out one of the eleven countries at a time for both the measurement error and no error models. SI Figure 3, which contains the results for the no error model, reveal that dropping France lowers the estimated effect on R_t for mask policies from approximately a 80% reduction to an approximately 50% reduction. During the sample period, France had a modest early increase in its masks score because it implemented a low-discrimination policy which apparently had little effect at reducing COVID infections. By comparison, SI Figure 4 shows that the model with measurement error experiences much less perturbation to removing a country from the data, with most estimates close to the average effect across countries.

Discussion

Making sense of the quantity and variation of government policies made in response to the pandemic is crucial for our understanding of the pandemic. To conceptualize our indices, we build on a simple, but reasonable, assumption that policymakers operate under fixed budget constraints and must thus consider a multitude of costs and benefits when committing resources across different possible COVID-19 policy responses. We further provide a framework for grouping different policies according to the theoretical mechanism through which they are expected to limit the spread of the pandemic. By using an ideal point

model (see Methods section below for more information), we are able to create indices which model the relative intensity a policymaker devotes to a given policy domain and we further show that these indices, while correlated with the underlying raw data from which they are derived, are able to capture latent traits beyond what can be measured by the raw data alone. We further show not only how these indices can be used to understand social, economic and political drivers of pandemic response, but by replicating Flaxman et al.’s influential study, how they can help address conceptual and methodological challenges in existing COVID-19 research.

Given the scale and complexity of this measurement problem, a further strength of our indices is that it leverage multiple datasets to be able to validate the underlying indicators. In this paper, we employ the CoronaNet and OxCGRT datasets to measure policy intensity because of their scope and complementary data coding strategies. The breadth of the CoronaNet data results in a large number of indicators while any particular indicator does not contain as much information about a given policy compared to the OxCGRT dataset. By contrast, the OxCGRT dataset contains relatively fewer indicators compared to CoronaNet, but each indicator represents a rich source of qualitative information about the extent of policy coverage. As such, pooling information from the two datasets allows for independent sources of information about policies to be combined. Indeed, it is much more difficult to pool information across datasets that adopt conceptually similar coding strategies because full taxonomies of all categories in the data must be manually mapped, assuming that such an exact mapping is possible.

Overall, our policy intensity scores address an important challenge in COVID-19 research: how should researchers adequately capture the flurry of policy-making activity made in response to COVID-19 in their research on the pandemic? When used as a predictor of COVID-19 cases and deaths, our indices can be seen as complements to using data on individual policy measures. Some of the individual policies have low associations with our indices, which signals that they may be adopted in a different pattern of intensity and timing as the others in the same policy domain. Hence, they will be more weakly associated with the latent variable and have less influence in the prediction of the relevant outcomes. It is possible that some of these individual policies are relevant in containing the pandemic spread, and policies affect contagion and death regardless of the intentions of the policymaker. For this reason, results based on our indices should not preclude further evaluation of individual policies given adequate care about collinearity issues. Meanwhile, if the purpose is to explain overall policy intensity across different countries within each of the six domains, our policy intensity latent variable present a clear methodologically advantage relative to using measures of raw individual policies.

Acknowledgments

We thank Seth Flaxman and Samir Bhatt for assistance with replicating their study. A full list of the hundreds of research assistants involved in compiling the CoronaNet data are in the supplemental information and we gratefully acknowledge each one for their role in this project. We acknowledge funding from New York University Abu Dhabi (Covid-19 Facilitator Research Fund), The National Council for Eurasian and East European Research (No. 832-06g), the European Union’s Horizon 2020 Research and Innovation Programme (No. 101016233), the Hochschule für Politik at the Technical University of Munich (TUM) and the TUM Chair of International Relations, and the Leibniz Research Alliance Group (Crises in a Globalised World). Technical infrastructure support has been generously provided by Overton, Slack Technologies and RStudio. The funders had no role in study design, data collection, analysis, or interpretation, nor in the decision to publish or in the preparation of the manuscript. We also thank Annika Kaiser and Enakshi Chatterjee for their help in identifying suitable literature to for the paper.

Author Contributions

Please see the contribution matrix below showing the contributions for each author. Authors affiliated with the CoronaNet project are marked with one asterisk and authors affiliated with the OxGCRT project are marked with two asterisks.

	Writing - review & editing	Writing - original draft	Visualization	Validation	Supervision	Software	Resources	Project Administration	Methodology	Investigation	Funding Acquisition	Formal Analysis	Data Curation	Conceptualization
RK*														
JB*														
RG**														
VG*														
TM*														
CS*														
CC*														
TH**														
AH*														
LM*														
AP**														
ST*														

Roles taken from the CRediT Contributor Roles Taxonomy (<https://casrai.org/credit/>).

Competing Interests

The authors have no competing interests to declare.

References

- ¹ Bo, Y. *et al.* Effectiveness of non-pharmaceutical interventions on COVID-19 transmission in 190 countries from 23 January to 13 April 2020. *International Journal of Infectious Diseases* **102**, 247–253 (2021). URL <https://www.sciencedirect.com/science/article/pii/S1201971220322700>.
- ² Bauch, C. T. Estimating the COVID-19 R number: a bargain with the devil? *The Lancet Infectious Diseases* **21**, 151–153 (2021). URL [https://www.thelancet.com/journals/laninf/article/PIIS1473-3099\(20\)30840-9/abstract](https://www.thelancet.com/journals/laninf/article/PIIS1473-3099(20)30840-9/abstract). Publisher: Elsevier.

- ³ Sharma, M. *et al.* How Robust are the Estimated Effects of Nonpharmaceutical Interventions against COVID-19? *arXiv:2007.13454 [cs, q-bio, stat]* (2020). URL <http://arxiv.org/abs/2007.13454>. ArXiv: 2007.13454.
- ⁴ Brauner, J. M. *et al.* Inferring the effectiveness of government interventions against COVID-19. *Science* (2020). URL <https://science.sciencemag.org/content/early/2020/12/15/science.abd9338>. Publisher: American Association for the Advancement of Science Section: Research Article.
- ⁵ Fetzer, T. R. *et al.* Global Behaviors and Perceptions at the Onset of the COVID-19 Pandemic. Tech. Rep. w27082, National Bureau of Economic Research (2020). URL <https://www.nber.org/papers/w27082>.
- ⁶ Varga, T. V. *et al.* Loneliness, worries, anxiety, and precautionary behaviours in response to the COVID-19 pandemic: A longitudinal analysis of 200,000 Western and Northern Europeans. *The Lancet Regional Health - Europe* **2**, 100020 (2021). URL <https://www.sciencedirect.com/science/article/pii/S266677622030020X>.
- ⁷ Yamada, Y. *et al.* COVIDiSTRESS Global Survey dataset on psychological and behavioural consequences of the COVID-19 outbreak. *Scientific Data* **8**, 3 (2021). URL <https://www.nature.com/articles/s41597-020-00784-9>. Number: 1 Publisher: Nature Publishing Group.
- ⁸ Zettler, I. *et al.* The Role of Personality in COVID-19-Related Perceptions, Evaluations, and Behaviors: Findings Across Five Samples, Nine Traits, and 17 Criteria. *Social Psychological and Personality Science* 19485506211001680 (2021). URL <https://doi.org/10.1177/19485506211001680>. Publisher: SAGE Publications Inc.
- ⁹ Demirgüç-Kunt, A., Lokshin, M. & Torre, I. The Sooner, the Better: The Early Economic Impact of Non-Pharmaceutical Interventions During the COVID-19 Pandemic. SSRN Scholarly Paper ID 3611386, Social Science Research Network, Rochester, NY (2020). URL <https://papers.ssrn.com/abstract=3611386>.
- ¹⁰ Maloney, W. & Taskin, T. *Determinants of Social Distancing and Economic Activity during COVID-19: A Global View*. Policy Research Working Papers (The World Bank, 2020). URL <https://elibrary.worldbank.org/doi/abs/10.1596/1813-9450-9242>.
- ¹¹ Kubinec, R., Lee, H. N.-K. & Tomashevskiy, A. How to Get Away with Spreading COVID-19: Political Connections and Pandemic Response. *SocArXiv* (2020). URL <https://osf.io/preprints/socarxiv/68fpr/>. Type: article.

- ¹² Brodeur, A., Gray, D., Islam, A. & Bhuiyan, S. J. A literature review of the economics of COVID-19. *IZA Institute of Labor Economics Discussion Paper Series* (2020). URL https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3636640.
- ¹³ Bonaccorsi, G. *et al.* Economic and social consequences of human mobility restrictions under COVID-19. *Proceedings of the National Academy of Sciences* **117**, 15530–15535 (2020). URL <https://www.pnas.org/content/117/27/15530>. Publisher: National Academy of Sciences Section: Social Sciences.
- ¹⁴ Ashraf, B. N. Economic impact of government interventions during the COVID-19 pandemic: International evidence from financial markets. *Journal of Behavioral and Experimental Finance* **27**, 100371 (2020). URL <https://www.sciencedirect.com/science/article/pii/S2214635020302422>.
- ¹⁵ Adams-Prassl, A., Boneva, T., Golin, M. & Rauh, C. Inequality in the impact of the coronavirus shock: Evidence from real time surveys. *Journal of Public Economics* **189**, 104245 (2020). URL <http://www.sciencedirect.com/science/article/pii/S0047272720301092>.
- ¹⁶ Adida, C. L., Dionne, K. Y. & Platas, M. R. Ebola, elections, and immigration: how politicizing an epidemic can shape public attitudes. *Politics, Groups, and Identities* **8**, 488–514 (2020). URL <https://doi.org/10.1080/21565503.2018.1484376>. Publisher: Routledge eprint: <https://doi.org/10.1080/21565503.2018.1484376>.
- ¹⁷ Woolhandler, S. *et al.* Public policy and health in the Trump era. *The Lancet* **0** (2021). URL [https://www.thelancet.com/journals/lancet/article/PIIS0140-6736\(20\)32545-9/abstract](https://www.thelancet.com/journals/lancet/article/PIIS0140-6736(20)32545-9/abstract). Publisher: Elsevier.
- ¹⁸ Pulejo, M. & Querubín, P. Electoral Concerns Reduce Restrictive Measures During the COVID-19 Pandemic. Tech. Rep. w27498, National Bureau of Economic Research (2020). URL <https://www.nber.org/papers/w27498>.
- ¹⁹ Painter, M. & Qiu, T. Political beliefs affect compliance with COVID-19 social distancing orders. *SSRN* (2020).
- ²⁰ Hart, P. S., Chinn, S. & Soroka, S. Politicization and polarization in COVID-19 news coverage. *Science Communication* (2020).
- ²¹ Gadarian, S. K., Goodman, S. W. & Pepinsky, T. B. Partisanship, health behavior and policy attitudes in the early stages of the COVID-19 pandemic. *SSRN* (2020).

- ²² Dave, D. M., Friedson, A. I., Matsuzawa, K., Sabia, J. J. & Safford, S. Black lives matter protests, social distancing, and COVID-19. *NBER* (2020).
- ²³ Bol, D., Giani, M., Blais, A. & Loewen, P. J. The effect of COVID-19 lockdowns on political support: Some good news for democracy? *European Journal of Political Research* **60**, 497–505 (2021). URL <https://doi.org/10.1111/1475-6765.12401>. Publisher: John Wiley & Sons, Ltd.
- ²⁴ Bargain, O. & Aminjonov, U. Trust and compliance to public health policies in times of COVID-19. *Journal of Public Economics* **192**, 104316 (2020). URL <https://www.sciencedirect.com/science/article/pii/S0047272720301808>.
- ²⁵ Borgonovi, F. & Andrieu, E. Bowling together by bowling alone: Social capital and COVID-19. *Social Science & Medicine* **265**, 113501 (2020). URL <https://www.sciencedirect.com/science/article/pii/S0277953620307206>.
- ²⁶ Fan, Y., Orhun, A. Y. & Turjeman, D. Heterogeneous actions, beliefs, constraints and risk tolerance during the COVID-19 pandemic. *NBER* (2020).
- ²⁷ Barceló, J. & Sheen, G. C.-H. Voluntary adoption of social welfare-enhancing behavior: Mask-wearing in Spain during the COVID-19 outbreak. *PloS one* **15**, e0242764 (2020). Publisher: Public Library of Science San Francisco, CA USA.
- ²⁸ UNDP and UN Women COVID-19 Global Gender Response Tracker. Dataset, United Nations Development Programme (2021). URL https://data.undp.org/wp-content/uploads/2021/03/COVID-19-Global-Tracker-Methodological-Note-v_March2021.pdf.
- ²⁹ Elgin, C., Basbug, G. & Yalaman, A. Economic policy responses to a pandemic: Developing the COVID-19 economic stimulus index. *Covid Economics* **1**, 40–53 (2020).
- ³⁰ Naqvi, A. COVID-19 European regional tracker. *medRxiv* 2021.02.15.21251788 (2021). URL <https://www.medrxiv.org/content/10.1101/2021.02.15.21251788v2>. Publisher: Cold Spring Harbor Laboratory Press.
- ³¹ Adolph, C., Amano, K., Bang-Jensen, B., Fullman, N. & Wilkerson, J. Pandemic Politics: Timing State-Level Social Distancing Responses to COVID-19. *Journal of Health Politics, Policy and Law* **46**, 211–233 (2021). URL <https://doi.org/10.1215/03616878-8802162>.
- ³² COVID-19 Government Measures Dataset. Tech. Rep., ACAPS (2020). URL <https://www.acaps.org/covid-19-government-measures-dataset>.

- ³³ Porcher, S. Response2covid19, a dataset of governments' responses to COVID-19 all around the world. *Scientific Data* **7**, 423 (2020). URL <https://www.nature.com/articles/s41597-020-00757-y>. Number: 1 Publisher: Nature Publishing Group.
- ³⁴ Grundy, C., Quinn, O. & Todowede, S. Global Dataset of Public Health and Social Measures. Tech. Rep., World Health Organization (2021). URL <https://www.who.int/emergencies/diseases/novel-coronavirus-2019/phsm>.
- ³⁵ Suryanarayanan, P. *et al.* AI-assisted tracking of worldwide non-pharmaceutical interventions for COVID-19. *Scientific Data* **8**, 94 (2021). URL <https://doi.org/10.1038/s41597-021-00878-y>.
- ³⁶ Haug, N. *et al.* Ranking the effectiveness of worldwide COVID-19 government interventions. *Nature Human Behaviour* **4**, 1303–1312 (2020). URL <https://www.nature.com/articles/s41562-020-01009-0>. Number: 12 Publisher: Nature Publishing Group.
- ³⁷ Desvars-Larrive, A. *et al.* A structured open dataset of government interventions in response to COVID-19. *Scientific Data* **7**, 285 (2020). URL <https://www.nature.com/articles/s41597-020-00609-9>. Number: 1 Publisher: Nature Publishing Group.
- ³⁸ Zheng, Q. *et al.* HIT-COVID, a global database tracking public health interventions to COVID-19. *Scientific Data* **7**, 286 (2020). URL <https://www.nature.com/articles/s41597-020-00610-2>. Number: 1 Publisher: Nature Publishing Group.
- ³⁹ Bauer, M. W. & Knill, C. A Conceptual Framework for the Comparative Analysis of Policy Change: Measurement, Explanation and Strategies of Policy Dismantling. *Journal of Comparative Policy Analysis: Research and Practice* **16**, 28–44 (2014). URL <https://doi.org/10.1080/13876988.2014.885186>. Publisher: Routledge eprint: <https://doi.org/10.1080/13876988.2014.885186>.
- ⁴⁰ Kruse, T. & Strack, P. Optimal Control of an Epidemic through Social Distancing. SSRN Scholarly Paper ID 3583186, Social Science Research Network, Rochester, NY (2020). URL <https://papers.ssrn.com/abstract=3583186>.
- ⁴¹ Mandel, A. & Veetil, V. The Economic Cost of COVID Lockdowns: An Out-of-Equilibrium Analysis. *Economics of Disasters and Climate Change* **4**, 431–451 (2020). URL <https://doi.org/10.1007/s41885-020-00066-z>.
- ⁴² Flaxman, S. *et al.* Estimating the effects of non-pharmaceutical interventions on COVID-19 in Europe. *Nature* **584**, 257–261 (2020). URL <https://doi.org/10.1038/s41586-020-2405-7>.

- ⁴³ Li, Y. *et al.* The temporal association of introducing and lifting non-pharmaceutical interventions with the time-varying reproduction number (R) of SARS-CoV-2: a modelling study across 131 countries. *The Lancet Infectious Diseases* **21**, 193–202 (2021). URL [https://www.thelancet.com/journals/laninf/article/PIIS1473-3099\(20\)30785-4/abstract](https://www.thelancet.com/journals/laninf/article/PIIS1473-3099(20)30785-4/abstract). Publisher: Elsevier.
- ⁴⁴ Liu, Y. *et al.* Aerodynamic analysis of SARS-CoV-2 in two Wuhan hospitals. *Nature* **582**, 557–560 (2020). URL <https://www.nature.com/articles/s41586-020-2271-3>. Bandiera_abtest: a Cg_type: Nature Research Journals Number: 7813 Primary_atype: Research Publisher: Nature Publishing Group Subject_term: Microbiology;Risk factors;SARS-CoV-2 Subject_term_id: microbiology;risk-factors;sars-cov-2.
- ⁴⁵ Spiegel, M. I. & Tookes, H. Business Restrictions and COVID Fatalities. SSRN Scholarly Paper ID 3725015, Social Science Research Network, Rochester, NY (2021). URL <https://papers.ssrn.com/abstract=3725015>.
- ⁴⁶ Southall, E. *et al.* An analysis of school absences in England during the Covid-19 pandemic. Tech. Rep. (2021). URL <https://www.medrxiv.org/content/10.1101/2021.02.10.21251484v2>. Company: Cold Spring Harbor Laboratory Press Distributor: Cold Spring Harbor Laboratory Press Label: Cold Spring Harbor Laboratory Press Type: article.
- ⁴⁷ Hellewell, J. *et al.* Feasibility of controlling COVID-19 outbreaks by isolation of cases and contacts. *The Lancet Global Health* **8**, e488–e496 (2020). URL [https://www.thelancet.com/journals/langlo/article/PIIS2214-109X\(20\)30074-7/abstract](https://www.thelancet.com/journals/langlo/article/PIIS2214-109X(20)30074-7/abstract). Publisher: Elsevier.
- ⁴⁸ Paganelli, A. I. *et al.* A conceptual IoT-based early-warning architecture for remote monitoring of COVID-19 patients in wards and at home. *Internet of Things* 100399 (2021). URL <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8023791/>.
- ⁴⁹ Cheng, C., Barcelo, J., Hartnett, A. S., Kubinec, R. & Messerschmidt, L. COVID-19 government response event dataset (CoronaNet v.1.0). *Nature Human Behavior* (2020).
- ⁵⁰ Hale, T. *et al.* A global panel database of pandemic policies (Oxford COVID-19 Government Response Tracker). *Nature Human Behaviour* 1–10 (2021). URL <https://www.nature.com/articles/s41562-021-01079-8>. Publisher: Nature Publishing Group.
- ⁵¹ Kubinec, R. Generalized ideal point models for time-varying and missing-data inference. *Open Science Foundation Preprints* (2019).

- ⁵² Soltesz, K. *et al.* The effect of interventions on COVID-19. *Nature* **588**, E26–E28 (2020). URL <https://www.nature.com/articles/s41586-020-3025-y>. Bandiera_abtest: a Cg_type: Nature Research Journals Number: 7839 Primary_atype: Research Publisher: Nature Publishing Group Subject_term: Computational science;Epidemiology;SARS-CoV-2 Subject_term_id: computational-science;epidemiology;sars-cov-2.
- ⁵³ Flaxman, S. *et al.* Reply to: The effect of interventions on COVID-19. *Nature* **588**, E29–E32 (2020). URL <https://www.nature.com/articles/s41586-020-3026-x>. Bandiera_abtest: a Cg_type: Nature Research Journals Number: 7839 Primary_atype: Research Publisher: Nature Publishing Group Subject_term: SARS-CoV-2;Statistical methods;Viral infection Subject_term_id: sars-cov-2;statistical-methods;viral-infection.
- ⁵⁴ Dave, D. M., Friedson, A. I., Matsuzawa, K. & Sabia, J. J. When Do Shelter-in-Place Orders Fight COVID-19 Best? Policy Heterogeneity Across States and Adoption Time. Tech. Rep. w27091, National Bureau of Economic Research (2020). URL <https://www.nber.org/papers/w27091>.
- ⁵⁵ Courtemanche, C., Garuccio, J., Le, A., Pinkston, J. & Yelowitz, A. Strong Social Distancing Measures In The United States Reduced The COVID-19 Growth Rate. *Health Affairs* **39**, 1237–1246 (2020). URL <https://www.healthaffairs.org/doi/full/10.1377/hlthaff.2020.00608>. Publisher: Health Affairs.
- ⁵⁶ Abouk, R. & Heydari, B. The Immediate Effect of COVID-19 Policies on Social-Distancing Behavior in the United States. *Public Health Reports* 0033354920976575 (2021). URL <https://doi.org/10.1177/0033354920976575>. Publisher: SAGE Publications Inc.
- ⁵⁷ Islam, N. *et al.* Physical distancing interventions and incidence of coronavirus disease 2019: natural experiment in 149 countries. *BMJ* **370**, m2743 (2020). URL <https://www.bmj.com/content/370/bmj.m2743>. Publisher: British Medical Journal Publishing Group Section: Research.
- ⁵⁸ Perra, N. Non-pharmaceutical interventions during the COVID-19 pandemic: A review. *Physics Reports* (2021). URL <https://www.sciencedirect.com/science/article/pii/S0370157321000624>.
- ⁵⁹ Carpenter, B. *et al.* Stan: A probabilistic programming language. *Journal of Statistical Software* **76** (2017).
- ⁶⁰ Clinton, J., Jackman, S. & Rivers, D. The statistical analysis of rollcall data. *American Political Science Review* **98**, 355–370 (2004). Tex.owner: saudiwin tex.timestamp: 2017.04.03.
- ⁶¹ Durbin, J. Errors in variables. *Review of the International Statistical Institute* **1**, 23–32 (1954).

- ⁶² Kurz, A. S. Missing Data and Other Opportunities. In *Statistical Rethinking with brms, ggplot2, and the tidyverse* (2020). URL https://bookdown.org/ajkurz/Statistical_Rethinking_recoded/missing-data-and-other-opportunities.html.
- ⁶³ Bizarro, F., Coppedge, M. & Pemstein, D. Incorporating v-dem’s uncertainty estimates in regression analysis. Working paper (2016). URL <https://kellogg.nd.edu/projects/vdem/files/2016/tutorial.pdf>. Tex.pagetotal: 32.
- ⁶⁴ Bürkner, P.-C. brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software* **80**, 1–28 (2017). Tex.encoding: UTF-8.
- ⁶⁵ Tang, F. & Ishwaran, H. Random forest missing data algorithms. Tech. Rep., Archiv (2017). URL <https://arxiv.org/pdf/1701.05305.pdf>.
- ⁶⁶ Stekhoven, D. J. & Bühlmann, P. MissForest–non-parametric missing value imputation for mixed-type data. *Bioinformatics* **28**, 112–118 (2012).
- ⁶⁷ Zhou, X. & Reiter, J. P. A Note on Bayesian Inference After Multiple Imputation. *The American Statistician* **64**, 159–163 (2010). URL <https://doi.org/10.1198/tast.2010.09109>. Publisher: Taylor & Francis _eprint: <https://doi.org/10.1198/tast.2010.09109>.
- ⁶⁸ Sebhatu, A., Wennberg, K., Arora-Jonsson, S. & Lindberg, S. I. Explaining the homogeneous diffusion of COVID-19 nonpharmaceutical interventions across heterogeneous countries. *Proceedings of the National Academy of Sciences* **117**, 21201–21208 (2020). URL <https://www.pnas.org/content/117/35/21201>. Publisher: National Academy of Sciences Section: Social Sciences.
- ⁶⁹ Petherick, A. *et al.* A worldwide assessment of changes in adherence to COVID-19 protective behaviours and hypothesized pandemic fatigue. *Nature Human Behaviour* 1–16 (2021). URL <https://www.nature.com/articles/s41562-021-01181-x>. Bandiera_abtest: a Cg_type: Nature Research Journals Primary_atype: Research Publisher: Nature Publishing Group Subject_term: Health humanities;Social policy Subject_term.id: health-humanities;social-policy.
- ⁷⁰ Phillips, T., Zhang, Y. & Petherick, A. A Year of Living Distantly: Trends in the Use of Stay-at-Home Orders Over the First 12 Months of the COVID-19 Pandemic. SSRN Scholarly Paper ID 3847818, Social Science Research Network, Rochester, NY (2021). URL <https://papers.ssrn.com/abstract=3847818>.
- ⁷¹ Brenner, M. H. & Bhugra, D. Acceleration of Anxiety, Depression, and Suicide: Secondary Effects of Economic Disruption Related to COVID-19. *Frontiers in Psychiatry* **11** (2020). URL <https://www.frontiersin.org/articles/10.3389/fpsy.2020.592467/full>. Publisher: Frontiers.

- ⁷² Coppedge, M. *et al.* V-dem dataset v7. Working paper, Social Science Research Network (2017). URL https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2968289. Url: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2968289.
- ⁷³ Edgell, A. B., Lachapelle, J., Lührmann, A. & Maerz, S. F. Pandemic backsliding: Violations of democratic standards during Covid-19. *Social Science & Medicine* **285**, 114244 (2021). URL <https://www.sciencedirect.com/science/article/pii/S0277953621005761>.
- ⁷⁴ Barceló, J., Kubinec, R., Cheng, C., Rahn, T. H. & Messerschmidt, L. Windows of Repression: Using COVID-19 Policies against Political Dissidents? Tech. Rep., SocArXiv (2020). URL <https://osf.io/preprints/socarxiv/yuqw2/>. Type: article.