A Global Panel Database of Pandemic Policies (Oxford COVID-19 Government Response Tracker)

Authors

Thomas Hale[1*], Noam Angrist[1], Rafael Goldszmidt[2], Beatriz Kira[1], Anna Petherick[1], Toby Phillips[1], Samuel Webster[3], Emily Cameron-Blake[1], Laura Hallas[1], Saptarshi Majumdar[1], Helen Tatlow[1]

Affiliations

1. Blavatnik School of Government, University of Oxford, Oxford, United Kingdom
2. Brazilian School of Public and Business Administration, Getulio Vargas Foundation, Rio de Janeiro, Brazil
3. Unaffiliated, London, United Kingdom

*Corresponding author: Thomas Hale (Thomas.Hale@bsg.ox.ac.uk)

Abstract

COVID-19 has prompted unprecedented government action around the world. We introduce the Oxford COVID-19 Government Response Tracker (OxCGRT), a dataset that addresses the need for continuously updated, readily usable, and comparable information on policy measures. From 1 January 2020, the data capture government policies related to closure and containment, health, and economic policy for more than 180 countries, plus several countries' subnational jurisdictions. Policy responses are recorded on ordinal or continuous scales for 19 policy areas, capturing variation in degree of response. We present two motivating applications of the data, highlighting patterns in the timing of policy adoption and subsequent policy easing and re-imposition, and illustrating how the data can be combined with behavioural and epidemiological indicators. This database enables researchers and policymakers to explore the empirical effects of policy responses on the spread of COVID-19 cases and deaths, as well as on economic and social welfare.

Main text

The rapid spread of COVID-19 globally has been met with an extraordinary range of government responses. These measures include school closings, travel restrictions, bans on public gatherings, emergency investments in healthcare facilities, new forms of social welfare provision, contact tracing and other interventions to contain the spread of the virus, augment health systems, and manage the economic consequences of these actions. The Oxford COVID-19 Government Response Tracker (OxCGRT) provides a systematic set of cross-national, longitudinal measures of government responses from 1 January 2020. The project tracks national and, for some countries, subnational governments' policies and interventions across a standardized series of indicators. It further creates a suite of composite indices to measure the extent of these responses. By providing open-access, near-real time data in an easily accessible, time-series format, the project offers a critical resource for policymakers and researchers to understand the effect of policies on disease spread, on socioeconomic welfare, and other outcomes of interest. Consequently, the data offer an objective grounding for debate as policymakers and publics deliberate over approaches to COVID-19.

While OxCGRT is continuously expanding, it at present includes 19 policy indicators covering closure and containment, health, and economic policies (Table 1). These data cover over 180 countries as well as subnational jurisdictions of the United States, Brazil, United Kingdom and Canada, with subnational units of more countries being added over time (Extended Data Figure 1). A team of more than four hundred volunteers around the world, affiliated with Oxford University and partners, have been working to collect and code the data in real time. Members of this large and diverse team use their contextual knowledge and expertise in 88 languages to parse reporting and government announcements. Coder training, testing, and weekly meetings that generate rules of thumb for edge-case policies ensure coding consistency, and every data point is reviewed by a second coder.

OxCGRT's design emphasizes comparability, legibility, and transparency. The data are published in multiple time-series formats for ease of use by non-experts and researchers alike, with legacy data available for continuity as we add new indicators. Several features underpin our approach. First, observations for most indicators are reported on monotonic ordinal scales, with others coded on continuous scales, allowing for quantitative analysis of the degree of government response. Second, the indicators are aggregated in different combinations into four composite indices (Table 2) that provide a snapshot of the number and degree of policies in place in a given area. Third, geographic scope is recorded for appropriate indicators. Fourth, source notes and archived links to original sources are included to support detailed interpretation of specific policies.

OxCGRT has been widely used (see examples below) in the first months of the pandemic, revealing that the global coverage, granularity of policy detail, and systematic structure of the data have been able to inform diverse literatures[1]. For instance, the data has been used by health policy experts and data scientists to calculate the levels of health care resources that are associated with different levels of transmission[2], to estimate the impact of combinations of physical distancing measures on disease incidence[3,4] and on the time-varying reproduction number (R_t)[5]. Environmental scientists have drawn upon the data to examine whether COVID-19 response policies affect air pollution levels[6,7]. Political scientists have considered whether policies vary by regime type[8,9], and assessed whether upcoming elections reduce the strength of responses[10]. Economists have used the data to explore how working from home has shifted countries' sectoral structures[11], to link stay-at-home policies to increasing food prices[12], and to identify knock-on effects of large countries' response policies on the GDP growth of smaller trade partners[13].

Many of those using the data have benefited from the specific features listed above. The ordinal indicator scales permit separate assessment of policy recommendations, as well as permissive and strict regulations[14]. While significant attention has focused on the closure and containment measures captured in the Stringency Index, some studies have used all four[15], or selected from among the Containment and Health Index (CHI), the more holistic Government Response Index (GRI)[16], and the Economic Support Index (ESI)[17], which provides an overall measure of financial assistance to households. Moreover, the coding of policies' geographic scope has enabled the analysis of strictly national policies[18], and comparison between national and localized approaches. These examples illustrate the value of OxCGRT data, and related datasets[19], in helping researchers—in addition to decisionmakers and publics—make sense of the effects of governments' responses to COVID-19 across different populations and contexts, as well as what leads governments to adopt different policies.

In the following sections, we describe patterns of global COVID-19 government responses with the OxCGRT data in order to demonstrate what kinds of questions the data can help researchers tackle. We describe cross-national patterns in the timing of containment and health policies, followed by a more detailed presentation of policy sequencing. We then combine the data with mobile-phone mobility data to relate policies to human behaviour[20] and review the potential for bringing together OxCGRT data with additional data sources in the Discussion. In the Methods section, we describe the individual indicators in more detail, along with the data collection process, data coverage and how we calculate the indices. We also briefly compare OxCGRT to related projects to highlight their complementarities.

Results

To motivate applications of the data, we present general trends and patterns in government responses in the first months of the pandemic. We focus here on cross-national patterns, though OxCGRT contains more granular data on subnational jurisdictions as well. First, we document a surprising degree of commonality across countries in the early months of the pandemic followed by growing divergence. We also note patterns in policy re-imposition and geographical scope, topics that have to-date been relatively underexplored in the literature, yet that have important implications for how countries manage each wave of the pandemic. Second, we consider associations between the OxCGRT indices and a key outcome of interest, individual mobility, to illustrate the potential for the data to be combined with other indicators to investigate economic, social, and epidemiological questions of interest.

What government responses do we observe?

The data reveal a striking degree of commonality in government responses to COVID-19 in the first months of the pandemic. We group the 19 indicators into themes of closure and containment, health, and economic support (Table 1), normalized to vary from 0 to 100 (for a full description, see Tables 3a and 3b). The Containment and Health Index (CHI) measures the number and intensity of closure and containment policies (e.g. school closings, stay-at-home measures) and policies toward disease surveillance (e.g. testing and contact tracing). Only a handful of countries had adopted strong containment (often referred to as “lockdown”) and health policies in early March, as Figure 1 shows, yet within a month the world had changed, and intensive policy responses had become a global phenomenon. In subsequent months, however, countries have lifted policy restrictions, and then, in some cases, reimposed policies in a policy “see-saw” as the epidemic waxes and wanes.

During the initial, global rise in policy responses, the data reveal a number of intriguing patterns. Most governments moved to a high level of response within a two-week period around the middle of March, showing remarkable clustering. Figure 2 displays this initial policy convergence across 183 countries, which is not observed in later policy decisions (such as rolling back measures). This initial clustering pattern in mid-March contrasts with what would be expected if countries reacted according to the local epidemiological progression of the pandemic. For instance, in most countries the sudden ramping up of response policies happened before they had experienced their tenth COVID-19 related death, while many other countries’ responses preceded even their tenth recorded case. Countries may have observed their neighbors, or the global response, and reacted in concert. This clustering then seems to dissipate in later months as countries’ responses diverge. This pattern has significant implications for the coordination of responses to global infectious diseases considering that the World Health Organization’s policy guidance to governments is tailored to the local progression of an infectious disease rather than potential “herd behaviour.”

Next, we examine specific policies, both during the initial process of policy adoption and in the months that followed as measures were either rolled back or maintained. The left panel of Figure 3 captures the ramping up of policies, showing the proportion of countries adopting a particular policy, with day zero representing the first day of

COVID-19 policy response in each country. The very limited crossing of the lines in this figure suggests that policies adopted by the median country (in terms of the speed of policy responses) occurred in approximately the same order as those adopted by countries in the first and third quartiles. In other words, the sequence of policy adoption is largely similar across countries. Specifically, there is more than a 50% chance that a randomly drawn country will have introduced public information campaigns, international travel controls, and testing policies within 20 days of the first government response of any policy type; there is a 40% chance of this within 10 days, and more than a 90% chance within two months. Economic support policies have tended to be established later than closure or containment and health policies, facial coverings aside.

A common pattern also characterizes policy reversal. The right panel of Figure 3 indicates the proportion of countries maintaining the highest point they reach on the ordinal scale for each policy area. The global rate of policy reduction is indicated by the slope of the lines. There is crossing over among policies, but little among those lines representing closure and containment policies, which have roughly similar rates of rollback. During the initial two months of policy easing, while closure and containment policies have been loosened, economic support policies and health policies have been maintained at countries' individual maximum strengths.

While we see similarities in what policies were adopted and relaxed, when, there is interesting variation across policies strength, geographic coverage, and the extent to which they were later reimposed. Figure 4 shows how frequently countries imposed the strongest possible policies, what portion of observed policies applied nationally versus sub-nationally, and which policies have been reduced and subsequently reimposed. Most closure policies were adopted nationally at some point, and in approximately 20% of countries, stronger closure and containment policies were re-imposed by the end of December 2020. In the case of workplace policies, approximately 80% of countries had reduced their restrictions by that point in the year, but 40% of all countries later reversed course.

The clustering of policy responses during the process of adoption has a critical implication for researchers. Analysis of individual policies is difficult because there is limited variation across and within countries, resulting in collinear relationships. This has meant that most analysis of government responses to date have had to focus on aggregate indices. However, in later periods we document substantially more variation. This variation enables more credible quasi-experimental analysis of individual policies, such as school re-opening, testing campaigns, and income support. Understanding the role of individual policies, in addition to aggregate government response levels, is of central importance for further research and policy action, which this database can enable.

<u>Motivating applications of the data: how do government responses relate to behaviour?</u>

A key application of the OxCGRT data is to understand how policies relate to human behaviour. A number of studies have used OxCGRT and similar data to try to estimate the effect of policies on behaviour and the spread of the disease. Here we do not aim to establish new estimates of causal effects, but rather seek only to demonstrate potential use cases to motivate further research. Figure 5 summarizes the results of linear panel regression models, comparing the strength of associations between the Containment and Health Index (CHI), Government Response Index GRI) and the Stringency Index (SI) with changes in citizen mobility over time, as recorded by mobile phone applications (the full results are presented in the Supplementary Tables 1 and 2). These models use standard techniques in the literature. We include country and date fixed effects in order to isolate within-country associations over time, accounting for seasonal and other calendar effects. In the supplementary information we include models that which also control for new daily deaths to identify the association between policies and mobility unconfounded by the relationship between the severity of the epidemic and mobility.

The coefficients and confidence intervals in Figure 5 show strong associations between OxCGRT indices and measures of behaviour. These associations are stronger the greater the number of policy indicators included in an index. All three indices shown contain eight indicators of closure and containment policies, plus at least one additional indicator. Increases in the GRI—our broadest index of government responses—are most strongly associated with increases in the percentage of time spent in residences, as well as with decreases in the frequency visits to groceries and pharmacies, workplaces, transit stations, places for retail and recreation, and parks. The CHI, which adds to closure and containment indicators all of our health-policy indicators—shows only slightly weaker associations. The SI, which brings together the containment and health indicators with just one additional indicator, public information campaigns, shows still less pronounced relationships in the same directions.

These analyses highlight the potential of OxCGRT data, combined with other datasets, to capture important changes in behaviour in response to historic government action. While the associations presented here are merely suggestive, researchers are already using the data for more in depth analyses[2,3,5]. Identifying causal effects of government policies is not straightforward due to many confounding factors and potential sources of endogeneity. Given these challenges, the rich nature of this database with day-by-day policy changes across a global distribution of countries and sub-national jurisdictions enables rigorous quasi-experimental analysis. This illustrative application is designed to motivate -further in-depth research and demonstrate the potential for policymakers and researchers to answer important public policy and epidemiological questions using OxCGRT data.

Discussion

Alongside epidemiological and behavioural data, measures of government response help researchers and decision-makers explore how best to address COVID-19. However, measuring government policies in a consistent and comparable way across jurisdictions and across time raises a number of methodological considerations, which can present difficult choices. In this section we review these trade-offs and outline strategies for addressing them.

First, while the OxCGRT ordinal scales distinguish, for example, a ban on gatherings of over 10 people from a ban on gatherings of over 100 people, the limitation of an ordinal approach is that it groups heterogenous observations into pre-established categories. For instance, both the UK and France had broadly similar stay-at-home orders during spring 2020, and both were categorized as the second-highest ordinal point on that indicator. However, French residents had to submit a form to authorities to leave their house, while UK residents did not. To mitigate the inevitable simplification that comes with codification, OxCGRT includes detailed notes and archived links to source materials for all observations in the dataset, helping researchers draw upon OxCGRT data in a more detailed way should it be required.

Second, the coding scheme loses granularity when applied to large jurisdictions with many heterogenous subunits. As described in the Methods section, OxCGRT data include three types of observations: those that describe all policies that apply to a given jurisdiction, those that describe policies put in place by a given level and lower levels of government, and those that describe only those instigated at a given level of government. Policies that apply only to a subunit of the given jurisdiction (e.g. a single state of a country being coded) are flagged as “targeted” while policies that apply to the whole jurisdiction are flagged as “general”. When both general and targeted policies exist simultaneously, OxCGRT always records the stricter policy. This choice may make the data more useful for evaluating the effect of policies on the spread of disease (since it records the stronger targeted measures that likely exist where there is a local outbreak) while reducing their ability to describe the overall state of policy across the country. For example, if a jurisdiction with many subunits has weak general policies and strong policies targeted at a single subunit, its overall coding will be high. In cases where this is frequently an issue, such as Brazil and the United States, OxCGRT has also comprehensively coded subunit jurisdictions (see Supplementary Figure 2). We encourage users to consider this granularity issue carefully when making cross-national comparisons, and to considering using subnational information for large, heterogenous jurisdictions where available.

Third, OxCGRT records policy interventions as a time series (the unit of observation is a “jurisdiction-day”) recording the intensity and scope of policy in place for a given indicator at that place and time. An alternative approach that has been pursued by comparable data projects is to record the start and end dates of individual policies (see more detailed comparison in methods section). While both options have merit, the time-series structure allows researchers to more easily match policy indicators to other time-series data, such as to case or death rates, mobile phone mobility data, and panel surveys. It also helps OxCGRT data collectors to capture government responses that do not take the form of discrete, formal policy interventions, but more ad hoc announcements, such as temporary limitations to internal movement during public holidays or religious events.

Fourth, the data are published continuously. Data-collection occurs weekly, which enables OxCGRT to provide up-to-date information on government responses. Given that SARS-Cov-2 can spread extremely quickly, this speed is essential for effective use of the data. However, our volunteers are not necessarily able to update every jurisdiction in every cycle, and consequently some data can be up to seven days out-of-date. To minimize recent gaps, we therefore publish our data in real-time so that it can be utilized as soon as it has been contributed. The trade-off to this speed is that the most recent data are published in advance of their final validation check (see the Methods section for further details), and may therefore be corrected in the review process or through external feedback (although in practice large revisions are rare; see methods section).

Fifth, OxCGRT relies on human judgement and contextual expertise, rather than automated data collection or coding, to provide the best possible degree of accuracy and consistency. The ordinal scales require individual contributors to carefully interpret various policies within each domain, in order to assign a code that best fits each indicator. For example, many countries have taken similar action to close workplaces, yet the types of workplaces that are required to close often differ from country to country. This means that each data contributor needs to assess the policy announcements in a country alongside detailed guidance material and apply judgment. Volunteers go through a training process to instill a high level of consistency and attend weekly meetings to discuss coding queries and standardize interpretations. Many of our contributors are specialists in the countries that they code, and understand the countries' culture, language and legal system in such a way that allows them to code with context and have access to local information to verify policies. While this "shoe leather science" approach is very human resource intensive, we have not found it possible to achieve comparable results with purely automated methods. Going forward, it may be fruitful to explore how different technological approaches can be combined with human coders.

Finally, and critically, the OxCGRT dataset records only the number and degree of government policies. It does not have a way to measure how well policies are implemented or enforced, nor does it measure the degree of compliance with official policies. OxCGRT data should therefore be considered one among several key elements in the broader puzzle of understanding governments policy adoption and the links between government interventions, human behavior, and the spread of COVID-19.

Methods

This section describes OxCGRT's design and structure, and the processes through which data are collected and confirmed. Because the project is continually evolving, adding further indicators and jurisdictions over time, users should always check the project website for the most current information. All OxCGRT data is available on GitHub and via an API and licensed under the Creative Commons Attribution CC BY standard.

We hope the methods outlined below will be of use not only to data users (the primary audience) but also to researchers who may be contemplating developing complementary measures or data collection projects for response to COVID-19 or other issues. In line with OxCGRT's "open source" ethos, we invite the scientific community to use and build on not just the data we collect but the methods and system described below.

<u>Indicators</u>

OxCGRT reports publicly available information on 19 indicators (see Table 1) of government response, as well as recording miscellaneous policies. The indicators capture all government measures related to a specific domain, including formally adopted laws, policies promulgated by executive or regulatory authorities, or softer guidance or advice. OxCGRT has added new indicators and refined old indicators as the pandemic has evolved· Future iterations may include further indicators or more nuanced versions of existing indicators. The indicators are of three types:

- **Ordinal**: These indicators measure policies on a simple scale of severity or intensity, allowing us to describe the degree or strength of government response in each category. For these indicators, the rank order of the different levels is significant, but we make no claims regarding the scale of the intervals. Instead, each level has a specific meaning, which allows the different values to also be used as categorical variables. These indicators are reported for each day a policy is in place (not the day it is announced).
    - Many have a further flag to note if they are "targeted", applying only to a sub-region of a jurisdiction, or a specific sector; or "general", applying throughout that jurisdiction or across the economy. For the newly added H7 (vaccination policy), the flag indicates whether the vaccine is being funded by the government or at cost to individuals.
- **Numeric**: These indicators measure a specific number, reporting fiscal values in USD. These indicators are only reported once, on the day they are announced.
- **Text**: This is a "free response" indicator that records other information of interest.

All observations also have a "notes" cell that reports sources and comments to justify and substantiate the designation.

<u>Data collection and reliability</u>

The initial set of data collectors in March 2020 were recruited largely from the postgraduate student body of the Blavatnik School of Government at the University of Oxford. Since then, additional contributors have been recruited through Oxford University departmental mailing lists, student societies and alumni email lists, as well as referrals from existing contributors. Subnational coders are mostly students or recent graduates from partner institutions in the countries we are collecting subnational data (for example the University of São Paulo, Fundação Getulio Vargas, and the State University of Pará in the case of Brazil). To date approximately 400 data collectors have contributed to OxCGRT.

New members of the data collection team undergo a series of training steps. First, they complete a self-directed tutorial of training slides and videos that explain how to search for data, interpret policies, and submit contributions through the online interface. New contributors are then given a short test for comprehension and understanding of the coding schema and collection process. After that, new data collectors are expected to attend a weekly all-contributor meeting, at which point they will start being included in the regular task allocation.

OxCGRT collects national data on a weekly schedule, during which new task allocations are sent to the data collection team. This allocation is based on a regular review of database coverage – prioritizing those countries that have not been updated within the last week. Most contributors are assigned to a list of 4-6 jurisdictions and will cycle through that list each allocation round, building up expertise in a small set of jurisdictions. The data is published in real-time as contributors enter it into the system.

After data is entered, it is marked ‘provisional’, which flags it for the review process. First, after each allocation round, a small team will do quick spot checks to ensure that data has been entered and there are no gross errors (for example, accidental deletion of a whole column can be noticed and fixed during this quick review). The provisional data is then queued for attention by a more thorough review team. This review team will examine the data entry and the original source, and either confirm its veracity or flag the data entry for escalation. The reviews process suggests a high degree of accurate in the initial data collection. As of 31 December 2020, 84.79 percent of all datapoints have never been changed, and, since 1 June 2020, 87.45 percent of data points have not required revision. Note these revisions include both post-hoc alterations to the coding scheme and factual errors. Meanwhile, just 0.41 percent of observations have been escalated by reviewers for adjudication, 0.25 percent since 1 June 2020. Of the 1.2 million data points captured between 1 June 2020 and 31 December 2020, 319,840 were reviewed or changed; of these 51 percent were confirmed without edits.

Data is collected from publicly available sources such as government press releases and briefings, international organizations reports, and trusted news articles. OxCGRT records the original source material using archived links so that coding can be checked and substantiated.

Coding different levels of government responses

OxCGRT includes data at country-level for nearly all countries in the world. It also includes subnational-level data for selected countries, currently Brazil (all states, the Federal District, state capitals and the next largest city that is not geographically connected to the state capital), the United States (all states plus Washington, DC and a number of territories), the United Kingdom (the four devolved nations), and Canada (all states and territories).

OxCGRT data are typically used in three ways. First, and primarily, to describe all government responses relevant to a given jurisdiction. Second, less commonly, to describe policies put in place by a given level and lower levels of government. And third, they are used to compare government responses across different levels of government.

To distinguish between these uses, different published versions of OxCGRT data are tagged in the database. The “TOTAL” label implies that all government responses relevant to the people in a given jurisdiction are included in the coding, regardless of whether those policies are set by national or subnational governments (these may also be presented without any jurisdiction label in some of our data products). The jurisdiction label “WIDE” refers to put in place by a given level and lower levels of government (eg. STATE-WIDE). “WIDE” observations therefore do not incorporate general policies from higher levels of government that may supersede local policies. For example, if a country has an international travel restriction that applies country-wide, this would not be registered in a STATE_WIDE record. The jurisdiction label GOV, indicates that observations include only policies instigated by a particular level of government; higher- or lower-level jurisdictions do not inform this coding.

In the main OxCGRT dataset, we show the total set of policies that apply to a given jurisdiction (the TOTAL policies described above). Specifically, in the main dataset, this means that we replace subnational-level responses with relevant national government (NAT_GOV) indicators when the following two conditions are met:

- The corresponding NAT_GOV indicator is general, not targeted, and therefore is applied across the whole country
- The corresponding NAT_GOV indicator is equal to or greater than the STATE_WIDE or STATE_GOV indicator on the ordinal scale for that indicator

In this way, national and subnational measures in the core dataset are comparable, in that they show the totality of policies in effect within a given jurisdiction.

Note that STATE_WIDE observations at the subnational level that code the totality of policies also capture policies that the national government may specifically target at a subnational jurisdiction. This is the case, for example, if a national government orders events to close in a particular city experiencing an outbreak. These kinds of policies are not inferred from NAT_GOV.

On our GitHub repositories, these different types data are available in three groups, summarized in Extended Data Figure 1:

1. The main repository: NAT_TOTAL for all countries, and STATE_TOTAL for Brazil, US, UK, and Canada
2. USA repository: NAT_GOV and STATE_WIDE
3. Brazil repository: NAT_TOTAL, NAT_GOV, STATE_WIDE, STATE_GOV, CITY_TOTAL, and CITY_WIDE

For large, heterogenous jurisdictions, users may wish to use a weighted average of subnational jurisdictions observations (e.g. STATE_WIDE) instead of national observations, NAT_TOTAL. See Supplementary Figure 2 for a comparison.

<u>Composite indices</u>

To make it easier to describe government responses in aggregate, OxCGRT calculates simple indices that combine individual indicators to provide an overall measure of the intensity of government response across a family of indicators. These indices are designed to provide a simple snapshot of the number and degree of government responses in a particular domain. Because we have not designed the indices for any specific analytic usage, we aim to make them as simple and transparent as possible. Those using the data to study the effect of government policies on outcomes of interest will therefore likely wish to modify the indices to suit the exact research questions they are seeking to answer (for example, selecting only the variables they believe to be relevant, or weighting those they believe to be of greater importance). In other words, we offer the indices as a convenient "prix fixe menu" option, but we urge users to tailor the data to their specific needs by ordering "a la carte."

As noted above, we stress that composite indices have strengths and weaknesses as descriptive and analytic tools. Governments' responses to COVID-19 exhibit significant nuance and heterogeneity. These issues create substantial measurement difficulties when seeking to compare national responses in a systematic way. Composite measures – which combine different indicators into a general index – inevitably abstract away from these nuances. Helpfully, cross-national measures allow for systematic comparisons across countries. By measuring a range of indicators, they mitigate the possibility that any one indicator may be over- or mis-interpreted. However, composite measures also leave out much important information, and make strong assumptions about what kinds of information "counts". If the information left out is systematically correlated with the outcomes of interest, or systematically under- or overvalued compared to other indicators, such composite indices may introduce measurement bias.

Broadly, there are two common ways to create a composite index: a simple additive or multiplicative index that aggregates the indicators, potentially weighting some; or a latent variable approach, in which observed indicators are used to predict an unobserved variable (i.e. the index). While there are several approaches to latent variable analysis, such as Factor Analysis, or Principal Component Analysis, Item Response Theory models are particularly suitable in this case due to the ordinal nature of most indicators. Each approach has advantages and disadvantages for different research questions.

OxCGRT uses simple, additive, unweighted indices because this approach is most transparent and easiest to interpret. Because the purpose of these indices is to describe the number and degree of government responses, we weight each indicator and each interval on the ordinal scale equally (within each indicator). In other words, the difference between a "1" and a "2" in a given indicator contributes as much to an index as the difference between a "2" and a "3". Again, this strong assumption will not be appropriate for all uses, so we encourage users to carefully consider what combinations and weightings of policies best capture the dimensions they are seeking to measure.

Despite this caveat, we find significant internal consistency within the indices. We used a latent variable approach, specifically Item Response Theory, as a robustness check for the Stringency Index (see Supplementary Table 3). IRT models have been extensively used in education to estimate the ability of a student (latent variable) based on the responses to individual test questions (observable indicators). In our case the individual policy levels (added to the geographic flag) are the observable indicators and the policy index, the unobservable variable. The scores generated by an IRT model were highly correlated to our linear index ($r = 0.98$), which reinforces the validity of our approach.

OxCGRT publishes four indices the group different families of indicators:

- Government response index (all categories)
- Stringency index (containment and closure policies, sometimes referred to as "lockdown" policies)
- Containment and health index (containment and closure and health policies)
- Economic support index (economic support measures)

Each index is composed of a series of individual policy response indicators. For each indicator, we create a score by taking the ordinal value and subtracting a half a point if the policy is targeted rather than general, if applicable. We then rescale each of these by their maximum value to create a score between 0 and 100, with a missing value contributing 0. These scores are then averaged to get the composite indices. This calculation is described in equation 1 below where k is the number of component indicators in an index and I_j is the sub-index score for an individual indicator.

We use a conservative assumption to calculate the indices. Where data for one of the component indicators are missing, they contribute "0" to the Index. An alternative assumption would be to not count missing indicators in the score, essentially assuming they are equal to the mean of the indicators for which we have data for. Our conservative approach therefore "punishes" countries for which less information is available, but also avoids the risk of over-generalizing from limited information.

$$(1) \qquad index = \frac{1}{k}\sum_{j=1}^{k} I_j$$

The different indices are comprised as described in Table 2.

To facilitate usage, two versions of each indicator are present in the database. A regular version which will return null values if there is not enough data to calculate the index, and a "display" version which will extrapolate to smooth over the last seven days of the index based on the most recent complete data.

Calculating sub-index scores for each indicator

All of the indices use ordinal indicators where policies are ranked on a simple numerical scale. The project also records five non-ordinal indicators – E3, E4, H4, H5 and M1 – but these are not used in our index calculations.

Some indicators – C1-C7, E1,H1, H6 and H7– have an additional binary flag variable that can be either 0 or 1. For C1-C7, H1 and H6 this corresponds to the geographic scope of the policy. For E1, this flag variable corresponds to the sectoral scope of income support. For H7, this flag variable corresponds to whether the vaccine is government funded or not.

The codebook has details about each indicator and what the different values represent. Because different indicators (j) have different maximum values (N_j) in their ordinal scales, and only some have flag variables, each sub-index score must be calculated separately. The different indicators are:

Each sub-index score (I) for any given indicator (j) on any given day (t), is calculated by the function described in equation 2 based on the following parameters:

- the maximum value of the indicator (N_j)
- whether that indicator has a flag (F_j=1 if the indicator has a flag variable, or 0 if the indicator does not have a flag variable)
- the recorded policy value on the ordinal scale ($v_{j,t}$)
- the recorded binary flag for that indicator ($f_{j,t}$)

This normalizes the different ordinal scales to produce a sub-index score between 0 and 100 where each full point on the ordinal scale is equally spaced. For indicators that do have a flag variable, if this flag is recorded as 0 (ie if the policy is geographically *targeted* or for E1 if the support only applies to *informal sector workers*) then this is treated as a half-step between ordinal values.

Note that the database only contains flag values if the indicator has a non-zero value. If a government has no policy for a given indicator (ie the indicator equals zero) then the corresponding flag is blank/null in the database. For the purposes of calculating the index, this is equivalent to a sub-index score of zero. In other words, $I_{j,t}$=0 if $v_{j,t}$=0.

$$(2) \quad I_{j,t} = 100\frac{v_{j,t} - 0.5(F_j - f_{j,t})}{N_j}$$

(if $v_{j,t}$=0 then the function F_j-$f_{j,t}$ is also treated as 0, see paragraph above.)

Data usage

The data is published in real time. Unless a country has been updated in the last 24 hours, there will be at least some gaps in coverage for the most recent days. In addition, if data is exported in the middle of an update, there can occasionally be missing data points in the time series. The dataset is also published with numbers of reported COVID-19 cases and deaths, drawn from open datasets at the European Centre for Disease Control and the John Hopkins University. Occasionally there have been missing days for some countries in these sources (for example, if a country has not updated their case data over a long weekend). For this reason, particularly when using the dataset for descriptive analysis, we usually interpolate to cover any single missing days, and use a ‘carryforward’ function to extend the latest value of a missing variable.

In addition, we caution users against over interpreting small fluctuations of single-digit changes in our index values. A small change in index value may not necessarily represent a substantive change in the country’s policy stance, it could, for example, just as easily represent a marginally different geographic coverage.

Comparison to related datasets

A number of datasets have tracked government response to COVID-19 since the start of the pandemic[21]. While it is beyond the scope of this article to describe all of them in detail, here we report similarities and differences with two sister projects, CoronaNet[19] and the Complexity Science Hub COVID-19 Control Strategies List (CCCSL)[22]. While these projects overlap with OxCGRT to some extent, allowing for direct comparisons, the three projects also offer complementary attributes, expanding the knowledge and options available to the research community.

The three projects have constructed datasets with a number of similar features but also points of difference.

Unit of analysis: Both CoronaNet and CCCSL record government policies or measures as the unit of analysis; OxCGRT instead uses the “jurisdiction-day.” While each approach can be converted into the other, the OxCGRT dataset is purpose-built as a panel. In contrast, CoronaNet and CCCSL are structured as an unbalanced panel, requiring additional steps to convert into a format that facilitates conventional analysis.

Coverage of jurisdictions and dates: The OxCGRT publishes data on 184 countries, and several subnational jurisdictions (50 United States, 13 Canadian Provinces and Territories, 27 Brazilian states and over 50 cities, and 4 UK devolved nations). CoronaNet publishes data on 195 countries, and the following subnational jurisdictions: Brazil, China, Canada, France, Germany, India, Italy, Japan, Nigeria, Russia, Spain, Switzerland and the United States. CCCSL publishes data on 56 countries, 33 of them European. All three datasets aim to update continuously, though at the time of writing only OxCGRT had up-to-date information for all jurisdictions.

Coverage of government responses: All three datasets broadly cover what we have termed closure and containment and health policies. In addition, OxCGRT and CCCSL record economic support measures. OxCGRT uniquely covers public transportation-related and vaccine policies. However, it does not include states of emergency or enforcement measures, as CoronaNet does, nor does it include receiving international help, measures to secure supply chains, crisis management plans, or port and ship restrictions, as CCCSL does.

Design of indicators: The OxCGRT 19 indicators are either ordinal or numeric, with an additional binary flag that records whether measures are general or targetted. CoronaNet considers different elements, such as the directionality of policies (e.g. inbound or outbound flights), the mechanism of travel (flights or trains), enforcement (mandatory or voluntary), and enforcers (national government or military). While the CCCSL covers fewer countries, their indicators are more granular split into four levels, without an ordinal scale. This more descriptive approach then needs further processing before it can be analyzed. While the detailed text descriptions enables rich qualitative analysis, it is less suited for quantitative analysis.

Data collection methods: All three datasets rely on hand-coded data entered by a large international pool of trained contributors into a central database. All three use publicly available sources, including policy documents and media reports. A key difference of the CoronaNet methodology is their use of the machine learning software instrument to extract data from news articles to aid contributors in their data collection. The CCCSL share their information sources in an open source Zotero library. From examination of the CoronaNet and CCCSL data and papers, it seems that OxCGRT is the only dataset to include archived weblinks to all original sources.

Looking at the data reveals further complementarities and differences between OxCGRT and related projects. OxCGRT most closely resembles CoronaNet, which also has global coverage for over 180 countries, and which produces a government policy activity index that can be compared quantitatively to the OxCGRT indices. Our database is highly correlated with CoronaNet within a given country. Supplementary Figure 3 shows the example of the United States, demonstrating how both indices track each other over time. Supplementary Figure 4 quantifies this relationship for all countries showing the average within country correlation between the CoronaNet and Oxford government response indices within a given country. The average correlation (Pearson's r) is high at .85. This suggests robustness across databases.

At the same time, the OxCGRT indices complements the CoronaNet index by providing greater comparability across countries. The CoronaNet index use item-response theory to model a given country's willingness to impose policies and adjusts for trends over time. This results in a relative rather than absolute measure. This results in scenarios, for example, where imposing a given policy when most countries have already imposed them will result in little change in the CoronaNet index. For example, both Nigeria and the United States closed primary schools in March, but the U.S. did so a few weeks earlier and has an index score that is a full standard deviation greater for the U.S. than Nigeria on the CoronaNet scale, despite being a similar policy action. This limits the ability to make cross-country comparisons on an absolute scale using the CoronaNet database. Figure 3 shows that the cross-country relationship between the Oxford and CoronaNet databases with a correlation (Pearson's r) of just 0.275 . This reveals that while the two database are highly consistent *within* countries, our database adds significant new information to also enable *across* country comparisons and analysis.

We note a few other distinctions. First, our absolute indices show variation more clearly. The CornaNet index falls, by and large, within 10 points on a 100-point scale with a standard deviation of 1.2 and stays nearly flat over time for an individual country. In contrast, countries in our database span the entire 100-point range across countries and over time with a standard deviation of 12. This granularity is particularly essential to capture important variation in waxing and waning of policies over time in addition to more sweeping lockdowns which can be captured with coarser measurement.

In summary, OxCGRT complements related efforts in a few dimensions. Our database has global coverage, enables comparable within and across country analysis, will be consistently updated and expanded, is publicly available, is built with a team of coders with contextual expertise in the respective countries in which they focus, and has a systematic panel data structure which has enabled merging with other databases and quantitative analysis.

Data availability

The most up-to-date OxCGRT data and documentation are available via the project github repository: https://github.com/OxCGRT/covid-policy-tracker.

References

1. Shuja, J., Alanazi, E., Alasmary, W. & Alashaikh, A. COVID-19 open source data sets: a comprehensive survey. *Appl Intell* (2020). https://doi.org/10.1007/s10489-020-01862-6

2. Tan-Torres, E. et al. Projected health-care resource needs for an effective response to COVID-19 in 73 low-income and middle-income countries: a modelling study. *Lancet Global Health* **8**(11), e1372-e1379 (2020).

3. Islam, N. et al. Physical distancing interventions and incidence of coronavirus disease 2019: natural experiment in 149 countries. *BMJ* **370,** m2743 (2020).

4. Hale T. et al. Global Assessment of the Relationship between Government Response Measures and COVID-19 Deaths. Preprint at: https://www.medrxiv.org/content/10.1101/2020.07.04.20145334v1 (2020).

5. Koh W. C., Niang, L. & Wong J. Estimating the impact of physical distancing measures in containing COVID-19: An empirical analysis. *Int. J. of Infect. Dis.*, **100,** 42-49 (2020).

6. Liu F., Wang M. & Zheng, M. Effects of COVID-19 lockdown on global air quality and health. Sci. *Total Environ*. **755,** 142533 (2020).

7. Venter, Z., Aunan, K., Chowdhury, S. & Lelieveld, J. COVID-19 lockdowns cause global air pollution declines. *Proc. Natl. Acad. Sci. U.S.A.* **117,** (2020).

8. Frey C.B., Chen C. & Presidente G. Democracy, Culture, and Contagion: Political Regimes and Countries Responsiveness to Covid-19. *Covid Economics* **18,** 1-20 (2020).

9. Hale, T. et al. Pandemic Governance Requires Understanding Socioeconomic Variation in Government and Citizen Responses to COVID-19. Available at SSRN: https://ssrn.com/abstract=3641927 (2020)

10. Pulejo, M. & Querubín, P. Electoral concerns reduce restrictive measures during the COVID-19 pandemic. *NBER Working Paper Series*, w27498 (2020).

11. Gottlieb C., Grobovsek J., Poschke M. & Saltiel F. Lockdown Accounting. IZA Discussion Paper No. 13397, Available at SSRN: https://ssrn.com/abstract=3636626 (2020)

12. Akter, S. The impact of COVID-19 related 'stay-at-home' restrictions on food prices in Europe: findings from a preliminary analysis. *Food Security* **12,** (2020).

13. Bonadio, B., Huo, Z., Levchenko, A. & Pandalai-Nayar, N. Global Supply Chains in the Pandemic. Centre for Economic Policy Research **3,** (2020).

14. Asfaw, A. Trade-offs Between Moral Hazard and Health: The Effect of Income Support Programs on Job Search and COVID-19. Available at SSRN: https://ssrn.com/abstract=3662711 (2020).

15. Berardi, C. et al. The COVID-19 pandemic in Italy: Policy and technology impact on health and non-health outcomes. *Health Policy and Technology*. https://doi.org/10.1016/j.hlpt.2020.08.019 (2020)

16. Michielsen, K. et al. International Sexual Health And REproductive Health (I-SHARE) survey during COVID-19: study protocol for online national surveys and global comparative analyses. *Sexually Transmitted Infections*. doi: 10.1136/sextrans-2020-054664 (2020).

17. Danquah, M., Schotte, S. & Sen, K. COVID-19 and employment: Insights from the Sub-Saharan African experience. *The Indian Journal of Labour Economics*, **63**, 23-30 (2020).

18. Sebhatu A., Wennberg K., Arora-Jonsson, S. & Lindberg, S.I. Explaining the homogeneous diffusion of COVID-19 nonpharmaceutical interventions across heterogeneous countries. *PNAS*, **117**(35), 21201–21208 (2020).

19. Cheng, C., Barceló, J., Hartnett, A., Kubinec, R. & Messerschmidt, L. COVID-19 Government Response Event Dataset (CoronaNet v.1.0). *Nature Human Behaviour* **4,** 756-768 (2020).

20. Oliver, N. et al. Mobile phone data for informing public health actions across the COVID-19 pandemic life cycle. *Sci. Adv.* **6,** eabc0764 (2020).

21. Daly, M., B. Ebbinghaus, L. Lehner, M. Naczyk, and T. Vlandas (2020). Oxford Supertracker: The Global Directory for COVID Policy Trackers and Surveys. Department of Social Policy and Intervention, University of Oxford.

22. Desvars-Larrive, A., Dervic, E., Haug, N. *et al.* A structured open dataset of government interventions in response to COVID-19. *Sci Data* **7,** 285 (2020).

Acknowledgements

Support for the Oxford COVID-19 Government Response Tracker has been provided by the Blavatnik Family Foundation, F. Hoffmann-La Roche AG, and the UK Cabinet Office. The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript. We thank Ethan Teo and Maria Luciano for valuable research assistance, and Katy Amar and Lucy Dixon for superb project management, Thomas Boby for database operations, as well as the Blavatnik School finance, communications, and IT teams. The project's Advisory Board has provided wise guidance throughout the development of this dataset. Most importantly, we are deeply grateful to the hundreds of volunteer data collectors who have contributed their time, energy, and expertise to sustaining the project. A complete and updated list of contributors can be found at: https://www.bsg.ox.ac.uk/research/research-projects/coronavirus-government-response-tracker.

Author contributions

Conceptualization: TH, BK, AP, TP, SW, EC-B, LH, SM, HT

Database design: TH, NA, BK, AP, TP, SW

Data collection: BK, EC-B, LH, SM, HT

Data analysis: TH, NA, RG, AP, TP

Project management: TH, BK, AP, TP, EC-B, LH, SM, HT

Competing interests

The authors declare no competing interests.

Figure legends

Figure 1. Containment and Health Index (CHI) by country on the first day of each month.

Figure 2. Clustering of initial government responses compared to the spread of COVID-19. This figure shows how 183 countries (each row), grouped by world region, ramped up their response policies to scores of 50 (out of 100) on the CHI within approximately the same fortnight in mid-March 2020 (as shown by the two vertical lines), despite the more scattered pattern of disease progression over time (as indicated by circles that mark the date a country experienced its tenth death) and in contrast to the greater divergence in policies observed in later months. Though the disease affected countries at different times, nearly all countries changed policy significantly in the same fortnight. The figure displays each country as a row and the x-axis dates are all in the year 2020.

Figure 3. Sequencing of initial policy adoption (from first policy enacted) and reduction (from first policy eased). This figure shows the ramping up (left panel) and rollback (right panel) of different kinds of government response policies. Days after first policy reduction is counted from the first reduction in the level of any policy for at least five consecutive days. A difference in the level of a policy line between the two panels in Figure 3—from 60 days into the ramp up (the far right of the left-hand panel) and the starting position of a policy line during rollback (the far left of the right-hand panel)—indicates the minimum percentage of countries that eased policies on the first day of policy reduction (if negative) or the percentage of countries adopting the policy after 60 days following the first measure, but before first policy reduction (if positive). For example, workplace closures have more commonly been eased at the initiation of rollback than regulations prohibiting public events. Countries that have increased policy intensity after five days of reducing policy strength are counted as countries that did not maintain their original maximum. The sample comprises 66,978 observations from 183 countries between 1 January 2020 and 31 December 2020.

Figure 4. Variation in the degree, geographic coverage, and permanence of policies. This figure shows the proportion of countries that, for various indicators, adopt any policies, adopt policies nationally (as opposed to geographically targeted policies), and adopt policies that correspond to the highest point on the ordinal scale. It also shows the percentage of countries that have at some point reduced and re-imposed such policies. A policy reduction is defined by a reduction in policy level sustained for at least five days. A policy reduction reversal is defined by an increase in policy level after a policy reduction that is sustained for at least five days. Adoption rates of some policies at any level (e.g. facial covering, income support and debt relief) are much higher than those depicted in the left panel of Fig.3. because countries often continue to adopt these policies after the first policy reduction. The sample comprises 66,978 observations from 183 countries between 1 January 2020 and 31 December 2020.

Figure 5. Associations between different combinations of government responses and aggregate population behaviour. The figure plots coefficients with 95% confidence intervals of the GRI, CHI and SI used as independent variables in separate panel regression models predicting changes in Google mobility data. The models use standard errors clustered by country and include country and date fixed effects (Supplementary Figure 1 plots coefficients of the same indices in models that control for daily deaths and Supplementary Tables 1 and 2 report full results). Note that change in the duration of time spent at home, as a proportion of the day, shows less variation than the other dependent variables, which capture change in the frequency of visits to different categories of locations.

Tables

Table 1: OxCGRT Indicators. See project website for detailed descriptions and coding information.

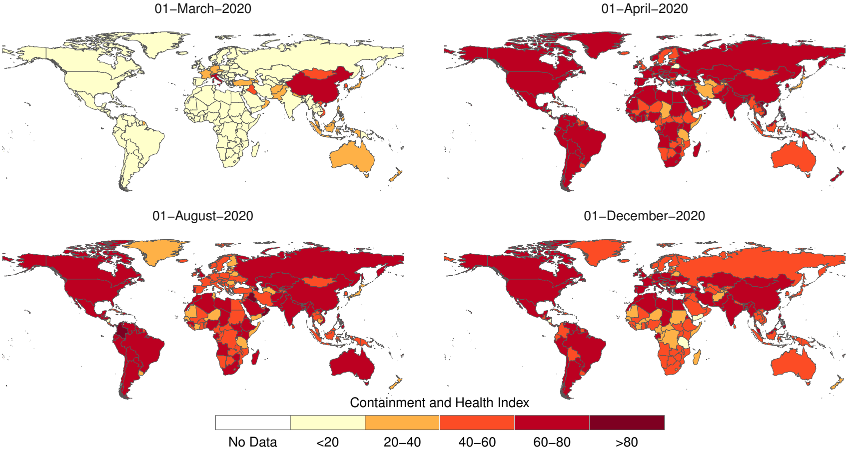
ID	Name	Type	Targeted/ General?
Containment and closure			
C1	School closing	Ordinal	Geographic
C2	Workplace closing	Ordinal	Geographic

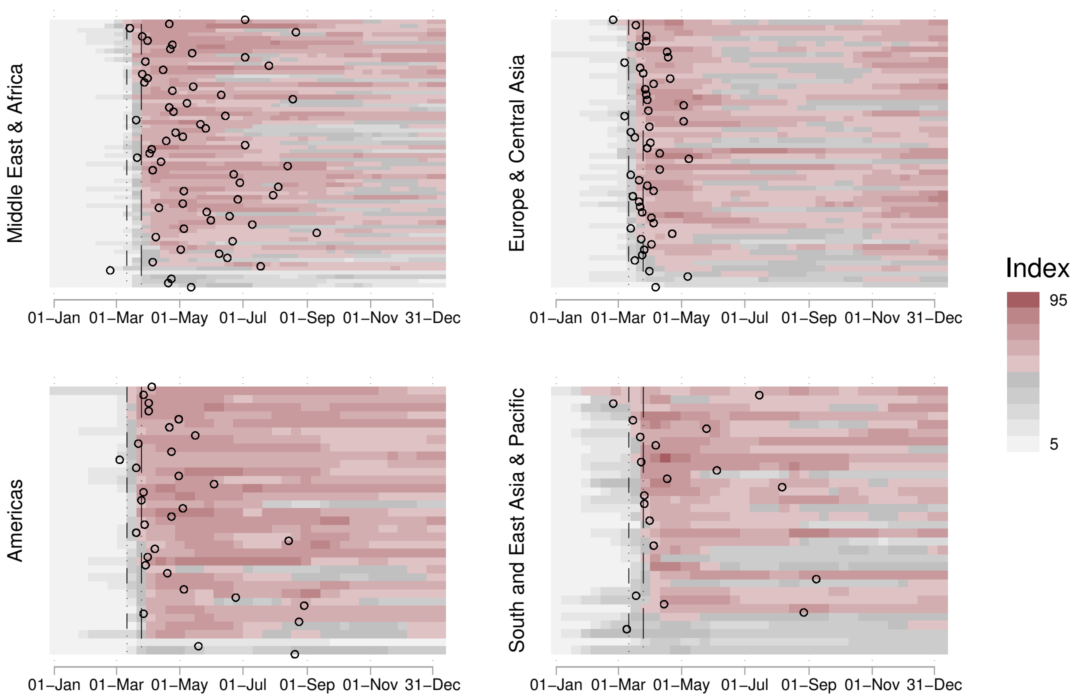
C3	Cancel public events	Ordinal	Geographic
C4	Restrictions on gathering size	Ordinal	Geographic
C5	Close public transport	Ordinal	Geographic
C6	Stay at home requirements	Ordinal	Geographic
C7	Restrictions on internal movement	Ordinal	Geographic
C8	Restrictions on international travel	Ordinal	No
Economic response			
E1	Income support	Ordinal	Sectoral
E2	Debt/contract relief for households	Ordinal	No
E3	Fiscal measures	Numeric	No
E4	Giving international support	Numeric	No
Health systems			
H1	Public information campaign	Ordinal	Geographic
H2	Testing policy	Ordinal	No
H3	Contact tracing	Ordinal	No
H4	Emergency investment in healthcare	Numeric	No
H5	Investment in Covid-19 vaccines	Numeric	No
H6	Facial Coverings	Ordinal	Geographic
H7	Vaccination Policy	Ordinal	Funding
Miscellaneous			
M1	Other responses	Text	No

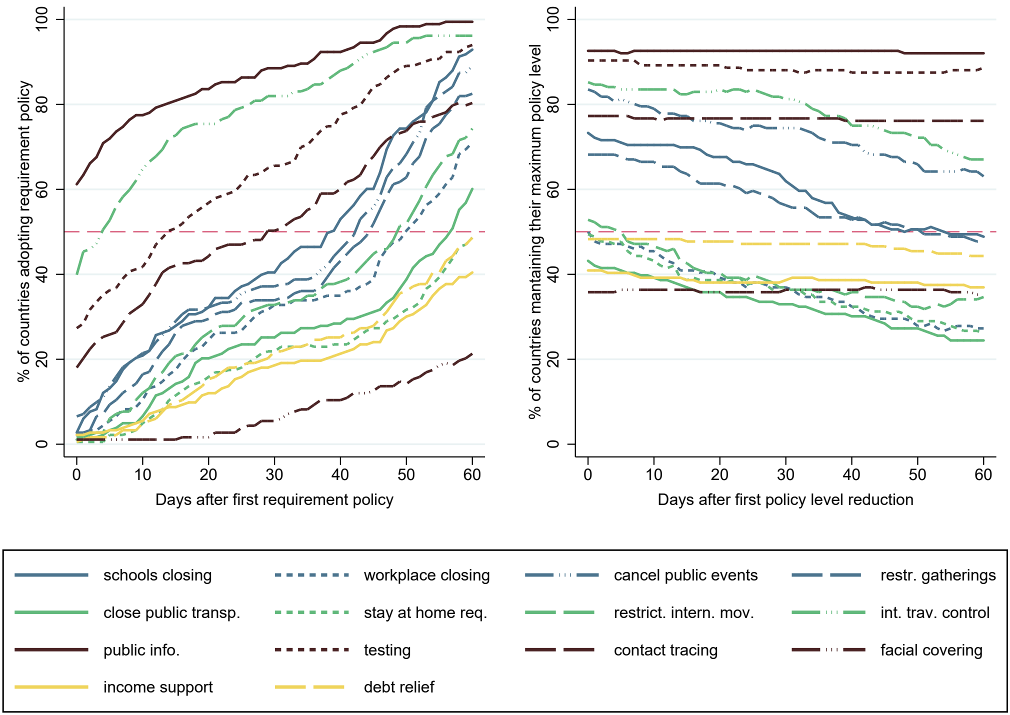
Table 2: OxCGRT index composition and values.

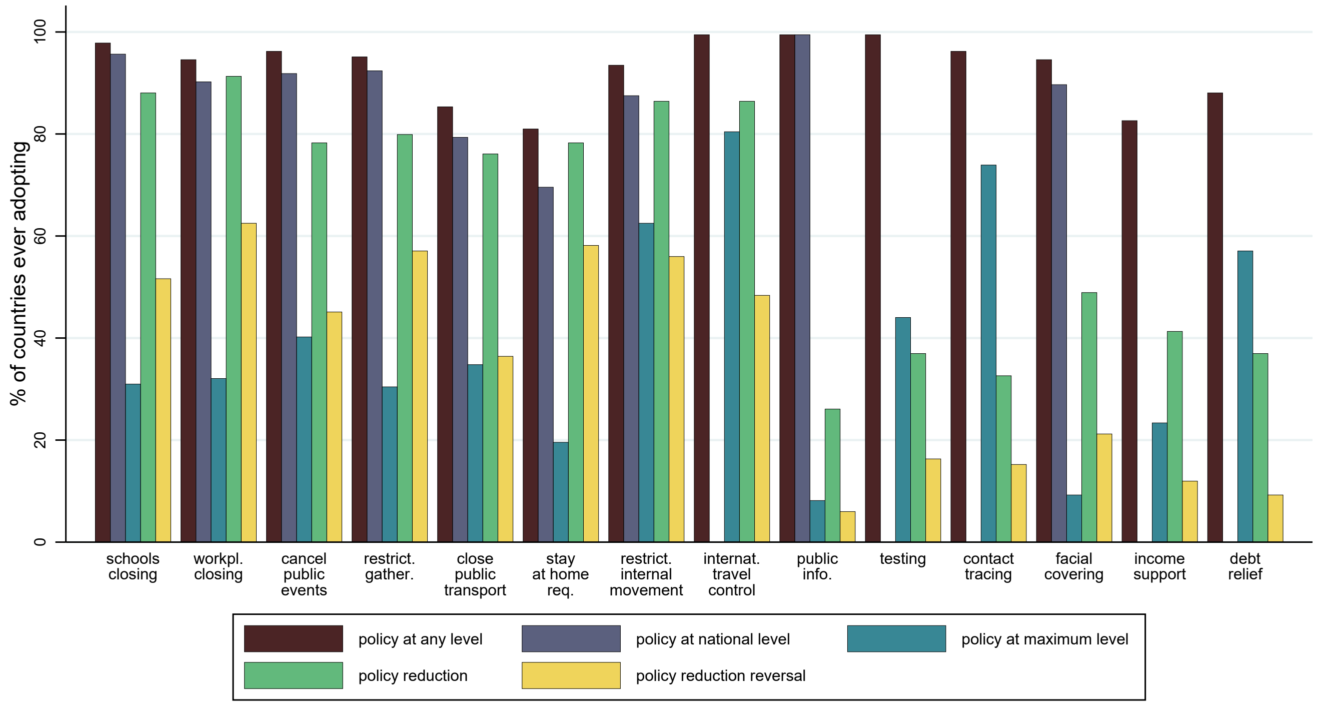
	Government response index	Containment and health index	Stringency index	Economic support index	Legacy stringency index	Max. value (N_j)	Flag? (F_j)
k	13	11	9	2	7		
C1	x	x	x		x	3 (0, 1, 2, 3)	yes=1
C2	x	x	x		x	3 (0, 1, 2, 3)	yes=1
C3	x	x	x		?	2 (0, 1, 2)	yes=1
C4	x	x	x		?	4 (0, 1, 2, 3, 4)	yes=1
C5	x	x	x		x	2 (0, 1, 2)	yes=1
C6	x	x	x		?	3 (0, 1, 2, 3)	yes=1
C7	x	x	x		?	2 (0, 1, 2)	yes=1
C8	x	x	x		x	4 (0, 1, 2, 3, 4)	no=0
E1	x			x		2 (0, 1, 2)	yes=1
E2	x			x		2 (0, 1, 2)	no=0
E3							
E4							
H1	x	x	x		x	2 (0, 1, 2)	yes=1
H2	x	x				3 (0, 1, 2, 3)	no=0
H3	x	x				2 (0, 1, 2)	no=0
H4							
H5							

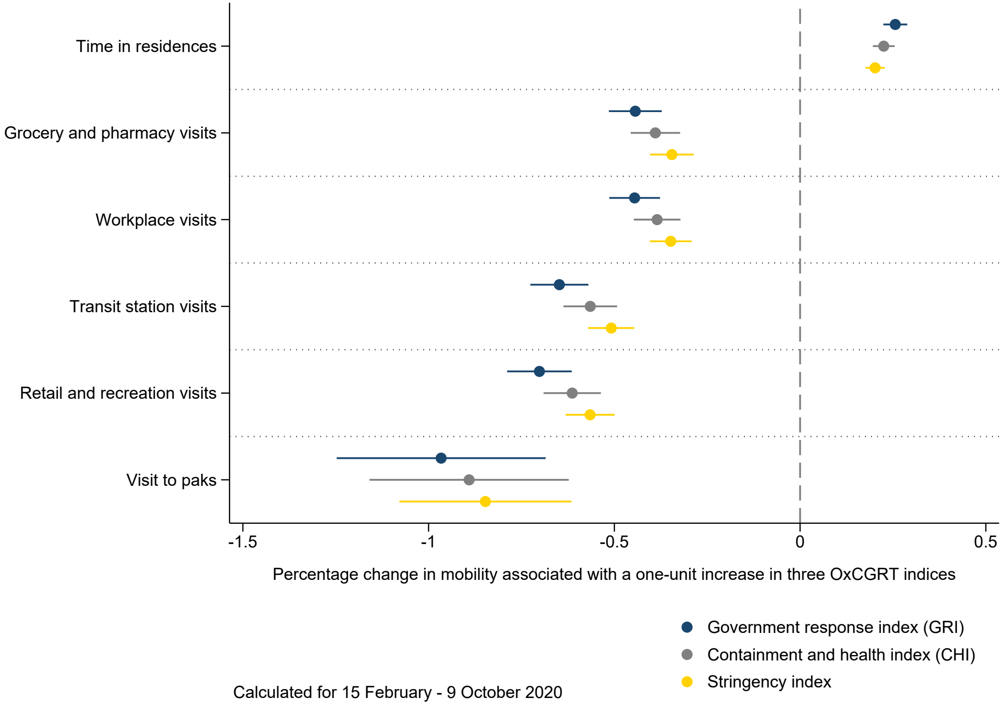
H6	x	x				4 (0, 1, 2, 3, 4)	yes=1
H7	x	x				5 (0, 1, 2, 3, 4, 5)	yes=1
M1							


01−March−2020
01−April−2020
01−August−2020
01−December−2020
Containment and Health Index
No Data
<20
20−40
40−60
60−80
>80


Middle East & Africa
Europe & Central Asia
Americas
South and East Asia & Pacific
01–Jan 01–Mar 01–May 01–Jul 01–Sep 01–Nov 31–Dec
Index
95
5


% of countries adopting requirement policy
Days after first requirement policy
% of countries mantaining their maximum policy level
Days after first policy level reduction
0
20
40
60
80
100
0
10
20
30
40
50
60
schools closing
workplace closing
cancel public events
restr. gatherings
close public transp.
stay at home req.
restrict. intern. mov.
int. trav. control
public info.
testing
contact tracing
facial covering
income support
debt relief


% of countries ever adopting
0
20
40
60
80
100
schools closing
workpl. closing
cancel public events
restrict. gather.
close public transport
stay at home req.
restrict. internal movement
internat. travel control
public info.
testing
contact tracing
facial covering
income support
debt relief
policy at any level
policy at national level
policy at maximum level
policy reduction
policy reduction reversal


Time in residences
Grocery and pharmacy visits
Workplace visits
Transit station visits
Retail and recreation visits
Visit to paks
-1.5
-1
-0.5
0
0.5
Percentage change in mobility associated with a one-unit increase in three OxCGRT indices
Government response index (GRI)
Containment and health index (CHI)
Stringency index
Calculated for 15 February - 9 October 2020

	TOTAL[1]	WIDE	GOV
National	180+ countries	Not applicable[2]	• USA federal government • Brazilian federal government • UK government (Westminster) • Canada federal government
State	• USA: 50 states and Washington DC • Brazil: 26 states and the Federal District • UK: 4 devolved nations • Canada: 13 provinces and territories	• USA: 50 states and Washington DC • Brazil: 26 states and the Federal District • UK: 4 devolved nations • Canada: 13 provinces and territories	• Brazil: 26 states and the Federal District
City	• Brazil: 26 state capital cities and 26 second cities	• Brazil: 26 state capital cities and 26 second cities	Not applicable[3]