Omega 122 (2024) 102966

Contents lists available at ScienceDirect

Omega

journal homepage: www.elsevier.com/locate/omega

International comparisons of COVID-19 pandemic management: What can be learned from activity analysis techniques?[☆,☆☆]

Víctor Giménez[a,1], Diego Prior[a,1], Claudio Thieme[b,1], Emili Tortosa-Ausina[c,1,*]

[a] Department of Business, Universitat Autònoma de Barcelona, 08193 Bellaterra (Barcelona), Spain
[b] School of Economics and Business, Universidad Diego Portales, Av. Santa Clara 797, Huechuraba, Santiago de Chile, Chile
[c] Department of Economics, IIDL and Ivie, Universitat Jaume I, Campus del Riu Sec, 12071 Castelló de la Plana, Spain

ARTICLE INFO

JEL classification:
C61
H51
I18

Keywords:
Data envelopment analysis
COVID-19
Composite indicators
Assurance region
Robust conditional convex frontier models

ABSTRACT

The global spread of COVID-19 since early 2020 has resulted in significant humanitarian costs. The pandemic has affected most countries to varying degrees, and governments have implemented diverse policies to minimise the impact on public health. However, these policies have varied across regions and even within countries. This study proposes a nonparametric activity analysis methodology to assess how different countries have managed the pandemic. Specifically, we assess the effectiveness of 61 countries nine months into the pandemic using a robust directional Benefit of Doubt (BoD) model according to expert opinion and conditional on country contextual factors. We then estimate the marginal impact of structural and discretionary contextual variables on effectiveness using nonparametric regression analysis, which shows that effectiveness is strongly influenced by socioeconomic and cultural factors. The results reveal three main groups of countries according to their level of effectiveness in pandemic management, suggesting that an accurate assessment of countries' management of the pandemic benefits greatly from operations research methods, as we obtain benchmarks and find out how these benchmarks (or best practices) vary when contextual factors are included in the analysis.

1. Introduction

The spread of the SARS-CoV-2 virus and its associated COVID-19 disease is the largest public health emergency since the 2002–2004 SARS outbreak in China, becoming a global health threat with unprecedented socioeconomic effects around the world. As the World Health Organisation (WHO) Emergency Committee on COVID-19 forecasts the pandemic will not be short lived, having evolved from a public health event into a transnational crisis, continued efforts to fend it off at the community, regional, national and global level should be sustained during the years ahead.

The unprecedented impact of the pandemic in most countries of the world has triggered rapid responses from multiple tiers of society, including governing bodies (both national and sub-national, depending on each country's level of decentralisation) and international organisations, in general, as well as professional and other communities. The scientific community has not remained on the sidelines, but the reactions across disciplines are notably divergent—partly due to their varying relationships with the causes and consequences of the pandemic.

Some of the scientific disciplines involved in evaluating COVID-19-related issues emerge naturally; epidemiology (prevention), medicine (treatment) and pharmacy (vaccines) stand out for obvious reasons to do with the public health implications of the pandemic. However, the social sciences are also playing a role, since epidemiologists' recommendations for lockdowns [1–4] and stay-at-home orders inevitably wreaked unprecedented economic havoc in many countries, particularly in those Western economies where the toughest stringency policies were implemented [5].

☆ Area: Data-Driven Analytics. This manuscript was processed by Associate Editor Victor Podinovski.

☆☆ All four authors are grateful to 145 respondents of the survey conducted for the study. We are particularly grateful to two anonymous reviewers, whose comments have contributed to the overall improvement of the paper. Emili Tortosa-Ausina acknowledges financial support by MCIN/AEI/10.13039/501100011033 (Ministerio de Ciencia e Innovación, PID2020-115450GB-I00) and Universitat Jaume I, Spain (UJI-B2020-27). Víctor Giménez and Diego Prior acknowledge financial support from AGAUR , reference 2021SGR00719. The usual disclaimer applies.

* Corresponding author.
E-mail addresses: victor.gimenez@uab.cat (V. Giménez), diego.prior@uab.cat (D. Prior), claudio.thieme@udp.cl (C. Thieme), tortosa@uji.es (E. Tortosa-Ausina).

[1] Authors contributed equally to the different aspects of the article.

https://doi.org/10.1016/j.omega.2023.102966
Received 13 November 2022; Accepted 9 September 2023
Available online 19 September 2023

Despite the substantial controversy over how lasting and restrictive the measures need to be, as well as their effectiveness,[2] there is no doubt that their impact on societies is huge and their effects manifold; these effects can be evaluated from the perspective of several social science disciplines. Among them, economics is at the forefront, due to the unprecedented decline in GDP caused by the imposition of harsh and lasting lockdowns and other social-distance measures (including closures in the hospitality sector, sports facilities, etc.).[3] This unprecedented blow to the world's economies has resulted in an equally unprecedented response from economics, both at the practitioner and academic levels.[4]

The range of economics-related issues studied has been wide, including not only obvious fields such as business [9], accounting [10] or banking and finance [11], but also interactions with other social science disciplines such as education [12] and geography [13–15], to name a few—although the links with psychology [16], sociology [17] and political science [18] are also evident. The literature from this economics perspective is burgeoning, since it now includes not only the working papers and other documents disseminated by prestigious institutions,[5] but also publications by respected journals in the field, dealing with issues related to the economic consequences of COVID-19 and government responses.[6]

It seems clear that a successful response to COVID-19 requires *coordinated* strategies from both epidemiologists and economists, since their advice might be mutually exclusive—i.e., lockdowns and business closure mandates are at odds with economists' advice to avoid recessions [19,20].[7] Some efforts have already attempted to understand how these coordinated strategies can help to improve governments' responses to COVID-19 [6], to ultimately provide joint advice on the benefits and costs of the different policies [see also21].

Each government, including sub-national levels within a country,[8] chose its own mix of social-distancing measures [22], which might include restrictions on mobility (on space and time, i.e., weekdays and weekends), curfews, schools and business closures, travel restrictions, compulsory face masks, different testing and contact-tracing strategies, etc., resulting in very heterogeneous combinations.[9] This is, *a priori*, surprising, since most territories implementing different measures are part of the same trade bloc (such as the European Union) or even the same country. Some explanations for these mixes might be related to health costs and temporary unemployment schemes covered by different layers of government, which could advocate opposing measures.

We argue that this variety of responses to the COVID-19 pandemic might be resulting in different degrees of success in governments' handling of the pandemic [24]. However, it is no easy task to measure governments' performance in tackling the pandemic. Nonetheless, this is a critical issue, since measuring the different performances with a certain degree of precision would identify some benchmarks or *best-practices*. In fact, there is a considerable body of literature concerned with benchmarking and best practice: the activity analysis literature [25] dealing with performance and efficiency evaluation via frontier methods. This literature is closely related to operations research, since many of the methodologies proposed to evaluate the performance of organisations (including government and public policies) and decision-making units (DMUs) in general come from this field.

We consider a broader view, with the aim of evaluating countries' responses to the pandemic. As indicated above, the approaches adopted by countries and subnational governments within countries have been heterogeneous in terms of timings and extents of lockdowns, processes to reopen the economy etc., and views about the onset of successive waves have been disparate and confusing [26]. As a consequence, the performances of different countries and regions also varied. Therefore, in order to identify the policies that work best, it is also necessary to identify the best practices or benchmarks, a task in which operations research methodologies have a long tradition of providing appropriate responses.

Consistent with this, several studies have analysed the performance of COVID-19 management by the health systems of different countries using activity analysis techniques. A review of the literature using Web of Science identifies at least 13 articles along these lines. All of them analyse performance from the perspective of efficiency, mainly using data envelopment analysis (DEA) models [27–37], and in specific cases stochastic frontier analysis (SFA) [38] and weighted stochastic imprecise data envelopment analysis (WSIDEA) [39].

This review reveals a significant number of studies using models that measure a country's performance in different phases of the pandemic [30,34,35,37]. For example, Lupu and Tiganasu [30] evaluate the performance of 31 European countries. They use DEA performed in three stages of the pandemic (first wave, relaxation and second wave), finding high inefficiency in the first phase, mainly in Western countries, which improved significantly afterwards. A second phase analysis shows that the factors influencing performance differ according to the phase of the pandemic in the country. Ordu et al. [35] use super efficiency DEA methods to assess the performance of 16 countries in each of the first five weeks of the pandemic. Their results evidence a sharp decrease in efficiency over time, and unlike most of the European countries evaluated, China and South Korea show an increase in efficiency. Ibrahim et al. [37] use DEA to evaluate the performance of 58 countries in two phases of the pandemic (pandemic control and treatment measurement), and consequently use two models. The outputs they used for the pandemic control phase were COVID-19 confirmed cases and for the treatment measurement phase COVID-19 related deaths and COVID-19 recovered cases. Their evaluation shows that 89.6% of the countries are inefficient in pandemic control and 79% in treatment measures. A sensitivity analysis highlights the critical nature of resources.

Similarly, two of these studies apply two-stage network DEA models, which use intermediate variables that act as hinges, being both an output of the first node and an input of the second node [27,28]. For example, Klumpp et al. [27] use a two-stage network DEA model to measure the efficiency of the health systems of 19 OECD countries over 12 time periods. Their model relates the number of tests applied, cases and deaths. The results indicate that, on the one hand, factors such as population size, population density and the country's stage of development did not play an important role in the success of pandemic management. Pre-pandemic health system policies were decisive, however. Health systems with a primary care orientation and a high ratio

[2] E.g., see the discussion on mask wearing in [6].

[3] These measures are likely to have other side effects in the medium to long term, including psychological issues among children due to school closures [7]. Although these are more difficult to measure, some studies are already quantifying those more directly related to the impact on the economy [8].

[4] Apart from the contributions published in academic journals, see also the "Covid Economics" initiative by the Centre for Economic Policy Research (CEPR, https://cepr.org/content/covid-economics-vetted-and-real-time-papers-0, "created to quickly disseminate fast-rising scholarly work on the Covid-19 epidemic", which has published several hundreds of documents on different aspects related to the economics of COVID.

[5] See https://cepr.org/content/covid-economics-vetted-and-real-time-papers-0.

[6] The number of contributions is increasing so rapidly to the point that a survey of the literature has already been published [3].

[7] One might also wonder about how much of the collapse resulted from government-imposed restrictions or people voluntarily choosing to stay at home, and some studies have found legal shutdown orders contribute modestly to changes in consumer behaviour [1].

[8] In several countries, decentralisation of powers has resulted in several competencies, particularly health and education, being transferred to sub-national layers of government.

[9] Enforcing measures also vary greatly from country to country [23].

of primary care physicians to specialists proved to be more efficient than systems with a medium level of resources that were partly publicly funded and characterised by a high level of access regulation. From an economic perspective, strategies without general blockages were identified as more efficient than the total blockage strategy.

In other aspects, these studies incorporate contextual variables directly as inputs to the evaluation model. The inputs used in many of these studies are non-discretionary contextual variables or short- or medium-term discretionary contextual variables. Examples include the studies by (i) Min et al. [28], who use population size, per capita gross national income; (ii) Doğan et al. [32], who include population density, population aged 65+, chronic diseases; (iii) Ordu et al. [35], who incorporate population, percentage of population aged 70+, median aged; (iv) Aydin and Yurdakul [39], who include a stringency index, extreme poverty, CVD death rate, diabetes prevalence, female smokers, male smokers, population, GDP; and (v) Ibrahim et al. [37], who incorporate population density, percentage of population aged 65+. Although we take a different methodological approach to the above studies, they do demonstrate the importance of incorporating this type of variable in order to compare countries adequately.

Finally, although many studies have attempted to evaluate country performance in managing the pandemic, very few attempt to explain the differences in performance found. In fact, only four of these studies provide a second-stage evaluation in this regard. These include [27], who perform a correlation analysis; and [28,30,36], who use Tobit regression as a *post hoc* analysis. For example, Min et al. [28] study the performance of 34 OECD member countries, relating health system performance outcomes to cultural aspects of the country. They use two-stage network DEA and conclude that a pervasive, less individualistic and uncertainty-avoiding culture positively influenced country performance.

In summary, a body of literature has evaluated the performance of health systems in different countries during the COVID-19 pandemic from an efficiency perspective. The diversity of variables used to model the phenomenon and evaluate the pandemic in its different stages reflects the diversity of research objectives pursued. However, more studies are still needed to understand the considerable complexity of this phenomenon and the impact of contextual variables on country performance. Similarly, to the best of our knowledge, no previous studies consider conditional models to evaluate health systems' performance of COVID-19 management in different countries from an effectiveness perspective.

In light of the above, in this paper we use composite indicators to evaluate the effectiveness with which countries manage the pandemic; these indicators are particularly appropriate for assessing complex phenomena that cannot be evaluated with a single variable [40], as in the case we are dealing with. Specifically, we propose a composite indicator combining several proposals in the field of nonparametric activity analysis methods: directional distance functions, benefit of the doubt, and robust conditional convex order-*m* estimations, a design which has been considered up to now in several studies, including [41–43], among others. In our case, we will also consider the opinion of the experts when choosing the directional vector of the distance functions, proposing a new way to determine the corresponding weights using Best–Worst Scaling. The model is refined in a second stage by considering a conditional model in which two types of variables enter the analysis, namely, structural contextual variables and discretionary contextual variables. The inclusion of these successive stages in the analysis is essential due to the existence of factors lying beyond governments' control (i.e., they are non-controllable) and that might clearly affect countries' performance, as well as other factors over which some control can be exercised (i.e., they are controllable). Since these factors differ in nature, they enter the model sequentially rather than simultaneously so as to assess their effect more precisely

The paper proceeds as follows. After this introduction, Section 2 presents the methodology and Section 3 describes the data. Section 4 then reports the results, and finally, Section 5 presents the conclusions of the study and implications for public policy.

2. The proposed methodology

2.1. Effectiveness composite indicator: directional BoD model

The evaluation of effectiveness uses different methods to create composite indicators (CI). These indicators have become popular and widely used metrics for assessing complex phenomena which cannot be evaluated with a single variable [40] as they globalise a process composed of multiple stages that generate different sub-indicators. Most of the methods available to build CI require previous information on the sub-indicators in order to determine the weight assigned to each of them [44]. Here we take an alternative approach by using data envelopment analysis methods in order to obtain endogenous weights without the need to require any previous information [45]. Specifically, we propose a composite indicator that combines alternative developments from the research on nonparametric activity analysis, namely, benefit of the doubt (BoD) [46–54]. BoD models are a variant of the nonparametric frontier models used to measure efficiency [55] where only outputs are considered and have been widely applied to construct composite indicators in different contexts. Some examples of applications are the Health System Performance Index [56], the Human Development Index [49], the Internal Market Index [46], the Technology Achievement Index [48], the Students' Evaluation of Teaching Indicator [50], the Competitiveness Index [57], the Quality of Life Indicator [52], the Digital Access Indicator [51], education system performance [58], Citizen Satisfaction with Local Police Effectiveness [59], the Life Satisfaction Index [60], the Sustainable Energy Index [45], the Environmental Performance Index [54], the Well-Being Indicator [61], spatial directional robust composite indicators in the presence of undesirable output [62] and the association between public spending effectiveness and the ideological orientation of regional governments [41].

There are three relevant issues to be considered when applying this type of model to our case. First, the compensability among different sub-indicators. The traditional BoD model assumes absolute freedom in determining the endogenous weights and, therefore, the total compensability/trade-off between sub-indicators. However, this assumption is not applicable when information regarding the preference structure is available from experts, as occurs in our case. For these cases, Fusco [63] proposes including a "directional" penalty for the indicators considered less important according to the experts' judgement using a directional distance function (DDF) [64,65]. This type of distance function has been widely used to evaluate environmental efficiency where good/bad outputs to be maximised/minimised coexist [54,66–68]. Consequently, DDFs are especially suitable for evaluating a set of indicators that improve by increasing their value simultaneously with others that have a contrary behaviour. This is common in the health field where it is desirable to minimise, for example, the number of deaths. The second issue to consider is the need to incorporate the environmental or contextual characteristics of each country to obtain the fairest evaluation possible. Finally, the risk of possible outliers increases when data from different countries is used. For all these reasons, we propose using a BoD model that is directional (incorporating expert opinion), robust (insensitive to the presence of outliers) and conditional (considering the contextual factors of each country). Regarding the impact of context on the evaluation, we also propose estimating the marginal impact of the structural contextual variables (those that, having an impact on the levels of effectiveness, are not discretionary as their levels are controlled by the authorities) and the marginal impact of the discretionary contextual variables (those that, having an impact on the levels of effectiveness, are discretionary because their level depends on decisions taken by the respective governments).

Let us assume that, for K countries, we have information on a set of J *desirable or good* indicators to maximise, $y \in \mathbb{R}_+^J$, as well as H

undesirable or *bad* indicators to minimise, $b \in \mathbb{R}_+^H$. The performance of any country can be measured through the following DDF [67,69,70]:

$$D(y,b) = max \left(\beta \mid \left(y + \beta g_y, b - \beta g_b\right)\right) \tag{1}$$

The above DDF determines respectively the maximum simultaneous increase and decrease (β) for the indicators y and b over the vector $g = (g_y, g_b) \in \mathbb{R}_+^{J+H}$ which defines the desirable directions for improvement for both types of indicator. Various methods can be used to calculate $D(y,b)$. In this study we use a nonparametric frontier model based on [67], although,to be coherent with the BoD formulation, inputs have disappeared. $D(y,b)$ is calculated by solving the following linear program.[10] for the evaluated country o:

$$\begin{aligned} &Max\ \theta(y_{oj}, b_{oh}) = \beta \\ &s.t. \quad \begin{array}{ll} \sum_{k=1}^{K} \lambda_k\, y_{kj} \geq y_{oj} + \beta g_y & j = 1 \ldots J, \\ \sum_{k=1}^{K} \lambda_k\, b_{kh} \leq b_{oh} - \beta g_b & h = 1 \ldots H, \\ \lambda_k \geq 0 & k = 1 \ldots K. \end{array} \end{aligned} \tag{2}$$

where y_{kj} represents the indicator j to be maximised for country k; b_{kh} is the indicator h to be minimised for country k; y_{oj} and b_{oh} are the observed levels of each indicator, respectively, for the country under assessment. β is the maximum simultaneously achievable increase (decrease) in the indicators to be maximised (minimised). Note that in the case of a country that has managed the pandemic in both dimensions effectively, $\beta = 0$.

To optimise the linear program (2), the vector $g = (g_y, g_b)$ must be defined. The literature usually follows [66] by defining $g_y = (y_{o1}, y_{o2}, \ldots, y_{oJ})$ and $g_b = (b_{o1}, b_{o2}, \ldots, b_{oH})$, assuming compensability among indicators.[11] However, when this assumption is not adequate because information is available on the importance of the indicators, Fusco [63] proposes weighing their components by defining $g_y = (y_{o1}w_1, y_{o2}w_2, \ldots, y_{oJ}w_J)$ and $g_b = (b_{o1}v_1, b_{o2}v_2, \ldots, b_{oH}v_H)$.

In our estimations, we propose using the Shepard output distance for the CI computation [42,63,72], since it is more easily interpreted:

$$CI(y_{oj}, b_{oh}) = \frac{1}{1 + \theta(y_{oj}, b_{oh})} \tag{3}$$

where $CI(y_{oj}, b_{oh}) \leq 1$, the higher the value of $CI(y_{oj}, b_{oh})$, the higher the level of effectiveness. Consequently, the countries with the best performance in managing the pandemic will obtain $CI(y_{oj}, b_{oh}) = 1$.

2.2. Conditional model

Previous models are labelled as *unconditional* as they do not restrict the comparison of units with significant differences in their environmental conditions. This implies that the value of the CI for inefficient units can be underestimated with regard to those units operating in a more favourable environment. To control for these circumstances, Daraio and Simar [73] developed a *conditional* approach to account for those variables in the same optimisation model. The basic rationale behind conditional models is to compare the unit under analysis with other units that present similar levels in their contextual or environmental factors. These conditional models are particularly relevant when dealing with units facing disparate environmental conditions, which is the case of cross-country studies.

The conditional approach has become very popular in the literature on performance measurement [e.g., 74,75].[12] To estimate the conditional scores, smoothing techniques are needed such that, in the reference samples, observations with comparable z-values have to be chosen. Therefore, this approach relies on the estimation of a nonparametric kernel function to select the appropriate reference partners and a bandwidth parameter b using a method with some bandwidth choice. As we use continuous z variables, we apply the method proposed by Bădin et al. [77], based on the least squares cross validation procedure (LSCV) developed in [78–80].

Following [81], we use the Epanechnikov kernel density function $K_{\hat{b}}(z, zi) = b^{-1}K[(Z - z_i)/b]$ where b is the smoothing parameter, or bandwidth. The conditional estimation implies the restriction of the technology to consider only those units that happen to have their contextual variables in between the two extremes delimited by the respective bandwidth:

$$(z - b) \leq z^o \leq (z + b) \tag{4}$$

The conditional formulation of program (2) is:

$$\begin{aligned} &Max\ \theta(y_{oj}, b_{oh}|z^o) = \beta \\ &s.t. \quad \begin{array}{ll} \sum_{k'=1}^{K'} \lambda_{k'}\, y_{k'j} \geq y_{oj} + \beta g_y & j = 1 \ldots J, \\ \sum_{k'=1}^{K'} \lambda_{k'}\, b_{k'h} \leq b_{oh} - \beta g_b & h = 1 \ldots H, \\ \lambda_{k'} \geq 0 & k' = 1 \ldots K'. \end{array} \end{aligned} \tag{5}$$

where k' only includes those units with contextual variables meeting condition (4), i.e.:

$$(z_{k'} - b) \leq z^o \leq (z_{k'} + b) \tag{6}$$

being the composite indicator:

$$CI(y_{oj}, b_{oh}|z^o) = \frac{1}{1 + \theta(y_{oj}, b_{oh}|z^o)} \tag{7}$$

After estimating the unconditional and conditional directional BoD models, it is possible to define the ratio Q^z to show the dependence on the contextual variables of the composite indicator by using a smoothed nonparametric regression [82] :

$$Q^z = \frac{CI(y_{oj}, b_{oh})}{CI(y_{oj}, b_{oh}|z^o)} \tag{8}$$

In our empirical application, contextual variables (z) can be classified as structural (or non-discretionary, z^s) and discretionary (z^d). A similar approach can be found in [42], who made the distinction between variables which can only be considered as exogenous in the short term and variables which can be considered as exogenous in both the short and the long term. According to our classification, we define a recursive process to determine the marginal importance of each group of variables by estimating model (5) twice. In the first step, we introduce only the structural environmental variables (z^s), which serves to determine the impact on CI due to the presence of the structural context variables:

$$Q^{z^s} = \frac{CI(y_{oj}, b_{oh})}{CI(y_{oj}, b_{oh}|z^{os})} \tag{9}$$

In the second step, and considering the results of (5) including both environmental variables z^s and z^d $[(z^s, z^d) \in z]$, we determine the ratio that shows the additional impact of the discretionary (z^d) context variables:

$$Q^{z^d} = \frac{CI(y_{oj}, b_{oh}|z^{os})}{CI(y_{oj}, b_{oh}|z^o)} \tag{10}$$

2.3. Providing robustness to the estimation model

One of the issues of nonparametric deterministic models is the lack of statistical properties and robustness when extreme values and outliers are present. In this context, Cazals et al. [83] propose a robust nonparametric estimation of the frontier: the order-m estimator [see, for instance,84, for a relevant application]. The idea behind it is to compare units against a sub-sample of randomly drawn m units, and repeat the process B times in order to determine the average value of

[10] There are other approaches based on the primal formulation of the program. See, for instance, [43].

[11] Indeed, one of the main criticisms underlying composite index methodologies is related to the trade-offs between the variables. See [71].

[12] Although the antecedents of conditional estimations had been introduced earlier by Ruggiero [76].

the CI that comes from the solution of B models that only integrates one sub-sample in each run. Although this robust approach is based on probabilistic formulation [82], this can also be computed by following a Monte-Carlo algorithm [85]. The adaptation to our case of the required steps to complete the algorithm is as follows:

1. Select a subsample of m countries, with replacement, among the K countries.
2. Solve programs (2) and (5) and estimate $CI(y_{oj}, b_{oh})$, $CI(y_{oj}, b_{oh}|z^{os})$ and $CI(y_{oj}, b_{oh}|z^{o})$.
3. Repeat steps 1 and 2 B times.
4. Determine the average values $\tilde{CI}(y_{oj}, b_{oh})$, $\tilde{CI}(y_{oj}, b_{oh}|z^{os})$ and $\tilde{CI}(y_{oj}, b_{oh}|z^{o})$.
5. Calculate the average ratios of the corresponding composite indicators $\tilde{Q}^{z}$, $\tilde{Q}^{z^s}$ and $\tilde{Q}^{z^d}$.

3. Data, sample and variables

3.1. Data sources and variables included in the model

For this worldwide cross-sectional study, we used data from several sources, all of them open access. We retrieved Covid-19 related data from the website `OurWorldInData.org` [86]. This website has compiled data from several important sources, such as the European Centre for Disease Prevention and Control, national government reports, the Oxford COVID-19 Government Response Tracker, the World Health Organisation, and the Global Health Observatory Data Repository, among others. It has documented daily Covid-19 case numbers, death numbers, test numbers and stringency indices, and demographic, health and human development indicators from more than 180 countries.[13]

We also consider information on several quality of government indicators. Specifically, the perceptions on the quality of government for the countries in our sample as of 2019 correspond to the Worldwide Governance Indicators (WGI), available via `www.govindicators.org` [for details on the methodological aspects, see87]. This is the most widely accepted database on quality of government indicators at the country level.[14] Finally, to account for the cultural aspects of each country, we used Hofstede's six dimensions model [88,89], data from which are available at `www.hofstede-insights.com`.

For the unconditional model, we consider indicators which could reflect the success of interventions against COVID-19. Given that the nature of the phenomenon involves different phases and different key actions, an evaluation from a holistic perspective is required. In this regard, [90], based on the Global Health Security Index (GHSI) developed by the World Health Organisation (2020), point out that the good performance of a country in dealing with the COVID-19 pandemic is based on its capacity to detect, contain and treat the infection. From the review of the literature, as well as websites and institutions that systematise the progress of the pandemic, such as Worldometer (https://www.worldometers.info/coronavirus/) and Statista (https://www.statista.com/), the five indicators of the unconditional model were defined as follows:

Detection:

Cumulative tests performed per million inhabitants (y_1). The higher the value, the higher the effectiveness [27,33,39,90].

Positivity rate (total confirmed cases/total tests performed) (b_1). The higher the value, the lower the effectiveness [38,90].

Containment:

Cumulative confirmed COVID-19 cases per million inhabitants (b_2). The higher the value, the lower the effectiveness [27,28,30–32,34,36,37,39,90].

Cumulative confirmed COVID-19 deaths per million inhabitants (b_3). The higher the value, the lower the effectiveness [27–30,32,33,35–39,90].

Treatment:

Case fatality rate (total deaths/total cases) (b_4). The higher the value, the lower the effectiveness [90–93].

The literature provides three alternatives to account for the start point of the pandemic: absolute (on a certain calendar date); relative (when certain conditions are met, for example reaching a specific number of infections or deaths); and a combination of the above (from a certain date, but the country is only considered in the sample if after a given number of days, they reach a minimum number of infections). All of them have advantages and disadvantages. In our case we use the relative option since we assume that the phenomenon studied (the management of the pandemic) will develop differently in each country, according to its degree of evolution. Hence, to compare the effectiveness of the action between countries under equal conditions, the start of the pandemic was calculated from the date of the first confirmed case per million inhabitants in each country. A start date relative to each country has also been used by Ordu et al. [35] and Su et al. [34].[15]

Thus, in our sample of 61 countries, the evaluation of the pandemic management of each country in this study is evaluated 270 days after the start of the pandemic, which corresponds to information in the range from November 15, 2020 (South Korea) to December 21, 2020 (Zambia). We report descriptive statistics for these indicators in Table 1.

Some stylised facts that emerge when examining the data indicate:[16]

- Luxembourg and Denmark are the countries with the highest **number of tests per million inhabitants** (2,131,356.5 and 1,247,560.9, respectively). In contrast, Nigeria and Senegal are the countries with the lowest number of tests per million inhabitants (4488.3 and 14,810.7, respectively).
- Mexico is the country with the highest **positivity rate (total confirmed cases/total tests performed)**, followed by Argentina (0.433 and 0.411, respectively). In contrast, New Zealand and Australia are the countries with the lowest positivity rate (0.0016 and 0.0028, respectively).
- The country in the sample with the highest **cumulative confirmed COVID-19 cases per million of inhabitants** on day 270 of the start of the pandemic is Luxembourg (50,295.8 cases per million inhabitants). In contrast, Thailand and Nigeria are the countries with the lowest number of confirmed cases per million inhabitants (58.8 and 405.4, respectively).
- Regarding the **cumulative confirmed COVID-19 deaths per million inhabitants** indicator, Spain is the country with the worst performance and Thailand and New Zealand are the countries with the best performance (0.86 and 5.20, respectively).
- Finally, with respect to the **case fatality rate indicator** (total deaths/total cases), Mexico and Iran are the countries with the worst performance (0.091 and 0.053, respectively). In contrast,

[13] The data from the OurWorldInData.org database. We were retrieved on January 4, 2020.

[14] The Quality of Government Institute (Stockholm, Sweden) provides data for lower territorial units.

[15] We are grateful to one of the reviewers of the paper for pointing out these alternatives.

[16] We do not report individual data for space reasons and because they are available from the websites referred to above.

Table 1
Descriptive statistics.

	Mean	Std.dev.	Min.	25% percentile	Median	75% percentile	Max.
Cumulative performed tests[a]	288,330.24	332,886.63	4,488.29	69,236.06	247,180.24	354,273.79	2,131,356.47
Positivity rate (confirmed cases/tests performed)	0.09	0.09	0.00	0.03	0.07	0.12	0.43
Cumulative confirmed COVID-19 cases[a]	17,298.99	13,687.85	58.84	4,088.28	14,108.62	27,512.41	50,295.86
Cumulative confirmed COVID-19 deaths[a]	339.40	315.37	0.86	70.03	268.35	550.11	1,409.80
Case fatality rate (total deaths/total cases)	0.02	0.01	0.00	0.01	0.02	0.02	0.09

[a] Per million inhabitants.

Sri Lanka and Iceland show the best figures for this indicator (0.0047 and 0.0049, respectively).

3.2. Expert opinions

As indicated in the previous section, a key element when dealing with composite indicators is to determine the importance for these indicators—or, more precisely, the subindicators that compose them. Unfortunately, there is no single valid process to achieve this objective [94]. In cases where an applicable theoretical model is unknown and it is required to objectify these limits appropriately, one of the strategies used is expert judgement [95,96]. However, this strategy commonly involves a reduced number of key informants and has innumerable methodological and validity disadvantages [97].

For this purpose, and unlike previous studies, in our case we used a survey based on Best–Worst Scaling (BW), which is a survey method to obtain the relative importance that people attach to the various attributes of a product or service [98]. This methodology corresponds to an extension of paired comparisons that models the cognitive process by which respondents identify their preferences over a set of attributes, organised into subgroups of three or more items. The advantages are that respondents only have to make discrete choices and that it provides a more discriminating way of measuring the degree of importance respondents attach to each item. Since respondents can only choose one most preferred and one least preferred item in each set of choices, they are necessarily required to make trade-offs among the advantages [99]. Second, BW avoids problems of ranking bias, since there is only one way to choose the most and least preferred item, regardless of the respondent's cultural background [100]. This is interesting, as it makes the BW method a powerful way of conducting cross-national studies on preferences [101,102].

To carry out this study we conducted a survey using the Sawtooth Software platform among 1228 academics whose output had been published in *BMJ Quality & Safety*, which ranks first (according to its impact factor) in the Web of Science category *Health Care Sciences & Services*. We collected 145 responses, of which 128 were considered valid. Respondents included healthcare professionals from all over the world and sub-disciplines of healthcare sciences.[17] Once the responses were received, the Sawtooth Software platform, using Hierarchical Bayes techniques, allowed us to estimate the relative importance that each respondent (expert) gave to each of the five indicators evaluated [103]. Descriptive statistics are reported in Table 2.

[17] Each respondent was contacted with the following request: "We are interested in assessing the overall performance of a country in COVID-19 management, nine months after the start of the pandemic. We would like to know your opinion on the relative importance of each of the most commonly used indicators. To do so, we will present six different sets of variables below and ask you to indicate for each of them the indicators that would be most and least important to you in the overall assessment of a country's COVID-19 management". Subsequently, they were asked: "If you were to evaluate a country's performance in managing COVID-19 nine months after it began, please indicate which variables would be the most and the least important to you for this evaluation?". An example of this choice is reported in Fig. 1.

3.3. Structural and discretionary contextual variables

In addition, and as mentioned earlier in the paper, we consider a total of six variables that affect the management of the pandemic by the governments of each country. In the first place, four variables are associated with demographic, economic and cultural aspects that are long-term structural aspects of the country and therefore cannot be modified discretionally (i.e., they are non-controllable) by governments in the short term. For instance, Lavigne et al. [43] took a similar approach. These variables were selected based on the previous studies on performance evaluation reviewed in the introduction section and studies that show the relationship between contextual variables and COVID-19 evolution. These variables are:

Population density: number of inhabitants per square kilometre of the country's surface area [30,32,34,37].

Human Development Index (HDI): summary measure of average achievement in key dimensions of human development: a long and healthy life, access to quality education, and having a decent standard of living [104–106].[18]

Individualism: measure of the degree to which a society emphasises the achievement of individual goals over collective goals. It corresponds to one of Hofstede's (2001) cultural dimensions [107,108].

Long-term orientation: reflects a society's perspective of achieving success and gratification in the long-term rather than in the immediate future. It emphasises persistence, perseverance, and long-term growth [88,107,109].

Secondly, two short and medium-term discretionary variables under government control were used: *rule of law* and the *stringency index*.

Rule of law: one of the six governance dimensions of the Worldwide Governance Indicators (WGI), it reflects perceptions of the extent to which agents have confidence in and abide by the rules of society, and in particular the quality of contract enforcement, property rights, the police, and the courts, as well as the likelihood of crime and violence [110,111].

Mean stringency index: a composite measure of nine of the response metrics: school closures; workplace closures; cancellation of public events; restrictions on public gatherings; closures of public transport; stay-at-home requirements; public information campaigns; restrictions on internal movements; and international travel controls. For each country, we averaged the daily stringency index across the 270 pandemic days considered in the study [39,112].

We report descriptive statistics for these indicators in Table 3.

[18] See http://hdr.undp.org/en/content/human-development-index-hdi.

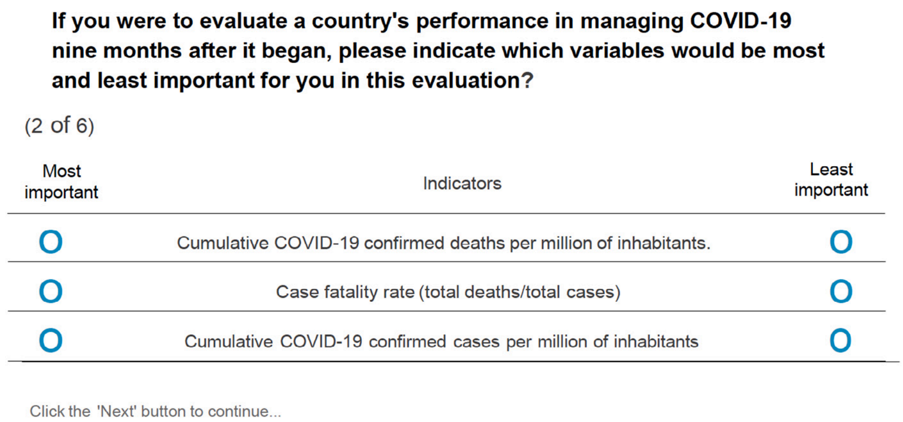

Fig. 1. Example of one of the survey questions.

Table 2
Results of the survey, descriptive statistics.

	Mean	Std.dev	Min.	Median	Max.	Lower 95.0% CL (median)	Upper 95.0% CL (median)
Cumulative performed tests[a]	10.87	12.60	0.25	5.11	46.68	3.26	8.33
Positivity rate (total confirmed cases/total performed tests)	14.51	16.24	0.15	7.03	54.57	5.29	10.63
Cumulative confirmed COVID-19 cases[a]	19.39	13.28	1.26	16.11	46.15	14.21	20.15
Cumulative confirmed COVID-19 deaths[a]	32.13	14.07	1.79	37.76	51.99	32.68	40.21
Case fatality rate (total deaths/total cases)	23.10	15.97	0.46	21.48	51.06	17.06	26.63

[a] Per million inhabitants.

Table 3
Contextual factors, descriptive statistics.

	Mean	Std. Dev.	Min.	25% percentile	Median	75% Percentile	Max.
Population density	145.90	209.75	3.08	24.72	83.48	209.59	1454.04
Human development index	0.82	0.11	0.51	0.77	0.85	0.91	0.95
Individualism	47.32	22.18	12.00	30.00	46.00	67.00	91.00
Long-term orientation	46.22	23.38	3.53	27.96	42.90	63.98	100.00
Rule of law	0.62	0.90	−0.90	−0.19	0.56	1.46	2.02
Stringency mean	59.96	11.19	36.21	52.51	58.49	67.23	88.10

4. Results

4.1. Effectiveness

The methodology described above involved calculating three models. All of them use the five outputs described in Section 3.1, but with different assumptions about the influence of contextual factors. The first model (unconditional) calculates $\tilde{CI}(y_{oj}, b_{oh})$, and assumes that each country's contextual factors do not affect its performance in managing the COVID-19 pandemic. The second model allows us to compute $\tilde{CI}(y_{oj}, b_{oh}|z^{os})$. It is a conditional model which incorporates the impact of the country's structural contextual factors in the calculation of the effectiveness index. The third model calculates $\tilde{CI}(y_{oj}, b_{oh}|z^{o})$, and additionally incorporates the discretionary factors under the government control. Table 4 summarises the contextual variables chosen in each model.

The calculation of these three models requires the definition of the directional vectors $g_y = (y_{o1}w_1, y_{o2}w_2, \dots, y_{oJ}w_J)$ and $g_b = (b_{o1}v_1, b_{o2}v_2, \dots, b_{oH}v_H)$, more specifically the set of the directional vector components weights $w_y = (w_1, w_2, \dots, w_J)$ and $v_b = (v_1, v_2, \dots, v_H)$[19]. Here, in accordance with the results of our survey we considered that an indicator is more important depending on the average weight assigned, and the consensus achieved among the experts. We therefore defined the directional vector components weights as $w_y = (\frac{\mu_{w_1}}{\sigma_{w_1}}, \frac{\mu_{w_2}}{\sigma_{w_2}}, \dots, \frac{\mu_{w_j}}{\sigma_{w_j}})$ and $v_b = (\frac{\mu_{v_1}}{\sigma_{v_1}}, \frac{\mu_{v_2}}{\sigma_{v_2}}, \dots, \frac{\mu_{v_h}}{\sigma_{v_h}})$, where μ is the average importance assigned by the experts to the indicator and σ is the standard deviation.[20] In order to normalise w_y and v_b with respect to y_1, we multiply each component of both vectors by $\frac{\sigma_{w_1}}{\mu_{w_1}}$. Consequently, and in accordance with the experts opinions, we set $w_y = (1)$ and $v_b = (1.035, 1.692, 2.646, 1.676)$.

Additionally, the calculation of any order-m model using the Monte-Carlo simulation requires the definition of the parameters m and B. There are different alternatives to determine the value of m. For instance, Bonaccorsi et al. [114] suggest choosing m such that the share of super-efficient observations is around 10% (6 units in our case). Simar [115], based on previous work by Barnett et al. [116], also suggests that the maximum percentage of outliers/super-efficient units should be $\frac{\sqrt{n}}{n}$ (8 units in our case). We have followed this criterion by setting $m = 400$.[21] For the parameter B, which determines the number of iterations for the bootstrap, Daraio and Simar [82] state that it is sufficient to set it to 200. Nevertheless, in our case, we have set $B = 500$.

Table 5 exhibits the average results for the effectiveness in managing the COVID-19 pandemic at both global and continental/subconti-

[19] It is important to differentiate the values used for defining the directional vector components from those assigned to each indicator in the primal program [see, for instance,43]. In our case, we orient the directional vector in accordance with the experts opinions, but we do not restrict the weights of the primal program. However, the results are still sensitive to the chosen weights for the components of the directional vector. This is a topic to be explored in future research.

[20] In finance, there is a widely used indicator (the Sharpe ratio) that has exactly the same definition. See [113].

[21] Daraio and Simar [82,85] noted that the order-m frontier converges to the full frontier when $m \to \infty$.

Table 4
Effectiveness models.

Context variable	Type	$\tilde{CI}(y_{oj}, b_{oh})$	$\tilde{CI}(y_{oj}, b_{oh}\|z^{os})$	$\tilde{CI}(y_{oj}, b_{oh}\|z^{o})$
Population density	Structural		X	X
Human development index	Structural		X	X
Individualism	Structural		X	X
Long-term orientation	Structural		X	X
Rule of law	Discretionary			X
Stringency	Discretionary			X

Table 5
Effectiveness results by continent and subcontinent.

Continent	$\tilde{CI}(y_{oj}, b_{oh})$	$\tilde{CI}(y_{oj}, b_{oh}\|z^{os})$	$\tilde{CI}(y_{oj}, b_{oh}\|z^{o})$	$\tilde{Q}^{z^s}$	$\tilde{Q}^{z^d}$	$\tilde{Q}^{z}$
Africa	0.843	0.939	0.974	0.902	0.965	0.869
Asia	0.880	0.963	0.991	0.916	0.970	0.887
Europe	0.837	0.964	0.963	0.870	1.000	0.870
North America	0.782	0.916	0.935	0.862	0.981	0.846
Oceania	1.000	1.000	1.000	1.000	1.000	1.000
South America	0.810	0.950	0.955	0.856	0.994	0.850
Total general	0.846	0.958	0.969	0.885	0.989	0.874

nental level. The higher the value of the score $\tilde{CI}(\bullet)$, the higher the level of effectiveness. Column $\tilde{CI}(y_{oj}, b_{oh})$, corresponding to the unconditional model, shows the results considering only the pandemic indicators, without considering any contextual variables. Columns $\tilde{CI}(y_{oj}, b_{oh}|z^{os})$ and $\tilde{CI}(y_{oj}, b_{oh}|z^{o})$ show the levels of effectiveness when countries that share similar contextual characteristics are compared. At the global level, the average performance is 0.846 when the particular environmental conditions of the countries are not considered, reaching a value of 0.958 when structural factors are considered and 0.969 when discretionary factors are also incorporated. The global impact of the environment is shown in column $\tilde{Q}^{z}$, standing at 0.874 (the lower the value, the greater the impact of the environment). Since $\tilde{Q}^{z^d} = 0.989$ and $\tilde{Q}^{z^s} = 0.885$, it is confirmed that a greater general impact can be attributed to structural variables.

The results at continent/subcontinent level show that the countries in the sample which best managed the pandemic belong to Oceania (the highest value corresponds to the highest performance), and appear efficient in all our estimations. These are countries that adopted restrictive measures from the beginning, with a strong negative impact on the economy. They are followed by countries in Asia and Africa, which show similar levels of effectiveness (0.880 and 0.843 respectively). In the African case, the good results could be due to a potential lower capacity to report the data correctly given the socio-economic characteristics of many of these countries—at least those in our sample. In the case of Asian countries, the results are consistent with the speed, forcefulness, duration and severity of the measures many of them adopted. They are followed by European and South American countries in terms of effectiveness in managing the pandemic, with an average value of 0.837 and 0.810, respectively. In these regions, the strategy was mostly to live with the virus, with many of them pursuing reactivation of the economy when health conditions improved. In last place are the North American countries with an effectiveness level of 0.782. As expected, differences between continents decrease when considering the contextual variables. The improvement of North and South American countries is especially remarkable, increasing their levels of effectiveness by 17%, from 0.782 to 0.916 and from 0.810 to 0.950, respectively. Indeed, $\tilde{Q}^{z}$ shows that North and South American countries were the most impacted by their environment, with a value of 0.846 and 0.850 respectively. This may be due to the more adverse contextual conditions or the strategies applied by many of the countries belonging to these subcontinents and their potential negative effect on virus control. At the opposite extreme, contextual factors showed no impact on the management of the pandemic in the countries of Oceania. In the rest of the areas, the impact of the environment remains at fairly similar levels, oscillating between 0.869 in Africa and 0.887 in Asia. The highest values of $\tilde{Q}^{z^d}$ for Oceania, Europe and South America mean that in these geographical areas structural factors had a greater impact.

At the country level (Table 6), Australia, Denmark, Estonia, Finland, Iceland, Israel, Luxembourg, Malaysia, New Zealand, Norway, Sri Lanka, Thailand and Uruguay stand out for their effective management of the pandemic without considering contextual conditions, followed by Nigeria (0.988), Ireland (0.952) and Ghana (0.908). In contrast, Colombia (0.741), Bulgaria (0.741), Chile (0.740), South Africa (0.739), Argentina (0.738), Sweden (0.729), Iran (0.728) and Mexico (0.727) are in the last positions. Sweden perhaps deserves special comment, since it opted for a lax application of measures to achieve rapid biological immunisation of its population. Although this merits further scrutiny, Sweden is ranked 59th out of the 61 countries considered, which suggests that results might not have corresponded to the expectations of the country's authorities. In contrast, other Scandinavian countries, some of which restricted their citizens' movement in and out of Sweden at some point, have much higher levels of effectiveness. This is the case of Norway, Denmark or Finland, all of which are top performers, as mentioned above. The United Kingdom also followed a strategy similar to that of Sweden in the first stage and, although it became stricter over time, it has often been more lax than other countries. The result has not been so positive, ranking 36th, with a value of 0.801.

The results per country in Table 6 reveal three large groups of countries. First, a select group of countries shows very high levels of effectiveness in the management of the pandemic, whether or not their contextual variables are taken into account. These 15 countries are Australia, Denmark, Estonia, Finland, Iceland, Israel, Luxembourg, Malaysia, New Zealand, Nigeria, Norway, Sri Lanka, Uruguay, Thailand and Ireland.

The second group includes the countries whose performance is ineffective after an initial evaluation (not considering contextual factors), but when compared with countries with similar characteristics (especially long-term structural factors) their evaluation improves considerably. It is striking that almost half of the countries in this group are European (15/31). The countries in this group are Ghana, Morocco, Senegal, Zambia, Japan, Kazakhstan, Pakistan, Saudi Arabia, South Korea, Turkey, Austria, Bulgaria, Germany, Greece, Hungary, Latvia, Lithuania, Malta, Netherlands, Portugal, Russia, Serbia, Slovakia, Spain, United Kingdom, Canada, Dominican Republic, United States, Chile, Colombia, and Paraguay. Consequently, these are countries whose governments, relative to others with similar levels of these background variables, have better managed the pandemic.

The last group corresponds to countries with low levels of effectiveness in a first evaluation, harmed by contextual variables and that maintain their low levels of effectiveness, even though they are compared with countries with similar contexts. This accounts for an inadequate performance of their governments. These 15 countries are: South Africa, Iran, Croatia, Italy, Romania, Slovenia, Ukraine, Argentina, Mexico, Switzerland, Sweden, Poland, the Czech Republic, the Philippines and Namibia.

4.2. Contextual effects

The above insights suggest that at least some contextual variables have a significant effect on the effectiveness of COVID-19 pandemic management. It is therefore of particular interest to analyse which variables have a statistically significant effect, and, if so, what kind of impact. To this end, two nonparametric regressions were performed [42, 82,85] following a similar approach to [117] of gradually introducing

Table 6
Effectiveness results by country.

Continent	Continent/subcontinent	$\tilde{CI}(y_{oj}, b_{oh})$	$\tilde{CI}(y_{oj}, b_{oh}\|z^{os})$	$\tilde{CI}(y_{oj}, b_{oh}\|z^{o})$	$\tilde{Q}^{z^s}$	$\tilde{Q}^{z^d}$	$\tilde{Q}^{z}$
Argentina	South America	0.738	0.787	0.810	0.937	0.972	0.911
Australia	Oceania	1.000	1.000	1.000	1.000	1.000	1.000
Austria	Europe	0.785	1.000	0.999	0.785	1.001	0.786
Bulgaria	Europe	0.741	1.005	1.034	0.736	0.972	0.716
Canada	North America	0.811	0.969	0.998	0.837	0.971	0.812
Chile	South America	0.740	0.964	0.964	0.768	1.000	0.768
Colombia	South America	0.741	1.000	1.000	0.741	1.000	0.741
Croatia	Europe	0.810	0.928	0.940	0.872	0.987	0.861
Czech Rep.	Europe	0.746	0.790	0.790	0.945	1.000	0.945
Denmark	Europe	1.001	1.000	1.000	1.001	1.000	1.001
Dominican Rep.	North America	0.806	0.966	1.000	0.834	0.966	0.806
Estonia	Europe	1.000	1.000	1.000	1.000	1.000	1.000
Finland	Europe	1.001	1.000	1.000	1.001	1.000	1.001
Germany	Europe	0.828	0.940	0.956	0.880	0.984	0.866
Ghana	Africa	0.908	1.000	1.000	0.908	1.000	0.908
Greece	Europe	0.814	1.000	1.000	0.814	1.000	0.814
Hungary	Europe	0.746	0.991	0.987	0.752	1.004	0.755
Iceland	Europe	1.001	1.000	1.000	1.001	1.000	1.001
Iran	Asia	0.728	0.765	0.905	0.951	0.846	0.804
Ireland	Europe	0.952	0.952	0.952	1.000	1.000	1.000
Israel	Asia	1.000	1.000	1.000	1.000	1.000	1.000
Italy	Europe	0.791	0.936	0.946	0.845	0.989	0.835
Japan	Asia	0.882	1.000	1.000	0.882	1.000	0.882
Kazakhstan	Asia	0.839	1.000	1.000	0.839	1.000	0.839
Latvia	Europe	0.879	1.000	0.999	0.879	1.001	0.880
Lithuania	Europe	0.804	1.000	1.000	0.804	1.000	0.804
Luxembourg	Europe	1.001	1.000	1.000	1.001	1.000	1.001
Malaysia	Asia	1.000	1.000	1.000	1.000	1.000	1.000
Malta	Europe	0.869	1.000	1.000	0.869	1.000	0.869
Mexico	North America	0.727	0.731	0.742	0.995	0.986	0.980
Morocco	Africa	0.798	0.999	1.000	0.799	0.999	0.798
Namibia	Africa	0.857	0.857	0.857	1.000	1.000	1.000
Netherlands	Europe	0.773	0.996	0.998	0.776	0.998	0.775
New Zealand	Oceania	1.000	1.000	1.000	1.000	1.000	1.000
Nigeria	Africa	0.988	1.000	1.000	0.988	1.000	0.988
Norway	Europe	1.001	1.000	1.000	1.001	1.000	1.001
Pakistan	Asia	0.810	1.000	1.000	0.810	1.000	0.810
Paraguay	South America	0.771	1.000	1.000	0.771	1.000	0.771
Philippines	Asia	0.763	0.791	0.990	0.964	0.799	0.770
Poland	Europe	0.765	0.779	0.824	0.983	0.945	0.929
Portugal	Europe	0.783	1.000	1.000	0.783	1.000	0.783
Romania	Europe	0.744	0.904	0.894	0.823	1.011	0.832
Russia	Europe	0.870	1.000	1.000	0.870	1.000	0.870
Saudi Arabia	Asia	0.807	1.000	1.000	0.807	1.000	0.807
Senegal	Africa	0.805	0.948	0.961	0.848	0.987	0.838
Serbia	Europe	0.791	1.000	1.000	0.791	1.000	0.791
Slovakia	Europe	0.870	1.055	1.030	0.825	1.024	0.845
Slovenia	Europe	0.748	0.931	0.927	0.804	1.004	0.807
South Africa	Africa	0.739	0.769	1.000	0.961	0.769	0.739
South Korea	Asia	0.877	1.000	1.000	0.877	1.000	0.877
Spain	Europe	0.754	1.000	1.000	0.754	1.000	0.754
Sri Lanka	Asia	1.000	1.000	1.000	1.000	1.000	1.000
Sweden	Europe	0.729	0.889	0.799	0.820	1.113	0.912
Switzerland	Europe	0.758	0.874	0.890	0.867	0.983	0.852
Thailand	Asia	1.000	1.000	1.000	1.000	1.000	1.000
Turkey	Asia	0.860	1.000	1.000	0.860	1.000	0.860
Ukraine	Europe	0.784	0.898	0.898	0.873	1.000	0.873
United Kingdom	Europe	0.801	1.000	1.000	0.801	1.000	0.801
United States	North America	0.784	1.000	1.000	0.784	1.000	0.784
Zambia	Africa	0.810	1.000	1.000	0.810	1.000	0.810
Uruguay	South America	1.062	1.000	1.000	1.062	1.000	1.062
Mean		0.846	0.958	0.969	0.885	0.989	0.874
Std Dev		0.099	0.076	0.064	0.091	0.047	0.093
Max		1.062	1.055	1.034	1.062	1.113	1.062
Min		0.727	0.731	0.742	0.736	0.769	0.716
# super-efficient		8	2	2			

the contextual variables.[22] In regression 1 (see Table 7), the Q^{z^s} ratio was used as the dependent variable and the structural environment variables as covariates. In regression 2, the Q^z ratio was used as the dependent variable and all the contextual variables as independent variables. The significance test of the variables was estimated via a bootstrap with 1000 iterations.

Table 7 shows the significance and type of impact of each contextual factor on the dependent variable. In both regressions all contextual variables, except rule of law and population density, were found to be

[22] Our approach also shares the same underpinnings, based on nonparametric regression, of one of the first contributions in the field which made an attempt to deal with similar issues [118].

Table 7
Nonparametric regression results.

Variable	Regression 1 ($\tilde{Q}^{z^s}$)		Regression 2 ($\tilde{Q}^{z}$)	
	p-value[a]	Impact	p-value[a]	Impact
Population density	0.122	(-)	0.112	(-)
Human development index	**0.098**	Favourable	**0.095**	Favourable
Individualism	**0.008**	Inverted U-shape	**0.033**	Inverted U-shape
Long-term orientation	**0.035**	Unfavourable	<2e-16	Unfavourable
Rule of law			0.282	(-)
Stringency			**0.006**	Unfavourable
R^2	0.308		0.448	

[a] p -values in bold indicate significance at least at the 10% level.

statistically significant at 10%. One of the advantages of this technique is that it graphically shows the gradient of each independent variable for its different values, which allowed us to determine that the marginal effects were different among the contextual variables. The scatter plot of the gradients[23] was calculated for the median in order to identify the marginal impact of each contextual variable on effectiveness. In both regressions, the human development index shows a favourable and significant effect. In other words, the countries with the highest income, educational level and life expectancy managed the pandemic more effectively, as was to be expected. Notably, and more directly related to the pandemic, countries with high levels of human development are probably at lower risk of previous pathologies and enjoy better levels of healthcare services.

Additionally, very high levels of the index may be associated with high levels of GDP per capita where a higher percentage of jobs are in the formal economy, more financial support was received from governments, it was easier to implement teleworking, and many citizens were not forced to go out to work or to interact socially to obtain the minimum income to survive. For the individualism variable, the generalised behaviour is an inverted U-shape in both regressions. At the individual level, it has a positive effect on effectiveness. However, when a high level is reached, the effect becomes negative. The explanation for this phenomenon may be that higher values of this characteristic are related to higher levels of free-rider behaviour, while at not very high levels it probably contributed to more effective respect for regulations implemented to reduce social interaction in many countries. Long-term orientation has an unfavourable effect for most of its range and a significant impact on effectiveness. One explanation for this behaviour may be that the higher the long-term orientation, the more difficult it may be for a society to achieve the levels of intensity and commitment in the short term that are required to combat the pandemic.

Finally, the marginal effect of the discretionary variables with respect to the structural variables is obtained from regression 2. Rule of law has an unfavourable impact, although not significant, on the effectiveness of pandemic management. The severity of the measures taken (stringency) has an unfavourable impact on effectiveness. This result may be contradictory and unexpected. It is important to note that the effect of the measures taken is seen weeks after their application. As there is no consensus on exactly how long this period is, we chose not to introduce any time lag in the variable. Consequently, a different interpretation is needed. In particular, the suggestion in our results that the greater the stringency, the less effective the management should be interpreted as an effect and not so much as a cause. In other words, countries mainly increased the severity of their measures when pandemic incidence indicators showed high levels. Therefore, it seems that the predominant strategy among the countries analysed was to favour coexistence with the virus, in an attempt to achieve the delicate balance between health and the economic situation, with varying degrees of success.

[23] Available upon request.

5. Concluding remarks

By February 2021, the number of COVID-19 cases had reached nearly 110 million worldwide, with approximately 2.5 million deaths, a human tragedy with enormous collateral damage to the economy and living conditions of the majority of the population.

Governments attempted to address the pandemic from various angles, with varying degrees of success. In this regard, while the literature on the consequences of COVID-19 and government responses from an economic standpoint is growing exponentially [119], the number of initiatives from an operations research position, despite their potential, is relatively limited and has focused on measuring point targets from an efficiency perspective, with few contributions analysing the impact of country contextual factors on the effectiveness of pandemic management. We attempt to fill this gap by proposing a novel method based on nonparametric activity analysis techniques developed in several stages. In the first stage, we propose a model of unconditional benefit of doubt (directional BoD), based on the construction of composite indicators that take into account expert judgment. We refine the model by proposing a conditional model operating in two stages, sequentially introducing structural and discretionary contextual variables (conditional directional BoD), since government actions may be strongly conditioned by factors beyond their control, at least in the short and medium term.

We adopted this approach because we believe that to properly evaluate countries' management of the pandemic, indicators that measure the phenomenon holistically should be used [120,121]. Thus, a country that manages the pandemic effectively is one that reports high results in terms of detection, control and treatment [90]. Similarly, a flexible methodology is required to guide the evaluation in terms of weights, but within certain normative parameters set by experts, and to ensure a fair allocation of responsibilities of managers.

Based on the results, we conclude that effective pandemic management by countries is a multifaceted task that involves not only short-term management measures but also the strengthening of long-term governance and structural conditions that favour detection, control and treatment. In this sense, this study establishes that a country's socioeconomic and cultural structural variables have a greater impact on the effectiveness of pandemic management than discretionary variables in the medium and short term.

Indeed, given the sophistication of the models used, the results can be explored from multiple perspectives. According to the results of the unconditional model, in which contextual variables are not taken into account, the region (continent) with the best pandemic management is Oceania, followed by Asia. In contrast, North America is in last place, followed by South America. These results are subject to further examination, particularly at the country level, as this assessment cannot be fully generalised. The inclusion of structural (non-controllable) contextual variables is revealing, as the assessment of some regions varies, particularly South America. This shows the negative impact of the disadvantageous situation of South American countries on the unconditional assessment that does not consider this aspect. The same outcome is seen, but less intensely, when the medium- and short-term contextual variables enter into the analysis. For example, there is no significant change in the evaluation for Europe. Even when all the contextual aspects are taken into account, the best evaluations of effectiveness are found in the countries of Oceania, followed by Asian countries. Europe lies in the middle of the table, above South America, but lower than Africa. North America is the worst performing region, considering its contextual restrictions.

At the country level, three major groups can be seen according to their performance in pandemic management: those that are highly effective (15 countries), those that become effective when considering contextual factors (31 countries) and those that maintain low levels of effectiveness, even when considering contextual variables (15 countries). This last group reflects the management weaknesses of their respective governments.

In contrast, the nonparametric regression analysis corroborates the results obtained by [106] and Zheng et al. [104] that socioeconomic aspects such as the human development index, which takes into account aspects of life expectancy, education and per capita income, have a positive impact on a country's performance in pandemic management.

The study also concludes that the individualism of a society is not linearly related to the effectiveness of a government, but rather has an inverted-*U* shape. That is, as a society's individualism increases, it has a positive effect on effectiveness, probably because it is easier to contain social interaction and there is more respect for the restrictions that many countries introduced in this regard. However, if individualism is high in society, to the extent that it is exacerbated, it can encourage selfish behaviour in certain segments of the population. This finding deepens the understanding of the impact of cultural factors such as individualism/collectivism, in line with, for example, Feng et al. [108], who concluded that people from more individualistic countries and regions were less likely to follow social distancing norms during the COVID-19 pandemic, or [122], who found that highly individualistic countries were unsuccessful in limiting the number of deaths and confirmed cases of COVID-19 (as people from such cultures prioritise their freedom and privacy).

Our findings also show that long-term orientation has a significant unfavourable impact on effectiveness, probably because the higher the long-term orientation, the more difficult it is for a society to achieve the levels of intensity and commitment needed to combat the pandemic in the short term. Along these lines, Chen and Biswas [122] note that nations with a short-term orientation are more likely to implement hypervigilance and pandemic prevention measures.

However, the nonparametric regression analysis performed to assess the importance of each contextual variable indicates that population density and the level of rule of law in a country are not shown to be significant factors in assessing pandemic management.

Thus, the implications of our study are manifold. Not only should we highlight the potential usefulness of our results, in terms of providing benchmarks of best practices, and the relevance of contextual factors, but also the potential of operations research to address various problems related to health policy management. In times of crisis, it is clear that governments must act quickly to introduce what can sometimes be unpopular short-term measures. However, this study shows that the effectiveness of a government's management depends on socioeconomic and cultural aspects of its society and therefore the measures taken must be analysed in terms of how that culture and socioeconomic situation may affect the behaviour of individuals.

There are limitations to assessing pandemic preparedness at the country level. The most obvious is the comparability of data at this level, given differences in reporting systems. These systems may vary not only between countries but also within countries, not only in what information is reported but also in its quality, depending on the degree of decentralisation—i.e. sub-national levels of government may report different information.

The possible extensions to this study are multiple. One of the most interesting ones would be to reflect the debate on the (possible) trade-off between public health *vis-à-vis* economic growth. Although several research initiatives have already addressed this issue [e.g., 123], the methods considered here would provide refreshing insights into the debate—not only because of the decline in GDP during the pandemic, but also because of the subsequent macroeconomic effects (e.g., high inflation levels or shortages of raw materials, among many others).

Declaration of competing interest

none

Data availability

Data will be made available on request.

References

[1] Goolsbee A, Syverson C. Fear, lockdown, and diversion: Comparing drivers of pandemic economic decline 2020. J Public Econ 2021;193:104311.
[2] Bourdin S, Jeanne L, Nadou F, Noiret G. Does lockdown work? A spatial analysis of the spread and concentration of Covid-19 in Italy. Reg Stud 2021;55(7):1182–93.
[3] Brodeur A, Clark AE, Fleche S, Powdthavee N. COVID-19, lockdowns and well-being: Evidence from Google Trends. J Public Econ 2021;193:104346.
[4] Pei J, de Vries G, Zhang M. International trade and Covid-19: City-level evidence from China's lockdown policy. J Reg Sci 2022;62:670–95.
[5] Aum S, Lee SYT, Shin Y. COVID-19 doesn't need lockdowns to destroy jobs: The effect of local outbreaks in Korea. Labour Econ 2021;70:101993.
[6] Murray EJ. Epidemiology's time of need: COVID-19 calls for epidemic-related economics. J Econ Perspect 2020;34(4):105–20.
[7] Yoshikawa H, Wuermli AJ, Britto PR, Dreyer B, Leckman JF, Lye SJ, et al. Effects of the global coronavirus disease-2019 pandemic on early childhood development: Short-and long-term risks and mitigating program and policy actions. J Pediatrics 2020;223:188–93.
[8] Bianchi F, Bianchi G, Song D. The long-term impact of the COVID-19 unemployment shock on life expectancy and mortality rates. Working Paper 28304, National Bureau of Economic Research; 2020.
[9] Juergensen J, Guimón J, Narula R. European SMEs amidst the COVID-19 crisis: Assessing impact and policy responses. J Ind Bus Econ 2020;47(3):499–510.
[10] Cohen S, Rossi FM, Caperchione E, Brusca I. Debate: If not now, then when? Covid-19 as an accelerator for public sector accrual accounting in Europe. Public Money Manag 2021;41(1):10–2.
[11] Andries AM, Ongena S, Sprincean N. The COVID-19 pandemic and sovereign bond risk. Technical report 20–42, Zurich: Swiss Finance Institute Research Paper; 2020.
[12] Agasisti T, Soncin M. Higher education in troubled times: On the impact of Covid-19 in Italy. Stud Higher Educ 2021;46(1):86–95.
[13] Dokhov R, Topnikov M. Everyday mobility as a vulnerability marker: The uneven reaction to coronavirus lockdown in Russia. Environ Plan A: Econ Space 2021;53(4):612–5.
[14] Harris JE. Geospatial analysis of a COVID-19 outbreak at the University of Wisconsin–Madison: Potential role of a cluster of local bars. Epidemiol Infect 2022;150.
[15] Ascani A, Faggian A, Montresor S. The geography of COVID-19 and the structure of local economies: The case of Italy. J Reg Sci 2021;61(2):407–41.
[16] Hamermesh DS. Life satisfaction, loneliness and togetherness, with an application to Covid-19 lock-downs. Rev Econ Household 2020;18(4):983–1000.
[17] Connell R. COVID-19/sociology. J Sociol 2020;56(4):745–51.
[18] Bol D, Giani M, Blais A, Loewen PJ. The effect of COVID-19 lockdowns on political support: Some good news for democracy? Eur J Political Res 2021;60(2):497–505.
[19] Glover A, Heathcote J, Krueger D, Ríos-Rull J-V. Health versus wealth: On the distributional effects of controlling a pandemic. Working paper 27046, National Bureau of Economic Research; 2020.
[20] Favero CA, Ichino A, Rustichini A. Restarting the economy while saving lives under Covid-19. Discussion paper DP14664, London: CEPR; 2020.
[21] Avery C, Bossert W, Clark A, Ellison G, Ellison SF. Policy implications of models of the spread of coronavirus: Perspectives and opportunities for economists. Working Paper 27007, National Bureau of Economic Research; 2020.
[22] Chen K, Pun CS, Wong HY. Efficient social distancing during the COVID-19 pandemic: Integrating economic and public health considerations. European J Oper Res 2022. [forthcoming].
[23] Kavanagh MM, Singh R. Democracy, capacity, and coercion in pandemic response: COVID-19 in comparative political perspective. J Health Politics, Policy Law 2020;45(6):997–1012.
[24] Haushofer J, Metcalf CJE. Which interventions work best in a pandemic? Science 2020;368(6495):1063–5.
[25] Farrell MJ. The measurement of productive efficiency. J R Stat Soc 1957;Ser.A,120:253–81.
[26] Nikolopoulos K, Punia S, Schäfers A, Tsinopoulos C, Vasilakis C. Forecasting and planning during a pandemic: COVID-19 growth rates, supply chain disruptions, and governmental decisions. European J Oper Res 2021;290(1):99–115.
[27] Klumpp M, Loske D, Bicciato S. COVID-19 health policy evaluation: Integrating health and economic perspectives with a data envelopment analysis approach. Eur J Health Econ 2022. [forthcoming].
[28] Min H, Lee CC, Joo S-J. Assessing the efficiency of the Covid-19 control measures and public health policy in OECD countries from cultural perspectives. Benchmarking: Int J 2022;29(6):1781–96.
[29] Sinimole KR. Healthcare efficiency and the best practices of health systems across the world during COVID-19 pandemic. Int J Healthc Manag 2022. [forthcoming].

[30] Lupu D, Tiganasu R. COVID-19 and the efficiency of health systems in Europe. Health Econ Rev 2022;12(1):1–15.
[31] Lee YH, Lim SYR. Efficacy in COVID-19 management: The case of ASEAN. Transform Gov: People, Process Policy 2022. [forthcoming].
[32] Doğan Mİ, Özsoy VS, Örkcü HH. Performance management of OECD countries on Covid-19 pandemic: A criticism using data envelopment analysis models. J Facil Manag 2021;19(4):479–99.
[33] Breitenbach MC, Ngobeni V, Aye GC. Global healthcare resource efficiency in the management of COVID-19 death and infection prevalence rates. Front Public Health 2021;9:638481.
[34] Su EC-Y, Hsiao C-H, Chen Y-T, Yu S-H. An examination of COVID-19 mitigation efficiency among 23 countries. Healthcare 2021;9(6):755.
[35] Ordu M, Kirli Akin H, Demir E. Healthcare systems and Covid19: Lessons to be learnt from efficient countries. Int J Health Plann Manag 2021;36(5):1476–85.
[36] Mourad N, Habib A, Tharwat A. Appraising healthcare systems' efficiency in facing COVID-19 through data envelopment analysis. Decis Sci Lett 2021;10(3):301–10.
[37] Ibrahim MD, Binofai FA, Mm Alshamsi R. Pandemic response management framework based on efficiency of COVID-19 control and treatment. Future Virol 2020;15(12):801–16.
[38] Ahmad WMAW, Zainon WMNW, Nawi MAA, Ghazali FMM. Malaysia's Efficiency in Dealing with COVID-19 Outbreaks Compared to Other Asian Countries by Using Stochastic Frontier Analysis (SFA). Ann King Edward Med Univ 2020;26(2):324–9.
[39] Aydin N, Yurdakul G. Assessing countries' performances against COVID-19 via WSIDEA and machine learning algorithms. Appl Soft Comput 2020;97:106792.
[40] Maricic M, Egea JA, Jeremic V. A hybrid enhanced scatter search—Composite I-Distance indicator (eSS-CIDI) optimization approach for determining weights within composite indicators. Soc Indic Res 2019;144(2):497–537.
[41] Titl V, De Witte K. How politics influence public good provision. Socio-Econ Plan Sci 2022;81:101000.
[42] Rogge N, De Jaeger S, Lavigne C. Waste performance of NUTS 2-regions in the EU: A conditional directional distance benefit-of-the-doubt model. Ecol Econom 2017;139:19–32.
[43] Lavigne C, De Jaeger S, Rogge N. Identifying the most relevant peers for benchmarking waste management performance: A conditional directional distance Benefit-of-the-Doubt approach. Waste Manag 2019;89:418–29.
[44] OECD-European Commission Joint Research Centre. Handbook on constructing composite indicators. Methodology and user guide. Paris: OECD publishing; 2008.
[45] Wang H. A generalized MCDA–DEA (multi-criterion decision analysis–data envelopment analysis) approach to construct slacks-based composite indicator. Energy 2015;80:114–22.
[46] Cherchye L, Lovell CK, Moesen W, Van Puyenbroeck T. One market, one number? A composite indicator assessment of EU internal market dynamics. Eur Econ Rev 2007;51(3):749–79.
[47] Melyn W, Moesen W. Towards a synthetic indicator of macroeconomic performance: Unequal weighting when limited information is available. Public Econ Res Pap 1991;1–24.
[48] Cherchye L, Moesen W, Rogge N, Van Puyenbroeck T, Saisana M, Saltelli A, et al. Creating composite indicators with DEA and robustness analysis: The case of the technology achievement index. J Oper Res Soc 2008;59(2):239–51.
[49] Despotis DK. A reassessment of the human development index via data envelopment analysis. J Oper Res Soc 2005;56(8):969–80.
[50] De Witte K, Rogge N. Accounting for exogenous influences in performance evaluations of teachers. Econ Educ Rev 2011;30(4):641–53.
[51] Gaaloul H, Khalfallah S. Application of the "benefit-of-the-doubt" approach for the construction of a digital access indicator: A revaluation of the "digital access index". Soc Indic Res 2014;118(1):45–56.
[52] Morais P, Camanho AS. Evaluation of performance of European cities with the aim to promote quality of life improvements. Omega 2011;39(4):398–409.
[53] Stumbriene D, Camanho AS, Jakaitiene A. The performance of education systems in the light of Europe 2020 strategy. Ann Oper Res 2019;288(2):577–608.
[54] Zanella A, Camanho A, Dias TG. Benchmarking countries' environmental performance. J Oper Res Soc 2013;64(3):426–38.
[55] Cooper W, Seiford L, Tone K. Data envelopment analysis: A comprehensive text with models, applications, references and DEA-solver software. 2nd ed.. New York: Springer US; 2007.
[56] Lauer JA, Lovell CK, Murray CJ, Evans DB. World health system performance revisited: The impact of varying the relative importance of health system goals. BMC Health Serv Res 2004;4(1):19.
[57] Bowen HP, Moesen W. Composite competitiveness indicators with endogenous versus predetermined weights: An application to the World Economic Forum's global competitiveness index. Compet Rev: Int Bus J 2011;21(2):129–51.
[58] Stumbriene D, Camanho AS, Jakaitiene A. The performance of education systems in the light of Europe 2020 strategy. Ann Oper Res 2019;1–32.
[59] Verschelde M, Rogge N. An environment-adjusted evaluation of citizen satisfaction with local police effectiveness: Evidence from a conditional Data Envelopment Analysis approach. European J Oper Res 2012;223(1):214–25.
[60] Guardiola J, Picazo-Tadeo AJ. Building weighted-domain composite indices of life satisfaction with data envelopment analysis. Soc Indic Res 2014;117(1):257–74.
[61] Peiró-Palomino J, Picazo-Tadeo AJ. OECD: One or many? Ranking countries with a composite well-being indicator. Soc Indic Res 2018;139(3):847–69.
[62] Fusco E, Vidoli F, Rogge N. Spatial directional robust Benefit of the Doubt approach in presence of undesirable output: An application to Italian waste sector. Omega 2020;94:102053.
[63] Fusco E. Enhancing non-compensatory composite indicators: A directional proposal. European J Oper Res 2015;242(2):620–30.
[64] Chambers RG, Chung Y, Färe R. Profit, directional distance functions, and Nerlovian efficiency. J Optim Theory Appl 1998;98:351–64.
[65] Färe R, Grosskopf S. Theory and application of directional distance functions. J Prod Anal 2000;13(2):93–103.
[66] Chung YH, Färe R, Grosskopf S. Productivity and undesirable outputs: A directional distance function approach. J Environ Manag 1997;51(3):229–40.
[67] Oh D-h. A metafrontier approach for measuring an environmentally sensitive productivity growth index. Energy Econ 2010;32(1):146–57.
[68] Picazo-Tadeo AJ, Reig-Martinez E, Hernandez-Sancho F. Directional distance functions and environmental regulation. Resour Energy Econ 2005;27(2):131–42.
[69] Chambers RG, Chung Y, Färe R. Benefit and distance functions. J Econom Theory 1996;70(2):407–19.
[70] Sueyoshi T, Goto M. Should the US clean air act include CO2 emission control? Examination by data envelopment analysis. Energy Policy 2010;38(10):5902–11.
[71] Greco S, Ishizaka A, Tasiou M, Torrisi G. On the methodological framework of composite indices: A review of the issues of weighting, aggregation, and robustness. Soc Indic Res 2019;141:61–94.
[72] D'Inverno G, De Witte K. Service level provision in municipalities: A flexible directional distance composite indicator. European J Oper Res 2020;286(3):1129–41.
[73] Daraio C, Simar L. Advanced robust and nonparametric methods in efficiency analysis. Methodology and applications. Studies in productivity and efficiency, New York: Springer; 2007.
[74] Cherchye L, De Witte K, Ooghe E, Nicaise I. Efficiency and equity in private and public education: A nonparametric comparison. European J Oper Res 2010;202(2):563–73.
[75] De Witte K, Kortelainen M. What explains the performance of students in a heterogeneous environment? Conditional efficiency estimation with continuous and discrete environmental variables. Appl Econ 2013;45(17):2401–12.
[76] Ruggiero J. On the measurement of technical efficiency in the public sector. European J Oper Res 1996;90(3):553–65.
[77] Bădin L, Daraio C, Simar L. Optimal bandwidth selection for conditional efficiency measures: A data-driven approach. European J Oper Res 2010;201(2):633–40.
[78] Racine JS, Li Q. Nonparametric estimation of regression functions with both categorical and continuous data. J Econometrics 2004;119(1):99–130.
[79] Hall P, Li Q, Racine JS. Nonparametric estimation of regression functions in the presence of irrelevant regressors. Rev Econ Stat 2007;89(4):784–9.
[80] Li Q, Racine JS. Nonparametric econometrics: Theory and practice. Princeton and Oxford: Princeton University Press; 2007.
[81] De Witte K, Kortelainen M. Blaming the exogenous environment? Conditional efficiency estimation with continuous and discrete environmental variables. Discussion papers 833, Leuven, Belgium: Center for Economic Studies; 2008.
[82] Daraio C, Simar L. Introducing environmental variables in nonparametric frontier models: A probabilistic approach. J Prod Anal 2005;24(1):93–121.
[83] Cazals C, Florens J-P, Simar L. Nonparametric frontier estimation: A robust approach. J Econometrics 2002;106:1–25.
[84] Cherchye L, De Witte K, Perelman S. A unified productivity-performance approach applied to secondary schools. J Oper Res Soc 2019;70(9):1522–37.
[85] Daraio C, Simar L. Conditional nonparametric frontier models for convex and nonconvex technologies: A unifying approach. J Prod Anal 2007;28(1):13–32.
[86] Roser M, Ritchie H, Ortiz-Ospina E, Hasell J. Coronavirus pandemic (COVID-19). Published online at OurWorldInData.org: OurWorldInData.org; 2020, Retrieved from https://ourworldindata.org/coronavirus.
[87] Kaufmann D, Kraay A, Mastruzzi M. The worldwide governance indicators: Methodology and analytical issues. Policy research working paper series 5430, The World Bank; 2010.
[88] Hofstede G. Culture's consequences. 2nd ed.. Thousand Oaks, CA: Sage; 2001.
[89] Hofstede G, Hofstede GJ. Cultures and organizations. Software of the mind. 2nd ed.. New York: McGraw-Hill; 2005.
[90] Kennedy DS, Vu V, Ritchie H, Bartlein R, Rothschild O, Bausch DG, et al. COVID-19: Identifying countries with indicators of success in responding to the outbreak. Gates Open Res 2021;4:62.
[91] Fan G, Yang Z, Lin Q, Zhao S, Yang L, He D. Decreased case fatality rate of COVID-19 in the second wave: A study in 53 countries or regions. Transbound Emerg Dis 2021;68(2):213–5.
[92] Bignami-Van Assche S, Ghio D. Comparing COVID-19 fatality across countries: A synthetic demographic indicator. J Popul Res 2022;39(4):513–25.

[93] Cao C, Li N, Liu L. Do national cultures matter in the containment of COVID-19? Int J Sociol Soc Policy 2020;40(9/10):939–61.
[94] Ueda T, Amatatsu H. Determination of bounds in DEA assurance region method: Its application to evaluation of baseball players and chemical companies. J Oper Res Soc Japan 2009;52(4):453–67.
[95] Giménez V, Thieme C, Prior D, Tortosa-Ausina E. Evaluation and determinants of preschool effectiveness in Chile. Soc-Econ Plan Sci 2022;81:100934.
[96] Waddock SA, Graves SB. The corporate social performance–financial performance link. Strateg Manag J 1997;18(4):303–19.
[97] Mumpower JL, Stewart TR. Expert judgement and expert disagreement. Think Reason 1996;2(2–3):191–212.
[98] Finn A, Louviere JJ. Determining the appropriate response to evidence of public concern: The case of food safety. J Public Policy Mark 1992;11(2):12–25.
[99] Cohen S. Maximum difference scaling: Improved measures of importance and preference for segmentation. In: Sawtooth software conference proceedings. Fir St., Sequim, WA: Sawtooth Software, Inc.; 2003, p. 61–74.
[100] Goodman S. An international comparison of retail consumer wine choice. Int J Wine Bus Res 2009;21(1):41–9.
[101] Auger P, Devinney TM, Louviere JJ. Using best–worst scaling methodology to investigate consumer ethical beliefs across countries. J Bus Ethics 2007;70(3):299–326.
[102] Lee JA, Soutar GN, Louviere J. Measuring values using best-worst scaling: The LOV example. Psychol Mark 2007;24(12):1043–58.
[103] Johnson RM. Understanding HB: An intuitive approach. Sequim, WA: Sawtooth Software Inc; 2000.
[104] Zheng T, Zhang C, Shi Y, Chen D, Liu S. Influencing factors and clustering characteristics of COVID-19: A global analysis. Big Data Min Anal 2022;5(4):318–38.
[105] Ameye SA, Ojo TO, Adetunji TA, Awoleye MO. Is there an association between COVID-19 mortality and Human development index? The case study of Nigeria and some selected countries. BMC Res Not 2022;15(1):186.
[106] Kavuran G, Gökhan Ş, Yeroğlu C. COVID-19 and human development: An approach for classification of HDI with deep CNN. Biomed Signal Process Control 2023;81:104499.
[107] Salvador CE, Berg MK, Yu Q, San Martin A, Kitayama S. Relational mobility predicts faster spread of COVID-19: A 39-country study. Psychol Sci 2020;31(10):1236–44.
[108] Feng Z, Zou K, Savani K. Cultural antecedents of virus transmission: Individualism is associated with lower compliance with social distancing rules during the COVID-19 pandemic. J Personal Soc Psychol 2023;124(3):461.
[109] Schlaegel C, Gunkel M, Taras V. COVID-19 and individual performance in global virtual teams: The role of self-regulation and individual cultural value orientations. J Organ Behav 2023;44(1):102–31.
[110] Hodzic S, Ravselj D, Alibegovic DJ. E-Government Effectiveness and Efficiency in EU-28 and COVID-19. Central Eur Public Adm Rev 2021;19:159–80.
[111] Martínez-Córdoba P-J, Benito B, García-Sánchez I-M. Efficiency in the governance of the Covid-19 pandemic: Political and territorial factors. Glob Health 2021;17:1–13.
[112] Cao Y, Hiyoshi A, Montgomery S. COVID-19 case-fatality rate and demographic and socioeconomic influencers: Worldwide spatial regression analysis based on country-level data. BMJ open 2020;10(11):e043560.
[113] Sharpe WF. The sharpe ratio. Streetwise–Best J Portfolio Manag 1998;3:169–85.
[114] Bonaccorsi A, Daraio C, Simar L. Advanced indicators of productivity of universities: An application of robust nonparametric methods to Italian data. Scientometrics 2006;66(2):389–410.
[115] Simar L. Detecting outliers in frontier models: A simple approach. J Prod Anal 2003;20:391–424.
[116] Barnett V, Lewis T, et al. Outliers in statistical data, vol. 3, no. 1. New York: Wiley; 1994.
[117] De Witte K, Schiltz F. Measuring and explaining organizational effectiveness of school districts: Evidence from a robust and conditional benefit-of-the-doubt approach. European J Oper Res 2018;267(3):1172–81.
[118] Balaguer-Coll MT, Prior D, Tortosa-Ausina E. On the determinants of local government performance: A two-stage nonparametric approach. Eur Econ Rev 2007;51(2):425–51.
[119] Brodeur A, Gray D, Islam A, Bhuiyan S. A literature review of the economics of COVID-19. J Econ Surv 2021;35(4):1007–44.
[120] Rose SM, Paterra M, Isaac C, Bell J, Stucke A, Hagens A, et al. Analysing COVID-19 outcomes in the context of the 2019 Global Health Security (GHS) Index. BMJ Global Health 2021;6(12):e007581.
[121] Omberg RT, Tabarrok A. Is it possible to prepare for a pandemic? Oxford Rev Econ Policy 2022;38(4):851–75.
[122] Chen Y, Biswas MI. Impact of national culture on the severity of the COVID-19 pandemic. Current Psychol 2022;1–14.
[123] Antràs P, Redding SJ, Rossi-Hansberg E. Globalization and pandemics. Amer Econ Rev 2023;113(4):939–81.