Efficiency of Healthcare Systems in the first wave of COVID-19 - a technical efficiency analysis

Breitenbach, Marthinus C and Ngobeni, Victor and Aye, Goodness

University of Pretoria, National Treasury Republic of South Africa, University of Pretoria

10 May 2020

Online at https://mpra.ub.uni-muenchen.de/101440/
MPRA Paper No. 101440, posted 01 Jul 2020 12:36 UTC
The first 100 days of COVID-19 coronavirus – How efficient did country health systems perform to flatten the curve in the first wave?

Marthinus C. Breitenbach* Victor Ngobeni** Goodness C. Aye***

Abstract

In this novel paper, we make use of a non-parametric method known as Data Envelopment Analysis (DEA) to analyse the 31 most infected countries during the first 100 days since the outbreak of the COVID-19 coronavirus for the efficiency in containing the spread of the virus – a question yet to be answered in the literature. Our model showed 12 of the 31 countries in our sample were efficient and 19 inefficient in the use of resources to manage the flattening of their COVID-19 contagion curves. Among the worst performers were some of the richest countries in the world, Germany, Canada, the USA and Austria, with efficiency between 50 and 60 per cent - more inefficient than Italy, France and Belgium, who were some of those hardest hit by the spread of the virus.

JEL Classification: C6, D2, I1

Keywords: Pandemic, COVID-19, Flattening the Curve, Data Envelopment Analysis, Non-Pharmaceutical Interventions, Healthcare, Technical Efficiency, Healthcare system efficiency barometer.

* Department of Economics, University of Pretoria. Address: University of Pretoria. Lynnwood Rd, Hatfield, Pretoria, 0002.
** National Treasury, Department of Finance, South Africa.
*** Department of Economics, University of Pretoria. Address: University of Pretoria. Lynnwood Rd, Hatfield, Pretoria, 0002.
Corresponding author email: martin.breitenbach@up.ac.za
Corresponding author Marthinus C Breitenbach ORCID ID: orcid.org/0000-0003-4623-2379
1. Introduction

The new coronavirus (Covid-19) has spread to nearly every country of the world since it first emerged in China in late December 2019 (Newey and Gualland, 2020). Within four months it had spread to 185 countries and regions of the world and the Philadelphia Inquirer of 28 April (Lubrano, 2020), reported that more than 3 million people were known to be infected and more than 211,000 deaths had been recorded at this time. In a very timely study, Mbuvha and Marwala (2020) point out that the rapid spread necessitates rapid responses to mitigate the spread of COVID-19 under conditions of extreme uncertainty. Even in these early stages of the spread of the virus, many analysts and researchers have started to model the spread of the virus, often with very scant and unreliable data and often accompanied by probable outcomes. Mbuvha and Marwala (2020), for example, use the susceptible-infected-recovered (SIR) compartmental model with South African data, using the initial conditions inferred from China and Italy to study the spread of the virus during the first three months of the virus in South Africa. They acknowledge that it is too early to come to a definitive conclusion, but that in the case of South Africa, where it may seem as if the curve is flattening, that either the pandemic is at very early stages of progression or mitigating measures might have resulted in a slowdown in the spread and severity.

In the epidemiology literature, the main pandemic spread mitigating interventions are referred to as non-pharmaceutical interventions (NPIs). Typically, NPIs promote social distancing, case isolation and public hygiene (Correia, et al., 2020) – similar to the global interventions with COVID-19. The difference with COVID-19 is the widespread, almost complete lockdown of economies around the world. Nevertheless, what was found when studying the 1918 Spanish Flu in the USA, also an influenza-type pandemic, in the regions that acted sooner rather than later with NPIs, there were significant reductions in peak mortality and moderate reductions in cumulative mortality (see e.g. Bootsma and Ferguson, 2007; Markel et al., 2007; Hatchett et al., 2007). As a matter of fact, a recent study by Correia, et al. (2020) of the 1918 Flu in the USA, has shown that there was substantial variation among cities in the speed and aggressiveness of the measures taken, and there was a direct link between the speed and aggressiveness of interventions and containment of the spread as well as mortality rates. Eichenbaum, et al. (2020), taking a completely different approach, show the link between the spread of a pandemic and interactions with economic decisions. In short, they find that there is an inevitable short-run trade-off between the severity of a short-run recession caused by a pandemic (as people withdraw from demand and supply activity), and its health consequences. In other words, the more people withdraw from economic activity (refrain from social contact) the higher the speed at which the pandemic is ended.
Gourinchas (2020), confirms the consensus among many epidemiology studies that containment of the COVID-19 pandemic is of utmost importance. In short, he explains that the health system of any country puts an upper bound on the number of patients that can receive proper treatment (available number of hospital beds, nursing staff, number of equipped intensive care units, etc.). If the spread of the coronavirus cannot be slowed or contained quickly enough, the threat is almost beyond comprehension and according to Gourinchas, “with a 2 per cent case fatality rate baseline for overwhelmed health systems and 50 per cent of the world population infected, 76 million people or 1 per cent of the world population would die”.

Sufficient evidence exists that the speed and effectiveness with which governments act to contain the spread of COVID-19 (mainly through NPIs, e.g. lockdowns, social distancing measures, sanitary measures, etc.) will determine to a large extent what the final outcome will be on the world. The strength of health systems would mean very little if the spread is not contained quickly. This has already proven to be the case with the COVID-19 pandemic.

On a daily basis new data is released and new analysis get published on the number of tests performed in each country, number of new cases, recoveries and deaths, etc. However, the data that is probably most covered in the media, are those related to the containment of the spread of COVID-19. This data is used to plot curves that all countries around the world watch with abated breath in the hope that these curves would “flatten” so that the economically devastating lockdowns can come to an end. Invariably, the measures taken by each country and the outcomes achieved, are compared – either relative to a country perceived to have done well or to one perceived to have done worse. Whether the resources used in flattening the COVID-19 curve is used efficiently or not is a question yet to be answered in the literature. This paper fills this void in the literature on COVID-19 applying a benchmarking tool widely used to compare the efficiency of healthcare systems across the world – a healthcare system efficiency barometer. We do this by scientifically analysing data related to resources and outcomes achieved over the first 100 days of COVID-19. In this novel paper, we make use of a non-parametric method known as Data Envelopment Analysis (DEA) to analyse the 31 most infected countries during the first 100 days since the outbreak, for the efficiency in their response to the outbreak and containing the spread of the virus - efficient use of available resources to flatten the curve (stabilise the rate of infection). At the time of writing this paper, 90 per cent of the more than 3 million infected people resided in these 31 countries (Worldometer, 2020).

---

2 Technically, it is 108 days from the first case reported outside of China in Thailand, on 13 January to the cut-off date of 30 April 2020.
We do acknowledge that national healthcare systems are different among countries because of differing cultural norms, market regulations, policies, etc. However, although there are differences in terms of infrastructure, patient numbers, funding, and governance between the healthcare systems, they face similar challenges and have common goals. Assessing and comparing the performance of several national healthcare systems, according to Nolte et al. (2006), provides an opportunity for policy makers to determine how well a particular national healthcare system is performing relative to its international peers, understand how it works in order to identify good and bad practices, and finally find more effective approaches to achieve sustainability and better quality. Our aim here is to see how the different national healthcare systems perform relative to their international peers (in our paper the peers are the sample of countries chosen with the highest infection rates), that is, how efficient resource utilisation was to reach the objective of flattening the curve.

We find that the average technical efficiency score is 83.3 per cent. This shows that not all the countries were efficient and on average were operating below the efficiency frontier. They would on average need to improve their efficiency by 16.7 per cent. Specifically, 12 of the sampled 31 countries implemented the COVID-19 lockdown measures very quickly and were efficient in the use of tests, doctors and health spending to manage the COVID-19 pandemic at prevailing output levels.

The rest of the paper is organised as follows: Section 2 deals with the literature, Section 3 with the model, Section 4 with the data, Section 5 with the results and Section 6 concludes.

2. Literature review

DEA has been used extensively to analyse efficiency in the health sector globally. We could only find one paper that used DEA to analyse the efficiency of health systems in the midst of an epidemic or pandemic and using variables related to resources used in times of an epidemic or pandemic, as well as outputs achieved during an epidemic or pandemic. We discuss this paper below. However, DEA has been applied extensively to compare efficiency of health care facilities within countries and between countries, and we briefly deal with some of that literature here.

Because our paper compares efficiency between countries, we do not deal with the literature on country studies. For literature on efficiency studies among different healthcare facilities within a country, see for example Ngobeni, et al. (2020) who analysed technical efficiency of provincial public healthcare in South Africa; Campanella et al. (2017) who assessed the technical efficiency of 50 Italian hospitals; Alhassan et al. (2015) used DEA to assess the technical efficiency of 64 health facilities in Ghana; DEA was also used by Jarjue et al. (2015)
to determine the technical efficiency of 41 secondary healthcare centres in Gambia. Also see Chowdhury et al. (2010); Gannon (2005); Marschall and Flessa (2009); Akazili et al. (2008); Masiye (2007); Zere et al. (2006); Kirigia et al. (2001); and Kirigia et al. (2000).

Literature on efficiency between health care facilities among countries

Available studies comparing healthcare efficiency among countries use either parametric or non-parametric analytical techniques such as the stochastic frontier analysis (SFA) model or the Data Envelopment Analysis (DEA), in which the healthcare systems are modelled as production units (see Giuffrida and Gravelle, 2001; Hollingsworth, 2003). As this study implements DEA as a method to compute efficiency across countries, the literature adopting DEA in this setting is discussed here. Bhat (2005) used DEA to assess the impact of financial and institutional arrangements on the national healthcare system efficiency in a sample of 24 OECD countries. He found countries having public-contract and public integrated- based healthcare systems are more efficient than public reimbursement-based systems. Lo Storto and Goncharuk (2017) employed DEA to measure the technical efficiency of 32 European (EU) countries. DEA was applied to compute two performance indices, measuring efficiency and effectiveness of these healthcare systems. The results of the study emphasize that the national healthcare systems achieve different efficiency and effectiveness levels. Comparing the efficiency and effectiveness scores, the authors identified a group of countries with the lowest performing healthcare systems that need to implement healthcare reforms aimed at reducing resource intensity and increasing the quality of medical services. Afonso and St Aubyn (2006) used a two-stage DEA to estimate a semi-parametric model of the healthcare systems in 30 OECD countries the years 1995 and 2003. Conventional and bootstrapped efficiencies are estimated in the first stage and corrected in the second stage by considering non-discretionary variables such as GDP per capita, education level, and health behaviour using a Tobit regression. Results show that a large amount of inefficiency is related to variables that are beyond government control.

Varabyova and Schreyögg (2013), used unbalanced panel data from OECD countries between 2000 and 2009 to compare the relative efficiency of healthcare systems. They took a different approach by performing two-step DEA and one-stage SFA and evaluate the internal and external validity of their findings by means of the Spearman rank correlations. They found that countries with higher health care expenditure per capita have on average a more efficient healthcare sector, and lower efficient healthcare is prevalent in countries with higher income inequality. Gonzalez et al. (2010), in a cross-sectional study using 2004 data, measured the technical and value efficiency of health systems in 165 countries. They used the amount of expenditure on health and education as inputs to the
healthcare system and data on healthy life expectancy and disability adjusted life years as health outcomes. Their study revealed that high income OECD countries have the highest efficiency indexes. De Cos and Moral-Benito (2014) estimated alternative measurements of efficiency using DEA and SFA between 1997 and 2009 ascertain the most important determinants of healthcare efficiency across 29 OECD countries. They provide empirical evidence that there are significant variance with respect to the level of efficiency in healthcare services provision among countries. Hadad et al. (2013) compared healthcare system efficiency of 31 OECD countries using conventional efficiency, super-efficiency and cross-efficiency and two model specifications, one including inputs under management control and the other inputs beyond management control. The results were ambiguous. Kim and Kang (2014), using a bootstrapped DEA, estimated the efficiency of healthcare systems in a sample of 170 countries. They divided the sample in four groups to obtain homogeneous sub-samples with respect to income. They found for a small number of the countries were able to manage their healthcare systems efficiently, that average efficiency in the high-income sub-sample was relatively high. Frogner et al. (2015), in a sample of 25 OECD countries, measured healthcare efficiencies between 1990 and 2010. They applied country fixed effects, country and time fixed effect models, and SFA including a combination of control variables reflecting healthcare resources, behaviours and economic and environmental factors. Rankings were found not to be robust due to different statistical approaches. Kim et al. (2016) estimated productivity changes in the healthcare systems of 30 national healthcare systems during 2002-2012. To analyse changes in productivity, efficiency and technology, they used the bootstrapped Malmquist index. They found for most countries in the sample, that recent policy reforms in the OECD countries stimulated productivity growth.

Literature on epidemics/pandemics and efficiency

Literature that deal specifically with the analysis of the outbreak of epidemics and pandemics, can be divided into the epidemiology literature and efficiency literature. In the epidemiology literature, the main focus is on the containment of the spread of an epidemic or pandemic; specifically emphasis is placed on the effectiveness of the NPIs. Correia, et al (2020), revisited the 1918 Flu in the USA, also an influenza-type epidemic, and found that in the regions that acted sooner rather than later with NPIs, there were significant reductions in peak mortality and moderate reductions in cumulative mortality (see e.g. Bootsma and Ferguson, 2007; Markel et al., 2007; Hatchett et al., 2007). Eichenbaum, et al. (2020), taking a completely different approach, show the link between the spread of an epidemic and interactions with economic decisions. In short, they found that there is an inevitable short-run trade-off between the severity of a short-run recession caused by an epidemic (as people withdraw from demand and supply activity), and the health consequences of the epidemic. In other words, the more
people withdraw from economic activity (refrain from social contact) the higher the speed at which the epidemic is ended. Mbuvha and Marwala (2020), found that in South Africa, the curve is flattening, that either the pandemic is at very early stages of progression or the NPIs resulted in a slowdown in the spread and severity. Gourinchas (2020), confirms the consensus among many epidemiology studies that containment of the COVID-19 pandemic is of utmost importance. In short, he explains that the health system of any country puts an upper bound on the number of patients that can receive proper treatment (available number of hospital beds, nursing staff, number of equipped intensive care units, etc.). If the spread of the coronavirus cannot be slowed or contained quickly enough, the threat is almost beyond comprehension.

In regard to the efficiency literature on epidemics, Jouzdani (2020), does a brief evaluation of the global fight against COVID-19 using confidence interval and the temporal confirmed, death, and recovered cases data. He presented a statistical method to visualize and distinguish the countries with conditions that call for close monitoring and international attention. He found that Iran, the United States, Iraq, and San Marino are the regions requiring more attention while Singapore, Malaysia, Vietnam, and Macau performed most effectively and efficiently in outbreak response management. Shirouyehzad et al. (2020), evaluate the efficiency of countries affected by COVID-19 considering their population density and health system infrastructure using Data Envelopment Analysis (DEA). The study is conducted in two steps. In the first step, considering their performance in contagion control of the disease, the efficiency values of the countries are estimated. In the second step, a comparison is made based on performance in medical treatment of the patients that could benefit from decreasing the number of death cases and increasing the number of recovered cases. The countries are classified into four classes based on their performance in contagion control and medical treatment and it was found that that Singapore, Vietnam, and Belgium are the countries with the highest efficiency in both aspects. Singapore with one of the highest population densities in the Southeast Asia, has the highest efficiency among the countries. In Europe, Italy is the least and Belgium the most efficient. In the Middle East, Egypt has been the least efficient in contagion control but most efficient in medical treatment, while Iran has been the most efficient in contagion control.

This literature review shows that empirical work pivot mostly on healthcare system performance based on the efficiency calculated as a ratio of a measure of some quality of life variable as an output and the physical health resources or expenditure on health as inputs. As will gradually become evident, our paper is similar in regard to the modelling approach followed in the literature, but different in respect to our choice of input and output variables. Although Shirouyehzad et al. (2020) uses DEA to analyse the efficiency of contagion of
COVID19, they focus on the number of deaths and recoveries as outcomes, while the express aim of this paper is to focus on flattening the curve as the main outcome/output.

3. Modelling Approach

In this paper, we use the variable returns to scale (VRS) approach reported by Gavurova et al. (2017) and developed in 1984 by Banker, Charnes and Cooper to allow for consideration of scale efficiency analysis. This is called the Banker, Charnes and Cooper (BCC) model. The terminology “envelopment” in DEA refers to the ability of the efficiency production frontier to tightly enclose the production technology (input and output variables). Cooper et al. (2007) and McWilliams et al. (2005) state that DEA was developed in a microeconomic setting and applied to firms to convert inputs into outputs. However, in efficiency determination, the term “firm” is often replaced by the more encompassing DMU. DEA is an appropriate method of computing the efficiency of institutions employing multivariate production technologies. Aristovnik (2012) and Martić et al. (2009) state that there are input-minimisation and output-maximisation DEA models. The former determines the quantity of inputs that could be curtailed without reducing the prevailing level of outputs and the latter expands outputs of DMUs to reach the production possibility frontier while holding inputs constant. However, the selection of each orientation is study-specific.

According to Taylor and Harris (2004), DEA is a comparative efficiency measurement tool that evaluates the efficiency of homogeneous DMUs operating in similar environmental conditions, for example, DMUs dealing with COVID-19 and where the relationship between inputs and outputs is unknown. Wang and Alvi (2011) report that DEA only uses the information used in a particular study to determine efficiency and does not consider exogenous factors. DEA measures the distance of production functions to determine the radial extent of DMUs to efficiency frontiers by categorising the DMUs into extremely efficient and inefficient performers. In terms of the DEA methodology, the current study uses the BCC model with the ratio of DMUs being 6 times the combined number of inputs and outputs to ensure the stability of the efficiency results. However, before explaining the BCC model, it is prudent to describe first the constant returns to scale (CRS) model, developed by Farrell in 1957 and enhanced in 1978 by Charnes, Cooper and Rhodes (also called the CCR model) to convert the fractional linear efficiency estimates into linear mathematical efficiency programmes under the constant returns to scale (CRS). These models are described in the following paragraphs.

Under the CCR model, suppose there are $C$ different number of inputs and $D$ different number of outputs for $N$ DMUs. These quantities are represented by column vectors $x_{ij}$ ($i = 1, 2, 3, \ldots C, j = 1, 2, 3, \ldots N$) and $q_{rj}$ ($r = 1, 2, 3, \ldots D, j = 1, 2, 3, \ldots N$) The $C \times N$ input matrix, $X$ and $D \times N$ output matrix, $Q$ represents the production technology for all the $N$ number of DMUs. For each
DMU, the ratio of all the output variables over all the input variables is represented by $u'q/v'x$. Where $u = D \times 1$ vector output weights and $v = C \times 1$ vector input weights. The optimal weights or the efficiency estimates are obtained by solving a mathematical problem. In the context of the CRS, an efficient DMU operates at technically optimal production scale (TOPS). Hence, the optimal weights or efficiency estimates are obtained by solving a mathematical problem that is reflected in equation 1.

$$\text{Tops} = \max_{u,v} (u'q/v'x)$$

Subject to:

$$u'q/v'x \leq 1$$

$$u, v \geq 0$$

Equation 1 shows the original linear programme, called the primal. It aims to maximise the efficiency score, which is represented by the ratio of all the weights of outputs to inputs, subject to the efficiency score not exceeding 1, with all inputs and outputs being positive. Equation 1, has an infinite number of solutions, if $(u,v)$ is a solution, so is $\alpha v, \alpha v$. To avoid this, one can impose a constraint $v'x = 1$, which produces equation 2.

$$\max_{u,v} (u'q)$$

Subject to:

$$v'x = 1$$

$$u'q - v'x \leq 0$$

$$u, v \geq 0$$

Equation 2 is an equivalent envelopment problem. An equivalent envelopment problem can be developed for the problem in equation 2, using duality in linear programming. The dual for $\max_{u,v} (u'q)$ is $\min \theta, \lambda \theta$. The value of $\theta$ is the efficiency score; it satisfies the condition $\theta \leq 1$; it is the scalar measure. Lauro et al. (2016) report that $\lambda$ is an $N \times 1$ vector of all constants representing intensity variables indicating necessary combinations of efficient entities or reference units (peers) for every inefficient DMU, it limits the efficiency of each DMU to be greater than 1. This results in equation 3, which represents the CCR-CRS model with an input minimisation orientation.
Avkiran (2001) states that the CRS postulates no significant relationship between DMU’s operational size and their efficiency. That is, under the CRS assumption, the large DMUs are deemed to attain the same levels of efficiency as small DMUs in transforming inputs to outputs. Therefore, the CRS assumption implies that the size of a DMU is not relevant when assessing technical efficiency. However, in most cases, DMUs have varying sizes and this becomes a factor when determining their efficiency. As a result, Gavurova et al. (2017) mention that in 1984, the CCR formulation was generalised to allow for the VRS. Aristovnik (2012) adds that, if one cannot assume the existence of the CRS, then a VRS type of DEA is an appropriate choice for computing efficiency. Gannon (2005) advises that the VRS should be used if it is likely that the size of a DMU will have a bearing on efficiency. As such, Yawe (2014) cautions that the use of the CRS specification when the DMUs are not operating at an optimal scale results in a measure of technical efficiency which is confounded by scale effects. The solution is to use the VRS as it permits for the calculation of scale inefficiency. The CRS linear programming problem can be modified to account for the VRS by adding the convexity constraint: $N'$\*l = 1 to equation 3, where $N'$ is a $N$x1 vector of ones to formulate equation 4. Equation 4 represents the BBC-VRS model with an input-minimisation orientation. Therefore, equations 1 to 3 represent the CRS models while equations 4 to 5 represents the VRS models.

Lauro et al. (2016) and Yuan and Shan (2016) report that the CCR and the BCC models only differ in the manner the latter includes convexity constraints. Since the current model considers the VRS, the restriction $\sum_{d=1}^{n} \lambda n = 1$ is introduced. Ramírez Hassan (2008) cautions that, if this restriction is not there, it would imply the application of the CRS model. The same analogy applies to all the inefficient DMUs in the sample. That is, the slacks and the radial
movements are calculated for all inefficient DMUs using equation 5. The BCC is adept to calculate pure technical efficiency and inefficiency and when applied with the CCR model, it also measures scale inefficiency. Where, $\sum_{i=1}^{I} \lambda l = 1$, a DMU is on a CRS frontier, if $\sum_{i=1}^{I} \lambda l < 1$, the DMU is located on the IRS frontier and if $\sum_{i=1}^{I} \lambda l > 1$, there is DRS. Given that this study has adopted both the CCR and the VRS with an input-minimisation orientation. The DEA models used in this study also consider the slack movements for the inefficient DMUs. As a result, the models account for the slacks in equation 5.

Min $\theta, \lambda j, S r^+, S i^-$

$$\theta - \varepsilon \left[ \sum_{i=1}^{C} S i^- + \sum_{r=1}^{D} S i^+ \right]$$

St.

$$\theta x_{i0} - \sum_{j=1}^{N} x_{ij} \lambda j - S i^- = 0,$$

$$\theta q_{r0} = \sum_{j=1}^{N} q_{rj} \lambda j - S r^+ = 0.$$

$$\sum_{j=1}^{N} \lambda j = 1$$

$$\lambda j, S r^+, S i^- > 0$$

Coelli et al. (2005) define slacks as input excesses and output shortfalls that are required over and above the initial radial movements to push DMUs to efficiency levels. Both the slack and radial movements are associated only with the inefficient DMUs. The radial movements are initial input contractions or output expansions that are required for a firm to become efficient. $S i^+$ and $S i^-$ in equation 5 are the output and input slacks respectively to be calculated with $\theta$, $\lambda n$. $\varepsilon$, is the non-Archimedean constant. Gavurova et al. (2017) hint that if the slack variables of a DMU are not equal to zero and the technical efficiency score is lower than one, it is necessary to perform a non-radial shift that is expressed by the slack variables to achieve technical efficiency. In equation 5, the slack variables determine the optimum level of inputs that DMUs would have to utilise and the outputs that they would have to produce to become efficient, provided that these DMUs are inefficient. Therefore, the slacks depict the under-produced outputs or overused inputs.
4. Data

We selected 31 countries (see Table 2) based on the fact that these countries collectively represented 90 per cent of COVID-19 cases reported at the time of this analysis, which was 100 days since the first case was reported outside of China. This study measures the efficiency of 31 countries in using the available resources to flatten the COVID-19 curve. Of the selected countries, excluding South Africa, 30 were selected as they were the top 30 countries with COVID-19 infections as at end 30 April 2020 (a cut-off period for the study which started on 31 December 2019, based on data from Worldometer (2020). South Africa was included as a country with most infections in Africa to represent the continent, as the sample had no African country. DMUs with an efficiency score of 1, are technically efficient and serve as benchmarks for inefficient countries with scores of less than 1.

The study uses an input-minimisation variable returns to scale (VRS) data envelopment analysis (DEA) model\(^3\). The inputs and outputs for the model are shown in Table 1. The inputs are number of days to lockdown, in line with the work of Correia, et al. (2020), Barro, et al. (2020), Garrett (2008), Bootsma and Ferguson (2007) and Markel, et al. (2007), number of doctors per 1000 of the population, total tests per 1 million population and spending on health as percentage of GDP. It is noted that with exception of the number of days to lock down, the other three inputs are not unique to COVID-19 though they are very standard input variables in the DEA health literature (see e.g. Giuffrida and Gravelle, 2001; Hollingsworth, 2003; Bootsma and Ferguson, 2007, Campanella et al., 2017; Anton, 2013; Marschall and Flessa, 2009; Markel, et al., 2007). For our main output variable, we were interested in data relating to the speed at which the curve was being flattened, i.e. how quickly countries are able to contain the rate at which the virus spreads. For this reason, our choice of output variable differs for example from Shirouyehzad et al. (2020) who chose as outputs and number of confirmed cases (stage 1) and in stage 2, two outputs, which are the number of the recovered cases and the number of death cases.

As the other input variables are standard, the analysis will focus mainly on number of days to lock down. Figure 1 shows a quick scatter plot relating the number of days it took countries to introduce lockdown and the number of days to flatten the curve. A quick inspection of the plot shows that there is a positive correlation, in other words, quicker introduction of lockdowns are generally associated with shorter periods flattening the curve. It took the 31 countries an average of 21 days to institute lockdown, and an average of 51 days to flattening the curve.

\(^3\) We also performed output maximization. The results were almost identical to the input minimization, which confirms the robustness of the results. The results are available on request from the authors.
The original output variable for this study was the critical point to persistently reduced COVID-19 infections; which is the number of days from the start of each country’s cycle up to a point where the COVID-19 infections start to fall consistently, showing that infections are substantially declining. For short, this is the inflection point on the COVID-19 trajectory curve. This was derived by computing 14-day rolling averages across the study period for each country; with the corresponding dates yielding the number of days at the critical points (where the exponential trends of the epi-curves ended and deceleration started) (see Appendix 1). DEA benchmarks data samples in such a way that high values represent efficient units. Therefore, the critical point for persistent COVID-19 reduced infections (the actual selected output variable) presented a challenge where countries reaching the critical point earlier would be represented by low values. To prevent DEA model misspecification and skewing the efficiency results, this output was replaced with an inverted measure, the number of days left in a cycle after a country reaches the critical point of persistently reduced COVID-19 infections. However, in the interpretation of the results either one of these variables or both can be used. This measure was calculated for each country by subtracting the days at the critical point of persistently reduced COVID-19 infections from the total days in a country’s cycle. This variable assigns higher output values to DMUs who swiftly reached the critical point and low values to those who delayed and zero to DMUs still showing a consistent pattern of increasing COVID-19 infections by 30 April 2020. Data were obtained from the International Monetary Fund, Johns Hopkins University, The World Bank and Worldometer.
Table 1: Analytical variables

<table>
<thead>
<tr>
<th>Model</th>
<th>DEA Model</th>
<th>Number of variables</th>
<th>Variable description</th>
</tr>
</thead>
</table>
| COVID-19 Model | VRS       | 5                   | Output 1: Number of days left in a cycle after reaching persistent COVID-19 reduced infections (spared days)  
Input 1: Number of days to lockdown  
Input 2: Number of doctors per 1000 population  
Input 3: Total tests per 1 million population  
Input 4: Spending on health % of GDP |

Source: Authors’ table.

5. Results

The results of the efficiency analysis are reported in Table 2 and Figure 2. The mean technical efficiency score of the 31 DMUs was 83.3 per cent as reflected in Table 2, showing that most of the sampled countries were operating close to the efficiency frontier. Of the 31 countries, 12 or 38.7 per cent were efficient in flattening the COVID-19 curve by instituting lockdown measures, conducting tests, using the available doctors and prevailing levels of health spending to gross domestic product (GDP). Put differently, these 12 countries minimised (optimised) the use of all inputs in the model or used inputs efficiently in the quest to flatten the infection curve. These 12 DMUs serve as benchmark and offer possible lessons in the optimal use of the production technology under consideration for the 19 or 61.3 per cent inefficient DMUs who should, on average, improve efficiency by 16.7 per cent.⁴

Table 2: Results of VRS input-minimisation

---

⁴ As mentioned by way of an earlier footnote, we also performed an output-maximisation DEA which confirms the efficiency scores obtained from the input-minimisation model. It confirms the robustness of our results and is available on request from the authors.
<table>
<thead>
<tr>
<th>Country</th>
<th>VRS efficiency score</th>
<th>Efficient lockdown days*</th>
<th>Actual number of days to flatten the curve</th>
<th>Spared days in a full cycle after flattening the curve</th>
</tr>
</thead>
<tbody>
<tr>
<td>Turkey</td>
<td>1.000</td>
<td>39</td>
<td>38</td>
<td>8</td>
</tr>
<tr>
<td>Iran</td>
<td>1.000</td>
<td>31</td>
<td>48</td>
<td>22</td>
</tr>
<tr>
<td>Brazil</td>
<td>1.000</td>
<td>12</td>
<td>50</td>
<td>0</td>
</tr>
<tr>
<td>China</td>
<td>1.000</td>
<td>24</td>
<td>47</td>
<td>75</td>
</tr>
<tr>
<td>Peru</td>
<td>1.000</td>
<td>13</td>
<td>49</td>
<td>0</td>
</tr>
<tr>
<td>India</td>
<td>1.000</td>
<td>14</td>
<td>58</td>
<td>0</td>
</tr>
<tr>
<td>Ecuador</td>
<td>1.000</td>
<td>1</td>
<td>54</td>
<td>0</td>
</tr>
<tr>
<td>Sweden</td>
<td>1.000</td>
<td>0</td>
<td>64</td>
<td>0</td>
</tr>
<tr>
<td>Singapore</td>
<td>1.000</td>
<td>60</td>
<td>62</td>
<td>0</td>
</tr>
<tr>
<td>Pakistan</td>
<td>1.000</td>
<td>15</td>
<td>64</td>
<td>0</td>
</tr>
<tr>
<td>Chile</td>
<td>1.000</td>
<td>14</td>
<td>59</td>
<td>0</td>
</tr>
<tr>
<td>South Africa</td>
<td>1.000</td>
<td>21</td>
<td>59</td>
<td>0</td>
</tr>
<tr>
<td>Russia</td>
<td>0.997</td>
<td>13</td>
<td>49</td>
<td>0</td>
</tr>
<tr>
<td>Saudi Arabia</td>
<td>0.960</td>
<td>20</td>
<td>59</td>
<td>0</td>
</tr>
<tr>
<td>Ireland</td>
<td>0.940</td>
<td>8</td>
<td>51</td>
<td>7</td>
</tr>
<tr>
<td>Switzerland</td>
<td>0.896</td>
<td>9</td>
<td>36</td>
<td>28</td>
</tr>
<tr>
<td>Portugal</td>
<td>0.881</td>
<td>12</td>
<td>18</td>
<td>38</td>
</tr>
<tr>
<td>Japan</td>
<td>0.857</td>
<td>23</td>
<td>84</td>
<td>10</td>
</tr>
<tr>
<td>Belarus</td>
<td>0.802</td>
<td>18</td>
<td>60</td>
<td>0</td>
</tr>
<tr>
<td>Mexico</td>
<td>0.757</td>
<td>11</td>
<td>59</td>
<td>0</td>
</tr>
<tr>
<td>Spain</td>
<td>0.703</td>
<td>11</td>
<td>38</td>
<td>25</td>
</tr>
<tr>
<td>Israel</td>
<td>0.670</td>
<td>23</td>
<td>40</td>
<td>22</td>
</tr>
<tr>
<td>Italy</td>
<td>0.646</td>
<td>11</td>
<td>61</td>
<td>9</td>
</tr>
<tr>
<td>UK</td>
<td>0.646</td>
<td>12</td>
<td>49</td>
<td>9</td>
</tr>
<tr>
<td>Belgium</td>
<td>0.634</td>
<td>11</td>
<td>38</td>
<td>22</td>
</tr>
<tr>
<td>Netherlands</td>
<td>0.634</td>
<td>10</td>
<td>47</td>
<td>15</td>
</tr>
<tr>
<td>France</td>
<td>0.606</td>
<td>9</td>
<td>41</td>
<td>19</td>
</tr>
<tr>
<td>Austria</td>
<td>0.588</td>
<td>14</td>
<td>34</td>
<td>28</td>
</tr>
<tr>
<td>USA</td>
<td>0.556</td>
<td>8</td>
<td>46</td>
<td>11</td>
</tr>
<tr>
<td>Canada</td>
<td>0.532</td>
<td>10</td>
<td>62</td>
<td>0</td>
</tr>
<tr>
<td>Germany</td>
<td>0.528</td>
<td>13</td>
<td>40</td>
<td>24</td>
</tr>
<tr>
<td>Mean</td>
<td>0.833</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>


Note: Zero spared days implies that the country has not yet flattened the COVID-19 curve and has been allocated the full duration from first reported cases to end April 2020. To prevent model misspecification, the number of spared days (total days less days when the curve was flattened) was used to ensure that underperforming countries are assigned a lower number.

Note: * Purely for illustrative purposes; although countries that are below the efficiency frontier could reach that frontier by optimising the use of inputs in many different combinations of resource reductions, we use the number of days to lockdown to illustrate the relative improvements required by each country to become efficient.
The 50 per cent cohort (4 or 13 per cent of DMUs)

Four countries, Germany, Canada, the United States of America and Austria had efficiency scores between 50 and 59 per cent, needing to improve the use of resources in the management of COVID-19 by 41 to 50 per cent. Germany had the lowest efficiency score of 52.8 per cent. In the 64 days period (27 February to 30 April 2020) pertaining to Germany’s timeline, to be efficient in flattening the COVID-19 curve, the country needed to use fewer resources, for example imposing lockdown measures 13 days from the first reported case that is precisely on 10 March 2020\(^5\). In the 62 days related to analysing Canada, the country did not manage to reach the critical point (zero spared days). Canada recorded an efficiency score of 53.2 per cent. To improve this performance relative to efficient peers in the sample, the country needed to improve efficiency by 46.8 per cent, e.g. by imposing lockdown measures within 10 days (saving 8 days) of reporting the first COVID-19 cases. The United States of America’s inefficiency score was 44.4 per cent. Relative to efficient countries, this inefficiency rate could have been improved by using fewer resources, for example by imposing lockdown

\(^5\) Purely for illustrative purposes; although countries that are below the efficiency frontier could reach that frontier by optimising the use of inputs in many different combinations of resource reductions, we use the number of days to lockdown to illustrate the relative improvements required by each country to become efficient.
measures within 8 days of reporting the first COVID-19 cases, instead of two weeks. Austria is the last and best-performing country in this cohort, realising a score of 58.8 per cent. Benchmarked against efficient peers, the country should improve efficiency in the use of resources by 41.2 per cent.

**The 60 per cent cohort (6 or 19.4 per cent of DMUs)**

France was the least efficient DMU in this category with a score of 60.6 per cent. To reach the optimal efficiency frontier relative to its efficient peers, France could improve the efficiency of resources by 39.4 per cent. Using lockdown days as illustration, it should have acted with lockdown in 9 days (opposed to 15 days) after reporting the first COVID-19 case. The Netherlands and Belgium had efficiency scores of 63.4 per cent, needing to improve relative efficiency by 36.6 per cent. This implies that the resources used in flattening the COVID-19 curve should have been fewer. The United Kingdom and Italy recorded efficiency scores of 64.6 per cent. To become relatively efficient in the use of resources for flattening the COVID-19 curve, the United Kingdom and Italy could have used resources 35.4 per cent more efficiently. Israel was the best performer in this cohort with a score of 67 per cent needing to improve efficiency by 13 per cent. To reach top efficiency, Israel would have for example, had to institute lockdown in 23 days, instead of 34 days after recording the first COVID-19 case.

**The 70 to 89 per cent cohort (6 or 19.4 per cent of DMUs)**

Spain, Mexico, Belarus, Japan, Portugal and Switzerland fall in this category. Spain was one of nations mostly affected by the pandemic. Its efficiency score is 70.3 per cent. To improve its efficiency relative to peers, it was supposed to have used fewer resources in flattening the COVID-19 curve such as implementing lockdown in 11 days (instead of 16 days) after reporting the first COVID-19. Mexico needed to target the same use of resources as Spain in respect of lockdown days. As it pertains to Belarus, it had a relative efficiency improvement target of 19.8 per cent. It could for example reach the frontier by reducing the number of days to lockdown from 22 to 18. Japan recorded an efficiency score of 85.7 per cent, while Portugal recorded an efficiency score of 88.1 per cent. They need to respectively improve the efficiency with which they use resources by 14.3 and 11.9 per cent. Switzerland recorded an efficiency score of 89.6 per cent; it maintained the critical point of persistently reducing the COVID-19 infections at 36 days, sparing 38 days. To be fully efficient, fewer resources could have been used. For instance, lockdown could have been introduced in 9 days (3 days earlier) after reporting the first COVID-19 case.

**The 90 to 99 per cent cohort (3 or 1 per cent of DMUs)**

Ireland, Saudi Arabia and Russia were very close to the optimal technical production scale. They had to respectively improve relative efficiency by 6, 4 and 0.3 per cent. Ireland had an
output of 7 spared days after reaching the critical point to persistently reducing the COVID-19 infections. To reach the maximum efficiency level, it needed to use fewer resources such as locking down 8 days after reporting the first COVID-19 case. Saudi Arabia and Russia had zero spared days as outputs. In their effort to reach efficiency, they had to lockdown within 20 (1 day less) and 13 days respectively – again as an illustration of relative improvements in efficiency required to reach the frontier.

6. Conclusions

The study analysed the technical efficiency with which health systems in our sample of 31 countries were able to contain the spread of COVID-19 infections (flattening the curve) using the input-oriented DEA methodology. This means that those countries that were on the frontier with an efficiency score of one, managed to reach that frontier by optimising the combination of inputs available to them. In this study, those inputs were health expenditure, speed at which lockdowns were implemented, number of doctors/thousand and number of tests performed/million of the population. The average technical efficiency score is 83.3 per cent. This shows that not all the countries were efficient and on average, are operating below the frontier. They would on average need to improve relative efficiency by 16.7 per cent, in other words, on average, the countries operating below the frontier, could have reached the same outcomes in terms of flattening the curve by using fewer resources such as locking down the economy earlier than they did.

Specifically, 12 of the 31 countries in our sample implemented the COVID-19 lockdown measures very quickly and were efficient in the use of resources to manage the flattening of their COVID-19 contagion curves. Given the objective to minimise the use of inputs in order to be efficient, the remaining 19 countries used their available resources inefficiently. Among the worst performers were some of the richest countries in the world, Germany, Canada, the USA and Austria, who obtained efficiency scores between 50 and 60 per cent. These countries were more inefficient applying their available resources to flatten the curve than countries like Italy, France and Belgium, who were some of those hardest hit by the spread of the virus.

In some sense these findings are in line with the recent study by Shirouyehzad et al. (2020), who found that some developing countries with a high population density and low International Health Regulations Core Capacity Scores (IHRCCS), were more efficient in contagion control of COVID-19, even though they may be worse in offering medical treatment. For example, Shirouyehzad et al. (2020) found that Singapore had the highest efficiency among the countries even with one of the highest population densities in the Southeast Asia, and was far ahead of others. In the Middle East, Iran has been the most efficient contagion control, and
although Egypt was worse, it was more efficient in medical treatment. This finding underscores the importance of using NPIs as efficiently as possible to stop the spread of an epidemic or pandemic (flattening the curve) in the shortest possible time, especially in countries with inferior health systems.

This study is limited in several ways. It should be noted that the results are obtained based on the data gathered during the (roughly) 100 days of what is perceived to be the first wave of the spread of COVID-19. Therefore, the results should not be generalised to other time periods and extrapolation should be done with caution. The selection of the indicators affects the outcomes of the model. Therefore, a different set of indicators may lead to a different collection of results and analyses.

References


https://coronavirus.jhu.edu/map.html


https://bmchealthservres.biomedcentral.com/articles/10.1186/1472-6963-7-58


https://www.researchgate.net/publication/320508371_Efficiency_vs_Effectiveness_a_Benchmarking_Study_on_European_Healthcare_Systems [accessed May 08 2020].


https://data.worldbank.org/indicator/SH.XPD.CHEX.GD.ZS
https://data.worldbank.org/indicator/SH.MED.PHYS.ZS


https://www.worldometers.info/coronavirus/


Appendix 1 Moving average of flattening COVID-19 case curve (infection trajectory curve)