UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
FACULDADE DE CIÊNCIAS ECONÔMICAS
PROGRAMA DE PÓS-GRADUAÇÃO EM ECONOMIA

ARTHUR SONNTAG KUCHENBECKER

MEASURING AND EXPLAINING EFFICIENCY OF DIFFERENT COUNTRIES RESPONSES TO COVID-19 PANDEMIC: A CONDITIONAL ROBUST NONPARAMETRIC APPROACH

Porto Alegre
2021

ARTHUR SONNTAG KUCHENBECKER

MEASURING AND EXPLAINING EFFICIENCY OF DIFFERENT COUNTRIES RESPONSES TO COVID-19 PANDEMIC: A CONDITIONAL ROBUST NONPARAMETRIC APPROACH

Dissertação submetida ao Programa de Pós-Graduação em Economia da UFRGS, como quesito parcial para obtenção do grau de Mestre em Economia.

Orientador: Prof. Dr. Flavio Augusto Ziegelmann

Porto Alegre
2021

CIP - Catalogação na Publicação

Sonntag Kuchenbecker, Arthur
Measuring and explaining efficiency of different countries responses to COVID-19 pandemic: a conditional robust nonparametric approach / Arthur Sonntag Kuchenbecker. -- 2021.
46 f.
Orientador: Flavio Augusto Ziegelmann.

Dissertação (Mestrado) -- Universidade Federal do Rio Grande do Sul, Faculdade de Ciências Econômicas, Programa de Pós-Graduação em Economia, Porto Alegre, BR-RS, 2021.

1. COVID-19 response. 2. Efficiency analysis. 3. Nonparametric methods. I. Ziegelmann, Flavio Augusto, orient. II. Título.

Elaborada pelo Sistema de Geração Automática de Ficha Catalográfica da UFRGS com os dados fornecidos pelo(a) autor(a).

ARTHUR SONNTAG KUCHENBECKER

MEASURING AND EXPLAINING EFFICIENCY OF DIFFERENT COUNTRIES RESPONSES TO COVID-19 PANDEMIC: A CONDITIONAL ROBUST NONPARAMETRIC APPROACH

Dissertação submetida ao Programa de Pós-Graduação em Economia da UFRGS, como quesito parcial para obtenção do grau de Mestre em Economia.

Aprovada em: Porto Alegre, 18 de outubro de 2021.

BANCA EXAMINADORA:

Prof. Dr. Flavio Augusto Ziegelmann (Orientador)
UFRGS

Prof. Dr. Hudson Torrent
UFRGS

Prof. Dr. Marcelo Fernandes
FGV EESP

Prof. Dr. Paulo Jacinto
UFPR

ABSTRACT

In this paper, we propose the use of a conditional nonparametric robust estimator to evaluate countries responses to the outburst of COVID-19 pandemic. We collect data for 105 countries (comprehending the initial period of pandemic through the end of may/2021), with variables regarding the death toll, economic indicators, demographic characteristics and non-pharmaceutical interventions. We create a novel framework for estimating efficiency of countries responses in more general terms than simply evaluating healthcare system performance. We use two distinct well-known second-stage approaches: regressing the conditional efficiency scores on the environmental variables, in order to compute measures of managerial efficiency to rank responses; and regressing the ratio of conditional and unconditional scores on conditioning factors, seeking to explore the relationship between non-pharmaceutical interventions and the estimated efficiencies. Our results indicate which countries and regions stood out for presenting efficient/inefficient responses and point to a negative relationship between the variables *median age*, *average stringency index* and *average retail and recreation visitors change* and efficiency estimates.

Keywords: COVID-19 response. Efficiency analysis. Nonparametric methods.

RESUMO

O objetivo deste trabalho é propor a aplicação de um estimador não paramétrico para avaliar a eficiência das respostas dos países à eclosão da pandemia de COVID-19. São coletados dados de 105 países (compreendendo o período inicial da pandemia até o final de maio de 2021), com variáveis que englobam o número de mortos, indicadores econômicos, características demográficas e intervenções não farmacológicas. Ao longo do texto são apresentadas as premissas utilizadas, que constituem um arcabouço inovador para estimar eficiência das respostas em termos mais gerais do que a simples avaliação da atuação do sistema de saúde. São implementadas duas técnicas de estimação em dois estágios, amplamente utilizadas na literatura: regressão dos scores de eficiência estimados contra as variáveis ambientais, com o objetivo de mensurar a eficiência gerencial e ranquear as respostas dos países; e uma regressão das razões dos scores condicionais e não condicionais contra os fatores condicionantes, buscando explorar a relação entre as medidas não farmacológicas e as estimativas de eficiência. Os resultados indicam quais países e regiões se destacaram, apresentando respostas mais eficientes/ineficientes, bem como apontam para uma relação negativo que as variáveis *idade mediana*, *índice médio de restrição* e *alteração média no número de visitantes em lojas e locais recreativos* tiveram nas estimativas de eficiência.

Palavras-chave: Respostas ao COVID-19. Análise de eficiência. Métodos não paramétricos.

LISTA DE ILUSTRAÇÕES

LISTA DE TABELAS

LISTA DE ABREVIATURAS E SIGLAS

CDF	Cumulative Distribution Function
DEA	Data Envelopment Analysis
DGP	Data Generating Process
DMU	Decision Making Unit
FDH	Free Disposal Hull
GDP	Gross Domestic Product
GGD	General Government Debt
IMF	International Monetary Fund
NCD	Non Communicable Diseases
OWID	Our World in Data
UN	United Nations
WEO	World Economic Outlook

SUMÁRIO

1 INTRODUCTION

Since the seminal work of Koopmans (1951) and Debreu (1951), which established the beginning of a literature commonly known as production efficiency analysis, several improvements emerged towards the construction of better estimators for the efficiency of *decision making units (DMU)* in a particular sample. In general, what every developed estimator has in common is the origin on microeconomic theory framework, regarding the assumptions securing the existence of a *transformation frontier* for the production set of a given economic process. This frontier is sometimes simply referred to as *efficient frontier*. The referred estimators are based on the idea of measuring distance between observed production *(output)* of each decision unit and an estimate of the maximum feasible production, given the observed amount of *inputs* and the existent technology (assumed to be described by the efficient frontier). If a production function can be analytically defined to represent the efficient frontier, relating a quantity x (input) to a quantity y (output), the efficiency (or inefficiency) of a DMU is straightforwardly obtained by some measure of geometric distance between observed production and the given function. Most frequent are the cases, however, when the researcher does not know an analytic function relating inputs and outputs. Rather, the empirical problem is given not only by trying to estimate an efficiency score for a DMU, but also by the attempt to estimate the maximum feasible production itself, evaluating an observable sample of DMUs.

Notwithstanding the fact that this framework is applied mainly for analyzing efficiency in formal production processes, primarily understood as transformation of factors like labor and capital into tangible goods and services, it is also common to find papers where researchers establish relations between different variables to measure efficiency or to find out what explains efficiency in much more abstract applications. Some examples (on which the present work vastly uphold) include measuring efficiency of police on crime clear-ups (NEPOMUCENO et al., 2020), explaining students performance in heterogeneous environments (WITTE; KORTELAINEN, 2013), explaining banks efficiency under different regulatory conditions (MINVIEL; BOUHENI, 2021), evaluating the impact of market risks on mutual funds performance (BĂDIN; DARAIO; SIMAR, 2014) or even exploring factors that affect the level of happiness across countries (CORDERO; SALINAS-JIMÉNEZ; SALINAS-JIMÉNEZ, 2017).

All those cited applications have in common the use of a nonparametric framework that evolved from traditional efficiency scores estimation, trying to explain how efficiency is affected by external factors, and not only by direct managerial skills (DARAIO; SIMAR, 2005). These factors are often called *environmental variables* and a very important contribution of this developed framework is that it allows researchers to rank DMUs taking into account the fact that sometimes they are exposed to very heterogeneous environments that may influence their results in terms of efficiency, or even alter the production frontier itself (BĂDIN; DARAIO; SIMAR, 2012). This methodology also allows one to consider specific actions that DMUs take as possibly explaining the level of efficiency and test whether and how it is related to the estimated efficiency scores. As an example, imagine a scenario where firms produce bread using some known inputs as grain, water etc., and have a wide range of production mechanisms to choose. Also, assume it is not possible to completely describe what production process each firm is using, so that the bare estimation of efficiency scores won't say much about what is the most efficient way to transform inputs into outputs. Even in this case, using the methodology described later in this paper, the researcher might be able to find out and test if some measurable indicators (say minutes in the oven, or the order of adding the ingredients) brings a higher level of efficiency, by considering this indicators as environmental variables.

Clearly though, the name *environmental variable* here loses some of its meaning, since actions took by a DMU manager are definitely not external factors. Ergo, in this framework, we could think of them as simply conditioning factors, which allow us to estimate *conditional efficiency scores*. This conditional scores can be roughly understood as the level of efficiency a DMU attains when operating under certain conditions (be those conditions actual external variables like climate, or just a choice of production process, like in the example above). The objective here is to present the existent nonparametric framework for robust conditional frontier estimation and contribute to the literature on COVID-19 responses by trying to measure what countries were most efficient on combating the outburst of cases and at what extent demographic characteristics (or the policies adopted by civil society and governments) can help explain the estimated efficiency levels.

The innovation of the proposed application is the fact we do not use traditional inputs for healthcare efficiency analysis, such as hospital infrastructure and number of medical staff. Because we want to evaluate the response of countries in terms of controlling death rates as well as not imposing a big socioeconomic cost on population, the proposed input variables are proxies for the economic impact of the pandemic. In a way, we can see this choice of variables as representing our attempt to measure which countries controlled deaths with minimal economic impacts (or *using* less economic downturn). The rest of this text is organized as follows. Chapter 2 presents the whole framework for nonparametric efficiency analysis and explains the choice for the specific methodology used for the proposed application. Chapter 3 presents the database and some descriptive analysis, alongside an explanation for the choice of input, output and environmental variables as well as the results for the proposed models. Finally, Chapter 4 brings discussion about the results, propose an effective rank for efficiency on combating the pandemic and outlines a possible explanation for what the data showed.

2 EFFICIENCY ESTIMATION

Following the work of Koopmans (1951) and Debreu (1951), Farrell (1957) established the first attempt to estimate a production frontier from a set of data containing input usage and results from agricultural production across US states. His model is based on the estimation of technical efficiency in terms of a *efficiency score*, derived directly from the definition of production frontier - see Equations (2.4) and (2.6). The first part of this Chapter follows basically the presentation from Simar e Wilson (2013) and is important to introduce notation, necessary assumptions and the definitions to be used in the rest of the text. The second part ends up describing the *robust conditional order-m frontier estimator*, which will be used in the proposed application.

2.1 MATHEMATICAL AND ECONOMIC ASPECTS

A productive process can be described as any combination of production factors (land, labor, capital, etc.) to be utilized in obtaining a final product (or output). This definition is intrinsically linked to the interpretation of economic science as the study of scarce resource allocation to meet relatively unlimited needs and desires of individuals and agents. Also, it is closely connected to the common understanding that individuals organize in firms (here DMUs) to avoid transaction costs, as firstly explored by Coase (1937), and intend to produce goods and services in the most efficient way, limited to existent technology. One consequence of the development of economic science, however, is that it can easily be adapted to analyze a major part of problems in social sciences, nohow being confined to those commonly known as economic problems (LAZEAR, 2000). In this sense, the interpretation given in the present work for terms like *productive process*, or even *decision making units* is not restricted to the customary understanding relating it to firms or corporations, which means the methodology herein presented can be applied to a vast scale of diverse and subjective problems of social life. The definitions and assumptions presented will be rigorous in terms of notation, but mostly abstract with regards to the *economic* approach. The first notion to be established is that any production process (i.e. transformation of inputs into outputs) is limited to a feasible set, which describes the existent technology.

Let $x \in \mathbb{R}^p_+$ and $y \in \mathbb{R}^q_+$, denote the vectors with quantities of p inputs and q outputs, and let

$$\Psi = \{(x, y) \mid x \text{ can produce } y\}. \tag{2.1}$$

We have that Ψ is the set of all possible combinations of x and y such that the quantities y of outputs can be produced using quantities x of inputs. The following assumptions derive directly from classic microeconomic formulation (SHEPHARD, 2015).

Assumption 1: Ψ is closed.

Assumption 2: *No Free Lunch*: $(x = 0) \wedge (y > 0) \Rightarrow (x, y) \notin \Psi$.

Assumption 3: *Free Disposability*: $(x, y) \in \Psi \Rightarrow (x', y') \in \Psi$ if $(x' \geq x \wedge y' \leq y)$.

Particularly important to efficiency measurement in empirical problems is **Assumption 1**, because it guarantees the existence of a subset of Ψ called *efficient*, for if the production of a DMU is contained in this subset, it is said to produce the given y using the least feasible x, or, for another point of view, with a given x, it produces the maximum y feasible. This subset is clearly given by the upper boundary of Ψ and is formally defined as

$$\Psi^\delta = \{(x, y) \in \Psi \mid (\gamma^{-1}x, \gamma y) \notin \Psi \ \forall \gamma \in (1, \infty)\}. \tag{2.2}$$

As described before, the set Ψ^δ is called *production frontier* and sometimes simply *technology*, because it completely defines the existent technology capable of transforming inputs into outputs. Another way of visualizing Ψ^δ is by noting it is the intersection between Ψ and $cl(\Psi^c)$. It is said that DMUs whose production processes are in the interior of Ψ are technically inefficient, whilst those who operate on Ψ^δ are technically efficient. The feasible set Ψ can also be described by its level sets, as below, representing the input amount needed to produce a fixed y (i.e. the set of all x capable of producing y):

$$\mathcal{X}(y) = \{x \in \mathbb{R}^p_+ \mid (x, y) \in \Psi\}. \tag{2.3}$$

The boundary set, with this notation, is defined as

$$\mathcal{X}^\delta(y) = \{x \mid x \in \mathcal{X}(y), \theta x \notin \mathcal{X}(y), \forall \theta \in (0, 1)\}. \tag{2.4}$$

Note that this is the definition regarding what the literature commonly refer to as *input-oriented* approach. Another way to define Ψ and its boundary is using the so-called *output-oriented* approach. The feasible set, then, is given by

$$\mathcal{Y}(x) = \{y \in \mathbb{R}^q_+ \mid (x, y) \in \Psi\}, \tag{2.5}$$

and the boundary

$$\mathcal{Y}^\delta(x) = \{y \mid y \in \mathcal{Y}(x), \lambda y \notin \mathcal{Y}(x), \forall \lambda \in (1, \infty)\}. \tag{2.6}$$

Equations (2.4) and (2.6) represent, by and large, the object of study of empirical papers seeking to measure and identify technical efficiency of DMUs. More specifically, the input-oriented Debreu-Farrell *technical efficiency score* is given by

$$\begin{aligned} \theta((x,y) \mid \Psi) &= \inf(\theta \mid (\theta x, y) \in \Psi) \\ &= \inf(\theta \mid \theta x \in \mathcal{X}(y)). \end{aligned} \tag{2.7}$$

Given a vector of output y and a vector x of inputs (representing a DMU), the efficient consumption of inputs is given by

$$x^{\delta} = \theta((x,y) \mid \Psi) \cdot x, \tag{2.8}$$

which is the projection of (x,y) on the efficient frontier Ψ^{δ} along the ray x and orthogonal to y. Made simple, for a given $(x,y) \in \Psi$, $\theta((x,y) \mid \Psi)$ represents the proportional reduction on inputs x for which it is still feasible to produce y. Conversely, it can be seen as the necessary reduction on x such that the DMU becomes efficient. By construction, $\forall (x,y) \in \Psi, \theta((x,y) \mid \Psi) \in (0,1]$ and a DMU represented by (x,y) is efficient if and only if $\theta((x,y) \mid \Psi) = 1$. Similarly, the output-oriented Debreu-Farrell *technical efficiency score* is given by

$$\begin{aligned} \lambda((x,y) \mid \Psi) &= \sup(\lambda \mid (x, \lambda y) \in \Psi) \\ &= \sup(\lambda \mid \lambda y \in \mathcal{Y}(x)). \end{aligned} \tag{2.9}$$

Now, $\lambda((x,y)$ represents the proportional raise on production level (output y) such that a DMU becomes efficient. Again, by construction $\forall (x,y) \in \Psi, \lambda((x,y) \mid \Psi) \in [1,\infty)$ and (x,y) is efficient if and only if $\lambda((x,y) \mid \Psi) = 1$. Given a quantity x of inputs, and a quantity y of outputs, the efficient level of production is given by

$$y^{\delta} = \lambda((x,y) \mid \Psi) \cdot y, \tag{2.10}$$

which is the projection of (x,y) on the efficient frontier Ψ^{δ} along the ray y and orthogonal to x.

2.2 STATISTICAL ASPECTS

Apart from the economic aspects and mathematical definitions, it is still necessary to point some considerations about the statistical characteristics of frontier estimators. In a conventional empirical problem, the only available information is given by the sample

$$\mathcal{A}_n = \{(X_i, Y_i), i = 1, \ldots, n\}, \tag{2.11}$$

which contains the input consumption and production levels of a set of n DMUs. From this sample, the development of any efficient frontier estimator should be grounded on answers to a couple of basic questions about consistence, bias and asymptotic properties. This answers, however, depend on establishing a formal statistical model. Simar e Wilson (2013) state that a model is composed by two different parts: (a) a probabilistic model, including the assumptions made about the production set Ψ and the distribution of (X_i, Y_i); and (b) a description of how the sample is obtained, following the probabilistic model.

In general, the statistical model provides a theoretical description regarding the mechanism supporting the generation of a sample $\mathcal{A}_n$, being sometimes called *data generating process* (*DGP*). One of the prime characteristics distinguishing different statistical models used for technical frontier estimation is the set of assumptions about the distribution of deviations from the frontier. In that sense, there is a large spectrum of models, ranging from fully parametric approach, which assumes a given form for the distribution function of deviations from the efficient frontier, to completely nonparametric approach, which assumes no parametric forms for the distribution function. As expected, parametric models are generally used when researcher has good information about the distribution of deviations from technical efficiency. In those cases, the sample is handled in order to estimate parameters of a given distribution function, completely describing the DGP. Fully parametric models have the advantage of allowing deviations from the frontier to be caused by stochastic noise, and not only by technical inefficiency, this tends to make estimators less sensitive to outliers. Thus, within a fully parametric framework, it is also possible to observe DMUs ranging outside the feasible production set Ψ, mainly as a consequence of sampling failures.

On the other side, there are fully nonparametric estimators, with the advantage that they assume no restrictive functional forms for any features of the model. This allows for flexibility on estimation and enhance what the data (sample) itself has to say about the production process. In some applications (like the one proposed in this paper), this characteristic is rather important because assuming functional forms for the DGP is often harmful when the assumptions are not properly tested and verified (which is a hard task to complete in abstract settings). The cost for this flexibility, however, is that all pairs (X_i, Y_i) must be considered as technically feasible. In other words, the hypothesis here is that all observations of the sample are iid from a population of DMUs whose input and output vectors are distributed in the interior of Ψ, following some probability distribution characterized by a density $f(x, y)$ or cumulative distribution function (CDF) given by $F(x, y) = P(X \leq x, Y \leq y)$, with the property

$$P((X_i, Y_i) \in \Psi) = 1. \tag{2.12}$$

The literature for technical efficiency often attribute to parametric models the name *stochastic approach*, since they are associated with the inclusion of stochastic errors allowing for violation of condition (2.12). Nonparametric models, on the other hand, are said to be part of the *deterministic approach*, because they respect condition (2.12), commonly known as deterministic hypothesis (KUMBHAKAR et al., 2007). This terminology, however, has been loosing significance since there are important developments within the literature of nonparametric estimators allowing for violation of (2.12), as, for example, in the work of Kumbhakar et al. (2007). Besides, Simar e Wilson (2013) state that this names can be confusing, because in both approaches the frontier itself is *a priori* unknown (thus, not determined), which is exactly what creates the necessity for estimation. As argued before, mainly because of the characteristics of the empirical problem to be dealt with, the presentation will now be narrowed to nonparametric estimators. Also, following this last observations, the remaining of this text will focus on the fact that they are *nonparametric* at the expense of the term *deterministic*.

2.3 NONPARAMETRIC ESTIMATORS

Adding to the three assumptions presented in Section 2.1, Kneip, Simar e Wilson (2008) state other three assumptions necessary for the construction of nonparametric estimators. The authors main goal is to derive the asymptotic distribution of estimators based on classical linear programming methods. For now, those assumptions will be adapted to complete the construction of a DGP commonly used in the literature to introduce *Data Envelopment Analysis* (*DEA*) and *Free Disposal Hull* (*FDH*) estimators.

Assumption 4: The n observations from $\mathcal{A}_n$ are realizations of iid random variables over Ψ.

Assumption 5: (a) The (X_i, Y_i) have joint probability density function f and compact support $\mathcal{D} \subset \Psi$; (b) f is continuous in $\mathcal{D}$; and (c) $f(\theta(x,y)x, y) > 0, \forall (x,y) \in \mathcal{D}$.

Assumption 6: The functions $\theta(x,y)$ and $\lambda(x,y)$ are twice continuously differentiable for all $(x,y) \in \mathcal{D}$.

Note that **Assumption 5** (c) is particularly important, for it states that the probability of observing DMUs operating inside any open ball containing points of the frontier set Ψ^{δ} is strictly positive.

2.3.1 Full Frontier

The most general nonparametric estimator was formally introduced by Deprins e Simar (1984) as a direct application of **Assumptions 1-6** for a given sample. Its name: *Free Disposal Hull* derives specifically from **Assumption 3** which guarantees the technical possibility of wasting resources (inputs) on any productive process, i.e. any DMU is capable of producing less with the same amount of inputs. The estimator $\hat{\Psi}_{FDH}$ is given by the set of points $(x,y) \in \mathbb{R}_+^{p+q}$ such that a DMU operating at (x,y) would be *wasting resources* when compared to at least one of the DMUs in sample $\mathcal{A}_n$. In other words, (x, y) is said to be dominated by at least one DMU. The formal definition is given by

$$\begin{aligned}\hat{\Psi}_{FDH} &= \{(x,y) \in \mathbb{R}_+^{p+q} \mid x \geq X_i, y \leq Y_i, (X_i, Y_i) \in \mathcal{A}_n\} \\ &= \bigcup_{(X_i,Y_i)\in\mathcal{A}_n} \{(x,y) \in \mathbb{R}_+^{p+q} \mid x \geq X_i, y \leq Y_i\}.\end{aligned} \tag{2.13}$$

As pointed by Simar e Wilson (2008), the nonparametric estimator for the efficiency score of a point (x, y) is obtained through the substitution of Ψ by $\hat{\Psi}_{FDH}$ in Equations (2.7) or (2.9), depending on the approach (input-oriented or output-oriented). On practical problems, the estimates (for the input-oriented approach) can be computed following two simple steps: (a) identify the set D of points dominating (x, y):

$$D(x,y) = \{i \mid (X_i, Y_i) \in \mathcal{A}_n, X_i \leq x, Y_i \geq y\}, \tag{2.14}$$

and (b) calculate

$$\hat{\theta}_{FDH}(x,y) = \min_{i\in D(x,y)} \max_{j=1,\ldots,p} \left(\frac{X_i^j}{x^j}\right), \tag{2.15}$$

where x^j denotes the j^{th} element of vector x. A major drawback of the most general FDH efficiency score estimator $\hat{\theta}_{FDH}(x, y)$, as stated before, is the fact that it necessarily envelops the whole sample $\mathcal{A}_n$, being this the reason it is said to be part of a class of *full frontier estimators*. Practically speaking, this brings some difficulties for the empirical researcher, regarding a high sensitivity of the estimator to presence of outliers. One solution for this problem was handled by Cazals, Florens e Simar (2002), who presented a new class of nonparametric frontier estimators that became known as *partial frontier estimators*.

2.3.2 Partial Frontier

For the presentation of partial frontier estimators, it is convenient to describe the DGP in a pure probabilistic formulation, putting aside for a moment the definition of production set, given in (2.13). This notation was first introduced by Cazals, Florens e Simar (2002) and its basic idea is that the pdf given in **Assumption 5** can be completely characterized by the following probability function

$$H_{XY}(x,y) = P(X \leq x, Y \geq y). \tag{2.16}$$

It is important to notice this function is not a common distribution function, since it uses cumulative form for x and the survival form for y. Summarizing, three interesting properties arise from (2.16).

a) $H_{XY}(x, y)$ represents the probability of a unit operating at (x, y) being *dominated*, i.e. the probability that at least one other unit exists producing more or the same as y, and using x or less inputs.

b) $H_{XY}(x, y)$ is non-decreasing in x and non-increasing in y.

c) The support of H_{XY} is exactly the set Ψ, i.e. $H_{XY}(x, y) = 0, \forall (x, y) \notin \Psi$.

From the definition of conditional probability, one can write

$$\begin{aligned} H_{XY}(x, y) &= P(Y \geq y \mid X \leq x) \cdot P(X \leq x) \\ &= S_{Y|X}(y|x) \cdot F_X(x), \end{aligned} \tag{2.17}$$

where $S_{Y|X}(y|x) = P(Y \geq y \mid X \leq x)$ denotes the conditional survival function. Since the support is Ψ and it is free-disposal, the frontier Ψ^δ can be completely defined in terms of the conditional distributions given above (DARAIO; SIMAR, 2007). In the output-oriented approach (the one to be used in our application), we have the efficiency score being defined as

$$\begin{aligned} \lambda(x, y) &= \sup\{\lambda' \mid S_{Y|X}(\lambda' y \mid x) > 0\} \\ &= \sup\{\lambda' \mid H_{XY}(x, \lambda' y) > 0\}. \end{aligned} \tag{2.18}$$

It is interesting the interpretation within this framework: the score represents the proportional increase in output necessary for a DMU operating at (x, y) to reach zero probability of being dominated by some other. When dealing with this approach in practical problems of estimation, it is necessary to rewrite Equation (2.18), replacing the true $H_{XY}(x, y)$ by the empirical distribution $\hat{H}_{XY,n}(x, y)$, which is given by

$$\hat{H}_{XY,n}(x, y) = \frac{1}{n} \sum_{i=1}^{n} \mathbb{I}(X_i \leq x, Y_i \geq y). \tag{2.19}$$

For any point $(x, y) \in \mathbb{R}_+^{p+q}$, $\hat{H}_{XY,n}(x, y)$ is simply the proportion of points in the sample $\mathcal{A}_n$ dominating (x, y). Deriving out of $\hat{H}_{XY,n}(x, y)$, one can also define $\hat{F}_{X|Y,n}(x \mid y)$ and $\hat{S}_{X|Y,n}(y \mid x)$, as

$$\hat{F}_X(x) = \hat{H}_{XY,n}(x, 0), \tag{2.20}$$

and

$$\hat{S}_{Y|X,n}(y|x) = \frac{\hat{H}_{XY,n}(x, y)}{\hat{H}_{XY,n}(x, 0)}. \tag{2.21}$$

Since the definitions given in Equations (2.9) and (2.18) coincide in theory, the Debreu-Farrell technical efficiency score can be estimated by *plugging-in* the above empirical correspondents of $S_{Y|X}(\lambda' y \mid x)$ or $H_{XY}(x, \lambda' y)$ in (2.18). This approach is commonly known as *plug-in principle* and it is also used when developing robust estimators for the efficient frontier, as we intend to show next. The basic intuition for partial frontier estimators is that they don't envelop the whole sample, being, thus, less susceptible to changing significantly in the presence of outliers. As described by Simar e Wilson (2013), there are two widely used class of partial frontier estimators: (a) order-m estimators; and

(b) α-quantile estimators, where the frontier will be seen as a α-quantile frontier, similarly to what we see in traditional quantile-regression. In practice, α-quantile estimators are given by substituting the zero in (2.18), with $(1-\alpha)$, for some $\alpha \in [0,1]$ (generally small).

Although rather different in nature, both estimators have similar properties, often yielding very alike results. Also, both had their consistency proved (CAZALS; FLORENS; SIMAR, 2002) (DAOUIA; SIMAR, 2007) and asymptotic properties described (CAZALS; FLORENS; SIMAR, 2002) (DARAIO; SIMAR, 2006). Daouia e Gijbels (2011) provide a theoretical background for comparison of both estimators and conclude none of them can be claimed to be preferable in all contexts. Nevertheless, in their specific tests, they found order-m estimators to be more efficient in terms of statistical properties. Given these considerations, and as an attempt to simplify presentation, we limit our framework to the use of order-m estimators, to be introduced in detail. As seen in Equation (2.18), the efficiency score can be completely characterized by $S_{Y|X}(y \mid x)$, which evaluates the behavior of DMUs using less inputs than x. Total frontier, within this framework, can be understood as the representation of the maximum possible output level for *all* units using x. This concept, following Simar e Wilson (2008), is rather extreme and somehow detached from any real world problems. Order-m estimators are defined to be estimators of an *expected* frontier, in the sense that any point (x, y) will be evaluated against the expected maximum y for m randomly chosen units, all operating with a level of input less then or equal to x. The first thing to be noted is that as $m \to \infty$, the order-m estimator coincide with the full frontier estimator, as presented above. As instructively pointed by Witte e Kortelainen (2013), the idea is to draw a sectional frontier depending on a random set of m units which consume maximally x resources. Taking the expectation of this less extreme benchmark, the order-m efficiency score $\lambda_m(x, y)$ is obtained. Cazals, Florens e Simar (2002) derived a closed form for this expectation, depending only on $S_{Y|X}(y \mid x)$, which by plug-in principle can be practically computed using the empirical $\hat{S}_{Y|X,n}(y|x)$:

$$\hat{\lambda}_{m,n}(x, y) = \int_0^\infty [1 - (1 - \hat{S}_{Y|X,n}(uy|x))^m] du. \tag{2.22}$$

Fortunately, this estimator involves the calculation of a univariate integral, which is easily implemented with numerical methods, and even for large p and q, the estimates $\hat{\lambda}_{m,n}(x, y)$ are easily computed (CAZALS; FLORENS; SIMAR, 2002).

2.3.3 Conditional Frontier

A classical research problem within the literature of technical efficiency is the attempt to explain inefficiency using a set of variables to be called *environmental factors*. The empirical approach investigates relationships between the estimated efficiency scores and these external factors and the problem of explaining inefficiency is mainly important for two different reasons: (a) accounting for heterogeneity in the conditions surrounding the sample DMUs, an issue that arises when one needs to rank or compare DMUs by estimating individual efficiency scores; and (b) actually make an attempt to determine what factors help to explain inefficiency in terms of high input usage or low output production. For this last matter, the researcher usually has a set of observable and separable factors that can be thought to be impacting the level of efficiency, and wants to test global interactions between these factors and the estimated scores.

The simplest procedures are given by merely augmenting the basic statistical model introduced earlier by treating the r environmental factors Z as free disposal inputs or outputs that contribute to defining the attainable set $\Psi \subset \mathbb{R}_+^{p+q} \times \mathbb{R}^r$, or simply regressing the unconditional efficiency scores for different DMUs, using (for instance) the scores defined in (2.22), against Z (SIMAR; WILSON, 2013). Both approaches have one or a couple of serious problems, carefully exposed by Simar e Wilson (2013) and lucidly summarized by Witte e Kortelainen (2013).

a) The effect of environmental factors is required to be monotone in the production process;

b) The researcher has to know *a priori* whether the variables will be treated as outputs or inputs, which is equivalent (as will be seen later) to determining if they have positive or negative impact on the production process; and

c) For the model to be well defined in terms of statistical and economic meaning, it is necessary to assume the so called separability condition.

This last problem, regarding the separability condition deserves to be explored in some more detail: the separability condition states the feasible production set Ψ is not affected by Z. In a sense, it states that Z affects only the distribution of inefficiencies, but the attainable maximum production (the efficient frontier) is the same for every DMU, despite any heterogeneous Z conditions they are exposed to. An easy way to interpret this is by noting

$$\Psi = \cup_{z \in \mathcal{Z}} \Psi(z) \subset \mathbb{R}_+^{p+q}, \tag{2.23}$$

where $\mathcal{Z} \subseteq \mathbb{R}^r$ is the support of Z, and

$$\Psi(z) = \{(X, Y) \mid Z = z, X \text{ can produce } Y\}. \tag{2.24}$$

Hence, only if $\Psi(z) = \Psi, \forall z \in \mathcal{Z}$, we have the separability condition satisfied. When this is not satisfied, though, it is kind of straightforward to notice any unconditional estimator (i.e. one that does not consider different conditions for different DMUs) will be deprived of economic meaning, once the attainable frontier is not the same for different DMUs. Practically speaking, one cannot effectively compare production processes between firms, being, thus, unable to estimate trustworthy efficiency scores. That is basically why two-stage approaches that involve regressing unconditional efficiency scores to external factors can only have meaning if this condition is satisfied (SIMAR; WILSON, 2011). Ultimately, deciding whether the separability condition holds is an empirical issue, and it is often the case where it does not hold. Although some recent developments have been made in the direction of testing the condition, it is also possible to use *conditional efficiency* estimators, which does not depend on it. Using the probabilistic approach presented above, Daraio e Simar (2005) proposed a completely nonparametric estimator, which is an extension of the order-m frontier estimator. Within this framework, environmental variables are included in the model by conditioning $S_{Y|X}(y \mid x)$ also to $Z = z$, defining a new conditional survival function

$$S_{Y|X,Z}(y \mid x, z) = P(Y \geq y|X \leq x, Z = z), \tag{2.25}$$

with support given by the production technology when $Z = z$, or $\Psi(z)$. Using again the plug-in principle, we have that a *robust order-m conditional estimator* can be expressed by

$$\hat{\lambda}_{m,n}(x, y|z) = \int_0^\infty [1 - (1 - \hat{S}_{Y|X,Z,n}(uy|x, z))^m]du. \tag{2.26}$$

Here, the practical estimation problem for $\hat{\lambda}_{m,n}(x, y|z)$ is considerably more complicated, since it is necessary to apply smoothing techniques due to the equality constraint $Z = z$. Cazals, Florens e Simar (2002) show the empirical analog to $S_{Y|X,Z}(y \mid x, z)$ is given by

$$\hat{S}_{Y|X,Z;n}(x|y, z) = \frac{\sum_{i=1}^{n} \mathbb{I}(X_i \leq x, Y_i \geq y) K\left(\frac{z-Z_i}{h}\right)}{\sum_{i=1}^{n} \mathbb{I}(X_i \leq x) K\left(\frac{z-Z_i}{h}\right)}. \tag{2.27}$$

As widely spread in the literature, $K(\cdot)$ is a kernel function with compact support and h is a bandwidth vector selected using some choice method (CAZALS; FLORENS; SIMAR, 2002) (DARAIO; SIMAR, 2005) (WITTE; KORTELAINEN, 2013) (BĂDIN; DARAIO; SIMAR, 2012) (CORDERO; SALINAS-JIMÉNEZ; SALINAS-JIMÉNEZ, 2017) (MINVIEL; BOUHENI, 2021).

2.3.4 Explaining Inefficiency

After estimating an efficiency score which acknowledge for the existence of heterogeneous conditions among different DMUs, we apply a widespread methodology for evaluating how environmental factors are related to the efficiency of DMUs. Summarized by Daraio e Simar (2007), this methodology is given by a nonparametric regression of the estimated ratio between conditional and unconditional efficiency measures on the environmental factors Z. This ratio is given by

$$\hat{R}^z = \frac{\hat{\lambda}_{m,n}(x, y|z)}{\hat{\lambda}_{m,n}(x, y)}. \tag{2.28}$$

Daraio e Simar (2005) propose the use of a smooth nonparametric kernel regression to estimate the model as $\hat{R}^z = \hat{f}(Z_i) + \varepsilon_i$. This approach is widely applied in empirical papers and allow the researcher to detect positive, negative, neutral or even non-monotonic effects of environmental factors on the production process (WITTE; KORTELAINEN, 2013) (BĂDIN; DARAIO; SIMAR, 2014) (MINVIEL; BOUHENI, 2021) (CORDERO; SALINAS-JIMÉNEZ; SALINAS-JIMÉNEZ, 2017). To understand the intuition on why evaluating this regression can help explain how environmental variables are linked to efficiency, we first need to remind that, since we are working under an output-oriented approach, larger efficiency scores means lower efficiency. Adapting the presentation given by Simar e Wilson (2013), consider two different scenarios:

(a) If Z is favorable to the production process, it can be seen as acting like an extra input, enhancing production. In this case, for **larger** values of z, $\hat{\lambda}_{m,n}(x, y|z)$ is larger than $\hat{\lambda}_{m,n}(x, y)$ (the firm operating at (x, y) is seen as less efficient when conditioning to $Z = z$). This happens because when estimating $\hat{\lambda}_{m,n}(x, y)$, firms operating under large values of z will have an unfair advantage over those operating under small values of z. On the other hand, when estimating $\hat{\lambda}_{m,n}(x, y|z)$, this unfair advantage will be controlled for, and the firms operating under large values of z will not anymore be seen as that much efficient. As a consequence, the ratios (2.28) will increase as z increases.

(b) Whenever Z is unfavorable, it can be seen as acting like an unavoidable output that adds no value to production, or simply as a condition aggravating inefficiencies within the production process. In this case, for **smaller** values of z, $\hat{\lambda}_{m,n}(x, y|z)$ is greater than $\hat{\lambda}_{m,n}(x, y)$ (the firm operating at (x, y) is seen as less efficient when conditioning to $Z = z$). This happens because when estimating $\hat{\lambda}_{m,n}(x, y)$, firms operating under small values of z will have an unfair advantage over the ones operating under large values of z. When estimating $\hat{\lambda}_{m,n}(x, y|z)$, though, this unfair advantage will be controlled for, and the firms operating under small values of z will not anymore be seen as that much efficient. As a consequence, the ratios (2.28) will decrease as z increases.

For final clarification, we examine an example from our proposed application: since COVID-19 is commonly known to be more severe (deadly) for elderly people, we expect the environmental variable defined as the *median age* of population to be considered unfavorable to our *production process*. If that is indeed the case, countries operating under small values for *median age* have an advantage over other countries when estimating efficiency. But then, these countries will present smaller $\hat{\lambda}_{m,n}(x, y)$ when compared to the conditional estimator $\hat{\lambda}_{m,n}(x, y|z)$, because for this last one the advantage is being controlled for. Summing up, we expect to see a decreasing relationship between the ratios (2.28) and the proposed variable *median age*.

2.3.5 Ranking DMUs under heterogeneous conditions

In general terms, the methodology described in Section 2.3.4 serves for evaluating the relationship between efficiency and the conditioning variables Z on the estimated ratios defined in (2.28), which can, ultimately, lead us to infer whether a factor has positive or negative impact on the production process. In our COVID-19 responses efficiency analysis, however, we are also interested in ranking countries responses from the most efficient to the least. To do so, we still make use of conditional efficiency estimates, but applying the second-stage approach proposed by Bădin, Daraio e Simar (2012). At first, we need to establish the concept of *managerial efficiency* as a measure of the ability of managers to allocate resources effectively, removing the effects of environmental factors (MINVIEL; BOUHENI, 2021). Here, as we did for most of the text, we generalize the terms *managers* and *resource allocation*, for our abstract purpose. As proposed by Bădin, Daraio e Simar (2012), we analyze the average effect of $\lambda(x, y|z)$ as a function of Z, regressing the conditional scores (and not the ratios, as we did before) on the environmental variables. Naturally, we can get results for a model specified as

$$\lambda(X, Y|Z) = \mu(z) + \sigma(z)\varepsilon, \tag{2.29}$$

where $\mu(z) = \mathbb{E}(\lambda(X, Y|Z)|z)$ and $\sigma(z) = \mathbb{V}(\lambda(X, Y|Z)|z)$. Within this model and using the conditional efficiency scores estimated as first stage, we can construct an estimate for managerial efficiency as

$$\hat{\varepsilon} = \frac{\hat{\lambda}(x, y|z) - \hat{\mu}(z)}{\hat{\sigma}(z)}, \tag{2.30}$$

by applying some common nonparametric regression procedure to estimate $\hat{\mu}(z)$ and $\hat{\sigma}(z)$. This estimated $\hat{\varepsilon}$ is viewed as the unexplained part of conditional efficiency, being interpreted as a pure efficiency measure of the unit operating at (x, y). In simple words, $\hat{\varepsilon}$ are good measures to compare DMUs after removing the effect of heterogeneous conditions, so we use them to effectively rank countries responses.

3 COVID-19 RESPONSES OVER COUNTRIES

Since the outburst of COVID-19 cases around the world, debate over what causes different impacts in different countries and which set of policies is optimal to prevent deaths, cases and other indirect (mainly socioeconomic) setbacks spread just as fast as the statistics kept showing us how damaging this virus is. Following the necessity of understanding dynamics of the pandemic, a massive amount of papers related to the topic materialized (CAMPOS et al., 2021). For a more robust literature review on every topic related to the spread of the virus, heterogeneous impacts, medical, social and economic consequences of the pandemic and even policy recommendations, we point to the works of Brodeur et al. (2021), Susskind e Vines (2020), as well as Padhan e Prabheesh (2021). As for other papers using similar methodology to the one proposed here, applied to COVID-19 framework, we found mostly applications for assessing efficiency of healthcare systems in a strict sense of using the available healthcare structure as input variables (number of medical doctors, nurses, hospital beds, or health expenditure in general) and recovery or death rates as outputs (BREITENBACH; NGOBENI; AYE, 2021), (HAMZAH; YU; SEE, 2021).

This is highly expected, since healthcare system and hospital efficiency evaluation is one of the most popular areas of application for FDH and DEA methods (LIU et al., 2013), and the mentioned variables are a very common choice for inputs and outputs. The novelty of this work, thus, lies on the fact that we propose a more abstract and general application to evaluate the responses of countries. The main objective is to determine how well each country dealt with the pandemic crisis in terms of reducing the death toll and preventing well-being loss. Given the lack of options for well established pharmaceutical treatment or vaccines (at least during 2020), governments and civil societies had to put their best efforts to avoid death escalation using so-called non-pharmaceutical interventions. These policies, however, usually implied a huge downturn in economic activity, which notably came associated with some extent of well-being loss, or (for a major part of countries) a record increase in debt levels. Because we want to measure efficiency in these general terms, i.e., trying to answer what countries suffered less from COVID-19 deaths, managing not to impose so big of a cost on its population, expressed by deteriorating socioeconomic indicators, it makes sense to use an abstract approach, and a nonparametric estimator to find efficiency scores.

Besides the extent of non-pharmaceutical interventions, some demographic characteristics (specially the age of population) are said to be determinants of good responses. To control for heterogeneous settings, we propose the use of a robust nonparametric conditional frontier estimator to find efficiency scores for 105 evaluated countries, taking into account the fact that each of them had different demographic structures affecting their performance on fighting COVID-19. The inclusion of environmental factors, hence, serves two purposes: (a) effectively rank countries and compare responses, considering heterogeneous settings, using the methodology presented in Section 2.3.5; and (b) highlight the extent at which the restriction policies seemed to be effective on combating the burden of the virus. For this second objective, we aim to assess if the most practiced non-pharmaceutical interventions (restrictions and lockdown policies) are positively correlated with the general efficiency scores for different countries, using the methodology from Section 2.3.4.

3.1 DATABASE

Accounting for what was just exposed, our choice for output, input and environmental variables goes as follows. The output is the number of deaths caused by COVID-19 per million people. As for most of proposed variables, the numbers come from COVID-19 database maintained by the initiative Our World in Data (OWID) (RITCHIE et al., 2020). Since deaths per million is clearly an undesirable output, we adjust it by applying a linear decreasing transformation, as proposed by Seiford e Zhu (2002). Specifically, we apply the following transformation

$$y_i^\star = (\bar{y} - y_i) - \min_{i \leq n}\{\bar{y} - y_i\} + 100, \quad i \in \{1, \ldots, n\}, \tag{3.1}$$

where y_i represents total deaths per million for country i, $y_i^\star$ is the final output variable for country i, and $\bar{y} = n^{-1}\sum_{i=1}^{n} y_i$. Applying this transformation, we end up with a set where the country with largest number for deaths per million, say country i, has $y_i^\star = 100$, and $y_j^\star \geq y_k^\star \iff y_j \leq y_k$.

For input variables, we propose the use of IMF World Economic Outlook Database (WEO), comparing end of 2021 projections from October 2019 edition (the last one before the outburst of cases) and April 2021 edition (more than a year after the first cases were recorded in most countries). The idea is to quantify how projections for Gross Domestic Product (GDP) and General Government Gross Debt (GGD) were affected by the pandemic. Aware of possible fragility in some cases, using difference in projections for the end of 2021, before and after the pandemic, we attempt to capture the differential effect of COVID-19 crisis on economic activity and indebtedness. Although one can argue this choice is not the most precise for evaluating impact on economic outcome - we could point, for example, to a growing research agenda using difference-in-difference designs to estimate causal effects (LECHNER, 2011) - it is convenient for our analysis because we need indicators for the largest possible number of countries. As we see, even though it wouldn't serve for a detailed assessment on causality, changes in IMF projections can be considered good proxies of how each country economic indicators responded to the pandemic. This choice for input variables comes in line with the objective presented earlier: *estimate which countries suffered less from COVID-19 deaths, managing not to impose so big of a cost on its population, expressed by deteriorating socioeconomic indicators*. On this abstract application for efficient frontier estimation, input usage can be seen as the *managing not to impose so big of a cost on its population* part. In other words, our measurement tends to point as most efficient the countries which presented least deaths per million statistics, *using* less economic downturn or debt increase. Formally, making use of WEO subject codes, we define

$$x_1 = \left(\frac{PPPGDP^{2021}_{apr2021} - PPPGDP^{2021}_{oct2019}}{PPPGDP^{2021}_{oct2019}} \right) \cdot 100, \tag{3.2}$$

where $PPPGDP^{y}_{my}$ is the purchasing power parity GDP projected by IMF in the my edition of WEO for the end of year y, and

$$x_2 = \left(\frac{GGXWDG^{2021}_{apr2021} - GGXWDG^{2021}_{oct2019}}{NGDP^{2021}_{oct2019}} \right) \cdot 100, \tag{3.3}$$

where $NGDP$ represents GDP at current prices expressed in national currency units and $GGXWDG$ represents GGD expressed in national currency units. Indexation follows the same logic as for x_1, and details about the methodology are available on IMF website. Note that, on defining x_2, we are interested in the size of debt (GGD) compared to GDP, but we don't want to capture the effect of GDP projection changes, since x_1 is meant to do that. Hence, we maintain GDP projection fixed in $oct2019$ as denominator.

As defined, x_1 and x_2 express merely percentage changes in predictions for GDP and GGD. Since we are interested in economic downturn to be considered as input usage, we need both x variables to be such that higher values represent higher economic struggle (higher x_1 has to denote more negative GDP variation and higher x_2 has to denote increasing debt levels). To achieve this, and transforming such that there are no negative values, we define

$$x^{\star}_{1i} = (\bar{x_1} - x_{1i}) - \min_{i \leq n}\{\bar{x_1} - x_{1i}\} + 100, \quad i \in \{1, \dots, n\}, \tag{3.4}$$

and

$$x^{\star}_{2i} = (x_{2i} - \bar{x_2}) - \min_{i \leq n}\{x_{2i} - \bar{x_2}\} + 100, \quad i \in \{1, \dots, n\}, \tag{3.5}$$

which are the final inputs to be used in estimation.

For the environmental factors we propose the use of variables collected from OWID database: (z_1) median age of population (UN projection for 2020), expressed in years; (z_2) government response stringency index (composite measure based on nine indicators including school and workplace closures, as well as travel bans, scaled to a value from 0

Tabela 1 – Output, Inputs and Environmental Factors

	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
y	105	781.272	797.196	5.392	69.680	1,287.902	3,077.839
x_1	105	−7.271	11.946	−48.131	−13.796	−0.160	22.375
x_2	105	9.053	7.731	−2.359	4.014	12.581	46.043
z_1	105	32.843	9.238	16.400	25.400	41.800	48.200
z_2	105	60.814	10.598	34.043	53.464	68.519	84.870
z_3	105	−24.634	12.547	−51.211	−33.693	−15.756	10.024

Fonte: Elaboração própria.

to 100, where 100 = strictest response); and (z_3) retail and recreation number of visitors change, measured by Google in its COVID-19 Community Mobility Reports, expressed in percentage change when compared to baseline days (the median value for the 5-week period from January 3 to February 6, 2020)[1]. Variable z_1 is the one used for ranking countries (applying the methodology from Section 2.3.5), since it is seen as the main environmental factor impacting COVID-19 hits severity. Variables z_2 and z_3, which summarize the average extent of non-pharmaceutical interventions, are added to the specification in order to apply the methodology presented in Section 2.3.4 and evaluate the relationship between efficiency and these policies. Actually, for z_2 and z_3, OWID and Google provide daily information, so we use the average number for the sample of days, meaning that we are considering how strict (and effectively adopted by citizens) restriction policies were on average during the evaluated period (from 2020/03/01 to 2021/05/31). For more details about these three variables, we point to OWID COVID-19 website. To summarize the distribution of the proposed variables, a descriptive analysis is set in Table 1. The sample size of 105 countries was reached after gathering information on every proposed variable and filtering them by number of cases of COVID-19 per million greater than 500.

3.2 ESTIMATION DETAILS

The fact that our application is fairly abstract immediately suggests the use of a nonparametric estimator, as proposed in previous Sections. More specifically, we propose the use of a conditional order-m efficiency estimator, which does not require any parametric assumptions that could be over-restrictive for our framework. To choose the size of m, we follow a widely used methodology where m is set to be the value for which the number of super-efficient ($\hat{\lambda}_m < 1$) observations decreases smoothly (DARAIO; SIMAR, 2007). After this analysis, our choice for m is 20.

Now for the smoothing of Z, it is necessary that we use a kernel with compact support given by $\{z : |z| \leq 1\}$, since for unbounded support kernels (like the gaussian kernel) the estimates will be unable to detect any influence of environmental variables

[1] For the application, we actually use negative values of the true variable. It helps on interpretation of results (larger values represent stricter responses) and does not affect efficiency estimates.

Tabela 2 – Bandwidth selection (model: $z = (z_1, z_2, z_3)$)

	z_1	z_2	z_3
h_{cv}	2.1705	5.5530	3.6990

Fonte: Elaboração própria.

Tabela 3 – Descriptive analysis for countries with $\hat{\lambda}(x, y) > 1$

	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
$\hat{\lambda}(x, y)$	79	2.156	3.621	1.008	1.081	1.967	31.679
$y^{\star}$	79	2,205.185	799.203	100.000	1,606.500	2,930.800	3,142
x_1	79	130.833	11.613	110.402	123.085	136.270	170.506
x_2	79	112.330	6.841	100.889	108.340	116.262	138.138

Fonte: Elaboração própria.

(DARAIO; SIMAR, 2005). Here, we use Epanechnikov kernel, a very natural choice, and highlight the fact that Daraio e Simar (2007) acknowledge it is well known in applications like this that the results are robust to the choice of function (provided it belongs to the class with support presented before). This is not the case for the choice of bandwidth vector h. For that matter, we use Least Squares Cross-Validation (LSCV), a data-driven method that selects bandwidth minimizing the integrated squared error of the distribution estimate, providing an optimal bandwidth for all z in the support of $K(z)$. For details on the methodology, we point to Li e Racine (2007). Applying LSCV method for the model with all environmental variables, the chosen bandwidths are given in Table 2. For the model that solely controls for z_1, we have $h_{cv} = 1.5907$.

Finally, for applying the methodologies presented in Sections 2.3.4 and 2.3.5, we use local constant regression for estimation (PAGAN et al., 1999). Also, following the approach by Witte e Kortelainen (2013), when estimating the relation of non-pharmaceutical interventions and the ratios, we will not delimit to descriptive analysis, as we will also test for significance. To achieve this, after estimating the conditional mean for each point of the sample, we apply bootstrap tests (which in this context work as a nonparametric analogous of standard t-tests in ordinary least squares regression), as proposed by Racine (1997).

3.3 RESULTS

First, we present a description of the estimates for the unconditional efficiency scores. Tables 3, 4 and 5 show statistics for the 105 countries, grouped by $\hat{\lambda}(x, y)$ greater, smaller or equal to 1 (inefficient, super-efficient and efficient). For clarification, because we are using order-m partial frontier estimator, the existence of super-efficient countries is possible, meaning they perform better than the average $m = 20$ countries they were benchmarked with.

Tabela 4 – Descriptive analysis for countries with $\hat{\lambda}(x, y) = 1$

	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
$\hat{\lambda}(x, y)$	7	1.000	0.000	1.000	1.000	1.000	1.000
$y^\star$	7	2,484.057	611.938	1,663.943	1,931.199	2,933.544	3,118.848
x_1	7	116.293	14.548	100.000	104.910	125.592	139.749
x_2	7	104.501	5.499	100.000	101.206	105.424	115.885

Fonte: Elaboração própria.

Tabela 5 – Descriptive analysis for countries with $\hat{\lambda}(x, y) < 1$

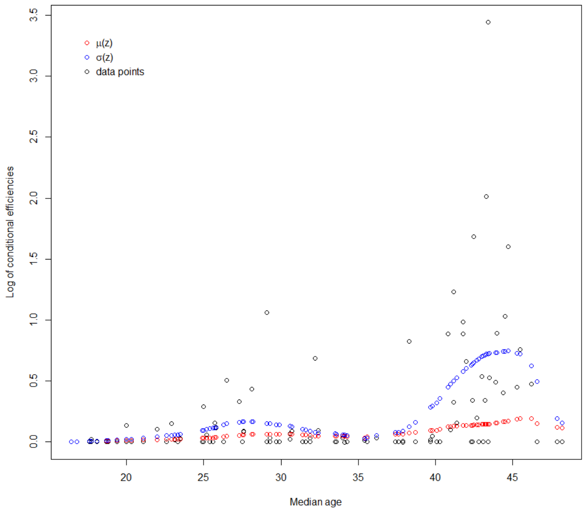
	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
$\hat{\lambda}(x, y)$	19	1.000	0.0003	0.999	1.000	1.000	1.000
$y^\star$	19	3,160.078	14.232	3,109.814	3,153.554	3,168.956	3,172.447
x_1	19	129.632	9.699	110.883	123.379	136.491	144.724
x_2	19	110.138	10.415	100.103	106.139	109.757	148.402

Fonte: Elaboração própria.

3.3.1 Ranking responses

As stated earlier, our specification for ranking countries responses considers the output variable $y^\star$, the two inputs $x_1^\star$ and $x_2^\star$, along with z_1, the median age of population as conditioning variable. Because we are already working with two input variables, and a limited sample size, we could not add as many environmental variables as we first intended, since efficiency estimates rapidly lost discriminatory power due to the *curse of dimensionality* (CHARLES; APARICIO; ZHU, 2019). Some other factors we considered for controlling heterogeneous settings included *population density*, *NCD prevalence* and *smoking habits*. For the last two, we did not find data for a large enough number of countries, which means adding them to the model would narrow the already restricted sample size, enhancing the discriminatory power problem. As for population density, we ran some tests with it as a second controlling variable and found mostly similar results, with less significance on the regression and a larger number of countries being rated as efficient ($\hat{\lambda}(x, y|z) = 1$). As discussed by Charles, Aparicio e Zhu (2019), this is the main sign of decreased discriminatory power of the estimation, which suggests a data set is not suited for frontier estimation with DEA (FDH) like methods. Trying to overcome this problem, we propose this restricted specification, controlling only for the median age of population. We base our decision on the fact that epidemiological evidence vastly points to the severity of infections being highly linked to the age of patients (CAMPOS et al., 2021). As for population density, since we are leaving it out as an environmental factor, we point to an interesting discussion about deaths and contamination rates in highly populated areas, see Hamidi, Sabouri e Ewing (2020) and Arbel et al. (2021). Before presenting the final ranking of countries responses, we exhibit some general results found when applying the methodology introduced in Section 2.3.5. After estimating the conditional efficiency scores $\hat{\lambda}(x, y|z_1)$ as first stage, we followed the second stage approach proposed by Bădin, Daraio e Simar (2012), and regressed its logarithm $\log \hat{\lambda}(x, y|z_1)$ on z_1. Within this framework,

Figura 1 – Second stage regression: effect of median age on conditional efficiencies



Fonte: Elaboração própria.

figures 1 and 2 are used to analyze the effect of z_1 on conditional efficiencies $\log \hat{\lambda}(x, y|z_1)$. At first, we note from Figure 1 that the regression line $\hat{\mu}(z_1)$ increases with larger values of z_1, indicating worst responses for higher levels of z_1 (median age). From the analysis of the histogram in Figure 2a, we note that managerial efficiencies are highly concentrated around zero, with mostly negative values (fewer inefficient countries than efficient). As for the scatter plot in Figure 2b, we find no specific pattern, indicating independence between them (correlation is calculated at 0.0051). This is important since it indicates our model has *cleaned* most of the effects of z_1 from our managerial efficiency measures, confirming them as good quantities to rank countries (MINVIEL; BOUHENI, 2021) (CORDERO; SALINAS-JIMÉNEZ; SALINAS-JIMÉNEZ, 2017). Finally, the ordering with results for all countries on the second-stage approach is presented in Tables 6 and 7.

3.3.2 Evaluating the relationship between efficiency and non-pharmaceutical interventions

Now for evaluating how non-pharmaceutical interventions are linked to the estimated scores, we apply the methodology presented in Section 2.3.4. For that purpose, besides controlling for z_1, we add variables z_2 and z_3 as proxies for the average non-pharmaceutical responses over countries. After estimating conditional efficiency scores within this specification, we calculate the ratios given in (2.28) for each country. Since we are using partial estimators (order-m), the goal here is to analyze the global correlation between Z and the efficient frontier as well as on the distribution of inefficiencies (BĂDIN; DARAIO; SIMAR, 2012). Controlling for heterogeneous characteristics (here, for median age of population), the reference sample for comparison is delimited to countries with similar environments

Tabela 6 – Ranking countries responses by $\hat{\varepsilon}$, the managerial efficiency estimates (most efficient response first)

Rank	Country	$y^\star$	$x_1^\star$	$x_2^\star$	z_1	$\log\hat{\lambda}(x,y)$	$\log\hat{\lambda}(x,y\|z_1)$	$\hat{\mu}(z_1)$	$\hat{\sigma}(z_1)$	$\hat{\varepsilon}$
1	KOR	3139.55	126.28	108.30	43.40	0.01	-0.00	0.72	0.15	-4.93
2	CHE	1928.80	113.30	102.37	43.10	0.00	0.00	0.71	0.14	-4.91
3	FIN	3005.30	121.26	108.71	42.80	0.05	-0.00	0.68	0.14	-4.83
4	SGP	3172.20	126.32	110.15	42.40	-0.00	-0.00	0.64	0.14	-4.65
5	DNK	2743.29	116.01	109.11	42.30	0.14	-0.00	0.63	0.14	-4.59
6	BLR	2876.12	128.15	100.11	40.30	0.00	0.00	0.36	0.10	-3.49
7	THA	3163.07	135.75	107.94	40.10	-0.00	-0.00	0.32	0.09	-3.45
8	EST	2234.03	120.98	118.30	42.70	0.35	0.20	0.67	0.14	-3.37
9	DEU	2120.34	122.53	116.99	46.60	0.40	-0.00	0.50	0.15	-3.27
10	KHM	3165.04	136.81	100.77	25.60	-0.00	0.00	0.11	0.03	-3.23
11	BOL	1933.60	123.03	102.30	25.40	0.00	0.00	0.11	0.03	-3.14
12	KGZ	2900.72	100.00	115.88	26.30	0.00	0.00	0.14	0.05	-3.11
13	NOR	3033.41	138.08	102.47	39.70	0.04	-0.00	0.29	0.09	-3.08
14	SWE	1746.94	125.33	107.78	41.00	0.59	0.10	0.48	0.13	-2.98
15	NPL	2924.34	111.36	114.09	25.00	0.08	0.00	0.10	0.03	-2.95
16	HND	2536.42	122.64	111.38	24.90	0.22	0.00	0.09	0.03	-2.90
17	LUX	1875.87	115.71	108.86	39.70	0.52	0.02	0.29	0.09	-2.89
18	BGD	3101.22	126.22	103.74	27.50	0.02	-0.00	0.16	0.06	-2.88
19	CAN	2501.46	125.74	132.30	41.40	0.23	0.15	0.52	0.13	-2.80
20	BRB	3014.29	150.76	121.18	39.80	0.05	0.05	0.29	0.09	-2.64
21	TJK	3168.40	124.15	102.36	23.30	-0.00	-0.00	0.06	0.02	-2.62
22	NLD	2133.36	124.62	111.22	43.20	0.40	0.34	0.71	0.14	-2.56
23	PNG	3159.73	127.79	107.91	22.60	-0.00	-0.00	0.05	0.02	-2.36
24	MEX	1443.85	129.81	103.57	29.30	0.78	0.00	0.15	0.06	-2.32
25	MMR	3118.71	160.34	103.83	29.10	0.02	0.00	0.15	0.06	-2.31
26	PAN	1701.28	121.82	107.16	29.70	0.60	0.00	0.14	0.06	-2.22
27	MYS	3091.45	142.77	106.90	29.90	0.02	-0.00	0.14	0.06	-2.19
28	MLT	2228.89	131.60	119.07	42.40	0.35	0.34	0.64	0.14	-2.19
29	GHA	3152.58	144.72	118.89	21.10	-0.00	-0.00	0.03	0.02	-2.17
30	AUT	2000.57	121.88	117.97	44.40	0.46	0.40	0.74	0.17	-2.05
31	IRL	2177.19	113.32	112.11	38.70	0.37	-0.00	0.16	0.08	-2.01
32	GEO	1981.35	110.40	117.82	38.70	0.45	-0.00	0.16	0.08	-2.01
33	RWA	3150.59	136.17	109.37	20.30	-0.00	-0.00	0.02	0.01	-1.83
34	IDN	2992.93	140.84	108.38	29.30	0.06	0.03	0.15	0.06	-1.78
35	JAM	2857.69	127.72	106.37	31.40	0.10	-0.00	0.10	0.06	-1.78
36	TUR	2614.32	116.25	110.17	31.60	0.19	-0.00	0.10	0.06	-1.76
37	SAU	2966.37	139.75	100.00	31.90	0.00	0.00	0.09	0.05	-1.72
38	GAB	3109.55	144.04	118.00	23.10	0.02	0.02	0.06	0.02	-1.71
39	KEN	3118.85	104.80	106.20	20.00	0.00	0.00	0.02	0.01	-1.70
40	KAZ	2993.94	135.04	109.31	30.60	0.06	0.02	0.13	0.06	-1.65
41	IND	2937.34	147.08	110.85	28.20	0.08	0.07	0.17	0.06	-1.64
42	ITA	1091.76	121.04	116.87	47.90	1.06	0.00	0.19	0.12	-1.62
43	LVA	1918.16	127.08	109.58	43.90	0.50	0.49	0.73	0.15	-1.57
44	GRC	2017.43	129.63	113.86	45.30	0.45	0.45	0.73	0.19	-1.50
45	PAK	3083.45	138.22	105.88	23.50	0.03	0.03	0.06	0.02	-1.47
46	TGO	3162.74	111.31	110.77	19.40	-0.00	-0.00	0.01	0.01	-1.46
47	NZL	3172.45	122.23	118.33	37.90	-0.00	-0.00	0.09	0.07	-1.43
48	BRA	1000.61	133.81	104.85	33.50	1.15	0.00	0.07	0.05	-1.43
49	CRI	2384.57	115.82	106.34	33.60	0.27	0.00	0.07	0.05	-1.39
50	SRB	2168.96	124.87	111.28	41.20	0.38	0.32	0.50	0.13	-1.37
51	DOM	2843.40	132.06	109.29	27.60	0.11	0.09	0.16	0.06	-1.36
52	LTU	1610.78	120.50	122.13	43.50	0.67	0.53	0.73	0.15	-1.34
53	JPN	3075.01	129.86	115.84	48.20	0.03	-0.00	0.15	0.12	-1.32

Fonte: Elaboração própria.

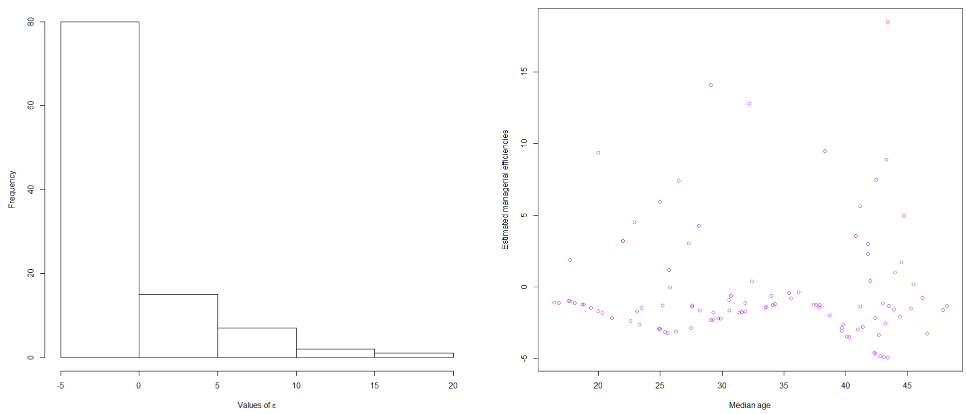
Tabela 7 – Ranking countries responses by $\hat{\varepsilon}$, the managerial efficiency estimates (most efficient response first)

Rank	Country	$y^\star$	$x_1^\star$	$x_2^\star$	z_1	$\log \hat{\lambda}(x,y)$	$\log \hat{\lambda}(x,y\|z_1)$	$\hat{\mu}(z_1)$	$\hat{\sigma}(z_1)$	$\hat{\varepsilon}$
54	PHL	2986.51	139.42	109.80	25.20	0.06	0.06	0.10	0.03	-1.29
55	SLV	2831.10	126.83	112.92	27.60	0.11	0.09	0.16	0.06	-1.28
56	LKA	3110.54	132.73	114.94	34.10	0.02	-0.00	0.06	0.05	-1.27
57	MDA	1663.94	105.02	104.65	37.60	0.00	0.00	0.08	0.06	-1.26
58	AUS	3142.15	127.39	132.29	37.90	0.01	0.01	0.09	0.07	-1.26
59	MUS	3163.68	143.90	107.48	37.40	-0.00	-0.00	0.08	0.06	-1.24
60	BEN	3169.51	129.95	109.33	18.80	-0.00	-0.00	0.01	0.01	-1.23
61	CMR	3129.81	133.20	102.39	18.80	0.01	0.00	0.01	0.01	-1.23
62	CIV	3166.28	110.88	108.04	18.70	-0.00	-0.00	0.01	0.01	-1.22
63	SEN	3109.81	141.09	100.10	18.70	-0.00	0.00	0.01	0.01	-1.22
64	BHS	2592.97	123.56	109.80	34.30	0.20	-0.00	0.05	0.05	-1.18
65	ROU	1602.18	117.44	114.66	43.00	0.68	0.54	0.70	0.14	-1.14
66	AGO	3154.53	122.61	148.40	16.80	-0.00	0.00	0.00	0.00	-1.13
67	NGA	3167.66	138.52	104.63	18.10	-0.00	-0.00	0.01	0.01	-1.13
68	QAT	2984.85	154.64	108.73	31.90	0.06	0.03	0.09	0.05	-1.12
69	MLI	3152.31	129.96	105.63	16.40	-0.00	0.00	0.00	0.00	-1.08
70	UGA	3169.93	128.91	107.85	16.40	-0.00	0.00	0.00	0.00	-1.08
71	MOZ	3151.09	131.27	106.65	17.70	-0.00	-0.00	0.01	0.01	-1.03
72	BFA	3169.90	120.65	108.01	17.60	-0.00	-0.00	0.01	0.01	-1.00
73	ISR	2437.04	120.47	113.41	30.60	0.26	0.07	0.13	0.06	-0.90
74	URY	1946.88	130.98	105.36	35.60	0.48	-0.00	0.03	0.04	-0.79
75	PRT	1508.18	122.50	117.20	46.20	0.74	0.48	0.63	0.19	-0.77
76	ARE	3007.98	138.87	116.95	34.00	0.05	0.03	0.06	0.05	-0.62
77	OMN	2718.63	161.84	101.89	30.70	0.15	0.09	0.13	0.06	-0.62
78	CHL	1645.11	134.07	105.14	35.40	0.65	0.01	0.03	0.03	-0.41
79	TTO	2824.14	147.81	104.43	36.20	0.11	0.03	0.05	0.06	-0.38
80	BWA	2816.81	136.90	113.48	25.80	0.12	0.12	0.12	0.04	-0.04
81	ESP	1467.79	128.74	115.22	45.50	0.77	0.76	0.72	0.19	0.18
82	BHR	2601.91	129.63	110.17	32.40	0.20	0.09	0.08	0.05	0.39
83	FRA	1554.35	123.53	112.66	42.00	0.71	0.66	0.60	0.14	0.42
84	HRV	1222.79	125.50	114.34	44.00	0.95	0.89	0.73	0.16	1.02
85	CPV	2703.01	144.71	100.89	25.70	0.16	0.16	0.12	0.04	1.19
86	SVN	1073.39	124.34	113.58	44.50	1.08	1.03	0.74	0.17	1.70
87	ZMB	3108.16	143.39	138.14	17.70	0.02	0.02	0.01	0.01	1.89
88	POL	1229.32	126.20	111.59	41.80	0.95	0.89	0.58	0.13	2.31
89	BEL	1024.62	119.48	117.28	41.80	1.13	0.98	0.58	0.13	3.01
90	ZAF	2225.09	135.15	108.41	27.30	0.35	0.33	0.16	0.06	3.06
91	NAM	2851.18	139.57	108.59	22.00	0.11	0.11	0.04	0.02	3.20
92	GBR	1291.66	127.81	120.77	40.80	0.90	0.89	0.45	0.12	3.58
93	ECU	2011.83	129.03	113.92	28.10	0.45	0.43	0.17	0.06	4.26
94	GTM	2722.09	129.86	104.99	22.90	0.15	0.15	0.05	0.02	4.52
95	BGR	630.51	129.98	110.03	44.70	1.61	1.60	0.75	0.17	4.96
96	SVK	917.07	135.64	113.99	41.20	1.24	1.23	0.50	0.13	5.63
97	BLZ	2362.99	153.40	121.11	25.00	0.29	0.29	0.10	0.03	5.96
98	PRY	1889.94	132.14	112.22	26.50	0.52	0.51	0.15	0.05	7.41
99	BIH	357.50	127.08	105.51	42.50	2.18	1.69	0.65	0.14	7.48
100	CZE	366.37	120.61	114.45	43.30	2.16	2.01	0.72	0.15	8.87
101	IRQ	2770.73	170.51	116.68	20.00	0.13	0.13	0.02	0.01	9.36
102	USA	1381.52	124.55	122.31	38.30	0.83	0.82	0.12	0.07	9.48
103	COL	1433.17	133.35	113.61	32.20	0.79	0.69	0.08	0.05	12.81
104	PER	1074.77	140.50	108.86	29.10	1.08	1.06	0.15	0.06	14.11
105	HUN	100.00	129.14	119.51	43.40	3.46	3.44	0.72	0.15	18.50

Fonte: Elaboração própria.

Figura 2 – Analysis of managerial efficiencies

(a) Histogram (b) Scatter plot against median age



Fonte: Elaboração própria.

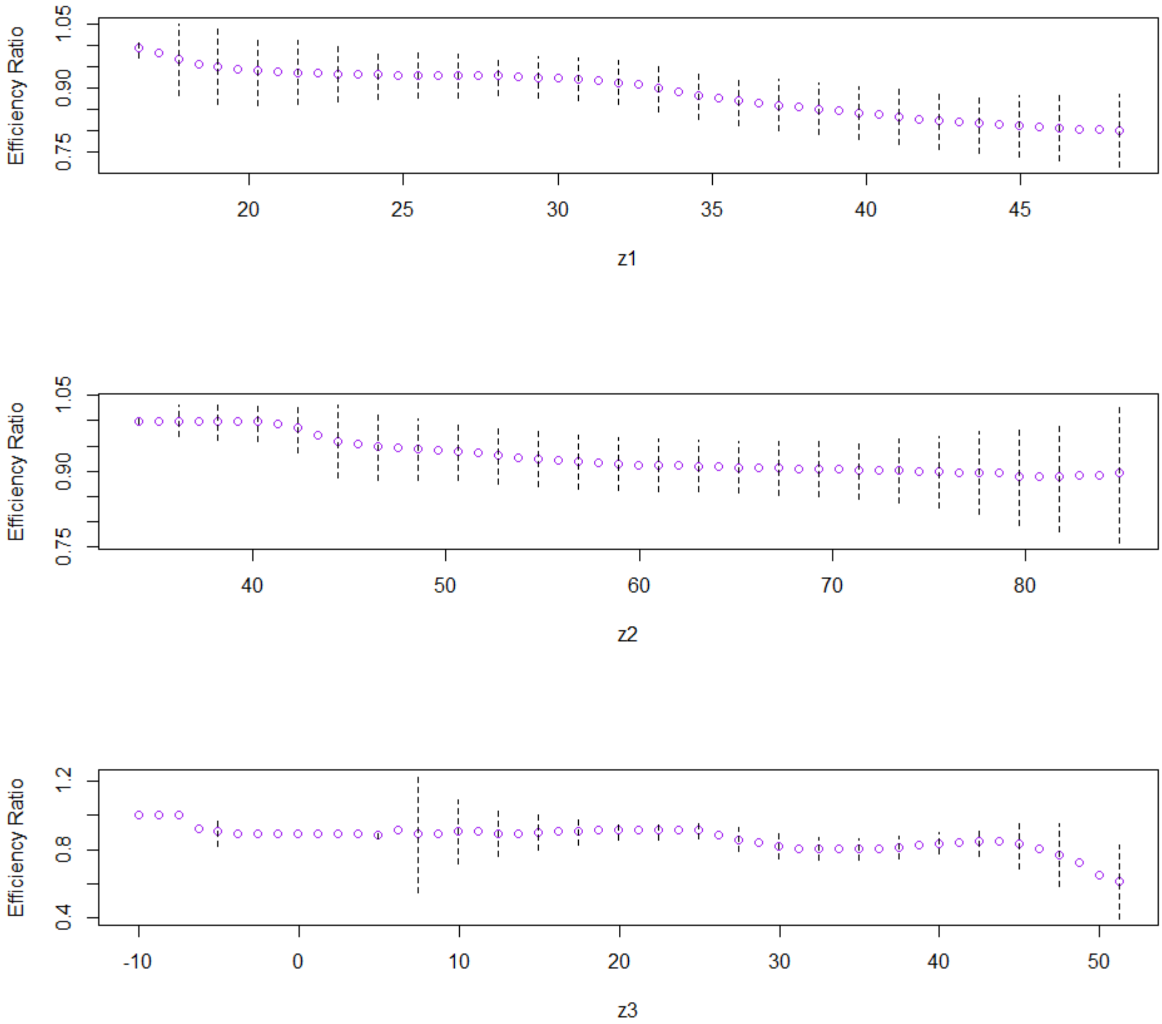
Tabela 8 – Descriptive analysis of ratios $\hat{R}^z$ (model: $z = (z_1, z_2, z_3)$)

	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
$\hat{R}^z$	105	0.851	0.185	0.113	0.688	1.000	1.001

Fonte: Elaboração própria.

and this can ultimately allow us to understand how z_2 and z_3 impact efficiency. Table 8 depicts statistics for the calculated ratios. As stated before, to assess the relationship between environmental factors and efficiency scores, we nonparametrically regress the ratios $\hat{R}^z$ on Z and evaluate partial regression plots of each factor against the given ratio. Practically, this analysis can help us understand the extent at which Z is affecting COVID-19 responses over countries, allowing us to detect (confirm) how the median age of population (variable z_1) intensified the stroke of the virus and ultimately serving as an indicative of effectiveness for non-pharmaceutical interventions (characterized by variables z_2 and z_3). For the nonparametric local constant regression, we use LSCV method to select bandwidths and Epanechnikov kernel function, just as we did for the estimation of $\hat{\lambda}(x, y|z)$. The partial regression plots are set in Figure 3. While evaluating for z_i we set all other $z_j, j \neq i$ on their median value, as usual in the literature for multivariate regression. The regression significance test results (p-values) are given in Table 9, alongside the description for the average effect, as observed from partial plots. For statistical inference, we used 1.000 bootstrap replications.

Figura 3 – Partial plots (z_1 = median age, z_2 = stringency index, z_3 = retail and recreation visitors change)



Fonte: Elaboração própria.

Tabela 9 – Nonparametric significance test (model: $z = (z_1, z_2, z_3)$)

	p-value	Average effect (from partial plot)
z_1 (Median Age)	$< 2e-16$	Negative
z_2 (Stringency Index)	0.046	Non-monotonic/Negative
z_3 (Retail and Recreation Visitors Change)	0.075	Non-monotonic/Negative

Fonte: Elaboração própria.

4 DISCUSSION

As stated before, the conditional structure of the proposed estimator implies that efficiency scores will be taking into account heterogeneous environments, so that countries responses to COVID-19 crisis will be compared to a narrower set of responses, from countries with similar characteristics. In general, though, even when applying the methodology described in Section 2.3.4, whether we can infer causal impacts of the environmental variables on efficiency scores is a matter of determining if all (observed and unobserved) variables causing heterogeneity between countries are being controlled for. Since it is inconceivable to assert we are controlling for every cause of heterogeneity just by adding *median age* as environmental factor, we leave the study of causal effects for future research. That being said, we focus solely on what the data showed within the framework we proposed, and try to outline some possible reasoning for the results. First, we analyze the results for each of the proposed environmental factors:

(z_1) As expected, it appears that countries with higher median age of population struggled more (on average) with the spread of the virus. This comes in line with the epidemiological evidences asserting the diseases severity and lethality is indeed age-dependent (CAMPOS et al., 2021), affecting disproportionately the elders. In that sense, countries with higher median age suffered with higher death rates, being unable to avoid the escalation of lethal infections even after imposing hard restriction policies that significantly impacted economic indicators.

(z_2) The stringency index summarizes information about government responses regarding school and workplace closures, cancellation of public events, restrictions on gatherings, closed public transportation services, restrictions on internal movement and international travel, test and contact tracing police and requirement of facial coverings. At first, we expected higher average indexes would help explain higher levels of estimated efficiency scores, but we did not observe that. In fact, analyzing the partial regression plot, it shows a negative (or neutral for larger values of z_2[1]) relationship between stricter responses and the calculated ratios. This is an indicative that duration of government restriction policies were not that effective on preventing deaths, whilst it is hard to assert they were not linked to a deterioration of economic indicators. In fact, a number of issues arise when evaluating results for this variable and we present some thoughts on why this is what the data showed. Countries with higher death rates were (almost by definition) the ones struggling more with the

[1] We even note an increasing behavior for $z_2 > 80$, but this is negligible, specially because confidence intervals become very large at this point.

pandemic, and this struggling demanded actions from governments, which responded by maintaining restriction policies for longer periods. This is the reasoning for a possible causality design operating contrary to what we were expecting. In other words, it is likely that higher death rates caused stricter responses, that caused economic downturn and debt increase, which explains the higher inefficiency levels. Summarizing, it is not as if our results are necessarily showing restriction policies caused inefficiency, but it is also not absurd to assert maintaining restrictions for longer periods, at least the way they were implemented on most countries, was relatively not effective on stopping death escalation, considering the cost in terms of activity reduction and debt increase[2]. In this sense, we could look at the anecdotal evidence from New Zealand. Commonly referred to as a role model for combating the pandemic, the island is, in our sample, indeed, the country with least deaths per million, and presents only the 103^{th} average strictest response during the evaluated period (out of 105 countries). Besides the convenience of not having terrestrial frontiers in a pandemic situation, another possible explanation for this is the success of assertive restriction policies, in the right time (contrary to long and unsteady political responses), that allowed the government to ease constrains as soon as the number of cases was completely under control. In essence, within this point of view, we could assert the more effective a country was on combating the virus with decisive non-pharmaceutical interventions, the sooner it was able to ease restrictions, presenting, thus, a small average stringency index for the whole evaluated period. In other words, our measurement does not take into account quality or timing of implemented policies, and this is probably a main issue on determining how successful they were. Any causal study on this matter must take these into account, most likely by expanding the study to a time series analysis of detailed implemented restrictions and direct observed results.

(z_3) The second variable related to non-pharmaceutical interventions is the average Google Mobility Trends retail and recreation visitors change during the evaluated period. The reasoning for including z_3 is the fact that (as always in public economics) population responses may significantly differ from the intention of policy-makers. Since the numbers represent percentage change in visitors, it is important to note that decreasing values can be seen as proxies for stricter responses, which is why we adjusted our sample (to ease interpretation) by taking negative values of the original variable. In Figure 3 the partial plot is already adjusted, so $z_3 = -10$, for example, means the number of visitors in retail and recreation increased by 10%. Again, regardless the non-monotonic shape of the curve, the plot indicates a negative

[2] It is really important to note we are not proposing a trade-off relation between economic burden and lives lost by COVID-19. Here, we simply infer that (when comparing responses) some countries were able to control the disease with less GDP reduction or GDD increase than others, and try to identify whether and how non-pharmaceutical interventions interacted with this process.

relationship between z_3 and efficiency estimates (specially for higher values of z_3). Consequently, we simply extend the previous analysis in terms of possible explanation for the results.

Besides explaining inefficiency with environmental variables, another objective of the proposed application is the effective ranking of countries responses, as we did in Tables 6 and 7. It is worth highlighting some of the observed results:

a) Hungary presented, in all measures, the least efficient response, mainly due to very high debt increase within IMF predictions for the end of 2021, and to the fact that it presented the higher deaths per million statistic on the end of the evaluated period (approximately 3.077 deaths per million in 2021/05/31);

b) Peru and Colombia would be ranked as 9^{th} and 15^{th} least efficient countries when evaluating unconditional scores and as 6^{th} and 14^{th} if we were to use the conditional scores alone. Besides, when we look at the unconditional scores, not only they were better ranked, but they also presented scores much closer to the average of countries (the average is 1.87, Peru has 2.95 and Colombia 2.21). When applying the methodology described in Section 2.3.5, though, ranking by $\hat{\varepsilon}$ (the *managerial efficiency estimate*), they stand out as 2^{nd} and 3^{rd} most inefficient, with $\hat{\varepsilon}$ of 14.11 and 12.81, contrasting with the average of -0.20. This is directly linked to the fact that both countries were in an advantageous positions in terms of population age, and the applied methodology clears the effect of environmental advantages, displaying the true level of inefficiency they attained.

c) When evaluating efficient responses, we observe some regions of the world standing out. As an example, the countries of our sample that belong to Eastern Asia and South-eastern Asia, presented mean values of -3.12 and -2.70 for $\hat{\varepsilon}$, meaning they were much more efficient than the estimated average, conditional on their values of z_1. On Table 10, we can compare the distributions of $\hat{\varepsilon}$ by world region and contrast, for example, what we just said about Eastern Asia and South-eastern Asia with the results for Eastern Europe (mean $\hat{\varepsilon}$ of 4.30), the most inefficient region on combating COVID-19.

4.1 CONSIDERATIONS ABOUT THE ESTIMATION

Having estimated the efficiency scores for each country, we bring some considerations about the conclusiveness of our findings. First, we point to the fact that in our exploratory analysis, we find a great deal of countries operating under low input usage and high output levels. Intuitively, this is a problem for estimating efficiency frontiers since higher input

Tabela 10 – Managerial efficiency estimates ($\hat{\varepsilon}$) distributions, by world regions

Region	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
Eastern Asia	2	-3.12	2.55	-4.93	-4.02	-2.22	-1.32
South-eastern Asia	7	-2.70	1.15	-4.65	-3.34	-1.99	-1.29
Central Asia	3	-2.46	0.74	-3.11	-2.87	-2.14	-1.65
Melanesia	1	-2.36		-2.36	-2.36	-2.36	-2.36
Northern Europe	9	-2.24	2.50	-4.83	-3.37	-1.57	3.58
Southern Asia	5	-2.04	0.81	-2.95	-2.88	-1.47	-1.27
Western Europe	7	-1.75	2.64	-4.91	-3.08	-0.82	3.01
Caribbean	5	-1.47	0.83	-2.64	-1.78	-1.18	-0.38
Middle Africa	3	-1.36	0.31	-1.71	-1.47	-1.18	-1.13
Australia and New Zealand	2	-1.34	0.12	-1.43	-1.39	-1.30	-1.26
Western Africa	9	-1.04	0.91	-2.17	-1.23	-1.08	1.19
Eastern Africa	6	-0.83	1.37	-1.83	-1.58	-1.04	1.89
Central America	7	0.05	3.61	-2.90	-2.27	1.62	5.96
Western Asia	9	0.11	3.55	-2.01	-1.72	-0.62	9.36
Southern Europe	9	0.33	2.98	-2.19	-1.50	1.02	7.48
Southern Africa	3	2.07	1.83	-0.04	1.51	3.13	3.20
Northern America	2	3.34	8.68	-2.80	0.27	6.41	9.48
South America	8	4.10	6.70	-3.14	-0.95	8.76	14.11
Eastern Europe	8	4.30	7.07	-3.49	-1.17	6.44	18.50

Fonte: Elaboração própria.

levels should translate into more output, and any significant deviations from this pattern probably means there are other factors impacting production. Of course, by controlling for the median age, we intend to tackle this, but there are certainly other circumstances influencing output and input levels on this abstract approach we propose. We could point to the very fact that our measurement of economic impacts may produce some gaps in estimation. For example, it is perfectly possible that a number of countries presented (in April 2021) substantially recovery signs, which led projections for the end of 2021 to be more optimistic, specially regarding the advance of vaccination policies. Unlike non-pharmaceutical interventions, which were implemented differently across countries, with heterogeneous results, vaccines have tested efficacy and are important tools to prevent deaths, working basically the same way around the globe. Including information about vaccination progress in countries, hence, could serve for isolating this factor on comparing countries and estimate better conditional scores for ranking them. Due to lack of available data for a sufficient amount of countries in our sample, we leave that information out of the proposed models.

In what concerns sample size, empirical applications of nonparametric frontier estimators usually demand a large number of observations, which, unfortunately, was not possible here. Since we propose analyzing countries efficiencies, we depart from an already reduced sample (total number of countries integrating UN is 193). At our best effort, we managed to gather the information needed for 114 countries. After filtering countries with less than 500 cases per million and removing outliers, we ended up with our sample of 105 countries. As a suggestion for future research, we leave the idea of estimating efficiency scores in similar frameworks for cities or states within countries. This should increase significantly the sample size, possibly leading to more trustworthy results and allowing for the inclusion of other controlling variables.

5 CONCLUDING REMARKS

In light of the different impacts of COVID-19 around the world, we propose an application of conditional nonparametric order-m partial frontier estimation to rank countries efficiencies on combating the pandemic and to evaluate the interaction between non-pharmaceutical interventions and the efficiency of responses. We apply two distinct second-stage approaches:

a) At first, we nonparametrically regress conditional efficiency scores on environmental variable *median age*, in order to rank countries by the residuals of this regression, widely understood as *managerial efficiency measures*;

b) To estimate how non-pharmaceutical interventions interacted with the efficiency scores, we regress the ratios of conditional and unconditional efficiency scores on the conditioning factors, and test significance. We find that variables *median age*, *stringency index* and *retail and recreation visitors change* were significant for having a negative relationship with the efficiency estimates. We point that variables *stringency index* (z_2) and *retail and recreation visitors change* (z_3) can be mostly considered to be measures of duration of restriction policies and social distancing, as they are not really informative on the quality and timing of interventions. Hence, our findings simply assert the maintenance of restrictions for longer periods was not effective (on average). Any causal study on this matter should consider the timing and differentiate types of actions took in different countries to reach more conclusive results.

Throughout the text, we approach some topics for future research. First of all, there is the causality issue. Witte e Kortelainen (2013) point that efficiency analysis papers are usually only concerned with correlation between environmental variables and efficiency outcomes, neglecting endogeneity and causality issues (this is indeed the case here). Another related point is the inclusion of other variables to be used as conditioning factors. This problem is mainly operational: due to a small sample size and the lack of established databases containing accurate information for a big number of countries, we ended up omitting other possibly important country-level characteristics impacting efficiency on combating COVID-19, like *non-communicable diseases (NCD) prevalence*, *smoking habits* and *number of vaccines delivered*. For future work, we propose the use of a similar framework for evaluating the interaction of efficiency estimates and environmental variables, only enlarging the number of observations by considering not countries, but cities or states information.

Finally, as far as we know, our work represents the first attempt to apply efficiency frontier estimation methods to COVID-19 framework in a multi-level approach, considering not only healthcare issues, but also the extent of production plunge and public finance deterioration that followed the outburst of the pandemic crisis.

REFERENCES

ARBEL, Y. et al. Do covid19 infection rates change over time and space? population density and socio-economic measures as regressors. *Cities*, p. 103400, 2021. ISSN 0264-2751. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0264275121002997>.

BĂDIN, L.; DARAIO, C.; SIMAR, L. How to measure the impact of environmental factors in a nonparametric production model. *European Journal of Operational Research*, v. 223, n. 3, p. 818–833, 2012. ISSN 0377-2217. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0377221712004833>.

BĂDIN, L.; DARAIO, C.; SIMAR, L. Explaining inefficiency in nonparametric production models: the state of the art. *Annals of Operations Research*, Springer, v. 214, n. 1, p. 5–30, 2014. Disponível em: <https://doi.org/10.1007/s10479-012-1173-7>.

BREITENBACH, M. C.; NGOBENI, V.; AYE, G. C. Global healthcare resource efficiency in the management of covid-19 death and infection prevalence rates. *Frontiers in Public Health*, Frontiers Media SA, v. 9, 2021.

BRODEUR, A. et al. A literature review of the economics of covid-19. *Journal of Economic Surveys*, v. 35, n. 4, p. 1007–1044, 2021. Disponível em: <https://onlinelibrary.wiley.com/doi/abs/10.1111/joes.12423>.

CAMPOS, E. L. et al. Multi-generational sir modeling: Determination of parameters, epidemiological forecasting and age-dependent vaccination policies. *Infectious Disease Modelling*, v. 6, p. 751–765, 2021. ISSN 2468-0427. Disponível em: <https://www.sciencedirect.com/science/article/pii/S2468042721000373>.

CAZALS, C.; FLORENS, J.-P.; SIMAR, L. Nonparametric frontier estimation: a robust approach. *Journal of Econometrics*, v. 106, n. 1, p. 1–25, 2002. ISSN 0304-4076. Disponível em: <https://www.sciencedirect.com/science/article/pii/S030440760100080X>.

CHARLES, V.; APARICIO, J.; ZHU, J. The curse of dimensionality of decision-making units: A simple approach to increase the discriminatory power of data envelopment analysis. *European Journal of Operational Research*, v. 279, n. 3, p. 929–940, 2019. ISSN 0377-2217. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0377221719305168>.

COASE, R. H. The nature of the firm. *Economica*, [London School of Economics, Wiley, London School of Economics and Political Science, Suntory and Toyota International Centres for Economics and Related Disciplines], v. 4, n. 16, p. 386–405, 1937. ISSN 00130427, 14680335. Disponível em: <http://www.jstor.org/stable/2626876>.

CORDERO, J. M.; SALINAS-JIMÉNEZ, J.; SALINAS-JIMÉNEZ, M. M. Exploring factors affecting the level of happiness across countries: A conditional robust nonparametric frontier analysis. *European Journal of Operational Research*, v. 256, n. 2, p. 663–672, 2017. ISSN 0377-2217. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0377221716305628>.

DAOUIA, A.; GIJBELS, I. Robustness and inference in nonparametric partial frontier modeling. *Journal of Econometrics*, v. 161, n. 2, p. 147–165, 2011. ISSN 0304-4076. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0304407610002447>.

DAOUIA, A.; SIMAR, L. Nonparametric efficiency analysis: A multivariate conditional quantile approach. *Journal of Econometrics*, v. 140, n. 2, p. 375–400, 2007. ISSN 0304-4076. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0304407606001400>.

DARAIO, C.; SIMAR, L. Introducing environmental variables in nonparametric frontier models: a probabilistic approach. *Journal of Productivity Analysis*, Springer, v. 24, n. 1, p. 93–121, 2005.

DARAIO, C.; SIMAR, L. A robust nonparametric approach to evaluate and explain the performance of mutual funds. *European Journal of Operational Research*, v. 175, n. 1, p. 516–542, 2006. ISSN 0377-2217. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0377221705005126>.

DARAIO, C.; SIMAR, L. *Advanced robust and nonparametric methods in efficiency analysis: methodology and applications*. [S.l.]: Springer, 2007.

DEBREU, G. The coefficient of resource utilization. *Econometrica*, [Wiley, Econometric Society], v. 19, n. 3, p. 273–292, 1951. ISSN 00129682, 14680262. Disponível em: <http://www.jstor.org/stable/1906814>.

DEPRINS, D.; SIMAR, L. Measuring labor efficiency in post offices. In: MARCHAND, P. P. M.; TULKENS, H. (Ed.). *The Performance of Public Enterprises: Concepts and Measurements*. Amsterdam: North-Holland, 1984. p. 243–267.

FARRELL, M. J. The measurement of productive efficiency. *Journal of the Royal Statistical Society: Series A (General)*, v. 120, n. 3, p. 253–281, 1957. Disponível em: <https://rss.onlinelibrary.wiley.com/doi/abs/10.2307/2343100>.

HAMIDI, S.; SABOURI, S.; EWING, R. Does density aggravate the covid-19 pandemic? *Journal of the American Planning Association*, Routledge, v. 86, n. 4, p. 495–509, 2020. Disponível em: <https://doi.org/10.1080/01944363.2020.1777891>.

HAMZAH, N. M.; YU, M.-M.; SEE, K. F. Assessing the efficiency of malaysia health system in covid-19 prevention and treatment response. *Health Care Management Science*, v. 24, n. 2, p. 273–285, 2021.

KNEIP, A.; SIMAR, L.; WILSON, P. W. Asymptotic and consistent bootstraps for dea estimators in nonparametric frontier models. *Econometric Theory*, Cambridge University Press, v. 24, n. 6, p. 1663–1697, 2008.

KOOPMANS, T. C. An analysis of production as an efficient combination of activities. In: KOOPMANS, T. C. (Ed.). *Activity analysis of production and allocation*. New York: John-Wiley and Sons, Inc., 1951. cap. 3, p. 33–97.

KUMBHAKAR, S. C. et al. Nonparametric stochastic frontiers: A local maximum likelihood approach. *Journal of Econometrics*, v. 137, n. 1, p. 1–27, 2007. ISSN 0304-4076. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0304407606000376>.

LAZEAR, E. P. Economic Imperialism*. *The Quarterly Journal of Economics*, v. 115, n. 1, p. 99–146, 02 2000. Disponível em: <https://doi.org/10.1162/003355300554683>.

LECHNER, M. *The estimation of causal effects by difference-in-difference methods*. [S.l.]: Now Hanover, MA, 2011.

LI, Q.; RACINE, J. S. *Nonparametric econometrics: theory and practice*. [S.l.]: Princeton University Press, 2007.

LIU, J. S. et al. A survey of dea applications. *Omega*, v. 41, n. 5, p. 893–902, 2013. ISSN 0305-0483. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0305048312002186>.

MINVIEL, J. J.; BOUHENI, F. B. Technical and managerial efficiency assessment of european banks using a conditional nonparametric approach. *International Transactions in Operational Research*, v. 28, n. 2, p. 560–597, 2021. Disponível em: <https://onlinelibrary.wiley.com/doi/abs/10.1111/itor.12872>.

NEPOMUCENO, T. C. C. et al. Exogenous crimes and the assessment of public safety efficiency and effectiveness. *Annals of Operations Research*, Springer, p. 1–34, 2020.

PADHAN, R.; PRABHEESH, K. The economics of covid-19 pandemic: A survey. *Economic Analysis and Policy*, v. 70, p. 220–237, 2021. ISSN 0313-5926. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0313592621000321>.

PAGAN, A. et al. *Nonparametric econometrics*. Cambridge University Press, 1999. ISBN 978-0-5215-8611-5. Disponível em: <https://doi.org/10.1017/CBO9780511612503>.

RACINE, J. Consistent significance testing for nonparametric regression. *Journal of Business & Economic Statistics*, Taylor & Francis, v. 15, n. 3, p. 369–378, 1997. Disponível em: <https://www.tandfonline.com/doi/abs/10.1080/07350015.1997.10524714>.

RITCHIE, H. et al. Coronavirus pandemic (covid-19). *Our World in Data*, 2020. Disponível em: <https://ourworldindata.org/coronavirus>.

SEIFORD, L. M.; ZHU, J. Modeling undesirable factors in efficiency evaluation. *European Journal of Operational Research*, v. 142, n. 1, p. 16–20, 2002. ISSN 0377-2217. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0377221701002934>.

SHEPHARD, R. W. *Theory of Cost and Production Functions:*. Princeton University Press, 2015. ISBN 978-1-4008-7108-7. Disponível em: <https://doi.org/10.1515/9781400871087>.

SIMAR, L.; WILSON, P. Statistical inference in nonparametric frontier models: recent developments and perspectives. In: FRIED C. LOVELL, S. S. H. (Ed.). *The measurement of productive efficiency*. Oxford: Oxford University Press, 2008. cap. 4, p. 421–521.

SIMAR, L.; WILSON, P. W. Two-stage dea: caveat emptor. *Journal of Productivity Analysis*, Springer, v. 36, n. 2, p. 205–218, 2011. Disponível em: <https://doi.org/10.1007/s11123-011-0230-6>.

SIMAR, L.; WILSON, P. W. Estimation and inference in nonparametric frontier models: Recent developments and perspectives. *Foundations and Trends® in Econometrics*, v. 5, n. 3–4, p. 183–337, 2013. ISSN 1551-3076. Disponível em: <http://dx.doi.org/10.1561/0800000020>.

SUSSKIND, D.; VINES, D. The economics of the covid-19 pandemic: an assessment. *Oxford Review of Economic Policy*, v. 36, n. 1, p. S1–S13, 08 2020. ISSN 0266-903X. Disponível em: <https://doi.org/10.1093/oxrep/graa036>.

WITTE, K. D.; KORTELAINEN, M. What explains the performance of students in a heterogeneous environment? conditional efficiency estimation with continuous and discrete environmental variables. *Applied Economics*, Routledge, v. 45, n. 17, p. 2401–2412, 2013. Disponível em: <https://doi.org/10.1080/00036846.2012.665602>.