Competitiveness of Nations and Inequality-Adjusted Human Development: Evaluating the Efficiency of Nations Using DEA and Random Forest Classification

Christopher Colin Campbell

Abstract Although popular indices like the Human Development Index (HDI) and Global Competitiveness Index (GCI) measure human development and competitiveness separately, no index directly considers their linkage, namely, the relative ability of countries to leverage their economic competitiveness to improve the human development of their citizens. This paper aims to combine data envelopment analysis and random forest classification to explore the relative performance of countries in terms of competitiveness and human development. In the first stage of the methodology, we evaluate 124 countries using data envelopment analysis (DEA), taking indicators from the GCI and IHDI (inequality-adjusted human development index) as input and output variables, respectively. In the methodology's second stage, we use random forest classification to identify the relative importance of input and output variables on the DEA results—specifically, whether countries were classified as efficient or inefficient. Our findings indicate that only 20 of 124 countries are efficient at using their competitiveness to generate human development, and that variables related to a country's innovation ecosystem are most important. The results suggest most countries fail to take full advantage of their economic resources amidst a period of rapid technological and social change; it also highlights huge disparities between different groups of countries (e.g. regions).

Keywords Competitiveness of nations · Human development · Random forests · Data envelopment analysis

---

The online version contains supplementary material available at [https://doi.org/10.1007/978-3-030-91851-4_5].

C. C. Campbell (✉)
HelloFresh, Birmingham, UK

Y. I. Topcu et al. (eds.), *New Perspectives in Operations Research and Management Science*, International Series in Operations Research & Management Science 326,
https://doi.org/10.1007/978-3-030-91851-4_5

1 Introduction

In this paper, we examine the relationship between human development and the competitiveness of nations using well-known indices—the IHDI and GCI—as their proxies.

'Human Development' is the idea that development should better peoples' lives by expanding the number of freedoms and opportunities at their disposal (Sen, 2001; Fukuda-Parr, 2003). The term is strongly associated with the work of Amartya Sen and, in particular, his capability approach. Still, topical literature suggests there are many sources of inspiration, including Aristotle (Seth & Santos, 2018), the Catholic Church (Seth & Santos, 2018), Jeremy Bentham, John Stuart Mill and John Rawls (Stanton, 2007). Even the United Nations Development Programme (UNDP), the institution which publishes the Human Development Index each year, describes the concept as having '[grown] out of global discussions . . . during the second half of the 20th Century' (HDRO, n.d.). These sweeping accounts imply that, while Sen's definition is central, human development has a fluid meaning that is subject to change and individual interpretation.

'Human development' first gained traction within the international development community during the 1990s among United Nations (UN) agencies and big international Non-Governmental Organizations (NGOs). At the time, there was mass recognition that development policy was too preoccupied with resource ownership and neoliberal economic theory; the resultant paradigm shift led to the creation of the Human Development Index, a measure which offers an alternative to development proxies, such as the Gross Domestic Product (GDP) or Gross National Income (GNI), which focus exclusively on economic productivity.

Today, the HDI is a leading tool for tracking and comparing countries' progress and, thus, we use it as our chosen proxy for human development in this paper. At the same time, however, readers should note that the HDI only represents a single interpretation of human development rather than a perfect or complete account. Indeed, the UNDP has never claimed their HDI is infallible. Quite the opposite, they deliberately designed the HDI to be 'flexible in both coverage and methodology [so that] gradual refinements [could be made]' (Kovacevic, 2010b, p. 1). Such refinements were made in 2010 when the UNDP updated the HDI's methodology and introduced its inequality-adjusted variant—the IHDI.

The HDI's current formulation looks at three dimensions—health, knowledge and standard of living (Conceição, 2019). And each of those dimensions is made up of either one or two indicators. 'Life expectancy at birth' is the indicator for health, 'Mean Years of Schooling' and 'Expected Years of Schooling' are the indicators for knowledge (50:50 weighting) and 'Gross National Income per capita' ($ PPP) is the indicator for 'standard of living'. After collecting the raw data, the UNDP normalizes them to produce indicator- and then dimension-scores that typically range from 0 to 1. The UNDP then obtains each country's final HDI score by taking the geometric mean of the separate dimension scores:

$$\text{HDI} = \sqrt[3]{I_{\text{Health}} \times I_{\text{Education}} \times I_{\text{Income}}}$$

where 'I_x' is one of the three dimensions.

Unlike the IHDI, HDI scores are not affected by the distribution of achievement across the given population (Conceição, 2019). HDI scores can appear misleading when a small proportion of the population is responsible for a high proportion of the achievement within a dimension, which can mask critically low levels of achievement elsewhere (Kovacevic, 2010a). The IHDI overcomes this by including a variable which penalizes inequality called 'A'. This is calculated as $A = 1 - g/\mu$, where g is the geometric mean and μ is the arithmetic mean of data for the given indicator (Kovacevic, 2010a). Subsequently, the final IHDI is obtained as shown below:

$$\text{IHDI} = \left(\sqrt[3]{(1 - A_{\text{Health}}) \times (1 - A_{\text{Education}}) \times (1 - A_{\text{Income}})}\right) \times \text{HDI}$$

Accounting for inequality can drastically change a country's human development profile. For example, in 2019, IHDI scores were 19.5% lower on average than HDI scores, and some countries' rankings differed dramatically too. One such country was Brazil, whose IHDI rank was 23 places lower than its HDI rank. The extent of these discrepancies demonstrates the importance of taking inequality into account when measuring human development and our study.

The second key concept in our paper is the 'Competitiveness of Nations'. The basic notion of nations competing can be traced back through several chapters of classical and neoclassical economic theory; from mercantilism, through Adam Smith's theory of absolute advantage, David Ricardo's comparative advantage, the Heckscher-Ohlin model of factor endowment and onwards (García Ochoa et al., 2017; Mashabela & Raputsoane, 2018). Notwithstanding some continuity between these paradigms, the concept has evolved significantly to where it is today. In general, the competitiveness of nations is no longer treated as a zero-sum game (Schwab, 2019), and while international trade remains an important factor, judgements of success are no longer based on the ratio of imports to exports alone (García Ochoa et al., 2017). Contemporary conceptualizations are generally more complex, taking into account more micro- and macro-level factors (Ketels, 2016). Although, it is worth noting that they often fail to consider the military power of countries as mercantilism once did.

In our paper, 'Competitiveness of Nations' specifically refers to countries' economic capabilities relative to one another, with a twin focus on the well-being of citizens (Ülengin et al., 2011). Önsel et al. (2008) state that a nation is competitive if 'it can, under free and fair market conditions, produce goods and services that meet the standards of international markets while simultaneously expanding the real income of its citizens, thus improving their quality of life' (p. 222). Note that this definition treats economic productivity as a 'means' rather than an 'ends'—an important parallel with human development. The difference is that while patrons of human development assert the equal importance of health, education and the economy (which is visible in the HDI/IHDI), patrons of 'Competitiveness of Nations' are primarily concerned with economic means of development (Im & Choi, 2018).

The Global Competitiveness Index (GCI), published annually by the World Economic Forum (WEF), is one of two major indices that rank nations according to competitiveness. The latest version— the GCI 4.0—considers 'factors and attributes [driving competitiveness] in the [context] of the Fourth Industrial Revolution' (Schwab, 2019, p. vii). It comprises 103 indicators covering 12 'pillars' of competitiveness:

1.	Institutions	7.	Product market
2.	Infrastructure	8.	Labour market
3.	ICT adoption	9.	Financial system
4.	Macroeconomic stability	10.	Market size
5.	Health	11.	Business dynamism
6.	Skills	12.	Innovation capability

Once again, readers should note that the GCI only represents a single interpretation of the factors driving economic competitiveness between nations. Like human development, proclaiming a universal definition of the concept is problematic because it is so subjective and pluralistic. The contentious nature of the key concepts in our paper is a notable limitation of our work.

Following the collection of raw data, the WEF computes the GCI over two stages. Firstly, the WEF normalizes raw data to produce scores that typically range from 0 to 100. A score approaching 100 means the country is near an ideal situation 'where the factor no longer represents a constraint on productivity'. Conversely, a score approaching 0 indicates 'a completely unsatisfactory situation' (Schwab, 2019, p. 13). The equation below summarizes the normalization process:

$$\text{Normalized indicator score} = \left(\frac{\text{Actual value} - \text{Minimum value}}{\text{Maximum value} - \text{Minimum value}}\right) \times 100$$

Readers should note that the minimum and maximum values in the equation above vary between indicators and may reflect policy targets, naturally occurring minimums and maximums, or percentile figures derived from statistical analysis (Schwab, 2019, p. 13). The normalized data are aggregated over successive phases, proceeding from indicator-level up until each country's fully composite GCI score is obtained. Their aggregation method is to find the arithmetic mean of relevant subcomponents (Schwab, 2019, p. 13). At the end of the process, each country receives a score ranging from 0 to100.

In this paper, we assume that the primary objective of the nation's economy is the human development of its citizens. Therefore, we also assume that it is essential to examine the performance of countries in terms of how well they use their economic competitiveness to generate human development for their citizens. Building on previous work by Ülengin et al. (2011) and drawing on data from the 2019 editions of the GCI and IHDI, we propose a hybrid methodology—made up of data envelopment analysis (DEA) and random forest classification—to achieve this function.

Our specific objectives are listed below:

1. Apply DEA to determine the relative efficiency of countries in terms of their ability to leverage economic competitiveness to generate human development, taking GCI and IHDI indicators/dimensions as input and output variables, respectively.
2. Apply random forest classification to determine the relative importance of DEA variables on country performance.
3. Identify possible explanations for the first and second stage results, the implications for practice and future research, and the limitations of the study.

The rest of the chapter is organized as follows. The second section provides the theoretical context for the reader, introducing the human development and competitiveness concepts. The third section outlines the methodology, from the selection of countries for the dataset, to our decision to use DEA and random forest classification instead of alternatives. The fourth section summarizes the findings of our model. The fifth section discusses the study's practical and theoretical implications and its limitations. Finally, the concluding chapter summarizes the key takeaways, reiterates the aim and value of our study, before recommending areas for future research.

2 Literature Review

Few studies have focused on the relationship between competitiveness of nations and human development, and fewer still have examined the relationship in a manner similar to what we propose here. Cetinguc et al. (2018) explored the relationship between the GCI and HDI but their theorization focused on how human development could be applied to produce competitiveness, rather than vice versa like our paper. Nonetheless, the paper's findings support the existence of a relationship between the GCI and HDI, with the authors concluding that countries should invest more in their human capital to foster the 'innovativeness' and competitiveness of their economies. Bucher (2018) confirmed a strong correlation between the HDI and GCI but did not speculate any further on the nature of the relationship. There was also no discussion of the relative performance of countries, or the relative importance of the underlying factors.

Very much the forerunner and inspiration of this study, Ülengin et al. (2011) used data from the 2005 editions of the GCI and HDI to build DEA models that assessed how well nations generated human development from economic competitiveness. Subsequently, they applied ANN analysis to investigate the relative importance of sub-factors (i.e. the sub-indices of the GCI and HDI). Like us, they prescribed a hierarchy and direction of influence between the variables, asserting—on philosophical grounds—that economic competitiveness is only meaningful if it improves human development. The study also employed a super-efficiency DEA model to issue distinct scores and ranks to efficient countries.

The results indicated that wealthier countries with more developed economies tended to have higher efficiency scores. Africa was the worst-performing region. Surprisingly, South American countries—including Venezuela and Argentina—performed exceptionally well, with an average score higher than Europe and North America. The authors noted that the average scores of Europe and North America, though high and approaching one, were lower than anticipated. The ANN analysis indicated that GDP per Capita was by far the most influential sub-factor, followed by Life Expectancy and Efficiency Enhancers (a GCI sub-index covering market size, adoption of technology, plus the quality of financial and labour markets).

Kılıç and Kabak (2019) proposed two DEA models, having assumed the possibility of a bi-directional relationship between economic competitiveness and human development. One model examined countries' efficiency at producing economic competitiveness from human development, while the other examined the opposite direction of influence. One of the distinguishing elements of the study was the assumption of a three-year time lag between cause and effect. Hence, the authors paired input data from 2007 to 2014 with output data from 2010 to 2017. In addition, they adopted a time window approach to calculate yearly scores, using averages that took neighbouring years into account. Finally, the study used cluster analysis to investigate the stability of countries' DEA performances over the period (2010–2017) and to provide additional insight about the relationship between the variables.

The results indicated that the GCI-to-HDI DEA model was more reliable than its HDI-to-GCI counterpart, leading Kılıç and Kabak (2019) to conclude that the predominant direction of influence is from competitiveness to human development. This assertion is significant when previous studies have elected to depict human development as a determinant of economic competitiveness (see: Cetinguc et al., 2018; Bucher, 2018).

Cluster analysis indicated high levels of stability, meaning countries tended to remain in the same band of achievement between 2010 and 2017. Kılıç and Kabak (2019) interpreted this as proof of the model's veracity, but it could also imply that nations' capabilities had become crystallized. This fits with the pattern of the DEA results, which echoed the findings of Ülengin et al. (2011) albeit with fewer surprises: the best-performing nations were almost exclusively highly developed and affluent.

In their other study, Kılıç and Kabak (2020) used DEA alongside Fuzzy Analytical Network Process (FANP). Again, their goal was to investigate a bi-directional relationship between economic competitiveness and human development. However, this time, the authors compared the results of two DEA models—one for each direction of influence—with those from a separate composite index weighted according to FANP. Once again, the authors assumed a time lag of 3 years between cause and effect, using GCI and HDI data from 2012 and 2015.

Kılıç and Kabak (2020) reasoned that FANP offered a means of incorporating the complexity inherent in multi-criteria decision-making, where inter-dependencies and hierarchies exist between and within variables, alongside uncertain human decision-makers. We agree that FANP is an exciting alternative to the prescriptive and somewhat rudimentary weighting of subcomponents in the GCI and HDI. That said, we also recognize its limitations. For example, in Kılıç and Kabak (2020)'s application, FANP involved respondents making highly complex determinations in reductive numerical terms. Moreover, the questions, which covered multiple social science disciplines, were directed at just two experts. We feel a more targeted consultation of a larger field of experts would make the approach more credible.

Kılıç and Kabak (2020)'s main conclusion was that competitiveness has a greater effect on human development than vice versa, reinforcing the findings of their 2019 paper. This was due to the authors observing a stronger correlation between the results of the GCI-to-HDI DEA model and the FANP index. It is also notable that the GCI-to-HDI model showed surprising high performers, including Algeria and Venezuela, alongside the likes of Australia and Norway. The FANP results were less surprising by comparison.

This study distinguishes itself from the existing literature in the following ways. Firstly, it uses data from the 2019 editions of the HDI and GCI, which is doubly significant because the indices' methodologies were updated in 2010 and 2019, respectively. Thus, it offers a fresh snapshot and analysis of country performance. This study also separates itself by using the inequality-adjusted HDI to populate the output side of the DEA model. As we have already noted, accounting for inequality can dramatically change one's perception of a country's progress. With that in mind, the adoption of the IHDI is a clear methodological advancement. Finally, this study is the first to combine random forest classification with DEA in this field. In our methodology below, we elaborate further on why we selected random forest classification over alternatives.

In addition to making an original contribution to topical literature, this paper also has practical relevance for policymaking, governance and commerce (alongside similar studies and indices in general). As world-leading indices, the HDI and GCI are already magnificent tools but, by using DEA to unify them, we extend their utility. Likewise, when combined with DEA, machine learning offers another convenient way of extracting further insight from their rich data. Ultimately, through the proffer of this tool, we hope to promote government accountability and efficacy via the benchmarking of nations' achievements or the identification of areas for further improvement. Equally, we wish to support businesses as they weigh the conditions of countries vying for their investment.

3 Methods

Our methodological framework combined DEA and random forest classification, adapting and improving on Ülengin et al. (2011). Our first objective was to use DEA to assess the relative efficiency of 124 countries, measuring their ability to convert economic competitiveness into human development. Our second objective was to identify the relative importance of variables affecting countries' DEA outcomes. To this end, we trained and deployed a random forest classification model. Figure 1 provides an overview of the methodology.

3.1 Selection of Countries

Our analysis included all 124 countries featuring in *both* the 2019 GCI and IHDI. The complete list of these countries is in the electronic companion which can be downloaded from the book's website, alongside respective scores and ranks. Unlike Ülengin et al. (2011), we decided not to filter or cluster the eligible countries further because choosing criteria for doing so would have been highly subjective and perhaps controversial. Furthermore, we did not feel it was imperative since the WEF and UNDP present all countries together on their indices.

A larger sample of countries offered methodological advantages too. Firstly, concerning DEA, it increased our chances of 'capturing high-performance units that would determine the efficient frontier and improve the discriminatory power' of our model (Sarkis, 2007, p. 1–2). It also facilitated having more input and output variables (Dyson et al., 2001). Concerning the second stage of our methodological framework—random forest classification—a larger base meant the training and test subsets could also be significant, which helped us train the model (Beleites et al., 2013). Finally, it also reduced the risk of 'overfitting' and, in general, gave us extra space to experiment with different parameters and techniques (Riley et al., 2020).

3.2 Stage 1: Data Envelopment Analysis

DEA is a linear programming technique for determining the relative efficiency of a set of entities, referred to as decision-making units (DMUs). It is a non-stochastic, non-parametric alternative to econometric models such as regression analysis (Ülengin et al., 2011). Convenience is one of its major advantages, since it does not require 'assumptions regarding the statistical properties of variables' and offers the researcher a significant amount of discretion (Ülengin et al., 2011, p. 19). In addition, it gives the researcher the control over extraneous constraints, plus the selection and weighting of variables, though this places an extra burden on the researcher to make the model meaningful (Ruggiero, 1998; Ülengin et al.,

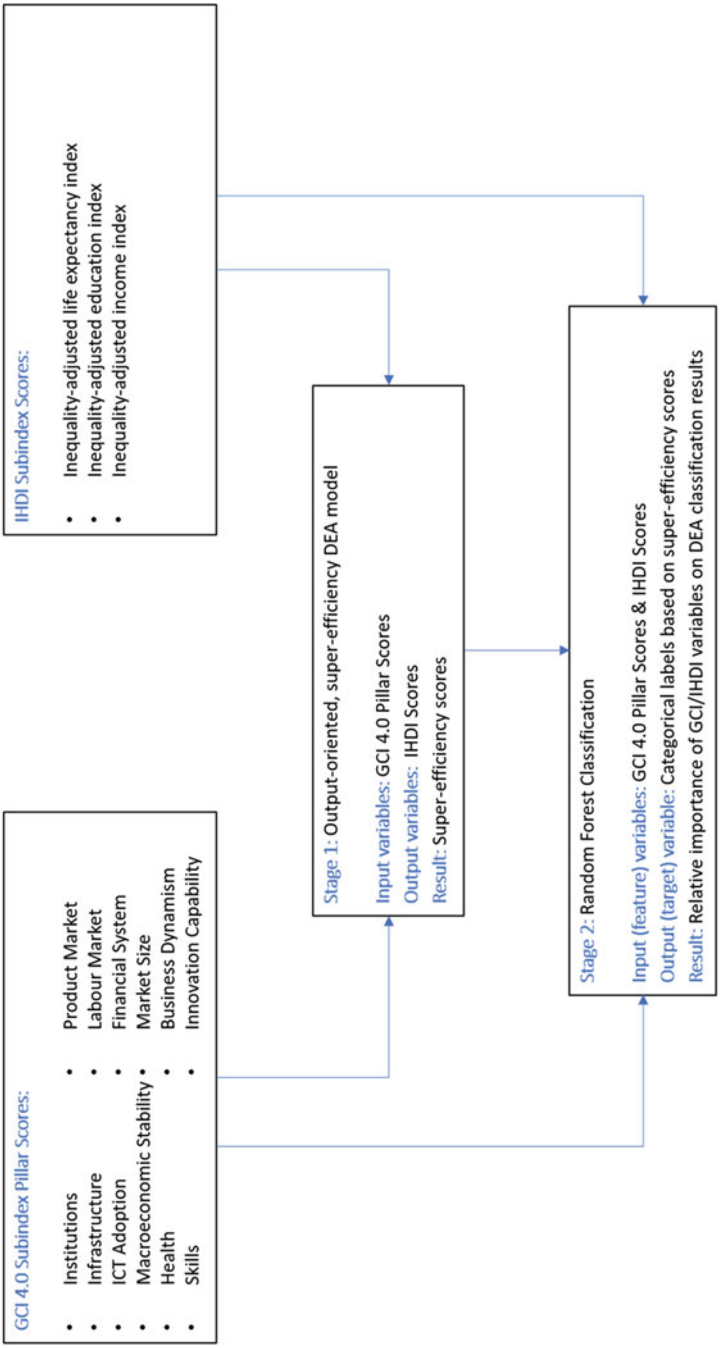

Fig. 1 Summary of the methodology

Table 1 GCI 4.0 pillars and categories

Enabling environment	**Markets**
1. Institutions	7. Product Market
2. Infrastructure	8. Labour Market
3. ICT Adoption	9. Financial System
4. Macroeconomic Stability	10. Market Size
Human capital	**Innovation ecosystem**
5. Health	11. Business Dynamism
6. Skills	12. Innovation Capability

Table 2 Revised DEA output variables

Dimension of human development	Ülengin et al. (2011)	This paper
Health	Life expectancy	Inequality-adjusted life expectancy index (life expectancy at birth)
Knowledge	Combined gross enrolment ratio (primary, secondary and tertiary)	Inequality-adjusted education index (expected years of schooling; mean years of schooling)
Standard of Living	GDP per capita	Inequality-adjusted income index (GNI per Capita)

2011). Nonetheless, DEA remains a popular option and continues to appear in many contexts, including this field of inquiry (see: Ülengin et al., 2011; Liu et al., 2013a; Mariano et al., 2015; Kılıç & Kabak, 2019, 2020).

Our DEA model consisted of 124 DMUs (countries), plus 12 input and 3 output variables from the GCI and IHDI, respectively. We view the output of human development as the goal of all nations, so we employed an output-oriented model. Regarding the input side, the WEF formerly grouped the 12 pillars of competitiveness (GCI) under three categories:

- Basic requirements
- Efficiency enhancers
- Innovation and sophistication factors.

Ülengin et al. (2011), Kılıç and Kabak (2019, 2020) adopted these categories as inputs for their DEA model, but they are obsolete since the emergence of the GCI 4.0. Today, the 12 pillars are arranged into the four categories as follows:

However, instead of adopting these categories as inputs, we used the 12 pillars for a more granular picture of relative importance. And in a significant departure from previous studies, we populated the output side of our model with variables from the IHDI instead of the HDI. Table 2 illustrates the difference versus Ülengin et al. (2011), whose output variables came from the pre-2010 HDI.

We used the IHDI to account for the distribution of achievement across the population. The IHDI's formulation penalizes countries in proportion to the unevenness of their outcomes: countries with higher levels of inequality are penalized more. The input and output data were from 2019 editions of their respective indices.

Fig. 2 Inversion of raw input data for DEA

The raw input and output data were already in index form, so there was no need for normalization or standardization to address issues related to scale and magnitude (Sarkis, 2007). However, since DEA is a measure of efficiency, models customarily reward lower consumption of inputs and higher production of outputs (Lewis & Sexton, 2004). In other words, input and output variables are usually defined as costs and benefits, respectively. While our output variables fulfilled this stipulation in their raw form, input variables from the GCI were inverted to reflect cost instead of benefit (see Fig. 2):

Any score below 1 after this transformation was assigned a score of 1 instead, to avoid input variables near or equal to zero distorting DEA results (see Appendix D).

Like Ülengin et al. (2011), we based our model on the classic Charnes–Cooper–Rhodes (CCR) blueprint (Charnes et al., 1978) and thus assumed constant returns to scale. The alternative Banker–Charnes–Cooper (BCC) framework (Banker et al., 1984) assumes variable returns to scale. 'Returns to scale' determines the shape of the efficiency frontier and affects the evaluation of DMUs. We took the view that having constant returns to scale was more appropriate because it implied that marginal gains in competitiveness should lead to equally proportional benefits for the citizens of the given country.

Our model also employed super-efficiency. DEA efficiency scores typically range from zero to one, with efficient units scoring one and inefficient units scoring less than one. So, while this facilitates the sortation of efficient DMUs (scores = 1) from inefficient ones (scores <1), efficient units are not differentiated (Andersen & Petersen, 1993). Super-efficiency models address this limitation by enabling individual scores of 1 *or* above for efficient units too.

Overall, we employed the same output-oriented, CCR-based, super-efficiency model used by Ülengin et al. (2011):

$$\text{Maximize } \eta_0, \text{ subject to :}$$

$$x_{i0} - \sum_{j=1;\mathrm{j}\neq 0}^{n} \mu_j x_{ij} \geq 0 \; i = 1, \ldots, m,$$

$$\eta_0 y_{r0} - \sum_{j=1;\mathrm{j}\neq 0}^{n} \mu_j y_{rj} \geq 0 \; r = 1, \ldots, s,$$

$$\mu_j \geq 0 \quad j = 1, \ldots, n$$

The model assumed there were n comparable DMUs, which all use m inputs x_{ij} ($i = 1, \ldots, m$) to produce r outputs y_{rj} ($r = 1, \ldots, s$). The super-efficiency value for DMU_0 was subsequently obtained as the value of $\frac{1}{\eta_o}$. Efficient DMUs obtained scores of 1 or above, while inefficient DMUs obtained score of less than 1. The model performed two functions. Firstly, it sorted DMUs into two groups, efficient and inefficient, and, secondly, it ranked all DMUs from most efficient to least.

The model was built in AMPL, using the CPLEX solver (see Appendix E). AMPL uses a syntax which '[closely matches] that of the algebraic, symbolic representation of a linear programming model' and thus readily accommodated our needs (Green, 1996, p. 1). The code for the model drew on an example from Cooper et al. (2007).

3.3 Stage 2: Random Forest Classification

In the second stage of the methodology, we trained a random forest to estimate the relative importance of the DEA variables on a country's classification as either efficient or non-efficient. A *random forest* is an ensemble learning algorithm that combines the predictive power of multiple, independently formed decision trees to perform classification or regression tasks (Breiman, 2001). Their 'randomness' has two sources. First, the decision trees that constitute the forest are trained with random subsets of data; second, each layer of the tree's node-splitting process uses a random subset of feature variables (Breiman, 2001). Hence, the trees are encouraged to take their own random approach to classification. Crucially, when classification outcomes differ between trees, the final prediction is based on the trees' aggregate decision, as displayed in Fig. 3.

As a classification tool, random forests boast several advantages. They perform better on classification tasks than alternatives like neural networks and support vector machines (Cutler et al., 2007; Fukuda et al., 2016), yet are remarkably user-friendly (Liu et al., 2013a). Like DEA, they do not require assumptions about the statistical properties of data (Liu et al., 2013b). Data pre-processing requirements are low (Liu et al., 2013b), and, aside from the configuration of parameters, they function with minimal human input (Lebedev et al., 2014). Other advantages include:

- resilience to overfitting
- their utility for determining the relative importance of feature variables (Cutler et al., 2007; Fukuda et al., 2016)
- 'their ability to model complex interactions between predictor variables' (Cutler et al., 2007, p. 2783)
- their modest computational processing requirements (Lebedev et al., 2014)

So, for these reasons and despite an artificial neural network featuring in Ülengin et al. (2011) to good effect, we preferred to use a random forest.

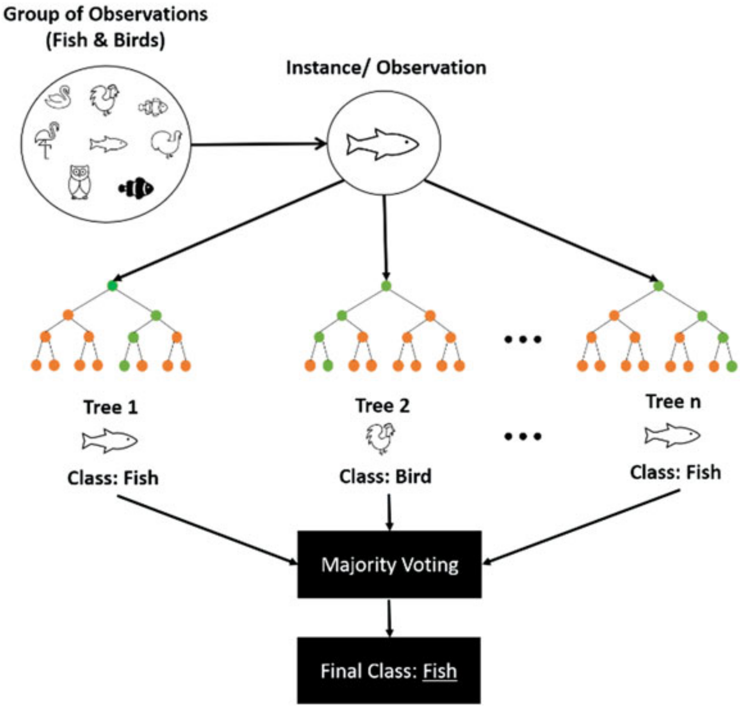

Fig. 3 Random forest classification

We built the random forest using the Python library 'Scikit-learn' (Pedregosa et al., 2011) (see Appendix F). Then, we trained it to predict whether a country was efficient or non-efficient using DEA variables and results. The model included 15 feature variables and one target variable. The feature variables were the input (GCI) and output (IHDI) variables used during DEA, while the target variable was an encoded version of countries' super-efficiency scores. Efficient countries were assigned scores of 1, while non-efficient countries were assigned scores of 0.

We used k-fold cross-validation to train the model and measure its accuracy, meaning the data was randomly split into k parts or overlaps, where the value of k is user-defined (Yadav & Shukla, 2016). We used stratification so that each fold preserved the same ratio of efficient to inefficient countries as the overall dataset. We then trained using '$k-1$' parts and used the remaining one part for testing. This process was repeated k times until each part had been used for training $k-1$ times and for testing once. Finally, we used the average testing performance as an indication of the model's overall accuracy.

When tuning the model for higher accuracy, we focused on two parameters—the number of folds (i.e. k-values) and the number of trees. After exploring the effects of different combinations, we decided that the optimum combination was 13 folds and 250 trees (see Fig. 4), which produced a cross-validation accuracy of 96.84%.

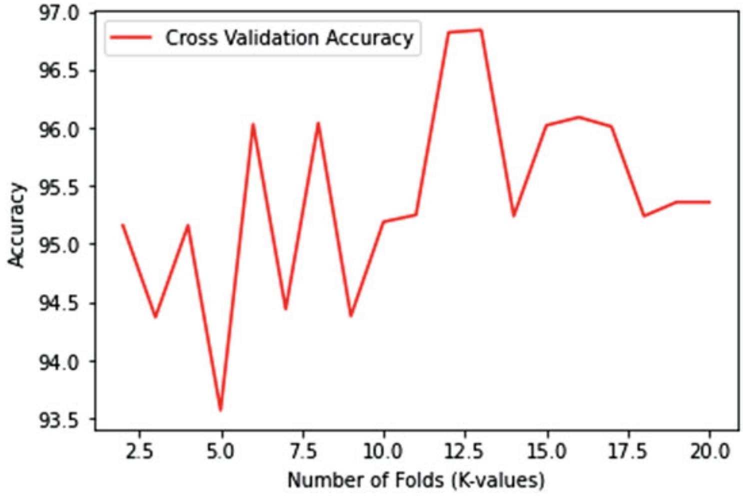

Fig. 4 Cross-validation accuracy with different *k*-values (number of trees = 250)

Following the advice of Probst et al. (2019) and given the modest size of our dataset plus the objectives of our analysis, we used Scikit's default settings for all other parameters (Pedregosa et al., 2011).

Last, we determined the relative importance of feature variables to the classification task using Scikit's designated 'feature importance' functionality (Pedregosa et al., 2011). The functionality offered two alternatives: computation could either be impurity- or permutation-based. The impurity-based option suited our study because our feature variables were continuous with similar, high levels of cardinality (Altmann et al., 2010).

4 Findings

4.1 Stage 1: Data Envelopment Analysis

The results revealed that most countries are inefficient at using their economic competitiveness to generate human development for their citizens. The median DEA score was 0.491, which means at least 50% of countries are less than halfway to reaching efficiency. Furthermore, there were only 20 efficient countries that obtained DEA scores greater than or equal to one. The remaining 104 countries all obtained scores of less than one, meaning they were inefficient. Appendix G provides a complete list of countries' super-efficiency scores and ranks.

Table 3 displays the best- and worst-performing countries according to the DEA results. The United States (USA) is the most efficient nation with an exceedingly

Table 3 Best- and worst-performing nations

			Bottom 20 countries		
Top 20 countries	Super-efficiency score (SE)	Super-efficiency rank		Super-efficiency score (SE)	Super-efficiency rank
United States	3.654	1	Guinea	0.261	105
Switzerland	1.650	2	Uganda	0.259	106
Japan	1.640	3	Eswatini	0.255	107
Korean Rep.	1.508	4	Lesotho	0.252	108
Germany	1.389	5	Malawi	0.251	109
Singapore	1.341	6	Zimbabwe	0.246	110
Netherlands	1.296	7	Ethiopia	0.234	111
China	1.276	8	Gambia	0.228	112
Sweden	1.230	9	Yemen	0.227	113
France	1.227	10	Angola	0.225	114
Hong Kong	1.215	11	Burundi	0.225	115
Finland	1.205	12	Mali	0.223	116
Denmark	1.153	13	Cameroon	0.215	117
Israel	1.112	14	Benin	0.214	118
Iceland	1.063	15	Burkina Faso	0.207	119
Norway	1.042	16	Mauritania	0.198	120
Australia	1.033	17	Mozambique	0.193	121
New Zealand	1.020	18	Haiti	0.183	122
United Kingdom	1.017	19	Congo (DR)	0.179	123
Canada	1.010	20	Chad	0.142	124

high score of 3.654. Switzerland (SE = 1.650) and Japan (SE = 1.64) are next. Chad is the least efficient nation with a score of 0.142, followed by Congo (SE = 0.179) and Haiti (SE = 0.183). The difference between a country's score and the efficiency frontier (SE = 1) signals the magnitude of a particular country's over- or under-achievement.

Table 4 provides further information on the best- and worst-performing countries. It shows that the best performers had better than average scores for every input and output variable and that the opposite was true for the worst performers.

Figure 5 shows the distribution of countries' efficiency scores over bands. With 22 countries, the most frequent band was $0.2 \leq SE < 0.3$. Next were the $1 \leq SE$ and $0.5 \leq SE < 0.6$ bands, with 20 countries falling into each. On the other hand, the least frequent bands were $0.8 \leq SE < 0.9$ (1 country) and $0.7 \leq SE < 0.8$ (2 countries). Overall, the distribution had a bimodal shape. Two distinct peaks were separated by a low-frequency trough covering the $0.6 \leq SE < 0.9$ range, containing only nine countries. It is unclear whether this shape is evidence of a deeper trend. If we assume it is, there are multiple possible explanations. One explanation is that an external influence, outside the scope of the model and beyond national control, keeps countries in two separate groups and makes it difficult for them to traverse the $0.6 \leq SE < 0.9$ range. It could also reflect a lack of internal resources and capabilities. Finally, there may be some other x-factor lacking among countries with scores below 0.6 but abundant in countries with scores above 0.9.

4.2 Correlation Analysis of Country Ranks

We used scatter plots and correlation analyses to investigate the degree of consistency between a country's rank on either the IHDI or GCI and its super-efficiency rank. We specifically used Kendall's test for the latter because ranking data is discrete and ordinal (Cliff, 1996).

The results indicate a strong relationship between a country's rank on either the GCI (Fig. 6; $\tau = 0.86$) or the IHDI (Fig. 7; $\tau = 0.75$) and its super-efficiency rank. The scatters shown in Figs. 6 and 7 are homoscedastic and portray positive linear relationships. Countries with better GCI or IHDI ranks generally obtained better super-efficiency ranks. As a result, these countries fell in the bottom-left corners of the graphs. On the other hand, countries with high (bad) GCI or IHDI ranks generally obtained high super-efficiency ranks and thus fell in the top-right corners of the graphs.

Table 5 shows the results of Kendall's test: the tau coefficient and p-values. Tau coefficients vary between -1 and 1 (Cliff, 1996). A negative value signifies an inverse relationship, while a positive value signifies a positive relationship. The tau coefficient reading was 0.86 for the relationship between GCI and super-efficiency ranks and 0.75 for IHDI and super-efficiency ranks. With p-values virtually equal to zero, these results indicate strong, positive associations. However, there is a stronger

Table 4 DEA variables for best- and worst-performing nations

	United States	Switzerland	Japan	Mean	Haiti	Congo (DR)	Chad
Input (GCI) variables							
Institutions	71.17	77.51	71.67	54.66	30.87	32.78	35.42
Infrastructure	87.90	93.16	93.16	64.83	26.88	29.23	30.53
ICT adoption	74.35	78.58	86.20	54.33	28.14	19.11	10.77
Macroeconomic stability	99.77	100.00	94.89	79.12	60.16	31.39	75.00
Health	83.02	99.94	100.00	74.83	50.82	41.60	35.87
Skills (Workforce)	82.47	86.72	73.28	60.88	41.48	42.30	29.04
Product market	68.55	63.80	70.36	54.90	37.81	44.73	35.43
Labour market	77.98	79.48	71.54	59.77	49.07	48.30	42.22
Financial system	90.99	89.72	85.94	61.77	44.02	42.53	37.31
Market size	99.53	66.23	86.86	54.93	33.92	43.27	37.09
Business dynamism	84.21	71.55	75.03	59.91	14.07	40.47	29.67
Innovation capability	84.15	81.20	78.31	43.09	18.90	17.97	22.66
Output (IHDI) variables							
Inequality-adjusted life expectancy index	0.85	0.94	0.96	0.72	0.46	0.40	0.31
Inequality-adjusted education index	0.85	0.88	0.84	0.58	0.28	0.35	0.16
Inequality-adjusted income index	0.70	0.82	0.85	0.56	0.21	0.23	0.31

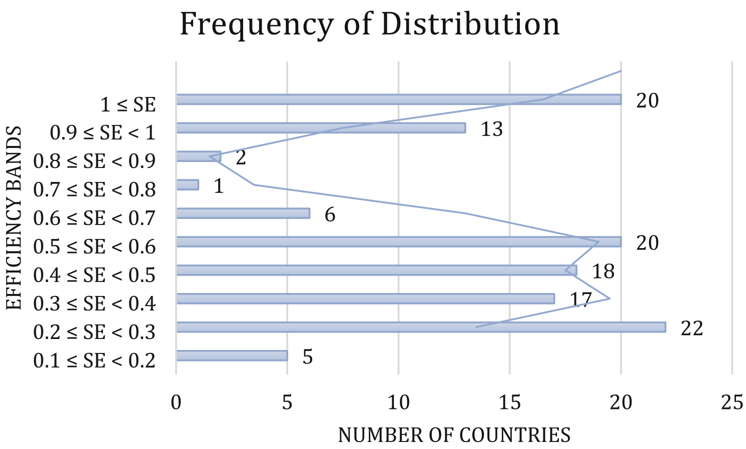

Fig. 5 Frequency of distribution across efficiency bands

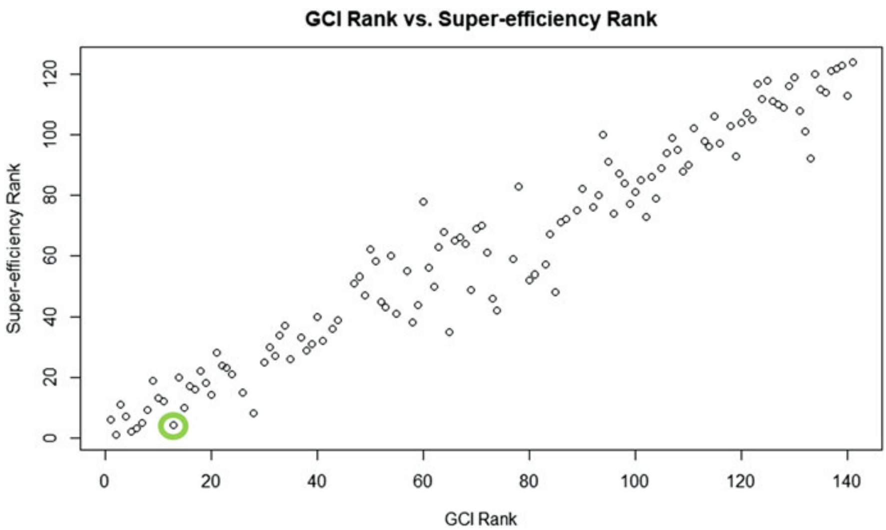

Fig. 6 Scatter plot of GCI rank and super-efficiency rank

Table 5 Results of Kendall's test

Relationship	Kendall's Tau coefficient (τ)	p-value (p)
GCI rank vs. super-efficiency rank	0.86	$p < 2.2e{-}16$
IHDI rank vs. super-efficiency rank	0.75	$p < 2.2e{-}16$

relationship between GCI rank and super-efficiency rank than between IHDI rank and super-efficiency rank, which explains why Fig. 6 is less scattered than Fig. 7.

Despite high consistency overall, super-efficiency ranks did not always correspond with GCI and IHDI ranks. China is the most extreme outlier. Its datum point

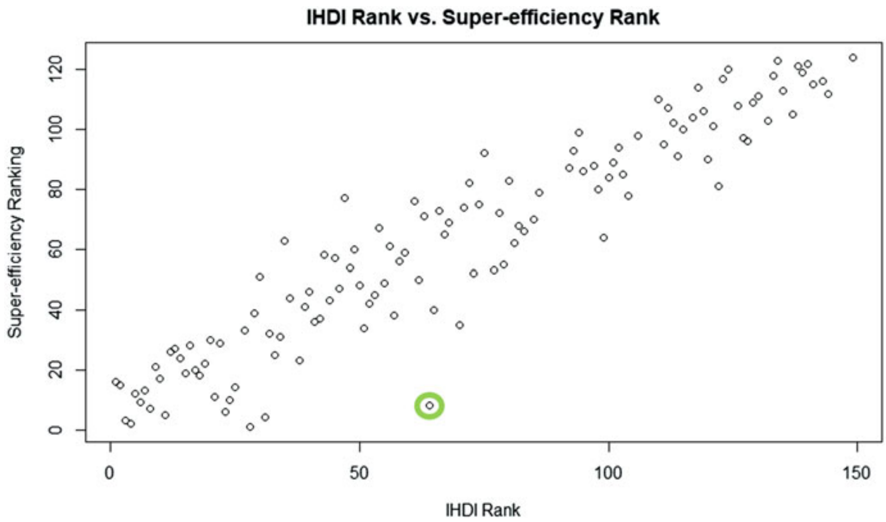

Fig. 7 Scatter plot of IHDI rank and super-efficiency rank

is circled in green in Figs. 6 and 7. Being on the bottom side of both scatters means China's DEA performance was better than expected, given its GCI and IHDI ranks. China obtained an eighth place super-efficiency rank but was only ranked 28th on the GCI and 64th on the IHDI. The implication is that China is highly efficient at leveraging its global competitiveness to achieve human development.

4.3 Analysing Group Performance

Although we avoided filtering or clustering countries in the pre-analysis, we considered group performance when interpreting and discussing results. Firstly, we considered geographic region, classifying countries according to their continent:

- Africa
- Asia
- Australia/Oceania
- Europe
- North America[1]
- South America

Secondly, we considered countries' affiliations with major intergovernmental organizations:

[1] Central America and the Caribbean were considered parts of North America.

- Emerging 7 (E7)
- Group of 7 (G7)
- Organization for Economic Co-operation and Development (OECD)

Thirdly, we considered countries 'Level of Human Development'. This classification was based on the 2019 HDR, with fixed cut-offs applied to HDI scores:

- 0.800 or above for very high human development
- 0.700–0.799 for high human development
- 0.550–0.699 for medium human development
- Less than 0.550 for low human development

Lastly, we considered the political regime of the countries. This classification was based on data from the V-Dem Project (Roser, 2013; Coppedge et al., 2019, 2021). Countries were grouped into 4 possible regime categories ranging from most to least democratic[2]:

- Liberal Democracy—Complete Democracy
- Electoral Democracy—Predominantly democratic with autocratic features
- Electoral Autocracy—Predominantly autocratic with democratic features
- Closed Autocracy—Complete Autocracy

Please refer to Appendix H for further detail on how specific countries were classified.

Table 6 summarizes group performance. The second and third columns show the ratio of efficient to inefficient countries within each group. The fourth column shows the geometric means of the groups' super-efficiency scores (GMSE). The use of geometric means (instead of arithmetic means) guaranteed that one country's good performance could not compensate for the bad performance of another, thus giving a better account of the group's central tendency.

Geographic Regions

Our analysis indicates that Africa is the least efficient geographic region. In addition to having no efficient countries, it also had the lowest average score (GMSE = 0.269). Its highest performing nation—Mauritius—only managed an efficiency score of 0.567 (Appendix G). Europe was the best-performing region. While having the highest number of countries at 42, the region had a high GMSE (0.805). Europe also had the most favourable ratio of efficient to inefficient countries outside of Australia/Oceania. Although Australia/Oceania had the highest GMSE (1.027), the group comprised two countries—Australia and New Zealand.

Intergovernmental Organizations

The performance of the E7 group was mediocre. The GMSE was 0.591, and China was the only efficient member of the group. We had anticipated slightly better results because economic efficiency is one of the connotations of 'E7' status.

[2] For further information on the features of democracy evaluated by the V-Dem Project, please refer to: Coppedge et al. (2021, pp. 254).

Table 6 Summary of DEA performance by group

	No. of countries	No. of efficient countries	Geometric mean of super-efficiency scores (GMSE)
Geographic region			
Africa	33	0	0.269
Asia	27	5	0.529
Australia/Oceania	2	2	1.027
Europe	42	11	0.805
North America	12	2	0.519
South America	8	0	0.463
Intergovernmental organization			
E7	7	1	0.591
G7	7	6	1.394
OECD	35	17	1.005
Level of human development			
Very High	48	19	0.914
High	34	1	0.479
Medium	21	0	0.323
Low	21	0	0.233
Political regime			
Closed Autocracy	9	2	0.500
Electoral Autocracy	36	1	0.343
Electoral Democracy	44	1	0.445
Liberal Democracy	35	16	0.912

With high scores and good ratios of efficient to inefficient countries (G7—6:7; OECD—17:16), the performances of the G7 and OECD groups were strong. These results felt unsurprising because the countries in these organizations tend to have wealthy societies, stable economies and high levels of human development.

Level of Human Development

As expected, the *Very High* group had the best mean performance (GMSE = 0.914) while the *Low* group had the worst (GMSE = 0.233). The *High* group also performed better than the *Medium* group.

Figure 8 shows how the average disparity in efficiency changed across consecutive levels of development. The most significant disparity was between countries with very high and high levels of human development. By contrast, the smallest disparity was between countries with medium and low levels of human development. The implication was that, as a country moves from low levels of human development to very high levels of human development, it becomes increasingly difficult to make the associated leaps in efficiency. This explanation also fits the bimodal distribution mentioned above in Fig. 5. The upper and lower bounds of this range (0.9 and 0.6) fit with the mean performances observed for the Very High (0.9) and High (0.5) groups. These findings suggest that the 0.6 to 0.9 range is critical.

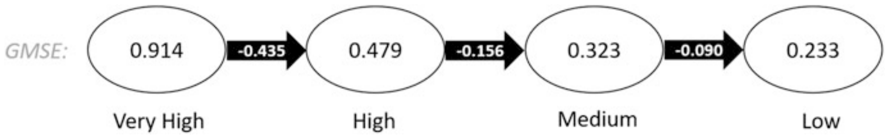

Fig. 8 Efficiency disparity across consecutive ‘levels of development’

Political Regime
According to our results, the two most efficient political regime types are Liberal Democracy and Closed Autocracy, while the least efficient are Electoral Autocracy and Electoral Democracy. Thus, the suggestion here is that efficiency benefits from either maximizing or minimizing democracy. As a caveat, it is worth highlighting that the sample size for closed autocracies was relatively small (9 countries), which meant China’s SE score could inflate the group’s average despite using the geometric mean.

4.4 Stage 2: Random Forest Classification

Using the same selection of 124 countries, we trained a random forest comprising 250 decision trees for binary classification. We trained the forest to separate efficient countries from non-efficient countries using the DEA model’s inputs and outputs. The forest achieved a cross-validation accuracy of 96.84% ($k = 13$).

After training the forest, we evaluated the relative importance of feature variables on the classification task. Figure 4 shows the calculation results, and Table 7 shows the full names of the feature variables. We found that GCI variables were far more influential in deciding a country’s classification as efficient or inefficient, accounting for 91.8% of ‘relative importance’. By contrast, IHDI variables only accounted for 8.2%. The single most influential variable was *Innovation Capability,* with a relative importance of 16.7%. *Business Dynamism* (15.7%) and *Institutions* (12.3%) followed. The least influential variables were *Health* (2.2%) and *Market Size* (2.1%).

We also considered the relative importance of the four GCI categories superseding the pillars. Table 8 shows *Innovation Ecosystem* had the highest relative importance with 32.4%, while the *Enabling Environment* was second with 26.8%.

5 Discussion

5.1 Implications for Policy and Business

The DEA model tended to reward countries with good GCI *and* IHDI performances; hence, super-efficiency rankings were strongly correlated with IHDI and GCI

Table 7 Key for horizontal labels in Fig. 9

GCI variables		IHDI variables	
INST	Institutions	INQLIFE	Inequality-adjusted life expectancy index
INFR	Infrastructure	INQEDU	Inequality-adjusted education index
ICT	ICT adoption	INQINC	Inequality-adjusted income index
MSTAB	Macroeconomic stability		
HEA	Health		
SKI	Skills (Workforce)		
PROD	Product market		
LAB	Labour market		
FIN	Financial system		
MAR	Market size		
DYN	Business dynamism		
INNOV	Innovation capability		

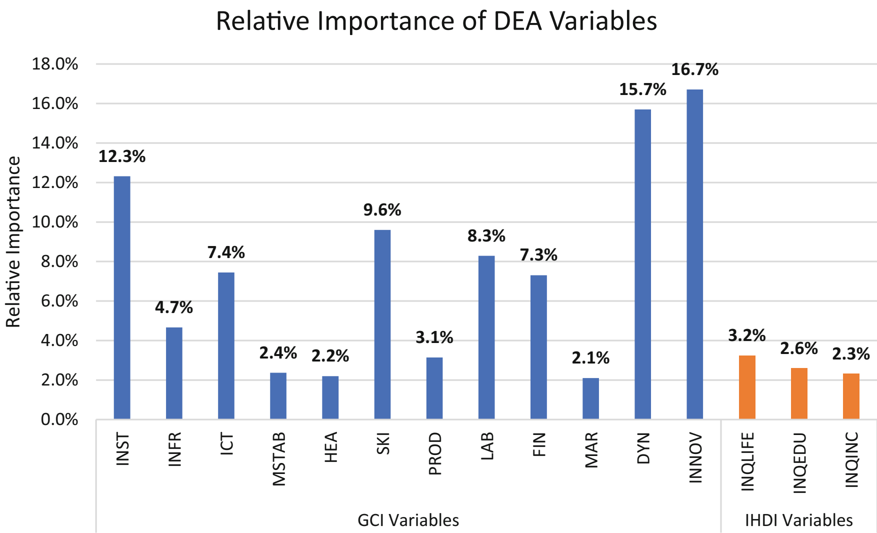

Fig. 9 Relative importance of criteria to DEA classification

Table 8 Relative importance of GCI categories

GCI super-category	Relative importance
Enabling environment	26.8%
Human capital	11.8%
Markets	20.8%
Innovation ecosystem	32.4%

rankings. However, there were exceptions like China, whose DEA performance outshone their IHDI and GCI performance. Overall, the model indicated that some countries are indeed better at leveraging their competitiveness to generate human development than others, reinforcing the findings of previous studies (see: Ülengin et al., 2011; Kılıç & Kabak, 2019, 2020).

Governments—and other parties invested in the betterment of a given country—should learn from the policies and practices of countries with better DEA performances than theirs, particularly those with similar characteristics. DEA models allow for the identification of 'peer units'—units that occupy a similar area of the efficiency frontier as a given unit under consideration (Bogetoft & Otto, 2010, p. 93; Ülengin et al., 2011, Kılıç & Kabak, 2019, 2020). Individuals and institutions may use 'peer units' or their discretion to determine the best candidates for comparison. For example, one might choose to focus on countries with similar GCI or IHDI scores; countries with similar resource endowments (see: Auty, 1998; Sachs, 1999); countries from the same region; or those with similar socio-cultural values, political systems and histories.

On the other hand, businesses could use the DEA model to direct corporate social responsibility (CSR) initiatives and other investments, using a country's efficiency score as a proxy for its potential.

The second part of our methodology identified the relative importance of variables underpinning DEA results. Notably, our results depicted trends *across* countries. Consequently, we cannot comment on how the relative importance of particular variables fluctuates between and within individual countries. Nonetheless, this information could also support decision-makers in government, policymaking and commerce.

The variables with the most significant influence on the DEA results were Innovation Capability and Business Dynamism. The GCI indicators that make up these variables (pillars) are shown in Table 9.

The high relative importance of a country's *Innovation Capability* and *Business Dynamism*—variables affecting the *Innovation Ecosystem*—seems timely given the advent of the fourth industrial revolution. Recent literature gives the impression that the relative importance of a country's innovation ecosystem will increase in the near

Table 9 Business Dynamism and Innovation Capability indicators (Schwab, 2019)

Business dynamism	Innovation capability
Diversity of workforce	Cost of starting a business
State of cluster development	Time to start a business
International co-inventions per million of the population	Insolvency recovery rate
Multi-stakeholder collaboration	Insolvency regulatory framework
Research and Development	Attitudes towards entrepreneurial risk
Scientific Publications Score	Willingness to delegate authority
Patent Applications per million of the population	Growth of innovative companies
Research and Development Expenditure	Companies embracing disruptive ideas
Research Institutions prominence	
Commercialization	
Buyer Sophistication	
Trademark Applications per million population	

future (Schäfer, 2018; Yang et al., 2019). There is an ongoing debate on whether the fourth industrial revolution is relevant to one of the most inefficient regions—Sub-Saharan Africa—given the region's socio-economic characteristics (Ayentimi & Burgess, 2019). However, the prevailing opinion appears to be that the region can benefit if governments direct their activities and resources towards their innovation ecosystems (Amankwah-Amoah et al., 2018; Asongu & Nwachukwu, 2018; Ogwo, 2018; Ayentimi & Burgess, 2019). All-in-all, our results reinforce the notion that countries should pay close attention to the fourth industrial revolution, showing that there are implications for human development.

Our results also emphasized the value of a countries' *Enabling Environment*, and in particular its *Institutions*. Again, this feels timely in the context of the fourth industrial revolution. Institutions are a clear theme in the literature cited above (Schäfer, 2018; Yang et al., 2019); they all press the idea that institutions (especially those from the public sector) will be especially crucial during the period of transformation ahead. Our results reinforce this view. In general, it is important to note that different areas of competitiveness are interdependent, so focusing exclusively those with the highest relative importance may not yield dividends.

Many in the West view democracy as a necessary ingredient for becoming an advanced society but our results challenge this. Although liberal democracies had the highest average DEA performance by far, electoral democracies—the most common political system in our dataset—performed very poorly on average. Although electoral democracies were superior to electoral autocracies, they did not outperform closed autocracies. Our results indicated that each type of political system is capable of producing efficiency.

5.2 *Stratified world system*

Our findings revealed two distinct classes of country, in terms of DEA performance. In Fig. 5, there was a bimodal distribution of efficiency scores, with two peaks separated by a trough spanning the 0.6–0.9 efficiency range. Figure 3 corroborated this. It showed that, as a country moves from low levels of human development to very high levels of human development, it becomes increasingly difficult to make the associated leaps in efficiency. The largest gap, by far, was between countries with very high human development, with an average DEA score of 0.914, and countries with high human development, which had an average of 0.479. Countries with high human development would have to increase their efficiency by approximately 91% to reach the standard of those with very high human development. Although further observation would help to corroborate this further, IHDI and GCI scores do not fluctuate greatly from year to year so it is likely that this is evidence of a trend rather than an anomaly (Conceição, 2019; Schwab, 2019).

The bimodality of our results throws into question the assumption of homogeneity that underpinned our selection of countries. It is possible that we should have followed the blueprint set by Ülengin et al. (2011), as well as Kılıç and Kabak (2019), by either clustering or filtering countries' pre-selection.

5.3 Limitations

The methodology entailed multiple assumptions and limitations. Firstly, it assumed that the countries under assessment were comparable in terms of internal resources and activities, as well as environmental factors. It also inherited the limitations and assumptions of the GCI and IHDI, having relied on these indices heavily as proxies for competitiveness and human development, respectively. Given the logistical challenge of collecting and verifying data across multiple countries simultaneously, both indices are likely to suffer from observational error. In previous sections, we also discussed how aspects of both indices are subjective, and therefore contestable. Generally, we viewed these flaws as inescapable by-products of decisions involved in the design of multidimensional indices. However, we were particularly concerned about the manner in which the GCI derives data from the Executive Opinion Survey because it produces highly subjective yet outwardly quantitative data. We were also briefly concerned that some of the DEA variables proxied the same or highly correlated phenomena. For example, 'health' is covered on both the input and output side of the model. Fortunately, it is extremely rare for inter-variable correlation to significantly impact DEA scores (Dyson et al., 2001).

6 Conclusions

This paper was predicated on the idea that the objective of a nation's economy competitiveness is to enhance the welfare of its citizens. It explored the relationship between human development, as measured by the IHDI, and competitiveness of nations, as measured by the GCI. In particular, it set out to determine (a) the relative performance of countries in leveraging their competitiveness to produce human development and (b) the relative importance of the factors, which facilitate that process.

The results were bleak. Most countries do not maximize the human-development-producing potential of their economies, and at least half are extremely inefficient, needing to more than double their current performance to become efficient. We found that the most important explanatory factors were related to the innovation ecosystem of the country, a discovery which felt timely given the ongoing buzz around the fourth industrial revolution.

The implications of our work are manifold. However, the main implication, particularly for policy, is that there is no substitute for tracking the connection between competitiveness and human development directly. Measuring human development

(IHDI) and competitiveness (GCI) in isolation is not the same as measuring 'the ability of a country to produce human development from its competitiveness'. In addition, we have also provided a practical tool, in the form of our hybrid DEA-random forest model, for such a task. While our results are deserving of attention and warrant further analysis, it is important to acknowledge the limitations of our methodology, in particular, those stemming from the HDI and GCI. They, like many other composite indices, are doomed to fallibility. This means our methodology is fallible too.

Further research is required to verify the underlying causes of some of the patterns we have observed. For instance, to explain the bimodality discovered in the frequency distribution of DEA/super-efficiency scores. Another area of intrigue concerns the surprising inefficiency of electoral democracies; more detailed analysis regarding the characteristics of political systems that nurture efficiency would be welcomed. Lastly, we recommend further studies of the relationship between countries' innovation ecosystems, their competitiveness and human development outcomes. With the fourth industrial revolution looming large, it is doubly important to understand this process so that as many countries as possible can take advantage of the opportunities to come.

References

Altmann, A., et al. (2010). Permutation importance: a corrected feature importance measure. *Bioinformatics*. Oxford University Press, *26*(10), 1340–1347.

Amankwah-Amoah, J., Osabutey, E. L. C., & Egbetokun, A. (2018). Contemporary challenges and opportunities of doing business in Africa: The emerging roles and effects of technologies. *Technological Forecasting and Social Change*. Elsevier, *131*, 171–174.

Andersen, P., & Petersen, N. C. (1993). A procedure for ranking efficient units in data envelopment analysis. *Management Science*. INFORMS, *39*(10), 1261–1264.

Asongu, S. A., & Nwachukwu, J. C. (2018). Openness, ICT and entrepreneurship in sub-Saharan Africa. In *Information Technology & People*. Emerald Publishing Limited.

Auty, R. M. (1998). *Resource abundance and economic development improving the performance of resource-rich countries*. UNU World Institute for Development Economics Research (UNU/WIDER).

Ayentimi, D. T., & Burgess, J. (2019). Is the fourth industrial revolution relevant to sub-Sahara Africa? *Technology Analysis & Strategic Management*. Taylor & Francis, *31*(6), 641–652.

Banker, R., Charnes, A., & Cooper, W. (1984). Some models for estimating technical and scale effficiencies in data envelopment analysis. *Managment Science, 30*, 1078–1092.

Beleites, C., et al. (2013). Sample size planning for classification models. *Analytica Chimica Acta*. Elsevier, *760*, 25–33.

Bogetoft, P., & Otto, L. (2010). *Benchmarking with dea, sfa, and r*. Springer Science & Business Media.

Breiman, L. (2001). Random forests. *Machine learning*. Springer, *45*(1), 5–32.

Bucher, S. (2018). The global competitiveness index as an indicator of sustainable development. *Herald of the Russian Academy of Sciences, 88*(1), 44–57. https://doi.org/10.1134/S1019331618010082

Cetinguc, B., et al. (2018). The relationships among the prominent indices: HDI-GII-GCI. In *Industrial engineering in the industry 4.0 era* (pp. 187–194). Springer.

Charnes, A., Cooper, W. W., & Rhodes, E. (1978). Measuring the efficiency of decision making units. *European Journal of Operational Research, 2*(6), 429–444. https://doi.org/10.1016/0377-2217(78)90138-8

Cliff, N. (1996). Answering ordinal questions with ordinal data using ordinal statistics. *Multivariate Behavioral Research*. Taylor & Francis*, 31*(3), 331–350.

Conceição, P. (2019). *Human Development Report 2019: beyond income, beyond averages, beyond today*. United Nations Development Programme.

Cooper, W. W., Seiford, L. M., & Tone, K. (2007). Data envelopment analysis: A comprehensive text with models, applications, references and DEA-solver software (2nd ed.). Springer.

Coppedge, M., et al. (2019). V-dem dataset V9. *SSRN Electronic Journal*. Elsevier BV. https://doi.org/10.2139/ssrn.3441052

Coppedge, M., Gerring, J., Knutsen, C. H., et al. (2019). V-dem country coding units v9 (V-dem): Varieties of democracy project. Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3441067

Coppedge, M., Gerring J., Knutsen, C. H., et al. (2021). V-dem methodology v9: Varieties of democracy (V-dem) project. Available at: https://www.v-dem.net/static/website/img/refs/methodologyv111.pdf

Cutler, D. R., et al. (2007). Random forests for classification in ecology. *Ecology*. Wiley Online Library*, 88*(11), 2783–2792.

Dyson, R. G., et al. (2001). Pitfalls and protocols in DEA. *European Journal of Operational Research*. Elsevier*, 132*(2), 245–259.

Fukuda, S., et al. (2016). Random forests hydrodynamic flow classification in a vertical slot fishway using a bioinspired artificial lateral line probe. In *International conference on intelligent robotics and applications* (pp. 297–307). Springer.

Fukuda-Parr, S. (2003). The human development paradigm: Operationalizing Sen's ideas on capabilities. *Feminist Economics, 9*(2–3), 301–317. https://doi.org/10.1080/1354570022000077980

García Ochoa, J. J., León Lara, J. d. D., & Nuño de la Parra, J. P. (2017). Propuesta de un modelo de medición de la competitividad mediante análisis factorial. *Contaduria y Administracion, 62*(3), 792–809. https://doi.org/10.1016/j.cya.2017.05.002

Green, R. H. (1996). DIY DEA: Implementing data envelopment analysis in the mathematical programming language AMPL. *Omega*. Elsevier*, 24*(4), 489–494.

Im, T., & Choi, Y. (2018). Rethinking national competitiveness: A critical assessment of governmental capacity measures. *Social Indicators Research*. Springer Netherlands*, 135*(2), 515–532. https://doi.org/10.1007/s11205-016-1494-z

Ketels, D. C. (2016). *Review of competitiveness frameworks* (pp. 1–58). Irish National Competitiveness Council. Available at: http://www.competitiveness.ie/Publications/2016/ReviewofCompetitivenessFrameworks.pdf

Kılıç, H., & Kabak, Ö. (2019). Analysis of relation between human development and competitiveness using fuzzy ANP and DEA. In *International conference on intelligent and fuzzy systems* (pp. 859–866). Springer.

Kılıç, H., & Kabak, Ö. (2020). Analysis of relation between human development and competitiveness using fuzzy ANP and DEA. In C. Kahraman, S. Cebi, & Cevik Onar, Sezi & Oztaysi, Basar & Tolga, A. Cagri & Sari, Irem Ucal (Eds.), *Intelligent and fuzzy techniques in big data analytics and decision making*, (pp. 859–866), Springer.

Kovacevic, M. (2010a). Measurement of inequality in Human Development–A review. *Measurement*, 35.

Kovacevic, M. (2010b). *Review of HDI critiques and potential improvements*. UNDP.

Lebedev, A. V., et al. (2014). Random Forest ensembles for detection and prediction of Alzheimer's disease with a good between-cohort robustness. *NeuroImage: Clinical*. Elsevier*, 6*, 115–125.

Lewis, H. F., & Sexton, T. R. (2004). Data envelopment analysis with reverse inputs and outputs. *Journal of Productivity Analysis*. Springer*, 21*(2), 113–132.

Liu, J. S., et al. (2013a). A survey of DEA applications. *Omega (United Kingdom)*. Elsevier*, 41*(5), 893–902. https://doi.org/10.1016/j.omega.2012.11.004

Liu, M., et al. (2013b). Comparison of random forest, support vector machine and back propagation neural network for electronic tongue data classification: Application to the recognition of orange beverage and Chinese vinegar. *Sensors and Actuators B: Chemical*. Elsevier, *177*, 970–980.

Mariano, E. B., Sobreiro, V. A., & Rebelatto, D. A. d. N. (2015). Human development and data envelopment analysis: A structured literature review. *Omega (United Kingdom)*. Elsevier, *54*, 33–49. https://doi.org/10.1016/j.omega.2015.01.002

Mashabela, J., & Raputsoane, L. (2018). Munich Personal RePEc Archive Important factors in a nations international competitiveness ranking Important factors in a nations international competitiveness ranking, (86477).

Ogwo, B. A. (2018). Re-visioning technical vocational education and training (TVET) for the youth in sub-Saharan Africa (SSA) and the Sustainable Development Goals (SDGs): Prospects and promises within the framework of the ubuntu paradigm. In *Re-Visioning Education in Africa* (pp. 155–173). Springer.

Önsel, Ş., et al. (2008). A new perspective on the competitiveness of nations. *Socio-Economic Planning Sciences, 42*(4), 221–246. https://doi.org/10.1016/j.seps.2007.11.001

Pedregosa, F., et al. (2011). Scikit-learn: Machine learning in Python. *The Journal of Machine Learning Research*. JMLR.org, *12*, 2825–2830.

Probst, P., Wright, M. N., & Boulesteix, A. (2019). Hyperparameters and tuning strategies for random forest. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*. Wiley Online Library, *9*(3), e1301.

Riley, R. D., Ensor, J., Snell, K. I. E., Harrell, F. E., Martin, G. P., Reitsma, J. B. et al. (2020). Calculating the sample size required for developing a clinical prediction model, *BMJ 368*, m441. https://www.bmj.com/content/368/bmj.m441

Roser, M. (2013). *Democracy, our world in data*. Available at: https://ourworldindata.org/democracy. Accessed 21 September 2020.

Ruggiero, J. (1998). Non-discretionary inputs in data envelopment analysis. *European Journal of Operational Research*. Elsevier, *111*(3), 461–469.

Sachs, J. D. (1999). Resource endowments and the real exchange rate: a comparison of Latin America and East Asia. In *Changes in exchange rates in rapidly developing countries: Theory, practice, and policy issues* (pp. 133–154). University of Chicago Press.

Sarkis, J. (2007). Preparing your data for DEA. In *Modeling data irregularities and structural complexities in data envelopment analysis* (pp. 305–320). Springer.

Schäfer, M. (2018). The fourth industrial revolution: How the EU can lead it. *European View*. SAGE Publications, *17*(1), 5–12. https://doi.org/10.1177/1781685818762890

Schwab, K. (2019). *The Global Competiitiveness Report 2019*. World Economic Forum. Available at: http://www3.weforum.org/docs/WEF_TheGlobalCompetitivenessReport2019.pdf

Sen, A. (2001). *Development as freedom*. Oxford Paperbacks.

Seth, S., & Santos, M. E. (2018). *Multidimensional inequality and human development, OPHI working paper 114*. University of Oxford. https://doi.org/10.2139/ssrn.3113294

Stanton, E. A. (2007). Human development index: A history. *Political Economy Research Institute Working Paper Series, 127*(February), 1–36.

Ülengin, F., et al. (2011). The competitiveness of nations and implications for human development. *Socio-Economic Planning Sciences*. Elsevier Ltd, *45*(1), 16–27. https://doi.org/10.1016/j.seps.2010.10.001

United Nations Human Development Report Office. (n.d.) *What is human development? | Human development reports*. Available at: http://hdr.undp.org/en/content/what-human-development. Accessed 6 July 2020.

Yadav, S., & Shukla, S. (2016). Analysis of k-fold cross-validation over hold-out validation on colossal datasets for quality classification. In *2016 IEEE 6th international conference on advanced computing (IACC)* (pp. 78–83). IEEE.

Yang, H., Kim, S. Y., & Yim, S. (2019). A case study of the Korean government's preparation for the fourth industrial revolution: public program to support business model innovation. *Journal of Open Innovation: Technology, Market, and Complexity*. Multidisciplinary Digital Publishing Institute, *5*(2), 35.