International Labour Organization | ACT/EMP

A Conceptual Framework for Measuring Business Resilience

First published 2023

Title: *A conceptual framework for measuring business resilience*

ISBN: 9789220385531 (web PDF)

Printed in Switzerland

A conceptual framework for measuring business resilience

July 2023

Bureau for Employers Activities (ACT/EMP)
International Labour Organization

Preface

In an era of unforeseeable shocks and disruptions, resilience is a crucial attribute for businesses of all sizes. It is in this context that our study, “A Conceptual Framework for Measuring Business Resilience,” emerges as a tool to systematically measure business resilience and offer insights to policymakers worldwide building upon the study we recently published under the title of “Determinants of Productivity and business resilience”. The COVID-19 outbreak, an extreme and widespread economic shock, underscored the urgency of comprehending the dynamics of business survival, particularly for start-ups and smaller firms that operate with limited resources. The Business Resilience Index (BRI) proposed in this report is an attempt to offer a comprehensive measure of resilience, which we believe is critical to informing policy decisions and interventions.

The report encompasses a detailed process of building the BRI, beginning from identifying the components of business resilience, outlining the challenges of data acquisition, to the application of various methods to synthesize a composite index. We also acknowledge that the current version of the BRI signifies the capacity for resilience rather than the outcome, making this study an essential building block for future research.

Our findings elucidate how different methods can impact the relative positions of countries in the BRI, the importance of tailored policy measures, and the relevance of business environment alongside organizational and managerial capabilities in influencing business resilience. We trust this research will guide governments, workers’ and employers’ organizations in developing policies that can enhance the resilience capacity of businesses, particularly in the critical areas of finance and technology.

The endeavour to develop a robust BRI would not have been possible without the collective efforts of our dedicated team. My sincere gratitude goes to Samuel Asfaha and José Luis Viveros Añorve, ILO-ACT/EMP officials, for conceptualising the research, leading the research process, reviewing several drafts of the report and providing valuable technical inputs. We extend our heartfelt appreciation to Hernan Viscarra Andrade for conducting an exhaustive literature review and drafting the report, displaying an exemplary commitment to thoroughness and clarity. We also wish to thank Roberto Leombruni of the University of Turin for his rigorous review of the report and invaluable contributions that substantially improved its quality. Finally, special thanks are extended to Ward Rinehart for his meticulous editing and keen attention to detail, greatly improving the clarity and coherence of this report.

We acknowledge the existence of certain limitations, including the reliance on secondary data and the challenge of capturing the organizational and managerial capabilities of enterprises that can significantly affect business resilience. We hope these limitations will not be seen as shortcomings but rather as opportunities for future researchers to delve deeper into these critical aspects. This study is, after all, one of the initial steps in a long journey to comprehensively understand, measure, and enhance business resilience in a global landscape increasingly susceptible to shocks and disruptions.

To our constituents, we hope this report offers valuable insights, stimulates thoughtful dialogue, and prompts effective action to bolster the resilience and survival of businesses around the globe, especially in the face of future adversities.

Deborah France-Massin

Director
Bureau for Employers’ Activities (ACT/EMP)
International Labour Office

▶ Contents

Tables

Figures

Introduction

Introduction

The purpose of this study is to build a business resilience index (BRI). This index is meant as a tool to help policymakers worldwide enhance the resilience and survival of firms – an issue brought to the fore by the COVID-19 outbreak.

The COVID-19 outbreak tested the ability of firms to adapt to changes and keep their operations afloat. Most businesses had to adapt their productive processes to deal with restrictions to transport and labour mobility and to workplace closures. The issue has been most challenging for start-ups and micro and small firms, which have more limited resources to cope with changing environments (Runyan 2006).

The advent of extreme and generalized economic shocks highlights the importance of understanding the drivers of firms' survival and resilience so as to avoid lasting losses in employment, investment and labour productivity. No country is immune to such disruptions in the future. This makes it important to be able to measure resilience in a comprehensive manner and to make policy decisions based on this evidence.[1]

As one of the first efforts to develop a BRI, this study not only seeks to provide policy makers with tools to help businesses but also could serve as a building block for future studies.

As one of the first efforts to develop a BRI, this study not only seeks to provide policy makers with tools to help businesses but also could serve as a building block for future studies.

1 Datta (2017) classifies disruptions into three categories: (i) unexpected events (natural disasters, vulnerability to technological change, new competition, etc.); (ii) internal practices or operational vulnerability; and (iii) complexity (strategic vulnerability due to industry type, globalization, outsourcing, technological intricacy, supplier dependency, etc.).

Chapter 1

1. Literature review

The earliest literature on business resilience started with the hypothesis that firms' survival depends significantly on productivity. The theory suggests that during downturns small and less efficient firms are the ones to exit the market (Schumpeter 1934 with the creative destruction theory, followed by Jovanovic 1982, Hopenhayn 1992, Melitz 2003, Melitz and Ottaviano 2008). In other words, less productive firms have lower chances of surviving than more efficient counterparts. More recently, however, several studies have shown that market imperfections can hurt productive firms disproportionally, preventing efficient firms from surviving disruptive events and leaving economies without their more productive businesses (Hallward-Driemeier and Rijkers 2013, Bosio, Djankov et al. 2020).

Earlier studies focused on planning-centric organizational resilience, which contends that having a formal plan to respond to crisis enhances resilience. In contrast, more recent studies have taken a capability-centric approach, which focus on the ability of a firm to adapt and recover successfully from adversity (resilience capacity) whether or not it has a plan.

Most of the literature defines organizational resilience as the ability to deal with challenging conditions by maintaining the functionality of a system when it is disrupted (Vogus and Sutcliffe 2007). The ISO 22316:2017 standard defines organizational resilience as: "the ability of an organization to absorb and adapt in a changing environment to enable it to deliver its objectives and to survive and prosper." ITIL 4 defines resilience as the "ability of an organization to anticipate, prepare for, respond to, and adapt to both incremental changes and sudden disruptions from an external perspective."

Fiskel (2006) and Hamel and Valikangas (2003) define business resilience as "the capacity for companies to survive, adapt and grow in the face of turbulent change." Fiskel states that resilient businesses can recover from disruptions and show adaptive capacity even in their business concept. Golan et al. (2020) state that "resilience implies a rapid and efficient response to minimize the consequences of disruptive events and maintaining or regaining a dynamically stable state".

Simeone (2015) points out that "business resilience enables organizations to adapt quickly to disruptions while maintaining sustainable business operations and protecting people, assets and overall brand equity". Dahles and Susilowati (2015) mention three different states of resilience: (i) returning to the previous state; (ii) recovering following an order of rescue; and (iii) returning to fundamentally different conditions that require new business models, operation methods, etc.

The BRI presented here intends to measure the ability of firms to adapt to disruptions based on features of firms and their context that reflect resilience capacity. However, to actually measure resilience outcome/performance, we might need to track firms over time and observe which ones were, indeed, able to recover from disruptions; that was not possible for this report. We assume that businesses that show higher resilience capacity are/were more likely to adapt and maintain business operations. In other words, a low resilience capacity would lead to a poor resilience performance. That is not always the case, however.

Aldianto et al. (2021) provide a conceptual framework for business resilience capacity based on the impact of technological and dynamic capabilities, knowledge stock and leadership agility as drivers of business resilience. The authors state that "through innovation, organizations can adapt to environmental changes and reduce the impact of threats and risks." Innovation ambidexterity, or the ability to balance exploration and exploitation, is a trait that enables organizations to become more resilient.[2]

According to Teece et al. (1997), dynamic capability includes the ability to reconfigure business competencies in a volatile environment. Sambamurthy and Zmud (1997) define technological capability as the ability to acquire, disseminate, combine and reconfigure technology resources to support and improve business strategies and processes. Moreover, Hamieda et al. (2018) show that flexible and adaptive leadership, encouraging continual feedback, learning and collaboration, is also needed to cope with challenges and disruptions.

2 For a detail explanation of innovation ambidexterity, see Cao et al. (2009).

Chapter 2

2. How to measure resilience?

This subsection provides an overview of the outcomes that might measure resilience, understood as preparedness, adaptation and recovery under disruptive conditions. Employment (ILO 2021), liquidity (Bosio, Djankov et al. 2020; Bosio, Jolevski et al. 2020) and productivity (ILO 2020) seem to be the variables that are severely affected in firms under disruptive situations. Other outcomes proposed in the literature as proxies for business resilience are the preservation of market share, performance after disruption, product depreciation, readiness,[3] expected cost of disruption, total direct losses, customer loss, delivery delay, fractional quantity loss and percentage of suppliers who lost capacity under disruption (Golan et al. 2020).

However, none of these variables by itself appropriately captures business resilience. This report contends that business resilience is, instead, a complex concept that cannot be adequately captured by a single proxy variable. The assumption is that business resilience is a latent characteristic that cannot be measured by a single variable but instead needs to be proxied by a combination of indicators selected from a theoretical perspective.[4]

For illustrative purposes, this report uses secondary data, but future efforts might require the collection of primary data at the firm level to include relevant indicators that are not available from secondary sources.

The construction of a more comprehensive business resilience measure requires the combination of several indicators into a single index. The first step in this effort is to clearly define the drivers and features that characterize business resilience capacity and to avoid the use of proxy indicators that do not adequately represent this concept.[5]

During the literature review, we did not find global-scale efforts by international organizations to build business resilience indices. However, there have been local efforts to capture business and organization resilience for small samples of firms and countries.[6]

For the construction of the BRI, we defined several subindices and, for each index, appropriate indicators. The selection of indicators was based both on a theoretical foundation to ensure that they capture resilience and on the availability of quality secondary data so that the index can be constructed periodically and for a large subset of countries. Relying on secondary data has its limitations, but we searched extensively for proxy indicators that capture resilience dimensions from several datasets provided by well-recognized multilateral and international organizations.[7]

During the selection process, we tried to avoid repetitive and irrelevant indicators that lack conceptual validity, but at the same time we tried to include sufficient indicators to capture as best possible all the dimensions of business resilience. The indicators used for the construction of the index come from

3 That is, how ready a firm is to face a disruption, classified into four stages: recognition, diagnosis, response development and response implementation. See Bode and Macdonald (2017).
4 For a discussion of the importance of firms' resilience, see Hamel and Valikangas (2003).
5 There are several methodological decisions that need to be made before building reliable composite indices. In the data transformation step, the developers of the index need to decide how to normalize/standardize the indicators and the techniques to deal with imputation of missing values, outliers, and other data transformation. During the weighting/aggregation step, the developers need to define the method to assign weights for each indicator and subindex and the method to aggregate them into the composite index.
6 Aldianto et al. (2021) collected semi-structured data through online interviews in Jakarta, Bogo, and Bandung to capture business resilience mainly for start-ups after COVID-19.
7 For more information see Caselli (2008) and https://warwick.ac.uk/fac/soc/pais/research/researchcentres/csgr/index/download/.

sources with different sample representativeness (for example, some indicators focused on the formal sector only), imputation and aggregation methods. Thus, precautions need to be taken when analysing the index scores.[8] Due to the limitations of secondary sources, conceptual validity is constrained by the availability of quality data with sufficient country coverage.

Considering the framework developed by Aldianto et al. (2021) we developed a framework to approximate business resilience.[9] The framework proposed by Aldianto et al. identifies five key drivers of business resilience: (i) dynamic capability, (ii) technology capability, (iii) agile leadership, (iv) knowledge stock and (v) innovation ambidexterity.[10] These are detailed in table 1.

▶ **Table 1.** Capabilities of business resilience

Capabilities	Explanation
Dynamic capability	Refers to the ability of businesses to reconfigure their internal and external resources and competencies to respond to changes in the business environment.
Technology capability	Refers to the ability of a firm to use and develop technology for product, process or management development to support business performance. Technology capabilities include technological infrastructure, human resources (consisting of technical and managerial skills) and intangibles such as knowledge assets, customer orientation and synergy (Bharadwaj 2000).
Agile leadership	Agile leadership is necessary to make firms more effective in adapting to changes. Agile leadership "describes the ability of a leader to be quick, adaptable and flexible in responding to unforeseen events in an unfamiliar circumstance" (Attar and Abdul-Kareem 2020).
Knowledge stock	Refers to the accumulated knowledge of a company and its human capital. This factor is related to innovation and learning. Rupietta and Backes-Gellner (2019) suggest an interesting categorization of types of knowledge: a) knowledge stock (that is, human capital) and knowledge flows (induced by human resource management practices).
Innovation ambidexterity	Innovation ambidexterity refers to the balance between exploration and exploitation. Jansen et al. (2005) contend that explorative and exploitative innovation are interdependent activities. Exploration is linked to the search for new opportunities, discovery and experimentation, while exploitation is "the refinement and expansion of existing competencies, technologies and paradigms with returns that are proximate and predictable" (March 1991).

Source: Summary based on Aldianto et al. (2021).

The framework proposed by Aldianto et al. (2021) serves as a general foundation for the selection of indicators. However, we also searched for practical methods to get a better sense of what indicators we should look for. The McKinsey consultancy firm developed a practical methodology to measure business resilience that is in agreement with the Aldianto proposal. As outlined in table 2, McKinsey (Fritz et al. 2021) considers six dimensions: financial, operational, technological, organizational, reputational and business model.

8 For example, some firm-level indicators that were selected in the index come from the World Bank Enterprise Surveys. The interpretation of the index must consider the sample limitations of these surveys. For example, the survey does not capture data from informal businesses and the agricultural sector, which might lead to lack of representativeness of the entire spectrum of country-level businesses resilience. In addition, the sample is collected only in the main cities and do not include micro-size firms with less than 5 employees. For more information, refer to https://www.enterprisesurveys.org/en/enterprisesurveys. For a detailed analysis of composite indicators and their weaknesses, see Freunderberg (2003) and Grupp and Mogee (2004).
9 The selection and grouping of indicators were both theory-grounded and empirically based (correlation analysis). The selection of indicators per subindex was supported by a correlation analysis to observe the association between variables within each subindex and subdimension.
10 For a detailed explanation of each driver, see Annex 1.

▶ **Table 2.** Business resilience dimensions – McKinsey consultancy

Dimensions	Explanation
Financial subindex	Refers to balance of short- and long-term financial management. The argument is that solid capital position and sufficient liquidity enable firms to manage drops in revenues, increases in costs or credit needs. Businesses must shield themselves from reduced access to capital, debt or equity.
Operational subindex	Refers to the production capacity of the organization to cope with both changes in demand and operational disruption without sacrificing quality and operational capacity. It is the ability of a firm to maintain operational capacity despite failures of individual suppliers or distributors, natural catastrophes and/or geopolitical events.
Technology subindex	Refers to strong, secure and flexible infrastructure to manage cyber threats and avoid technology breakdown. This index intends to measure the use, maintenance and implementation of infrastructure and IT/digital solutions to respond to continuously changing customer need and regulatory requirements.
Organizational subindex	Refers to the presence of a diverse and skilled workforce. It measures the ability of firms to recruit the best talent, develop talent equitably and upskill or reskill employees and the competencies to implement strong processes and a good working environment.
Reputational subindex	Refers to the existence of a strong and shared mission, values and purpose that guide the firm's actions, flexibility and openness in listening to and communicating with stakeholders, anticipating and addressing societal expectations, and responding to criticism of firm behaviour.
Business model subindex	Refers to the capacity to adapt to significant shifts in customer demand, the competitive landscape, technological changes and regulatory requirements – in other words, evolving business models capable of responding to disruptions.

Source: Fritz, Pancaldi et al. 2021.

One missing piece of these frameworks is the effect of macro-level factors on business resilience. The capabilities/dimensions identified are entirely at the firm level, but we know that business environment barriers vary across countries and might affect business resilience through different channels, such as: (i) quality of institutions, (ii) political and macroeconomic conditions, (iii) constraints in the financial sector; (iv) uncertainty and volatility, among others. Therefore, using secondary data, we included business environment indicators as an additional dimension for the construction of the BRI, which we called "institutions".

Reviewing the literature, we did not find specialized studies that assess the role of macro- and meso-level factors on business resilience. However, our intuition suggests that two average firms with the same adaptive ability but operating in countries with different business friendliness environments (for example, the United States of America and the Plurinational state of Bolivia) might exhibit different resilience outcomes.

Doern et al. (2019) examine the effects of crises on entrepreneurship. The study suggests that research should focus on examining differential impacts of and responses to international crises where other factors (beyond entrepreneurs or organizations) can play an important role in recovery from the crisis. The authors analysed several articles describing situations where external and regional factors might affect business performance and also the response of businesses to crisis. For example, in the study "Knowledge Diversity and Entrepreneurship Following an Economic Crisis: An Empirical Study of Regional Resilience in Great Britain", Bishop (2018) finds a positive relationship between regional factors and business creation and resilience in several parts of Great Britain after the 2007–2008 global financial crisis.

Other studies also show the importance of macroeconomic factors to business resilience. According to the report *The Impact of the Global Crisis on SME and Entrepreneurship Financing and Policy Responses* by the Organisation for Economic Co-operation and Development (OECD) (2009), "access to finance has been the most significant challenge for the creation, survival and growth of SMEs, especially innovative ones...", suggesting that survival depends not only on firms' characteristics but also on external factors beyond of their control, such as sectoral and/or macroeconomic factors.

Chapter

3

3. BRI construction methodology

For this analysis business resilience was disaggregated into subindices, or dimensions, and each subindex into several indicators. The BRI was constructed using secondary data from the World Bank Enterprise Survey Data, the Digital Adoption Index 2016, Doing Business Index 2019, Global Competitiveness Index 2019, Global Digital Readiness Index 2019 and Global Innovation Index (GII).[11]

Initially, 56 indicators were selected and classified in six dimensions: institutions, human capital, technology capability, financial status, innovation and management practices. The final selection of indicators considered a correlation analysis within and across dimensions to avoid overlapping indicators that do not add much insight. Due to the high correlation between the technology and innovation indicators, we decided to use only one dimension, "technology and innovation", and to drop the highly correlated indicators.

In the end the BRI was developed using 44 indicators (table 3), which collected data from 50 economies for 2019.[12] Considering the methodological analysis, we divided the BRI into five dimensions: *institutions,* with ten indicators; *human capital,* with eight indicators; *technology and innovation,* with 12 indicators; *financial status,* with six indicators; and *management practices,* with four indicators. The indicators used for the construction of the index a combination of "hard" and "perception" data. Actionable indicators, or hard data (national representative surveys or proxy variables), are based on direct measurement of the outcomes, while perception-based indicators come from opinion surveys (public opinion and expert assessment).[13]

Most of the selected indicators are updated yearly for many countries from reliable sources. The BRI country coverage is not as extensive as some other indices because we avoided working with partial data that could bias the results.[14] The selected indicators are produced periodically and without large time lags.

11 See Annex 2 for a detailed explanation of the indices.
12 Most of the data sources collect data from more than 100 economies. The limitation comes from the use of the World Bank Enterprise Survey, which collects the firm-level data needed to measure business resilience and covers just 44 economies.
13 The main criticisms found in the literature of the use of opinion survey data are: (i) actionable indicators are less prone to be influenced by changes in perception of individuals, while perception-based indicators depend on responses that could differ systematically due to variations in perceptions of the same phenomenon; (ii) responses may be influenced by others' assessments, resulting in correlated perception errors undermining the validity of the weighting scheme; (iii) the small sample sizes of opinion surveys could bias the results; and (iv) experts may not be able to adequately answer the diversity of questions asked in opinion surveys. However, the use of perception-based indicators is necessary to capture complex concepts that cannot be measured with hard data. Imprecision and biases do not disqualify the use of perception data but rather highlight the importance of selecting the right method to address them (Foa and Tanner 2012).
14 The Americas Quarterly, Multi Poverty Index, Technology Achievement Index, The Index of Economic Well-Being and the African Gender and Development Index of the UN Commission for Africa are examples of indices that suffer from replicability problems and imbalanced datasets that might affect cross-country comparability. For more information see Rippin (2011) and Osberg and Sharpe (2005).

Table 3. Business resilience indicators[15]

Dimensions	Subdimension	Indicators
Institutions	**Political**	The Political, Legal, Operational or Security Risk Index
		Government Effectiveness Index
	Regulatory	Regulatory Quality Index from the Worldwide Governance Indicators
		Rule of Law Index
		Cost of redundancy dismissal
	Business	Ease of starting business
		Ease of resolving insolvency
		Ease of protecting minority investors
		Flexibility of wage determination
		Applied tariff rate, weighted average, all products (%)
Human capital	**Knowledge stock and flows**	Tertiary enrolment, % gross
		Extent of staff training
		On-the-job training
		Charges for use of intellectual property, for example, payments (%, total trade, three-year average)
		Telecommunications, computer and information (ICT) services imports (% of total trade)
	Knowledge creation and impact	Number of resident patent applications filed at a given national or regional patent office (per billion PPP$ GDP)
		Number of scientific and technical journal articles (per billion PPP$ GDP)
		Total computer software spending (% of GDP)
Technology capability	**Research and development**	Gross expenditure on R&D (GERD) (% GDP)
		Average score of the top three universities according to the QS world university ranking
		The extent to which businesses and universities collaborate on R&D
	Infrastructure	Information and Communication Technologies Access Index from the GII calculations based on the World Telecommunication/ICT Indicators Database
		Information and Communication Technologies Use Index
		Quality of port infrastructure
		Quality of airport infrastructure
		Logistics Performance Index
	Innovation	Production process sophistication
		State of cluster development and depth
		Number of patent families filed in at least two offices (per billion PPP$ GDP)
		Capacity for innovation

15 See Annex 3 for a detailed explanation of the sources for each indicator.

Dimensions	Subdimension	Indicators
Financial	**Liquidity**	Financial obstacles
		Overdraft facility
	Funding	Ease of Getting Credit Index
		Trustworthiness and confidence
		Financing equity market
		Domestic credit to private sector (% of GDP)
Management		Management Index
		Efficient use of talent
		Extent to which ICTs enable new organizational model
		Capacity to attract and retain talent

The methodology proposed for construction of the BRI was based on an extensive literature review of the advantages and disadvantages of the most important methods and techniques used in the development of composite indices. As discussed in the next sections, we decided to use the Data Envelopment Analysis (DEA) for the development of the BRI. In addition, we used alternative methods to check the robustness of the composite indicator. Two weighting and aggregation options were considered: (i) no weights, with geometric averages, and (ii) weights based on principal component analysis with arithmetic aggregation. The following sections explain the process used to construct the BRI and the results of the robustness analysis.

▶ 3.1 Missing data and outliers

There are three methods for dealing with missing values: case deletion; univariate imputation and multivariate imputation.[16] Each of these methods has its own advantages and disadvantages. Case deletion drops variables that are incomplete from the full or partial set of unit observations, as explained by Foa and Tanner (2012). Case deletion restricts the composite index to a smaller set of variables and so might restrict the number of countries that can be analysed. Moreover, dropping non-random missing values might bias the index scores. The alternative is to impute missing values using univariate and/or multivariate statistical approaches.[17] However, imputation techniques are not immune to technical and political criticisms.

Considering that this study is a first effort to construct a BRI, we decided to base our approach on case deletion at the country level.[18] The countries that did not have all the indicators were excluded from the analysis.

16 See Annex 4 for a more detailed explanation.
17 Univariate feature imputation uses statistics such as the mean, median and mode of each indicator to replace the missing values. More sophisticated approaches estimate missing values as a function of other indicators using an iterative process or using other techniques such as the nearest neighbors' imputation algorithm, cross-country OLS, Expected Maximization Imputation, or multiple imputation techniques such as Markov Chain Monte Carlo, among others.
18 According to Foa and Tanner (2012), most international institutions prefer to drop observations rather than imputing values because use of the former protects them from criticisms.

However, in some cases, we used univariate imputation methods. The gross loan portfolio (percentage of GDP) and firms offering formal training (percentage of firms) were imputed using the mean values from their income group and region.[19]

For the analysis we constructed an initial database merging the Digital Adoption Index (183 countries), the Global Competitiveness Index 2017-2018 (152 countries), the Global Digital Readiness Index 2019 (141 countries) and the Global Innovation Index (GII) 2019 (129 countries). All of these datasets provide data at the country level.

After merging all the country-level datasets, we obtained a combined dataset of 168 observations. However, we did not have firm-level data, which are crucial for the BRI, and so we decided to use the firm-level Enterprise Survey datasets available for only 65 economies for the years 2017-2020. After observations with missing values in most variables were dropped, the final dataset consisted of 50 observations.

After that process 12 variables still exhibited a percentage of missing values ranging from 2–4 per cent of total observations (one or two observations per variable).[20] These variables were imputed using region and economic group mean values.[21] The summary statistics before and after imputation show practically no changes in the distribution (see Annex 6).

Finally, to control for outliers, the indicators were winsorized using the fifth and 95th percentiles. Around two per cent of observations were dropped. The average correlation between the variables before and after winsorization was 99 per cent, which suggests that it caused no significant changes in the distribution of variables.

The "hard data" indicators from the GII were treated according to the recommendations of the Joint Research Centre and the European Commission's Competence Centre on Composite Indicators and Scoreboards (JRC-COIN).[22] The outlier strategy for perception data from the Global Competitiveness Index involved the exclusion of observations where respondent gave the same answer to at least 80 per cent of the questions; surveys with a completion rate below 50 per cent; respondents who are not based in the same country as the Partner Institute; and respondents who do not have the required level of seniority.

▸ 3.2 Normalization and standardization

Several normalization and standardization methods were considered. The most common methods are z-score standardization and min-max goalposts, distance to a reference country, categorical scales, indicators above or below the means, cyclical indicators, balance of opinions, and percentage of annual differences over consecutive years.[23]

19 Outliers can be addressed by univariate or multivariate methods or a combination of both. The former identifies outliers looking at values in a single feature space, while multivariate methods are undertaken in an n-dimensional space.

20 Tertiary enrollment; gross expenditure on research and development; ICT access; ICT use; logistics performance; gross capital formation; applied tariff rate; patents filed; high tech imports; business–government investment; human capital; and startup environment.

21 Income groups: high income, low income, lower middle income, and upper middle income. Regions: East Asia and Pacific, Europe and Central Asia, Latin America and Caribbean, Middle East and North Africa, North America, South Asia, and sub-Saharan Africa.

22 In summary, the first step used by the developers of the GII was to determine problematic indicators using skewness (abs>2.25) and kurtosis (abs>3.5). Indicators with one to five outlier observations were winsorized up to the specified absolute values of the skewness and/or kurtosis limits. Indicators with five or more outliers and with skewness and kurtosis over the specified ranges were transformed using natural logarithms. For reference to a data quality assessment framework, see IMF (2012).

23 For a detailed review of advantages and disadvantages, see Annex 5. For a detailed explanation, see (Freudenberg, 2003).

After an extensive review, we decided to use the min-max goalpost method, which is one of the most widely used in the literature.[24] The method requires the definition of the minimum and maximum values to be used in the normalization formula to rescale the values to a range from 0 to 100.

$$score_{i,c} = \left(\frac{v_{i,c} - wp_i}{f_i - wp_i}\right) * 100$$

where $v_{i,c}$ is the unnormalized value for country c of indicator i; wpi is the worst performance indicator *i*, while f_i is the best possible outcome, or frontier.

For illustrative purposes, we used the minimum and maximum values of the entire dataset, which covers 47 economies. In the case of indicators where a higher value corresponds to a worse outcome, the normalized score was reoriented so that higher values reflect better outcomes. Firm-level indicators used in the index that came from the Enterprise Survey were scaled to a range from 0 to 100, and the aggregation to country-level scores was performed using a weighted average of all answers in each country using the following formula:

$$\sum_{\substack{i=1 \\ j=1,64}}^{N_j} \frac{f_{ij} * w_{ij} * x_{ij}}{N_j}$$

f_{ij} refers to the survey weights, w_{ij} to the number of permanent employees, x_{ij} to the indicators of observation i and country j; and N_j to the total number of employees per country.

▶ 3.3 Weighting

In this section we use equal weighting methods to estimate a baseline index and an analytical-based weighting method for the construction of the BRI. Comparison between the two might give us information about the possible biases of considering all variables to have the same importance.

For the baseline we used equal weighting for each level (from indicators to subindices and from subindices to the BRI). In other words, we summed all the values and divided by the total number of elements. The aggregation was performed using normalized variables after handling missing values and outliers.

Equal weighting is a simple method that has been extensively used by international organizations for the construction of composite indices.[25] However, it suffers from several problems that can affect the interpretation of the indices. One of the weaknesses is referred as the double counting issue, which is weighting an indicator higher than appropriate due to collinearity between indicators not addressed

24 The American Human Development Index (HDI) applies the min-max goalposts method to normalize the indicators to a scale of 0 to 10, while the Child Development Index (CDI), Ibrahim Index of African Governance (IIAG), the Social Inclusion Index (SII) and the Human Poverty Index for Developing Countries (HPI-I) normalize the indicators to a scale of 0 to 100 using observed minimum and maximum goalposts. The Technology Achievement Index indicators are transformed into a scale of 0 to 1.

25 For instance, the Commitment to Development Index (CDI), the Environmental Sustainability Index (SEI), Ibrahim Index of African Governance (IIAG), Summary Innovation Index (SII) and Technology Achievement Index (TAI) use an unweighted arithmetic average. The Human Development Index (HDI), the Inequality-adjusted HDI (IHDI), the Gender Inequality Index (GII) and other methods use unweighted geometric means. The Multidimensional Poverty Index (MPI) uses equal dimension weighting and equal indicator weighting within dimensions. Equal weighting can be used with standardized or normalized methods such as rankings, z-scores, min-max goalposts and distance to a reference, among others. For instance, in the Information and Communication Technologies and the Doing Business Index, in which the index is calculated as the average of country percentile rankings on each of ten topics, the ranking on each topic is the simple average of the percentile rankings of its component indicators.

during aggregation.[26] Moreover, most of the time a theoretical foundation to justify equal weighting is lacking.

To address these issues, we used analytical-based methods to set weights. Other methods can be used, such as participatory methods, but were not considered in the analysis.[27] Although participatory methods are not suitable for aggregating a large number of indicators (for example, more than ten), they could be used at the subindex level. In further efforts, participatory methods can be utilized to obtain a consensus about the relative importance of the six dimensions used to build the BRI, set minimum weights for subindices and take other decisions about weighting and aggregation methods.

The literature review found that the most common analytical methods used to set weights are Principal Component Analysis (PCA), Factor Analysis (FA), Cronbach coefficient alpha, cluster analysis, regression-based methods and other non-parametric methods such as Data Envelopment Analysis (DEA) and the Unobserved Component Method (UCM).

Each of these methods has its advantages and disadvantages. For example, the PCA/FA methods derive weights based on the correlation structure among variables. The resulting weights do not necessarily correspond to the actual linkages among indicators and might perform poorly if the largest variations in the indicators are not informative.[28] Moreover PCA/FA might be sensitive to outliers if they are not properly treated.[29] Regression-based methods need a well-defined variable that adequately represents the latent variable meant to be measured. The non-parametric methods might be difficult to explain to policymakers and sometime could be difficult to implement.

Regression-based methods were discarded because we do not have any good proxy of business resilience to use as the dependent variable. The UCM was a good candidate to determine the weight vector, but we did not want to depend on any assumption of the error terms during the optimization process.

After assessing advantages and disadvantages of different methods, we decided to use a combination of PCA and DEA–BoD (Benefit of the Doubt).[30] We used PCA as a dimensionality reduction algorithm to aggregate indicators into subindices and subindices into dimensions. The DEA–BoD was utilized to derive weights to aggregate dimensions into the composite index.[31]

The Development Envelope Analysis is a non-parametric approach that uses linear programming to set weights of indicators and subindices. The DEA defines a frontier (best practice) that serves as a benchmark and limit of what a country can achieve. The indicator's performance is then given by the relative distance between the actual observed performance and the nearest benchmark. Figure 1 shows the frontier constructed by points A, B and C and the distance D^T between the actual value D and the frontier. The DEA constructs an envelope for the observed indicator combinations for all countries and allows classification of countries into best-performing units (if they are at the frontier) or worse-performing units if they lie below the frontier.

26 There is always some positive correlation between variables that might affect weighting. The analyst should utilize a threshold to define what level of correlation could be deemed double counting (OECD 2008).

27 Participatory methods incorporate the opinions of various stakeholders to assign weights. This type of weighting method has received much criticism due to its subjectivity and insufficient clarity as to how the experts have set the weights. However, the literature presents several participatory methods that reduce subjectivity and make the weighting decision more transparent (for example, Budget Allocation Process and Analytic Hierarchy Process). These methods are not suitable when composite indices are constructed from a large number of indicators (for example, more than ten), as in the case of the BRI, because they lead to inconsistencies.

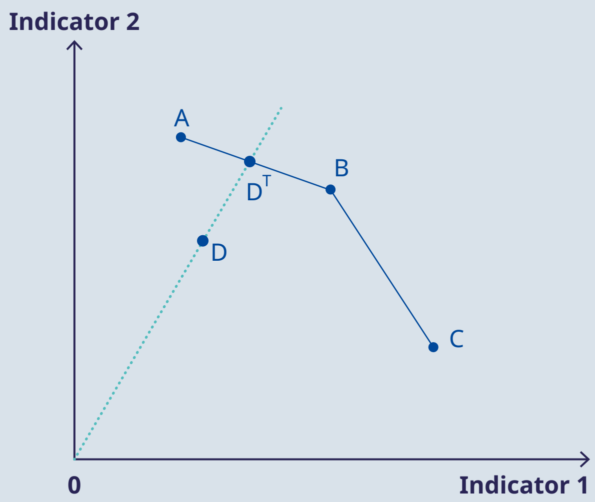
28 For more information see Saisana and Tarantola (2002) and Becker at al. (2017).

29 For a review of robust PCA that is less sensitive to outliers, see Hubert et al. (2005).

30 For a detailed comparison of weighting methods, see Annex 7.

31 See Andrews (2021) for the use PCA and DEA-BoD and Shen et al. (2012) and Tsaples and Papathanasiou (2020) for multi-layer BoD applications.

▶ **Figure 1.** DEA envelope: example of the construction of an index with two indicators



Source: Mahlberg and Obersteiner, 2001.

As mentioned, DEA faces mainly two limitations: the boundary problem and the differential weighting problem.

The boundary problem refers to the possibility of maximizing the objective function by setting weights of some variables to zero or one, which is not desirable. To address the boundary problem by adding weight restrictions to the model, we set upper and lower bounds (non-negative constraints and assurance regions) for each indicator, ensuring that no weight is either zero or one.[32] Thus, we avoid the exclusion of variables by allowing the method to set weights not equal to zero attached to any variable or setting the weight of only one variable to 1. To avoid dimensions with zero weights or contribution, we specified a minimum constraint of one per cent, and a maximum constraint of 30 per cent. Ideally, weight restrictions could be set using participatory methods to find agreement among experts on the minimum and maximum weights (Cherchye et al. 2008).

The problem of differential weighting refers to not having a common set of weights for all units, which prevents comparisons across units (Despostis 2005b). As explained, DEA derives an optimal set of non-unique weights for each observation in order to maximize each unit's performance or produce the most favourable score (Greco et al. 2019). These differential weights might reflect the fact that the dimensions' weights used for the construction of the BRI index are not homogeneous across units.

Although the overall ranking of the BRI produced with these optimal weights would still serve as an assessment of the relative efficiency of countries, it would not be possible to rank countries along a single scale. The results could generate multiple "champions" with similar scores.

To address differential weighting, we used a Multi Criteria Decision Analysis approach known as Benefit of the Doubt (BoD), which was originally used for evaluating macroeconomic performance. The BoD approach consists of creating a dummy input with the weighted indicators/dimensions as outputs.[33] The main difference between conventional DEA and the DEA used to construct composite indicators is that the latter focus on maximizing achievements without considering the inputs.

32 Allen et al. (1997) gives a good overview of DEA models with weight restrictions. In the literature this technique is known as "Type I Assurance Regions". For example, Hemans et al. (2008) use the responses of a panel of experts to assign weights to several indicators, using their opinions as binding constraints.
33 A number of techniques have been applied to arrive at a common weighting scheme, such as the Multi Criteria Decision Analysis – Data Envelopment Analysis (Hatefi and Torabi 2010). The Data Envelopment Analysis (DEA) has also been applied for the determination of HDI weights. Human Development Index (HDI) weights were estimated by Mahlberg and Obersteiner (2001), in Despostis (2005a).

As Cherchye et al. (2008) point out, there is no clear consensus in the literature on the objective of having common weights across units. The decision depends on the purpose of the study. When the objective is to compare scores across economies along a unique scale, we should address differential weighting. Otherwise, if a single ranking is not the main objective, the information provided by the DEA–BoD could be used to gain insights from benchmarking countries against optimal performance.

Considering the different information provided by addressing or not addressing differential weights, we decided to report both results and observe the difference in rankings. Additionally, as an individual benchmark between optimal and common-set weights, the standardized weights of each country were compared with the scores produced by their optimal weights.

The BoD method used in this report is based on Cherchye, Moesen and Rogge (2008), which is equivalent to the original DEA input-oriented method with all indicators considered as outputs and a "dummy input" equal to one for all decision units, weights $w_d = (d{=}1, \ldots, D)$ being the variables of the model:

$$CI_o = Max \sum_{d=1}^{D} w_d y_{dj0}$$

s.t.

$$\sum_{d=1}^{D} w_d y_{dj} \leq 1 \;\; j = 1, \ldots, n$$

$$\sum_{d=1}^{D} w_{dj} = 1 \;\; j = 1, \ldots, n$$

$$u_d \geq 0 \;\; r = 1, \ldots, D$$

The rationale of the BoD method relies on the fact that a priori we do not know the criteria to set weights for each decision-making unit (DMUs). The BoD solves this problem by letting each decision unit select weights that optimize the output value(s) such that, when DMU1 under assessment gets a higher value for the composite indicator than other DMUs, the former outperforms the others.

$$CI_o = Min \sum_{r=1}^{s} z_j$$

s.t.

$$\sum_{r=1}^{s} u_r y_{rj} + z_j = CI_0 \;\; j = 1, \ldots, n$$

$$u_r \geq 0 \;\; z_j \geq 0$$

As an illustration, the weights optimization was specified for the six subindices of the BRI. The BoD method can also be applied to set weights at lower levels, such as subdimensions and indicators, as shown above in the linear programming specification. As recommended in the literature, we used a linear programming model that minimizes the deviations (Z) in relation to the composite indicator (CI_0) obtained in the previous programming model, as in Baptista Teixeira de Morais (2011), to derive a common set of weights and, thus, to be able to compare the scores across economies.

▶ 3.4 Aggregation

According to the OECD handbook (2008) on constructing composite indices, aggregation methods are divided into two distinct categories: compensatory (linear and geometric) and non-compensatory. As for the weighting methods, there is no "perfect aggregation" scheme, as Arrow and Raynaud (1986) explain. Ideally, the selection of the aggregation method should allow for partial and/or non-compensatory effects, should consider overlapping inequalities in the distributions and should include individual preferences across dimensions.[34]

Compensatory aggregation can be understood as an additive, utility-based approach that assumes total or partial trade-offs between indicators, as pointed out in OECD (2008). The most common compensatory methods are (un)weighted arithmetic and geometric means. The (un)weighted arithmetic mean allows perfect substitutability, while the (un)weighted geometric mean aggregation allows imperfect substitutability between dimensions.[35]

Several well-known indices use arithmetic mean aggregation despite its problems.[36] However, other well-known indices, such as the Human Development Index (HDI), the Inequality-adjusted Human Development Index (IHDI) and the Gender Inequality Index, decided to change from arithmetic to geometric average aggregation to allow for only partial substitutability.

The use of geometric averages comes with its own limitations. One of the most relevant is the presence of zero values.[37] The treatment of zero and negative values might likely influence the results of the composite index and rankings, making the results less robust. For example, sensitivity analysis of the IHDI reveals that rankings are, indeed, affected by different choices of outlier replacement techniques.[38]

Geometric average aggregation does not solve the compensatory problem. Non-compensatory approaches were introduced in the literature to deal with substitutability.[39] However, they are not widely used by international organizations for the construction of composite indices. In fact, we did not find any well-known index that uses non-compensatory methods for the aggregation process, probably due to difficulties during implementation, but, most importantly, due to problems in explaining and interpretating these methods to policymakers.[40]

To avoid the zero-imputation issue of the geometric average and the complexity of the non-compensatory methods, we decided to use arithmetic average mean for the construction of the baseline index. The decision also took into account that most of the indicators selected from the secondary data are already normalized, using min-max goalposts, to the range from zero to one, which complicates the use of geometric averages.

34 For an application of the use of non-compensatory methods to the construction of composite indices, see Mazziota and Pareto (2016).
35 Perfect substitutability refers, for example, to a low achievement in one dimension that is linearly compensated for by a higher achievement in another dimension.
36 For example, the Human Poverty Index for Developing Countries (HPI-I) and the Global Competitiveness Index are based on successive aggregations of scores, from the indicator level to the overall index, by taking the arithmetic means at each level.
37 The HDI estimates the missing values using cross-country regression models. Other indices replace zeros with the minimum non-zero value of the distribution or add one to all the values, among other techniques.
38 Sensitivity analysis of the IHDI is described in Kovacevic and García Aguna (2010).
39 See Greco et al. (2019) for a more detailed explanation.
40 See Munda and Nardo (2009) for more information. Munda (2012) applied non-compensatory methods in the case of the Environmental Sustainability Index. The author found that there were significant differences in ranking positions (mainly in the middle of the distribution) when linear and non-compensatory methods are compared.

For the construction of the BRI, we used PCA as a data dimensionality reduction algorithm to derive the scores of subindices and dimensions. This approach increases robustness to collinearity and reduces the effects of double counting. PCA methods set weights based on the correlation structure underlying the variables. The largest weights are assigned to the indicators with the largest variation, which are the ones that explain differences in performance. PCA avoids double-counting underweighting those factors that depend on the same underlying factor (for example, weights are corrected for overlapping information between correlated indicators), which addresses one of the main weaknesses of the equal weighting method.

Due to the great amount of categorical and ordinal data, we used a variation of the PCA method deemed the Polychoric PCA. This method does not assume a linear relationship, and it can handle categorical and/or ordinal data, using a latent normally distributed variable that underlies the ordinal categorical variable.[41] We used the first two principal components to estimate the scores, which were averaged using as weights the proportion of the variance explained by each component.

The principal components were calculated using the resulting correlation matrix. The criteria to select the components were based on explaining about 70 per cent of the variance and with eigenvalues higher or equal to one. The variables were previously standardized using the min-max goalposts method.

41 For more information see Greyling and Tregenna (2017).

Chapter 4

4. Results

The frontier (score equal to one) is the maximum value obtained by a country by selecting the weights that optimize the output. With the first maximization equation, we obtained optimal weights for each economy that maximize the output score, considering actual values of each dimension (Table 4). The second maximization equation derives equal weights for all the economies, considering the restrictions mentioned in the previous section. Equal weights are used to compare countries along a unique scale assuming same preferences over dimensions for all countries while optimal weights reflect the maximum value that can be achieved by each country (see annexes 8 and 9).

Table 4. Equal weights results

Dimension	Weights
Institutions	0.30
Human capital	0.10
Technology and innovation	0.20
Financial	0.10
Management	0.30

Table 4 shows the equal weights for all countries. The results show that the institutional and management dimensions exhibit the highest importance in the final composite index, followed by technology and innovation. These weights are used to calculate the DEA scores (normalized from 0 to 1).

A score of zero means that the country exhibits the worst performance in that dimension, while a score of one is attributed to the country with the best performance (Annex 9 shows the scores by dimension for each country).

To understand commonalties across countries, we created quintiles using the optimal DEA score of all countries considered in the study. Table 5 explains the grouping and the percentage of countries in each group.

Table 5. Grouping of countries by quintiles

Quintile 1: Exceptional Resilience	This group represents the top 10% of countries with the highest resilience scores. These countries exhibit exceptional resilience and can serve as benchmarks for best practices and success stories.
Quintile 2: High Resilience	This group represents the next 20% of countries, falling between the 70th and 90th percentile. These countries demonstrate a high level of resilience and have strong capabilities to withstand and recover from various challenges.
Quintile 3: Moderate Resilience	This group represents the middle 40% of countries, ranging from the 30th to the 70th percentile. These countries have average or moderate resilience scores, indicating a balanced level of preparedness and adaptability.
Quintile 4: Low Resilience	This group represents the next 20% of countries, falling between the 10th and 30th percentile These countries face challenges and have lower resilience scores, indicating a need for improvement in building resilience capacities.
Quintile 5: Critical Resilience	This cluster represents the bottom 10% of countries with the poorest resilience scores. These countries face significant challenges and require urgent attention and support to enhance their resilience capacities.

The quintiles were utilized to create a worldwide and continental heatmaps using the official World Wank country boundaries shapefiles updated on March 19, 2020. Figure 2 shows that **most of the countries that exhibit exceptional resilience are in Europe while the majority that exhibit low to critical resilience are in Africa and South America.**

▶ **Figure 2.** World Resilience Heatmap overall and by continent

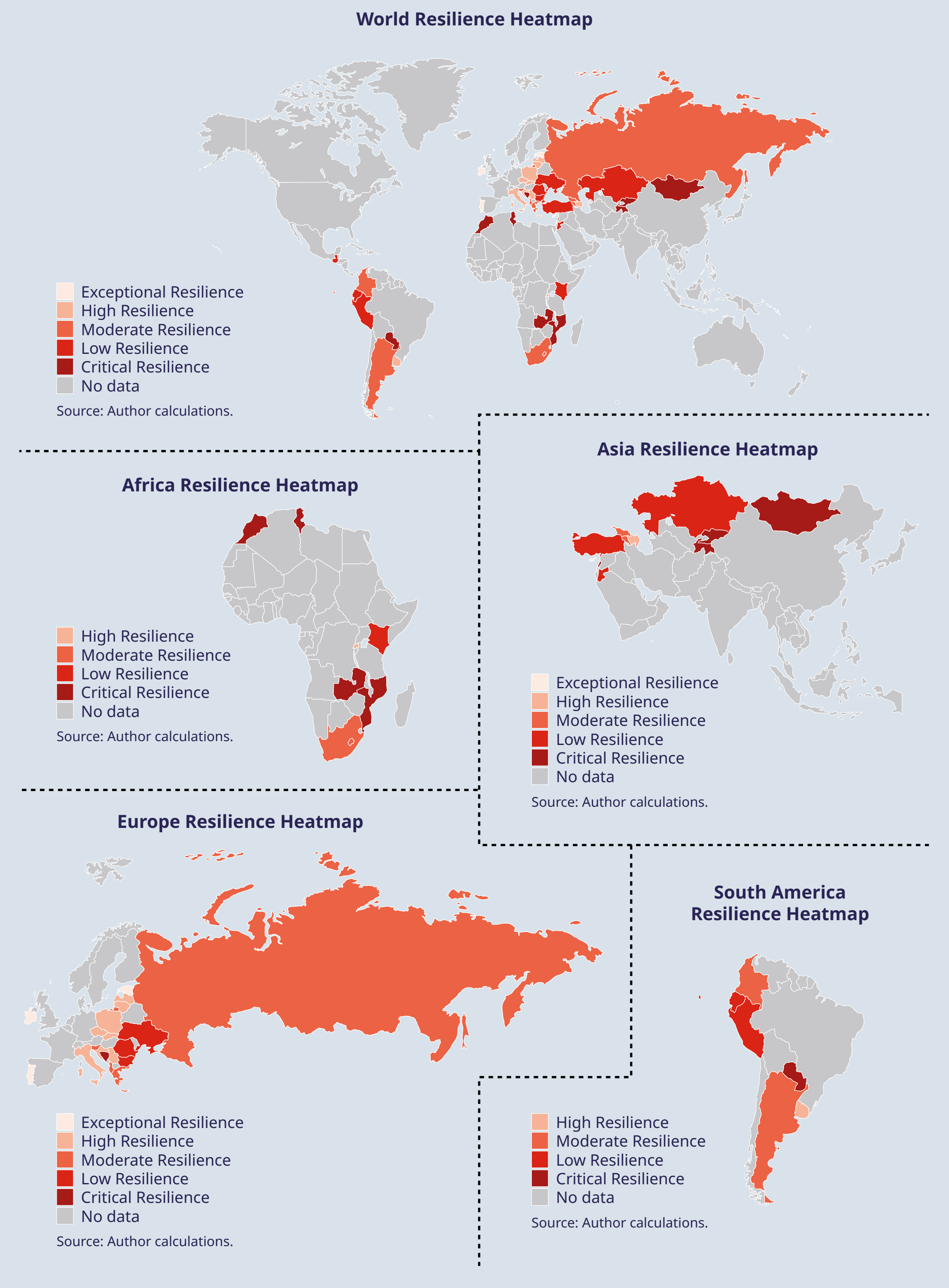
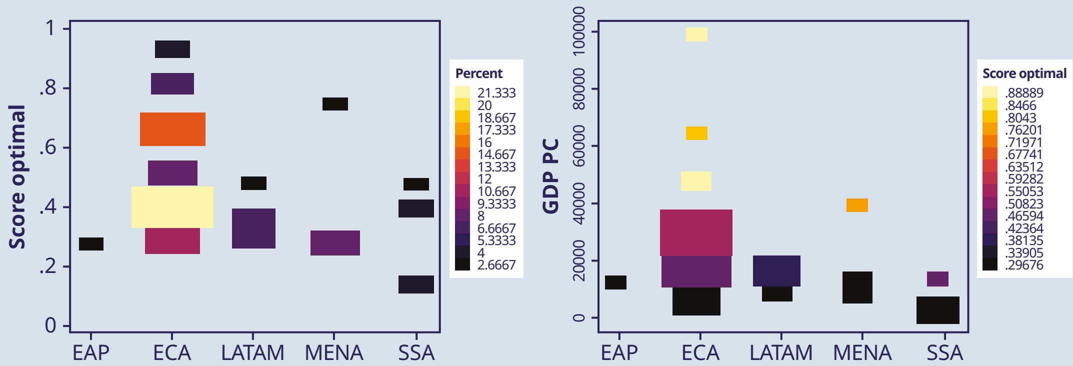


Figure 3 shows a heatmap of the optimal DEA score by region.[42] The highest frequency of countries across regions exhibits a score between 0,3 and 0,5 (left plot) and the majority with resilience over the media are in Europe and Central Asia (ECA). Finally, **we observe that GDP per capita is positively correlated with the resilience score** (right plot).

Figure 3. Heatplot Score Optimal by region and GDP per capita



Source: Author calculations and World Bank 2020.

In addition, to further understand commonalties across dimension and countries, we created five clusters or groups of similar countries, using the K-means algorithm. The grouping was obtained using as input the scores of the five dimensions for each country.

The selection of the clusters' number was derived from the Elbow Curve Analysis, specifically the elbow point, which measures the point where the improvement in clustering performance starts to diminish, in other words, the addition of more clusters does not significantly reduce the within-cluster sum of squares (inertia).[43]

Table 6 summarizes the composition of each cluster. Luxembourg, Estonia, Malta and Montenegro exhibit the highest objective function values and are part of cluster 2. The lowest values are exhibited by Bosnia and Herzegovina, Ecuador, Guatemala, Kyrgyz Republic, Lebanon, Mongolia, Mozambique, Paraguay, Tajikistan, Zambia, which are part of cluster 5.

Table 6. Clusters of countries

Clusters	Countries
1	Belgium, Czech Republic, Ireland, Italy, Netherlands, Portugal, Slovenia
2	Cyprus, Estonia, Luxembourg, Malta, Montenegro
3	Argentina, Armenia, Bulgaria, Colombia, Georgia, Greece, Hungary, Jordan, Kazakhstan, Kenya, Morocco, Poland, Peru, Romania, Russian Federation, South Africa, Tunisia, Turkey, Ukraine
4	Albania, Azerbaijan, Croatia, Latvia, Lithuania, Rwanda, Serbia, Slovak Republic, Uruguay
5	Bosnia and Herzegovina, Ecuador, Guatemala, Kyrgyz Republic, Lebanon, Mongolia, Mozambique, Paraguay,Tajikistan, Zambia

42 (EAP-> East Asia & Pacific; ECA->Europe & Central Asia; LAC-> Latin America and the Caribbean; MENA-> Middle East & North Africa; SSA->Sub-Saharan Africa).
43 The number of clusters, derived from the Elbow curve analysis, was set at five. For reference see https://www.analyticsvidhya.com/blog/2021/01/in-depth-intuition-of-k-means-clustering-algorithm-in-machine-learning/.

Table 7 shows the average scores per cluster and dimension. Clusters 1 and 2 exhibit the highest average scores, and they are composed of the best-performing economies, while cluster 5 has the lowest average scores.

▶ **Table 7.** Average scores per cluster and dimension

Cluster	Institutions	Human capital	Technology	Finance	Management
1	0.86	0.72	0.76	0.22	0.51
2	0.70	0.59	0.49	0.84	0.80
3	0.47	0.32	0.33	0.10	0.22
4	0.60	0.31	0.32	0.41	0.49
5	0.15	0.12	0.12	0.36	0.21

Tailored policy measures are essential for addressing the unique concerns of each cluster. For instance, countries belonging to clusters 1 and 3 exhibit relatively lower average scores in finance and management. On the other hand, cluster 5 demonstrates subpar performance across multiple dimensions, with human capital and technology emerging as the weakest areas. These shortcomings can be attributed to the selected indicators that reflect the resilience of businesses within each thematic area in such countries.

Chapter 5

5. Robustness analysis

This section presents uncertainty and sensitivity analyses. The uncertainty refers to the changes that are observed in the outcome (the composite index value) when different choices are made in the "inputs" (parameters and techniques chosen in the development steps). The sensitivity analysis measures how much of the variance in the overall output is attributable to those uncertainties.

Both analyses are needed to have a complete picture of the robustness of the composite index.[44] Naïve and/or incompatible choices could result in meaningless scores and a ranking that is not suitable for policy purposes. Neglecting the limitations of weighting and aggregation techniques might distort the primary intention of the index, which is to represent accurately the factors that influence the resilience of businesses.

The missing value imputation and outlier treatment was performed at the country level and affected less than two per cent of the data. Therefore, we did not include changes in the imputation and outlier methods in the robustness analysis.

For the robustness analysis, we first constructed a baseline using unweighted arithmetic averages to aggregate all levels into the composite index. Second, we changed the PCA used to aggregate indicators into subindices to average arithmetic means with equal weights to assess the change in the relative importance of dimensions and also to observe the effects of collinearity.

For the DEA–BoD we changed the maximum and minimum boundaries to assess resulting changes in the relative weights of each economy. In addition, we changed the linear programming specification to use the Despostis (2005b) method to assess how these changes affect the overall index when equal weights are applied to all economies (second maximization specification).[45]

▶ 5.1 Baseline

The results of the baseline analysis show that Netherlands, Belgium, Luxembourg and Ireland are the countries at the top, and Kyrgyz Republic, Bosnia and Herzegovina, Paraguay and Mozambique are at the bottom of the index (table 8).

▶ **Table 8.** Baseline scores

Country	Region	Income group	Baseline score
Netherlands	Europe & Central Asia	High income	1.00
Belgium	Europe & Central Asia	High income	0.90
Luxembourg	Europe & Central Asia	High income	0.87
Ireland	Europe & Central Asia	High income	0.80
Estonia	Europe & Central Asia	High income	0.78
Czech Republic	Europe & Central Asia	High income	0.67
Malta	Middle East & North Africa	High income	0.64
Portugal	Europe & Central Asia	High income	0.61
Cyprus	Europe & Central Asia	High income	0.56

44 See Saisana et al. (2005) for an example using the Environmental Sustainability Index and Permanyer (2011) for an example of the UNDP Human Development Index, the Gender-related Development Index, and the Human Poverty Index.
45 See Cherchye et al. (2008) for composite indicators with DEA and Robustness Analysis.

Country	Region	Income group	Baseline score
Slovenia	Europe & Central Asia	High income	0.53
Latvia	Europe & Central Asia	High income	0.52
Lithuania	Europe & Central Asia	High income	0.51
Poland	Europe & Central Asia	High income	0.49
Azerbaijan	Europe & Central Asia	Upper middle income	0.48
Montenegro	Europe & Central Asia	Upper middle income	0.48
Slovak Republic	Europe & Central Asia	High income	0.47
Rwanda	Sub-Saharan Africa	Low income	0.47
Italy	Europe & Central Asia	High income	0.46
South Africa	Sub-Saharan Africa	Upper middle income	0.46
Hungary	Europe & Central Asia	High income	0.44
Kenya	Sub-Saharan Africa	Lower middle income	0.42
Colombia	Latin America & Caribbean	Upper middle income	0.39
Turkey	Europe & Central Asia	Upper middle income	0.39
Russian Federation	Europe & Central Asia	Upper middle income	0.37
Uruguay	Latin America & Caribbean	High income	0.37
Bulgaria	Europe & Central Asia	Upper middle income	0.34
Georgia	Europe & Central Asia	Upper middle income	0.34
Armenia	Europe & Central Asia	Upper middle income	0.34
Jordan	Middle East & North Africa	Upper middle income	0.33
Greece	Europe & Central Asia	High income	0.32
Serbia	Europe & Central Asia	Upper middle income	0.30
Croatia	Europe & Central Asia	High income	0.29
Peru	Latin America & Caribbean	Upper middle income	0.29
Romania	Europe & Central Asia	Upper middle income	0.28
Albania	Europe & Central Asia	Upper middle income	0.27
Kazakhstan	Europe & Central Asia	Upper middle income	0.26
Morocco	Middle East & North Africa	Lower middle income	0.25
Tunisia	Middle East & North Africa	Lower middle income	0.24
Argentina	Latin America & Caribbean	Upper middle income	0.23
Ukraine	Europe & Central Asia	Lower middle income	0.23
Guatemala	Latin America & Caribbean	Upper middle income	0.22
Mongolia	East Asia & Pacific	Lower middle income	0.21
Zambia	Sub-Saharan Africa	Lower middle income	0.21
Lebanon	Middle East & North Africa	Upper middle income	0.17
Tajikistan	Europe & Central Asia	Lower middle income	0.13
Ecuador	Latin America & Caribbean	Upper middle income	0.11
Kyrgyz Republic	Europe & Central Asia	Lower middle income	0.10
Bosnia and Herzegovina	Europe & Central Asia	Upper middle income	0.10
Paraguay	Latin America & Caribbean	Upper middle income	0.09
Mozambique	Sub-Saharan Africa	Low income	0.00

This section presents uncertainty and sensitivity analyses. The uncertainty refers to the changes that are observed in the outcome (the composite index value) when different choices are made in the "inputs" (parameters and techniques chosen in the development steps). The sensitivity analysis measures how much of the variance in the overall output is attributable to those uncertainties.

As expected, the countries' scores change depending on the weighting and aggregation methods. However, these differences do not radically alter the relative positions of countries.

5.2 Arithmetic mean aggregation with equal weights + DEA

As mentioned, the PCA methods set weights based on the correlations underlying the structure among variables. This helps deal with double-counting underweighting factors that depend on the same underlying factor, which is one of the main weaknesses of the equal weighting method.

For comparison, we used equal weight arithmetic mean aggregation to estimate subindices from indicators and dimensions from subindices. For the final step, which is calculating the BRI, we used the DEA–BoD with equal weights.

Table 9. DEA–BoD weights comparison

	Dimensions weights using PCA				
Equal weights	w1	w2	w3	w4	w5
w1	0.92				
w2		0.59			
w3			0.58		
w4				0.21	
w5					0.78

Table 9 indicates the correlation of weights derived from the PCA and average arithmetic mean method for the institutions (w1), human capital (w2), technology (w3), financial (w4) and management (w5) dimensions. The correlation between the weight matrix used to build the BRI (PCA + DEA) and the one constructed using arithmetic mean aggregation with equal weights is strong for some dimensions (institutions and management) and weak for others (financial). Table 10 compares the score rankings that result from the two weighting methods.

Table 10. Relative positions of countries compared, based on weighting method (order from low risk to high risk)

PCA+DEA	Arithmetic mean + DEA
Luxembourg	Netherlands
Netherlands	Luxembourg
Belgium	Belgium
Ireland	Estonia
Estonia	Ireland
Malta	Malta
Cyprus	Czech Republic
Slovenia	Rwanda
Azerbaijan	Azerbaijan
Portugal	Cyprus
Czech Republic	Portugal
Montenegro	Latvia
Lithuania	Montenegro
Latvia	Kenya
Italy	South Africa
Slovak Republic	Lithuania
Uruguay	Slovak Republic
Rwanda	Colombia
Croatia	Poland
Poland	Hungary
Hungary	Slovenia
Serbia	Russian Federation
South Africa	Italy
Albania	Georgia
Colombia	Turkey
Georgia	Bulgaria
Greece	Armenia
Armenia	Peru
Bulgaria	Jordan
Kenya	Uruguay
Kazakhstan	Guatemala
Argentina	Romania
Russian Federation	Albania
Jordan	Zambia
Peru	Morocco
Romania	Serbia
Turkey	Kazakhstan
Tunisia	Croatia
Mongolia	Mongolia

PCA+DEA	Arithmetic mean + DEA
Tajikistan	Greece
Guatemala	Tunisia
Morocco	Ukraine
Kyrgyz Republic	Lebanon
Ukraine	Argentina
Bosnia and Herzegovina	Tajikistan
Ecuador	Paraguay
Paraguay	Kyrgyz Republic
Zambia	Bosnia and Herzegovina
Lebanon	Ecuador
Mozambique	Mozambique

We observe some changes in the positions of the economies, mainly in the middle of the distribution, and less at the top and bottom of the distribution. In the middle of the distribution, we find that some countries, such as Slovenia, Italy, Uruguay, Serbia and Croatia, among others, gain several positions when the arithmetic mean is applied instead of PCA, while others, such as Rwanda, South Africa and Kenya, lost several positions in the ranking.

▶ **Table 11.** Weights comparison

Dimension	Weights PCA/DEA	Weights AV/DEA
Institutions	0.30	0.13
Human capital	0.10	0.10
Technology	0.20	0.17
Financial	0.10	0.30
Management	0.30	0.30

Table 11 shows the weights derived from the two different methods. The preferred first method, PCA + DEA, groups the indicators into subindices and subindices into dimensions using PCA, while the second uses unweighted arithmetic means. The former gives much more weight to institutions and less to the financial dimensions, while the latter give more importance to the financial dimension. The differences might be associated with the correlation between variables (double counting problem) not addressed by the arithmetic average.

▶ 5.3 DEA uncertainty analysis

$$Min\ t * \left(\frac{1}{47}\right) * \sum_{r=1}^{s} d_j + (1-t) * z$$

$$\text{s.t.}$$

$$\sum_{r=1}^{s} (u_r y_{rj} + d_j) = CI_0 \quad j = 1, \dots, n$$

$$u_r \geq 0 \quad d_j \geq 0$$

$$d_j - z \leq 0$$

For the DEA uncertainty analysis, we used the method proposed by Despostis (2002)[46] to better discriminate between efficient units. The model above estimates common weights minimizing the distance between the DEA scores (optimal weights) and the global scores (equal weights) $Min\ t * \left(\frac{1}{47}\right) * \sum_{r=1}^{s} d_j + (1\text{–}t) * z$. The distance is measured by two norms given by d_j and z. The coefficient *t* can vary from 0 to 1 providing a different set of weights for each dimension. When $t=1$, the linear programming function assumes L1 norm, or mean deviation, between DEA scores and global scores; when $t=0$, the function assumes an infinite norm.

▶ **Table 12.** Weights deviation for different t-values

t values	w1	w2	w3	w4	w5
0.0	0.24	0.30	0.12	0.12	0.22
0.2	0.30	0.25	0.13	0.14	0.18
0.4	0.30	0.25	0.13	0.134	0.18
0.6	0.30	0.10	0.21	0.09	0.30
0.8	0.30	0.10	0.21	0.09	0.30
1.0	0.30	0.08	0.30	0.02	0.30
Standard deviation	0.02	0.10	0.07	0.04	0.06

To obtain different set of weights, the programming model was optimized for different sets of *t* values (table 13). The results show the largest variations in the weights for the human capital (w2) and technology (w3) subindices. However, the ranking outcomes remain similar across different specifications.

46 (Despostis D. , 2002).

Conclusions

Conclusions

The literature review did not identify efforts by well-recognized institutions to build and periodically update a business resilience index for a high proportion of the world's countries. This study seeks to illustrate the steps necessary to construct such an index as well as the limitations of data and country coverage.

One of the main obstacles to construction of the BRI was the limited availability of quality secondary data with good country coverage that capture the diverse dimensions of business resilience. Collection of primary data was not possible for this study. Secondary data provide only a proxy of business resilience, and much effort needs to be made to collect primary data to better capture it.

The BRI does not capture resilience outcome but rather resilience capacity. The indicators do now show if a business survived a shock or not (resilience outcome). However, the BRI indicates the resilience capacity of businesses as it depends on internal and macroeconomic/institutional characteristics.

There are plenty of methods and techniques to deal with the most important steps in the construction of a composite index, and all of them have their own advantages and disadvantages. There are no perfect methods, and each of them has been the subject of criticisms. Developers of composite indices must be aware of their limitations and how the use of different methods affects the final scores.

The results derived from this analysis show that the relative positions of countries could change depending on the weighting and aggregation methods selected to construct the index. After an extensive review, we selected the PCA to aggregate from indicators to subindices and subindices to dimensions and DEA–BoD to aggregate from dimensions to the composite index. For the analysis, we divided the countries into clusters using the scores obtained for each dimension. To strengthen firms' resilience capacity, each cluster might need tailored policy measures that respond to specific concerns.

The results derived from this analysis show that the relative positions of countries could change depending on the weighting and aggregation methods selected to construct the index.

The COVID-19 outbreak brought to light the importance of business resilience. The pandemic affected productive and non-productive, large, small and medium-sized enterprises, especially in sectors that could not easily adapt to remote work and online, contactless provision of services. Future shocks are not unlikely, and firms need to be prepared to respond and adapt appropriately if they want to survive and maintain operations. Moreover, policymakers need to be aware of firms' deficiencies and limitation so as to focus their effort on the factors that could most increase the probability of adaptation and survival.

Aldianto et al. (2021) proposed the conceptual model of business resilience that was used as a basis for the construction of the BRI. The authors contend that business resilience depends on the ability of companies (which, in turn, depends on their knowledge stock) to adapt internal and external resources in response to changing market needs; use and develop technologies to support business strategies and processes and create new opportunities; and exploit existing competencies and explore new opportunities to drive innovation.

For the authors, business resilience is strongly associated with organizational and managerial capabilities, but they do not consider the context in which companies operate and what governments can do to support business efforts. The BRI includes several variables and dimensions (for example, institutions and finance) that incorporate the business environment in which business operate.

In this context governments should prioritize both policies that facilitate organizational and managerial capabilities and policies to improve the business environment in which companies operate. For example, governments should focus R&D expenditure and strengthen businesses–university collaboration. More importantly, governments need to ensure the provision of core technological and physical infrastructure throughout the country. In addition, governments should work on developing their financial institutions to ease access to credit for investment and to cover working capital needs as well as ensure the provision of a quality education system (secondary and tertiary) attuned to market needs.

Considering that all these factors are interconnected, governments should tackle them in a coordinated manner. Our results show that, overall, the finance and technology dimensions are the ones that exhibit the lowest average values across clusters and, thus, are those that need more attention. However, countries need to identify where the weakest areas are to focus their efforts on them especially if they operate with limited fiscal space. Specifically, countries in clusters 1 and 3 should focus on improving the financial dimension, which mainly includes easing access to credit to fund business operations and new investments and developing financing equity markets, among other actions. Countries in clusters 2 and 4 need to focus their efforts on improving access to technology and foster human capital to have an adequately skilled workforce. Meanwhile, countries in cluster 5 need to enhance performance across all dimensions, paying particular attention to human capital and institutions, that is, by promoting skills development according to market needs and creating a conducive environment for sustainable enterprise development.

Government can help with policies to support staff and worker training, improve the efficient use of talent, provide a high-quality education system that respond to the market needs and help businesses with their capacity to attract and retain talent.

As mentioned, this study has several limitations. The BRI does not capture resilience outcomes and assumes that outcomes are highly correlated with resilience capacity. Moreover, the BRI relied on secondary data, mainly at the national level. The collection of primary data would be needed to properly capture organizational and managerial capabilities (for example, agile leadership and dynamic capabilities as well as technological capabilities within businesses) that might have a significant effect on business resilience. These efforts could be the basis for longitudinal studies and further research to investigate the effects of these factors on the resilience outcome of businesses before and after facing critical conditions.

References

Abdul-Kareem, Aleem. (2020). The Role of Agile Leadership in Organisational Agility. 10.1108/978-1-80043-380-920201011.

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

Afuah, Allan. "Mapping Technological Capabilities into Product Markets and Competitive Advantage: The Case of Cholesterol Drugs." *Strategic Management Journal* 23, no. 2 (2002): 171–79. http://www.jstor.org/stable/3094440.

Aldianto, L., G. Anggadwita, A. Permatasari, I. Mirzanti, and I. O. Williamson. 2021. "Toward a Business Resilience Framework for Startups." *Sustainability* 13: 3132.

Allen, R., A. Athanassopoulos, R. Dyson, et al. 1997. "Weights Restrictions and Value Judgements in Data Envelopment Analysis: Evolution, Development and Future Directions". *Annals of Operations Research* 73: 13–34. https://doi.org/10.1023/A:1018968909638.

Andrews, A. 2021. "An Application of PCA-DEA with the Double-Bootstrap Approach to Estimate the Technical Efficiency of New Zealand District Health Boards." *Health Economics, Policy and Law* 1-25.

Arrow, K.J. and H. Raynaud. 1986. *Social Choice and Multicriterion Decision-making.* Cambridge: MIT Press.

Attar, M. and A Abdul-Kareem. 2020. "The Role of Agile Leadership in Organisational Agility." In *Agile Leadership Methods for Indusry 4.0*, edited by Bulent Akkya, 171-191. Bingley, UK: Emerald Publishing Limited.

Baptista Teixeira de Morais, P. A. 2011. *Evaluation of Performance of European Cities with the Aim of Increasing Quality of Life.* Porto: Faculdade de Engenharia da Universidade do Porto.

Becker, W., M. Saisana, P. Paruolo, and I. Vandesteele. 2017. "Weights and Importance in Composite Indicators: Closing the Gap. *Ecological Indicators* 12-22.

Bharadwaj, A. S. 2000. "A Resource-Based Perspective on Information Technology Capability and Firm Performance: An Empirical Investigation." *MIS Quaterly* 169-196.

Bishop, P. 2018. "Knowledge Diversity and Entrepreneurship Following an Economic Crisis: An Empirical Study of Regional Resilience in Great Britain." *Entrepreneurship & Regional Development* 31 (5-6): 1-20.

Blasius, J. and Greenacre, M. (2006) Correspondence Analysis and Related Methods in Practice. In: Greenacre, M. and Blasius, J., Eds., Multiple Correspondence Analysis and Related Methods, Chapman & Hall/CRC Press, London, 4-40. http://dx.doi.org/10.1201/9781420011319.ch1.

Bode, C., and J. R. Macdonald. 2017. "Stages of Supply Chain Disruption Response: Direct, Constraining, and Mediating Factors for Impact Mitigation". *Decision Sciences*, 48: 836-874. https://doi.org/10.1111/deci.12245.

Bosio, E, S. Djankov, F. Jolevski, and R. Ramalho. 2020. "Survival of Firms during Economic Crisis." World Bank Policy Research Working Paper 9239.

Bosio, E., F. Jolevski et al. 2020. "Survival of Firms in Developing Economies During Economic Crisis". In *COVID-19 in Developing Economies*, edited by S. Djankov and U. Panizza, 157-174. London: CEPR Press.

Cao, Q., E. Gedajlovic, and H Zhang. 2009. "Unpacking Organizational Ambidexterity: Dimensions, Contingencies, and Synergistic Effects." *Organizational Science* 20 781–796.

Caselli, M. 2008. "Measuring... What? Notes on Some Globalization Indices." *Globalizations* 5 (3): 383-404. https://www.tandfonline.com/doi/abs/10.1080/14747730802252529.

Chaudhary, Sanjay. (2019). Knowledge stock and absorptive capacity of small firms: the moderating role of formalization. *Journal of Strategy and Management*. 12. 10.1108/JSMA-09-2018-0100.

Cherchye, L., W. Moesen, N. V. Rogge, et al. 2008. "Creating Composite Indicators with DEA and Robustness Analysis: The Case of the Technology Achievement Index". *Journal of the Operational Research Society* 59: 239-251.

Dahles, H. and T. Susilowati. 2015. "Business Resilience in Times of Growth and Crisis. *Annals of Tourism Research* 51.

Datta, P. 2017. " "Supply Network Resilience: A Systematic Literature Review and Future Research." *The International Journal of Logistics Management.*

Despostis, D. K. 2002. "Improving the Discriminating Power of DEA: Focus on Globally Efficient Units." *Journal of the Operational Research Society* 53*:* 314-323.

———. 2005a. "A Reassessment of the Human Development Index via Data Envelopment Analysis". *Journal of the Operational Research Society* 56: 969–980.

———. 2005b. "Measuring Human Development via Data Envelopment Analysis: The Case of Asia and the Pacific." *Science Direct* 385-390.

Doern, R., N. Williams, and T. Vorley. 2019. "Special Issue on Entrepreneurship and Crises: Business as Usual? An Introduction and Review of the Literature." *Entrepreneurship & Regional Development* 31 (5-6): 400-412.

Fiskel, J. 2006. "Sustainability and Resilience: Toward a Systems Approach." *Sustainability: Science, Practice and Policy* 2: 14-21.

Foa, R., and J. C. Tanner. 2012. "Methodology of the Indices of Social Development." International Institute of Social Studies Working Paper No. 2012-4.

Freudenberg, M. 2003. "Composite Indicators of Country Performance: A Critical Assessment". OECD Science, Technology and Industry Working Papers 2003/16.

Fritz N., L. Pancaldi, T. Poppensieker, et al. 2021. "The Resilience Imperative: Succeeding in Uncertain Times" McKinsey and Company. https://www.mckinsey.com/capabilities/risk-and-resilience/our-insights/the-resilience-imperative-succeeding-in-uncertain-times.

Golan, M. S., L. H. Jernegan, and I. Linkov. 2020. "Trends and Applications of Resilience Analytics in Supply Chain Modeling: Systematic Literature Review In the Context of the COVID-19 Pandemic." *Environment Systems and Decisions* 40: 222-243.

Greco, S., A. Ishizaka, M. Tasiou, et al. 2019. "On the Methodological Framework of Composite Indices: A Review of the Issues of Weighting, Aggregation, and Robustness." *Springer* 141: 61-94.

Greyling, T., and F. Tregenna. 2017. "Construction and Analysis of a Composite Quality of Life Index for a Region of South Africa." *Social Indicators Research* 131 (3): 887-930.

Grupp, H., and M. Mogee. 2004. Indicators for National Science and Technology Policy: How Robust are Composite Indicators? *Research Policy* 33 (9), 1373-1384.

Hallward-Driemeier, M., and B. Rijkers. 2013. "Do Crises Catalyze Creative Destruction? Firm-level Evidence from Indonesia." *The Review of Economics and Statistics* 95 (5): 1788-1810.

Hamel, G., and L. Valikangas. 2003. Why Resilience Matters. *Harvard Business Review* 81: 56-57.

Hatefi, S.M, and S.A Torabi. 2010. "A Common Weight MCDA–DEA Approach to Construct Composite Indicators." *Ecological Economics* 70 (1): 114-120.

Hemans, E., F. Van den Bossche, and G. Wet. 2008. "Combining Road Safety Information in a Performance Index. ." *Accident Analysis and Prevention* 40 (4): 1337-1344.

Hopenhayn, H. 1992. "Entry, Exit, and Firm Dynamics in Long Run Equilibrium," *Econometrica* 60 (5): 1127-1150.

Hubert, M., P. Rousseeuw, P., and K. Vanden Branden. 2005. "ROBPCA: A New Approach To Robust Principal Component Analysis". *Technometrics 47(1)*, 64-79.

Jacobs, R., P. Smith, and M. Goddard. 2004. "Measuring Performance: An Examination of Composite Performance Indicators". Centre for Health Economics, Technical Paper Series 29.

Jansen, Justin J. P., Frans A. J. Van Den Bosch, and Henk W. Volberda. "Exploratory Innovation, Exploitative Innovation, and Performance: Effects of Organizational Antecedents and Environmental Moderators." *Management Science* 52, no. 11 (2006): 1661–74. http://www.jstor.org/stable/20110640.

Joiner, B. and Josephs, S. (2007), "Developing agile leaders", *Industrial and Commercial Training*, Vol. 39 No. 1, pp. 35-42. https://doi.org/10.1108/00197850710721381.

ILO. 2020. *Driving Up Productivity - A Guide for Employer and Business Membership Organizations*. Regional Office for Latin America and the Caribbean, Bureau for Employers' Activities.

———. 2021. *COVID-19 and the world of work. Updated estimates and analysis.* ILO Monitor.

IMF (International Monetary Fund). 2012. *Data Quality Assessment Framework (DQAF) for National Account Statistics.*

Jansen, J.J., F.A. Van den Bosch, and H.W Volberda. 2005. "Exploratory Innovation, Exploitative Innovation, and Ambidexterity: The Impact of Environmental and Organizational Antecedents." *Schmalenbach Business Review* 57: 351-363.

Jovanovic, B. 1982. "Selection and the Evolution of Industry," *Econometrica* 50 (3): 649- 670.

Kovacevic, M. and C. G. Aguna. 2010. "Human Development Report". UNDP Research Paper 2010/47.

Lu, Ying, and K. (Ram) Ramamurthy. "Understanding the Link Between Information Technology Capability and Organizational Agility: An Empirical Examination." *MIS Quarterly* 35, no. 4 (2011): 931–54. https://doi.org/10.2307/41409967.

Mahlberg, B., and M. Obersteiner. 2001. "Remeasuring the HDI by Data Evelopment Analysis: Interim Report". International Institute for Applied Systems Analysis. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1999372.

March, J.G. 1991. "Exploration and Exploitation in Organizational Learning." *Organization Science* 2: 71-87.

Mazziota, M., and A. Pareto. 2016. On a Generalized Non-Compensatory Composite Index for Measuring Socio-Economic Phenomena. *Social Indicators Research* 127(3), 983-1003.

Melitz, M. 2003. "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity," *Econometrica* 71 (6): 1695-1725.

Melitz, M., and G. Ottaviano. 2008. "Market Size, Trade, and Productivity". *Review of Economic Studies* 75 (1): 295-316.

Munda, G. 2012. "Choosing Aggregation Rules for Composite Indicators." *Social Indicators Research* 109 (3): 337-354.

OECD (Organisation for Economic Co-operation and Development). 2008. *Handbook on Constructing Composite Indicators: Methodology and User Guide.* Paris: OECD Publishing.

———. 2009. *The Impact of the Global Crisis on SME and Entrepreneurship Financing and Policy Responses*. Paris: OECD Centre for Entrepreneurship, SMEs and Local Development.

Osberg, L., and A. Sharpe. 2005. "How Should We Measure the "Economic" Aspects of Well Being?" *Review of Income and Wealth* 51 (2): 311-336.

Papa, A., Dezi, L., Gregori, G.L., Mueller, J. and Miglietta, N. (2020), "Improving innovation performance through knowledge acquisition: the moderating role of employee retention and human resource management practices", *Journal of Knowledge Management*, Vol. 24 No. 3, pp. 589-605. https://doi.org/10.1108/JKM-09-2017-0391.

Permanyer, I. 2011. Assessing the Robustness of Composite Indices Ranking. *Review of Income and Wealth* 57 (2): 306-326.

Rippin, N. "A response to the weaknesses of the multidimensional poverty index (MPI): The Correlation Sensitive Poverty Index (CSPI)". Briefing Paper, 2011. https://www.die-gdi.de/uploads/media/BP_19.2011.pdf.

Runyan, R.C. 2006. "Small Business in the Face of Crisis: Identifying Barriers to Recovery from a Natural Disaster." *Journal of Contingencies Crisis Management* 14: 12-26.

Rupietta, C., and U Backes-Gellner. 2019. "Combining Knowledge Stock and Knowledge Flow to Generate Superior Incremental Innovation Performance—Evidence from Swiss Manufacturing." *Journal of Business Research* 94: 209-222.

Sahn, D.E. and Stifel, D. (2003), Exploring Alternative Measures of Welfare in the Absence of Expenditure Data. Review of Income and Wealth, 49: 463-489. https://doi.org/10.1111/j.0034-6586.2003.00100.x.

Saisana, M., M. Nardo, and A. Saltelli. 2005. "Uncertainty and Sensitivity Analysis of the 2005 Environmental Sustainability Index." In *Environmental Sustainability Index: Benchmarking National Environmental Stewardship*, edited by D. Esty, T. Srebotnjak and A. de Sherbinin, 75-78. New Haven: Yale Center for Environmental Law and Policy.

Saisana, M., and S. Tarantola. 2002. *State-of-the-Art Report on Current Methodologies and Practices for Composite Indicator Development.* Ispra (VA) Italy: European Commission. Joint Research Centre EUR 20408 EN.

Sambamurthy, V., and R.W Zmud. 1997. *At the Heart of Success: Organizationwide Management Competencies. In Steps to the Future: Fresh Thinking on the Management of IT-Based Organizational Transformation*. San Fransisco: Jossey-Bass: 143-163.

Shen, Y. and E. Hermans, and T. Brijs, et al. 2012. “Data Envelopment Analysis for Composite Indicators: A Multiple Layer Model”. *Social Indicators Research* 114.

Schumpeter, J. 1934. *The Theory of Economic Development*. Cambridge, MA: Harvard University Press.

Simeone, C.L. 2015. “Business Resilience: Reframing Healthcare Risk Management.” *Joint Healthcare Risk Management* 35: 31-37.

Teece, D.J., G. Pisano, and A Shuen. 1997. “Dynamic Capabilities and Strategic Management.” *Strategic Management* 18: 509-533.

Tsaples, G. and J. Papathanasiou. 2020. “Multi-level DEA for the Construction of Multi-dimensional Indices. *MethodsX* 7: 101169.

Vogus, T.J., and K.M Sutcliffe. 2007. “The Safety Organizing Scale: Development and Validation of a Behavioral Measure of Safety Culture in Hospital Nursing Units.” *Medical Care* 45: 46-54.

Wold, H. (1966) Estimation of Principal Components and Related Models by Iterative Least Squares. In: Krishnaiah, P.R., Ed., Multivariate Analysis, Academic Press, New York.

Annexes

▶ Annex 1. Key drivers of business resilience – Aldianto et al. (2020) framework

Technology capability	Technology capability refers to the ability of a firm to use and develop technology for product, process or management development to support business performance. Bharadwaj, 2000 suggests that technological capabilities include technological infrastructure, human resources (consisting of technical and managerial skills) and intangible things such as knowledge assets, customer orientation and synergy. According to Afuah (2002), technological capability refers to the "ability that enables companies to use and develop various technologies by involving technology development, product development, production processes, manufacturing procedures, and technology estimates". Lu and Ram (2011) define technological capability as "the extent to which companies are good at managing information technology resources to support and improve business strategies and processes".
Agile leadership	According to Aldianto at al., agile leadership is necessary to make firms more effective in collaborating and adapting to changes. Abdul-Kareem (2020) defined that agile leadership as "the ability of a leader to be quick, adaptable, and flexible in responding to unforeseen events in an unfamiliar circumstance". Joiner and Josephs (2007) identified four key competencies of successful agile leaders in an unstable business environment: context-setting agility, stakeholder agility, creative agility and self-leadership agility.
Knowledge stock	Knowledge stock refers to accumulated knowledge of a company and its human capital. This factor is related to innovation and learning. Knowledge stock plays a role in increasing service or product innovation. In fact, Papa et al. (2018) found that firm innovation is driven by the knowledge possessed by employees. Chaudhary (2019) stated that the knowledge stock of a firm influences the development of its ability to acquire, assimilate and exploit external knowledge. Rupietta and Backes-Gellner (2019) suggested categorizing types of knowledge as either a) knowledge stock (i.e., human capital) or knowledge flows (governed by human resource management practices).
Innovation ambidexterity	Innovation ambidexterity refers to the balance between exploration and exploitation. Jansen et al. (2006) contended that explorative and exploitative innovation are interdependent activities. According to March 1991, exploration is linked to innovation in the sense of searching for new opportunities, discovery and experimentation, while exploitation refers to "the refinement and expansion of existing competencies, technologies and paradigms with returns that are proximate and predictable."

▶ Annex 2. Indices data used for construction of the BRI

Data source	Description
Digital Adoption Index – 2016	Measures digital adoption in 180 countries across three dimensions: people, government and business. The Digital Adoption Index (business dimension) is the simple average of four normalized indicators (0–1 scales): the percentage of businesses with websites, the number of secure servers, the speed of download and 3G (third generation) wireless coverage in the country.
Doing Business Index – 2019	Based on 41 indicators and 10 dimensions. The indicators are normalized using the min-max goalposts. The highest score represents the best regulatory performance. For more information about the subindices and indicators, see https://www.doingbusiness.org/en/custom-query.
Global Competitiveness Index 4.0 – 2019	Covering 141 economies, this index seeks to measure the drivers of total factor productivity using 103 individual indicators organized in 12 dimensions or pillars: institutions, infrastructure, ICT adoption, macroeconomic stability, health, skills, product market, labour market, financial system, market size, business dynamism and innovation capability.
Global Digital Readiness Index – 2019	Measures digital readiness for 141 economies, with scores ranging on a scale from 0 to 25. The GRDI is based on seven dimensions: basic needs, human capital, ease of doing business, business and government investment, start-up environment, technology infrastructure, and technology adoption.
The Global Innovation Index – 2019	A composite index estimated for 132 economies using 81 indicators (63 hard-data indicators, 15 composite indicators and three survey questions from the WEF Executive Opinion Survey).

▶ Annex 3. Business resilience indicators

Dimensions	Subdimension	Indicators	Source
Institutions	**Political**	*The Political, Legal, Operational or Security Risk Index 2020.* This index measures the likelihood and severity of political, legal, operational or security risks affecting business operations. *(GII score 0–100, Direction: +)*	IHS Markit, Country Risk Scores, aggregated for end Q1, Q2, Q3 and Q4 2020.
		Government Effectiveness Index 2019. This index reflects the quality of public services, civil service and independence from political pressures. *(GII score 0–100, Direction: +)*	IHS Markit, Country Risk Scores, aggregated for end Q1, Q2, Q3 and Q4 2020.
	Regulatory	*Regulatory Quality Index 2019.* This index reflects the ability of the government to formulate and implement policies that promote private-sector development. *(GII score 0-100, Direction: +)*	World Bank, Worldwide Governance Indicators, 2019 update.
		Rule of Law Index, 2019. This index reflects confidence in and the quality of contract enforcement, property rights, police and courts and the likelihood of crime and violence. *(GII score 0–100, Direction: +)*	World Bank, Worldwide Governance Indicators, 2019 update.
		Cost of Redundancy Dismissal. This measures the cost of advance notice requirements, and severance payments due, when terminating a redundant worker. *(GII score 0–100, Direction: -)*	World Bank, Doing Business 2019
	Business	*Ease of starting business.* *(GII score 0–100, Direction: +)*	World Bank Doing Business 2019
		Ease of Resolving Insolvency. *(GII score 0–100, Direction: +)*	World Bank Doing Business 2019
		Ease of protecting minority investors, 2019. *(GII score 0–100, Direction: +)*	World Bank Doing Business 2019
		Flexibility of wage determination. Response to the survey question, "In your country, how are wages generally set?" 2018–2019 weighted average or most recent period available. (GCI: 1 = by a centralized bargaining process; 7 = by each individual company)	World Economic Forum, Executive Opinion Survey, 2017–2018
		Applied tariff rate, weighted average, all products (%) 2019 *(GII score 0-100, Direction: -)*	World Bank, World Development Indicators database
Human capital	**Knowledge stock and flows**	*Tertiary enrolment, % gross 2018.* *(GII, Score 0-100, Direction: -)*	UNESCO Institute for Statistics (UIS) online database (2010–20)
		Extent of staff training Response to the survey question, "In your country, to what extent do companies invest in training and employee development?" 2018–2019 weighted average or most recent period available. *(GCI: 1 = not at all; 7 = to a great extent)*	World Economic Forum, Executive Opinion Survey, 2017–2018
		On the job training *(GCI: Direction +)*	World Economic Forum, Executive Opinion Survey, 2017–2018

Dimensions	Subdimension	Indicators	Source
Human capital	**Knowledge stock and flows**	*Charges for use of intellectual property*, i.e., payments (%, total trade, three-year average), 2019. *(GII score 0–100, Direction: +)*	World Trade Organization, Trade in Commercial Services database; values based on the classification of the sixth (2009) edition of the International Monetary Fund's Balance of Payments and International Investment Position Manual and Balance of Payments database.
		Telecommunications, computer, and information (ICT) services imports (% of total trade), 2019. *(GII score 0–-100, Direction: +)*	World Trade Organization, Trade in Commercial Services database; values based on the classification of the sixth (2009) edition of the International Monetary Fund's Balance of Payments and International Investment Position Manual and Balance of Payments database.
	Knowledge creation and impact	*Number of resident patent applications* filed at a given national or regional patent office (per billion PPP$ GDP) 2019. *(GII score 0–100, Direction: +)*	World Intellectual Property Organization, Intellectual Property Statistics; International Monetary Fund, World Economic Outlook Database, October 2020
		Number of scientific and technical journal articles (per billion PPP$ GDP) 2020. *(GII score 0–100, Direction: +)*	Clarivate, Web of Science, accessed 15 March 2021; International Monetary Fund, World Economic Outlook Database, October 2020
		Total computer software spending (% of GDP), 2020. *(GII score 0–100, Direction: +)*	IHS Markit, Information and Communication Technology Database 2020
Technology capability	**Research and development**	*Gross expenditure on R&D* (GERD), % GDP 2019. *(GII score 0–100, Direction: +)*	UNESCO Institute for Statistics (UIS) online database; Eurostat, Eurostat database; OECD, Main Science and Technology Indicators (MSTI) database, 2021
		Average score of the top three universities according to the QS world university ranking 2020. *(GII score 0–100, Direction: +)*	QS Quacquarelli Symonds Ltd, QS World University Ranking, Top Universities
		University–industry R&D collaboration. Average answer to the survey question, "In your country, to what extent do businesses and universities collaborate on research and development (R&D)?" [1 = not at all; 7 = to a great extent] *(GII score 0–100, Direction: +)*	World Economic Forum, Executive Opinion Survey 2020
	Infrastructure	*Information and communication technologies access index, 2019.* Composite index that equally weights fixed telephone subscriptions per 100 inhabitants, mobile cellular telephone subscriptions per 100 inhabitants, international internet bandwidth per internet user, percentage of households with a computer and percentage of households with internet access. *(GII score 0–100, Direction: +)*	GII calculations based on the World Telecommunication/ICT Indicators Database (released January 2020), following the methodology of the ITU ICT Development Index 2017

Dimensions	Subdimension	Indicators	Source
Technology capability	**Infrastructure**	*Information and communication technologies use index, 2019.* Composite index that weights five ICT indicators (20% each): (1) Fixed telephone subscriptions per 100 inhabitants; (2) Mobile cellular telephone subscriptions per 100 inhabitants; (3) International Internet bandwidth (bit/s) per Internet user; (4) Percentage of households with a computer; and (5) Percentage of households with Internet access. *(GII score 0–100, Direction: +)*	GII calculations based on the World Telecommunication/ICT Indicators Database (released January 2020), following the methodology of the ITU ICT Development Index 2017.
		Quality of port infrastructure. *(GCI: 1 worst –7 best)*	World Economic Forum's calculations, 2017–2018
		Quality of airport infrastructure. *(GCI: 1 worse –7 best)*	World Economic Forum's calculations, 2017–2018
		Logistics performance index 2018. Composite indicator that includes customs (efficiency and border management clearance), infrastructure (quality of trade and transport infrastructure), international shipments (ease of arranging competitively priced shipments), service quality of logistic services, tracking and tracing of consignments, and timeliness of meeting expected delivery times. *(GII score 0–100, Direction: +)*	World Bank and Turku School of Economics, Logistics Performance Index 2018; Arvis et al., 2018, Connecting to Compete 2018: Trade Logistics in the Global Economy – The Logistics Performance Index and its Indicators
	Innovation	Production process sophistication. *(GCI: Direction +)*	World Economic Forum, Executive Opinion Survey, 2017–2018
		State of cluster development and depth. Average answer to the survey question, "In your country, how widespread are well-developed and deep clusters (geographic concentrations of firms, suppliers, producers of related products and services and specialized institutions in a particular field)?" [1 = non-existent; 7 = widespread in many fields]. *(GII score 0–100, Direction: +)*	World Economic Forum, Executive Opinion Survey 2020 (2018–20)
		International co-inventions Number of patent family applications with co-inventors located abroad per million population, 2013–2015 average	Organisation for Economic Co-operation and Development (OECD), STI Micro-data Lab: Intellectual Property database
Financial	**Liquidity**	*Financial obstacles*	Enterprise Surveys, World Bank
		Overdraft facility	Enterprise Surveys, World Bank
	Funding	*Ease of getting credit index, 2019* A composite indicator combining the strength of the legal rights index and the depth of credit information index. The former set measures the characteristics that facilitate lending. The latter measures the coverage, scope and accessibility of credit information available through credit reporting service providers, such as credit bureaus or credit registries. *(GII score 0–100, Direction: +)*	World Bank, Doing Business 2020

Dimensions	Subdimension	Indicators	Source
Financial	**Funding**	*Trustworthiness and confidence* A composite indicator that combines the soundness of banks, regulation of securities exchanges and legal right index.	World Economic Forum, Executive Opinion Survey, 2017–2018
		Financing through local equity market Average answer to the survey question: "In your country, to what extent can companies raise money by issuing shares and/or bonds on the capital market?" [1 = not at all; 7 = to a great extent]	World Economic Forum, Executive Opinion Survey, 2017–2018
		Domestic credit to private sector (% of GDP), 2019. *(GII score 0–100, Direction: +)*	International Monetary Fund, International Financial Statistics and data files; World Bank and OECD GDP estimates; extracted from the World Bank's World Development Indicators database
Management	**Management**	*Management index, 2019.* Considers the manner in which businesses manage problems, the number of production indicators, achievement of production targets, communication of production targets within the firm, basis for managers' bonuses and basis for promotion of managers.	World Bank Enterprise Survey, 2019
		Efficient use of talent. *(GCI: Direction +)*	World Economic Forum, Executive Opinion Survey, 2017–2018
		Extent to which ICTs enable new organizational models, 2018. Average answer to the question, "In your country, to what extent do ICTs enable new organizational models (e.g., virtual teams, remote working, telecommuting) within companies?" [1 = not at all; 7 = to a great extent] *(GII score 0–100, Direction: +)*	World Economic Forum, Executive Opinion Survey 2019
		Capacity to attract and retain talent *(GCI: Direction +)*	World Economic Forum, Executive Opinion Survey, 2017–2018

▶ Annex 4. Advantages and disadvantages of normalization/standardization methods

Method	Strengthens	Weaknesses
Ranking[1]	- Simple method to apply and explain to policymakers - Not affected by outliers - Allows comparisons across time in terms of relative positions	- Loss of absolute-level information
Standardization[2]	- Takes care of heteroskedasticity issues, scaling the variance to one. - The range of the scores varies for each indicator. (It gives greater weight to an indicator in units with extreme values.)	- Extreme values have a greater effect on the composite indicator. - Does not allow for global mean comparisons over time
Min–max goalposts	- Preserves absolute values differences - Identical range for all indicators. (The range for indicators with very little variation is increased.) - Do not force the mean of the distribution to be zero every year	- Extreme values or outliers could distort the transformed indicator. - Min–max normalization could widen the range of indicators lying within a small interval, increasing the effect on the composite indicator more than the z-score transformation. - Definition of the goalposts (minimum and maximum values) might change every year, which might affect year-to-year comparisons. - Scaling does not always make indicators comparable. For example, two indicators might be in the scale of 0–10, but in practice most of the scores of one indicator are concentrated in the 7–10 range, while, in the other indicator, responses are spread out over 0–10.
Distance to reference	- Measures the relative position of a given indicator with respect to a reference point (e.g., a reference country). It could be the average or the maximum (frontier)	- The approach is based on extreme values, which could be unreliable outliers. - Not very common lately
Discretization[3]	- Splits up variables in categories based on the percentile distribution	- Difficult to follow increases over time - Exclude information about the variance of the transformed indicators - Splits the data irrespective of the underlying distribution even when there is very little variability
Indicators above or below the mean	- Simple method that is not affected by outliers	- Arbitrariness of the threshold level - Omission of absolute-level information - Not very common lately
Percentage of annual differences over consecutive years	- Simple method that is not affected by outliers	- It can be used only for indicators that are available for a number of years.

1 **Ranking** is one of the simplest methods and, despite its limitations, it is still used by several institutions. For example, the Information and Communications Technology Index, the Medicare Study on Healthcare Performance, Doing Business use unweighted aggregation of indicator rankings.

2 **Standardization** refers to the process of rescaling the data to have a mean of zero (mean removal) and variance scaling (standard deviation of 1 − unit variance), while normalization refers to the process of rescaling values into a range between zero and one. Depending on data properties, it can use various methods, such as max absolute scaler, robust scalers and non-parametric methods.

3 **Discretization** is useful when the index combines categorical and continuous variables. For example, the Social Institutions Gender Index developed by the OECD discretizes continuous variables, aggregating data by percentiles (i.e., quintiles). There are several ways to partition the data, such as one hot encoding, k-bins discretization and feature binarization, among others. The selection depends on the data properties of the indicators and the weighting and aggregation methods to be used to construct the index.

▶ Annex 5. Comparison of missing imputation methods

	Advantages	Limitations
Case deletion	- Avoids criticisms from countries that are rated poorly based on estimated or imputed data	- Produces bias estimates if deleted records are not a random sub-sample of the original sample - Standard errors will be larger in smaller samples (higher uncertainty).
Imputation	- Could lead to the minimization of bias and reduces the effort to collect data - Underestimates the variance, partially reflecting imputation uncertainty - Could lead to erroneous results when a significant proportion of data is missing - Difficult to defend against criticisms from countries that are rated poorly based on estimated or imputed data	- Allows data to influence the type of imputation, leading to substantial bias - Subject to criticisms when scores are based on imputed or calculated data.[1] - Imputation could be susceptible to omitted variables bias when data is not missing at random. - Imputation could be unreliable in cases where appropriate models cannot be determined from available data. - Imputation could lead to erroneous results when data is missing for a large number of countries and/or indicators (more than 5%) (OECD, 2008)).

1 A problem could arise when a country' score is obtained from imputed data, which leads to uncertainty in the rankings results that is subject to criticisms.

Annex 6. Missing values analysis and summary statistics

Missing values

Variable	Missing	Total	Percent missing
Tertiary enrolment – GII	1	50	2
Gross Expenditure R&D – GII	2	50	4
ICT access – GII	1	50	2
ICT use –p GII	1	50	2
Logistics performance –, GII	2	50	4
Gross capital formation – GII	1	50	2
Applied tariff rate – GII	1	50	2
Patents filled – GII	2	50	4
High-tech imports	1	50	2
Business and government investment – GDRI	1	50	2
Human capital – GDRI	1	50	2
Startup environment – GDI	1	50	2

Summary statistics

Variable	Obs	Mean	Std dev	Min	Max
Tertiary enrolment – GII	50.0	43.1	20.3	2.6	100.0
Gross expenditure R&D – GII	50.0	15.2	12.4	0.3	56.5
ICT access – GII	50.0	66.4	16.0	20.8	94.2
ICT use – GII	50.0	55.7	17.6	15.6	84.8
Logistics performance – GII	50.0	43.2	19.5	12.9	92.4
Gross capital formation – GII	50.0	38.0	13.9	13.7	84.0
Applied tariff rate – GII	50.0	81.5	13.3	31.3	96.2
Patents filled – GII	48.0	9.8	20.4	0.0	85.9
High-tech imports	50.0	25.5	13.5	0.0	63.6
Business and government investment – GDRI	50.0	1.4	0.3	0.8	2.5
Human capital – GDRI	50.0	2.7	0.4	1.4	3.3
Startup environment – GDI	50.0	0.5	0.4	0.2	2.6

▶ Annex 7. Weights methods comparison

Method	Strengths	Weaknesses
Equal weighting	- Simplicity of construction, especially when there is a lack of theoretical structure to justify unequal weighting. - It might work well if indicators are uncorrelated or they are all highly correlated. - In some cases give the same results as more sophisticated methods such as PCA or UCM (e.g., Doing Business).	- Double counting, which refers to the issue of implicitly weighting an indicator higher than the desired level due to collinearity between indicators not addressed during aggregation.[1] - The composite indices might down-weight dimensions with a greater number of indicators.[2] - Total compensation/substitutability between indicators, which might not be desirable.
Principal component analysis (PCA)/factor analysis (FA)	- Analyses the underlying structure of the data, thus avoiding misinterpretation. - Largest factor loadings (i.e., weights) are assigned to the indicators with the largest variation across countries, which are the ones that explain differences in performance. - PCA avoids double counting by underweighting those factors that depend on the same underlying factor (i.e., weights are corrected for overlapping information between correlated indicators).	- Correlations do not necessarily represent the relative importance of the indicators in the composite index or do not necessarily (due to confusion between correlation and causality) correspond to the underlying relationships. - Saisana and Tarantola (2002) suggest that resulting weights do not necessarily correspond to the actual linkages among indicators and do not necessarily reflect a sound theoretical framework. - Sensitive to modifications in the data (revisions and updates), outliers, and small-sample datasets. Does not allow missing values. See Hubert et al. (2005). - May perform poorly if the largest variations in the indicators are not informative, which occurs when observed variables contain large measurement errors or variations coming from other latent variables.[3] - Requires the selection of hyperparameters (e.g., the number of components/factors, the rotation method, the choice of an extraction method, etc.), and the outcome is very sensitive to their selection. - The weights are inconsistent over time (changing with the inclusion of new data/indicators), making the comparison difficult to make. - Infeasible in certain cases due to either negative weights or very low correlation between indicators. - Assumes the presence of continuous indicators and a linear relationship between them.[4] - Does not handle categorical or ordinal data.
Cronbach coefficient alpha	- Measures the internal consistency in a set of indicators.	- Correlation does not necessarily represent the relative importance of the indicators in the composite index. - Meaningful only when the composite indicator is computed as a “scale” (as the sum of the individual indicators). - These algorithms may perform poorly if the largest variations in the indicators are not informative.
Cluster analysis	- Data-driven approach to group observations based on distance and gives some insights on the structure of the dataset.	- Selection of hyperparameters makes the results somewhat subjective (a priori selection of the number of clusters, the distance measure) - Significance tests are not valid for testing differences between clusters. - Always produces a grouping and, depending on the selection of parameters, gives different results.

Method	Strengths	Weaknesses
Factorial k-means analysis	- Combines k-means cluster analysis with aspects of FA and PCA to identify the best partition of the objects. - Achieves data reduction and synthesis simultaneously. - Fast algorithm.	- Suffers from several of the same weaknesses as PCA/FA.
Polychoric PCA	- Does not assume a linear relationship. This method uses non-linear methods (e.g., categorical PCA, polychoric PCA) Greyling and Tregenna (2017). - Handles categorical or ordinal data, assuming that there is a latent normally distributed variable that underlies the ordinal categorical variable, and does not assume a specific underlying structure of the data.	- Suffers from several of the same weaknesses as PCA/FA.
Regression-based approaches	- Perform better than PCA when a highly valid and reliable measure of the latent variable exists. - Find weights that are more relevant to a particular response variable, in contrast to PCA, which is an all-purpose method. - Perform better than PCA when the largest variations come from measurement errors or irrelevant factors.	- Greyling and Tregenna (2017) claim that regression approaches need a highly valid and reliable measure of the latent variable (that describes the concept intended to be measured). - Assumes strict linearity, which is not common in composite indices. See Saisana et al. (2005). - The regression method offers no solution in cases where data may be missing for particular countries. Imputation is needed. - Vulnerable to multicollinearity (When there is not an exact linear relationship among the predictors, but they are close to 1, the variances of the predictors will be overestimated.)
Partial least squares (PSL)	- PSL is not vulnerable to multicollinearity of covariates. - Finds weights that are more relevant to a particular response variable than PCA, which is an all-purpose method.[5] - PSL will perform better than PCA when the largest variations come from measurement errors or irrelevant factors.	- PSL can be done when a highly valid and reliable measure of the latent variable exists (that describes the concept intended to be measured). - The PSL algorithm was developed mainly for continuous variables scenarios, and this is a limitation for the construction of composite indices that are composed of non-metric variables.
Data envelopment analysis (DEA)[6]	- The composite indicator values are independent of the units of the indicators and the composite indicator meets the property of units' invariance, making any normalization stage redundant. Fixed weighting methods do not meet the units' invariance property.	- Differential weighting makes comparison across countries difficult. The axiom of neutrality from social theory claims that all alternatives (e.g., countries) must be treated equally (OECD 2008).[7] - The model considers any country supporting the frontier to be equally well performing even if it is superior with respect to one indicator but performs poorly with respect to other (the boundary problem). In an extreme case, the DEA could assign a weight of one to an indicator and zero to the others.

Method	Strengths	Weaknesses
Unobserved components model (UCM)	- Weights do not depend on ad hoc restrictions. - Estimates the uncertainty associated with the process of aggregating indicators, providing margins of error to avoid misinterpretation of small differences that are not statistically significant.[8] - The problem is formulated as a signal extraction problem, allowing the inclusion of several disparate data sources to get the best possible signal of the underlying concept. - The UCM maintains the cardinality of the data, in contrast to methods based on rankings, which do not preserve the absolute values information. - Provides a data-driven, precision weighting for indicators, extracting signal from noisy data. - The UCM weeds out outlier scores. The criterion that varies the most across observations carries less weight in the aggregating rating.	- Reliability and robustness of results depend on the availability of sufficient data. - Highly correlated individual indicators could lead to identification problems. - Rewards the absence of outliers (the weights are a decreasing function of the variance of individual indicators). - If each country has a different number of individual indicators, the weights are country-specific. - The UCM requires z-score standardization, which does not allow the comparison of global averages across years because the mean is set to zero. - The UCM assumes that the errors are independent across indicators. - The UCM assumes that any correlation between indicators is because they are both measuring the same underlying unobserved component. - Since not all countries have data on all individual indicators, the weight vector could be unit-specific. This might lead to non-comparability of country values in the composite index.
Matching percentiles	- Handles many missing values (imputation is not necessary).[9] - This method tends to be easier to understand than other methods, such as the UCM. - The score of each observation is independent of other observations, which reduces the chances of omitted variable bias. - The method does not rely on the linearity assumption, which is often unrealistic. The functional form is non-parametric; the matching percentiles rely on ordinal rather than cardinal information.	- Relies only on rankings rather than cardinal values. - Not commonly used.

1 There is always some positive correlation between variables that might affect weighting. The analyst should utilize a threshold to define which correlation level could be deemed double counting (OECD, 2008).

2 To address this issue, indices such as the Child Well-being index use the same number of indicators by dimension. For the Global Innovation Index, the World Intellectual Property Organization uses a weighted arithmetic mean for aggregation to achieve a balanced contribution of indicators and sub-indices. (The numbers of indicators per sub-pillar and of sub-pillars per pillar are not the same.)

3 Tends to fail when there is no correlation between indicators or the variation is very small. For example, to measure corruption from survey datasets, the values may be influenced not only by corruption but also by other latent factors, such as the attitude of respondents.

4 In some cases, the results using equal weights and PCA/FA/MCA weights tend not to differ substantially. For instance, the Doing Business Indicators 2005 report conducts such a comparison and shows minimal differences.

5 PCA, Fact Analysis and MCA are all-purpose methods in the sense that they are meant to extract the largest variation of the indicators used to build the composite index regardless of what one wishes the composite index to measure (Wold, 1966).

6 DEA was proposed by Charnes et al. (1978). HDI weights were estimated by Mahlberg and Obersteiner (2001) in Despotis 2005a).

7 However, according to Greco et al. (2019), the differential weighting scheme for units could be desirable for policymakers because each unit chooses its own weight in such a way as to maximize its performance. To address differential weighting, a number of techniques have been applied to arrive at a common weighting scheme, such as multi-criteria decision analysis – data envelopment analysis (see, among others, Despostis (2005) and Hatefi and Torabi (2010).

8 There are other indices that provide margin errors in their score estimates, such as the Global Integrity Index.

9 When the matching percentiles method as described here encounters a country with a missing value for a particular variable, that observation is skipped, and no matched score is calculated for that country variable.

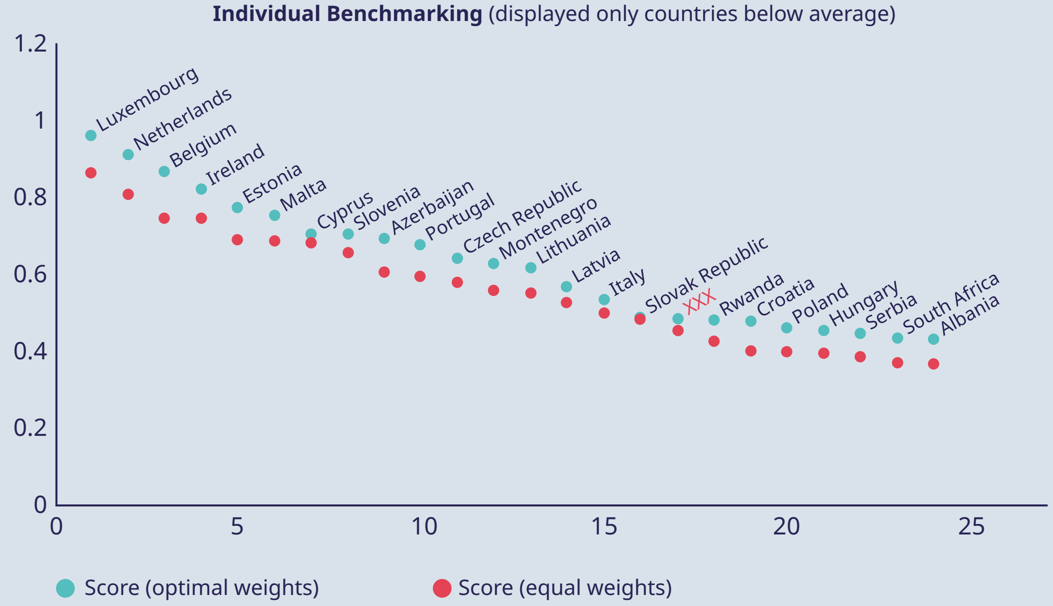
▶ Annex 8. DEA scores using optimal and equal weights by country

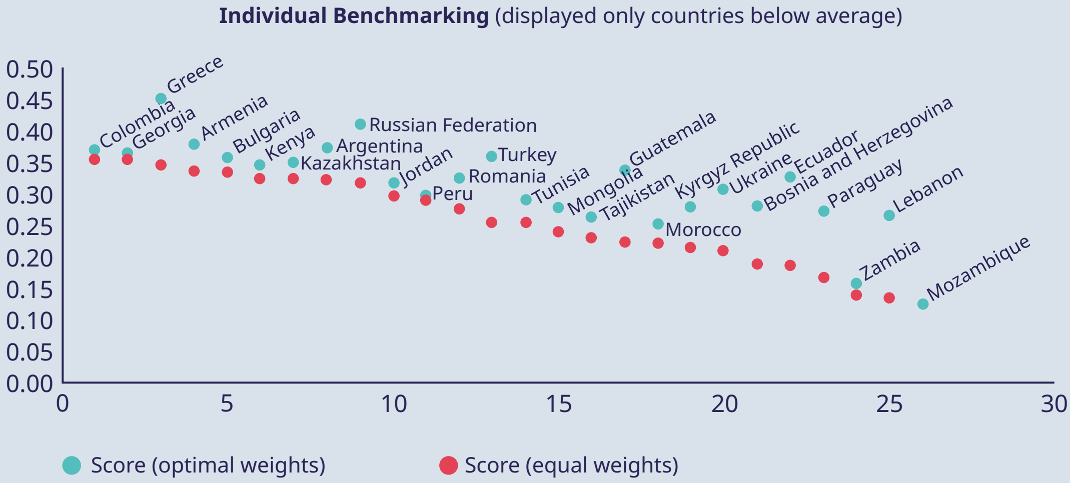
Country	Score (Optimal weights)	Score (Equal weights)	Rank (Equal weights)
Luxembourg	0.91	0.86	1
Netherlands	0.96	0.80	2
Belgium	0.86	0.74	3
Ireland	0.82	0.74	4
Estonia	0.70	0.69	5
Malta	0.75	0.68	6
Cyprus	0.77	0.68	7
Slovenia	0.70	0.65	8
Azerbaijan	0.64	0.60	9
Portugal	0.67	0.59	10
Czech Republic	0.62	0.57	11
Montenegro	0.69	0.56	12
Lithuania	0.56	0.55	13
Latvia	0.53	0.52	14
Italy	0.61	0.50	15
Slovak Republic	0.48	0.48	16
Uruguay	0.48	0.45	17
Rwanda	0.48	0.42	18
Croatia	0.43	0.40	19
Poland	0.49	0.39	20
Hungary	0.43	0.39	21
Serbia	0.46	0.38	22
South Africa	0.45	0.37	23
Albania	0.43	0.36	24
Colombia	0.37	0.35	25
Georgia	0.36	0.35	26
Greece	0.45	0.34	27
Armenia	0.38	0.34	28

Country	Score (Optimal weights)	Score (Equal weights)	Rank (Equal weights)
Bulgaria	0.36	0.33	29
Kenya	0.35	0.32	30
Kazakhstan	0.35	0.32	31
Argentina	0.37	0.32	32
Russian Federation	0.41	0.32	33
Jordan	0.31	0.29	34
Peru	0.29	0.29	35
Romania	0.32	0.27	36
Turkey	0.36	0.25	37
Tunisia	0.29	0.25	38
Mongolia	0.28	0.24	39
Tajikistan	0.26	0.23	40
Guatemala	0.33	0.22	41
Morocco	0.25	0.22	42
Kyrgyz Republic	0.28	0.21	43
Ukraine	0.31	0.21	44
Bosnia and Herzegovina	0.28	0.19	45
Ecuador	0.32	0.18	46
Paraguay	0.27	0.17	47
Zambia	0.16	0.14	48
Lebanon	0.26	0.13	49
Mozambique	0.12	0.10	50

The table shows DEA optimization weights for each economy and their corresponding scores as well as the scores estimated using equal weights for all units, as derived by the second maximization formula. The DEA results show that Luxembourg ranks first, with the highest score. When using the optimal scores to rank countries, we find that Luxembourg (first position), Netherlands, Belgium and Ireland are the countries at the top, while Paraguay, Zambia, Lebanon, and Mozambique (last position) are at the bottom of the ranking.

Individual benchmarking displaying only countries above average score (top) and countries below average (bottom)





The gap between optimal and equal weights shows the differences on individual and common preferences over the importance of each dimension.

▶ Annex 9. Country scores by dimension, ordered from high to low based on the score of the composite index

Country	Income group	Institutions	Human Capital	Technology	Financial	Management
Luxembourg	High Income	0.76	0.76	0.77	1.00	1.00
Netherlands	High Income	1.00	1.00	1.00	0.15	0.63
Belgium	High Income	0.89	0.83	0.97	0.24	0.58
Ireland	High Income	0.98	0.73	0.82	0.27	0.61
Estonia	High Income	0.79	0.64	0.58	0.60	0.70
Malta	High Income	0.64	0.56	0.50	0.83	0.84
Cyprus	High Income	0.80	0.59	0.37	0.84	0.74
Slovenia	High Income	0.87	0.70	0.58	0.40	0.55
Azerbaijan	High Income	0.53	0.15	0.42	0.48	0.99
Portugal	High Income	0.83	0.63	0.63	0.03	0.48
Czech Republic	High Income	0.72	0.60	0.60	0.30	0.49
Montenegro	High Income	0.52	0.42	0.23	0.92	0.72
Lithuania	High Income	0.72	0.40	0.47	0.30	0.57
Latvia	High Income	0.71	0.38	0.40	0.32	0.53
Italy	High Income	0.70	0.56	0.72	0.17	0.23
Slovak Republic	High Income	0.61	0.34	0.35	0.36	0.52
Uruguay	High Income	0.60	0.36	0.31	0.46	0.42
Rwanda	Low Income	0.56	0.12	0.18	0.42	0.55
Croatia	High Income	0.57	0.38	0.26	0.38	0.34
Poland	High Income	0.63	0.48	0.44	0.04	0.21
Hungary	High Income	0.57	0.37	0.41	0.17	0.29
Serbia	High Income	0.48	0.50	0.24	0.45	0.32
South Africa	High Income	0.53	0.35	0.55	0.13	0.17
Albania	High Income	0.58	0.21	0.21	0.52	0.25
Colombia	High Income	0.49	0.24	0.31	0.15	0.35
Georgia	High Income	0.72	0.21	0.18	0.14	0.22
Greece	High Income	0.56	0.51	0.38	0.16	0.11

Country	Income group	Institutions	Human Capital	Technology	Financial	Management
Armenia	High Income	0.45	0.38	0.22	0.10	0.36
Bulgaria	High Income	0.52	0.27	0.33	0.14	0.24
Kenya	Low Income	0.37	0.20	0.34	0.10	0.38
Kazakhstan	High Income	0.62	0.17	0.27	0.00	0.21
Argentina	High Income	0.40	0.39	0.36	0.13	0.26
Russian Federation	High Income	0.39	0.44	0.48	0.03	0.19
Jordan	High Income	0.28	0.26	0.39	0.17	0.30
Peru	High Income	0.33	0.20	0.17	0.14	0.40
Romania	High Income	0.46	0.35	0.22	0.11	0.16
Turkey	High Income	0.42	0.34	0.43	0.04	0.02
Tunisia	Low Income	0.43	0.31	0.17	0.17	0.14
Mongolia	Low Income	0.33	0.20	0.03	0.22	0.30
Tajikistan	Low Income	0.17	0.08	0.15	0.28	0.37
Guatemala	High Income	0.07	0.18	0.16	0.57	0.31
Morocco	Low Income	0.40	0.13	0.27	0.05	0.09
Kyrgyz Republic	Low Income	0.32	0.15	0.00	0.34	0.23
Ukraine	Low Income	0.27	0.44	0.28	0.01	0.09
Bosnia and Herzegovina	High Income	0.25	0.14	0.08	0.51	0.10
Ecuador	High Income	0.00	0.13	0.25	0.58	0.21
Paraguay	High Income	0.09	0.00	0.06	0.57	0.23
Zambia	Low Income	0.21	0.05	0.04	0.14	0.15
Lebanon	High Income	0.04	0.28	0.33	0.26	0.00
Mozambique	Low Income	0.03	0.04	0.09	0.11	0.19

According to these results, Netherlands is the country with the best performance in the institutional dimension while Ecuador has the worst performance. In the human capital and technology dimensions, Netherlands exhibits the highest scores, while Paraguay has the lowest human capital score, and Kyrgyz Republic has the lowest technology score. In the financial dimension, Luxembourg exhibits the highest score, while Kazakhstan has the lowest. In management Luxembourg has the highest score, while Lebanon exhibits the lowest score.

ilo.org

International Labour Organization
Route des Morillons 4
1211 Geneva 22
Switzerland