Journal of the Operational Research Society

ISSN: 0160-5682 (Print) 1476-9360 (Online) Journal homepage: www.tandfonline.com/journals/tjor20

A Beta-Bernoulli composite index

Mariateresa Ciommi , Paolo Giudici , Francesca Mariani , Gloria Polinesi & Maria Cristina Recchioni

To cite this article: Mariateresa Ciommi , Paolo Giudici , Francesca Mariani , Gloria Polinesi & Maria Cristina Recchioni (14 Sep 2026): A Beta-Bernoulli composite index, Journal of the Operational Research Society, DOI: 10.1080/01605682.2026.2731154

To link to this article: https://doi.org/10.1080/01605682.2026.2731154

© 2026 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group

View supplementary material

Published online: 14 Sep 2026.

Submit your article to this journal

Article views: 175

View related articles

View Crossmark data

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=tjor20

JOURNAL OF THE OPERATIONAL RESEARCH SOCIETY
https://doi.org/10.1080/01605682.2026.2731154

OPEN ACCESS Check for updates

A Beta-Bernoulli composite index

Mariateresa Ciommi[a], Paolo Giudici[b], Francesca Mariani[a], Gloria Polinesi[c] and Maria Cristina Recchioni[a]

[a]Marche Polytechnic University, Ancona, Italy; [b]University of Pavia, Pavia, Italy; [c]University of Macerata, Macerata, Italy

ABSTRACT

Recent studies highlight the need for composite indices capable of integrating individual component indicators (or simply indicators) with expert opinions. To address this challenge, we propose a method for constructing composite indices that incorporates expert opinions in a fully endogenous manner through a Beta-Bernoulli process. The process parameters determine the performance of each unit in the phenomenon of interest across the different indicators. These parameters are estimated from the observed indicator values, which are scaled using expert opinion-based goalposts, independently of the units under consideration. The composite index is obtained as the mean of the Beta-Bernoulli variables, but penalised for the degree to which they deviate from the mean—captured in this model by their correlation. By varying the goalposts, the approach enables identification of which indicators play a crucial role in improving the composite index—namely, in improving unit performance—thus providing an effective monitoring tool.

PRACTITIONER SUMMARY

- The paper proposes a new methodological approach to build composite indices, that can incorporate expert opinions into their calculation. The proposal offers three key advantages.
- First, it provides a composite index rooted in a consistent statistical framework, that of Beta-Bernoulli processes.
- Second, the methodology defines a composite index which can embed expert opinions from practitioners in a transparent manner, without resorting to exogenous variables and hidden assumptions.
- Third, it provides a procedure for assessing the effect of expert opinions, expressed by policy makers, on the mean performance of a set of units.
- This can enable policy makers and investors, among others, to set realistic targets for the units under study, which can thereby be employed as motivation factors.

ARTICLE HISTORY
Received 1 August 2025
Accepted 4 September 2026

KEYWORDS
Beta-Bernoulli process; composite index; expert opinions; scaled indicators

1. Background

In recent years, researchers and policymakers have shown growing interest in consolidating multiple indicators into a single composite index. Major governmental institutions have developed such composite indices to capture information about complex phenomena. The methodologies available in the literature for defining and constructing composite indices vary according to their specific applications and contexts. Despite these variations, researchers commonly employ certain approaches to ensure robustness against data anomalies and to maintain the reliability and meaningfulness of the resulting classifications (see for details Hoffmann et al., 2008).

These approaches require defining a theoretical framework that determines the selection and combination of individual indicators to accurately reflect the dimensions and structure of the phenomenon under investigation (see Mazziotta & Pareto, 2017; Salzman, 2003).

Each framework requires normalising, weighting, and aggregating the selected indicators. Multiple approaches have been proposed for each of these three steps (Greco et al., 2019). For normalisation, practitioners can employ minimum-maximum scaling, distance-to-reference methods, or percentage change approaches. Aggregation weights can be determined through statistical methods, expert analysis, or participatory decision models (see, for

CONTACT Paolo Giudici giudici@unipv.it University of Pavia, Pavia, Italy

Supplemental data for this article can be accessed online at https://doi.org/10.1080/01605682.2026.2731154.

© 2026 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

example, Ceriani et al., 2016; Chen et al., 2023; Chen & Li, 2010; Shemshadi et al., 2011). Finally, the available aggregation techniques include linear methods (weighted sums) or nonlinear methods (complex aggregation functions).

Several well-knowon global indices illustrate different approaches to indicator aggregation and weighting. The Global Innovation Index (GII) employs a weighted arithmetic mean to combine various indicators into a single composite measure. The Human Development Index (HDI) uses a geometric mean of three normalised dimensions: health, education, and income. The Global Competitiveness Index (GCI) aggregates multiple pillars—including infrastructure and macroeconomic stability—using specific weights determined through both empirical and theoretical analysis. The Environmental Performance Index (EPI) normalises indicators on a 0–100 scale, applies relative importance weights, and aggregates them linearly. Finally, the Ease of Doing Business Index (EDBI) derives its weights from regression analysis based on the economic impact of individual indicators.

The recent literature highlights the importance of the aggregation phase and explores alternatives, such as unit-dependent weights, computed through the Benefit of Doubt (see, for example, Blancard & Hoarau, 2013; Karagiannis, 2017; Walheer, 2024); using the Data Envelopment Analysis approach (see, for example, Blancas et al., 2013; Cherchye et al., 2008; Zanella et al., 2015); or using penalised means (as Mariani et al., 2024; Verbunt & Rogge, 2018). All of these alternatives aim to better capture the specificity of individual units, enhancing the reliability of the composite indices by increasing their robustness to small data variations.

The importance of appropriately selecting normalisation, weighting, and aggregation techniques in constructing composite indices is emphasised in a recent review by Kuc-Czarnecka et al. (2020), which also examines the advantages and disadvantages of employing composite indices in real-world policymaking. Among the advantages, the work highlights their ability to summarise and simplify complex phenomena to aid decision-making, assess progress over time, and provide rankings of the units under examination.

Among the disadvantages, the authors warn against potentially misleading or overly simplistic policy messages, the possibility of political contention in selecting indicators and their weights, and the risk of implementing inappropriate policies when difficult-to-measure dimensions are not adequately considered.

Furthermore, Kuc-Czarnecka et al. (2020) emphasise that composite indices suffer from a degree of subjectivity and can potentially become "an end" rather than "a means"—that is, improving the index may become an end in itself, regardless of actual outcomes. This highlights the need to construct a composite index that properly accounts for indicator values while ensuring robustness to small data variations, and aggregates them in as endogenous manner as possible—that is, independently of subjective external factors.

In this paper, we propose a composite index defined, for each unit as the mean of latent variables underlying the individual indicators of that unit, penalised by their dispersion from the mean (i.e., the correlation), see for details, Mariani et al. (2024), Mazziotta and Pareto (2018), and Polinesi et al. (2026). The latent performance is captured by a Beta-Bernoulli conjugate model, where the latent variable is Bernoulli-distributed with a Beta prior on the probability of success. The observed values of the success probabilities are obtained by scaling the original indicator through min-max goalposts defined by policymaking experts according to the scenario of interest. The sequence of the latent variables follows a Beta-Bernoulli process and contributes to measuring the overall performance of the unit.

In doing so, we provide a fully endogenous aggregation method that ensures robustness to small data variations, penalises highly correlated performances, and embeds expert opinions through the specification of goalposts chosen independently of the units under consideration.

Our paper is closely related to the recent work of Chen et al. (2023), which represents one of the few attempts to incorporate expert opinions into a composite index. In that paper, the authors propose a composite index defined as a linear combination of individual indicators, with weights that incorporate expert opinions, computed by solving an optimisation problem for principal components with an additive penalty term that reflects expert opinions and their reliability. Along this line, Ravanos and Karagiannis (2021) combine the Benefit of Doubt model with Value Efficiency Analysis, encoding external preferences and value judgments through a "model" decision-making unit that serves as a benchmark in the Benefit of Doubt weighting scheme.

In line with these contributions, we incorporate expert opinions but in a fundamentally different manner under mild assumptions. Specifically, for each individual component indicator, alternative expert beliefs are considered by modifying the goalposts of the indicator's range (minimum and maximum values), chosen independently of the units under analysis. The individual indicators are then scaled using these pairs of goalposts to obtain scores in the interval $[0, 1]$. The scores

associated with a unit are assumed to be realisations of a Beta distribution with unit-dependent parameters.

In our approach, the only exogenous element in normalisation/weighting/aggregation methodology concerns the specification of the goalposts (i.e., the expert opinions), whereas the remaining components of the approach are entirely endogenous.

Summarising the proposed methodology offers several advantages. First, the composite index accommodates min-max ranges that reflect different expert opinions at a given time for cross-sectional data and across different time periods for panel data. Second, by varying the min-max ranges, we can obtain the sample distribution of the composite index, enabling comparison of how alternative policy specifications of the goalposts affect the composite index. Third, the weighting scheme is endogenously determined, with the importance (weight) of each indicator based on its actual performance rather than on any a priori exogenous choices.

To illustrate the proposed Beta-Bernoulli composite index, we apply it to compute composite indices that measure and compare the financial sustainability of a set of small and medium-sized enterprises over the period 2020–2022.

Two empirical analyses are conducted. The first shows that the composite index captures findings about enterprise performance already established in the liability theory of Fichman and Levinthal (1991) regarding the "honeymoon period" and of Lukason et al. (2016) concerning the need for micro firms to overcome post-foundation challenges to survive.

The second analysis focuses on the effect of enterprise objectives expressed through suitable pairs of goalposts. Empirical analysis shows that increasing the maximum goalposts leads to higher composite index values only up to a certain point (corresponding to an increase of 35%), after which the values decrease. This finding is consistent with the established principle that "overly ambitious demands can be discouraging, whereas modest ones may not capture attention effectively." This concept is supported by the Yerkes-Dodson Law (Yerkes & Dodson, 1908).

The remainder of this paper is organised as follows. Section 2 introduces the proposed methodology, consisting of the scaling approach to embed expert opinions and the proposed composite index. Section 3 illustrates the empirical results derived from applying the approach to a set of small and medium enterprises. Finally, Section 4 presents concluding remarks while the online Appendix presents several robustness checks for the results of Section 3.

2. Methodology

Our aim is to develop a methodology that can combine K indicators, describing individual characteristics of a population of interest, into a single one-dimensional construct: a composite index that can adequately summarise and rank the N units of the population.

For example, the indicators could include income, health, and education levels measured across N regions of a country to assess and compare quality of life. At the company level, indicators might comprise financial ratios measured for N companies to evaluate financial sustainability. Finally, at the individual level, indicators could describe the mathematical, programming, and communication skills of N graduates to assess and compare their potential for data science positions.

In all the described examples, it seems intuitive to consider individual indicators that are standardised, that is, expressed in a similar scale; and that have a positive polarity, that is, having a positive effect on the composite index: the higher the level of each indicator, the higher the level of the composite index.

From a mathematical viewpoint, standardisation and positive polarity can be achieved by means of a linear transformation of the indicators in the set of positive real numbers, that will be described in the next subsection. In the following subsection we will then introduce a probabilistic model to build a composite index.

2.1. Scaling the individual indicators

In this subsection, we introduce a linear transformation of the indicators[1] to standardise them and ensure they all have positive polarity.

Denote the observed value of the j-th indicator for the i-th unit as y_j^i, for $j = 1, 2, ..., K$, and $i = 1, 2, ..., N$.

The most common way to obtain standardised indicators with positive polarity is to subtract each indicator's minimum value from all its values and divide by its range, as follows:

$$p_{j,0}^i = \frac{y_j^i - m_j}{M_j - m_j}, \tag{1}$$

where $m_j = \text{minimum}\left\{y_j^1, y_j^2, ..., y_j^N\right\}$ and $M_j = \text{maximum}\left\{y_j^1, y_j^2, ..., y_j^N\right\}$. In this way we obtain a transformed set of K indicators, $p_{j,0}$, all defined in $(0, 1)$.

The "classic" transformation in Equation (1) is convenient, as all transformed indicators assume values in the same unit range. There may be situations, however, in which the normalisation in

Equation (1) is "too strong." For example, when comparing different regions in terms of quality of life, we may have regions with much larger values, leading to a normalisation that does not allow for comparing regions with small values among each other. When comparing different companies in terms of financial sustainability, there may be companies with very low values of all ratios and, in this case, it becomes difficult to distinguish among the remaining companies. Finally, when comparing individual skills for a data science position, we may want to focus the comparison on the top profiles, but their normalised values may be very close to each other if the normalisation is carried out using the minimum and the maximum.

These considerations suggest generalising the classic transformation in Equation (1). Consider different expert opinions embedded in R alternative specifications. For each expert opinion, $r = 1, 2, \ldots, R$, replace m_j and M_j in Equation (1) with the following pair of "goalposts":

$$m_{j,r} = m_j(1 - \Delta_{m,r}), \tag{2}$$

$$M_{j,r} = M_j(1 + \Delta_{M,r}), \tag{3}$$

where m_j, M_j are positive while $\Delta_{m,r}$ and $\Delta_{M,r}$ are non-negative constants, whose values range between zero and one. Note that, when $\Delta_{m,r}$ and $\Delta_{M,r}$ are both equal to zero, the goalposts simplify to the minimum and the maximum observation, as in Equation (1).

The goalposts in Equation (2) and (3) are employed to obtain different transformations of the individual indicators, that guarantee values in $[0, 1]$ and positive polarity while allowing for a more flexible range definition, specified in terms of expert opinions, $r = 1, 2, \ldots, R$, for each indicator $j = 1, 2, \ldots, K$, and unit $i = 1, 2, \ldots, N$, as follows:

$$p^i_{j,r} = \frac{y^i_j - m_{j,r}}{M_{j,r} - m_{j,r}}. \tag{4}$$

The following proposition shows when the transformed values $p^i_{j,r}$ are higher (or lower) than those obtained with the classic transformation $p^i_{j,0}$ in Equation (1).

Proposition 2.1. Let $p^i_{j,0}$ be the classic min-max transformation in Equation (1). If $y^i_j > \frac{m_j M_j(\Delta_{m,r}+\Delta_{M,r})}{M_j\Delta_{M,r}+m_j\Delta_{m,r}}$ then $p^i_{j,r} < p^i_{j,0}$; otherwise, if $y^i_j \leq \frac{m_j M_j(\Delta_{m,r}+\Delta_{M,r})}{M_j\Delta_{M,r}+m_j\Delta_{m,r}}$ then $p^i_{j,r} \geq p^i_{j,0}$

Proof. The proof is obtained through a simple computation by rewriting $p^i_{j,r} < p^i_{j,0}$ as follows:

$$\frac{y^i_j - m_j(1 - \Delta_{m,r})}{M_j(1 + \Delta_{M,r}) - m_j(1 - \Delta_{m,r})} < \frac{y^i_j - m_j}{(M_j - m_j)}.$$

□

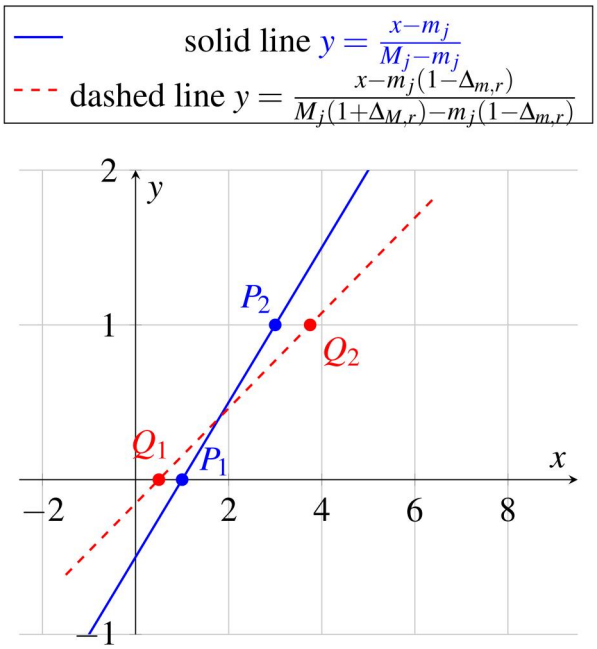


Figure 1. Graphical representation of the scaling in the (x, y) plane. The x-axis shows the values of the j-th indicator while the y-axis is the scaled indicator. The blue-solid line is the scaling corresponding to $\Delta_m = 0 = \Delta_M = 0$ while the red-dashed line is the scaling with a non-zero goalposts. The points have the following coordinates: $P_1 = (m_j; 0)$, $P_2 = (M_j; 1)$, $Q_1 = (m_j(1 - \Delta_{m,r}); 0)$ and $Q_2 = M_j(1 + \Delta_{M,r}); 1)$.

Figure 1 illustrates Proposition 2.1 from a graphical viewpoint. It shows the straight lines representing the transformed values for $\Delta_{m,r} = \Delta_{M,r} = 0$ (solid blue line) and for $\Delta_{m,r} = 0.5$, $\Delta_{M,r} = 0.25$ (dashed red line). The value $\overline{x}^{j,r} = \frac{m_j M_j(\Delta_{m,r}+\Delta_{M,r})}{M_j\Delta_{M,r}+m_j\Delta_{m,r}}$ corresponds to the x-value where the two lines intersect.

Figure 1 shows that, when $\Delta_{m,r}$ and $\Delta_{M,r}$ are greater than zero, the slope of the dashed line becomes lower than the slope of the solid line. The difference between the two slopes increases when $\Delta_{M,r}$ and $\Delta_{m,r}$ increase.

From an interpretational viewpoint, goalposts can be interpreted in terms of decision making. When the upper goalpost $(\Delta_{M,r})$ is greater than zero, the decision maker sets a limit for each indicator that is higher than the maximum ever reached by the units under study. This reduces the probability that a given unit achieves a high performance.

Conversely, when the lower goalpost $(\Delta_{m,r})$ is greater than zero, the decision maker sets a limit for each indicator that is lower than the minimum ever reached by the considered units. This increases the probability that a given unit achieves a high performance.

The advantage of using goalposts different from the minimum and the maximum allows to evaluate the effects of alternative policy targets, set by the decision makers.

We remark that the transformation described above can be usefully extended to panel data. In that context, each individual indicator is usually scaled using the smallest and largest values across different years. However, as demonstrated by the COVID-19 pandemic, exogenous events can generate outliers that may distort the scaling, flattening the importance of time periods in which extreme values were not observed, thereby compromising the reliability of composite indices. The transformation in Equation (4) can address this issue selecting, for each indicator j, the intervals $[m_{j,r}; M_{j,r}]$ as the annual, monthly, or daily min-max ranges, where the values of r, $r = 1, 2, ..., R$, represent the years, months, or days considered, as illustrated in Section 3.2.2.

2.2. Building a composite index

In this subsection, we define a probabilistic model that leads to the definition of a composite index as a function of the scaled indicators introduced in the previous subsection. To achieve this, we recall the Beta distribution and Beta-Bernoulli process.

A Beta distribution for random proportion P, with left parameter $\alpha > 0$ and right parameter $\beta > 0$ has a probability density function f given by:

$$f(p; \alpha, \beta) = \frac{p^{\alpha-1}(1-p)^{\beta-1}}{B(\alpha, \beta)}, \quad p \in (0, 1), \tag{5}$$

where $B(\alpha, \beta)$ is a normalising constant, B is the Beta function, $B(\alpha, \beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$, and Γ the Gamma function. For the Beta distribution, the mean and variance are given by $E(P) = \frac{\alpha}{\alpha+\beta}$, and $\text{var}(P) = \frac{\alpha\ \beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$, respectively.

The Beta-Bernoulli process is as follows.

Definition 2.1. Siegrist (2022) Let P be a random variable drawn from a Beta distribution with left and right parameters α and β, respectively. The Beta-Bernoulli process of parameters α and β is a sequence of binary (0–1) random variables, denoted by $\underline{X} = (X_1, X_2, ...)$, with the property that, conditional on $P = p$, the sequence $\underline{X}$ is a sequence of independent Bernoulli trials such that:

$$X_l | P = p \sim \text{Bernoulli}(p), l = 1, 2, ..., L, \tag{6}$$

or equivalently:

$$P(X_l = 1 | P = p) = p, l = 1, 2, ..., L. \tag{7}$$

In other words, once a value p of P is drawn from the Beta distribution, the process $\underline{X}$ is a sequence of Bernoulli trials with probability of success $P = p$.

In our context, we assume that the distribution of "successes" across individual indicators, for a set of N units, are N Beta-Bernoulli processes. Thus, different goalposts assigned by R different experts generate $R \times N$ replications of the indicators.

In this way, each of the K indicators, for a given unit i, and an expert opinion r, can be a "success" with a common probability of success $P = p$, unknown, and to be learned from the observations. What is observed, for a given unit i, and an expert opinion r, are a set of K realisations of the Beta-Bernoulli random variables or the set of the K probabilities of success defined in Equation (4).

Before proceeding, it is useful to recall the properties of a Beta-Bernoulli process $\underline{X}$. It can be shown that:

$$E(X_l) = \frac{\alpha}{\alpha+\beta}, \quad l = 1, 2, ..., L, \tag{8}$$

$$\text{var}(X_l) = \frac{\alpha}{\alpha+\beta} \cdot \frac{\beta}{\alpha+\beta}, \quad l = 1, 2, ..., L, \tag{9}$$

$$\text{cov}(X_l, X_s) = \frac{\alpha}{\alpha+\beta} \cdot \frac{\beta}{\alpha+\beta} \cdot \frac{1}{\alpha+\beta+1}, \quad l \neq s, \quad l, s = 1, 2, ..., L, \tag{10}$$

$$\text{cor}(X_l, X_s) = \frac{1}{\alpha+\beta+1}, l \neq s, \quad l, s = 1, 2, ..., L, \tag{11}$$

where $E(\cdot)$ denotes the expectation, $\text{var}(\cdot)$ the variance, $\text{cov}(\cdot, \cdot)$ the covariance, and $\text{cor}(\cdot, \cdot)$ the correlation.

Remark 2.1. Note that $\alpha + \beta$ represents the concentration of the random variable P and measures the reliability of the mean. Specifically, as the concentration goes to infinity and the ratio $\frac{\alpha}{\alpha+\beta}$ approaches a constant, the Beta distribution reduces to a Dirac delta distribution; in this limit, P becomes deterministic and equal to this constant. Therefore, the correlation in Equation (11) is an inverse measure of concentration: the higher the correlation, the higher the dispersion from the mean; conversely, lower correlation implies higher concentration around the mean.

From the above properties, we note several key features. First, $E(X_l)$ equals the mean of the Beta distribution of P. Second, the variance is a function of the mean. Third, the covariance equals the variance multiplied by a correction factor. Finally, the correlation is positive, which relates to the fact that the process is exchangeable due to the expression of the joint probability mass function of the Beta-Bernoulli process:

$$P(X_1 = x_1, X_2 = x_2, ..., X_L = x_L) = \frac{\alpha^{\sum_{l=1}^{L} x_l} \beta^{L - \sum_{l=1}^{L} x_l}}{(\alpha+\beta)^L} \tag{12}$$

which depends only on the sum of the observations $x_1, x_2, ..., x_L$, and not on their order.

All of the above moments (see, Equations (8)–(11)) are functions of the two Beta distribution parameters, α and β, and can therefore be estimated once these parameters are inferred from the data.

We associate with each unit i a vector of success performances $\underline{X}^i = (X_1^i, X_2^i, ..., X_K^i)$, where X_j^i represents the impact of the $j-th$ indicator on unit i. Specifically, $X_j^i = 1$ indicates the best possible performance for that indicator, while $X_j^i = 0$ indicates the worst.

Thus, we assume that the vector of the success performances $\underline{X}^i = (X_1^i, X_2^i, ..., X_K^i) \sim BBeta(\alpha_i, \beta_i)$, where *BBeta* indicates a Beta-Bernoulli process with success probability $P^i \sim Beta(\alpha_i, \beta_i)$[2].

For each unit i, the scaled values $p_{j,r}^i$ in Equation (4) are assumed to be realisations of the probability of success P^i.

We now define a composite index named Beta-Bernoulli Composite index (BBC for short) for the unit i as follows:

$$I_i = E(X^i)\left(1 - \frac{1}{K-1}\sum_{k\neq j} \mathrm{cor}(X_j^i, X_k^i)\right), \quad \forall i = 1, ..., N. \tag{13}$$

The BBC index is a composite index defined as the mean of the Beta-Bernoulli process that defines the phenomenon, penalised by the correlation among the random Beta-Bernoulli variables. The penalty is a function of the indicators' correlation, i.e., a dispersion from the mean, and, as a consequence, units with lower correlation among indicators (i.e., the success performances) are preferred because they lead to a higher value of the composite index.

The correlation among indicators plays a crucial role because it reduces the impact of extreme events, such as COVID-19. When such events occur, all indicators become highly correlated, and therefore a composite index that is not corrected becomes biased. Our composite index, in contrast, incorporates a correction that causes it to decrease when extreme events increase the correlation.

We now prove some useful properties of the proposed composite index.

Proposition 2.2. Equation (13) can then be rewritten as follows:

$$I_i = \mu_i\left(1 - \frac{1}{\nu_i + 1}\right), \quad \forall\ i = 1, ..., N, \tag{14}$$

where μ_i is the mean of P^i and ν_i represents its concentration, given by

$$\mu_i = \frac{\alpha_i}{\alpha_i + \beta_i}, \quad \nu_i = \alpha_i + \beta_i, \quad \forall i = 1, ..., N. \tag{15}$$

Proof. From Definition 2.1 we have:

$$E(X_j^i | P^i) = P^i. \tag{16}$$

Therefore, from Equation (16) using the law of iterative expectations, we have:

$$E(X_j^i) = E(E(X_j^i | P^i)) = E(P^i) = \frac{\alpha_i}{\alpha_i + \beta_i} = \mu_i. \tag{17}$$

Bearing in mind that according to Equation (11)

$$\frac{1}{K-1}\sum_{k\neq j} cor(X_j^i, X_k^i) = \frac{1}{K-1}\sum_{k\neq j}\frac{1}{\alpha_i + \beta_i + 1} = \frac{1}{\nu_i + 1}. \tag{18}$$

This concludes the proof. □

Note that, when ν_i approaches infinity, the beta distribution concentrates at its mean μ_i. When $\mu_i = 1$ and $\nu_i \to +\infty$, unit i is successful in terms of all indicators, while when $\mu_i = 0$ and $\nu_i \to +\infty$, unit i is always unsuccessful.

More generally, the mean μ_i can be interpreted as the mean performance of unit i (i.e., its mean probability of success, across all indicators), while the concentration ν_i can be interpreted as a reliability measure for this mean performance: the higher the concentration, the higher the reliability.

Based on Proposition 2.2, we summarise the assumptions underlying the composite index: (1) distributional: each unit's indicator is assumed to follow a Beta distribution on $[0, 1]$, with parameters $\alpha_i > 0$ and $\beta_i > 0$ estimated *via* maximum likelihood; (2) monotonicity conditional on concentration: higher values of each indicator reflect better performance only when the concentration parameter $\nu_i = \hat{\alpha}_i + \hat{\beta}_i$ is held constant across units. As a direct consequence, smaller concentration (i.e., higher dispersion from the mean) penalises the index, so that two units with the same mean $\hat{\alpha}_i/(\hat{\alpha}_i + \hat{\beta}_i)$ will be ranked differently if their concentration differs; (3) comparability: units can be meaningfully ranked against one another on a common scale; (4) positive polarity: all indicators are oriented so that larger values are preferable; (5) statistical reliability: a sufficiently large number of observations per unit is required to ensure the stability of the MLE estimates $\hat{\alpha}_i$ and $\hat{\beta}_i$.

These assumptions depart from those of other composite indices in several respects. Methods such as the Benefit of the Doubt (BoD) and Data Envelopment Analysis (DEA) are non-parametric and assign unit-specific weights that are most favourable to each unit, requiring no assumption on polarity beyond the choice of outputs versus inputs, and no distributional assumption whatsoever. However, they do not penalise units for

inconsistency or volatility in their indicator scores, and the resulting rankings can be sensitive to the choice of the frontier estimation methodology. Principal Component Analysis (PCA) based composite index instead extracts orthogonal linear combinations of the indicators that maximise explained variance, implicitly assuming that indicators with higher variance carry more information; it requires no assumption on concentration, but the resulting composite indices are often difficult to interpret substantively and polarity must be verified post hoc. The PCA-based index is sensitive to scale, so indicators are typically standardised before running it, which itself is an assumption about equal weighting of variance across variables.

Proposition 2.3. The index in Equations (13) and (14) belongs to the interval $[0,1]$, with 1 reached when all indicators are successful with probability one, and 0 when all indicators are not successful with the same probability.

Proof. Recall that $\mu_i = \frac{\alpha_h}{\alpha_i+\beta_i}$, with α_i and β_i being positive values. Therefore, it follows that $\mu_i \in (0,1)$. The endpoints can be approached as limits when the parameters approach zero or infinity. The quantity $\frac{1}{1+\alpha_i+\beta_i}$ is positive and less than one, so $1-\frac{1}{1+\alpha_i+\beta_i}$ is also positive and smaller than one. When $\alpha_i \to +\infty$, the index approaches 1 for any value of β_i. In contrast, when $\alpha_i \to 0^+$ and $\beta_i \to +\infty$, the index approaches 0, indicating that all distributions are concentrated at zero. □

The composite index I_i in Equation (14) is a theoretical quantity that depends on the unknown parameters of the Beta Bernoulli distribution, α_i and β_i. These parameters are estimated from the available data, which for a given unit i consists of the K scaled values in Equation (4) for each expert evaluation r, or alternatively, the observed values of the indicator random variables. In this regard, the following properties hold.

Proposition 2.4. A maximum likelihood estimator for $\widehat{I}_i$ can be obtained as:

$$\widehat{I}_i = \left(\frac{\hat{\alpha}_i}{\hat{\alpha}_i+\hat{\beta}_i}\right)\left(1-\frac{1}{\hat{\alpha}_i+\hat{\beta}_i+1}\right), \quad i=1,2,\ldots,N. \tag{19}$$

where the quantities $\hat{\alpha}_i$ and $\hat{\beta}_i$ are the maximum likelihood estimates of the parameters α_i and β_i.

Recalling that the mean and the variance of P^i are given by $\mu_i = \frac{\alpha_i}{\alpha_i+\beta_i}$ and $\mathrm{var}_i = \frac{\alpha_i\beta_i}{(\alpha_i+\beta_i)^2(\alpha_i+\beta_i+1)}$, respectively, and nothing that the ratio $\mathrm{var}_i/(\mu_i(1-\mu_i)) = 1/(\alpha_i+\beta_i+1)$, the composite index (Equation (14)) can be equivalently rewritten as:

$$\hat{I}_i = \overline{P}_i\left(1-\frac{\widehat{\mathrm{var}}_i}{\overline{P}_i(1-\overline{P}_i)}\right), \tag{20}$$

where $\overline{P}_i$ is the sample mean of P^i:

$$\overline{P}_i = \frac{1}{KR}\sum_{j=1}^{K}\left(\sum_{r=1}^{R} p^i_{j,r}\right), i=1,2,\ldots,N, \tag{21}$$

and $\widehat{\mathrm{var}}_i$ is the sample variance of P^i:

$$\widehat{\mathrm{var}}_i = \frac{1}{(KR-1)}\sum_{j=1}^{K}\left(\sum_{r=1}^{R}(p^i_{j,r}-\overline{P}_i)^2\right), \quad i=1,2,\ldots,N. \tag{22}$$

It is important to determine the marginal effect of each indicator on the composite index. This is derived in the following proposition.

Proposition 2.5. The marginal effect of the j-th indicator on the estimated composite index $\hat{I}_i$ is given by:

$$\frac{\partial \hat{I}_i}{\partial y^i_j} = \frac{1}{KS_j}\left\{\left[1-\frac{\widehat{\mathrm{var}}_i}{(1-\overline{P}_i)^2}\right] - \frac{2RK}{(RK-1)}\frac{1}{(1-\overline{P}_i)}\sum_{r=1}^{R}\left[w_{j,r}(p^i_{j,r}-\overline{P}_i)\right]\right\} \tag{23}$$

where S_j is the harmonic mean of the interval sizes of the j-th indicator:

$$S_j = \left[\frac{1}{R}\sum_{r=1}^{R}\frac{1}{(M_{j,r}-m_{j,r})}\right]^{-1}, \tag{24}$$

while $w_{j,r}$, $r=1,2,\ldots,R$, are positive weights summing one:

$$w_{j,r} = \frac{1}{R}\frac{S_j}{M_{j,r}-m_{j,r}}. \tag{25}$$

Proof. The proof follows using Equations (4), (21), and (22) with a straightforward computation:

$$\begin{aligned}\frac{\partial \hat{I}_i}{\partial y^i_j} &= \sum_{r=1}^{R}\frac{\partial p^i_{j,r}}{\partial y^i_j}\frac{\partial \hat{I}_i}{\partial p^i_{j,r}} = \sum_{r=1}^{R}\left[\frac{1}{M_{j,r}-m_{j,r}}\right] \\ &\quad\left\{\left[\frac{1}{RK}-\frac{1}{RK}\frac{\widehat{\mathrm{var}}_i}{(1-\overline{P}_i)^2}\right]-2\frac{(p^i_{j,r}-\overline{P}_i)}{(1-\overline{P}_i)}\right\} \\ &= \frac{1}{KS_j}\left[1-\frac{\widehat{\mathrm{var}}_i}{(1-\overline{P}_i)^2}\right]-\frac{2}{(RK-1)}\frac{1}{(1-\overline{P}_i)} \\ &\quad\sum_{r=1}^{R}\left[\frac{(p^i_{j,r}-\overline{P}_i)}{(M_{j,r}-m_{j,r})}\right]\end{aligned}$$

$$= \frac{1}{KS_j}\left[1 - \frac{\widehat{\mathrm{var}}_i}{(1-\overline{P}_i)^2}\right] - \frac{2}{(RK-1)}\frac{1}{(1-\overline{P}_i)} \sum_{r=1}^{R}\left[\frac{y_j^i - m_{j,r} - \overline{P}_i(M_{j,r} - m_{j,r})}{(M_{j,r} - m_{j,r})^2}\right]. \quad (26)$$

This concludes the proof. □

The marginal effect described in Equation (23) provides several important insights. First, the smaller the mean range of indicator j, denoted S_j, the greater its effect. This occurs because the inverse of the harmonic mean of the range acts as a multiplicative factor. We recall that the range of a given indicator embeds expert opinion and is independent of the specific unit. A decision maker can thus affect the relative importance of each indicator by varying the goalposts, which determine the range.

Second, the marginal effect is linear in the deviation of the j-th indicator from the sample mean $\overline{P}_i$. Specifically, the term $\sum_{r=1}^{R}\left[w_{j,r}(p_{j,r}^i - \overline{P}_i)\right]$ represents a weighted average of the deviations of the observed performances $p_{j,r}^i$ from the average overall performance $\overline{P}_i$. This means that greater importance is assigned to deviations of observations associated with smaller ranges. A decision maker can thus affect the relative importance of an extreme deviation: the smaller the range, the larger the impact. While in terms of the previous comment a small range is beneficial, in this case it is not.

In this perspective, the expert specification of $\Delta_{M,r}$ and $\Delta_{m,r}$ and, consequently, of the two goalposts, is a key aspect, as it affects the range. The net effect of the specification depends on the examined data, and will be examined in the next Section.

Note also that, bearing in mind (Equations (4) and (23)) can be expressed as:

$$\frac{\partial \hat{I}_i}{\partial y_j^i} = \frac{1}{KS_j}\left\{\left[1 - \frac{\widehat{\mathrm{var}}_i}{(1-\overline{P}_i)^2}\right] - \frac{2RK}{(RK-1)}\frac{S_j}{(1-\overline{P}_i)} \sum_{r=1}^{R}\left[\frac{1}{(M_{j,r} - m_{j,r})^2}(y_j^i - m_{j,r} - \overline{P}_i(M_{j,r} - m_{j,r}))\right]\right\}. \quad (27)$$

From Equation (27) note that, when the range S_j is small, the marginal effect of an indicator is driven by the term $\left[1 - \frac{\widehat{\mathrm{var}}_i}{(1-\overline{P}_i)^2}\right]$ which depends on the unit but not on the specific indicator. This indicates that the marginal effect has a weak dependence on the value of the indicators.

As we have already remarked, what matters for the composite index are the mean value of the indicators and their correlations, rather then the specific values of the indicators.

We finally remark that, in a panel data setting, the harmonic mean S_j of the range can be interpreted as the average variability of the j-th indicator over the considered period. The choice of different periods, or time frequencies, in which to calculate the composite indices can thus be seen, in a panel data setting, as an expert based strategy similar to the choice of goalposts.

3. Application

3.1. The composite index

In this subsection we illustrate the effectiveness of the proposed composite index, employing as a case study a set of 239 manufacturing companies belonging to the ATECO code 15 of the Marche region, whose financial performance data for the years 2020–2022 are organised in a publicly available database, at https://www.leanus.it/. We, we focus on financial performance as the primary metric.

Manufacturing companies will be analysed separately based on both age (year founded) and size (number of employees). Specifically, we divide companies in "old firms" and "young firms" with the latter including those founded after 2010. Each group is further divided into "micro size" (less than 10 workers), "small size" (10 to 49 workers), and "medium-large size" (50 or more workers). Thus, the analysis includes 33 medium-large firms (among which 26 old and 7 young), 161 small firms (118 old and 43 young), and 45 micro firms (26 old and 19 young).

The database contains ten individual component indicators that concern different key aspects of company performance[3]. The list of all indicators is provided in Table 1.

From Table 1 note that the indicators can be grouped in three main categories: (i) Profitability ratios, which include Return on Assets (ROA), Return on Equity (ROE), Return on Sales (ROS), Return on Investment (ROI), and margins; (ii) Liquidity ratios, such as the Current Assets to Current Liabilities (CACL) ratio, and the Cash (or quick assets) to Current Liabilities (CCL) ratio; (iii) Financial leverage ratios, which include the total debt to total equity ratio and the percentage of total long-term debts.

To construct the composite index defined in Equation (19), we first define the samples on which its replications will be based. To this end, we set $\Delta_{m,r}$ at 0 and $\Delta_{M,r}$, $r = 1, 2, \ldots, R$, as an ordered sequence of R distinct numbers within the interval $[0, 1]$, reflecting various expert opinions. Specifically, we consider $\Delta_{M,r} = (r-1) \times 0.05, r = 1, 2, \ldots, R$, where $R = 21$.

Table 1. List of variables.

Category	Indicators
Profitability	P (or L)/TA (ROA), P (or L)/TE (ROE), EBIT/S (ROS), EBIT/TA (ROI), GP/S, EBITDA
Liquidity	CA/CP (CACL), QA/CL (CCL)
Financial leverage	TD/TE (Debt Ratio Leverage), % LD

Note: the acronyms have the following meanings: P (or L), balance profit or loss; TA, total assets; CA, current assets; QA, quick assets; TD, total debt; LD, total long-term debt; CL, current liabilities; TE, total equity; RE, retained earnings; S, sales revenue; GP, gross profit; EBIT, earnings before interest and taxes; EBITDA, earnings before interest, taxes, amortisation and depreciation.

Table 2. Summary statistics of the composite index for the years 2020 (a), 2021 (b) and 2022 (c).

(a)

	Large-medium size		Small size		Micro size	
	Old	Young	Old	Young	Old	Young
Min	0.3325	0.3047	0.3268	0.3072	0.3376	0.3150
1st Qu.	0.3411	0.3100	0.3347	0.3123	0.3459	0.3206
Median	0.3480	0.3165	0.3411	0.3172	0.3533	0.3274
Mean	0.3470	0.3152	0.3398	0.3165	0.3520	0.3258
3rd Qu.	0.3542	0.3207	0.3465	0.3213	0.3594	0.3310
Max	0.3569	0.3223	0.3486	0.3229	0.3618	0.3331

(b)

	Big-medium size		Small size		Micro size	
	Old	Young	Old	Young	Old	Young
Min	0.3053	0.2789	0.2940	0.2713	0.3069	0.2786
1st Qu.	0.3158	0.2880	0.3026	0.2785	0.3170	0.2866
Median	0.3274	0.2982	0.3125	0.2863	0.3282	0.2954
Mean	0.3282	0.2986	0.3130	0.2864	0.3288	0.2954
3rd Qu.	0.3404	0.3091	0.3232	0.2948	0.3405	0.3043
Max	0.3524	0.3191	0.3335	0.2999	0.3512	0.3109

(c)

	Large-medium size		Small size		Micro size	
	Old	Young	Old	Young	Old	Young
Min	0.3709	0.3465	0.3607	0.3412	0.3727	0.3534
1st Qu.	0.4038	0.3757	0.3926	0.3714	0.4103	0.3927
Median	0.4115	0.3818	0.3996	0.3779	0.4177	0.4003
Mean	0.4083	0.3797	0.3974	0.3756	0.4147	0.3971
3rd Qu.	0.4170	0.3873	0.4058	0.3828	0.4241	0.4056
Max	0.4193	0.3890	0.4077	0.3844	0.4264	0.4076

We then compute the ten scaled indicators using Equation (4) for each of the $R = 21$ expert settings and all 239 examined companies[4].

The summary statistics of the proposed composite index by different types of firms are reported in Table 2.

Table 2 shows that, on average, micro firms exhibit higher performance across the considered indicators compared to small and medium-sized firms, regardless of their age. Since we are analysing young but financially mature firms, this finding aligns with Lukason et al. (2016), who demonstrate that micro firms must overcome post-foundation challenges to survive. This also supports the "honeymoon period" of normal financial performance described by the liability theory of Fichman and Levinthal (1991). This trend is also evident in Figure 2, where the right panel shows that micro firms have higher composite index values than large and small enterprises when considering the young and old categories separately.

Further insights into the comparison across different companies can be obtained by examining the estimated density of the composite index for the six considered groups in Figure 3.

Figure 3 shows that, across different government policies, firms—regardless of size and age—improve their performance in 2022 (blue dotted line), following the negative trend observed in 2021 (red dashed line).

For the year 2022, the bimodal shape of the density also reveals the existence of "two-speed companies," referring to a situation in which businesses within an economic system or industry develop at different rates (i.e., the two modes). This typically highlights a gap between companies that are more advanced and competitive and those struggling to keep pace with new technologies, market changes, or global economic dynamics.

We further investigate the capability of the index to distinguish well-performing from poorly performing firms by clustering them according to the mean value of their composite index. We focus on small firms, both young and old, as they form the backbone of the considered manufacturing sector. Firms are grouped into three contiguous classes based on the first and ninth deciles of the mean values of the composite index. Specifically, older firms are clustered as worst-performing $[0, 0.37]$, middle-performing $(0.37,\ 0.427)$, and best-performing $[0.427, 1]$, while younger firms fall into worst-performing $[0, 0.33]$, middle-performing $(0.33,\ 0.412)$, and best-performing $[0.412, 1]$.

We then assess whether the obtained clusters align with the actual nature of the firms using indicators not included in the composite index within a predictive framework. Figure 4 presents summary plots of two indicators not used in constructing the composite index: the debt-to-equity ratio and total short-term debt (as a percentage), for the worst-, middle-, and best-performing firm clusters.

Figure 4 clearly shows that the best-performing firms have lower levels of total short-term debt and debt-to-equity ratios compared to the middle- and worst-performing firms. This indicates that the

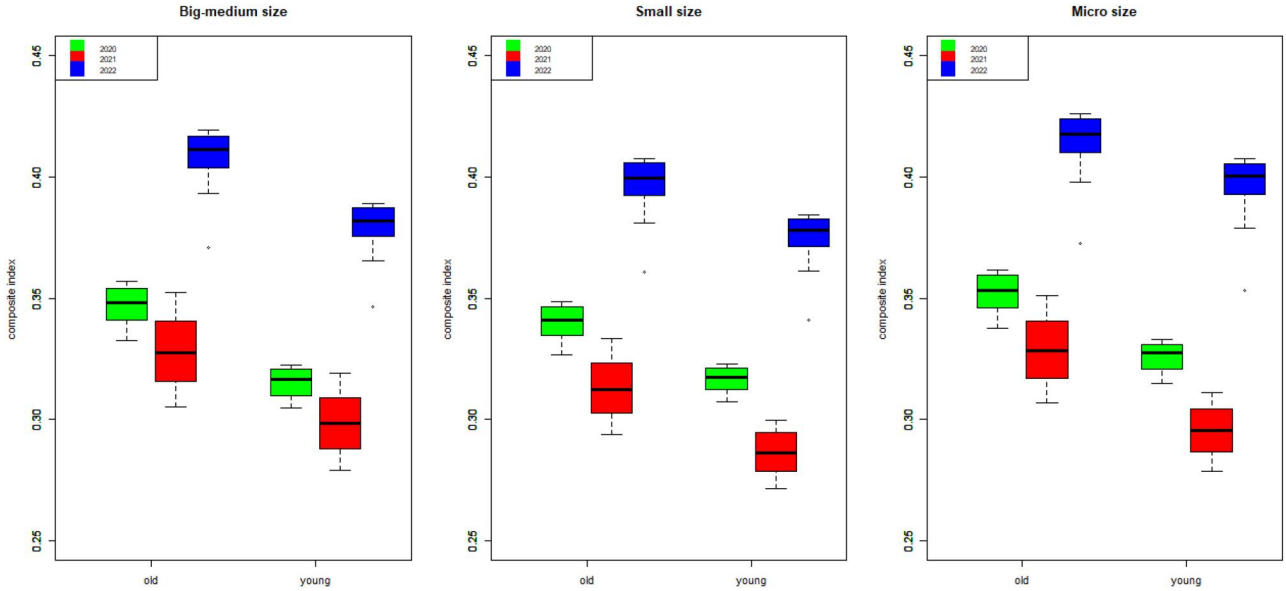


Figure 2. Box plots of the composite index in the specific firm class. Colours refer to different years of the analysis: 2020 (green), 2021 (red), 2022 (blue).

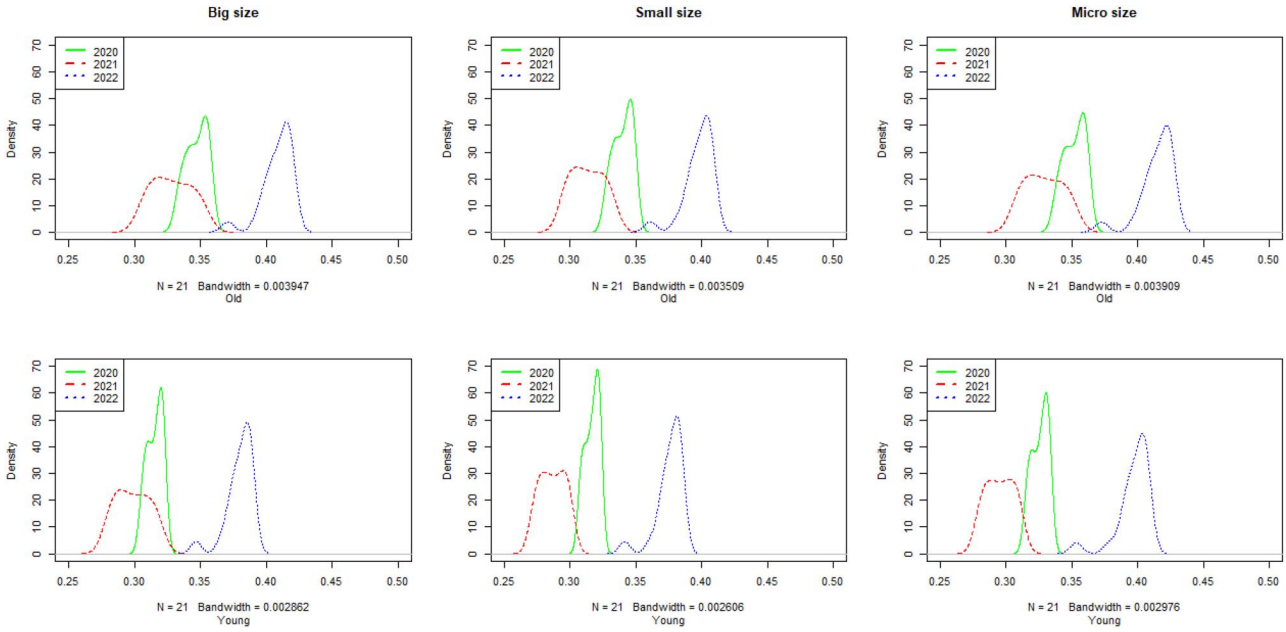


Figure 3. Density of the composite index by groups of firms. Colours and lines refer to different years of the analysis: 2020 (green solid line), 2021 (red dashed line), 2022 (blue dotted line).

composite index is a good predictor of the actual financial state of the firms.

3.2. The impact of expert opinions

In this subsection, we analyse the impact of expert opinions on the composite indices. As previously discussed, expert opinions can affect the probability of success of a unit (a firm in this application) by varying the success performances. It is of interest to understand which expert settings are more beneficial for achieving higher mean values of the composite index. We examine the impact of expert opinions first in a cross-sectional framework and then in a panel context.

3.2.1. Cross-sectional framework

We focus on the year 2022 to analyse the impact of different expert opinions and, specifically, of different choices of Δ_m and Δ_M.

Figure 5 presents four panels, each displaying the value of the composite index in correspondence to the $R = 21$ scaling approaches described in the previous subsection. We focus on the companies belonging to two sub-categories of ATECO code 15 (manufacture of footwear and manufacture of footwear and leather parts), and we also distinguish them by size.

In Figure 5 the horizontal axis represents the pairs $(\Delta_{m,r}, \Delta_{M,r})$, where $\Delta_{m,r} = 0$, and $\Delta_{M,r} = (r-1) \times 0.05$ for $r = 1, 2, ..., 21$. These pairs can be

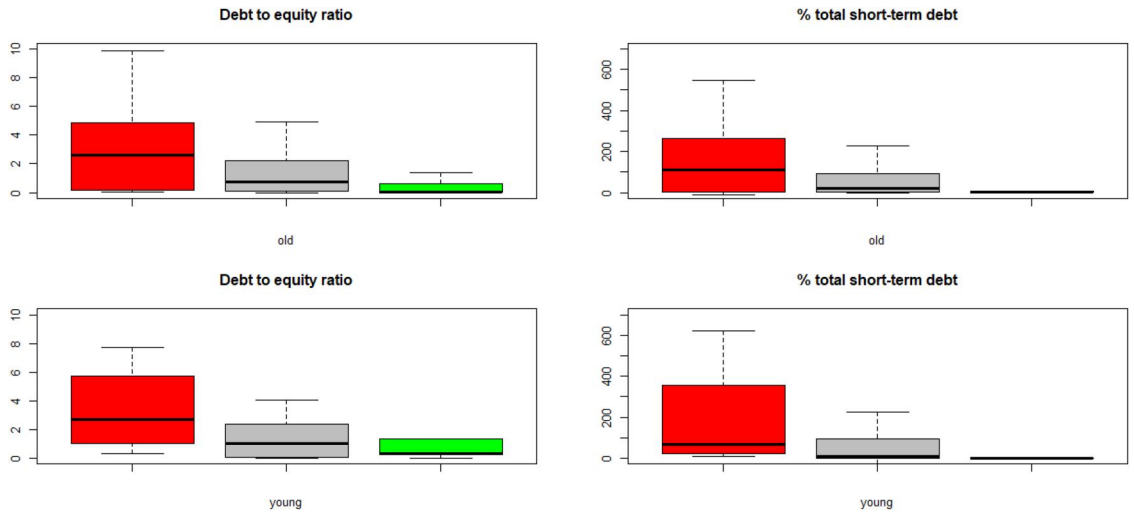


Figure 4. Summary plots of total short-term debt (in percentage—right) and debt to equity ratio (left) for old (upper panel) and young (bottom panel) small firms in 2022. Colours indicate best (green), middle (grey) and worst (red) performing firms.

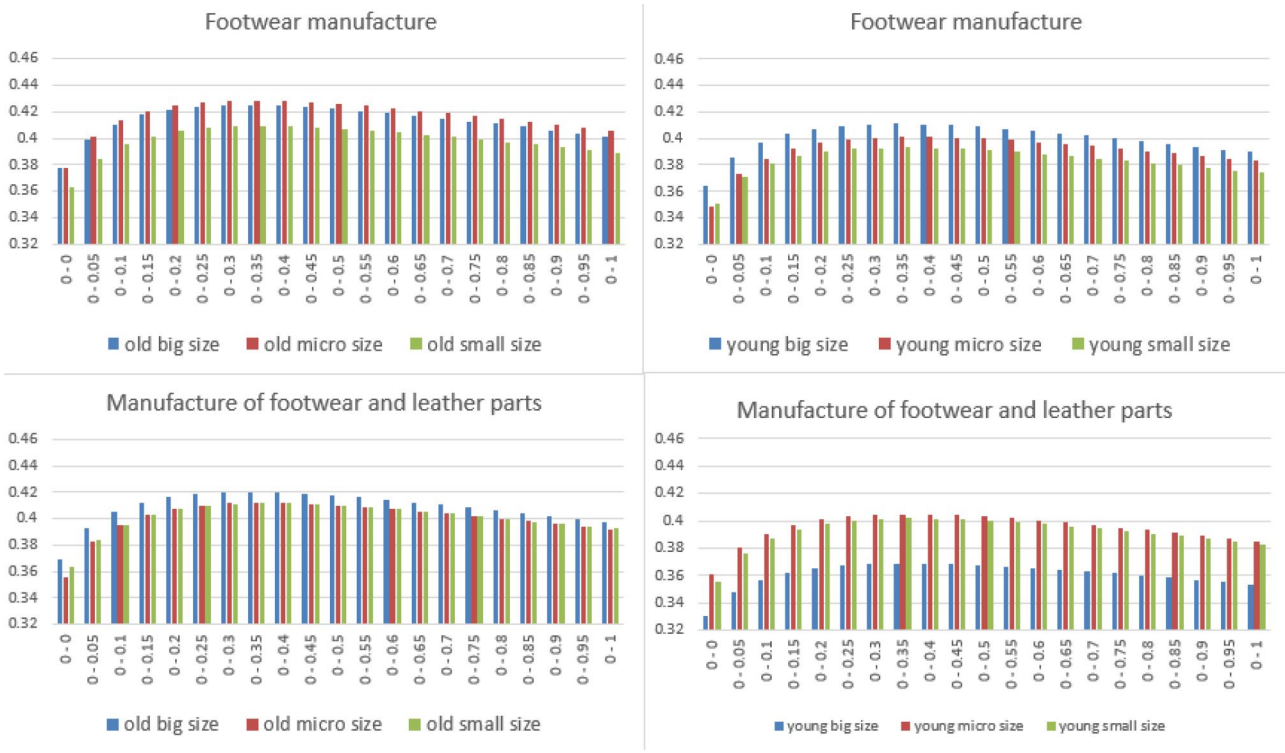


Figure 5. Composite index values for two different ATECO 15 sectors in 2022, divided by firm age and size. Colours represent firm size: small (green), micro (red), and large (blue). The left panel shows old firms, while the right panel displays young firms. The *x*-axis represents the pairs $(\Delta_{m,r}, \Delta_{M,r})$, where $\Delta_{m,r} = 0$, and $\Delta_{M,r} = (r-1) \times 0.05$ for $r = 1, 2, \ldots, 21$. The *y*-axis represents the value of the composite index.

interpreted as targets set by policymakers (such as public authorities or market investors). The choice of $\Delta_{m,r} = 0$ indicates no reduction in the minimum performance, whereas $\Delta_{M,r}$, ranging from 0 to 1, indicates increasing performance improvement over the observed maximum from 0% to 100%.

3.2.2. Panel data framework

Here, we focus on panel data to evaluate whether the company ranking provided by the composite index is robust to different specifications of expert opinions. To achieve this goal, we calculate the composite index for three different years and consider two different approaches for setting expert opinions. The first approach scales each indicator using the minimum and maximum values observed across all three years. The second approach scales each indicator using the minimum and maximum values observed in each specific year. Table 3 shows the obtained composite index for each group of firms.

Table 3 highlights that the ranking of firms is robust: it does not vary when changing expert opinions, with micro firms showing higher values of the

Table 3. Composite index values according to (a) minimum and maximum values computed annually and (b) minimum and maximum values computed across all years.

	Large-medium size		Small size		Micro size	
	Old	Young	Old	Young	Old	Young
(a)	0.3695	0.3384	0.3564	0.3288	0.3700	0.3399
(b)	0.3696	0.3472	0.3623	0.3407	0.3735	0.3549

composite index regardless of their age. Note that using the minimum and maximum values across all years as goalposts results in slightly higher composite index values.

4. Concluding remarks

In this paper, we have proposed a new methodological approach for composite indices that can incorporate expert opinions into their calculation. The proposal offers three key advantages.

First, it provides a composite index rooted in a new statistical framework, that of Beta-Bernoulli processes. This leads to estimated composite index values that are simple to understand, being the mean of the scaled indicators penalised by their correlation for any given unit. The empirical analysis in the application section provides evidence of the effectiveness of the proposed composite index in identifying units (firms) with different performance levels, including best-performing units and those experiencing difficulties.

Second, the methodology defines a composite index which embeds expert opinions in a transparent manner, without resorting to exogenous variables and hidden assumptions. It does so employing a Beta-Bernoulli process estimated from real data combined with easy to understand goalposts, which can reflect expert opinions, and have a light impact, as demonstrated by the robustness analysis presented in the panel data framework considered in the paper.

Third, it provides a procedure for assessing the effect of the expert opinions, expressed by policy makers, on the mean performance of a set of units. It does so embedding expert opinions in the construction of the composite index, without resorting to any optimisation problem (as Chen et al., 2023 does), but simply scaling indicators according to the goalposts specified by the experts, and using the scaled values to calculate the composite index. The empirical evidence presented in the cross sectional framework of the paper shows how to actually calibrate expert opinions to improve the overall mean value of the indicators. This can enable policy makers and investors, among others, to set realistic targets for the units under study, which can thereby be employed as motivation factors.

The methodology is well-suited for case studies involving a large number of component indicators. Finally, while the Beta-Bernoulli process represents one possible approach, alternative processes merit further exploration.

Notes

1. We select indicators defined on the set of positive real numbers.
2. Note that the vector of success performances $\underline{X}^i$ could be known or unknown.
3. In our application, the realizations of the random indicator variable, X_j^i, are unknown. We only observe the values y_j^i of the component indicator and choose the goalposts to scale them.
4. For illustrative purposes, we select $\Delta_{m,r}$, $r = 1, 2, \ldots, R$, equal to 0 and $R = 21$ expert opinions. However, the results are not limited to this specific choice and can be generalized to any number of expert beliefs.

Author contributions

CRediT: **Mariateresa Ciommi**: Writing – original draft; **Paolo Giudici:**Supervision and draft revision**; Francesca Mariani**: Methodology; **Gloria Polinesi**: Software, Writing – original draft; **Maria Cristina Recchioni**: Supervision and methodology

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This research was funded by the European Commission (GRINS PE00000018)

References

Blancard, S., & Hoarau, J.-F. (2013). A new sustainable human development indicator for small island developing states: A reappraisal from data envelopment analysis. *Economic Modelling*, *30*, 623–635. https://doi.org/10.1016/j.econmod.2012.10.016

Blancas, F., Contreras, I., & Ramírez-Hurtado, J. (2013). Constructing a composite indicator with multiplicative aggregation under the objective of ranking alternatives. *Journal of the Operational Research Society*, *64*(5), 668–678. https://doi.org/10.1057/jors.2012.90

Ceriani, L., & Gigliarano, C., et al. (2016). Multidimensional well-being: A Bayesian networks approach. *Social Indicators Research*, *152*, 237–263.

Chen, R., Ji, Y., Jiang, G., Xiao, H., Xie, R., & Zhu, P. (2023). Composite index construction with expert opinion. *Journal of Business & Economic Statistics*, *41*(1), 67–79. https://doi.org/10.1080/07350015.2021.2000418

Chen, T.-Y., & Li, C.-H. (2010). Determining objective weights with intuitionistic fuzzy entropy measures: A comparative analysis. *Information Sciences*, *180*(21), 4207–4222. https://doi.org/10.1016/j.ins.2010.07.009

Cherchye, L., Moesen, W., Rogge, N., Van Puyenbroeck, T., Saisana, M., Saltelli, A., Liska, R., & Tarantola, S. (2008). Creating composite indicators with DEA and robustness analysis: The case of the technology achievement index. *Journal of the Operational Research Society*, *59*(2), 239–251. https://doi.org/10.1057/palgrave.jors.2602445

Fichman, M., & Levinthal, D. A. (1991). Honeymoons and the liability of adolescence: A new perspective on duration dependence in social and organizational relationships. *The Academy of Management Review*, *16*(2), 442–468. https://doi.org/10.2307/258870

Greco, S., Ishizaka, A., Tasiou, M., & Torrisi, G. (2019). On the methodological framework of composite indices: A review of the issues of weighting, aggregation, and robustness. *Social Indicators Research*, *141*(1), 61–94. https://doi.org/10.1007/s11205-017-1832-9

Hoffmann, A., Giovannini, E., Nardo, M., Saisana, M., Saltelli, A., & Tarantola, S. (2008). *Handbook on constructing composite indicator: Methodology and users guide*. OECD Publishing.

Karagiannis, G. (2017). On aggregate composite indicators. *Journal of the Operational Research Society*, *68*(7), 741–746. https://doi.org/10.1057/jors.2015.81

Kuc-Czarnecka, M., Lo Piano, S., & Saltelli, A. (2020). Quantitative storytelling in the making of a composite indicator. *Social Indicators Research*, *149*(3), 775–802. https://doi.org/10.1007/s11205-020-02276-0

Lukason, O., Laitinen, E. K., & Suvas, A. (2016). Failure processes of young manufacturing micro firms in europe. *Management Decision*, *54*(8), 1966–1985. https://doi.org/10.1108/MD-07-2015-0294

Mariani, F., Ciommi, M., & Recchioni, M. C. (2024). A new class of composite indicators: The penalized power mean. *European Journal of Operational Research*, *317*(3), 1015–1035. https://doi.org/10.1016/j.ejor.2024.04.032

Mazziotta, M., & Pareto, A. (2017). Synthesis of indicators: The composite indicators approach. *Complexity in society: From indicators construction to their synthesis* (pp. 159–191). Springer. https://doi.org/10.1007/978-3-319-60595-1_7

Mazziotta, M., & Pareto, A. (2018). Measuring well-being over time: The adjusted Mazziotta–Pareto index versus other non-compensatory indices. *Social Indicators Research*, *136*(3), 967–976. https://doi.org/10.1007/s11205-017-1577-5

Polinesi, G., Ciommi, M., Mariani, F., & Gigliarano, C. (2026). Unveiling multidimensional poverty across italian provinces using small area estimation and penalized power means. *Annals of Operations Research*, *363*(1), 69–86. https://doi.org/10.1007/s10479-025-06580-1

Ravanos, P., & Karagiannis, G. (2021). A VEA benefit-of-the-doubt model for the HDI. *Social Indicators Research*, *155*(1), 27–46. https://doi.org/10.1007/s11205-020-02589-0

Salzman, J. (2003). *Methodological choices encountered in the construction of composite indices of economic and social well-being*. Centre for the Study of Living Standards.

Shemshadi, A., Shirazi, H., Toreihi, M., & Tarokh, M. J. (2011). A fuzzy vikor method for supplier selection based on entropy measure for objective weighting. *Expert Systems with Applications*, *38*(10), 12160–12167. https://doi.org/10.1016/j.eswa.2011.03.027

Siegrist, K. (2022). *Probability, mathematical statistics, and stochastic processes*. LibreTexts, University of Alabama.

Verbunt, P., & Rogge, N. (2018). Geometric composite indicators with compromise benefit-of-the-doubt weights. *European Journal of Operational Research*, *264*(1), 388–401. https://doi.org/10.1016/j.ejor.2017.06.061

Walheer, B. (2024). A sequential benefit-of-the-doubt composite indicator. *European Journal of Operational Research*, *316*(1), 228–239. https://doi.org/10.1016/j.ejor.2024.01.029

Yerkes, R., & Dodson, J. (1908). The relation of strength of stimulus to rapidity of habit-formation. *Journal of Comparative Neurology and Psychology*, *18*(5), 459–482. https://doi.org/10.1002/cne.920180503

Zanella, A., Camanho, A. S., & Dias, T. G. (2015). Undesirable outputs and weighting schemes in composite indicators based on data envelopment analysis. *European Journal of Operational Research*, *245*(2), 517–530. https://doi.org/10.1016/j.ejor.2015.03.036

Appendix

A.1. Proof of the properties of Beta-Bernoulli process

In this appendix, we prove Equations (8)–(11) of Section 2. Keeping in mind that X_i for any $P = p$ is a Bernoulli random variable with success probability p, we have:

$$E(X_l) = \int_0^1 E(X_l|P = p)f(p;\alpha,\beta)dp = \int_0^1 (1 \cdot p + 0 1 - p))f(p;\alpha,\beta)dp \tag{A.1}$$

Since $B(\alpha,\beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$, $B(\alpha+1,\beta) = \frac{\Gamma(\alpha+1)\Gamma(\beta)}{\Gamma(\alpha+\beta+1)}$ and $\Gamma(x+1) = x\Gamma(x)$ we get:

$$E(X_l) = \frac{B(\alpha+1,\beta)}{B(\alpha,\beta)} = \frac{\alpha\Gamma(\alpha)\Gamma(\beta)}{(\alpha+\beta)\Gamma(\alpha+\beta)} \cdot \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} = \frac{\alpha}{\alpha+\beta}. \tag{A.2}$$

To compute variance end covariance we first compute the second order moments. We have:

$$E(X_l^2) = \int_0^1 E(X_l^2|P = p)f(p;\alpha,\beta)dp = \int_0^1 (1 \cdot p + 0 1 - p))f(p;\alpha,\beta) \qquad dp = \int_0^1 \frac{p^\alpha(1-p)^{\beta-1}}{B(\alpha,\beta)}dp = \frac{B(\alpha+1,\beta)}{B(\alpha,\beta)} = \frac{\alpha}{\alpha+\beta}, \tag{A.3}$$

therefore, the variance in Equation (9) is equal to

$$\mathrm{var}(X_l) = E(X_l^2) - (E(X_l))^2 = \frac{\alpha}{\alpha+\beta} - \left(\frac{\alpha}{\alpha+\beta}\right)^2 = \frac{\alpha}{\alpha+\beta} \cdot \frac{\beta}{\alpha+\beta}. \tag{A.4}$$

The mixed moment of the variables which are identically distributed is given by:

$$\begin{aligned} E(X_l X_s) &= \int_0^1 \left(1 \cdot p^2 + 0(2p(1-p) + (1-p)^2\right)f(p;\alpha,\beta)dp \\ &= \int_0^1 \frac{p^{\alpha+1}(1-p)^{\beta-1}}{B(\alpha,\beta)}dp = \frac{B(\alpha+2,\beta)}{B(\alpha,\beta)}, \end{aligned} \tag{A.5}$$

that is:

$$\begin{aligned} E(X_l X_s) = \int_0^1 p^2 f(p;\alpha,\beta)dp &= \frac{\Gamma(\alpha+2)\Gamma(\beta)\Gamma(\alpha+\beta+2)}{\Gamma(\alpha)\Gamma(\beta)\Gamma(\alpha+\beta)} \\ &= \frac{(\alpha+1)\alpha\Gamma(\alpha)\Gamma(\beta)\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)(\alpha+\beta+1)(\alpha+\beta)\Gamma(\alpha+\beta)} \\ &= \frac{(\alpha+1)\alpha}{(\alpha+\beta+1)(\alpha+\beta)} \end{aligned} \tag{A.6}$$

The proof of Equation (10) is:

$$\begin{aligned} \mathrm{cov}(X_l, X_s) = E(X_l X_s) - E(X_l)E(X_s) \quad &= \frac{(\alpha+1)\alpha}{(\alpha+\beta+1)(\alpha+\beta)} - \frac{\alpha}{\alpha+\beta} \cdot \frac{\alpha}{\alpha+\beta} = \\ &= \frac{\alpha}{\alpha+\beta}\left(\frac{(\alpha+1)}{(\alpha+\beta+1)} - \frac{\alpha}{(\alpha+\beta)}\right) \\ &= \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}. \end{aligned} \tag{A.7}$$

The proof of Equation (11) follows recalling the $\mathrm{cor}(X_l, X_s) = \mathrm{cov}(X_s, X_s)/\sqrt{\mathrm{var}(X_l)\mathrm{var}(X_s)}$ with $\mathrm{var}(X_l) = \mathrm{var}(X_s)$.