INSTITUTE
OF ECONOMICS

Scuola Superiore Sant'Anna

LEM | Laboratory of Economics and Management

Institute of Economics
Scuola Superiore Sant'Anna

Piazza Martiri della Libertà, 33 - 56127 Pisa, Italy
ph. +39 050 88.33.43
institute.economics@sssup.it

LEM
WORKING PAPER SERIES

E pluribus, quaedam. Gross domestic product out of a dashboard of indicators

Mattia Guerini [a,b,c,d]
Fabio Vanni [e,b,c]
Mauro Napoletano [b,f,g,c]

[a] University of Brescia, Italy.
[b] GREDEG, CNRS, Université Côte d'Azur, Nice, France.
[c] Institute of Economics, Scuola Superiore Sant'Anna, Pisa, Italy.
[d] Fondazione Eni Enrico Mattei (FEEM), Milan, Italy.
[e] Università degli Studi dell'Insubria, Varese, Italy.
[f] Sciences Po, OFCE, Sophia Antipolis, France.
[g] SKEMA Business School, Sophia Antipolis, France.

2022/15 May 2022
ISSN(ONLINE) 2284-0400

E pluribus, quaedam. Gross domestic product out of a dashboard of indicators

Mattia Guerini[1,2,3,4] Fabio Vanni[5,2,3] Mauro Napoletano[2,6,7,3]

[1] *University of Brescia*

[2] *GREDEG, CNRS, Université Côte d'Azur*

[3] *Institute of Economics, Sant'Anna School of Advanced Studies*

[4] *Fondazione Eni Enrico Mattei (FEEM)*

[5] *Università degli Studi dell'Insubria*

[6] *Sciences Po, OFCE*

[7] *SKEMA Business School*

May 11, 2022

Abstract

Is aggregate income enough to summarize the well-being of a society? We address this long-standing question by exploiting a novel approach to study the relationship between gross domestic product (GDP) and a set of economic, social and environmental indicators for nine developed economies. By employing dimensionality reduction techniques, we quantify the share of variability stemming from a large set of different indicators that can be compressed into a univariate index. We also evaluate how well this variability can be explained if the univariate index is GDP. Our results indicate that univariate measures, and GDP among them, are doomed to fail in accounting for the variability of well-being indicators. Even if GDP would be the best linear univariate index, its quality in synthesizing information from indicators belonging to different domains is poor. Our approach provides additional support for policy makers interested in measuring the trade offs between income and other relevant socio-economic and ecological dimensions. Furthermore, it adds new quantitative evidence to the already vast literature criticizing GDP as the most prominent measure of well-being.

JEL codes: C43, I30, I31.

Keywords: Gross domestic product, well-being indicators, data reduction techniques.

*All authors received financial support by the European Union Horizon 2020 research and innovation program with grant agreement No. 822781 - GROWINPRO.

1 Introduction

In this paper we quantify the ability of Gross Domestic Product (GDP henceforth) to summarise well-being, as defined by a broad set of economic, social, environmental sustainability and demographic indicators. One of the main characteristics that allowed GDP to gain a central role among all the possible economic measures of well-being, is its quantitative, monetary and synthetic nature. Together with the system of national accounting, GDP can be calculated across different countries on the same scale, thus allowing comparisons among them (Hoekstra, 2019). However, the monetary nature of GDP is also its main limitation. It is indeed difficult to assign a monetary value to all the activities for which a regular market does not exist but affect living standards and well-being. Furthermore, even where markets do exist they might be imperfect. And prices might not be able to incorporate all the possible effects that the production and consumption of goods and services generate, especially when these activities generate significant externalities. For all these reasons a number of scholars and policy making institutions have challenged the idea of relying uniquely upon GDP as a tool to quantify well-being and to evaluate policies (Michalos, 1982; Stiglitz et al., 2009).

As a consequence, *dashboards of indicators* have been created to better evaluate well being in a society, by taking into account dimensions such as environmental sustainability, social equity, cultural development, economic vulnerability as well as the demographic dimension (see OECD, 2011; Pinar et al., 2014; Roser, 2014; Fitoussi et al., 2018a,b; Ferran et al., 2018; Bacchini et al., 2020; Kalimeris et al., 2020, among the others). All these works focus on the multifaceted and complex nature of well being and, at the normative level, do not imply the complete replacement of GDP with the new metrics. They rather suggest that GDP shall be accompanied by alternative statistics.

We contribute to this literature by quantitatively evaluating the ability of GDP to capture the information embedded in a large set of social, economic and ecological indicators, which are constitutive of well-being. We make use of a widely known dimensionality reduction technique, namely generalized Principal Component Analysis (PCA). Moreover, we combine PCA with the Random Matrix Theory (RMT) approach, which allow us to single out the components that are statistically significant (see e.g. Onatski, 2010). More precisely, this procedure involves comparing the empirically estimated eigenvalues with the distribution of eigenvalues that is generated by a Gaussian random model with only spurious correlations.[1] Using the leading component or the GDP as univariate measures of well-being, we reconstruct two alternative synthetic series for all the indicators. By comparing the synthetic series with the original counterparts, we quantitatively measure the ability of GDP to summarize the variability of all the indicators. We apply this strategy to nine advanced OECD economies and our findings suggest that univariate measures, and GDP among them, are only imperfect proxies of well-being. With respect to the ability of GDP to approximate single indicators, substantial heterogeneity is found at the country level with the possibility of poor performance, especially over the demographic and social equity dimensions. Overall, our results, confirm that one shall rely upon multivariate composite indices of well-being, which are more apt at capturing the interactions between different indicators also pertaining to heterogeneous domains.

[1] This technique has also been recently adopted in the business cycle and financial economics literature.(see respectively Guerini et al., 2019; Barbieri et al., 2021). The main advantage of RMT is that it provides more precise and accurate information about a panel of time-series compared to basic PCA analysis, which does not allow one to distinguish between factors reflecting spurious correlations obtainable with a finite number of observations and those that instead contain relevant information about the similarity of the series.

2 Methodology

We start from N time series of well-being indicators observed for T periods and all sampled at the same frequency Δt. We denote the matrix of time-series by $\tilde{\boldsymbol{X}}(t)$ and their complex Hilbert transformation by $\boldsymbol{X}(t)$.[2]

2.1 Dimensionality reduction

According to the generalized PCA (Ng et al., 2001) the time series of the indicators can be expressed as:

$$\boldsymbol{X}(t) = \boldsymbol{A}(t)\boldsymbol{V} \tag{1}$$

where $\boldsymbol{A}(t)$ is a $T \times N$ loading matrix and $\boldsymbol{V}$ is a $N \times N$ matrix of eigenvectors.[3] These eigenvectors are associated to the N-dimensional vector of eigenvalues $\boldsymbol{\lambda}$, computed from the spectral decomposition:

$$\boldsymbol{C}\boldsymbol{V} = \boldsymbol{\lambda}\boldsymbol{V} \tag{2}$$

where the correlation matrix $\boldsymbol{C}$ can be estimated by $\widehat{\boldsymbol{C}} = \frac{1}{N}\boldsymbol{X}(t)\boldsymbol{X}(t)'$.[4] The correlation matrix $\widehat{\boldsymbol{C}}$ is positive semi-definite and bears N non-negative and distinct eigenvalues $\boldsymbol{\lambda}$ with their associated eigenvectors $\boldsymbol{V}$. According to Principal Component Analysis (PCA, Vidal et al., 2016), each eigenvalue can be expressed as a linear combination of the original series and corresponds to a principal component, also explaining a portion of the total variance of the data proportional to its magnitude. Thus, the empirical density function of the eigenvalues can be expressed as:

$$\widehat{\rho}(\boldsymbol{\lambda}) = \frac{dn(\boldsymbol{\lambda})}{d\boldsymbol{\lambda}} \tag{3}$$

where $n(\boldsymbol{\lambda})$ indicates the number of eigenvalues larger than $\boldsymbol{\lambda}$.

To focus solely on principal components which are statistically significant one can compare the empirical density function $\hat{\rho}(\boldsymbol{\lambda})$ with a theoretical benchmark distribution of eigenvalues that would have been generated under a known null-hypothesis. The random matrix theory (RMT) provides a well-specified theoretical null-hypothesis for such a statistical significance test (Onatski, 2010). But for the RMT to hold, it is required that no autocorrelation exists in the series and that the series are infinitely dimensional – in the sense that both $N, T \to \infty$, with $Q = \frac{T}{N}$ finite. To be free from these two tight restrictions, one can alternatively rely upon the less demanding rotational random shuffling (RRS) simulations which, in the limit, converge to the same theoretical distribution of RMT (Iyetomi et al., 2011; Aoyama et al., 2017; Kichikawa et al., 2020), represented by the Marchenko-Pastur distribution:

$$\rho(\boldsymbol{\lambda}) = \begin{cases} \frac{1}{2\pi\boldsymbol{\sigma}^2}\frac{\sqrt{(\lambda_M-\boldsymbol{\lambda})(\boldsymbol{\lambda}-\lambda_m)}}{\boldsymbol{\lambda}} & \text{if } \lambda_m \leq \boldsymbol{\lambda} \leq \lambda_M \\ 0 & \text{else} \end{cases} \tag{4}$$

[2]The Hilbert transformation on the series is useful because it will allow us to also capture the correlation between similar time series displaying time shifts in their co-movements.

[3]The t symbol in parenthesis is made explicit for matrices representing time series.

[4]Without loss of generality we assume all series to be stationary and standardized. This implies $\tilde{\boldsymbol{X}}(t) \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{1})$.

where $\lambda_m = \boldsymbol{\sigma}^2 \frac{\left(1-\sqrt{Q}\right)^2}{Q}$ and $\lambda_M = \boldsymbol{\sigma}^2 \frac{\left(1+\sqrt{Q}\right)^2}{Q}$ represent the lower and upper bounds. Deviations between the empirical distribution $\hat{\rho}(\boldsymbol{\lambda})$ and the theoretical one $\rho(\boldsymbol{\lambda})$, indicate the presence of some statistically significant components which can summarize the co-movements between the empirical indicators. This is exactly the reason why PCA is a considered a dimensionality reduction technique. In particular, the number of significant components is equivalent to the number of eigenvalues which exceed the theoretical (or simulated with the RRS) upper bound λ_M (Laloux et al., 2000).

2.2 Construction of Synthetic Indicators

Once the significant principal components have been selected according to the above-described procedure, one can construct a synthetic indicator of the original time series ("synthetic PC" henceforth) as follows:

$$\widehat{\boldsymbol{X}}_J(t) = \boldsymbol{A}_J(t)\boldsymbol{V}_J \tag{5}$$

where the index J is an integer indicating the number of significant eigenvalues, $\widehat{\boldsymbol{X}}_J(t)$ is a $T \times N$ matrix with the synthetic series, as generated using the J leading principal components (thus the index J), $\boldsymbol{V}_J$ is a $J \times N$ matrix of the estimated complex eigenvectors associated to the J significant eigenvalues, and $\boldsymbol{A}_J(t)$ is a $T \times J$ matrix with the associated loadings. The synthetic series will be different from the original ones, and they represent the indicators that would have been observed if the noisy component of each of the original indicators would have been ignored.[5]

Furthermore, one can evaluate the quality of GDP at summarizing information about well-being (as provided by the large set of the N original series) by assuming that the GDP is the leading component that summarizes well-being. Formally, this corresponds to assuming that GDP replaces the component loadings $\boldsymbol{A}_J(t)$, under the condition $J = 1$. With this assumption, one can obtain an alternative synthetic indicator ("synthetic GDP" henceforth) of the original time series as follows:

$$\widehat{\boldsymbol{X}}_G(t) = \alpha \boldsymbol{A}_G(t)\boldsymbol{V}_1 \tag{6}$$

where $\widehat{\boldsymbol{X}}_G(t)$ is the $T \times N$ matrix with the GDP-based synthetic series (thus the index G), $\boldsymbol{A}_G(t)$ is a $T \times 1$ matrix with the Hilbert transform of the GDP series, $\boldsymbol{V}_1$ is the $1 \times N$ eigenvector associated to the largest eigenvalue λ_1 and α is a rescaling factor measuring the deviation scalar from the dominant eigenvector.[6]

Once the synthetic indicators $\widehat{\boldsymbol{X}}_J(t)$ and $\widehat{\boldsymbol{X}}_G(t)$ are available, one can evaluate the quality of the matching with the original indicators $\boldsymbol{X}(t)$. For that, we use the root mean squared error (RMSE) and the Pearson correlation coefficient between the series.[7]

[5] The noisy components are here intended to be all the non-significant ones according to the procedure developed in Section 2.1.

[6] According to the small perturbation theory $\alpha = \frac{A_1(t)A_G(t)^*}{A_G(t)A_G(t)^*}$ and $A_G(t)^*$ is the complex conjugate of the GDP series (Stewart and Sun, 1990; Ng et al., 2001).

[7] Other more complicated alternatives are possible, for example the dynamic time-warp (DTW). However, it is not the aim of this paper to evaluate the quality of different similarity measures.

3 Empirical Application

We employ 42 different time-series indicators capturing economic, environmental, social equity and demographic dimensions.[8] All the indicators are sampled at annual frequency and cover the 1995-2015 period forming a balanced panel dataset for each country. Over the cross-section dimension, our analysis is performed on nine different advanced OECD economies.[9] The dataset format ensures a perfect comparability between the sampled countries.

After having transformed the indicators into stationary series by means of the first difference transformation, we apply the generalized PCA and the RMT procedures to test for the statistical significance of the estimated principal components. For most countries (all but Great Britain) we find that only the largest eigenvalue exceeds the RMT upper bound. This implies that one can significantly summarize a certain fraction of the data variance by means of a single variable. In particular, the fraction of variance captured by the first principal component is called the *absorption rate* and is reported in Table 1 (mid column). The leading principal component explains between 32% and 45% of the variance provided by the original data in the countries considered. Spain (ESP) has the highest absorption rate (45%) whereas Great Britain (GBR) has the lowest (31%). In addition, Great Britain is the unique country for which the leading component is not statistically significant.[10]

Overall, measuring well-being only by means of the best linear univariate predictor, as represented by the leading principal component, would embed a loss of about 55% to 70% of information from the original series. This is a relevant amount of the overall variation and is an indication for the failure of univariate measures to account for the complex and multivariate relationships between all the different indicators.

Table 1: Fraction of variance explained by the dominant eigenvalue of the generalized PCA (absorption rate) and correlation between GDP growth rate and the principal component. For all countries except GBR (starred), only the leading principal component is significant.

Country code	PCA absorption rate	Correlation ($\rho_{GDP,PC}$)
DEU	32.19%	88.26%
ESP	44.88%	89.51%
FRA	32.77%	90.31%
GBR*	31.13%	83.90%
ITA	40.37%	96.82%
JPN	35.19%	86.97%
NLD	32.83%	89.88%
SWE	32.16%	93.48%
USA	41.83%	91.01%

The reminder of the paper aims at evaluating the ability of GDP and the leading principal component to reconstruct the original indicators. We focus our analysis mostly on the USA, while extensive results for the other countries are presented in the on-line supplementary material.[11]

Figure 1 compares the evolution of GDP and of the leading principal component (leading PC henceforth). The correlation between these series is summarized in Table 1 (last column, which also report the correlation values for the other countries in the sample). The correlation is outstandingly high for the USA (91%), suggesting that the idea of employing GDP as a first order approximation of well-being might not be completely far fetched. In fact, it is fair to

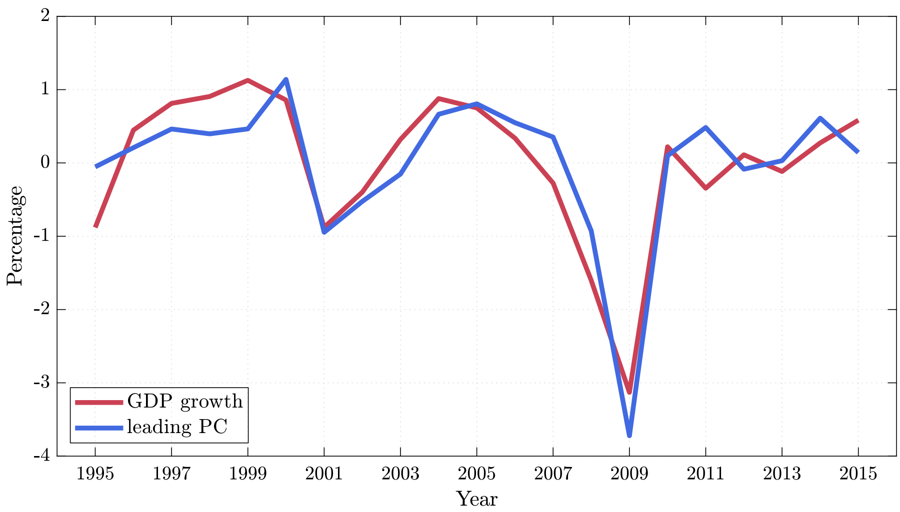
[8] The complete list is presented in Appendix A and all series are publicly available at the OECD i-Library.

[9] The nine countries and their short labels are Germany (DEU), Spain (ESP), France (FRA), Great Britain (GBR), Italy (ITA), Japan (JPN), Netherlands (NLD), Sweden (SWE) and United States (USA).

[10] The latter result points to the impossibility to summarize the information stemming from the indicators in a low dimensional space and warning against the usage of a univariate measure of well-being for this country.

[11] The on-line supplementary material is available at the authors' github pages.

Figure 1: GDP growth and leading principal component time series for the United States ($\rho = 0.9101$).



affirm that GDP growth mimics the dynamic of the best univariate linear approximation, which maximizes the absorption rate. However, as already mentioned, even the leading PC accounts for only about one third of the total variance.

Table 2: Comparison of original and synthetic indicators. For each indicator constructed with the leading PC or with the GDP growth, we report the Root Mean Squared Error in percentage terms with respect to the original indicator, the correlation between the series and the p-value of the correlation coefficient.

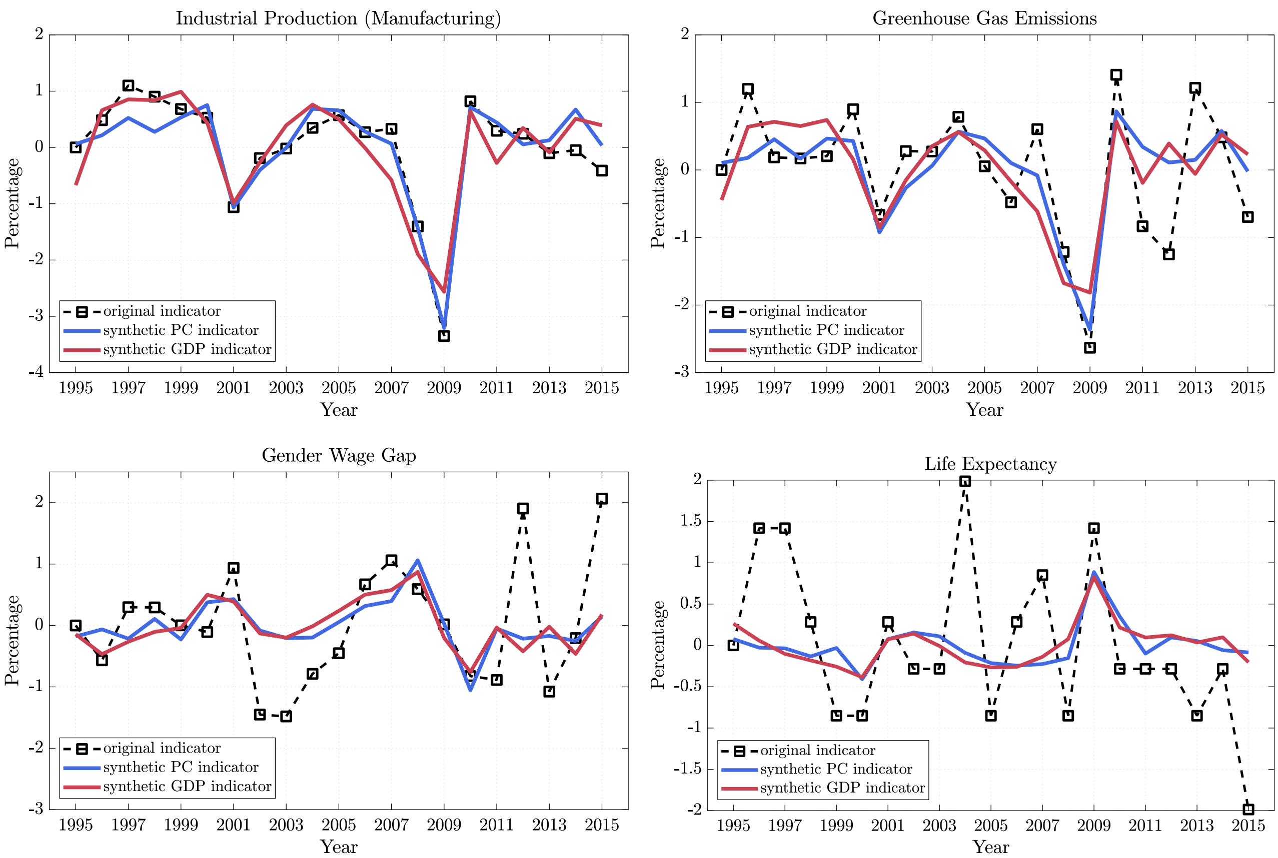
Indicator	RMSE (in %)		Correlation		Correlation p-value	
	PC	GDP	PC	GDP	PC	GDP
Industrial Production	7.03%	9.52%	0.94671	0.90082	< 0.001	< 0.001
Greenhouse Gas Emissions	14.23%	15.94%	0.75817	0.68303	< 0.001	< 0.001
Gender Wage Gap	19.89%	20.59%	0.41125	0.33877	0.0640	0.1397
Life Expectancy	21.02%	20.92%	0.26792	0.29519	0.2403	0.2075

To further quantify the loss of information when using a univariate measure of well-being we proceed with the construction of the synthetic indicators as described in Equations (5) and (6) (respectively "Synthetic PC" and "Synthetic GDP") for all the 42 original indicators in our sample. Results about four of these indicators, pertaining to different domains (economic, environmental, social and demographic) are presented in Figure 2. In addition, in Table 2 we also report the Root Mean Squared Error (RMSE) for each combination of original and synthetic indicators, the correlation coefficient between them, and its correspondent p-value. Figure 2 reveals that the two synthetic indicators display a similar behaviour in all the domains considered. This is coherent with the high degree of Pearson correlation shown in Table 1 (last column). However, their performance in tracking representative varies wildly across domains. For instance, the economic domains, represented by the percentage change of industrial production in the manufacturing sector, is reproduced with high precision (with an error of about 7%-9%). A similar qualitative result (but quantitatively different, cf. Table 2) is observed in the environmental domain, represented by the percent change of greenhouse gas emissions.[12]

In contrast, the synthetic indicators do not track well the evolution of variables selected as representative of social equity (the percent variation in the gender wage gap, with a RMSE of about 20%) and demography (measured by the percent variation in life expectancy), as clearly visible from the bottom panels of Figure 2. This can be explained by

[12]A similar performance holds true also for the CO_2 emissions indicator. This is also coherent with the empirical literature reporting a cointegration relationship (possibly time-varying) between GDP and emission of global pollutants (Mikayilov et al., 2018).

Figure 2: Selected original indicators (dashed black lines with squares) and synthetic indicators constructed starting from the leading PC (blue lines) and the GDP (red lines) for the USA.



the fact that these variables are *slow moving* and therefore they contribute less to the total variance of the dataset. Accordingly, the PCA assigns them lower weights in the loading matrix $A(t)$. The consequence is that neither GDP nor the leading PC is effective in predicting them. As a result, also the correlation coefficients between the original series and the synthetic ones are low and not significant (see Table 2).

Table 3: Cross country comparison of original and synthetic indicators. We report the Root Mean Squared Error in percentage terms with respect to the original indicator.

Indicator	PC RMSE (in %)				GDP RMSE (in %)			
	Ind. Prod.	Greenh. Gas	Gender Wage Gap	Life Exp.	Ind. Prod.	Greenh. Gas	Gender Wage Gap	Life Exp.
DEU	9.57%	17.89%	21.44%	21.35%	8.29%	20.40%	21.85%	21.01%
ESP	10.04%	11.31%	21.81%	21.34%	11.04%	14.54%	21.77%	21.75%
FRA	8.82%	19.92%	20.81%	20.29%	7.48%	20.71%	21.57%	20.93%
GBR	17.00%	15.25%	16.26%	20.57%	15.13%	21.08%	21.79%	22.12%
ITA	11.65%	11.31%	14.50%	21.52%	11.18%	12.90%	15.22%	21.33%
JPN	13.55%	12.53%	20.41%	21.45%	10.17%	15.28%	22.48%	22.25%
NLD	15.32%	20.62%	21.19%	21.65%	13.83%	20.83%	21.91%	21.99%
SWE	8.46%	19.93%	21.01%	21.79%	7.29%	18.88%	21.24%	21.85%
USA	7.03%	14.23%	19.89%	21.02%	9.52%	15.94%	20.59%	20.92%

Finally, Table 3 reports the percentage RMSE of the same four indicators for all countries in our sample. We find similar heterogeneity across indicators, with substantially larger RMSE for the social equity and demographic domains. Overall these results confirm that univariate measures of well-being are doomed to fail in their attempt of summarizing the information stemming from a large set of different indicators and pertaining to heterogeneous domains.

4 Conclusions

We can draw two main conclusions from our work. First, we find that, among the univariate alternatives aimed at summarizing a multitude of dimensions related to well-being, the GDP can be considered to be a good choice. It delivers a similar performance with respect to the leading principal component of the series (i.e. the best linear indicator). This militates in support of the usage of GDP as unique measure of well-being as also suggested by Malay (2019). At the same time, however, our results also suggest that one should not rely upon univariate measures of well-being. This is because the best linear estimator is only able to capture a relatively small portion of the indicators' variance (about 30% to 45%). Overall, this suggests that univariate measures of well-being are doomed to fail and one shall rely also upon multivariate composite indices of well-being (Bacchini et al., 2020; Kalimeris et al., 2020) and sustainability (Pinar et al., 2014; Luzzati and Gucciardi, 2015). These measures are more apt at capturing the complex interactions between different indicators also pertaining to very heterogeneous domains.

This work could be extended in several ways. First, one might enlarge the number of series in the sample, especially with respect to the social, equity and environmental dimensions. This might lead to different estimates for the loadings $A(t)$, assigning different weights to the single indicators when constructing the synthetic indicators based on principal component analysis. The effect of the inclusion of new variables is *a priori* unclear. Clearly, the variance explained by the first factor might increase or decrease, depending upon the degree of correlation between the new indicators with the one already in our sample. Furthermore, when new indicators are included, the number of significant factors might also vary, forcing one to account also for the higher order principal components. Second, one might enrich the analysis by considering economies at a different stages of the country development process. In particular, this type of research could be useful to detect whether the usage of GDP, interpreted as a univariate indicator of well-being, is more or less appropriate in developed vs. developing economies. For both these extension, however, the main difficulty lies in data availability, especially in relation to domains different from the economic one.

References

Aoyama, H., Y. Fujiwara, Y. Ikeda, and W. Souma (2017). *Macro-econophysics: new studies on economic networks and synchronization*. Cambridge University Press.

Bacchini, F., B. Baldazzi, and L. Di Biagio (2020). The evolution of composite indices of well-being: An application to italy. *Ecological Indicators 117*, 106603.

Barbieri, C., M. Guerini, and M. Napoletano (2021). The anatomy of government bond yields synchronization in the eurozone. *LEM Working Paper Series 2021/07*.

Ferran, P. H., R. Heijungs, and J. G. Vogtländer (2018). Critical analysis of methods for integrating economic and environmental indicators. *Ecological Economics 146*, 549–559.

Fitoussi, J.-P., M. Durand, et al. (2018a). *Beyond GDP Measuring What Counts for Economic and Social Performance: Measuring What Counts for Economic and Social Performance*. OECD Publishing.

Fitoussi, J.-P., M. Durand, et al. (2018b). *For Good Measure Advancing Research on Well-being Metrics Beyond GDP: Advancing Research on Well-being Metrics Beyond GDP*. OECD Publishing.

Guerini, M., D. T. Luu, and M. Napoletano (2019). Synchronization patterns in the european union. *LEM Working Paper Series 2019/33*.

Hoekstra, R. (2019). Replacing GDP by 2030. *Cambridge Books*.

Iyetomi, H., Y. Nakayama, H. Yoshikawa, H. Aoyama, Y. Fujiwara, Y. Ikeda, and W. Souma (2011, September). What causes business cycles? Analysis of the Japanese industrial production data. *Journal of the Japanese and International Economies 25*(3), 246–272.

Kalimeris, P., K. Bithas, C. Richardson, and P. Nijkamp (2020). Hidden linkages between resources and economy: A beyond-gdp approach using alternative welfare indicators. *Ecological Economics 169*, 106508.

Kichikawa, Y., H. Iyetomi, H. Aoyama, Y. Fujiwara, and H. Yoshikawa (2020, February). Interindustry linkages of pricesAnalysis of Japans deflation. *PLOS ONE 15*(2).

Laloux, L., P. Cizeau, M. Potters, and J.-P. Bouchaud (2000, July). Random matrix theory and financial correlations. *International Journal of Theoretical and Applied Finance 03*(03), 391–397. Publisher: World Scientific Publishing Co.

Luzzati, T. and G. Gucciardi (2015). A non-simplistic approach to composite indicators and rankings: an illustration by comparing the sustainability of the eu countries. *Ecological economics 113*, 25–38.

Malay, O. E. (2019). Do beyond gdp indicators initiated by powerful stakeholders have a transformative potential? *Ecological Economics 162*, 100–107.

Michalos, A. C. (1982). *North American social report: a comparative study of the quality of life in Canada and the USA from 1964 to 1974*, Volume 2. Springer Science & Business Media.

Mikayilov, J. I., F. J. Hasanov, and M. Galeotti (2018, December). Decoupling of CO2 emissions and GDP: A time-varying cointegration approach. *Ecological Indicators 95*, 615–628.

Ng, A. Y., A. X. Zheng, and M. I. Jordan (2001). Link analysis, eigenvectors and stability. In *International Joint Conference on Artificial Intelligence*, Volume 17, pp. 903–910. Citeseer.

OECD (2011). Measuring well-being and progress. *How's Life? 2020: Measuring Well-being URL: https://doi.org/10.1787/9870c393-en.*

Onatski, A. (2010). Determining the number of factors from empirical distribution of eigenvalues. *The Review of Economics and Statistics 92*(4), 1004–1016.

Pinar, M., C. Cruciani, S. Giove, and M. Sostero (2014). Constructing the feem sustainability index: A choquet integral application. *Ecological Indicators 39*, 189–202.

Roser, M. (2014). Human development index (HDI). *Our World in Data*.

Stewart, G. W. and J.-g. Sun (1990). *Matrix perturbation theory.* Boston: Academic Press.

Stiglitz, J. E., A. Sen, J.-P. Fitoussi, et al. (2009). Report by the commission on the measurement of economic performance and social progress.

Vidal, R., Y. Ma, and S. S. Sastry (2016). *Generalized principal component analysis*, Volume 5. Springer.

A List of indicators

Table 4: List of indicators' names, as sourced from the OECD i-Library database.

Name	Short
Agricultural land (Total), Thousand hectares	Land
Average wages (Total), US dollars	Wages
Carbon dioxide (CO2), Tonnes per capita	CO2
Crude oil production (Total), Thousand of toe	Oilprod
Electricity generation (Total), Gigawatt-hours	ElectrGen
Employment rate (Total), Thousand persons	EmployRate
Fertility rates (Total), Children per woman	Fertility
Gender wage gap, Percentage	WageGap
Greenhouse gas (GHG), Tonnes per capita	GHG
Gross insurance premiums (Total), Million US dollars	Insurance
Gross national income (Total), US dollars per capita	GNI
Health spending (Total), US dollars per capita	Healthspend
Hospital beds (Total), Per 1 000 inhabitants	HospBeds
Hours worked (Total), Hours per worker	HoursWorked
Household disposable income Net	HouseIncome
Household spending (Total), Million US dollars	HouseSpend
Housing prices, Real house prices	HousePrices
Housing prices, Rent price	RentPrices
Industrial production (Manufacturing)	IndustrialProd
Infant mortality rates (Total), Deaths live births	Mortality
Inflation of consumer price index (Total), Annual growth rate	CPI
Labour force (Total), Thousand persons	LabourF
Investment (Total), Million	Investment
Life expectancy at birth (Total), Years	LifeExpect
Municipal waste (Total), Thousand tonnes	Waste
Nutrient balance (Nitrogen), Kilograms per hectare	Nitrogen
Nutrient balance (Phosphorus), Kilograms per hectare	Phospho
Pharmaceutical spending (Total), US dollars per capita	PharmaSpend
Population (Total), Million persons	POP
Primary energy supply (Total), Million toe	EnergySupply
Producer price indices (Manufacturing), domestic market, Annual growth rate	PPI
Renewable energy (Total), Thousand toe	RenEnergy
Researchers (Total)	Research
Share prices (Total)	SharePrices
Social spending (Public), US dollars per capita	SocialSpend
Suicide rates (Total), Per 100000 persons	Suicide
Tax revenue (Total), Million US dollars	Taxrev
Trade in goods and services (Exports), Million US dollars	TradeExp
Trade in goods and services (Imports), Million US dollars	TradeImp
Triadic patent families (Total)	Patents
Unemployment rate (Total), of labour force	Unemploy
Unit labour costs, by persons employed	UnitLabforce

toe = tonnes of oil equivalent