European Commission

JRC TECHNICAL REPORTS

The JRC Statistical Audit of the 2018 European Skills Index (ESI)

Norlén, Hedvig

Saisana, Michaela

2018

EUR 29352 EN

Joint Research Centre

This publication is a Technical report by the Joint Research Centre (JRC), the European Commission's science and knowledge service. It aims to provide evidence-based scientific support to the European policymaking process. The scientific output expressed does not imply a policy position of the European Commission. Neither the European Commission nor any person acting on behalf of the Commission is responsible for the use that might be made of this publication.

Contact information

Competence Centre on Composite Indicators and Scoreboards
jrc-coin@ec.europa.eu

JRC Science Hub
https://ec.europa.eu/jrc

JRC 113225

EUR 29352 EN

PDF ISBN 978-92-79-93710-1 ISSN 1831-9424 doi: 10.2760/803749

Luxembourg: Publications Office of the European Union, 2018

How to cite this report: Norlén Hedvig, Saisana Michaela, *The JRC Statistical Audit of the European Skills Index (ESI),* European Commission, Joint Research Centre, Ispra, 2018, ISBN 978-92-79-93710-1, doi:10.2760/803749, JRC 113225

Contents

Abstract

Increased connectivity between nations, technological development, changes in work organisation and demographic trends have profound effects on the future of work and workplaces. Policies focusing on skills development and human capital are essential to turn these structural changes into an opportunity for all, by increasing productivity levels and quality of life in the EU. The Commission services have developed an EU tailor-made monitoring framework – the European Skills Index (ESI) – that measures the performance of a country's skills system taking into account its multiple facets from continually developing the skills of the population to activating and effectively matching these skills to the needs of employers in the labour market. The European Skills Index builds on three pillars: skills development, skills activation and skills matching. These pillars are used to organise and aggregate 15 individual indicators into a single summary measure. This framework inevitably entails both conceptual and practical challenges. The statistical audit discussed in this note was conducted by the European Commission's Joint Research Centre, and it aims at maximising the reliability and transparency of the European Skills Index ([1]). It should enable policy analysts and researchers alike to draw more relevant, meaningful and useful conclusions on the national skills systems in the EU.

([1]) The JRC statistical audit is based on the recommendations of the OECD & JRC (2008) Handbook on Composite Indicators, and on more recent research from the JRC. Generally, JRC audits of composite indicators and scoreboards are conducted upon request of their developers, see https://ec.europa.eu/jrc/en/coin and https://composite-indicators.jrc.ec.europa.eu/

1 Introduction

The European Skills Index (ESI) aims at measuring the performance of a country's skills system taking into account three main aspects: skills development, skills activation and skills matching. These pillars are used to organise and aggregate 15 individual indicators into a single summary measure. The index is developed by the European Centre for the Development of Vocational Training (CEDEFOP) with the technical expertise of Cambridge Econometrics.

This audit was performed by the European Commission's Competence Centre on Composite Indicators and Scoreboards at the Joint Research Centre (JRC) and was conducted upon invitation of the index developers. The analysis herein aims at shedding light on the transparency and reliability of the ESI model and thus to enable policymakers to derive more accurate and meaningful conclusions, and to potentially guide their choices on priority setting and policy formulation.

The JRC assessment of the ESI 2018 focuses on two main issues: the statistical coherence of the hierarchical structure of indicators and the impact of key modelling assumptions on the ESI ranking. The JRC analysis complements the reported country rankings for ESI with confidence intervals in order to better appreciate the robustness of these ranks to the computation methodology (in particular estimation of missing data, normalisation method, use of goalposts for the indicators, and weights and aggregation formula at the pillar level).

The ESI 2018 updates and refines the work undertaken for the Making Skills Work Index (MSWI), first published in 2016. Earlier versions of the index were evaluated by the JRC in May and in December 2017. Consequently, the theoretical and conceptual framework for the ESI building on 3 pillars and 6 sub-pillars has remained unchanged. Yet, several improvements were introduced by the index developers (see Table 1):

1) The aggregation layer on "indicator groups" in the MSWI was removed because it was not offering clear-cut policy insights and it was not statistically supported by the data;
2) Eleven indicators in the MSWI were either removed or replaced by four indicators that were found to be more relevant both conceptually and statistically (there are now 15 indicators in the ESI as opposed to 22 indicators in its predecessor MSWI);
3) To ease communication, the normalisation method was altered, from z-scores to a min-max normalisation;
4) The aggregation method at pillar level was changed from weighted arithmetic average to weighted geometric average in order to emphasise that the level of priority given to an ESI pillar should not be invariant to the level of attainment.

Table 1. The European Skills Index: Conceptual framework (right) and earlier working version (left).

Making Skills Work Index version 2017

Pillar (P)	Sub-pillar (SP)	Indicator group (IG)	Indicator (ind)	
P1: Skills Develop-ment	SP1: Compulsary education	IG1: Participation to compulsory education	Pre-primary participation	ind.01
			Upper secondary participation (aged 15-17)	ind.02
		IG2: Attainment from compulsory education	Upper secondary attainment (aged 15-64)	ind.03
			Reading, maths & science scores (aged 15)	ind.04
	SP2: Post-compulsory education and training	IG3: Participation in post-compulsory education and training	Recent training	ind.05
			Lifelong learning (employees)	ind.06
			Lifelong learning (aged 25-64)	ind.07
			VET students	ind.08
			Training deficit	ind.09
		IG4: Attainment from post-compulsory education and training	Tertiary attainment	ind.10
			High computer skills	ind.11
P2: Skills Activation	SP3: Transition from education to work	*(IG missing)*	Early leavers from training	ind.12
			NEETs	ind.13
			Recent graduates in employment	ind.14
	SP4: Activity rates		Activity rate (aged 15-24)	ind.15
			Activity rate (aged 25-54)	ind.16
			Activity rate (aged 55-64)	ind.17
P3: Skills Matching	SP5: Unemploy-ment and vacancies	IG5: Unemployment	Long-term unemployment	ind.18
		IG6:Vacancies	Structural vacancies	ind.19
		IG7:Under-employment	Underemployed part-time workers	ind.20
	SP6: Skills mismatch	*(IG missing)*	Skills obsolescence	ind.21
			Higher education mismatch	ind.22

European Skills Index version 2018

Indicator (ind)		Sub-pillar (SP)	Pillar (P)
Pre-primary pupil-to-teacher ratio	ind.1	SP1: Compulsary education	P1: Skills Develop-ment
Upper secondary attainment (aged 15-64)	ind.2		
Reading, maths & science scores (aged 15)	ind.3		
Recent training	ind.4	SP2: Post-compulsory education and training	
VET students	ind.5		
High computer skills	ind.6		
Early leavers from training	ind.7	SP3: Transition from education to work	P2: Skills Activation
Recent graduates in employment	ind.8		
Activity rate (aged 25-54)	ind.9	SP4: Activity rates	
Activity rate (aged 20-24)	ind.10		
Long-term unemployment	ind.11	SP5: Unemploy-ment and vacancies	P3: Skills Matching
Underemployed part-time workers	ind.12		
Higher education mismatch	ind.13	SP6: Skills mismatch	
ISCED 5-8 proportion of low wage earners	ind.14		
Qualification mismatch	ind.15		

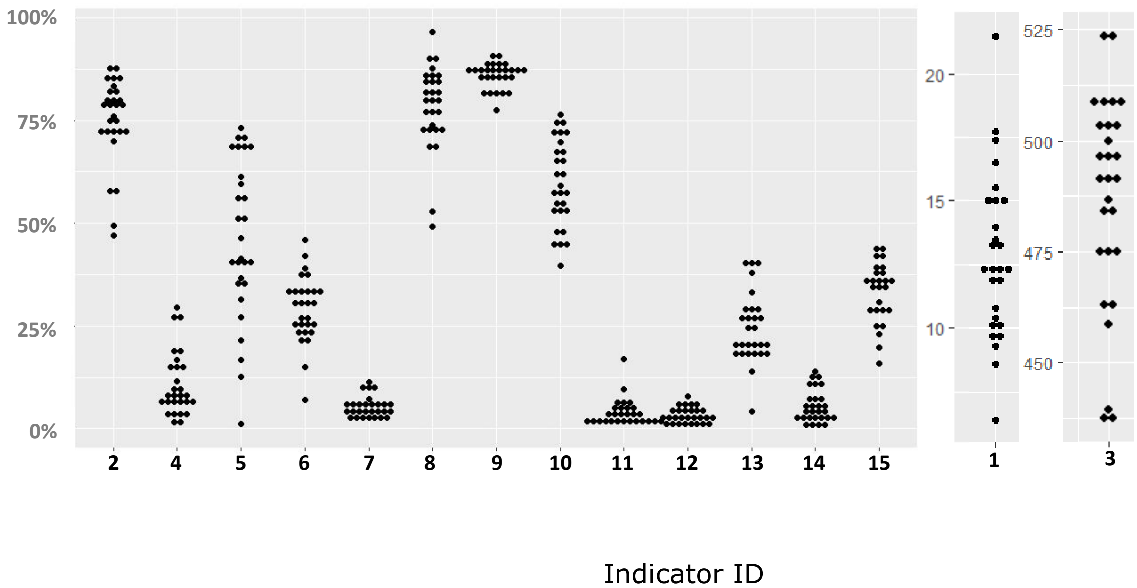
Notes: Making Skills Work Index (left) was an earlier beta-version of the European Skills Index (right). Eleven indicators (in red, left table) were either removed or replaced with four indicators (in green, right table).

Source: European Commission, Joint Research Centre, 2018 (based on the European Skills Index report).

2 Data analysis and rationale for the choices underpinning the ESI construction

Relevance to the ESI framework. Fifteen indicators were selected by the ESI developers for their relevance to a specific pillar, capturing skills development, skills activation or skills matching, on the basis of the literature review, expert opinion, country coverage and timeliness. Figure 1 illustrates the distributions of the indicators in the form of dot plots. All indicators are expressed in percentages (share of population), except for two indicators measuring the number of pre-primary pupils per teacher (ind 1) and the average PISA scores for 15 year old students in reading, maths and science (ind 3). The former indicator may be seen as a proxy for the quality of teaching at pre-primary education level; the latter indicator is a proxy for key outcomes from initial education which build the foundation for long-term economic growth of societies and social inclusion of individuals.

Figure 1. Dot plots for the 15 indicators in the ESI framework for the EU



Notes: Indicator names appear in Table 1. Indicators 2, 4-15 are expressed in share of population (%) (left panel). Indicator 1 expresses the number of pre-primary pupils per teacher (middle panel) and indicator 3 is the average PISA score for reading, maths and science (right panel).

Source: European Commission, Joint Research Centre, 2018.

Data availability. The most recently released data within the period 2013-16 were used for each country: 85% of the available data refer to 2015 or 2016. Table 2 offers summary statistics for the ESI indicators. The dataset has excellent data coverage; only three values are missing - Ireland's value on pre-primary pupil to teacher ratio (ind 1), and Croatia's and Malta's values on Qualification mismatch (ind 15). The ESI developers, for transparency and replicability, opted not to estimate missing data for these three cases.

On the desirable direction of performance, for eight indicators (ind 2, 3, 4, 5, 6, 8, 9 and 10) the higher the indicator value the better the performance, whereas for the remaining seven indicators (ind 1, 7, 11, 12, 13, 14, 15) the opposite holds true. For example, higher values are desirable for the share of population with upper secondary education (ind 2), whilst lower values are desirable for long-term unemployment (ind 11).

Outlier detection. Potentially problematic indicators that could bias the overall index results were identified on the basis of two measures related to the shape of the distributions: skewness and kurtosis. A practical rule suggested by the JRC is that country values should be treated if the indicators have absolute skewness greater than 2.0 and kurtosis greater than 3.5.(2) As shown in Table 2 and in

Figure 2, only one value may result problematic: the very high long-term unemployment rate (17%) for Greece (ind 11). To avoid that this value becomes an unintended benchmark and introduces bias in the aggregation with other indicators, the value for Greece is assigned to the second highest value from Spain (9.5%).

Normalisation. Next, the raw data for the fifteen ESI indicators were put in a common scale from 0.0 (lowest performance) to 1.0 by using the min-max normalisation method with fixed bounds (goalposts). Minimum and maximum values (goalposts) were chosen by the developers to act as the "logical worst case" and "logical best case" (or else aspirational targets), respectively, from which the ESI indicators are normalised (see Table 2). The main reason for the choice to use fixed bounds, as opposed to adopting the observed minimum and maximum values, is the need to benchmark performance over time. Keeping time-invariant the lower and upper bounds for the ESI indicators allows benchmarking over time. Detailed explanations on the rationale for the bounds for each indicator are offered in the European Skills Index 2018 report.

One simplification in the ESI calculation emerges at this point: winsorising Greece's value (from 17% to 9.5%) for long-term unemployment rate is not required given that by adopting the goalposts during the normalisation step the lower bound (worst case) for that indicator is set at 10%, which is very close to Spain's value (9.5%). The JRC recommendation is to consider simplifying the ESI development – simpler communication to the wider audience – by removing the winsorisation step but to keep on monitoring in next releases if the normalised (with the use of goalposts) indicator values satisfy the double criterion for skewness and kurtosis.

([2]) Groeneveld and Meeden (1984) set the criteria for absolute skewness above 1 and kurtosis above 3.5. The skewness criterion was relaxed in the ESI case after having conducted ad-hoc tests in the ESI 2010-2016 timeseries.

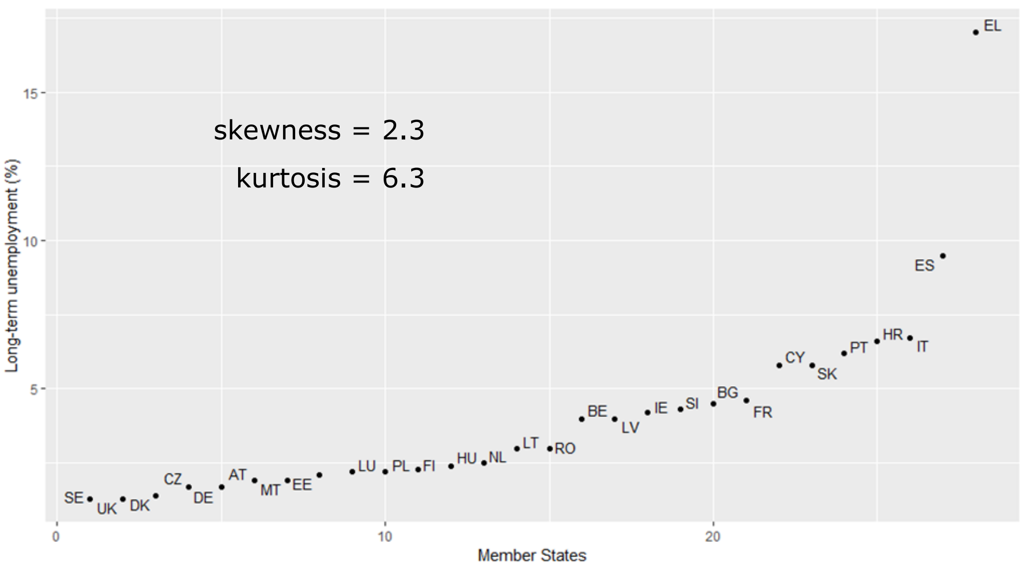
Table 2. Summary statistics of ESI indicators (raw data) and goalposts for the normalisation step

Pillar	Sub-pillar	Indicator		Direction	Nr of obs	Mean	Median	Skewness	Kurtosis	Observed best-lowest cases		Goalposts used at the normalisation step	
										Lowest	Best	Lower bound	Upper bound
Skills Development	Compulsory training	Pre-primary pupil-to-teacher ratio (students per teacher)	1	-	**27**	12.8	12.4	0.5	0.1	21.5	6.4	22.0	6.0
		Share of pop (aged 15-64) with at least upper secondary education (%)	2	+	28	75.1	78.2	-1.2	0.7	47.1	87.6	50.0	90.0
		Reading, maths & science scores (aged 15) (PISA)	3	+	28	486.9	491.8	-0.6	-0.5	437.5	524.3	440.0	550.0
	Training and tertiary education	Recent training (%)	4	+	28	10.8	8.4	1.0	0.0	1.2	29.6	1.0	30.0
		VET students (%)	5	+	28	46.2	43.9	-0.4	-0.8	1.2	73.2	10.0	75.0
		High computer skills (%)	6	+	28	29.2	30.0	-0.4	0.5	7.0	46.0	7.0	46.0
Skills Activation	Transition to work	Early leavers from training (%)	7	-	28	5.2	4.6	1.2	0.3	11.4	2.1	10.0	2.0
		Recent graduates in employment (%)	8	+	28	78.4	79.9	-1.0	1.2	49.2	96.6	55.0	95.0
	Activity rates	Activity rate (aged 25-54) (%)	9	+	28	86.1	87.1	-0.9	0.2	77.5	90.9	80.0	90.0
		Activity rate (aged 20-24) (%)	10	+	28	59.7	58.5	-0.1	-1.2	39.7	76.5	40.0	78.0
Skills Matching	Unemployment	Long-term unemployment (%)	11	-	28	4.1	3.0	**2.3**	**6.3**	17.0	1.3	10.0	1.0
		Underemployed part-time workers (%)	12	-	28	3.3	3.3	0.4	-0.7	7.8	0.5	7.0	1.0
	Skills mismatch	Higher education mismatch (%)	13	-	28	24.6	22.7	0.2	-0.1	40.7	4.2	40.0	10.0
		ISCED 5-8 proportion of low wage earners (%)	14	-	28	5.6	4.7	0.7	-0.8	13.8	0.2	14.0	0.0
		Qualification mismatch (%)	15	-	**26**	33.3	35.2	-0.6	-0.5	44.1	16.0	44.0	16.0

Notes: Raw data refer to the latest year available. Practical JRC rule for outlier detection: Indicators with |skewness|>2 and kurtosis >3. N=28 EU Member States.

Source: European Commission, Joint Research Centre, 2018.

Figure 2. Greece's outlier performance in long term unemployment rate in 2016



Source: European Commission, Joint Research Centre, 2018.

Aggregation. A hybrid aggregation approach was adopted to build the ESI components. Weighted *arithmetic* average was used at the first two aggregation levels (from indicators to sub-pillars, and from sub-pillars to pillars), and weighted *geometric* average was used at the third aggregation level (from pillars to an overall index). The rationale for this choice is the following. Weighted arithmetic averages are easy to interpret and allow perfect compensability between indicators, whereby a high score on one indicator can fully offset low scores in other indicators. At the lowest aggregation levels (indicators and sub-pillars), the assumption of perfect compensability of scores is considered reasonable and adequate. Yet, in the context of monitoring the performance of a country's skills system, adopting an arithmetic averaging at the highest aggregation level (where skills development, skills activation and skills matching are at play) it would have been problematic because it would have implied that the level of priority given to an ESI pillar is invariant to the level of attainment. Instead, the geometric average gives more incentives for improvement to low values (concave function). Thus a country with scores at 0.6, 0.9, and 0.3 for the skills development, skills activation and skills matching, respectively, would have more incentives to improve on skills matching than on any of the two other pillars.

Weights. The developers choice of the ESI weights was guided by the rationale that all ESI indicators should be equally informative with respect to the theme covered by the sub-pillar. The same rationale applied to all aggregation levels. To this end, an iterative process was adopted for the calculation of the weights: starting with equal weights within and across all ESI components, the weights of the indicators, sub-pillars and pillars were then calibrated by using information from the PCA factor loadings. Less weight was given to more correlated components (indicators, sub-pillars or pillars) and similarly more weight was given to less correlated components. Figure 3 illustrates the different weights and aggregation methods employed in the framework.

Figure 3. Aggregation methods and weights used in the ESI framework.

Indicator	Indicator Weights	Weighted Arithmetic Average	Sub-pillar	Sub-pillar Weights	Weighted Arithmetic Average	Pillar	Pillar Weights	Weighted Geometric Average	Index	Index Weights
Ind 01	0,4	→	Compulsory training	0,5	→	Skills Develop-ment	0,3	→	European Skills Index	0,06
Ind 02	0,3									0,05
Ind 03	0,3									0,05
Ind 04	0,3	→	Training and tertiary education	0,5						0,05
Ind 05	0,35									0,05
Ind 06	0,35									0,05
Ind 07	0,7	→	Transition to work	0,5	→	Skills Activation	0,3			0,11
Ind 08	0,3									0,05
Ind 09	0,5	→	Activity rates	0,5						0,08
Ind 10	0,5									0,08
Ind 11	0,4	→	Unemploy-ment	0,4	→	Skills Matching	0,4			0,06
Ind 12	0,6									0,10
Ind 13	0,4	→	Skills mismatch	0,6						0,10
Ind 14	0,1									0,02
Ind 15	0,5									0,12

Source: European Commission, Joint Research Centre, 2018.

3 Statistical coherence of the ESI framework

The reliability of the European Skills Index depends - among other things – on the degree of coherence between the conceptual framework and the statistical structure of the data. The more the ESI conceptual framework embraces the statistical structure, the higher the reliability of the ESI will be. The coherence of the ESI framework was assessed using two tests: (a) analysing the extent to which the ESI indicators can explain a sufficient amount of variation in the aggregated scores (be those sub-pillars, pillars or the overall index) by means of correlation, cross-correlation and principal component analysis, and (b) analysing the impact on the ESI country ranks when the least influential indicators (as identified in the first test) are omitted from the ESI framework.

Given that the present statistical analysis of the European Skills Index is in part, though not exclusively, based on correlations, the correspondence of the ESI to a real-world phenomenon needs to be critically addressed by experts in the field because 'correlations need not necessarily represent the real influence of the individual indicators on the phenomenon being measured'.([3])

In a nutshell, the argument is that the validity of the ESI framework relies on the combination of both statistical and conceptual soundness. In this respect, the ESI framework has been developed following an iterative process that went back and forth between the theoretical understandings of EU Member States' skills formation and matching systems on the one hand, and data observations on the other.

3.1 First statistical coherence test for the ESI framework

Starting with the simplest approach, correlation and cross-correlation analysis was used to assess to what extent the data collected support the ESI conceptual framework.

There is no redundancy of information in the ESI framework given the lack of highly collinear (i.e. Pearson correlation coefficients greater than 0.92) pairs of indicators within the same sub-pillar. The ESI framework is instead highly multifaceted, whereby indicators within any ESI sub-pillar exhibit generally modest to low correlations between them.

A more detailed analysis of the ESI correlation structure within and across the six sub-pillars confirms the expectation that the indicators are more associated with their own sub-pillar than to any of the other sub-pillars (Table 3). The same holds true for the associations within and across the three pillars. This result suggests that the allocation of the ESI indicators to the specific sub-pillar, and allocation of sub-pillars to pillars, is consistent both from conceptual and statistical perspectives. Furthermore all correlation coefficients within an ESI sub-pillar are close to or greater than 0.70, which suggests that roughly 50% (or more) of the variance in the ESI sub-pillar scores can be explained by an underlying indicator. The ESI indicators can also explain a significant share of the variance in the pillar scores as well. Most correlation coefficients between indicators and the pillar they belong to are also close to or greater than 0.7. The only indicator that is not significantly correlated to its own pillar (although it was sufficiently related to its own sub-pillar) is the Proportion of low wage earners (#14) in the Skills Matching. The majority of the ESI indicators (ten out of 15) are also positively and significantly correlated with the overall index. The five indicators that are found not to be sufficiently related to the overall index: Pre-primary pupil-to-teacher ratio (#1), VET students (#5), Activity rate (aged 20-24) (#10), Underemployed part-time workers (#12), and Proportion of low wage earners (#14). However, given that these five indicators are influential at the first and second aggregation levels (sub-pillars and pillars), their

([3]) OECD & EC JRC (2008).

inclusion in the ESI framework is corroborated by the analysis. All ESI sub-pillars correlate strongly with the respective pillars (correlation coefficients close to 0.85 or greater) and all three ESI pillars correlate strongly and in a balanced way with the ESI (correlations ranging between 0.71 and 0.77). This confirms the developers choice to use uneven weights for the three pillars (0.3, 0.3 and 0.4) in order to ensure that all three pillars are placed on equal footing when it comes to calculating a summary measure for the performance of a country's skills system.

Besides the statistical confirmation for many of the ESI choices made thus far, correlation analysis has evidenced an outcome worthy of further reflection: there is good statistical association between the first two pillars, namely Skills Development and Skills Activation (correlation 0.62) and no association between these two pillars and the Skills Matching pillar. This outcome will be discussed in more detail in Section 5.

Next, principal component analysis (PCA) was used to confirm whether there is a single statistical dimension in each ESI component, which would give the "statistical justification" for aggregating indicators into one number. Technically, the expectation here is that there is only one principal component with eigenvalue greater than 1.0. Indeed, PCA results corroborate the presence of a single latent dimension in each of the six ESI sub-pillars that captures between 47% (Sub-pillar 1 Compulsory training) up to 72% (Sub-pillar 3 Transition to work) of the total variance in the underlying indicators. PCA analysis at the pillar level confirms unidimensionality in each of the three pillars: the single latent dimension captures 74% in Pillar 1 Skills Development, 81% in Pillar 2 Skills Activation and 86% in Pillar 3 Skills Matching of the total variance of the underlying sub-pillars. Finally, the three ESI pillars share a single statistical dimension that summarizes 58% of the total variance. This latter result supports the aggregation of three pillars into one number.

Table 3. Statistical coherence in the ESI

(a) correlations between indicators and other ESI components

Pillar	Sub-pillar	Indicator	SP1: Compul-sory education	SP2: Training and tertiary education	SP3: Transition to work	SP4: Activity rates	SP5: Unemploy-ment	SP6: Skills mismatch	P1: Skills Develop-ment	P2: Skills Activation	P3: Skills Matching	Index
P1: Skills Develop-ment	SP1: Compulsory training	ind.1	0,71	0,22	0,30	0,17	0,39	0,22	0,51	0,26	0,32	0,43
		ind.2	0,70	0,19	0,43	0,27	0,41	0,18	0,50	0,39	0,30	0,52
		ind.3	0,64	0,62	0,38	0,55	0,08	0,01	0,73	0,51	0,04	0,52
	SP2: Training and tertiary education	ind.4	0,41	0,81	0,36	0,66	-0,03	0,07	0,72	0,56	0,03	0,49
		ind.5	0,20	0,61	0,08	-0,03	0,19	0,46	0,48	0,03	0,37	0,38
		ind.6	0,48	0,75	0,41	0,67	-0,09	-0,02	0,72	0,60	-0,06	0,48
P2: Skills Activation	SP3: Transition to work	ind.7	0,53	0,36	***0,96***	0,53	0,13	0,16	0,51	0,83	0,16	0,65
		ind.8	0,36	0,23	0,66	0,60	0,52	0,48	0,34	0,70	0,53	0,72
	SP4: Activity rates	ind.9	0,40	0,51	0,45	0,81	-0,02	0,06	0,53	0,69	0,03	0,54
		ind.10	0,37	0,37	0,56	0,81	-0,01	-0,06	0,43	0,76	-0,04	0,46
P3: Skills Matching	SP5: Unemploy-ment	ind.11	0,47	0,32	0,51	0,45	0,72	0,49	0,45	0,54	0,63	0,72
		ind.12	0,28	-0,11	0,05	-0,29	0,91	0,68	0,08	-0,13	0,84	0,42
	SP6: Skills mismatch	ind.13	0,26	0,46	0,39	0,25	0,64	0,78	0,42	0,36	0,78	0,74
		ind.14	-0,23	0,32	-0,15	-0,16	0,05	0,50	0,06	-0,17	0,34	0,12
		ind.15	0,23	0,13	0,11	-0,12	0,63	0,84	0,21	0,00	0,81	0,48

(b) correlations between sub-pillars and other ESI components

ESI sub-pillars and pillars	SP1: Compul-sory education	SP2: Training and tertiary education	SP3: Transition to work	SP4: Activity rates	SP5: Unemploy-ment	SP6: Skills mismatch	P1: Skills Develop-ment	P2: Skills Activation	P3: Skills Matching	Index
SP1: Compulsory education	1.00						0.84	0.57	0.29	0.70
SP2: Training and tertiary education	0.48	1.00					0.88	0.50	0.20	0.62
SP3: Transition to work	0.55	0.37	1.00				0.53	0.90	0.29	0.76
SP4: Activity rates	0.47	0.54	0.62	1.00			0.59	0.89	-0.01	0.61
SP5: Unemploy-ment	0.41	0.06	0.27	-0.02	1.00		0.26	0.14	0.90	0.64
SP6: Skills mismatch	0.16	0.28	0.28	0.00	0.73	1.00	0.26	0.16	*0.95*	0.68
P1: Skills Development							1.00			0.77
P2: Skills Activation							0.62	1.00		0.76
P3: Skills Matching							0.28	0.16	1.00	0.71

Notes: Numbers represent the Pearson correlations coefficients between the ESI components and the underlying indicators (for the 28 EU Member States). Correlations that are not significant at the significance level of α = 0,01 are left blank. Grey boxes show the conceptual grouping of the indicators. Very strong correlations (*i.e.* Pearson correlation coefficients greater than 0,92) are marked in italic.

Source: European Commission, Joint Research Centre, 2018.

Concluding, the first statistical coherence test corroborated the three-level structure in the ESI framework and the unidimensionality of all ESI components (sub-pillars, pillars, index). Furthermore, all fifteen indicators were found to be influential at least at the first aggregation level (sub-pillars) and for ten out of the 15 indicators, this influence arrives up to the overall index. This is a highly desirable outcome as it suggests that the information content in the majority of the underlying indicators is maintained at all levels of aggregation in the ESI framework.

3.2 Second statistical coherence test in the ESI framework

A second coherence test aims at assessing whether the five indicators that were found not to be significantly correlated with the overall index, hence they do not explain a sufficient amount of variation in the ESI scores, are important in a different way in the overall index, for example by influencing the overall ESI ranking. These five indicators are two indicators related to skills development (pre-primary pupil-to-teacher ratio, share of population attending vocational training), one indicator related to skills activation (activity rate aged 20-24) and two indicators related to skills matching (underemployed part-time workers and proportion of low wage earners).

The test consists of assessing how country ordering changes when these indicators are omitted one-at-a-time from the calculation. Table 4 presents the results of this second coherence test.

Two indicators are found to be somewhat more influential in this test: the share of population attending vocational training (VET students) and underemployed part-time workers. When omitting *VET students* from the ESI framework, Croatia loses five positions, while Malta gains four positions. Four more countries – Poland, Lithuania, Slovakia and Hungary – change 3 positions. When omitting *underemployed part-time workers*, Croatia loses five positions while the Netherlands and the United Kingdom gain five positions. Three more countries (Denmark, Poland and Austria) change 3 positions.

Excluding ISCED 5-8 proportion of low wage earners from the ESI framework does not have a noteworthy impact on the ESI ranking but it impacts the Skills Matching results for several countries: Croatia loses 8 positions, Lithuania gains five positions, and Estonia and Latvia gain four positions.

Finally, excluding the remaining two indicators – pre-primary pupil-to-teacher ratio, activity rate aged 20-24 – from the ESI framework has no significant impact on the ESI ranks.

Table 4. Second statistical coherence test in the ESI: excluding one-at-a-time selected indicators

		European Skills Index without:						**Skills Matching pillar without:**
	Rank order	**Pre-primary pupil-to-teacher ratio**	**VET students**	**Activity rate (aged 20-24)**	**Under-employed part-time workers**	**ISCED 5-8 proportion of low wage earners**	Rank order	**ISCED 5-8 proportion of low wage earners**
Higher level of skills	CZ	0	-2	0	-2	0	CZ	0
	FI	0	0	-3	0	0	MT	-2
	SE	-1	2	-1	2	-1	LU	1
	LU	1	0	2	-1	1	HU	1
	SI	0	-2	2	-1	-1	PL	0
	EE	0	1	0	-2	1	FI	-2
	DK	0	1	0	3	0	HR	-8
	PL	0	-3	-1	-3	0	SK	2
	DE	-1	0	-2	0	0	SI	0
	AT	1	0	-2	3	-1	BG	0
	LT	0	3	3	-2	1	EE	4
	HR	-1	-5	-2	-5	-3	RO	1
	SK	-1	-3	3	-2	0	SE	-3
	LV	-1	1	1	2	2	BE	-3
	NL	3	0	-1	5	1	DE	1
	MT	-4	4	-2	0	-1	LV	4
	HU	1	3	2	-1	1	DK	-1
	BE	1	-1	1	-1	0	LT	5
	UK	1	1	0	5	0	AT	0
	FR	1	0	0	0	0	IT	-1
	PT	0	0	0	0	0	NL	1
	IE	0	0	-1	0	0	FR	0
	BG	0	-1	1	-2	0	PT	-1
	CY	0	1	-1	1	-1	UK	1
	RO	0	0	1	1	1	IE	0
	IT	0	0	-2	0	0	CY	0
Lower level of skills	EL	-1	-1	1	-1	0	ES	0
	ES	1	1	1	1	0	EL	0
		Number of countries shifting 3 positions or more						
		2	***6***	***3***	***6***	***1***		***6***

Notes: The five indicators that were found not to be statistically related to the overall index (although statistically related to their own sub-pillar) are included in this analysis. Numbers represent shifts in rank in ESI when an indicator is excluded from the framework. Positive numbers imply improvement in a country's position; negative numbers imply deterioration in a country's position. Changes equal to 3 positions or greater are highlighted.

Source: European Commission, Joint Research Centre, 2018.

3.3 JRC recommendations based on the two statistical coherence tests

Generally, the results of the two statistical coherence tests (correlation analysis and impact of excluding from the ESI framework the least influential indicators one-at-a time) suggest that the conceptual grouping of the indicators into six sub-pillars and three pillars is statistically confirmed, and that the index is in general influenced by most underlying indicators. Ten out of the 15 indicators are positively and significantly correlated with the overall index. The remaining five indicators that are found to be the least influential at the index level are: Pre-primary pupil-to-teacher ratio (#1), VET students (#5), Activity rate (aged 20-24) (10), Underemployed part-time workers (#12), and Proportion of low wage earners (#14). However, given that these five indicators are influential at the first two aggregation levels (sub-pillars and pillars), their inclusion in the ESI framework is to a large extent corroborated by the analysis. Interestingly, four out of the five indicators discussed are the newly introduced indicators in the framework (see previously Table 1). Although conceptually enriching and statistically informative up to the second aggregation level, the information content of these indicators does not arrive sufficiently at the index level.

The second coherence test offered an additional perspective on the impact of these five indicators, showing that three of them – pre-primary pupil-to-teacher ratio in Skills Development, activity rate aged 20-24 in Skills Activation, and ISCED 5-8 proportion of low wage earners in Skills Matching have a low impact on the ESI country ordering; nevertheless the latter indicator on the low wage earners has an impact on the Skills Matching ranks.

The JRC recommendation to the ESI development team is to carefully monitor how these three indicators behave in the coming releases of the index and eventually to fine-tune the framework in this respect.

4 Impact of modelling assumptions in the ESI

An important part of the ESI statistical audit is to assess the effect of varying modelling assumptions inside plausible ranges. The rationale for the choices made in the development of the European Skills Index is manifold:

- Expert opinion, literature review and statistical analysis are behind the selection of the fifteen ESI indicators, 6 sub-pillars and 3 pillars;
- common practice and ease of interpretation suggests the use of a min-max normalization approach in the [0–1] range;
- the use of fixed bounds in the min-max normalisation allows for monitoring progress over time;
- the treatment of outliers is driven by statistical analysis and aims at avoiding polarised scores;
- simplicity and parsimony criteria seem to advocate for not imputing missing data;
- the use of calibrated weights aims at ensuring that each indicator is roughly equally informative with respect to the theme covered by the pillar;
- and finally adopting geometric averaging at the highest aggregation level, where skills development, skills activation and skills matching are at play, is desirable because it implies that the level of priority given to an ESI pillar is not invariant to the level of attainment.

Despite the well-founded rationale for these choices made during the ESI development, there is an unavoidable subjectivity (or uncertainty) in these choices, which is accounted for in the robustness assessment carried out by the JRC. More precisely, the uncertainly analysis is conducted herein in order to allow for the **joint** analysis of the impact of the modelling choices on the ESI results, resulting in error estimates and confidence intervals calculated for the ESI 2018 country ranks. This analysis complements and extends the uncertainty analysis conducted by the ESI developers as it helps to evidence whether the space of alternatives explored by the developers (three assumptions tested and 5 scenarios run) is wide enough to draw robust inference when benchmarking the performance of EU Member States skills systems.

As suggested in the relevant literature on composite indicators,([4]) the robustness assessment was based on Monte Carlo simulation and multi-modelling approaches, applied to 'error-free' data where eventual errors and typos have already been corrected in a preliminary stage. In particular, the five key modelling issues considered in the assessment of the ESI were the treatment of missing data, the normalisation method, the bounds used in the min-max normalisation, the aggregation formula at the pillar level and finally the pillar weights (see Table 5 for a summary of the five types of uncertainties considered).

Missing values. The ESI developers, for transparency and replicability and following common practice on composite indicator development, opted not to estimate missing data for three cases: Ireland's value on pre-primary pupil to teacher ratio (ind 1), and Croatia's and Malta's values on Qualification mismatch (ind 15). Technically, the 'no imputation' choice in an average is equivalent to replacing an indicator's missing value for a given country with the respective sub-pillar score. Hence, the available data (indicators) in the incomplete pillar may dominate, sometimes biasing the ranks up or down. Furthermore, the 'no imputation' choice might encourage countries not to report low data values. To test the impact of the 'no imputation' choice, the JRC estimated the

([4]) Saisana et al., 2005; Saisana et al., 2011 ; Vértesy 2016; Vértesy and Deiss, 2016 three missing values in the ESI dataset using the Expectation Maximization (EM) algorithm that was applied in the entire set of 15 indicators. ([5])

Normalisation method. Raw data for the fifteen ESI indicators were put in a common scale from 0.0 (lowest performance) to 1.0 by using the min-max normalisation method ([6]). The rationale for this choice was to ease communication with the general public. The previous version of the index (called the Making Skills Work Index) was based on the z-scores approach ([7]). To assess the impact on the ESI ranks of the normalisation method, the JRC included both the min-max and z-scores approach in the uncertainty analysis.

Bounds in the normalisation. The min-max normalisation method with fixed bounds was adopted by the ESI developers in order to allow benchmarking performance of countries skills systems over time. Minimum and maximum values (goalposts) were chosen to act as the "logical worst case" and "logical best case" (or else aspirational targets), respectively, from which the ESI indicators are normalised. To test the impact of using fixed bounds as opposed to the observed minimum and maximum values over the 7 year period (2010-2016), both options were included in the analysis.

Aggregation formula. Regarding the aggregation formula at the pillar level, the ESI team opted for the geometric averaging of the three pillars which is a partially compensatory approach that rewards countries with balanced profiles and motivates countries to improve in the ESI pillars in which they perform poorly, and not just in *any* ESI pillar.([8]) This choice is in line with relevant literature that challenges the use of simple arithmetic averages at higher aggregation levels because of their fully compensatory nature, in which a comparative high advantage on a few indicators can compensate a comparative disadvantage on many indicators.([9]) To assess the impact of this compensability issue, the JRC included in the analysis both the geometric and the arithmetic averaging of the three pillars.

Weights. Monte Carlo simulation comprised 1,000 runs of different sets of weights for the three ESI pillars: Skills Development, Skills Activation and Skills Matching. The weights were assigned to the pillars based on uniform continuous distributions centred in the reference values (plus/minus 25%). As a result, the limit values of uncertainty for the three pillars are 22.5%–37.5% for Skills Development and Skills Activation, and 30%-50% for Skills Matching. Note that the equal weights assumption is included herein.

Twelve models were tested based on the combination of no imputation versus EM imputation, min-max versus z-scores normalisation, fixed bounds versus observed bounds (only applicable in the min-max normalisation), and geometric versus arithmetic averaging at pillar level. Combined with 1,000 simulations per model (perturbed weights versus fixed weights), a total of 12,000 simulations for the ESI were run.

([5]) The Expectation-Maximization (EM) algorithm (Little and Rubin, 2002; Schneider, 2001) is an iterative procedure that finds the maximum likelihood estimates of the parameter vector by repeating two steps. Step 1: The expectation E-step: Given a set of parameter estimates, such as a mean vector and covariance matrix for a multivariate normal distribution, the E-step calculates the conditional expectation of the complete-data log likelihood given the observed data and the parameter estimates. Step 2: The maximization M-step: Given a complete-data log likelihood, the M-step finds the parameter estimates to maximize the complete-data log likelihood from the E-step. The two steps are iterated until the iterations converge.

([6]) With the min-max normalization method, a country's score for each indicator is calculated by subtracting the minimum value (or lower bound) across all countries and dividing by the difference between the maximum and minimum values (or upper and lower bounds).

([7]) With the z-scores method, a country's score for each indicator is calculated by subtracting the average value (across all countries) and dividing by the standard deviation.

([8]) In the geometric average, pillars are multiplied as opposed to summed in the arithmetic average. Pillar weights appear as exponents in the multiplication. All pillar scores were greater than zero, hence there was no reason to rescale them to avoid zero values that would have led to zero geometric averages.

([9]) Munda, 2008.

Table 5. Uncertainty parameters in the ESI: Missing values, normalisation, goalposts, aggregation, weights

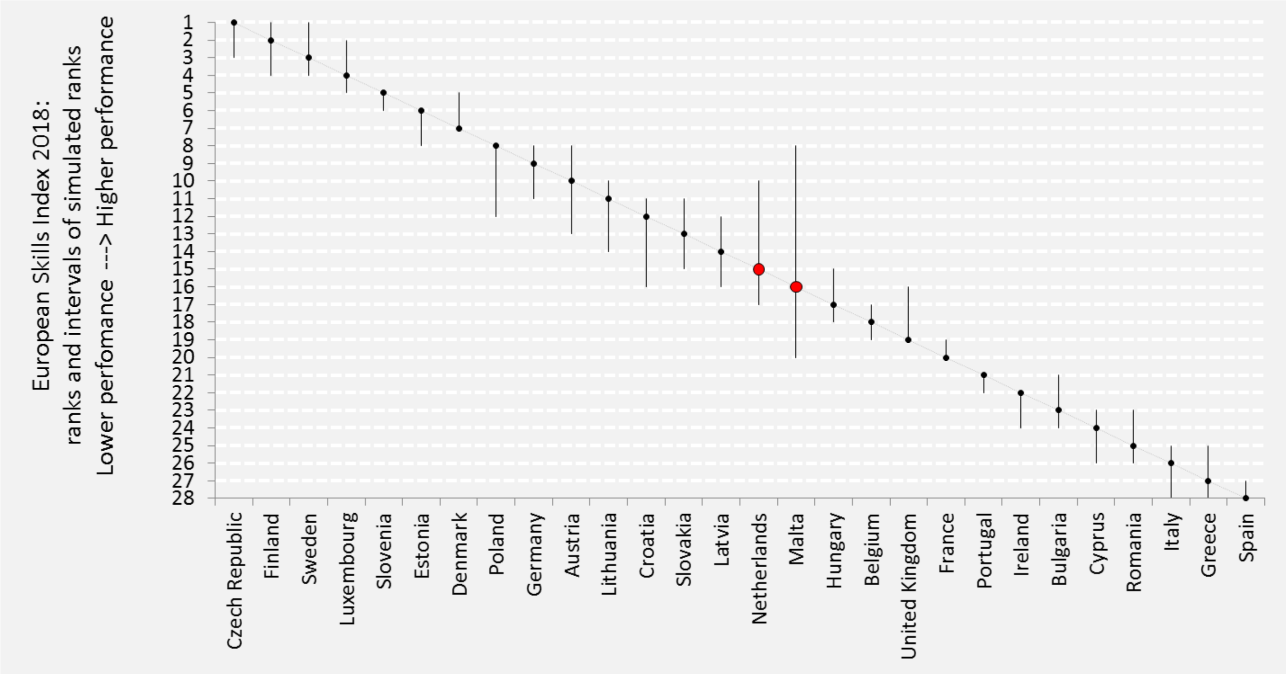
		Reference	**Alternative**
I. Uncertainty in the treatment of missing values		No estimation of missing data	Expectation Maximization (EM)
II. Uncertainty in the normalisation method		Min-Max	z-scores
III. Uncertainty in the bounds used in normalisation		Fixed bounds (goalposts)	Observed minimum and maximum values
IV. Uncertainty in the aggregation formula at pillar level		Geometric average	Arithmetic average
V. Uncertainty intervals for the ESI pillar weights		Reference value for the weight	Distribution assigned for robustness analysis
	1. Skills Development	0.3	U[0.225, 0.375]
	2. Skills Activation	0.3	U[0.225, 0.375]
	3. Skills Matching	0.4	U[0.300, 0.500]

Source: European Commission, Joint Research Centre, 2018.

The main results of the robustness analysis for the ESI are presented in Figure 4, which shows the distribution of ESI ranks over the 12,000 Monte Carlo simulations performed. The height of the error bars in the plots represents the uncertainty in the country ranks associated to the five types of uncertainty, namely in the treatment of missing values, the normalisation method, the bounds used in the min-max method, the aggregation formula and the weights at pillar level. The dot represents the baseline scenario (original ESI rank) for each EU Member State.

Overall, the magnitude of uncertainty in the ESI ranks is modest given that for most EU Member States the simulated intervals are narrow enough for meaningful inferences to be drawn: compared to the baseline rank there is a shift of 3 positions or less for 24 of the 28 countries. However, it is also true that two country ranks vary significantly with changes in the ESI modelling assumptions. These two countries — Malta and the Netherlands — have 90% confidence interval widths of respectively 12 and 7 positions. Consequently, their ESI ranks — 16th for Malta and 15th for the Netherlands— should be interpreted cautiously and certainly not taken at face value. Follow, Austria and Croatia with confidence interval widths of 5 positions. As expected and commonly in similar contexts of benchmarking through indices, the few countries with "sensitive ranks" are found in the middle of the distribution and have very similar baseline scores; thus, small changes in the country scores can have a very high impact on the respective ranks.

Figure 4. Robustness analysis (ESI rank, 90% confidence intervals)



Notes: Intervals (90% confidence intervals) are calculated over 12,000 simulated scenarios based on imputing (or not) missing values, normalization method, bounds used in the min-max normalization method, random weights plus/minus 25% around the reference weights, and aggregation formula at pillar level, as shown in Table 5). Malta and the Netherlands are the two countries with most sensitive ESI ranks (in red); Austria and Croatia follow.

Source: European Commission Joint Research Centre, 2018.

The robustness results presented here and the conclusions drawn are very much in line with the results and discussions offered by the ESI developers in the technical report accompanying the European Skills Index. As expected, given the higher number of scenarios tested (12,000), there are some countries for which the uncertainty intervals are slightly wider in this analysis compared to the analysis conducted in the ESI technical report (for example Finland, see Table 6). Exploring a high number of modelling scenarios has helped to confirm that the five scenarios considered in the ESI technical report, although very limited in number, they are representative of a much wider space of uncertainties.

When completing the big picture with the uncertainties around the country ranks, it is possible to distinguish five groups of countries: top performers varying within the top 7 positions (with scores above 0.67); a small group of three upper-middle countries follows; a big group of middle performers varying approximately between the 11th and the 21st positions (with scores 0.45-0.61); a group of lower-middle performers varying between the 22nd and the 25th position (with scores 0.31-0.36); and finally a small group of lower performing countries (with scores 0.23-0.25).

Table 6. ESI 2018: Original scores, ranks, intervals and JRC 90% confidence intervals

	Original score	Original rank	Original interval of the 5 scenarios in the ESI report	90% Interval over the 12 000 JRC scenarios	Performance group
Czech Republic	0.75	1	[1 ,3]	[1 ,3]	Higher
Finland	0.72	2	[2 ,2]	[1 ,4]	
Sweden	0.72	3	[1 ,4]	[1 ,4]	
Luxembourg	0.71	4	[3 ,5]	[2 ,5]	
Slovenia	0.69	5	[5 ,6]	[5 ,6]	
Estonia	0.68	6	[6 ,8]	[6 ,8]	
Denmark	0.67	7	[4 ,7]	[5 ,7]	
Poland	0.62	8	[8 ,12]	[8 ,12]	Upper middle
Germany	0.62	9	[9 ,11]	[8 ,11]	
Austria	0.62	10	[7 ,10]	[8 ,13]	
Lithuania	0.61	11	[10 ,15]	[10 ,14]	Middle
Croatia	0.60	12	[12 ,18]	[11 ,16]	
Slovakia	0.59	13	[11 ,15]	[11 ,15]	
Latvia	0.59	14	[13 ,16]	[12 ,16]	
Netherlands	0.58	15	[10 ,17]	[10 ,17]	
Malta	0.56	16	[9 ,19]	[8 ,20]	
Hungary	0.55	17	[16 ,17]	[15 ,18]	
Belgium	0.53	18	[18 ,19]	[17 ,19]	
United Kingdom	0.52	19	[15 ,19]	[16 ,19]	
France	0.48	20	[20 ,20]	[19 ,20]	
Portugal	0.45	21	[21 ,21]	[21 ,22]	
Ireland	0.36	22	[22 ,24]	[22 ,24]	Lower middle
Bulgaria	0.33	23	[22 ,24]	[21 ,24]	
Cyprus	0.32	24	[23 ,26]	[23 ,26]	
Romania	0.31	25	[23 ,25]	[23 ,26]	
Italy	0.25	26	[25 ,28]	[25 ,28]	Lower
Greece	0.23	27	[27 ,28]	[25 ,28]	
Spain	0.23	28	[26 ,28]	[27 ,28]	

Notes: Intervals (90% confidence intervals) are calculated over 12,000 simulated scenarios based on imputing (or not) missing values, normalization method, bounds used in the min-max normalization method, random weights plus/minus 25% around the reference weights, and aggregation formula at pillar level, as shown in Table 5). Malta and the Netherlands are the two countries with most sensitive ESI ranks (in red); Austria and Croatia follow. Countries are ordered from the higher to the lower performance on the European Skills Index.

Source: European Commission Joint Research Centre, 2018.

Complementary to the uncertainty analysis, sensitivity analysis has been used to identify which of the ESI's five modelling assumptions have the highest impact on the four countries with the most volatile ESI ranks (Table 7). The estimation of missing values using the EM approach has an impact only on Malta, changing 3 positions from 16th down to 19th place if the missing value on Qualification mismatch is estimated statistically. Ireland's and Croatia's ESI rank is not affected when their single missing value is estimated via the EM approach. Malta, because of its diverse performance across the three pillars (0.29 for Skills Development, 0.62 for Skills Activation and 0.86 for Skills Matching) is affected by all assumptions in the ESI development. The Netherlands and Austria are affected by the choice of weighting and aggregation at pillar level. Finally, Croatia is affected modestly by the choice of the normalisation method (from 12th down to 16th had the z-scores approach been used).

Table 7. Sensitivity analysis: impact of uncertainties on four countries with most sensitive ESI ranks

	Malta	Netherlands	Croatia	Austria
I. Uncertainty in the treatment of missing values	YES			
II. Uncertainty in the normalisation method	YES		YES	
III. Uncertainty in the bounds used in normalisation	YES			
IV. Uncertainty in the aggregation formula at pillar level	YES	YES		YES
V. Uncertainty intervals for the ESI pillar weights	YES	YES		YES

Source: European Commission Joint Research Centre, 2018.

All in all, the published ESI 2018 ranks are reliable and for the vast majority of EU Member States (24 out of 28) the simulated 90% confidence intervals are narrow enough for meaningful inferences to be drawn. ESI ranks for Malta and the Netherlands in particular, and to some extent for Croatia and Austria should be considered with caution when developing narratives around those ranks. Given the sensitivity of Malta's ESI rank to the estimation of the missing value on Qualification mismatch, the JRC recommendation to the index developers is to find a suitable way for approximating the missing value (for example by contacting Malta's national statistical office). For the readers and users of the ESI 2018 report, the recommendation is to consider country ranks in the ESI 2018 not only at face value but also within the 90% confidence intervals in order to better appreciate to what degree a country's rank depends on the modelling choices.

5 Skills Development, Activation and Matching: From three concepts to one single number or are we missing something in-between?

This section aims at touching upon the added value of the European Skills Index as a summary measure of the three pillars and at discussing how the statistical associations between the three pillars can be used to inform policies at national level in the EU.

Table 8 shows that the ESI ranking and any of the three pillar rankings differ by 7 positions or more for at least 15% of the Member States. This finding suggests that there is an added value in referring to the ESI results in order to identify aspects of countries' skills system that do not directly emerge by looking into the three pillars separately. At the same time, this outcome points to the value of examining individual pillars (and all underlying ESI components) on their own merit in order to see which aspects (indicators) of a skills system are driving a Member State's performance.

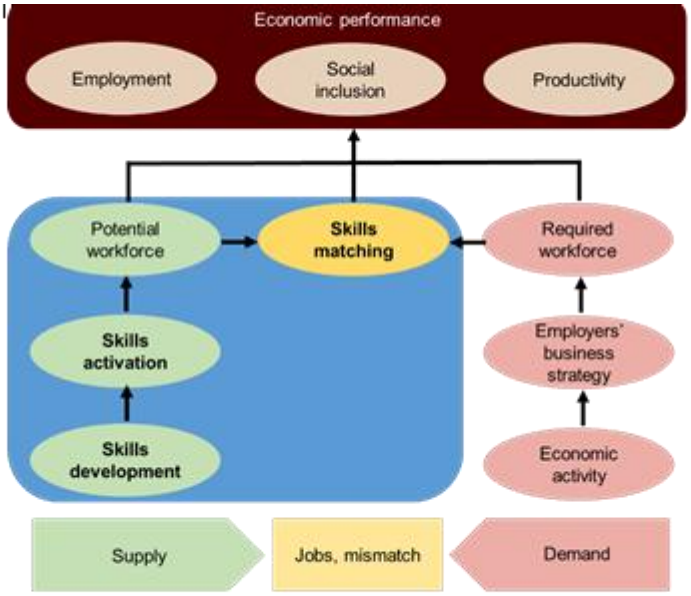
Table 8. Distribution of differences between pillars and ESI rankings

Shifts with respect to ESI	Skills Development pillar	Skills Activation pillar	Skills Matching pillar
more than 10 positions	4%	11%	14%
7-10 positions	11%	18%	14%
4-6 positions	18%	21%	36%
1-3 positions	64%	43%	32%
0 positions	4%	7%	4%
Total	100%	100%	100%
more than 7 positions	***15%***	***29%***	***28%***

Source: European Commission Joint Research Centre, 2018.

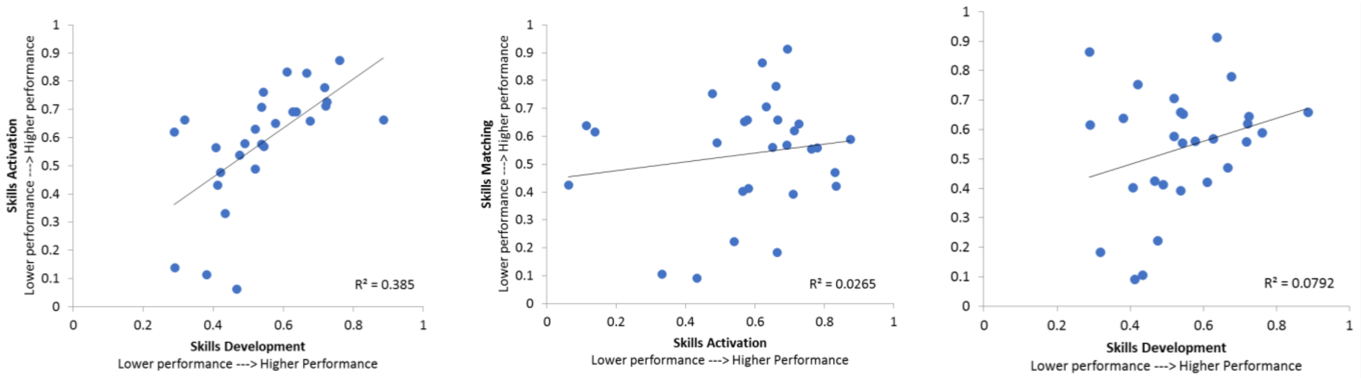
The theoretical framework for the skills system underpinning the European Skills Index places skills development and skills activation under the same building block (the supply side), whilst skills matching belongs to another building block that lies between the supply and the demand side (Figure 5). This theoretical framework receives statistical confirmation through the way the three pillars have been calculated in the ESI model. In fact, Figure 6 shows that there is a good linear relationship between the ESI skills development scores and the ESI skills activation scores (left panel), against a weak and diffuse pattern between the ESI skills matching scores and either the ESI skills development or the ESI skills activation scores.

Figure 5. Theoretical framework for the skills system



Source: Cambridge Econometrics (European Skills Index 2018 – technical report), 2018.

Figure 6. ESI 2018: Skills Development vs Skills Activation vs Skills Matching



Notes: The dots represent country scores for the 28 EU Member States for the three ESI pillars: skills development, skills activation, skills matching.

Source: European Commission Joint Research Centre, 2018.

Building further on the association between skills development and skills activation, we combine the two pillars in one – what we call here the Skills Formation – adopting the ESI logic whereby geometric averaging is more suitable at high aggregation level and assigning equal weights to both pillars. We then plot Skills Formation versus Skills Matching in Figure 7.

First, the plot shows a diffuse scatter of points, suggesting a negligible association between these two main elements of a skills system ([10]). Hence aiming to kill two birds with one stone by identifying policies that can promote skills formation and skills matching at the same time may not necessarily produce many results.

([10]) The Pearson correlation coefficient between Skills Formation and Skills Matching is only 0.23, which is not statistically signficant at 1% level for a sample size of 28 countries.

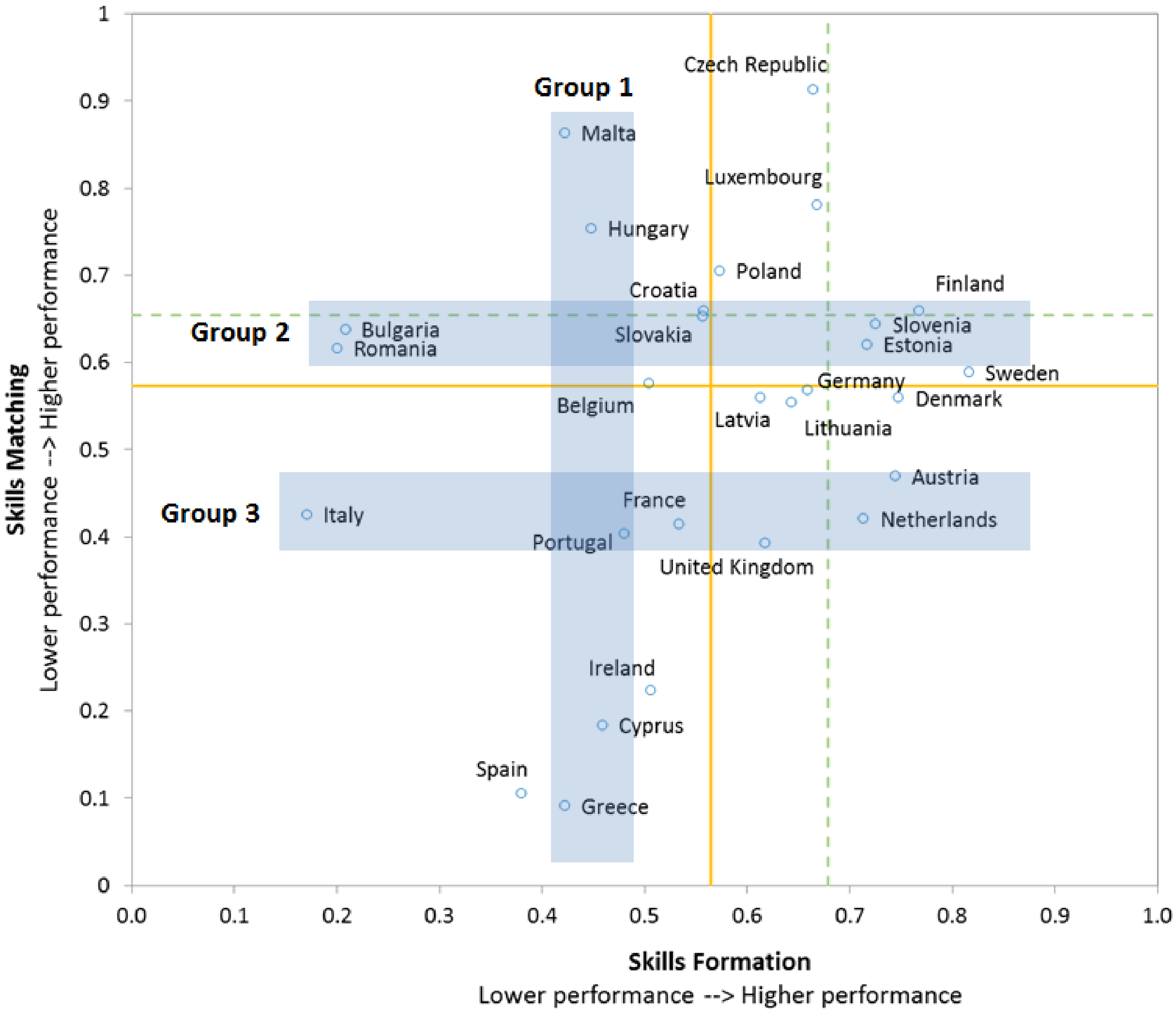
Second, several policy lessons may emerge by analysing countries on the top right and bottom left quadrants. The solid lines in the plot represent the median values of the scores in each series across the 28 EU Member States; the dashed lines represent the 75^{th} percentiles. Countries close to or beyond the two dashed lines at the top right side – Czech Republic, Luxembourg, Finland, Slovenia, Estonia – are those countries where most good practices for both Skills Formation and Skills Matching are to be found. These countries are all in the top positions in the overall ESI as well. Interestingly, Finland, although it ranks 2^{nd} behind Czech Republic in the overall ESI index, it has a more balanced profile, compared to Czech Republic, where both Skills Formation and Skills Matching are in the top 25% of the best scores. On the other hand, countries on the low left side of the graph may need to take action to adopt policies for promoting skills formation and skills matching; and it is very likely that there few, if any at all, policies that can achieve both objectives at the same time.

Third, analysing EU Member States at similar levels of skills formation or skills matching, in Figure 7, provides interesting policy insights and comparisons:

- Group 1 countries have very similar skills formation scores but rather different skills matching scores. Malta and Greece stand at the two sides of this group. Hence, there may be policies related to skills matching in Malta (the country scores close to 0.9 on this aspect) that can inspire action in Greece and Cyprus.
- Groups 2 and 3 consist of countries with similar levels of skills matching but very different levels of skills formation. In Group 2, effective policies on skills formation in Finland, Slovenia and Estonia may be helpful for gauging how policies can be shaped in Bulgaria and Romania. In Group 3, Austria and the Netherlands may be used as good examples for “what works” policies on advancing skills formation in Italy.

All in all, the JRC recommendation for the best strategy to be adopted in order to get further insights on policies that work and where bottlenecks exist in the EU when it comes to skills development, skills activation and skills matching is to use the entire ESI framework of indicators, sub-pillars and pillars, together with the Skills Formation component proposed herein, and under the umbrella of the goals and actions included in the 2016 New Skills Agenda for Europe.

Figure 7. Skills Formation vs Skills Matching



Notes: Skills Formation (x-axis) is calculated by the JRC taking the geometric average of the ESI Skills Development and ESI Skills Matching scores. Solid lines represent median values (across the 28 EU Member States). Dashed lines represent 75th percentiles.

Source: European Commission Joint Research Centre, 2018.

6 Conclusion

The Commission services developed the European Skills Index (ESI) with a view to measure the performance of national skills systems in the EU. The JRC statistical audit has delved around in the workings of the ESI framework to assess the statistical properties of the data, and the methodology used in the index construction. Overall the ESI framework is well-constructed, into which a lot of thought has clearly been put. One of the greatest strengths is the amount of original research into the multiple facets of skills systems in the EU Member States, as well as the transparency and detail of all data populating the ESI framework and the rationale for all choices made. This transparency and detail in the source information lends considerable credibility to the European Skills Index as an ensemble of carefully selected indicators and opens the data and the ESI components for use by policy analysts and researchers alike.

The key findings of the statistical assessment conducted herein are the following:

First, two statistical coherence tests suggest that the **conceptual grouping** of the 15 indicators into six sub-pillars, three pillars and an overall index is statistically confirmed, and that the index is equally influenced by the three main pillars: Skills development, Skills activation and Skills matching. Ten out of the 15 indicators in the ESI framework are also found to be influential all the way up to the index level. Nevertheless, three indicators – *pre-primary pupil-to-teacher ratio* in Skills Development, *activity rate aged 20-24* in Skills Activation, and *proportion of low wage earners* in Skills Matching have a low impact on the ESI country ordering and can explain only a small (negligible) amount of variation in the ESI scores. Although these indicators are conceptually enriching the ESI framework and their statistical impact arrives up to the first and/or second aggregation levels (thanks to the ESI developers' choice to calibrate the weights), it is recommended to carefully monitor how these three indicators behave in the coming releases of the index and eventually to fine-tune the framework in this respect.

Second, the results offer **statistical justification for the theoretical framework** underpinning the European Skills Index, which places skills development and skills activation under the same building block (the supply side), whilst skills matching belongs to another building block that is found between the supply and the demand side. This statistical justification comes from the good linear relationship between the ESI skills development scores and the ESI skills activation scores; instead there seems to exist a weak and diffuse pattern between the ESI skills matching scores and either the ESI skills development or the ESI skills activation scores.

Third, the ESI dataset has very good **data coverage** and 85% of the data refer to 2015 or 2016. Only three values are missing: Ireland's value on pre-primary pupil to teacher ratio in Skills Activation, and Croatia's and Malta's values on Qualification mismatch in Skills Matching. Uncertainty and sensitivity analysis have shown that it is important to find a reliable estimate for Malta's value on Qualification mismatch because of the impact on Malta's ESI rank. Ireland's and Croatia's ESI rank is not affected by the way missing values are estimated.

Forth, treating the outlier value for Greece for long-term unemployment rate (capping it from 17% down to 9.5%) is not required given that by adopting the goalposts during the normalisation step the lower bound (worst case) for that indicator is set at 10%. To ease communication to the wider audience, this winsorisation step can be removed; yet it is important to monitor in next releases if the normalised (with the use of goalposts) indicator values satisfy the double criterion for skewness and kurtosis.

Fifth, the developers choice to adopt the min-max normalisation method with a view to ease communication with the wider public, compared to the z-scores used in the previous beta-version of the index, does not affect significantly the overall ESI results (there is a modest impact on Malta's and Croatia's ESI ranks).

Sixth, the developers choice to calibrate the **weights** for the three pillars (0.3, 0.3 and 0.4) helps to ensure that all three pillars – Skills Development, Skills Activation and Skills Matching – are placed on equal footing when it comes to calculating a summary measure for the performance of a country's skills system. Furthermore, adopting a suitable **aggregation** formula (geometric averaging) to combine the three pillars allows for the level of priority given to an ESI pillar to depend on the level of attainment (more priority given to the pillar where country has lower performance).

Seventh, the tests helped to single out two countries — Malta and the Netherlands — with ESI ranks that are very **sensitive to the modelling choices** and hence these ranks should be interpreted cautiously. Some caution, though much less, is also needed for the ESI ranks for Croatia and Austria. On the other hand and compared to the baseline ESI rank, there is a shift of 3 positions or less for 24 of the 28 countries when varying five key assumptions in the ESI development over 12,000 simulations. Thereafter, the ESI framework allows to draw meaningful inferences on the performance of skills systems in the vast majority of EU countries. Furthermore, exploring a high number of modelling scenarios, and their joint effect, has helped to confirm that the five scenarios considered in the ESI technical report, although very limited in number, they are representative of a much wider space of uncertainties.

Eighth, when analysing ESI country ranks in the realm of the inherent uncertainties, it is possible to distinguish **five performance groups**: top performers varying within the top 7 positions (with scores above 0.67); a small group of three upper-middle countries follows; a big group of middle performers varying approximately between the 11th and the 21st positions (with scores 0.45-0.61); a group of lower-middle performers varying between the 22nd and the 25th position (with scores 0.31-0.36); and finally a small group of lower performing countries (with scores 0.23-0.25). Hence, these five performance groups are worthy discussing in detail when communicating the ESI results.

Ninth, results show that there is an **added value in referring to the ESI results** in order to identify aspects of countries' skills system that do not directly emerge by looking into the three pillars separately. In fact, the ESI ranking and any of the three pillar rankings differ by 7 positions or more for 15% up to 29% of the Member States.

Tenth, **relevant and actionable policy insights may emerge** when analysing EU Member States that have similar levels of skills formation or skills matching. Skills Formation is an additional component of a country's skills system proposed herein, which is calculated by aggregating together the two ESI pillars that belong to the supply side: Skills Development and Skills Activation. Best practices and policies related to skills matching in Malta may inspire action in Greece and Cyprus. Effective policies on skills formation in Finland, Slovenia and Estonia may be helpful for gauging how policies can be shaped in Bulgaria and Romania. Finally, Austria and the Netherlands may be used as good examples for "what works" policies on advancing skills formation in Italy.

All things considered, the present JRC audit findings confirm that the European Skills Index 2018 meets international quality standards for statistical soundness, which indicates that the ESI framework offers a sound starting point for more informed discussions on skills systems at the country level in the EU. The readers and policy analysts of the European Skills Index should hence go beyond the overall index scores and duly take into account the individual indicators and pillars on their own merit. By doing so, country-specific strengths and challenges in developing, activating or matching skills to the job market can be identified and serve as an input for data-informed policy analysis. The European Skills Index cannot possibly serve as the ultimate and definitive yardstick of EU countries skills systems. Instead, the ESI best represents an ongoing attempt by CEDEFOP to help focus the policy discussions on the multiple facets of national skills systems in the EU, continuously adapting the European Skills Index framework to reflect the improved availability of statistics and the theoretical advances in the field.

7 References and related reading

Becker, W., M. Saisana, P. Paruolo, and I. Vandecasteele. 2017. 'Weights and Importance in Composite Indicators: Closing the Gap'. *Ecological Indicators* 80: 12–22.

Groeneveld, R. A. and G. Meeden. 1984. 'Measuring Skewness and Kurtosis'. *The Statistician* 33: 391–99.

Little, R. J. A. and D. B. Rubin. 2002. *Statistical Analysis with Missing Data*. 2nd edition. Hoboken, NJ: John Wiley & Sons, Inc.

Munda, G. 2008. *Social Multi-Criteria Evaluation for a Sustainable Economy.* Berlin Heidelberg: Springer-Verlag.

OECD/EC JRC (Organisation for Economic Co-operation and Development/European Commission, Joint Research Centre). 2008. *Handbook on Constructing Composite Indicators: Methodology and User Guide.* Paris: OECD.

Paruolo, P., M. Saisana, and A. Saltelli. 2013. 'Ratings and Rankings: Voodoo or Science?' *Journal of the Royal Statistical Society* A 176 (3): 609–34.

Saisana, M., B. D'Hombres, and A. Saltelli. 2011. 'Rickety Numbers: Volatility of University Rankings and Policy Implications'. *Research Policy* 40: 165–77.

Saisana, M., A. Saltelli, and S. Tarantola. 2005. 'Uncertainty and Sensitivity Analysis Techniques as Tools for the Analysis and Validation of Composite Indicators'. *Journal of the Royal Statistical Society* A 168 (2): 307–23.

Saltelli, A., M. Ratto, T. Andres, F. Campolongo, J. Cariboni, D. Gatelli, M. Saisana, and S. Tarantola. 2008. *Global Sensitivity Analysis: The Primer*. Chichester, England: John Wiley & Sons.

Schneider, T. 2001. 'Analysis of incomplete climate data: Estimation of mean values and covariance matrices and imputation of missing values. Journal of Climate, 14, 853–871.

Vertesy, D., Deiss, R. 2016. The Innovation Output Indicator 2016. Methodology Update. EUR 27880. European Commission, Joint Research Centre.

Vértesy, D. (2016, July). A Critical Assessment of Quality and Validity of Composite Indicators of Innovation. Paper presented at the OECD Blue Sky III Forum on Science and Innovation Indicators. Ghent, 19-21 Sept 2016.

List of figures

List of tables

KJ-NA-29352-EN-N

JRC Mission

As the science and knowledge service of the European Commission, the Joint Research Centre's mission is to support EU policies with independent evidence throughout the whole policy cycle.

EU Science Hub
ec.europa.eu/jrc

@EU_ScienceHub

EU Science Hub - Joint Research Centre

Joint Research Centre

EU Science Hub

Publications Office

doi:10.2760/803749

ISBN 978-92-79-93710-1