On-line Appendix

Robustness of the WHR Framework for Estimating Well-being Efficiency to the Exclusion of Inputs

We start our robustness checks by dropping our current inputs one at a time from the baseline model. Our aim is to check whether models with a partial set of inputs from the WHR framework provide significantly different results.

The results indicate that our well-being efficiency scores do not depend on the inclusion of one input or another, they are remarkably robust to dropping inputs. Appendix Table 1 reports the coefficients of Spearman's rank test between the scores from the baseline model, and those from the trimmed models. The Spearman's rank test checks whether the ranking of countries resulting from two variables are statistically related. We find that in the worst case scenario, when we omit freedom of choice, the coefficient is 0.914 (significant at 1 per cent). In all other cases the coefficients range between 93 per cent and 97 per cent. We also estimated the correlations between trimmed models and found the coefficients are still above 83 per cent (this part of the correlation matrix has been omitted for brevity).

The results are similar when we use the standard Pearson's correlation test: the correlation coefficients are all above 96 per cent (significant at 1 per cent), except when we exclude freedom of choice (model CRS_TE_3) for which the correlation coefficient is 92 per cent (significant at 1 per cent). In sum, the well-being efficiency scores are stable to variations in inputs.

Additional inputs

Although the WHR framework provides empirical guidance to identify relevant variables to explain subjective well-being worldwide, the list may be incomplete: after all, 25 per cent of the variance of subjective well-being remains unexplained in the WHR regression model. Omitted variables, such as inequality, optimism, unemployment, and education, could contribute meaningfully. Education in particular was included in both Cordero *et al.* (2021) and Nikolova and Popova (2021).

To account for this possibility, we check how total well-being efficiency changes when we expand the baseline model with additional variables one at a time. The additional variables are those used in section 4 where we study the correlates of well-being efficiency.

The results indicate that adding inputs does not significantly affect our well-being efficiency ranks or scores. Appendix Table 1 reports the coefficients of Spearman's rank test and Pearson's correlation between the baseline well-being efficiency scores and new scores produced with additional inputs (listed in rows). Both tests provide fairly high coefficients. The smallest coefficient of the Spearman's rank test is 72 per cent when we include social expenditures and the population dependency ratio

Appendix Table 1: Spearman's Rank Test and Pearson's Correlations Between the Results of the Baseline Model and Trimmed Models.

Omitted inputs Observations	Correlation Coefficients	
	Spearman	Pearson
GDP per capita	0.932	0.968
Social support	0.937	0.970
Healthy life expectancy	0.970	0.980
Freedom of choice	0.914	0.921
Generosity	0.945	0.972
Absence of Corruption	0.948	0.967

Note: All coefficients are statistically significant at 1%. The number of obeservation is 126.

Source: authors' own elaboration. Data sourced from WHR 2021.

Appendix Table 2: Sensitivity of Well-Being Efficiency Scores to the Inclusion of Additional Inputs

Added Inputs	Correlation coefficients		Observations
	Spearman	Pearson	
Unemployment rate	0.89	0.96	126
Gini	0.97	0.99	126
Years of School	0.98	0.99	111
Optimism	0.99	0.99	126
Quality of Governance	0.81	0.92	126
Social Expenditures	0.72	0.9	120

Note: All coefficients are statistically significant at 1%

Source: authors' own elaboration. Data sourced from WHR 2021 and others that are documented in Section 4.

in the model. All coefficients are statistically significant at 1 per cent. The number of observations used to compute efficiency scores changes because of missing data. In those cases, we recomputed the baseline well-being efficiency scores in order to compute correlations on the same set of observations.

Excluding outliers

A potential pitfall of DEA is that extreme values in the data can have large impacts on the computed scores. To address this concern we repeat our analysis after dropping outlying values.

We analyse two cases in which we consider first the middle 98 per cent and then the middle 80 per cent of the distributions of each considered variable. In the first case, we drop all observations with values in the top or bottom 1 per cent of any of the variables. This is why the sample reduces from 126 to 115 observations. In the most conservative case, we drop all the observations with values in the top or bottom 10 per cent of any of the variables. Consequently, the sample available for the analysis drops to 39 countries. Further cuts are not possible because this would lead to samples that are too small.

The results are not sensitive to dropping outlying countries. The correlation between well-being efficiency before and after excluding outliers is remarkably high (see Appendix Table 3). In the most conser-

Appendix Table 3: Sensitivity of the Results to Outlying Values.

	Spearman's rank test		Pearson's correlation test	
	Coefficient	Obs.	Coefficient	Obs.
middle 98%	0.96	115	0.95	115
middle 80%	0.99	39	0.97	39

Note: All coefficients are statistically significant at 1%

Source: authors' own elaboration. Data sourced from WHR 2021.

vative case (dropping the top and bottom 10 per cent), the Pearson correlation coefficient is 97 per cent (significant at 1 per cent), and Spearman's correlation is 99 per cent (significant at 1 per cent). When we restrict the analysis to the middle 98 per cent, the Pearson coefficient is 95 per cent (significant at 1 per cent), and the Spearman's is 96 per cent (significant at 1 per cent).