Competitiveness Review

# Towards a new measure of a country's competitiveness: applying canonical correlation

Lars Wenzel André Wolf

# Towards a new measure of a country's competitiveness: applying canonical correlation

Lars Wenzel and André Wolf
*Hamburg Institute of International Economics, Hamburg, Germany*

Received 12 September 2014
Revised 17 December 2014
5 March 2015
12 June 2015
Accepted 3 July 2015

## Abstract

**Purpose** – This paper aims to develop and implement a new approach of ranking countries according to their level of economic competitiveness.

**Design/methodology/approach** – Competitiveness is conceptualized as the degree of participation in global integration, measured by the levels of exports, capital inflows and immigration. Based on a wide range of development indicators, composite indicators are constructed by means of a principal component analysis. In turn, these composites enter a canonical correlation analysis and estimated canonical weights are used as weights in the calculation of index scores.

**Findings** – Measures of institutional and political quality are shown to be most closely connected to a country's ability to export and to attract foreign factors of production. Inflows of foreign capital turn out to correlate much less with our weighting composites than export volumes.

**Research limitations/implications** – Regarding prospects for future research, a broad range of further applications of this approach in the macro area is conceivable, including measures of a country's quality of schooling, the quality of its health care system or of its governmental institutions.

**Originality/value** – In comparison to prevailing macroeconomic indices, this method of index construction is superior in at least two respects. First, direct measures of a country's performance are methodologically separated from mere performance drivers, contributing to higher clarity concerning the concept to be measured. Second, weights of performance drivers are not determined arbitrarily or based on subjective judgment, but emerge from a transparent and well-established statistical procedure.

**Keywords** Canonical correlation, Competitiveness, Index construction

**Paper type** Research paper

## 1. Introduction

The number of composite indicators (CI) ranking nations according to various criteria has risen sharply in the past couple of years, intending to facilitate direct comparison between countries. Virtually every aspect of a country has been ranked from its standard of living to its economic competitiveness or educational attainments. Without a doubt, there are benefits to creating composite country indicators. Much like other indices, ease of comparison and attributing numerical values to abstract concepts can be extremely useful, especially in political communication. Clearly, this approach also suffers from shortcomings, but the analytical simplification of summarizing a set of intertwined indicators into a single one (or a few) is appealing.

Governments have started taking note of these indicators and some consider them relevant reflections of their economies. As countries can be ranked, their performance can be compared and competition between economies can thus be induced. This point is highlighted by the rise of countries like Georgia or Malaysia (9th and 12th, respectively,

Competitiveness Review
Vol. 26 No. 1, 2016
pp. 87-107

1059-5422
DOI 10.1108/CR-09-2014-0030

in 2012) through the World Bank's "Doing Business" rankings. Intriguingly, this is part of a larger strategy to attract investment into Georgia and Malaysia. Better placement in the table is thus an explicit policy goal and close cooperation with the World Bank led these countries to address particular sub-indicators of the "Doing Business" index (Høyland *et al.*, 2012). The effects of catering to indicators can be manifold. Georgia might genuinely be a very pleasant place for enterprises nowadays, but this may come at the cost of low health and safety standards. For better or worse, this documents that country indicators do matter in the real world.

Unfortunately, the methodology applied in the construction of existing CI is often not fully transparent or theoretically sound. This applies equally to the selection and weighting of chosen variables. As a consequence, the latent variable to be proxied often remains vague. Given the perceived policy relevance of these CI, such a lack of transparency and scientific foundations is unacceptable. The present paper addresses this shortcoming by exploring a statistically well-founded method for the construction of country indices, based on canonical correlation analysis (CCA), where selection and weighting of indices are performed in a transparent way. The method will be illustrated via the construction of an index reflecting national competitiveness in the world markets. It is composed of an innovative two-step procedure, which first uses principal component analysis (PCA) to extract factors from a broad set of indicators and then CCA to attribute weightings based on the correlation in the data.

The remainder of the paper will commence with Section 2 discussing the measurement of competitiveness. In Section 3, the application of CCA to the problem at hand is elucidated. Section 4 presents the estimation results, while Section 5 presents the resulting ranking and associated discussion. Section 6 presents a brief sensitivity analysis and Section 7 concludes the paper.

## 2. Measuring competitiveness

The notion of national competitiveness has drawn strongly on ideas from management theory. This legacy is still noticeable today and Porter's (1990) "diamond framework" is widely acknowledged as a seminal concept in competitiveness research. Therein, competitiveness is understood as productivity and a strong focus on costs is a direct consequence. With the advent of CI reflecting national economic competitiveness (i.e. Global Competitiveness Report), a conceptual debate intensified in which the sensibility of such measures was questioned and their construction deemed "misleading or even dangerous" (Krugman, 1994). In particular, the contradiction with basic insights from international trade and the resulting policy recommendations received criticism.

Further criticism from researchers (Lall, 2001) has induced some discipline into competitiveness CI generation and caused a development in the statistical and theoretical underpinnings. Some years on, Aiginger (2006) proposed that a conceptual consensus on competitiveness was emerging that included the "ability to create welfare" as part of its definition. This refers to outcome competitiveness and suggests that the most competitive countries are those, which generate the highest standard of living for their citizens. This outcome-based approach has also been established in CI (e.g. the Economic Complexity Index by Hausman *et al.*, 2013) and provides different insights to the traditional productivity-based measures. Clearly, outcome-based measures also suffer from shortcomings, as it is seldom clear that observed variation in market outcomes is related to the concept envisaged rather than alternative influences.

As Aiginger (2006) pointed out that "the attempt to define the term 'competitiveness of nations' has reached the phase of decreasing returns", this will not be the focus of this paper and the interested reader is referred to the existing literature (e.g. Aiginger *et al.*, 2013). For the intents and purposes of this paper, a working definition of competitiveness will be adapted that is deliberately focused. Competitiveness will be understood as a country's ability to benefit from the global exchange of goods and factors. This is more closely related to a measure of globalization than to productivity or welfare. In what follows, we use and elaborate on this alternative concept that is trying to avoid some of the issues related to pure input or pure outcome measures.

Aside from conceptual issues, there are also problems associated with the implementation of CI and thus the methodology of index construction deserves attention. CI can be aggregated in a number of different ways and the Organisation for Economic Co-operation and Development (OECD) "Handbook on Constructing Composite Indicators" (2008) provides an overview of the different approaches available. There is also extensive literature covering the shortcomings of existing CI (Lall, 2001). Criticisms include the unnecessary use of qualitative indicators, insufficient multivariate analysis or conflating outcome and explanatory variables.

While the methodology explored in this paper provides several advantages, its crucial contribution is arguably the endogenous determination of weights for the different elements of the CI. Recent advances in this area were primarily concerned with attributing different weights to different (groups of) countries (Global Competitiveness Report, Bowen and Moesen, 2011), as it was argued that the relevant factors for a country's competitiveness may vary depending on the phase of development. While this is sensible, it still relies on weights chosen ex-ante. The method presented here goes further by providing a data-led approach to determining appropriate weights. These are in turn applied to all countries, but application to country groups is conceivable given sufficient data.

Using predetermined (often equal) weights for different subcomponents is common practice for most competitiveness CI. This has practical advantages including ease of communication, stability and comparability across years. In turn, this also makes CI powerful, albeit one-sided, policy tools. From an analytical perspective, this can be seen as unsatisfactory, especially when inadequate multivariate analysis has been performed. For one, the importance of certain factors toward competitiveness is likely to vary across time. Additionally, CI like the Index of Economic Freedom (The Heritage Foundation, 2014) arguably over-represents certain dimensions by combining multiple highly correlated variables. Weighting then becomes secondary, as little information is to be gained by combining highly correlated variables. However, the very point of CI is to combine the multiple distinct dimensions of a concept and variables reflecting similar concepts should be bundled within a single sub-indicator whose appropriate weight in the CI needs to be determined.

For weighting in this paper, we chose to make use of a range of fundamental economic variables that can be expected to stand in close relationship with the envisioned performance measure. A method to achieve this already exists and has been widely applied within other fields of Social Sciences: CCA. We use this method to construct an index measuring a country's level of competitiveness in the world market, where weights of index variables are determined according to their linkages to a set of

development indicators. To the best of our knowledge, this method of index construction has not been applied to a macroeconomic ranking before.

## 3. Applying canonical correlation analysis

The CCA method dates back as far as Hotelling (1936) and is a widely used tool in behavioral and consumer research (Pappu *et al.*, 2007; Liu *et al.*, 2009). The basic CCA scenario consists of two latent variables, which are both proxied by linear combinations of multiple observable indicators. A priori the weights within both sets are unknown, but can be determined (up to an arbitrary factor) by maximizing the correlation between the two. Consequently, single indicators which exhibit a relatively strong correlation with the other set are assigned relatively large weights, indicating that they are primarily responsible for the shared variance between these sets.

When applying CCA to index construction, it is essential to clearly distinguish between *outcome variables* on the one side and *weighting variables* on the other. Outcome variables enter the index and thus need to reflect the latent variable(s) the index is intended to proxy. Our objective is to derive a multidimensional measure of a country's competitiveness on globally integrated markets. Competitiveness is seen here as a country's ability to benefit from the global exchange of goods and factors. One sign of high competitiveness would be a large inflow of foreign factors of production, as it signals that this country has attracted particular investment or labor and can subsequently benefit from these foreign factor endowments. Another sign would be a large outflow of own final products, as it indicates a high sales potential abroad for domestic producers. As outcome variables, the following indicators are therefore selected: the country's level of exports, the inflow of foreign capital and the level of immigration from abroad[1]. To achieve comparability among countries of different size, these variables are all expressed in per capita terms.

To capture the multidimensionality of development, the measures chosen as weighting variables span a wide range from health- and education-related measures to indicators of economic freedom and infrastructure quality. Measures of economic performance like GDP are deliberately excluded: based on our framework, they could neither be unambiguously classified as outcome variables nor as development indicators. The potentially large number of available weighting variables requires an objective method of selection in order not to over-represent particular dimensions. To maintain the self-determining nature of the data a rotated factor analysis is performed to distil distinct composites of indicators. This way it is ensured that all distinct dimensions of development are able to exert influence on the final index. Figure 1 summarizes our multi-step approach.

Country data for this paper have been retrieved from several international databases, including the World Bank Development Indicators, the IMF World Economic Outlook, the Heritage Foundation and the Penn World Tables. The data set spans the time period from 2006 to 2011, which both guarantees a focus on recent developments and a sufficiently large sample. A requirement for countries to be included is a population exceeding one million. As not all variables are updated each year for each country, the panel set comprises a range of missing values. Single missing values are replaced by simple averages of preceding and subsequent values. Countries with indicators for which less than two values are reported during that time span are not included in model

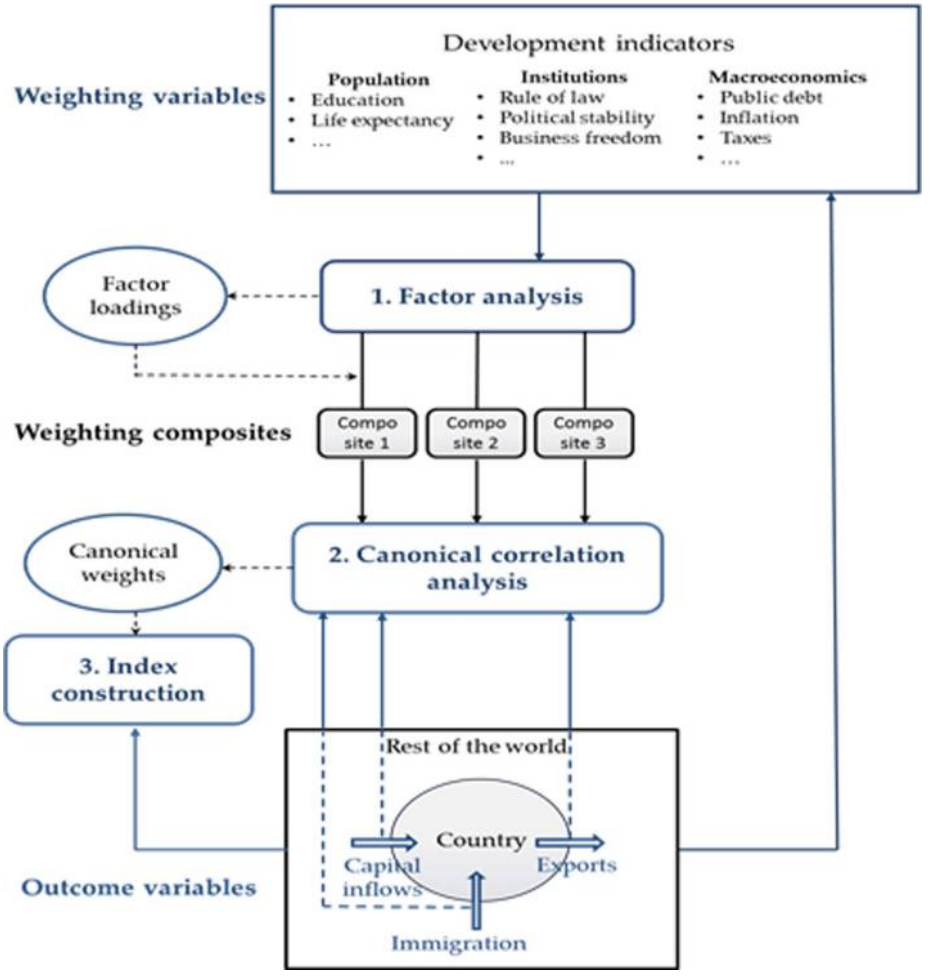



**Figure 1.**
Method of index construction

estimation. A closer description of the variables and their measurement is provided in Table AI.

The estimated factor loadings are used to group weighting variables into composites based on the approach of Nicoletti *et al.* (2000): each composite represents a weighted average of the variables with the highest loadings on a given factor. The weights are determined as the share of a variable's squared factor loading in the sum of squared loadings of the included variables[2]. This procedure requires choices on the number of factors to extract and on the number of variables to include in each composite. Following a common rule of thumb (Stevens, 1992), only variables whose factor loadings exceed an amount of 0.4 are considered to be sufficiently linked to the factor. To determine the number of factors considered, we apply the 90 per cent criterion in our baseline model, according to which enough factors to account for 90 per cent of total variation in the data are included. In a subsequent sensitivity analysis, this will be replaced by the Kaiser

criterion (Kaiser and Dickman, 1959), according to which only factors with eigenvalues above 1.0 are considered.

In the second step, the constructed composites enter a CCA. They appear as one linear combination (canonical variate) in the analysis, while the three outcome variables appear as another. As a result, canonical weights for each variable maximizing the correlation between the two sets are generated. In this way, a weighting scheme for outcome variables dependent on their correlation with composites of weighting variables is attained. In addition to this aggregate version, alternative model versions are estimated where two of the three dimensions of competitiveness are further split into their subcomponents.

In the last step, the canonical weights are applied to observed values of outcome variables for the individual countries. To omit the influence of short-run fluctuations and avoid data limitations, each value is chosen as the simple average of annual figures from the time period 2006 to 2011. The weighted sum of these values then serves as an index score for a country. Given that the set of weights is only unique up to an arbitrary factor, rescaling is allowed. To have some anchor point, the scaling here is chosen such that the country at the top of the ranking exhibits an index score of 100.

It is important to stress that this model formulation and the interpretation of its estimation results do not rest on the assumption of any kind of causal relationship between the two sets. In estimation, the two composites are completely interchangeable. The objective of the analysis is merely to provide a statistically meaningful basis for country rankings.

## 4. Estimation results

Table I documents the results of the PCA in the form of factor loadings and degrees of uniqueness for the single weighting variables. Factor loadings represent the correlation coefficient between the variables and a factor, while the degree of uniqueness measures the share of variance in a variable that is not accounted for by any of the considered factors. Factor loadings above the threshold of $|0.4|$ are highlighted.

Of the four factors reported, only the first three are worth considering according to the 90 per cent criterion; two, according to the Kaiser criterion. The first factor captures about 50 per cent of the total variation in the weighting variables. Of the 18 variables, 13 will be considered part of the first composite. This demonstrates that a large number of development variables are strongly correlated and appropriate multivariate analysis is paramount. The largest weights are on political and institutional indicators like rule of law, control of corruption and regulatory quality. Composite 2 includes eight variables with the largest weights on social and health indicators of life expectancy, number of physicians and tertiary education. The first two composites share several variables. In contrast, Composites 3 and 4 consist only of the real exchange rate (Rodrik, 2008) and tax rate, respectively. Their explained variance is quite limited.

Using these results, the constructed composites enter the CCA. Table II displays the estimation results in the form of canonical weights and canonical correlations obtained for the first pair of canonical variates for each of three different model versions. Model 1 is at the highest level of aggregation, while Model 3 disaggregates the outcome variables as far as possible given the available data.

Throughout all model versions, Composite 1 is assigned the dominant weight. Weights for Composites 2 and 3 are negative, but remain near zero. Hence, it is primarily

| Variable | Factor 1 | Factor 2 | Factor 3 | Factor 4 | Uniqueness |
|---|---|---|---|---|---|
| Adjusted real exchange rate | −0.559 | −0.268 | −0.580 | 0.052 | 0.267 |
| Business freedom | 0.627 | 0.499 | 0.008 | −0.040 | 0.247 |
| Control of corruption | 0.911 | 0.297 | 0.129 | −0.032 | 0.054 |
| Inflation | −0.394 | −0.063 | −0.304 | −0.078 | 0.668 |
| Internet access | 0.691 | 0.598 | 0.187 | −0.026 | 0.104 |
| Investment freedom | 0.696 | 0.192 | 0.055 | 0.015 | 0.231 |
| Labour freedom | 0.308 | 0.181 | −0.164 | −0.014 | 0.658 |
| Life expectancy | 0.459 | 0.779 | 0.017 | −0.060 | 0.135 |
| Number of physicians | 0.274 | 0.785 | 0.019 | −0.009 | 0.263 |
| Political stability | 0.677 | 0.239 | 0.093 | −0.020 | 0.326 |
| Public debt | 0.073 | 0.093 | 0.133 | 0.371 | 0.754 |
| Regulatory quality | 0.855 | 0.373 | 0.084 | −0.091 | 0.026 |
| Rule of law | 0.920 | 0.345 | 0.075 | −0.022 | 0.027 |
| Tax rate | −0.170 | −0.011 | −0.076 | 0.490 | 0.720 |
| Tertiary education | 0.435 | 0.794 | 0.066 | 0.020 | 0.162 |
| Trade freedom | 0.511 | 0.462 | 0.062 | −0.266 | 0.335 |
| Transport infrastructure | 0.724 | 0.478 | 0.140 | 0.163 | 0.141 |
| Urbanization | 0.414 | 0.650 | 0.364 | 0.114 | 0.227 |
| Eigenvalue | 6.247 | 3.881 | 0.715 | 0.513 | |
| Explained variance (%) | 0.520 | 0.323 | 0.060 | 0.043 | |
| Cumulative explained variance (%) | 0.520 | 0.843 | 0.902 | 0.945 | |
| Highlighted variables enter | Composite 1 | Composite 2 | Composite 3 | Composite 4 | |
| *N* | 629 | | | | |

**Source:** Own calculations

**Table I.**
Rotated factor loadings

indicators of political and institutional quality (composite 1) which are closely related to our concept of competitiveness and thus serve as main determinants of the weights assigned to outcome variables. Concerning these outcome variables, the most notable result for Model 1 is the negative sign of the weight for immigration and capital inflows. This indicates a negative relationship of immigrant and capital flows with the set of development factors when controlling for exports.

By splitting up exports and capital flows into their prime subcomponents, Model 2 shows that relevance varies with investment strategy and type of good traded. Finally, Model 3 further decomposes exports into their subcategories. It highlights manufactures and agricultural goods among the commodities and travel and transport among the services as sources of the general dominance of exports. In all three model versions, an interesting result is that factor flows prove to be much less relevant than exports. In some variants, their weights are even negative. This seems to contradict the notion of highly developed countries as main attractors of mobile capital and workers. A correlation table (Table AII) reveals that partial correlations with the dominant Composite 1 are indeed substantial and positive. However, in comparison, the correlation of Composite 1 with exports is much stronger.

For the purpose of index construction, the presence of negative weights creates some difficulties, as it complicates index interpretation. Applying results of Model 2 would mean that in a comparison of two countries with identical export activities, the one with the lower level of capital inflows would actually be ranked higher. This direct

| Variable | Model 1 | Model 2 | Model 3 |
|---|---|---|---|
| *Outcome variables* | | | |
| Capital inflows | −0.075 | | |
| FDI | | −0.048 | −0.085 |
| Portfolio | | −0.208 | 0.007 |
| Exports | 1.215 | | |
| Commodities | | 0.563 | |
| Food | | | 0.002 |
| Fuel | | | −0.072 |
| Manufactures | | | 0.264 |
| Metals | | | 0.262 |
| Raw agriculture | | | 0.301 |
| Commercial services | | 0.663 | |
| Communication | | | −0.022 |
| Finance | | | 0.118 |
| Transport | | | 0.172 |
| Travel | | | 0.318 |
| Immigration | −0.428 | −0.126 | 0.136 |
| *Weighting variables* | | | |
| Composite 1 | 1.029 | 1.001 | 1.042 |
| Composite 2 | −0.061 | −0.061 | −0.100 |
| Composite 3 | −0.031 | −0.058 | −0.051 |
| Correlation coefficient | 0.743 | 0.761 | 0.837 |
| *N* | 581 | 581 | 533 |

**Table II.** Standardized canonical weights

**Source:** Own calculations

application, however, would be misleading, given that the partial correlations of capital inflows with the (positively signed) weighting variables are clearly positive. As an alternative, a restricted version of CCA featuring a non-negativity constraint on the weights of outcome variables would therefore be desirable. Das and Sen (1994) have proposed an iterative procedure for determining canonical weights under this restriction, where outcome measures with insufficient roots in a country's development are automatically deleted from the index.

Table III reports the results of the restricted estimation for all three model versions. Dropping Immigration in Model 1 renders capital inflows positive. Weights of the undeleted outcome measures are marginally affected by the imposed constraint, while weights attached to weighting variables are also quantitatively similar. Finally, we have thus obtained a weighting scheme, which can be utilized for an illustrative country ranking. Furthermore, the need for appropriate weighting is highlighted, particularly in the more disaggregated model versions.

## 5. Ranking and discussion

By applying the estimated restricted weights to observed country values for 2011, index scores are obtained and a ranking emerges. Due to missing data, some countries are dropped from the ranking. Depending on the model specification, 120 to 139 countries are ranked. Figure 2 displays the global distribution of index scores for both versions.

| Variable | Model 1 | Model 2 | Model 3 |
|---|---|---|---|
| *Outcome variables* | | | |
| Capital inflows | 0.013 | | |
| FDI | | 0 | 0 |
| Portfolio | | 0 | 0.001 |
| Exports | 0.994 | | |
| Commodities | | 0.539 | |
| Food | | | 0 |
| Fuel | | | 0 |
| Manufactures | | | 0.230 |
| Metals | | | 0.240 |
| Raw agriculture | | | 0.306 |
| Commercial services | | 0.541 | |
| Communication | | | 0 |
| Finance | | | 0.089 |
| Transport | | | 0.117 |
| Travel | | | 0.349 |
| Immigration | 0 | 0 | 0.079 |
| *Weighting variables* | | | |
| Composite 1 | 1.008 | 1.014 | 1.024 |
| Composite 2 | −0.123 | −0.113 | −0.086 |
| Composite 3 | −0.122 | −0.107 | −0.059 |
| Correlation coefficient | 0.699 | 0.741 | 0.835 |
| *N* | 581 | 593 | 533 |

**Source:** Own calculations

**Table III.** Standardized canonical weights based on restricted estimation

As the ranking in both cases turns out to be quite similar, we base our discussion in the following on version 2 (Table AIV). The distribution of index scores is noticeable, as only a dozen countries achieve scores above 30, against the top country's score of 100. This occurs because the scores are a cardinal index resulting from the use of metric measures which themselves exhibit fairly skewed distributions involving several outliers. From a theoretical perspective, a more even distribution of index scores might be viewed as preferable, as this better reflects perceived differences in economic performance among the countries at the top end of the ranking. Section 6 addresses this point by analyzing the impact of a variable transformation.

On a global level, the ranking is clearly dominated by high-income countries. The Top 3 consist of Singapore, Ireland and Hong Kong. In general, top places are occupied by comparatively small countries. In contrast, large developed countries like the USA (Rank 41) and Japan (Rank 42) do not perform well, due to their mediocre performance in exports per capita. The large emerging economies Russia (Rank 56), Brazil (84), China (85) and India (108) are ranked relatively low. This is in line with our concept of competitiveness as a measure of participation in international goods and factor trade. Due to their size, small economies are less self-sufficient by nature and thus face higher pressure to integrate. Primarily African countries occupy the lower ranks.

In addition to its transparency, the ranking method explored also has another merit. The fact that the ranking as such is only influenced by the actual level of goods and

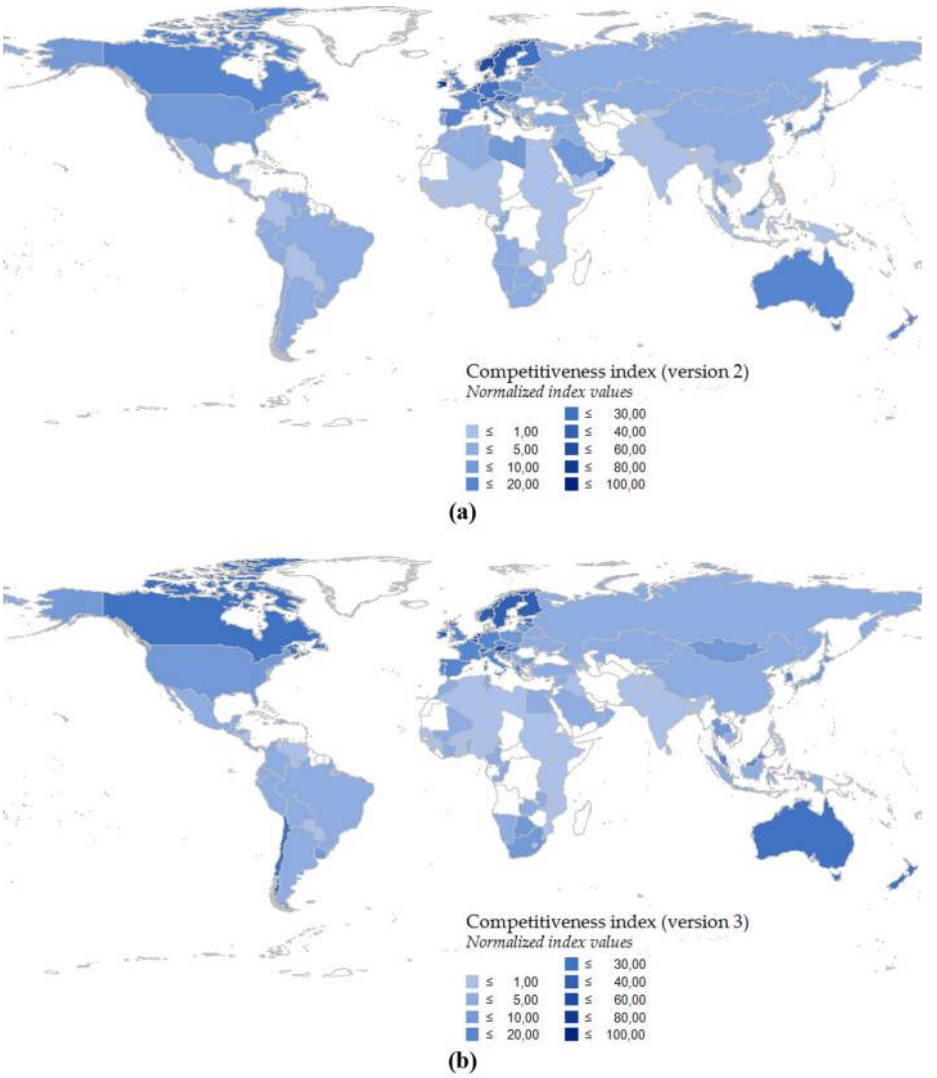



**Figure 2.**
Index scores of countries worldwide (model versions 2 and 3). Blanks indicate missing countries

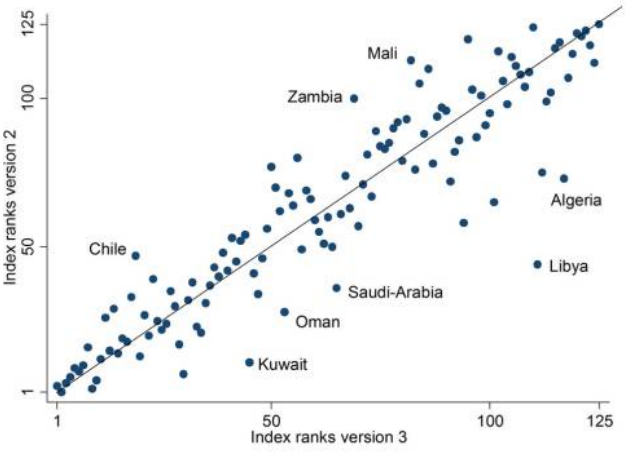
**Source:** Own calculations

factor flows reduces the scope for policy-based manipulation strategies. In indices measuring competitiveness based on institutional designs, manipulation incentives can represent a serious issue. By adjusting its laws to the specific requirements defined by the index, a country can artificially improve its position in the ranking while not necessarily improving its position in international competition. However, the method makes it difficult to derive clear policy recommendations for countries and should thus not be confused with a tool for policy analysis.

Another feature of our method, the data-based determination of indicator weights, has both advantages and disadvantages. It requires a regular updating of weights over time, thereby reduces the comparability of rankings across years. In contrast to other indices, however, the updating strategy is already specified by the general methodology and thus not subject to a priori judgments. One drawback related to the estimation outcome can be seen in the apparent dominance of exports. While this is an interesting result *per se*, it reduces the outward appeal of the index as a multidimensional measure. The inclusion of other dimensions of market outcomes could be of help here. In the end, this is primarily a question of data availability. The impact of a further differentiation within our data set is discussed in the next section.

## 6. Sensitivity analysis

While our explorative method of index construction is rooted in well-established statistical techniques, the eventual ranking is undoubtedly sensitive to the selection of indicators. This concerns both the choice of outcome and of weighting variables. Comparing estimation results for model versions 2 and 3, the weighting of exports compared to factor flows is broadly similar and the overall distribution of index scores worldwide is as well. Nevertheless, as version 3 introduces additional refinement in the weighting of export subcategories, a few economies with highly specialized export patterns do experience a serious shift in their position. In line with our aim of emphasizing development-related performance, these are primarily oil-exporting countries from the Arab region. Since weights of fuel exports are driven down to zero in our constraint-based estimation, these countries are downgraded by several ranks compared to model version 2 (Figure 3). Interestingly, the same does not apply to resource exports in the form of metals and ores. For this reason, economies with significant trade in precious metals like Chile and Mali are able to rise significantly in the ranking. This demonstrates that correlation patterns between trade performance and the distinct dimensions of development are complex enough to deserve a disaggregated analysis.




**Source:** Own calculations

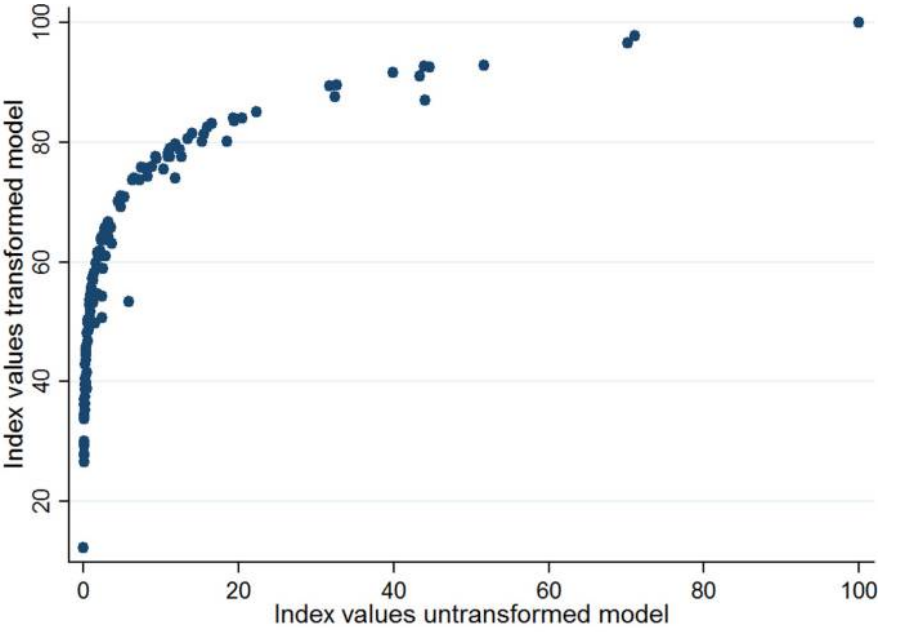
**Figure 3.**
Comparison of model versions 2 and 3

Regarding the choice of weighting composites, rules other than the 90 per cent criterion could be implemented. For instance, the Kaiser criterion would only include the first two composites. Corresponding estimation results are documented in Table AIII, where weights of outcome variables are hardly affected. Moreover, a transformation of outcome variables before conducting the CCA could be considered appropriate. As documented in Table AI, the distributions of the single outcome measures exhibit a considerable degree of skewness. This results into index scores that themselves are highly asymmetric. Transforming outcome variables to reduce skewness is one way to prevent this. It can also help to reduce the sensitivity of index scores towards outliers. On the downside, any transformation renders the interpretation of index scores less straightforward and intuitive.

Due to the presence of negative values, a simple log transformation is not an option. Instead, we apply the related "neglog" transformation explained in Whittaker *et al.* (2005) to our outcome variables. Table AV displays the estimation results obtained by using the transformed variables for the CCA. Establishing a ranking based on these transformations has the expected effect that variation in index values shrinks, as is evident by the comparison of country scores with and without the transformation procedure in Figure 4. Nevertheless, the relative ranking of countries is only affected to a minor degree.

## 7. Conclusion

Our achievement in this paper was to develop and implement an explorative approach of ranking countries according to their level of competitiveness. Competitiveness was conceptualized as the degree of participation in global integration, measured by the levels of




**Source:** Own calculations

**Figure 4.**
Distributions of country index scores with and without transformation (model version 2)

exports, capital inflows and immigration. Based on a wide range of development indicators, weighting composites were constructed by means of a PCA. These composites entered a CCA as one set of variables and the index variables as another. Estimated canonical weights were used as weights in the calculation of index scores. In this way, an index was constructed where weights are assigned based on an element's roots in the stage of development. In comparison to prevailing macroeconomic indices, our method of index construction is superior in at least two respects. First, direct measures of a country's performance are methodologically separated from mere performance drivers, contributing to higher clarity concerning the concept measured. Second, weights of performance drivers are not determined arbitrarily or based on subjective judgment, but emerge from a transparent and well-established statistical procedure.

The application of our procedure to recent worldwide country data has provided interesting insights. Among the development indicators measures of institutional and political quality attained the highest weight in the dominant composite. Concerning the weighting of outcome variables, the imposition of a non-negativity constraint implies that capital flows and immigration are dropped from the index in Model 2. Within the data at hand, inflows of foreign capital have turned out to correlate much less with our weighting composites than export volumes. However, a further disaggregation has revealed that using aggregate exports alone would be misguided. In fact some product categories are much closer linked to development indicators than others and should receive larger weights, while other groups like fuel are dropped entirely. This strong asymmetry in index weights can be seen as a justification of our estimation-based approach, as any a priori weighting would most probably misfit the underlying correlation structure.

As this paper provides a first exploration of applying CCA to competitiveness, several points could be addressed in future research. For one, the necessity and effect of transforming outcome variables might warrant further investigation. Furthermore, the choice of per capita outcome variables arguably leads to a small country bias. Lastly, the procedure presented here has proven relatively data-intensive and this might prove an obstacle for further applications. Nonetheless, the CCA based approach has offered some fresh insights into national competitiveness and addressed several of the ailments of existing indices. In the future, the method could similarly be applied to a number of topics seeking quantification, such as a country's quality of schooling, the quality of its health care system or of its governmental institutions.

## Notes

1. Exports are measured in terms of annual volumes. Capital inflows are computed as the sum of annual foreign direct investments (FDI) and portfolio investments. Immigration is measured as the annual change in the stock of foreign-born citizens, i.e. the net inflow of immigrants.
2. In case of negative loadings, squared loadings enter with a negative sign.

## References

Aiginger, K. (2006), "Competitiveness: from a dangerous obsession to a welfare creating ability with positive externalities", *Journal of Industry, Competition and Trade*, Vol. 6 No. 2, pp. 161-177.

Aiginger, K., Bärenthaler-Sieber, S. and Vogel, J. (2013), "Competitiveness under new perspectives", WWWForEurope Working Paper No. 44, WIFO, Vienna.

Bowen, H. and Moesen, W. (2011), "Composite competitiveness indicators with endogenous versus predetermined weights: an application the World Economic Forum's global competitiveness index", *Competitiveness Review*, Vol. 21 No. 2, pp. 129-151.

Das, S. and Sen, P. (1994), "Restricted canonical correlation", *Linear Algebra and its Applications*, Vol. 210, October, pp. 29-47.

Hausman, R., Hidalgo, C.R., Bustos, S., Coscia, M., Simoes, A. and Yildirim, M.A. (2013), *The Atlas of Economic Complexity-Mapping Paths Towards Prosperity*, MIT Press, Cambridge, MA.

Hotelling, H. (1936), "Relations between two sets of variates", *Biometrika*, Vol. 28 Nos 3/4, pp. 321-377.

Høyland, B., Moene, K. and Willumsen, F. (2012), "The tyranny of international index rankings", *Journal of Development Economics*, Vol. 97 No. 1, pp. 1-15.

Kaiser, H.F. and Dickman, K. (1959), "Analytic determination of common factors", *American Psychologist*, Vol. 14 No. 1, pp. 425-441.

Krugman, P. (1994), "Competitiveness: a dangerous obsessions", *Foreign Affairs*, Vol. 73 No. 2, pp. 28-44.

Lall, S. (2001), "Competitiveness indices and developing countries: an economic evaluation of the global competitiveness report", *World Development*, Vol. 29 No. 9, pp. 1501-1525.

Liu, Z., Zeng, F. and Su, C. (2009), "Does relationship quality matter in consumer ethical decision-making?", *Evidence from China, Journal of Business Ethics*, Vol. 88 No. 3, pp. 483-496.

Nicoletti, G., Scarpetta, S. and Boylaud, O. (2000), "Summary indicators of product market regulation with an extension to employment protection legislation", Economics Department Working Papers No. 226, OECD, Paris.

Pappu, R., Quester, P.G. and Cooksey, R.W. (2007), "Country image and consumer-based brand equity: relationships and implications for international marketing", *Journal of International Business Studies*, Vol. 38 No. 5, pp. 726-745.

Porter, M.E. (1990), "The competitive advantage of nations", *Harvard Business Review*, March-April 1990.

Rodrik, D. (2008), "The real exchange rate and economic growth", Brookings Papers on Economic Activity, The Brookings Institution, Washington, DC, Vol. 39, pp. 365-439.

Stevens, J.P. (1992), *Applied Multivariate Statistics for the Social Sciences*, 2nd ed., Erlbaum, Hillsdale, NJ.

The Heritage Foundation (2014), *2014 Index of Economic Freedom*, Heritage Foundation, Washington, DC.

Whittaker, J., Whitehead, C. and Somers, M. (2005), "The neglog transformation and quantile regression for the analysis of a large credit scoring database", *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, Vol. 54 No. 5, pp. 863-878.

## Further reading

IMD (2014), *World Competitiveness Yearbook 2014*, International Institute for Management Development, Lausanne.

World Bank (2014), "Doing business 2014 – understanding regulation for small and medium-size enterprises", World Bank Group, Washington, DC.

World Economic Forum (2014), "Global competitiveness report 2014-2015: full data edition", World Economic Forum, Geneva.

## Appendix

**Table AI.**
Data sources

| Variable | Description | Measurement | Source | Mean | SD | Skewness |
|---|---|---|---|---|---|---|
| *Outcome variables* | | | | | | |
| FDI | Annual inflows of foreign direct investment per capita | US dollars PPP | UNCTAD | 608.07 | 1,805.17 | 5.79 |
| Portfolio | Annual inflows of portfolio investment per capita | US dollars PPP | IMF | 268.96 | 1,959.49 | 8.49 |
| Exports (including subcategories) | Annual export volumes of commodities and services per capita | US dollars PPP | WTO | 8,962.06 | 16,093.97 | 3.29 |
| Immigration | Average annual net increase in the number of reported foreign-born citizens between 2006 and 2011 relative to total population size | Number of people | World Bank | 0.002 | 0.007 | 6.726 |
| *Weighting variables* | | | | | | |
| Adjusted real exchange rate | The average real exchange rate of the national currency to US-Dollar in a year corrected for differences in real GDP per capita (Balassa-Samuelson effect; methodology adopted from Rodrik (2008)) | Dimensionless | Penn World Tables; own estimations | 0.000 | 0.306 | −0.517 |
| Business freedom | Index measuring the ability to start, operate, and close a business | Interval scale between 0 (worst) and 100 (best) | Heritage Foundation | 63.414 | 16.219 | −0.212 |
| Control of corruption | Index measuring perceptions of the extent to which public power is exercised for private gain | Interval scale between −2.5 (worst) and +2.5 (best) | World Bank | −0.001 | 1.005 | 0.618 |
| Inflation | Annual per cent change in the level of consumer prices | Percentage | Penn World Tables | 6.370 | 5.970 | 2.354 |
| Internet access | Number of citizens per 1000 people with regular access to the internet | Number of people | World Bank | 22.762 | 24.848 | 1.074 |

(*continued*)

Table AI.

| Variable | Description | Measurement | Source | Mean | SD | Skewness |
|---|---|---|---|---|---|---|
| Investment freedom | Index measuring the absence of constraints on the flow of investment capital | Interval scale between 0 (worst) and 100 (best) | Heritage Foundation | 51.279 | 20.36 | −0.090 |
| Labor freedom | Index measuring the quality of the legal and regulatory framework of a country's labor market | Interval scale between 0 (worst) and 100 (best) | Heritage Foundation | 61.080 | 17.071 | −0.263 |
| Life expectancy | Average life expectancy of citizens in years | Number of years | World Bank | 68.327 | 10.293 | −0.830 |
| Number of physicians | Number of physicians per 1000 people | Number of people | World Bank | 1.997 | 1.542 | 0.337 |
| Political stability | Index measuring the likelihood that the government will be destabilized or overthrown by unconstitutional means | Interval scale between −2.5 (worst) and +2.5 (best) | World Bank | 0.003 | 0.994 | −0.655 |
| Public debt | Ratio of government debt to GDP | Percentage | IMF | 50.853 | 47.402 | 6.501 |
| Taxes | Ratio of the sum of total taxes and contributions payable by businesses to commercial profits | Percentage | World Bank | 49.829 | 41.298 | 4.278 |
| Tertiary education | Ratio of total enrollment in tertiary schooling to the size of the relevant age group | Percentage | World Bank | 37.028 | 26.566 | 0.410 |
| Trade freedom | Index measuring the absence of tariff and non-tariff barriers | Interval scale between 0 (worst) and 100 (best) | Heritage Foundation | 72.480 | 13.947 | −1.679 |
| Transport infrastructure | Logistics performance index measuring the quality of transport infrastructure | Interval scale between 1 (worst) and 5 (best) | World Bank | 2.603 | 0.722 | 0.795 |
| Urbanization | Ratio of number of persons living in urban areas to total population size | Percentage | World Bank | 56.952 | 24.289 | −0.172 |

| Variable | Composite 1 | Composite 2 | Composite 3 | Capital inflows | Exports | Immigration |
|---|---|---|---|---|---|---|
| Composite 1 | 1 | | | | | |
| Composite 2 | 0.8742 | 1 | | | | |
| Composite 3 | −0.7752 | −0.6485 | 1 | | | |
| Capital inflows | 0.3353 | 0.2575 | −0.2644 | 1 | | |
| Exports | 0.6952 | 0.5852 | −0.5758 | 0.4744 | 1 | |
| Immigration | 0.1814 | 0.1272 | −0.2200 | 0.1120 | 0.5647 | 1 |

**Source:** Own calculations

**Table AII.** Correlation coefficients for model 1

| Variable | Model 1 | Model 2 | Model 3 |
|---|---|---|---|
| *Outcome variables* | | | |
| Capital inflows | 0.016 | | |
| FDI | | 0 | 0 |
| Portfolio | | 0 | 0.000 |
| Exports | 0.992 | | |
| Commodities | | 0.536 | |
| Food | | | 0 |
| Fuel | | | 0 |
| Manufactures | | | 0.231 |
| Metals | | | 0.242 |
| Raw agriculture | | | 0.313 |
| Commercial services | | 0.543 | |
| Communication | | | 0 |
| Finance | | | 0.092 |
| Transport | | | 0.115 |
| Travel | | | 0.343 |
| Immigration | 0 | 0 | 0.075 |
| *Weighting variables* | | | |
| Composite 1 | 1.120 | 1.110 | 1.077 |
| Composite 2 | −0.140 | −0.127 | −0.090 |
| Correlation coefficient | 0.697 | 0.740 | 0.835 |
| *N* | 581 | 593 | 533 |

**Source:** Own calculations

**Table AIII.** Standardized canonical weights based on Kaiser criterion

| Rank | Country | Score |
|---|---|---|
| 1 | Singapore | 100 |
| 2 | Ireland | 71.1 |
| 3 | Hong Kong SAR, China | 70.23 |
| 4 | Belgium | 51.68 |
| 5 | Norway | 44.59 |
| 6 | Qatar | 44.1 |
| 7 | Denmark | 43.99 |
| 8 | Netherlands | 43.35 |
| 9 | Switzerland | 39.96 |
| 10 | Austria | 32.71 |
| 11 | Kuwait | 32.46 |
| 12 | Sweden | 31.74 |
| 13 | Finland | 22.36 |
| 14 | Germany | 20.51 |
| 15 | Slovenia | 19.48 |
| 16 | Bahrain | 19.28 |
| 17 | Cyprus | 18.52 |
| 18 | United Kingdom | 16.61 |
| 19 | Estonia | 16.04 |
| 20 | Canada | 15.54 |
| 21 | Czech Republic | 15.3 |
| 22 | Israel | 14.01 |
| 23 | France | 13.5 |
| 24 | Slovak Republic | 12.69 |
| 25 | Hungary | 12.47 |
| 26 | Australia | 11.98 |
| 27 | Spain | 11.81 |
| 28 | Oman | 11.8 |
| 29 | Republic of Korea | 11.15 |
| 30 | New Zealand | 11.11 |
| 31 | Italy | 10.95 |
| 32 | Greece | 10.94 |
| 33 | Lebanon | 10.34 |
| 34 | Trinidad and Tobago | 10.02 |
| 35 | Croatia | 9.42 |
| 36 | Saudi Arabia | 9.38 |
| 37 | Portugal | 9.35 |
| 38 | Lithuania | 8.82 |
| 39 | Malaysia | 8.32 |
| 40 | Latvia | 7.96 |
| 41 | United States | 7.43 |
| 42 | Japan | 7.19 |
| 43 | Panama | 6.59 |
| 44 | Mauritius | 6.3 |
| 45 | Libya | 5.86 |
| 46 | Poland | 5.33 |
| 47 | Costa Rica | 4.81 |

**Table AIV.**
Ranking (Model 2)

(*continued*)

| Rank | Country | Score |
|---|---|---|
| 48 | Chile | 4.77 |
| 49 | Bulgaria | 4.51 |
| 50 | Kazakhstan | 3.63 |
| 51 | Belarus | 3.51 |
| 52 | Thailand | 3.3 |
| 53 | Romania | 3.27 |
| 54 | Russian Federation | 3.17 |
| 55 | Uruguay | 3.14 |
| 56 | Jamaica | 3.04 |
| 57 | Botswana | 2.9 |
| 58 | Azerbaijan | 2.84 |
| 59 | Macedonia, FYR | 2.73 |
| 60 | Jordan | 2.73 |
| 61 | Mexico | 2.49 |
| 62 | Turkey | 2.47 |
| 63 | Tunisia | 2.45 |
| 64 | Venezuela, RB | 2.38 |
| 65 | Angola | 2.36 |
| 66 | Serbia | 2.33 |
| 67 | Bosnia and Herzegovina | 2.31 |
| 68 | Albania | 2.22 |
| 69 | Swaziland | 2.18 |
| 70 | Argentina | 2.05 |
| 71 | Namibia | 2.03 |
| 72 | South Africa | 1.94 |
| 73 | Dominican Republic | 1.85 |
| 74 | Ukraine | 1.84 |
| 75 | Congo, Republic | 1.84 |
| 76 | Algeria | 1.72 |
| 77 | Morocco | 1.56 |
| 78 | Iraq | 1.51 |
| 79 | Honduras | 1.32 |
| 80 | Mongolia | 1.28 |
| 81 | Ecuador | 1.2 |
| 82 | Peru | 1.2 |
| 83 | El Salvador | 1.19 |
| 84 | China | 1.17 |
| 85 | Brazil | 1.16 |
| 86 | Syrian Arab Republic | 1.05 |
| 87 | Georgia | 1 |
| 88 | Egypt, Arab Republic | 0.91 |
| 89 | Paraguay | 0.9 |
| 90 | Moldova | 0.9 |
| 91 | Colombia | 0.9 |
| 92 | Armenia | 0.82 |
| 93 | Vietnam | 0.81 |
| 94 | Guatemala | 0.8 |

(*continued*)

**Table AIV.**

| Rank | Country | Score |
|---|---|---|
| 95 | Philippines | 0.79 |
| 96 | Papua New Guinea | 0.77 |
| 97 | Bolivia | 0.68 |
| 98 | Indonesia | 0.65 |
| 99 | Sri Lanka | 0.6 |
| 100 | Kyrgyz Republic | 0.6 |
| 101 | Cambodia | 0.57 |
| 102 | Cote d'Ivoire | 0.56 |
| 103 | Nicaragua | 0.47 |
| 104 | Nigeria | 0.46 |
| 105 | Zambia | 0.43 |
| 106 | Ghana | 0.39 |
| 107 | Yemen, Rep | 0.39 |
| 108 | Lesotho | 0.38 |
| 109 | India | 0.35 |
| 110 | Senegal | 0.35 |
| 111 | Cameroon | 0.34 |
| 112 | Lao PDR | 0.34 |
| 113 | Kenya | 0.27 |
| 114 | Sudan | 0.26 |
| 115 | Tajikistan | 0.24 |
| 116 | Togo | 0.22 |
| 117 | Benin | 0.2 |
| 118 | Gambia, The | 0.2 |
| 119 | Tanzania | 0.17 |
| 120 | Mali | 0.17 |
| 121 | Afghanistan | 0.16 |
| 122 | Liberia | 0.16 |
| 123 | Mozambique | 0.15 |
| 124 | Myanmar | 0.15 |
| 125 | Pakistan | 0.14 |
| 126 | Guinea | 0.13 |
| 127 | Guinea-Bissau | 0.13 |
| 128 | Bangladesh | 0.11 |
| 129 | Haiti | 0.11 |
| 130 | Uganda | 0.1 |
| 131 | Burkina Faso | 0.07 |
| 132 | Rwanda | 0.07 |
| 133 | Sierra Leone | 0.07 |
| 134 | Timor-Leste | 0.07 |
| 135 | Malawi | 0.07 |
| 136 | Nepal | 0.07 |
| 137 | Niger | 0.07 |
| 138 | Ethiopia | 0.07 |
| 139 | Burundi | 0.01 |

**Table AIV.**

| Variable | Model 1 | Model 2 | Model 3 |
|---|---|---|---|
| *Outcome variables* | | | |
| Capital inflows | 0.069 | | |
| FDI | | 0 | 0 |
| Portfolio | | 0.021 | 0.017 |
| Exports | 0.974 | | |
| Commodities | | 0.387 | |
| Food | | | 0.088 |
| Fuel | | | 0.098 |
| Manufactures | | | 0.133 |
| Metals | | | 0.057 |
| Raw agriculture | | | 0.116 |
| Commercial services | | 0.642 | |
| Communication | | | 0.100 |
| Finance | | | 0.216 |
| Transport | | | 0.270 |
| Travel | | | 0.082 |
| Immigration | 0 | | 0.030 |
| *Weighting variables* | | | |
| Composite 1 | 0.137 | 0.244 | 0.284 |
| Composite 2 | 0.707 | 0.665 | 0.626 |
| Composite 3 | −0.229 | −0.154 | −0.149 |
| Correlation coefficient | 0.925 | 0.942 | 0.942 |
| *N* | 581 | 581 | 533 |

**Table AV.**
Standardized canonical weights when applying neglog transformations to outcome variables

**Corresponding author**
Lars Wenzel can be contacted at: wenzel@hwwi.org