December 2012

151 Slater Street, Suite 710
Ottawa, Ontario K1P 5H3
613-233-8891, Fax 613-233-8250
csls@csls.ca

CENTRE FOR THE STUDY OF LIVING STANDARDS

AN ASSESSMENT OF WEIGHTING METHODOLOGIES FOR COMPOSITE INDICATORS: THE CASE OF THE INDEX OF ECONOMIC WELL-BEING

Final Draft

Andrew Sharpe and Brendon Andrews

CSLS Research Report No. 2012-10
December 2012

An Assessment of Weighting Methodologies for Composite Indicators: The Case of the Index of Economic Well-being

ABSTRACT

The Index of Economic Well-Being (IEWB) – a composite indicator consisting of consumption, wealth, equality, and economic security – underwent several changes in the weighting of its components. For example, the final aggregation of the IEWB was changed to equal weighting after the IEWB was criticized for having a bias against sustainability; however, all weighting schemes have both advantages and shortcomings. To isolate the preferred ordinal ranking for the results, strong and weak dominance rules were established for countries across several observed weighting schemes, and each of these rules were ranked in all possible ways. An 'iterative dominance equilibrium' was computed for comparison to observed weighting schemes. Constrained data envelopment analysis (CDEA) performed best, yet CDEA is not ideal for comparisons across countries. Among explicit weights, the original weights of the IEWB were best. Although the original weights are supported, they were controversial – a shift to equal weights mitigated this controversial – it appears equal weights remain least objectionable.

An Assessment of Weighting Methodologies for Composite Indicators: The Case of the Index of Economic Well-being

Table of Contents

An Assessment of Weighting Methodologies for Composite Indicators: The Case of the Index of Economic Well-being

EXECUTIVE SUMMARY

The Index of Economic Well-being (IEWB), developed by Lars Osberg of Dalhousie University and Andrew Sharpe of the Centre for the Study of Living Standards (CSLS), offers an alternative measure of well-being along economic and social dimensions. The measure is an aggregate of four components: consumption, wealth, equality, and economic security. Two versions of the IEWB are produced by the CSLS: one for Canada and the provinces and one for fourteen selected OECD countries. This report focuses on the results from the IEWB for OECD countries. In 2009, Norway placed first with an IEWB value of 0.799 and Spain, with an IEWB value of 0.451, was the worst performing nation. These results are based on equal weights of 0.25 placed on each component of the IEWB. These weights are controversial – changing the weights on the components changes the outcome of the IEWB.

The ideal weighting for the IEWB will be transparent, will objectively capture societal valuations, and will produce comparable index values across sections and time. The current weighting of the IEWB does not satisfy the second condition. To determine whether another weighting methodology satisfies all three conditions, we analyze several explicit weights (including equal weighting), factor analysis, data envelopment analysis (DEA), constrained DEA, common weights DEA, and compromise solution DEA. The rankings obtained from each weighting method were used to determine the optimal ordinal ranking. Comparison of robustness with this ranking indicated that constrained DEA and the original weights used in the aggregation of the IEWB produced the most preferable results.

The history and purpose of the IEWB indicates that neither of these weighting methodologies works in practice. Constrained DEA allows for variation of weights across countries. Although an interesting analysis, this is not optimal for cross-sectional comparison. The original weights of the IEWB were replaced by equal weights in order to mitigate disagreement over the weighting of the IEWB. Although there is evidence to support the original weights, history indicates they also fail in practice. Equal weighting remains the best option for weighting the IEWB as it allows simple comparisons and it mitigates disagreement. The other weighting methodologies are good analyses, but do not possess optimal qualities for weighting the IEWB in practice.

List of Figures

List of Tables

List of Appendix Tables

An Assessment of Weighting Methodologies for Composite Indicators: The Case of the Index of Economic Well-being

INTRODUCTION

The Index of Economic Well-Being (IEWB) has considered many options for the aggregation of its components, learning valuable lessons about the weighting of composite indicators. The changes made by the IEWB are in line with work discussing the societal differences in valuations of indices, but the literature on the weighting of composite indicators is continually expanding and so a new method might perform better. The authors explore a variety of weighting schemes for the IEWB under the guidance of the lessons already learned, thereby allowing for a discussion of the suitability of these weighting techniques in practice.

The analysis finds that equal weighting is both far from optimal and the most convenient method to use. Although alternative methods of weighting such as constrained data envelopment analysis and the original weighting scheme of the IEWB prove more robust than equal weighting, the requirements and experience of the IEWB indicate that neither of these methods, in practice, is more desirable than equal weighting. These alternatives appear to be credible sensitivity analyses but are not suited for the role of weighting the baseline index. Given the lack of a fitting alternative, the IEWB must continue to use equal weights.

The first section of the paper offers a detailed description of the IEWB, a summary of some key estimates from the most recent update of the IEWB (Osberg and Sharpe, 2011), and outlines the ideal methodology for the future of the weighting of the IEWB. The paper then analyzes various weighting methods subject to the data, structure, and history of the IEWB. The second section focuses on explicit weighting: equal weighting, expert weighting, user weighting, and survey weighting. The third section focuses on statistical methods of weighting: factor analysis, regression analysis, along with an analysis of the methodology used by the Composite Learning Index (CLI). The fourth section explores various forms of data envelopment analysis. The fifth section is based on the consumption-equivalent measure. The sixth section summarizes our findings on the optimality of weighting methods and the seventh section offers dominance rankings in order to determine the most robust weighting method. The final section concludes, offering best practices based on the lessons learned from the IEWB and the analyses performed.

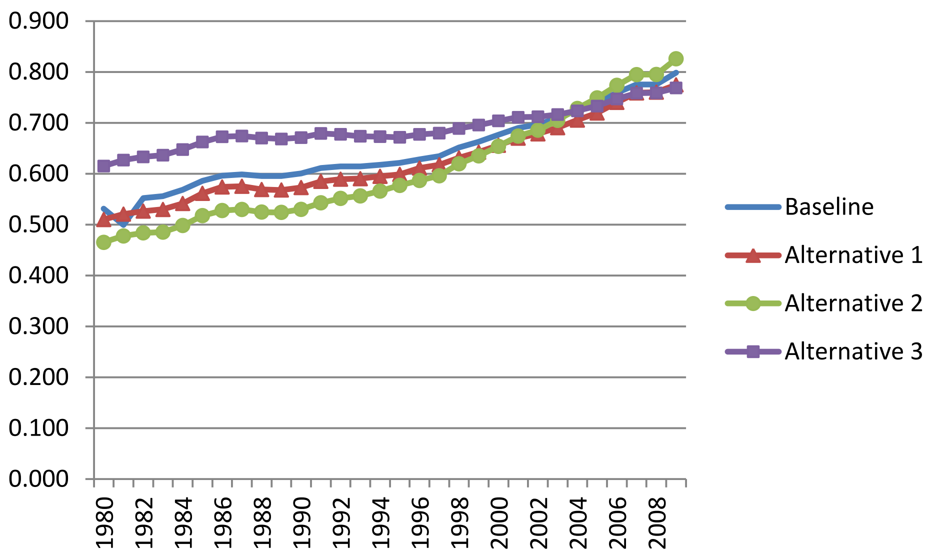
I. THE INDEX OF ECONOMIC WELL-BEING

The IEWB for selected OECD countries[1] measures economic well-being for 14 OECD nations: Australia, Belgium, Canada, Denmark, Finland, France, Germany, Italy, the Netherlands, Norway, Spain, Sweden, the United Kingdom, and the United States. Four components of well-being are measured: consumption flows, stocks of wealth, economic equality, and economic security. Each component of the IEWB is composed of numerous variables. The consumption component consists of: personal consumption per capita (adjusted by

[1] The IEWB was developed by Lars Osberg of Dalhousie University and Andrew Sharpe of the Centre for the Study of Living Standards. See Osberg and Sharpe (2002, 2005, 2009a, 2011) for the IEWB for OECD countries and Osberg and Sharpe (2009b) for the IEWB for Canada and the provinces.

family size), adjusted relative cost of leisure per capita, and government final consumption expenditures per capita, all adjusted for life expectancy. The wealth component consists of: total net stock of gross fixed capital per capita, stock of gross expenditures on R&D per capita, total net international investment position per capita, human capital stock per capita, and greenhouse gas emission cost per capita. The equality component consists of overall poverty intensity and the Gini coefficient, where poverty intensity is defined as the relative poverty rate multiplied by the poverty gap. Finally, the economic security component consists of the risk of poverty in old age, the risk of single parent poverty, the risk imposed by illness, and the risk imposed by unemployment. Many of these indicators are themselves aggregations of variables. For a complete discussion of the variables and the performance of the fourteen selected countries in each variable, see Osberg and Sharpe (2009a, 2011).[2]

Figure 1: The IEWB for Norway, 1980-2009



Baseline Weights: Consumption (0.25), Wealth (0.25), Equality (0.25), Security (0.25)
Alternative 1 Weights: Consumption (0.4), Wealth (0.1), Equality (0.25), Security (0.25)
Alternative 2 Weights: Consumption (0.33), Wealth (0.33), Equality (0), Security (0.33)
Alternative 3 Weights: Consumption (0.2), Wealth (0.1), Equality (0.4), Security (0.3)

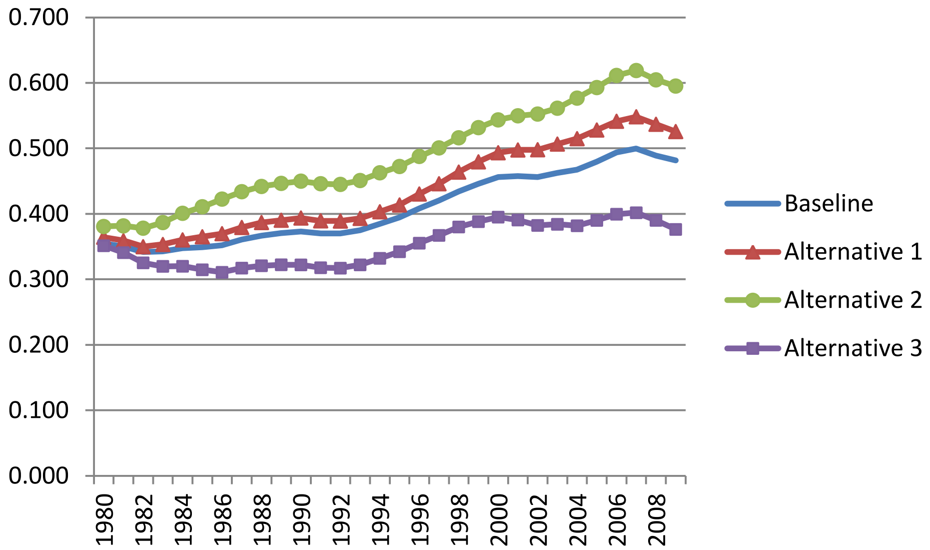
In the most recent update of the IEWB for selected OECD Countries (Osberg and Sharpe, 2011), the top performing nation in 2009 was Norway with an index value of 0.799 and the worst performing nation was Spain with an index value of 0.451. The index value of the United States was 0.482, placing the United States in thirteenth (or second last) place. The growth of the IEWB was positive for all countries from 1980 to 2009.

In Figure 1 above and Figure 2 below, the values of the IEWB for Norway and the United States from 1980 to 2009 are presented for four different weighting schemes. These charts are taken from Osberg and Sharpe (2011). Note that the United States consistently ranks below Norway under all four offered sets of weights. Under the baseline, each component of the IEWB is assigned equal weight (0.25). Under alternative 1, consumption is weighted at 0.4, wealth at

[2] Osberg and Sharpe (2009a) is available at: http://www.csls.ca/iwb/oecd.asp

0.1, equality at 0.25, and security at 0.25. Under alternative 2, consumption, wealth, and security are weighted at 0.33, while equality is assigned a weight of 0. Finally, under alternative 3, consumption is weighted at 0.2, wealth at 0.1, equality at 0.4, and security at 0.3. Norway and the United States will act as reference nations at various occasions in the paper.

Figure 2: The IEWB for the United States, 1980-2009



The method of aggregation of indicators and variables in the IEWB varies. In the consumption domain, the variables that compose the index of consumption are aggregated by a simple sum of prices with and adjusted by relative life expectancy. In the index of wealth domain, the variables that compose the index are also simply summed to obtain the value of stocks of wealth in the nation. Both of these methods are simple and justifiable - because variables are measured in prices, stocks, or expenditures, these can simply be aggregated to find the total measured net stocks or expenditures in dollars. The index of equality explicitly assigns a weight of 0.75 to poverty intensity and a weight of 0.25 to the Gini coefficient. The index of security is more complex. The individual risks associated with this index are aggregated using the percentage of the population who are considered at risk of each plight.[3] For the risk of unemployment, the scaled value of unemployment is given an explicit weight of eighty per cent and the scaled value of the replacement rate is assigned an explicit weight of twenty per cent. The risk of illness is simply a scaled value of medical expenses as a percentage of personal disposable income. The risk of single parent poverty is based on the product of poverty intensity for single women with children and the divorce rate per thousand people. The risk of poverty in old age is the scaled value of poverty intensity for the elderly. Therefore, a variety of weighting techniques are used at multiple levels after much scaling; however, the majority of weights assigned in the IEWB appear to be explicit weights.

Many papers (for example, Salzman and Sharpe (2003) and Maggino and Zumbo (2011)) consider different methodological choices in the construction of composite indices. In this paper,

[3] This method of calculating the overall risk or insecurity is noted as a subset of "Frequency-based weights" in Decancq and Lugo (2008). There, it is noted as the example where a large percentage of the population subject to some variable (in this case, risk) is given a larger weight share.

the only issue being addressed is the weighting of individual components. This weighting is assumed to occur after any scaling techniques or other methodological tools are employed. For simplicity, all weighting methods are applied to the values of the components of the IEWB after each has been scaled using the linear scaling technique. Furthermore, the examples used assume the weighting of the subcomponents is accurate, although this might not be the case. This final aggregation of an index is the most visible, as it is the first set of weights that users encounter. Although index creators must first develop the weighting of sub-indices, the user of an index works in the opposite direction. In order to understand an index, one must first know the values of the components. Only then will the question of how these values are calculated arise. For this reason, the only step being considered in this paper is the very last step of creating a composite indicator: the *final* aggregation. Any results of this review of the final aggregation can easily be applied to subcomponents if deemed necessary; however, there is no reason to assume that the optimal method of weighting the four main components is the same as the optimal method of weighting the subcomponents.

Table 1: Weighting Schemes of the IEWB

	Consumption	**Wealth**	**Equality**	**Security**
IEWB	0.25	0.25	0.25	0.25
Alternative 1	0.4	0.1	0.25	0.25
Alternative 2	0.33	0.33	0	0.33
Alternative 3	0.2	0.1	0.4	0.3

The explicit weights used in the final aggregation of the components of the IEWB have been criticized as not being representative of the true valuations of society. In order to address this concern, the Osberg and Sharpe offer an IEWB based on equal weights alongside three alternative weightings, which are depicted in Table 1.

Table 2: Ordinal Rankings of the IEWB Under Four Weighting Schemes

	Baseline	Alternative 1	Alternative 2	Alternative 3
Highest well-being	Norway	Norway	Norway	Norway
	Denmark	Denmark	Netherlands	Denmark
	Germany	Belgium	Denmark	Sweden
	Belgium	Sweden	Germany	Finland
	Netherlands	Netherlands	Belgium	Belgium
	Sweden	France	Canada	Germany
	Finland	Germany	United States	France
	France	Finland	United Kingdom	Netherlands
	Canada	Australia	Australia	Australia
	United Kingdom	United Kingdom	France	United Kingdom
	Australia	Canada	Sweden	Canada
	Italy	Italy	Finland	Italy
	United States	United States	Italy	Spain
Lowest well-being	Spain	Spain	Spain	United States

Osberg and Sharpe generally report that the rankings of the IEWB are robust amongst these alternatives. In the most recent update of the IEWB, the top performing country was always Norway, and Spain was always in the bottom two. Table 2 (Exhibit 6 in Osberg and Sharpe (2011)) lists the rankings from of the countries under all the alternatives and confirms that the ordinal rankings of nations are robust under these changes. Unfortunately, each alternative weighting scheme is simply an individual set of valuations.[4] As such, the weighting of the IEWB must be further justified and other options for weighting must be analyzed in order to ensure the best possible practice is used for the weighting of the IEWB. For example, it may be more appropriate to allow the data to offer relative measures that reflect the overall preferences of society rather than explicitly choosing weights. The ideal weighting for the IEWB will have three characteristics:

1. Understandable Procedures (Transparency)
2. Objective Capturing of Societal Valuations[5]
3. Comparable Index Values across Regions and Time

The current IEWB satisfies the first and last condition but fails the objectivity test. Through analysis of different methodologies, an appropriate alternative for the weighting of the IEWB might be discovered. This discussion should help determine which methods of weighting are, *in practice*, easily applicable to the weighting of composite indices. We also offer comments on the different methods of determining weights in an attempt to discover which weighting schemes are truly objective from a statistical point of view, which have made progress in eliminating individual influence on the determination of weights, and which have merely hidden individual influence in statistical methods.

The lessons learned and to be learned are important when weighting composite indicators. The IEWB has discovered several 'best practices' in organizing its data, as it has augmented its weighting to that best perceived by users and critics. These lessons are applied to different methodologies in order to analyze the suitability of each to an index such as the IEWB.

II. EXPLICIT WEIGHTING

a. Equal Weighting

Hagerty and Land (2007) have shown that where data on subjective weightings of a particular index are not available, the methodology that results in the lowest level of disagreement among large variance in individuals' weightings is the system of equal weights. Therefore, the use of equal weighting is justified when surveys of the weights people place on the components of an index are not available. In practice, these weights will rarely be available, as most indices contain a unique set of components (otherwise an index would be irrelevant, as it would be a replication of another index). Hagerty and Land also make two other major conclusions. First, that "When correlations among social indicators are all positive...then

[4] Although this cross-section of valuations might be representative of the major views held by society, it is also certainly possible that the true societal valuations of the components are not included in any of these alternatives.
[5] Societal preferences are inherently subjective. Indeed, the weights that reflect the valuation society places on each components of an index are simply the average of the individual valuations of every person. Here, we attempt to objectively determine approximations of these subjective measures.

agreement will be high regardless of the variation in weights," (485). Secondly, that "disagreement is much rarer than expected and occurs only when the distribution of individuals' weights is (a) bimodal and (b) negatively correlated," (485). The first conclusion is not satisfied for the IEWB (see Table 3) because consumption is negatively correlated with both equality and security. Therefore, it is not necessary under the first condition that agreement be high.

Table 3: Correlation of Variables in the IEWB for 14 OECD Countries, 2009

	Consumption	Wealth	Equality	Security
Consumption	1	0.456	-0.584	-0.566
Wealth	*	1	0.214	0.098
Equality	*	*	1	0.785
Security	*	*	*	1

The second condition imposed by Hagerty and Land is slightly more abstract. The Centre for the Study of Living Standards has not conducted a survey on the weights people assign to each indicator in the IEWB. Therefore, it is impossible to know the exact distribution of weights; however, there is reason to believe that these weights are not bimodal. For instance, although certain people may hold extreme views concerning consumption and equality, it is unlikely that these extreme views dominate the balanced approach. There is not a strong basis for a negative weight consumption inequality when consumption is considered irrelevant. Furthermore, if consumption is not important, we assume future consumption (wealth) is also not important under the condition that the valuations of this individual remain unchanged. The opposite does not hold - an argument can be made that equality is not important based on societal incentives. Under these two arguments, upward pressure is placed on the importance of consumption at the fringe of valuations. If this accurately represents reality, the distribution of weights is unlikely to be bimodal; however, it is very possible this does not hold. Indeed, a survey will be required to ensure that it does indeed hold - and if it does not, then disagreement will continue to blossom.[6]

Overall, the results presented by Hagerty and Land are in line with the reasoning behind choosing equal weighting for the IEWB. The original weighting scheme used in the calculation of the index was what is currently presented as alternative 1. Osberg and Sharpe (2009a, 2011) changed this after being "criticized for a bias against sustainability...in favour of material goods," (vi, 10). The experience of the IEWB therefore supports the conclusion of Hagerty and Land concerning the minimization of disagreement. Furthering their arguments, if the value of an index is determined by how many people agree with an index's weighting scheme, the IEWB can further improve its weighting scheme through the use of survey data collected to measure the relative importance society places on each component; however, this raises a number of issues. First, should this survey be based on the overall population or on the population of each geographical unit covered by the index? Values are likely to range across countries (or even the provinces of Canada), and the IEWB for both Canada and OECD countries is likely to be markedly different under the two different weightings. In the IEWB for OECD countries, survey weights will place a large weight on the valuations of the typical American. These valuations might differ from the valuations of a small nation, such as Norway. If this is the case, the

[6] Furthermore, if a survey is conducted, it would be best to use the data from the survey to weight the components, as the results of Hagerty and Land (2007) indicate (and as reason would suggest).

weighting of Norway's IEWB will be influenced much more by the valuations of Americans than by the valuations of Norwegians, which will not accurately represent the state of economic well-being in Norway as perceived by Norwegians. Therefore, it might be more reasonable to allow weights to vary across countries. This option will be discussed in various sections in this paper.[7]

b. Expert Weighting

The OECD Handbook on Constructing Composite Indices (2006) describes expert weighting as a 'Budget Allocation Process' (BAP). In reality, these terms are used interchangeably; however, the term 'BAP' implies only that there is a 'budget' of weights that must be distributed amongst indicators, whereas 'expert weighting' implies that the process is completed by professionals familiar with the field being studied. Therefore, although the OECD Handbook mentions that 'experts' might in fact be the entire electorate, the term 'expert weighting' will be used in this paper to stress the fact that weighting decisions were made based on the valuations offered by professionals.

The first weighting of the IEWB was based on 'expert weighting.' In essence, this weighting boils down to decisions based on theories or observed behaviour recognized by the experts. Originally, the creators of the IEWB explicitly assigned the set of weights described by alternative 1. The reasoning given by Osberg and Sharpe (2009a, 2011) for this choice of weights was that "these weights reflected observed aggregate proportions for consumption and savings," (10). These valuations were critiqued. The experience of the IEWB, therefore, suggests that a reasonable basis for weights is not always satisfactory. These weights are still chosen somewhat subjectively by the creators of the index and are therefore often at odds with the valuations of a particular user.

c. User-Weighting - The Example of the OECD Better Life Index

A simple solution to this difference in valuations is to allow each individual to weight the components of an index according to their own preferences. The IEWB has an online tool[8] which allows users to weight the four components of the IEWB according to their preferences, but Osberg and Sharpe continue to compute estimates for reports using base weights (equal weights).

In contrast, the weighting scheme for the OECD Better Life Index is very simple: it provides *only* an online tool. It does not attempt to weight indicators; however, variables within each index are combined and scaled according to the methodology of the researchers. The same is true of the IEWB online weighting tool. Therefore, complete control of the weighting scheme is still not completely in the hands of the users.

Another problem quickly arises: what are the policy implications of the results? User-weighting allows each individual to obtain index values based on their personal valuations. Therefore, the [perfect] weighting for each individual is obtainable through the online weighting tool; however, a summary index for policy use is desirable but not calculable if there is no agreed

---

[7] Specifically, basic data envelopment analysis assigns different weights to different countries.

[8] The online weighting tool for OECD countries is available alongside the paper and the excel database at: http://www.csls.ca/iwb/oecd.asp. The link to the weighting tool can be found on the left side of the screen amongst other IEWB-related links.

upon medium for the determination of the results.[9] For example, if policy makers choose the valuations, these might not represent the valuations of society.

The best solution to the weighting issue for a summary index would be to obtain a large, representative sample of the population being observed. As noted by Hagerty and Land (2007), this is the best process for increasing the degree of acceptance of a baseline weighting scheme. On the other hand, the collection of data would not come without its own set of complexities. An important issue becomes how one should weight the raw data on the valuations of each individual. Should this be done by the population of each country? If so, this may skew the results in favour of the more populous nations, as it is conceivable that many people will hold valuations similar to the results in their country. At the same time, this may be desirable if this is what the majority of people sincerely believe. Another easily implementable weighting scheme based on survey data could be developed based on the online weighting tool. The OECD is planning a similar action. The frequently asked questions on the Better Life Index webpage confirm that the OECD is planning to develop a weighting based on the aggregation of weightings from individuals:

> "Every time you create an Index, it will enter a publicly accessible database. Over time, this will help the OECD to build up a picture of what you, citizens from across the world, believe shapes a good life. In the future, we will use this data to allow online visitors to cross-compare their Better Life Indexes with other people from around the world."[10]

Unfortunately, this weighting might not accurately reflect the views of society, as the population that visits the OECD website is likely a recognizable subset of the population. Self-selection bias therefore exists. Also, there exists potential for people to manipulate the contents of the database in order to create a given bias. Finally, people are likely to test the robustness of their valuations against various other valuations.[11] These attempts to compare differing values may distort the database being prepared by the OECD with multiple hypothetical valuations per user. The database is at risk of becoming a non-representative sample.

The second option for providing a baseline index is more problematic: it raises the question of which statistical procedure to use. The online weighting tool is useful and should definitely be maintained; however, the development of baseline weights is possible using one of the many procedures discussed in this paper. The choices involved in the development of this weighting scheme are wide; however, the solution to the issue of weighting is not to avoid the issue altogether when presenting the index. For this reason, the OECD should be commended for including an online weighting tool and attempting survey weighting; however, further contemplation is required concerning the baseline estimates. The database of inputted weights is

---

[9] It should be noted that the OECD website offers each user the choice to begin with equal weights; however, this does not constitute a baseline index.

[10] Response to the FAQ "Why should I share my index?" available along with other FAQ at http://www.oecdbetterlifeindex.org/about/better-life-initiative/

[11] In fact, this is why an online weighting tool is particularly useful - it allows people to compare valuations. The authors of the IEWB "stressed the subjectivity of value judgments and have provided access to Microsoft Excel spreadsheets so that readers can assess for themselves the implications of differing value judgments," (Osberg and Sharpe, 2009a, 2011; p. vi).

unlikely to produce representative valuations of the Better Life indicators. As such, the OECD should consider the variety of options explored in this paper.

d. Survey Weighting

The weighting of an index should reflect the valuations of society. The simplest way to replicate these valuations relatively accurately is to collect a representative sample of the population. Furthermore, the conclusions of Hagerty and Land (2007) indicate that this is the best practice when the goal is to minimize the amount of disagreement concerning the weighting of an index. Unfortunately, the Centre for the Study of Living Standards does not have the resources to conduct such a survey at the moment - the costs and energy required to conduct such a survey are enormous. The database of peoples' valuations of the Better Life Index being developed by the OECD is a good attempt at conducting a survey of peoples' weights but will be subject to a large bias - it will not be a representative sample. For this reason, survey weighting remains the optimal solution to the weighting issue, but it is rarely achievable. The next section highlights statistical methods designed to proxy for this societal valuation.

III. STATISTICAL METHODS

There are a number of statistical methods that are used in the derivation of weights for composite indicators. In many instances, an explicit weighting scheme is not optimal. Although explicit weights have the favourable quality of being transparent, the options available often do not satisfy other necessary conditions.[12] Many models of endogenous weight assignment exist which might be beneficial to the construction of a composite index such as the IEWB. Given that the weights are chosen based on patterns in the data, the researchers cannot be criticized for a particular bias in their assignment of weights; however, there are potential criticisms concerning the nature of the weights and the models being used to assign them.

In this light, various statistical techniques are applied to the IEWB in order to gauge their practical use.[13] The first subsection considers the use of factor analysis in weighting the components of the IEWB. Although there are many algorithms for the assignment of weights based on factor analysis, this paper focuses on the method used in the OECD Handbook on Constructing Composite Indicators (2006).[14] The second subsection briefly describes the possibility of regression analysis in the weighting of the IEWB. The third subsection looks at the results of the Composite Learning Index (CLI), produced by the Canadian Council on Learning. The implications of the methodology used by the CLI are discussed and applied to the context of

[12] For example, the IEWB does not have the resources to collect a representative sample and therefore cannot employ 'survey weighting.' Although 'user-weighting' is easy to implement, this tool is not useful as a summary index and therefore cannot be used for policy-making. 'Expert weighting' was employed by the IEWB but has been replaced by equal weighting due to criticisms arising concerning the weights chosen. Furthermore, the equal weighting of the IEWB is simply another individual set of subjective weights which are unlikely to represent the true valuations of society. Therefore, all methods of explicit weighting have proven to have some conflict with the requirements of the IEWB. Although equal weighting proposes the least conflict for a summary index, this approach must be justified or replaced.

[13] Not all statistical methods of weighting are considered here.

[14] With the slight difference that we do not consider the application of principal components analysis in this paper and therefore do not combine the two methods.

the IEWB. Due to the many extensions of data envelopment analysis, this statistical method follows in its own section.

a. Factor Analysis

Factor analysis is a simple way to endogenously determine the weights to be used in the summation of a composite index. This methodology aims to describe the data with a set of orthogonal factors which are considerably large. This evaluation is generally at the discretion of the researcher; however, as Rummel (1967) notes, the most important factors will generally be included despite differences in the criteria used for selection. For example, whereas Rummel uses a 0.50 as the minimum value for eigenvalues in his example, the OECD Handbook on Constructing Composite Indicators suggests using the condition that an eigenvalue be larger than one. Under Rummel's condition four factors are isolated in his example, whereas under the OECD's condition only two factors would have been isolated. These two factors are the largest of the four factors used by Rummel. Indeed, the total variance explained by each factor in Rummel's example is 40.9, 22.5, 9.1, and 7.6 per cent, respectively. The difference in total explained variation in this example is only 16.7 per cent[15]. For simplicity, the guidelines of the OECD are followed in the application of factor analysis to the IEWB data (the factors and their respective eigenvalues and shares of the variance of the data are presented in Table 4). The OECD Handbook suggests three qualifications for a factor:

1. Eigenvalue > 1
2. Explains more than 10 per cent of the variance in the data set.
3. The basket of factors chosen must explain more than 60 per cent of the variance in the data set.

Under these conditions, we accept the first two factors. The factor loadings generated by this analysis, along with rotated factor loadings[16] and scaled, squared factor loadings of the IEWB are presented in Table 5.

Table 4: Eigenvalues of the Factors of the IEWB Data Set

	Eigenvalue	Proportion	Cumulative
Factor 1	1.99	0.754	0.754
Factor 2	0.91	0.347	1.101
Factor 3	-0.09	-0.034	1.067
Factor 4	-0.18	-0.067	1.00

[15] 9.1 + 7.6 = 16.7

[16] The factor rotation matrix used to obtain this rotated factor loadings result is:

$$\begin{bmatrix} 0.9618 & -0.2736 \\ 0.2736 & 0.9618 \end{bmatrix}$$

This was an orthogonal rotation using the varimax method. As such, many features of the non-rotated matrix remain unchanged (such as variance share); however, the factors become orthogonal

Table 5: IEWB under Factor Analysis

Factor Loadings of the IEWB

	Factor 1	**Factor 2**
Consumption	-0.7680	0.4404
Wealth	-0.0591	0.7816
Equality	0.8583	0.2813
Security	0.8133	0.1758

Rotated Factor Loading of the IEWB

	Factor 1	**Factor 2**
Consumption	-0.6182	0.6337
Wealth	0.1570	0.7680
Equality	0.9025	0.0357
Security	0.8303	-0.0534
Explained VARIANCE	1.91078	0.99546
Proportion of Variance	0.6575	0.3425

Squared Factor Loading of the IEWB Scaled to sum to 1

	Factor 1	**Factor 2**
Consumption	0.200014	**0.403381**
Wealth	0.0129	**0.592475**
Equality	**0.426281**	0.00128
Security	**0.360805**	0.002864

Table 6: Weights Assigned to Each Component of the IEWB

	Domain Weight	*Weight of the Respective Factor*	*Weight Score* (w_i)	*Resulting Weight* $(\sum w_i = 1)$
Consumption	0.403381	0.3425	0.138158	0.160913
Wealth	0.592475	0.3425	0.202923	0.236344
Equality	0.426281	0.6575	0.28028	0.326442
Security	0.360805	0.6575	0.237229	0.276301

Following the path outlined by the OECD Handbook, we use the approach of Nicoletti et al. (2000) and form a value for each factor. Here factor 1 is formed by Equality and Security, with weights of 0.426 and 0.361, respectively. Factor 2 is formed by consumption and wealth, with weights of 0.403 and 0.592, respectively. Continuing this methodology, these values are then weighted based on the proportion of variance explained by the respective factor. The results of this summation are provided in Table 6 above.

Note from Table 7 that the ordinal rankings of the IEWB are robust under this sensitivity analysis. The largest single change in rankings was an increase from sixth place to third place by Sweden. This can be explained the high valuation of equality under the weights determined by factor analysis. The high valuation can be explained by the large variance in the equality domain.[17]

[17] A quick calculation determines that the variance of the consumption domain is 0.013, the variance of the wealth domain is 0.023, the variance of the equality domain is 0.038, and the variance of the security domain is 0.018. The equality domain therefore has, by far, the largest variance of the four domains of the IEWB.

Table 7: Results of the IEWB under Factor Analysis Weights

	Rank	*New Index Value*	*Equal Weights Rank*
Norway	1	0.793331	1
Denmark	2	0.706239	2
Sweden	3	0.662832	6
Germany	4	0.661541	3
Finland	5	0.657619	7
Belgium	6	0.652863	4
Netherlands	7	0.623788	5
France	8	0.617981	8
Canada	9	0.557551	9
UK	10	0.552215	10
Australia	11	0.543669	11
Italy	12	0.528727	12
Spain	13	0.448947	14
USA	14	0.408883	13

The merits of factor analysis are hotly debated. Salzman and Sharpe (2003) critique the situation[18] where a component which varies little is belittled in favour of a component with more impressive variance. Nicoletti et al. (2000) support factor analysis precisely for this reason. The authors argue that where the variance of an indicator is low, that indicator cannot be responsible for any significant variance in outcomes. In the case of the IEWB, which measures the relative well-being of the citizens of fourteen OECD countries, this certainly holds; if a particular component doesn't vary, it cannot be possible for any given change in economic well-being. We therefore consider the results of this factor analysis among potential new methodologies for the IEWB.

b. Regression Analysis

Regression analysis is a natural method of weighting composite indicators. Multiple linear regression is used to estimate a set of weights which 'best fit' the data by the criteria of least squares. Therefore, regression can be used to estimate the weights of a set of variables when there is a dependent variable. Unfortunately, according to Decancq and Lugo (2008), it is often difficult to find a suitable dependent variable. For example, in the case of the IEWB, there is no dependent variable to which it seems natural to weight economic well-being.

c. Case Study: The Composite Learning Index

The Composite Learning Index (CLI), produced by the Canadian Council on Learning (CCL), uses a methodology that conveniently delegates the assignment of weights to a statistical procedure involving several techniques. This statistical procedure uses principal components analysis, factor analysis, and regression analysis. According to the Canadian Council on Learning, the determination of the weighting of the CLI "pillars"[19] (or sub-indices) is based on a

[18] In their analysis, Salzman and Sharpe (2003) focus on principal components analysis

[19] The components of the CLI include: 'Learning to Be', 'Learning to Do', 'Learning to Know', and 'Learning to Live Together'. Each of these is composed of several indicators. In total, there are 26 observed 'measures' which form 17 'indicators.' For a complete description of the variables included in the CLI, please visit http://www.cli-

procedure which maximizes the correlation the of the CLI scores with scores of the an index which measures desirable economic and social results, the Economic and Social Well-Being Index (ESWBI)[20]. This procedure is more statistically complicated than the methodology of most indices, and the following summary of steps is adapted from a complete review of the index by Saisana (2008).

First, the indicators from each pillar were readjusted for direction and converted to z-scores to allow for comparability. Second, the researchers applied factor analysis to isolate uncorrelated unobserved variables which explain more than 85% of the variation of other indicators. The researchers then applied the same factor analysis to isolate a common unobserved factor from the ESWBI. A multiple regression analysis was conducted with the isolated variables from the learning indicators as independent variables and the isolated variable from the ESWBI as the dependent variable. This produces the weightings of the isolated learning indicators within each pillar which produce the strongest correlation with the isolated factor from the ESWBI, as the least squares analysis of the regression ensures the strongest correlation (or highest R-squared). These weights were applied to produce a value for each pillar of learning, which were once more converted to z-scores for comparability. The researchers applied Principal Components Analysis to isolate orthogonal values for the pillars of learning. Another multiple regression analysis similar to previous one was conducted to calculate the weights to be assigned to each pillar of learning. The independent variables were the orthogonal values for the pillars of learning and the dependent variable was once again the isolated variable from the ESWBI. These weights were applied to produce a value for the CLI in each community, and these values were then scaled.

Unfortunately, this approach does not lend itself to the calculation of every index. In this case, it seems natural to link learning data to economic and social data as skills and knowledge ultimately affect economic and social outcomes through human capital. In other instances, there might not be indicators which lend themselves as natural dependent variables (Decancq and Lugo, 2008). For example, the Index of Economic Well-Being measures the economic and social outcomes to which the CLI links. It would therefore be unreasonable to link the IEWB to another set of economic and social variables.

Furthermore, this method involves some level of implicit subjective weighting. The choice of indicators of the ESWBI as the dependent variables in the regression analyses implies a choice of desired economic and social results. In this light, the CLI weights its pillars in the manner that produces the strongest correlation with a single subjective evaluation of desirability. Conceptually, this is the same as weighting the pillars by coefficients determined in an arbitrary manner. An exogenous change in the variables chosen for the ESWBI will shift the data points of the dependent variable in the regression and will therefore change the optimal weighting of the pillars. The choice of the 'desirability' of outcomes is quite similar to the choice of which pillars

---

ica.ca/en/about/about-cli/methodology.aspx and select 'four pillars of learning' for detailed information on these components.

[20] The ESWBI is composed of seven indicators (CCL, 2010): income, unemployment, adult literacy, early childhood development, population health, environmental responsibility, voter participation. The first two are listed as economic outcomes and the last five are listed as social outcomes.

are most important. The choice of variables in the ESWBI therefore acts as an implicit weighting scheme in the calculation of CLI scores.

The methodology of the Composite Learning Index solves some problems related to individual subjective evaluation, given that the weights are not explicitly chosen, yet has the potential to create new sources of bias in the construction of an index. Although the process is more statistical, it is not completely free of the bias inherent in the selection of variables in the ESWBI. Therefore, the framework used in the calculation of the Composite Learning Index is an intriguing tool to submerge weighting decisions in statistical analysis; however, the net result is a system that can still be influenced by decisions as to what is desirable and, given the complex nature of the derivation of weights, is considerably less transparent than traditional methods of weighting.

This does not mean that the framework used in the weighting of the CLI is not useful. Assuming that the variables chosen by the researchers for the ESWBI are the ideal variables, the calculations are accurate and reflect the optimal weighting scheme under the criterion that the index be as closely correlated as possible to the desired outcomes. The error, as explained, lies in the subjective choice of variables in the ESWBI. In situations where there is only one possible indicator to which an index can be linked in a regression analysis, the methodology of the CLI might be optimal.

The framework of the CLI offers a useful new tool for weighting indices but more research is required to determine the optimal dependent variable set. Assuming an optimal dependent variable set can be established, perhaps by a survey which estimates what people feel are most important, the framework of the CLI will produce weights for each of its domains without subjectivity. Otherwise, the framework simply complicates matters, making the process less transparent, and the weights will continue to be shrouded by the subjectivity of the decisions of the researchers.

IV. DATA ENVELOPMENT ANALYSIS

This method, explained briefly in the OECD Handbook on Constructing Composite Indicators, uses the best performing observations in each indicator to create a ‘boundary’ of feasible performance which is then used to measure the score of each observation. It was first developed in its modern form by Charnes et al. (1978) and presented by Melyn and Moesen (1991) in the field of macroeconomics. The top-ranking country in each measured component is given the score 1 and the data point is incorporated into the boundary. This restriction ensures that no judgment is made about the relative importance of each of the components.[21] Therefore, the weighting of each component will be determined uniquely for each observation in the data. Consider this simple example from the IEWB for OECD Countries:

[21] In the assignment of the value '1' to top performers, we are arguing that a 'best performance' in any indicator is optimal (in terms of efficiency), regardless of the trade-off between this indicator and the other variables. This judgment can be augmented by certain restriction and extensions described in later sections.

Assume the only relevant components are the Index of Security and the Index of Consumption. Then, the top performances belong to the United States in consumption (0.909) and Norway in security (0.829). Following the rule that a top performance in any given domain is efficient, both the United States and Norway are assigned the value of 1 in this hypothetical index. When comparing the results of Canada, which has a consumption value of 0.674 and a security value of 0.661, to these countries, find the intersection of the line from the origin to the data point for Canada and this boundary line, if it exists.[22] The result is a score of 0.869 for Canada[23], compared to scores of 1.000 for both Norway and the United States.

a. DEA in n Dimensions

Of course, n variables could also be combined using DEA by applying the process n-1 times. This may be cumbersome, but the larger problem with this is the number of countries that might be ranked as equivalent in order to satisfy the conditions of DEA. An n-dimensional extension of DEA used in constructing composite indicators is the Benefit-of-the-Doubt Approach (BOD). This methodology simplifies the analysis required to complete the process. An overview of the description of BOD in the OECD Handbook on Constructing Composite Indicators is offered below (this description is adapted from the OECD Handbook - equations are taken directly from the text):

The value produced by the BOD approach is the ratio of a country's score to the score of an imaginary country I* which "is the score of the hypothetical country that maximises the

[22] The line that represents the boundary line, where data points are represented by (Consumption, Security), is (0.909, 0.280) + t(0.153, -0.549). The line that represents the segment from the origin to the Canadian data point is simply u(0.674, 0.661).

$$(0.909, 0.280) + t(-0.153, 0.549) = u(0.674, 0.661)$$
$$(0.909 - 0.153t, 0.280 + 0.549t) = (0.674u, 0.661u)$$

Then two equations must be satisfied:

$$0.909 - 0.153t = 0.674u$$
$$\frac{0.909}{0.674} - \frac{0.153}{0.674}t = u$$

and:

$$0.280 + 0.549t = 0.661u$$
$$0.280 + 0.549t = 0.661\left(\frac{0.909}{0.674} - \frac{0.153}{0.674}t\right)$$
$$0.280 + 0.549t = 0.661\left(\frac{0.909}{0.674} - \frac{0.153}{0.674}t\right)$$
$$t = 0.874713$$

Then:

$$u = \frac{0.909}{0.674} - \frac{0.153}{0.674}(0.874713)$$
$$u = 1.150102$$

Given that the intersection of these lines occurs where u = 1.150102, the ratio $\frac{1}{1.150102}$ will give the relative score of Canada in comparison to the boundary line created by the best performances in the two measured domains.

[23] The underlying weights for this result are consumption at 0.782 and security at 0.218. Interestingly, these are the same weights that are optimal for Canada under DEA at the four dimensional level as well, the results of which are displayed in Table 10.

overall performance (defined as the weighted average), given the (unknown) set of weights," (93):

$$CI_c = \frac{\sum_{c=1}^{M} I_{qc} w_{qc}}{\sum_{c=1}^{M} I^*_{qc} w_{qc}}$$

where I_{qc} is the score of a variable scaled using max-min for a particular country, c, and a particular variable, q. Similarly, w_{qc} is the weight assigned to this particular variable, q, for the particular country, c.

The explanation of the method is based largely on Cherchye, Moesen, and Van Puyenbroeck (2004), whom the OECD Handbook recognizes as the developers of the benchmark:

$$I^* = I^*(w) = \underset{I_k, k \in \{1, \ldots, M\}}{argmax} \left(\sum_{q=1}^{Q} I_{qk} w_q \right)$$

It is notable that k can be any country in the sample and this benchmark country is potentially different for each individual country being observed, as the weights chosen for each country are different, and are chosen to optimize the score of the country being considered.[24]

The end result of this approach will be quite similar to the 2-dimensional results of DEA. The countries that perform best in any particular indicator will have an indexed score of 1. As in DEA, this ensures that we are not favouring one variable over another in the calculation of index scores[25]. It is uncertain which method of weighting is optimal, and this methodology allows each country to use weights that are optimal under the conditions in that particular country. This does create potential issues. For example, in the Index of Economic Well-Being for selected OECD countries, the top ranking countries are the United States (in consumption), Finland (in equality), and Norway (in both wealth and security). We can therefore infer, without any calculations, that the indexed score of all 3 of these countries will be 1. This is not informative, as economic well-being is likely not the same in the United States, Finland, and Norway.

Table 8: Reported Results of the IEWB for Norway, Finland, and the United States

	Equal Weights	Alternative 1	Alternative 2	Alternative 3
Norway	0.799	0.775	0.826	0.768
Finland	0.626	0.621	0.564	0.684
United States	0.482	0.526	0.595	0.376

Baseline Weights: Consumption (0.25), Wealth (0.25), Equality (0.25), Security (0.25)
Alternative 1 Weights: Consumption (0.4), Wealth (0.1), Equality (0.25), Security (0.25)
Alternative 2 Weights: Consumption (0.33), Wealth (0.33), Equality (0), Security (0.33)
Alternative 3 Weights: Consumption (0.2), Wealth (0.1), Equality (0.4), Security (0.3)

[24] This optimization problem must be constrained so that weights are either equal to zero or positive. It is in most instances irrational to have negative weights. For instance, in the case of the IEWB, it is not likely somebody prefers consuming less, becoming less wealthy, more inequality, or more insecurity, although this is intuitively possible.
[25] The resulting weights for a given country might favour one variable over another; however, the index is calculated so that countries with specializations in a particular component are not penalized for this. In other words, the process allows a wide array of possible weighting based on optimality - countries are not judged on a set valuation of weights.

None of the explicit weights chosen for the IEWB imply that these three countries perform equally well. In general, Norway consistently out-performs Finland and the United States. If it is possible for the IEWB to be equal in all three countries under explicit weighting, there must be some unique weights that make this true. Using the values of the four domains of the IEWB available in figure 2, we can easily calculate the weights that would be required: Consumption (1.508), Wealth (-0.562), Equality (1.379), and Security (-0.699). This result is illogical – people certainly value wealth and economic security. Otherwise, social security systems would not be in place in any of these countries. Therefore, there is no rational solution to the system of equations which satisfies equality in index scores for these three countries.

Table 9: Values of the Four Domains of the IEWB for Norway, Finland, and the United States

	Consumption	Wealth	Equality	Security
Norway	0.756	0.917	0.692	0.829
Finland	0.468	0.500	0.793	0.742
United States	0.909	0.614	0.123	0.280

The BOD approach therefore produces 'false' results at the high end. Although it may provide useful comparisons between countries at the low end of the spectrum, comparisons between countries in the high end are not possible. Consequently, any comparisons between countries at the low end of the index and countries at the high end of the index might also be skewed. For example, if Norway and the United States have the same index score, then the comparison between Canada and the United States and the comparison between Canada and Norway will yield the same results. If the comparison between Norway and the United States fails, these comparisons with Canada must also fail. Nevertheless, Table 10 provides results of the BOD approach for the IEWB for selected OECD countries. Note that five countries tie for first place under this method and that every country applies a weight of zero to at least one variable. These are extreme values which are not likely to represent the true valuations of people in these countries – yet another argument to augment the DEA approach.

An index should be used for comparisons both across observations and time. Otherwise, the index loses a lot of value. A change in BOD procedure will therefore be required to ensure that comparability is not lost. At the same time, a goal of the index is to maintain transparency and not be subject to a specific set of valuations. For this reason, some options to restore comparability between Norway, Finland, and the United States must be discarded. The first option - to readjust the scores based on a given set of valuations - is not optimal simply because the reasoning behind using the BOD approach is to remove this type of subjectivity. Another set of options - to link the score of each of Norway, Finland, and the United States back to the original data - is preferred.

Table 10: DEA/BOD Approach Results for the IEWB

	Consumption*	Wealth*	Equality*	Security*	Result	New Rank	EW Rank
Australia	0.784	0	0.040	0.176	0.908	**9**	11
Belgium	0.306	0	0.694	0	0.966	**6**	4
Canada	0.782	0	0	0.218	0.869	**13**	9
Denmark	0	0	0.545	0.455	1	**1**	2
Finland	0	0	1	0	1	**1**	7
France	0.306	0	0.694	1.56×10^{-9}	0.925	**7**	8
Germany	0	0.205	0.795	0	0.921	**8**	3
Italy	0	0	0	1	0.878	**12**	12
Netherlands	0.788	0	0.212	0	0.897	**10**	5
Norway	0.250	0.25	0.25	0.25	1	**1**	1
Spain	0	0	0	1	0.696	**14**	14
Sweden	0.018	0	0.747	0.234	1	**1**	6
United Kingdom	0	0	0	1	0.888	**11**	10
United States	0.856	0.119	0	0.025	1	**1**	13

*denotes Weight

There are two possibilities regarding this latter case: adjustments can be made to every index score, or adjustments can be made only to the scores of the top ranking countries. In the latter scenario, the top countries might be repositioned in order to give them comparable values. Consequently, the meaning of the comparison between these countries and the other countries will be lost, as the index scores will no longer have been calculated using the same methodology. As well, this will not solve the issue of extreme valuations of the domains. Therefore, a readjustment should be made to all index scores if a readjustment is deemed necessary.

b. Benefit-of-the-Doubt/DEA with Common Weights

The literature on data envelopment analysis has recently expanded significantly. A main focus of this recent discussion has been on the reassignment of weights to each country based on restrictions which ensure some level of commonality. This might be possible by weighting with an average of the optimal weights (Doyle, 1995) or otherwise creating some sort of common weights from the previous analysis, thereby continuing to be objective in our choice of weights while permitting meaningful comparison across all countries.

The simplest scenario is to take the average optimal weight (suggested by Doyle, 1995); however, if one country is ahead or behind by a large margin in a given indicator, the weightings of the other countries might be a lot lower in the case of an extremely good performance and might be a lot higher in the case of an extremely poor performance. This may very well create a bias against countries that perform either very poorly or very well in an individual indicator. Furthermore, there is no reason why the each country should be given equal weight in the weighting decision, resulting in another weighting conundrum. This circular problem will not solve the weighting issue. Other extensions of the DEA model which consider this additional criterion must therefore be considered.

Table 11: DEA Results Under Common Weights

	Index under Average Weight[26]	***RANKING***	***Index under 'Compromise Solution'***	***RANKING***
Australia	0.600789	9	0.701972	2
Belgium	0.670923	4	0.66001	8
Canada	0.575008	11	0.667881	7
Denmark	0.712596	2	0.672482	5
Finland	0.67033	5	0.595887	13
France	0.663831	6	0.671087	6
Germany	0.655081	7	0.629281	11
Italy	0.554356	12	0.632176	10
Netherlands	0.627122	8	0.676271	4
Norway	0.762711	1	0.790473	1
Spain	0.486558	13	0.53708	14
Sweden	0.695555	3	0.647899	9
United Kingdom	0.593844	10	0.688512	3
United States	0.421555	14	0.615388	12

Kao and Hung (2005) introduced the 'compromise solution' to the spectrum of common weights in DEA. Under this methodology, a common set of weights, which minimizes the differences between the common set of weights and each observation's individual set of weights, is derived. Note in Table 11 that the weight assigned to consumption under the compromise solution is 53.3 per cent and the weight assigned to security is 46.7 per cent. Both equality and wealth are assigned compromise weights of 0. This can be explained by the low valuation of wealth in unconstrained DEA[27] and the fact that security had the fewest valuations of zero in the same analysis. As noted earlier in the paper, consumption and security have a negative correlation whilst equality and security have a positive correlation. Consumption and equality have a negative correlation. For this reason, weighting equality above zero would increase the performance of those performing well under security whilst hindering the performance of those already performing poorly for economic security. In order to balance this effect, or create a compromise for the inclusion of security, there therefore exists a large weight on consumption.

There are many different versions of these weighting restrictions. For example, Makui et al. (2008) recently introduced what they believe to be an improvement upon the model of Kao

[26] The average optimal weights are consumption (29.2 per cent), wealth (4.1 per cent), equality (35.6 per cent), and security (31.1 per cent).
[27] The low valuation of wealth in baseline DEA can be explained by the fact that Norway performs very well in Wealth. Placing a high value on wealth will make Norway the baseline against which the country is compared, which will produce a low estimate. In order to maximize this estimate, the wealth domain is therefore given a relatively small weight. Given the relatively small weight, the weights that have the smallest average distance to each set of optimal weights will also place a relatively small weight on the wealth domain, in this case, 0.

and Hung (2005). Indeed, essentially any weighting method discussed in this paper can be used to weight the weighting produced by basic data envelopment analysis.

c. Benefit-of-the Doubt/DEA with Weight Constraints

In some cases it is not the different weights that cause an issue. It might be preferable to weight indices based on these differing weights but with simple adaptations. Cherchye et al. (2008) use panel information to set weight restrictions (ranges) for the components of the Technology Achievement Index. Interestingly, under the new restrictions, Singapore ranks lowest, whilst under the original DEA approach, Singapore ranked as one of the top nations (242-243). Clearly, the restrictions on weights in DEA analysis are crucial.

Another problem with researcher interference has therefore been created in order to make comparable a method that might be employed to reduce the interference of the researcher in the output of a composite indicator. Note that this interference is less restrictive than the original concept of fixed weights; however, if the optimal analysis is to satisfy the three conditions set at the beginning of the paper, there must be no valuations based on the whims of the researchers. Results for a constrained DEA analysis are provided in Table 12 using the maximum and minimum values used currently used in sensitivity analyses for the IEWB – a broad base for selection of weights.

Table 12: DEA/BOD Approach Results for the IEWB with Weight Constraints

	Cons.*	Wea.*	Equ.*	Sec.*	Result	New Rank	DEA Rank	EW Rank
Australia	0.4	0.1	0.167	0.333	0.800	**9**	9	11
Belgium	0.25	0.1	0.4	0.25	0.873	**5**	6	4
Canada	0.4	0.1	0.167	0.333	0.782	**11**	13	9
Denmark	0.2	0.1	0.4	0.3	0.939	**2**	1	2
Finland	0.2	0.1	0.4	0.3	0.890	**4**	1	7
France	0.2	0.1	0.4	0.3	0.849	**7**	7	8
Germany	0.2	0.1	0.4	0.3	0.863	**6**	8	3
Italy	0.4	0.1	0.167	0.333	0.732	**13**	12	12
Netherlands	0.4	0.1	0.25	0.25	0.831	**8**	10	5
Norway	0.25	0.25	0.25	0.25	1.000	**1**	1	1
Spain	0.4	0.1	0.167	0.333	0.628	**14**	14	14
Sweden	0.2	0.1	0.4	0.3	0.909	**3**	1	6
United Kingdom	0.4	0.1	0.167	0.333	0.789	**10**	11	10
United States	0.4	0.333	0	0.267	0.776	**12**	1	13

* Consumption weight between 0.2 and 0.4, Wealth weight between 0.1 and (1/3), Equality weight between 0 and 0.4, and Security weight between 0.25 and (1/3). These are based on the minimum and maximum valuations assigned to each variable in the four alternatives currently offered by the IEWB.

Interestingly, the countries group themselves into two major groups, with the exception of the United States, Belgium, the Netherlands, and Norway, which have unique weightings. The first group includes Australia, Canada, Italy, Spain, and the United Kingdom, which weights consumption at its maximum (0.4), wealth at its minimum (0.1), equality at 0.167, and security at its maximum (0.333). The second group includes Denmark, Finland, France, Germany, and Sweden, and weights consumption at its minimum (0.2), wealth at its minimum (0.1), equality at its maximum (0.4) and security at 0.3. This process allows each country to weight each indicator according to a reasonable range and has the desirable effect of both allowing variations in valuations across countries and eliminating the majority of extreme values. In this extension of DEA for the IEWB, there are no valuations of 100% and in only one instance (the valuation of equality by the United States) is a weight of 0% realized. Furthermore, these results are robust when compared with the general rankings of equal weighting and the simple DEA analysis. In only one comparison does a country switch from being a top-six country to being a bottom-six country or vice versa (the Netherlands from fifth in equal weighting to eighth in constrained DEA) and in this instance, the gap is only 3 positions (which is also the value of the largest gaps in the comparisons). The top ranked country is always Norway and the bottom ranked country is always Spain.

Therefore, whether or not DEA analysis is optimal compared to equal weighting appears to have little significant value on the ordinal results of the IEWB. These results are robust. As such, this constrained DEA analysis is well-suited for a sensitivity analysis; however, the fact that the weights must be constrained in some way does not eliminate the influence of the researchers. As well, the issue of comparability between countries with different weightings arises. Although we can compare these countries on the definition of 'efficiency' under optimal weights, this method does not allow a direct comparison along an individual set of weights. The inclusion of an online weighting tool alongside this baseline method might mitigate this issue.

V. CONSUMPTION-EQUIVALENT VALUATIONS

Although indexed values might be comparable across observations and time, the value of an index is often of little significance when analyzed in isolation. For example, suppose there is an index with a range of [0, 1], where a value of 1 is the best possible value and 0 is the worst possible value. Suppose further that country x has an index value of 0.461 and country y has an index value of 0.654. When compared, it is clear that, relative to country x, country y is performing $\frac{I_y}{I_x} = \frac{0.654}{0.461}$ times better than country x (approximately 1.419 or 41.9 per cent). On the contrary, it is impossible to infer much from the value $I_y = 0.654$ without information concerning the performance of other countries or the performance of this country in the past. This isolated value does have meaning: the country has an indexed value at 65.4% of the maximal index value. Unfortunately, this does not provide a real method of comparison. A solution to this problem is to convert this measure to an easily understandable equivalent - consumption. In "Beyond GDP? Welfare across Countries and Time," Charles Jones and Peter Klenow use a consumption-equivalent measure to calculate welfare, producing the desired effect of an easily communicable measure of welfare.

Jones and Klenow (2010) present measures of consumption-equivalent welfare based on four domains: life expectancy, consumption, leisure, and inequality. The results of their analysis are meant to be "simple summary statistic for a nation's flow of welfare," (2). As such, their model is unique and many extensions of their model cannot by generalized for all indices. The values calculated by the authors are based on a utility function defined by:

$$V(e,c,l,\sigma) = e(\bar{u} + \log c + v(l) - \frac{1}{2}\sigma^2)$$

where e is life expectancy, c is consumption, l is leisure, and σ is inequality. The question asked by the authors is then, "By what factor, λ_i, must we adjust Rawls' consumption in the United States to make him indifferent between living in the two countries?" (9) where the second country is country i and Rawls is a random person living in the United States.

In order to answer this question, certain restrictive assumptions are placed on the data used by the authors. This is done in order to make calculations based on the utility function possible. For example, a value of leisure must be calculated in this utility function. Therefore, an assumption is made that people sleep 8 hours per day. This results in an overall time allocation of 16 hours per day, or 5840 hours per year (13). This is done to split sleep from leisure, which is probably a reasonable choice, as sleeping is a necessary bodily function and is therefore not equivalent to other leisure activities. At the same time, given that people value sleep, the trade-offs between sleep, leisure, and work are likely different across countries given the different labour markets and cultures. The assumption that people in all countries sleep an average of 8 hours per day is not realistic, but it must be accepted for reasonable calculations to be made. Jones and Klenow also state that "expected utility for Rawls is... $\frac{e \cdot u(C,l)}{100}$" (8). This is another unrealistic assumption - that mortality is the same across all ages. Fortunately, Jones and Klenow note this inaccuracy and correct this problem for a small set of countries when using micro-data.

Nevertheless, the measures produced by Jones and Klenow are based on subjective and restrictive assumptions on the functional form as well. The utility function used by the authors weights the domains based on evaluations of the equivalence of each domain to consumption. Consider the solution to the utility function presented by the authors:

$$\log \lambda_i = \frac{e_i - e_{us}}{e_{us}}\left(\bar{u} + \log c_i + v(l_i) - \frac{1}{2}\sigma_i^2\right) + (\log c_i - \log c_{us}) + \left(v(l_i) - v(l_{us})\right) - \frac{1}{2}(\sigma_i^2 - \sigma_{us}^2)$$

The first term represents life expectancy, the second term represents consumption, the third term represents leisure and the fourth term represents inequality. The authors also present a solution for compensating variation:

$$\log \lambda_i^{cv} = \frac{e_i - e_{us}}{e_i}\left(\bar{u} + \log c_{us} + v(l_{us}) - \frac{1}{2}\sigma_{us}^2\right) + (\log c_i - \log c_{us}) + \left(v(l_i) - v(l_{us})\right) - \frac{1}{2}(\sigma_i^2 - \sigma_{us}^2)$$

The terms in this function represent the same domains as their respective terms in the first solution equation, which is the authors' depiction of equivalent variation. There is an economic basis for the choices made when developing the functional form, and these rationales are explained in the text. For example, basing their utility function on the literature, Jones and Klenow assume leisure "takes a form that implies a constant Frisch elasticity of labor supply," (15). These reflections, therefore, offer no argument against the measure developed by Jones and Klenow - it is the purpose of this paper to discuss functional forms, not the various measures of welfare that are being developed. Furthermore, the estimation of welfare in terms of a consumption-equivalent measure is certainly a potential improvement in the construction of indices; however the use of this measure certainly does not solve the weighting issue for any given index.[28]

Overall, the methodology of the paper offers two potential solutions to the weighting crisis. Clearly, the same consumption-equivalent utility function cannot be transferred from this paper and used in the weighting of arbitrary indices as each index will be comprised of different components; however, consumption-equivalence is a good measure for an index as it is easily understandable and offers immediate resonance about the conditions implied by an index. Therefore, if it is desirable, the consumption-equivalent utility function must be adjusted for each index in order to incorporate each of the variables that are included. This could be a difficult task. The functional form chosen for the utility function will necessarily weight the variables. The solution to this would be to weight variables according to how much people actually value them - survey-based weighting. On the other hand, it is not feasible to collect data on the valuations of each variable that might be included in an index. There are far too many contexts in which a person might be asked for their valuation of a variable and an infinite number of possible variables. For example, the relative valuation of life expectancy and consumption will be dependent on the initial level of both variables. Therefore, an inference of the relative valuations must be made in some manner.

The result is a set of two options. First, the utility function can be weighted using any of the other methods in this paper. The coefficients used in the calculation of an index can be determined based on factor analysis, regression-based estimation, explicit weighting, or data envelopment analysis. These methods, as discussed earlier in the paper, each have their own setbacks that might make their use in weighting a utility function undesirable.[29] Second, an index can use a consumption-equivalent utility function and allow for observers to change parameters with an online tool. A complete set of robustness checks such as the ones performed by Jones and Klenow validates this option and produces understandable values of welfare in each country. In this case, it is important that each assumption and parameter be stated explicitly alongside results so that accurate inferences can be made from the data. The second option is either to convert index values to consumption values or make no attempt at consumption-equivalence. Although an interesting measure that resonates with readers, there is fundamentally no difference between scaled values and consumption-equivalent values of the same variables if they are

---

[28] Given the theoretical backing of the utility function currently used, the implicit weighting derived from the utility function can be described as 'expert weighting.' These weights can be described by the partial derivatives of the function. Therefore, if we compute the Jacobian matrices for the solutions at $(e_i^*, c_i^*, l_i^*, \sigma_i^*)$ we obtain the results offered in Appendix A. The simplest manner of testing that null hypothesis that these are indeed the correct valuations would be to collect survey data on valuations people place on each level of consumption, inequality, life expectancy, and leisure. Unfortunately, undertaking this analysis would be unreasonable as it would involve the use of data that is too complex to be available. This problem is quite similar to the problem discussed in the section on 'survey weighting' and is the reason analyses of different weighting methods are required.

[29] On the other hand, many functional form changed that might be considered desirable (such as logs) can easily be performed on the data before these analyses are conducted.

weighted in the same way (see Appendix A). Therefore, as long as the optimal weighting scheme has been chosen, the unit of measurement will be irrelevant.

Is it possible to calculate the Index of Economic Well-Being in consumption equivalent measures? Certainly. The question becomes: how do we value wealth, economic security, and economic equality? It is furthermore unclear what the standard should be for each domain - at what value of each domain the effect on consumption-equivalence would be zero. For example, concerning economic equality - should we begin with a base of consumption and increase consumption-equivalence as equality goes up or should we begin with the same base and decrease consumption-equivalence as inequality increases from perfect equality? Should value be calculated from some weighted average?[30] If anything, the decision to switch to consumption-equivalence increases the number of questions that must be answered in order to compute a final value for each observation. This, therefore, increases the subjectivity of the results. Notably, Jones and Klenow opt to use only the relevant measures of compensating variation and equivalent variation; however, due to these directional issues, it is much simpler and more transparent to simply weight a composite index with explicit weights[31]. In the end, the consumption equivalent measure is a useful unit for index measurement, but it does not solve the weighting issue – it is simply another (potentially accurate) form of expert weighting.

VI. COMPARISON OF WEIGHTING METHODS

Multiple weighting methods were adapted for use in the weighting of the IEWB. Appendix C and Table 13 summarize the results of the various analyses that were attempted in this paper, in order of their robustness with the dominance rankings presented in the following section. These are followed by weighting methodologies discussed but not adapted to the IEWB.

Table 13: Summary of the Optimality of Weighting Methodologies

Methodologies that satisfy any of the three optimal qualities for the IEWB:
1. Understandable Procedures (Transparency) *Equal Weights, Alternative 1, Alternative 2, Alternative 3, User-Weighting, Survey Weighting*
2. Objective Determination of Societal Valuations *DEA, Common (Avg.) Weights DEA, Compromise Solution, User-Weighting, Survey Weighting*
3. Comparable Index Values Across Regions and Time *Equal Weights, Alternative 1, Alternative 2, Alternative 3, User-Weighting, Survey Weighting, Consumption-Equivalent Measure*
Methodologies that satisfy all of the three optimal qualities for the IEWB:
User-Weighting, Survey Weighting

[30] It should be noted that this issue is only relevant in the calculation of an absolute value for consumption-equivalent welfare. When calculating relative values of consumption-equivalent welfare, the 'benchmark' for the base value of inequality is irrelevant - a relative increase in inequality will require a decrease in relative consumption-equivalent welfare.

[31] Of course, these weights can be estimated from theory to measure their relative contribution to welfare compared to consumption. If this is the case, the relative values of this indicator will be roughly equivalent to the consumption-equivalent measure, as the utility function defining consumption equivalence implicitly assigns weights to each indicator.

From Table 13, we see that user-weighting and survey weighting are the two optimal weighting methodologies under the three criteria specified at the beginning of the paper. It must be noted that each of these methodologies are optimal under a restrictive condition. User-weighting is only optimal under the condition that the user weighting the components is the target of the indicator - the only person whose valuation counts. Survey weighting is only optimal under the condition that these are obtained from a representative sample. As discussed previously, this can be extremely difficult to obtain. We therefore conclude that there is no perfect solution to the weighting issue.

VII. ROBUSTNESS - DOMINANCE CRITERIA

Researchers usually perform sensitivity analyses to ensure their results hold under various conditions. These robustness checks can be performed on the various methods of weighting and aggregating indicators. Furthermore, ordinal rankings of observations can be established from the robustness checks. For example, Saisana et al. (2011) derive the *median* rankings of universities under the Academic Ranking of World Universities (ARWU) and the Times Higher Education Supplement (THES) using seventy different methods involving data envelopment analysis, equal weighting, and factor analysis. Similarly, in a discussion of the different methodologies used to weight the HDI, Cherchye et al. (2008b) rank countries in the HDI based on dominance. Cherchye et al. propose to rank each observation based on the dominance of the observation over other observations under "*every possible normalization*," (12). They note that the number of possible methods is large and that agreement on this may not be possible; however, by increasing the number of methods used in the analysis, "a country's raw data...become increasingly more important in assessing that country's relative position in the sample," (12). Furthermore, this methodology reduces the subjective decisions of the creators of an index further, as the impact of decisions on methodology and weighting are minimized, and the rankings are based on robustness. Given that there is no optimal weighting methodology under the criteria established at the beginning of this paper, a similar analysis is attempted here.

Table 14: Ordinal Rankings of the IEWB for All Presented Weighting Methods

RANK	Equal Weights	Alt. 1	Alt. 2	Alt. 3	DEA
1	Norway	Norway	Norway	Norway	Denmark Finland Norway Sweden United States (5-way tie)
2	Denmark	Denmark	Netherlands	Denmark	
3	Germany	Belgium	Denmark	Sweden	
4	Belgium	Sweden	Germany	Finland	
5	Netherlands	Netherlands	Belgium	Belgium	
6	Sweden	France	Canada	Germany	Belgium
7	Finland	Germany	United States	France	France
8	France	Finland	United Kingdom	Netherlands	Germany
9	Canada	Australia	Australia	Australia	Australia
10	United Kingdom	United Kingdom	France	United Kingdom	Netherlands
11	Australia	Canada	Sweden	Canada	United Kingdom
12	Italy	Italy	Finland	Italy	Italy
13	United States	United States	Italy	Spain	Canada
14	Spain	Spain	Spain	United States	Spain

RANK	Constrained DEA	Compromise Solution (DEA)	Common Weights (Avg.)	Factor Analysis
1	Norway	Norway	Norway	Norway
2	Denmark	Australia	Denmark	Denmark
3	Sweden	United Kingdom	Sweden	Sweden
4	Finland	Netherlands	Belgium	Germany
5	Belgium	Denmark	Finland	Finland
6	Germany	France	France	Belgium
7	France	Canada	Germany	Netherlands
8	Netherlands	Belgium	Netherlands	France
9	Australia	Sweden	Australia	Canada
10	United Kingdom	Italy	United Kingdom	United Kingdom
11	Canada	Germany	Canada	Australia
12	United States	United States	Italy	Italy
13	Italy	Finland	Spain	Spain
14	Spain	Spain	United States	United States

This method of ranking encompasses all the types of weighting, as observations will be ranked according to how many other observations they dominate. A summary of the ordinal rankings of the IEWB for each of the weighting methods analyzed is provided in Table 14. Cherchye et al. (2008b) propose both strong and weak dominance rules by which to order the results:[32]

x strongly dominates **y** for some p, $0 \leq p \leq \frac{1}{2}$, denoted $\boldsymbol{x} \succ_*^p \boldsymbol{y}$, if and only if

$$(1)_*^p \quad \forall \boldsymbol{b} \in \mathbb{B}^p \; and \; \forall I \in \mathbb{I}: I\left(\frac{x_1}{b_1}, \dots, \frac{x_n}{b_n}\right) \geq I\left(\frac{y_1}{b_1}, \dots, \frac{y_n}{b_n}\right)$$

$$(2)_*^p \quad \exists \boldsymbol{b} \in \mathbb{B}^P and \; \exists I \in \mathbb{I}: I\left(\frac{x_1}{b_1}, \dots, \frac{x_n}{b_n}\right) > I\left(\frac{y_1}{b_1}, \dots, \frac{y_n}{b_n}\right)$$

and

x weakly dominates **y** for some p, $0 \leq p \leq \frac{1}{2}$, denoted $\boldsymbol{x} \succ^p \boldsymbol{y}$, if and only if

$$(1)^p \quad \exists \boldsymbol{b} \in \mathbb{B}^P such\ that\ \forall I \in \mathbb{I}: I\left(\frac{x_1}{b_1}, \dots, \frac{x_n}{b_n}\right) \geq I\left(\frac{y_1}{b_1}, \dots, \frac{y_n}{b_n}\right)$$

$$(2)^p \quad \forall \boldsymbol{b} \in \mathbb{B}^p : \exists I \in \mathbb{I}\ such\ that\ I\left(\frac{x_1}{b_1}, \dots, \frac{x_n}{b_n}\right) > I\left(\frac{y_1}{b_1}, \dots, \frac{y_n}{b_n}\right)$$

Here, we adapt several elements of these dominance rules. Cherchye et al. (2008b) base their rankings of the Human Development Index based on the notion that agreement on the weights might vary according to the extremes represented by the percentile 'p'. A larger value of p represents a smaller range of valuations. This paper does not attempt this analysis. Further, we do not consider various normalizations - we only consider the aggregation of the components. Thus, b, which is the 'benchmark vector' determined from some normalization, is constant. So:

x strongly dominates **y**, denoted $\boldsymbol{x} \succ_* \boldsymbol{y}$, if and only if

$$(1)_* \quad \forall I \in \mathbb{I}: I(x_1, \dots, x_n) \geq I(y_1, \dots, y_n)$$

[32] These rules are taken directly from the text of Cherchye et al. (2008b). The rule for strong domination is taken from page 13 and the rule for weak domination, page 14. I is a particular aggregation and $\mathbb{I}$ is the set of aggregations.

$$(2)_* \quad \exists I \in \mathbb{I}: I(x_1, \dots, x_n) > I(y_1, \dots, y_n)$$

and

x weakly dominates **y**, denoted $\boldsymbol{x} \succ \boldsymbol{y}$, if and only if

$$(1) \quad \forall I \in \mathbb{I}: I(x_1, \dots, x_n) \geq I(y_1, \dots,, y_n)$$
$$(2) \quad \exists I \in \mathbb{I} : I(x_1, \dots, x_n) > I(y_1, \dots, y_n)$$

Note that these rules are now equivalent. This occurs because dominance was originally established across normalizations rather than across aggregations. We must therefore differ from the guidelines of Cherchye et al. (2008b) and present very simple rules for dominance:[33]

x strongly dominates **y**, denoted $\boldsymbol{x} \succ_* \boldsymbol{y}$, if and only if

$$\forall I \in \mathbb{I}: I(x_1, \dots, x_n) > I(y_1, \dots, y_n)$$

and

x weakly dominates **y**, denoted $\boldsymbol{x} \succ \boldsymbol{y}$, if and only if

$$\forall I \in \mathbb{I}: I(x_1, \dots, x_n) \geq I(y_1, \dots,, y_n)$$

Strong and weak dominance charts based on these rules are presented for the IEWB under the nine tested weighting methods in Table 13.[34]

Table 15: Dominance Patterns of the 14 IEWB Countries

Strong Dominance

	Aus	Bel	Can	Den	Fin	Fra	Ger	Ita	Net	Nor	Spa	Swe	UK	US	Total
Australia		0	0	0	0	0	0	1	0	0	1	0	0	0	2
Belgium	0		0	0	0	0	0	1	0	0	1	0	0	0	2
Canada	0	0		0	0	0	0	0	0	0	1	0	0	0	1
Denmark	0	1	1		0	1	1	1	0	0	1	0	0	0	6
Finland	0	0	0	0		0	0	0	0	0	1	0	0	0	1
France	0	0	0	0	0		0	1	0	0	1	0	0	0	2
Germany	0	0	0	0	0	0		0	0	0	1	0	0	0	1
Italy	0	0	0	0	0	0	0		0	0	1	0	0	0	1
Netherlands	0	0	1	0	0	0	0	1		0	1	0	0	0	3
Norway	1	1	1	0	0	1	1	1	1		1	0	1	0	9
Spain	0	0	0	0	0	0	0	0	0	0		0	0	0	0
Sweden	0	0	0	0	0	0	0	1	0	0	1		0	0	2
United Kingdom	0	0	0	0	0	0	0	1	0	0	1	0		0	2
United States	0	0	0	0	0	0	0	0	0	0	0	0	0		0
TOTAL	1	2	3	0	0	2	2	8	1	0	12	0	1	0	***32***

[33] Note that the basic reasoning for producing these dominance rankings remains the same: instead of increasing the number of normalizations, we can increase the number of alternative weighting methods used in the calculation of the index. In doing so, the influence of the researchers is diminished in the same manner as increasing the number of normalizations under the original model.

[34] The rows represent the countries being considered whereas the columns represent the countries being compared to the row country (across all measures). A value of '1' indicates that the row country dominates the column country. The 'Total' column therefore calculates the number of countries that a row country dominates whereas the 'Total' row calculates the number of countries that dominate the respective column country.

Weak Dominance

	Aus	Bel	Can	Den	Fin	Fra	Ger	Ita	Net	Nor	Spa	Swe	UK	US	Total
Australia		0	0	0	0	0	0	1	0	0	1	0	0	0	2
Belgium	0		0	0	0	0	0	1	0	0	1	0	0	0	2
Canada	0	0		0	0	0	0	0	0	0	1	0	0	0	1
Denmark	0	1	1		1	1	1	1	0	0	1	1	0	1	9
Finland	0	0	0	0		0	0	0	0	0	1	0	0	0	1
France	0	0	0	0	0		0	1	0	0	1	0	0	0	2
Germany	0	0	0	0	0	0		0	0	0	1	0	0	0	1
Italy	0	0	0	0	0	0	0		0	0	1	0	0	0	1
Netherlands	0	0	1	0	0	0	0	1		0	1	0	0	0	3
Norway	1	1	1	1	1	1	1	1	1		1	1	1	1	13
Spain	0	0	0	0	0	0	0	0	0	0		0	0	0	0
Sweden	0	0	0	0	1	0	0	1	0	0	1		0	0	3
United Kingdom	0	0	0	0	0	0	0	1	0	0	1	0		0	2
United States	0	0	0	0	0	0	0	0	0	0	0	0	0		0
TOTAL	1	2	3	1	3	2	2	8	1	0	12	2	1	2	***40***

Cherchye et al. (2008b) rank countries based on net dominance, with the country that is least dominated by other countries ranking higher when there is a tie. Following these guidelines, Table 14 indicates that the ordinal rankings provided by the equal weighting scheme are reasonably robust. The top two countries and bottom country under the strong dominance rule, the weak dominance rule, and equal weighting, are the same. Moreover, these countries maintain their exact ordinal ranking. The largest change exhibited is by Germany, where both the strong and weak dominance rules rank Germany in the bottom tier of countries whereas equal weighting promotes Germany to third place. This can be explained by the poor performance of Germany under the 'compromise solution' (see Table 14). Overall, constrained DEA and alternative 1 tied for the most robust weighting methodology under these two dominance patterns.

On the other hand, these net dominance rankings are only two out of the eight possible dominance rankings that can be calculated from the two established dominance rules. Under each dominance rule, four rankings are possible. First, one can rank countries by net dominance and break ties by choosing the country that is least dominated by others. Second, one can rank countries by net dominance and break ties by choosing the country that dominates the most other countries. Third, one can rank countries by which country dominates the most other countries. Finally, one can rank countries by which is least dominated by other countries.

Table 16: Results of Dominance Ranking

<table>
<tr><th></th><th>Net Strong Dominance</th><th>Net Dominance</th><th>Net Weak Dominance</th><th>Net Dominance</th><th>Equal Weights</th></tr>
<tr><td>1</td><td>Norway</td><td>9</td><td>Norway</td><td>13</td><td>Norway</td></tr>
<tr><td>2</td><td>Denmark</td><td>6</td><td>Denmark</td><td>8</td><td>Denmark</td></tr>
<tr><td>3</td><td>Sweden</td><td>2</td><td>Netherlands</td><td>2</td><td>Germany</td></tr>
<tr><td>4</td><td>Netherlands</td><td>2</td><td rowspan="2">Australia and UK (Tied)</td><td>1</td><td>Belgium</td></tr>
<tr><td>5</td><td>Finland</td><td rowspan="2">1</td><td rowspan="2">1</td><td>Netherlands</td></tr>
<tr><td>6</td><td rowspan="2">Australia and UK (Tied)</td><td>Sweden</td><td>Sweden</td></tr>
<tr><td>7</td><td rowspan="2">1</td><td rowspan="2">Belgium and France (Tied)</td><td rowspan="2">0</td><td>Finland</td></tr>
<tr><td>8</td><td>United States</td><td>France</td></tr>
<tr><td>9</td><td rowspan="2">Belgium and France (Tied)</td><td>0</td><td>Germany</td><td>-1</td><td>Canada</td></tr>
<tr><td>10</td><td>0</td><td>United States</td><td rowspan="2">-2</td><td>United Kingdom</td></tr>
<tr><td>11</td><td>Germany</td><td>-1</td><td rowspan="2">Canada and Finland (Tied)</td><td>Australia</td></tr>
<tr><td>12</td><td>Canada</td><td>-2</td><td>-2</td><td>Italy</td></tr>
<tr><td>13</td><td>Italy</td><td>-7</td><td>Italy</td><td>-7</td><td>United States</td></tr>
<tr><td>14</td><td>Spain</td><td>-12</td><td>Spain</td><td>-12</td><td>Spain</td></tr>
</table>

Table 17: First Iteration of Dominance Rankings for the IEWB

<table>
<tr><th></th><th>Net Strong Dominance (With least dominated by)</th><th>Net Strong Dominance (With most dominance)</th><th>Strong Dominance</th><th>Least Strong Dominated</th><th>Net Weak Dominance (With least dominated by)</th><th>Net Weak Dominance (With most dominance)</th><th>Weak Dominance</th><th>Least Weak Dominated</th></tr>
<tr><td>1</td><td>Norway</td><td>Norway</td><td>Norway</td><td>Norway</td><td>Norway</td><td>Norway</td><td>Norway</td><td>Norway</td></tr>
<tr><td>2</td><td>Denmark</td><td>Denmark</td><td>Denmark</td><td>Denmark</td><td>Denmark</td><td>Denmark</td><td>Denmark</td><td>Denmark</td></tr>
<tr><td>3</td><td>Sweden</td><td>Netherlands</td><td>Netherlands</td><td>Sweden</td><td>Netherlands</td><td>Netherlands</td><td>Netherlands</td><td>Netherlands</td></tr>
<tr><td>4</td><td>Netherlands</td><td>Sweden</td><td>Sweden</td><td>Finland</td><td rowspan="2">Australia and UK (Tied)</td><td>Sweden</td><td>Sweden</td><td rowspan="2">Australia & UK (Tied)</td></tr>
<tr><td>5</td><td>Finland</td><td rowspan="2">Australia and UK (Tied)</td><td rowspan="2">Australia and UK (Tied)</td><td>United States</td><td rowspan="2">Australia and UK (Tied)</td><td rowspan="2">Australia and UK (Tied)</td></tr>
<tr><td>6</td><td rowspan="2">Australia and UK (Tied)</td><td>Netherlands</td><td>Sweden</td><td>Sweden</td></tr>
<tr><td>7</td><td>Finland</td><td rowspan="2">Belgium and France (Tied)</td><td rowspan="2">Australia and UK (Tied)</td><td rowspan="2">Belgium and France (Tied)</td><td rowspan="2">Belgium and France (Tied)</td><td rowspan="2">Belgium and France (Tied)</td><td rowspan="2">Belgium and France (Tied)</td></tr>
<tr><td>8</td><td>United States</td><td rowspan="2">Belgium and France (Tied)</td></tr>
<tr><td>9</td><td rowspan="2">Belgium and France (Tied)</td><td>Finland</td><td rowspan="2">Belgium and France (Tied)</td><td>Germany</td><td>Germany</td><td>Germany</td><td>Germany</td></tr>
<tr><td>10</td><td>United States</td><td>Germany</td><td>United States</td><td rowspan="2">Canada and Finland (Tied)</td><td rowspan="2">Canada and Finland (Tied)</td><td>United States</td></tr>
<tr><td>11</td><td>Germany</td><td>Germany</td><td>Canada</td><td>Germany</td><td rowspan="2">Canada and Finland (Tied)</td><td rowspan="2">Canada and Finland (Tied)</td></tr>
<tr><td>12</td><td>Canada</td><td>Canada</td><td>Italy</td><td>Canada</td><td>United States</td><td>Italy</td></tr>
<tr><td>13</td><td>Italy</td><td>Italy</td><td>United States</td><td>Italy</td><td>Italy</td><td>Italy</td><td>United States</td><td>Italy</td></tr>
<tr><td>14</td><td>Spain</td><td>Spain</td><td>Spain</td><td>Spain</td><td>Spain</td><td>Spain</td><td>Spain</td><td>Spain</td></tr>
</table>

Note that the countries that dominate well are usually not dominated by many other countries. For this reason, if one country is high in one of these eight rankings, it is usually high in the other seven rankings. Also note that if a country dominates well under these rankings, it dominates well under the original rankings produced by the various methods of weighting. Therefore, the determination of a nation's dominance under the dominance rules will proxy well for the determination of a nation's general dominance under the original rankings. In the interest of not selecting a particular method of ranking the dominance of a country, we conduct an iterative procedure on the calculation of a nation's dominance. That is, in the event of disequilibrium, calculations of a nation's dominance are taken on the preceding dominance tables (see Appendix B). The iterations continue until equilibrium is reached: where all eight methods of ranking the countries by dominance produce the same results. This occurs after six iterations.[35]

The strong and weak dominance rankings are the most comprehensive lists possible to include as ordinal rankings in a paper. The iterative equilibrium of dominance rankings is the optimal choice in dominance rankings, as it preserves the overall dominance structure and considers all possible dominance rankings under evaluated dominance rules. Considering robustness checks, these rankings are a natural extension of the process which conveniently summarize the results obtained. The results furthermore allow us to compare results from each procedure and determine which are the most robust. The ordinal ranking for the IEWB considered most desirable in this light is provided in Table 18.

Table 18: Iterative Equilibrium of Dominance Rankings for the IEWB

1	Norway
2	Denmark
3	Sweden and Netherlands (Tied)
4	
5	Australia, UK, and Finland (Tied)
6	
7	
8	Belgium and France (Tied)
9	
10	United States
11	Germany
12	Canada
13	Italy
14	Spain

In terms of the IEWB, the most robust under the original dominance rankings are constrained DEA for statistical procedures and alternative 1 for explicit procedures[36]; this does

[35] The second through sixth iterations are provided in Appendix B alongside the original iteration.
[36] In fact, based on the average difference between rankings of each method and the two original dominance rankings, constrained DEA and alternative 1 tied for first place with an average distance of 2.04 positions. Note that

not change under the iterative equilibrium. If robustness is based on the average difference between a country's observed ordinal ranking and its ideal ordinal ranking under the iterative equilibrium, the different methods of weighting are listed in descending order from most robust to least robust in Table 17.

Therefore, constrained data envelopment analysis remains the most robust method of calculating weights and alternative 1 remains the most robust explicit set of weights. Equal weighting, on the other hand, does not appear to be relatively robust when considering the original dominance structures or the iterative equilibrium.

Table 19: Robustness of Various Methods of Weighting (Similarity to Iterative Dominance)

1	Constrained DEA
2	Alternative 1
3	DEA
4	Alternative 3 and Common (Avg.) Weights (Tied)
5	
6	Factor Analysis
7	Equal Weights
8	Compromise Solution
9	Alternative 2

VIII. CONCLUSIONS AND FURTHER RESEARCH

The criteria for the optimal weighting of the Index of Economic Well-being are that the weighting methodology is understandable, that it is objective (or minimizes the subjectivity on the part of the researchers), and that it produces comparable results. The most consistent problem experienced by the IEWB is that of subjectivity on the part of the creators in weighting the index.

A quick review of methodological choices has enabled the Centre for the Study of Living Standards (CSLS) to analyze the options available. The ideal solution would be to weight the IEWB based on survey results; however, the collection of a representative sample is a task for which the CSLS does not have the resources. Fortunately, it appears that the results of the index are largely robust to any changes made to the method of weighting used to aggregate the index at the final level. The results of the dominance checks appear to be the best methodological tool to rank the IEWB in ordinal terms; however, the IEWB also requires cardinal rankings so that index values may be compared across time and across countries in relative terms. It may be possible to produce cardinal rankings from this ordinal ranking. Given that the ranking was produced using a plethora of weighting methodologies, this same set of methods may be weighted to produce cardinal rankings. In simple terms, it may be possible to weight the weighting methodologies to produce a set of weights; however, this is not attempted in this paper.[37]

---

this is the average of the average distance for the weak dominance structure and the average distance for the strong dominance structure.

[37] This set of weights might be estimated based on the level of similarity between each method of weighting considered and the final rankings based on the iterative dominance equilibrium. Unfortunately, this would decrease the transparency of the weighting methodology.

Constrained data envelopment analysis appears to be the most robust method of weighting considered. This statistical procedure also has the desirable property that the influence of the index creators is somewhat lessened. Maggino and Zumbo (2011) extend the argument that principal components analysis supports an incorrect belief in the objectivity of the weighting process which is nearly impossible to achieve in measures of social outcomes (Salzman and Sharpe: 2003) across all statistical methods. This certainly applies to the model of constrained data envelopment analysis: the maximum and minimum possible values of each weight were determined from the original 'expert weighting' alternatives. This methodology is not completely objective. This problem is compounded by the fact that constrained data envelopment analysis results in different weights for different countries. Although common weight alternatives were explored, these methods of weighting rank only fourth and eighth (of nine) in terms of robustness. We discard these options in favour of the top three methods: constrained DEA, alternative 1, and DEA. Further discarding both constrained DEA and DEA on the grounds that different weights are used for different countries, alternative 1 appears to be the most robust method of weighting that satisfies the basic needs of the IEWB. The arguments for this particular 'expert weighting' discussed by Osberg and Sharpe (2009a, 2011) are therefore supported.[38]

Unfortunately, the disagreement surrounding this original weighting cannot be dismissed. As confirmed by Hagerty and Land (2007), equal weights reduce the amount of conflict surrounding the weighting of an index, and has certainly done so for the IEWB. Furthermore, if we redefine robustness based on the number of *exact* ordinal positions maintained, then equal weighting places first alongside four other weighting methodologies (constrained DEA, 'compromise solution' DEA, factor analysis, and DEA). Alternative 1 ranks only sixth out of nine countries under this definition. Indeed, there is no method of weighting considered for which it can be demonstrably shown as always more desirable than equal weighting.

The availability of an online weighting tool allows the users of an index to weight the index according to their valuations. If such a tool accompanies an index, the results are both transparent and adjustable to each individual's preferences. Although the index may not produce a baseline value truly representative of society's preferences, the method of equal weights reduces tension over the weighting of the index. Therefore, the equivalent weighting method appears to be a useful proxy for the valuations of society when data concerning these valuations do not exist. The constrained DEA approach produces the most robust rankings of the IEWB; however, these rankings are based on weights that vary across countries and are therefore not convenient when attempting to compare countries on the same basis. Alternative 1 is the most robust explicit weighting option; however, the experience of the IEWB indicates that this alternative is not as agreeable with the general perception of what the weighting should be. Therefore, although the IEWB should include many sensitivity analyses (especially constrained DEA and alternative 1) and perhaps produce a more robust ordinal ranking using dominance rules, the cardinal baseline index ought to remain weighted by equal weights.

---

[38] It should be noted that this result might change if more weighting methods are added to the analysis. As more and more weighting methods are added, the more informative the iterative dominance equilibrium will be. Nine weighting methods is a relatively small number of methods to base the procedure upon. As such, the analysis ought to be replicated with more weighting methods in order to confirm the results obtained here. For now, we take the results as given.

APPENDIX A - Weights Under the Jones and Klenow (2010) Utility Function

If we compute the Jacobian matrices for the solutions at $(e_i^*, c_i^*, l_i^*, \sigma_i^*)$ and denote the Jacobian matrix for equivalent variation $DF^{ev}(e_i^*, c_i^*, l_i^*, \sigma_i^*)$ and for compensating variation $DF^{cv}(e_i^*, c_i^*, l_i^*, \sigma_i^*)$, we obtain the result:

$$DF^{ev}(e_i^*, c_i^*, l_i^*, \sigma_i^*) = \left(\frac{\partial F^{ev}}{\partial e_i}(e_i^*, c_i^*, l_i^*, \sigma_i^*) \quad \frac{\partial F^{ev}}{\partial c_i}(e_i^*, c_i^*, l_i^*, \sigma_i^*) \quad \frac{\partial F^{ev}}{\partial l_i}(e_i^*, c_i^*, l_i^*, \sigma_i^*) \quad \frac{\partial F^{ev}}{\partial \sigma_i}(e_i^*, c_i^*, l_i^*, \sigma_i^*)\right)$$

For convenience, let us represent this as a gradient:

$$\nabla F^{ev} = \begin{pmatrix} \left[\frac{1}{e_{us}}\left(\bar{u} + \log c_i + v(l_i) - \frac{1}{2}\sigma_i^2\right)\right](e_i^*, c_i^*, l_i^*, \sigma_i^*) \\ \left[\frac{e_i - e_{us}}{e_{us} \cdot c_i} + \left(\frac{1}{c_i}\right)\right](e_i^*, c_i^*, l_i^*, \sigma_i^*) \\ [\frac{e_i - e_{us}}{e_{us}} v'(l_i) + v'(l_i)](e_i^*, c_i^*, l_i^*, \sigma_i^*) \\ [-\frac{e_i - e_{us}}{e_{us}} \sigma_i - \sigma_i](e_i^*, c_i^*, l_i^*, \sigma_i^*) \end{pmatrix}$$

and

$$DF^{cv}(e_i^*, c_i^*, l_i^*, \sigma_i^*) = \left(\frac{\partial F^{cv}}{\partial e_i}(e_i^*, c_i^*, l_i^*, \sigma_i^*) \quad \frac{\partial F^{cv}}{\partial c_i}(e_i^*, c_i^*, l_i^*, \sigma_i^*) \quad \frac{\partial F^{cv}}{\partial l_i}(e_i^*, c_i^*, l_i^*, \sigma_i^*) \quad \frac{\partial F^{cv}}{\partial \sigma_i}(e_i^*, c_i^*, l_i^*, \sigma_i^*)\right)$$

Again, for convenience, let us represent this as a gradient:

$$\nabla F^{cv} = \begin{pmatrix} \left[-\frac{1}{e_i}\left(\bar{u} + \log c_{us} + v(l_{us}) - \frac{1}{2}\sigma_{us}^2\right)\right](e_{us}^*, c_{us}^*, l_{us}^*, \sigma_{us}^*) \\ \left[\frac{e_i - e_{us}}{e_i \cdot c_{us}} - \left(\frac{1}{c_{us}}\right)\right](e_{us}^*, c_{us}^*, l_{us}^*, \sigma_{us}^*) \\ [\frac{e_i - e_{us}}{e_i} v'(l_{us}) - v'(l_{us})](e_{us}^*, c_{us}^*, l_{us}^*, \sigma_{us}^*) \\ [-\frac{e_i - e_{us}}{e_i} \sigma_{us} + \sigma_{us}](e_{us}^*, c_{us}^*, l_{us}^*, \sigma_{us}^*) \end{pmatrix}$$

Given these matrices, we can estimate differences between the marginal changes in the equivalent variation and the compensating variation at different values of life expectancy, consumption, leisure, and inequality. These marginal changes in consumption equivalence can be considered the 'weights' of each variable at the given values of each variable.

APPENDIX B - Iterative Calculation of Dominance Rankings

From Table 13 we can rank each dominance rule according to four different ranking patterns. First, one can rank countries by net dominance and break ties by choosing the country that is least dominated by others. Second, one can rank countries by net dominance and break ties by choosing the country that dominates the most other countries. Third, one can rank countries by which country dominates the most other countries. Finally, one can rank countries by which is least dominated by other countries. This produces eight different rankings of countries that can be established under the two dominance rules.

Appendix Table 1: First Iteration

<table>
<tr><th></th><th>Net Strong Dominance (With least dominated by)</th><th>Net Strong Dominance (With most dominance)</th><th>Strong Dominance</th><th>Least Strong Dominated</th><th>Net Weak Dominance (With least dominated by)</th><th>Net Weak Dominance (With most dominance)</th><th>Weak Dominance</th><th>Least Weak Dominated</th></tr>
<tr><td>1</td><td>Norway</td><td>Norway</td><td>Norway</td><td>Norway</td><td>Norway</td><td>Norway</td><td>Norway</td><td>Norway</td></tr>
<tr><td>2</td><td>Denmark</td><td>Denmark</td><td>Denmark</td><td>Denmark</td><td>Denmark</td><td>Denmark</td><td>Denmark</td><td>Denmark</td></tr>
<tr><td>3</td><td>Sweden</td><td>Netherlands</td><td>Netherlands</td><td>Sweden</td><td>Netherlands</td><td>Netherlands</td><td>Netherlands</td><td>Netherlands</td></tr>
<tr><td>4</td><td>Netherlands</td><td>Sweden</td><td>Sweden</td><td>Finland</td><td rowspan="2">Australia and UK (Tied)</td><td>Sweden</td><td>Sweden</td><td rowspan="2">Australia & UK (Tied)</td></tr>
<tr><td>5</td><td>Finland</td><td rowspan="2">Australia and UK (Tied)</td><td rowspan="2">Australia and UK (Tied)</td><td>United States</td><td rowspan="2">Australia and UK (Tied)</td><td rowspan="2">Australia and UK (Tied)</td></tr>
<tr><td>6</td><td rowspan="2">Australia and UK (Tied)</td><td>Netherlands</td><td>Sweden</td><td>Sweden</td></tr>
<tr><td>7</td><td>Finland</td><td rowspan="2">Belgium and France (Tied)</td><td rowspan="2">Australia and UK (Tied)</td><td rowspan="2">Belgium and France (Tied)</td><td rowspan="2">Belgium and France (Tied)</td><td rowspan="2">Belgium and France (Tied)</td><td rowspan="2">Belgium and France (Tied)</td></tr>
<tr><td>8</td><td>United States</td><td rowspan="2">Belgium and France (Tied)</td></tr>
<tr><td>9</td><td rowspan="2">Belgium and France (Tied)</td><td>Finland</td><td rowspan="2">Belgium and France (Tied)</td><td>Germany</td><td>Germany</td><td>Germany</td><td>Germany</td></tr>
<tr><td>10</td><td>United States</td><td>Germany</td><td>United States</td><td rowspan="2">Canada and Finland (Tied)</td><td rowspan="2">Canada and Finland (Tied)</td><td>United States</td></tr>
<tr><td>11</td><td>Germany</td><td>Germany</td><td>Canada</td><td>Germany</td><td rowspan="2">Canada and Finland (Tied)</td><td rowspan="2">Canada and Finland (Tied)</td></tr>
<tr><td>12</td><td>Canada</td><td>Canada</td><td>Italy</td><td>Canada</td><td>United States</td><td>Italy</td></tr>
<tr><td>13</td><td>Italy</td><td>Italy</td><td>United States</td><td>Italy</td><td>Italy</td><td>Italy</td><td>United States</td><td>Italy</td></tr>
<tr><td>14</td><td>Spain</td><td>Spain</td><td>Spain</td><td>Spain</td><td>Spain</td><td>Spain</td><td>Spain</td><td>Spain</td></tr>
</table>

Note that the countries that dominate well are usually not dominated by many other countries. For this reason, if one country is high in one of these eight rankings, it is usually high in the other seven rankings. Also note that if a country dominates well under these rankings, it dominates well under the original rankings produced by the various methods of weighting. Therefore, the determination of a nation's dominance under the dominance rules will proxy well for the determination of a nation's general dominance under the original rankings. In the interest of not selecting a particular method of ranking the dominance of a country, we conduct an iterative procedure on the calculation of a nation's dominance. If the rankings are not identical, the same procedure is applied to these dominance rankings. The iterations continue until equilibrium is reached: where all eight methods of ranking the countries by dominance produce the same results. In the case of this paper, this occurs after six iterations.

Appendix Table 2: Second Degree Dominance Patterns

Strong Dominance

	Aus	Bel	Can	Den	Fin	Fra	Ger	Ita	Net	Nor	Spa	Swe	UK	US	TOTAL
Aus		1	1	0	0	1	1	1	0	0	1	0	0	0	6
Bel	0		1	0	0	0	1	1	0	0	1	0	0	0	4
Can	0	0		0	0	0	0	1	0	0	1	0	0	0	2
Den	1	1	1		1	1	1	1	1	0	1	1	1	1	12
Fin	0	0	0	0		0	0	1	0	0	1	0	0	0	2
Fra	0	0	1	0	0		1	1	0	0	1	0	0	0	4
Ger	0	0	1	0	0	0		1	0	0	1	0	0	0	3
Ita	0	0	0	0	0	0	0		0	0	1	0	0	0	1
Net	1	1	1	0	0	1	1	1		0	1	0	1	0	8
Nor	1	1	1	1	1	1	1	1	1		1	1	1	1	13
Spa	0	0	0	0	0	0	0	0	0	0		0	0	0	0
Swe	0	1	1	0	1	1	1	1	0	0	1		0	1	8
UK	0	1	1	0	0	1	1	1	0	0	1	0		0	6
US	0	0	0	0	0	0	0	0	0	0	1	0	0		1
TOTAL	3	6	9	1	3	6	8	11	2	0	13	2	3	3	70

Weak Dominance

	Aus	Bel	Can	Den	Fin	Fra	Ger	Ita	Net	Nor	Spa	Swe	UK	US	TOTAL
Aus		1	1	0	0	1	1	1	0	0	1	0	1	0	7
Bel	0		1	0	0	1	1	1	0	0	1	0	0	0	5
Can	0	0		0	0	0	0	1	0	0	1	0	0	0	2
Den	1	1	1		1	1	1	1	1	0	1	1	1	1	12
Fin	0	0	1	0		0	0	1	0	0	1	0	0	0	3
Fra	0	1	1	0	0		1	1	0	0	1	0	0	0	5
Ger	0	0	1	0	0	0		1	0	0	1	0	0	0	3
Ita	0	0	0	0	0	0	0		0	0	1	0	0	0	1
Net	1	1	1	0	0	1	1	1		0	1	0	1	0	8
Nor	1	1	1	1	1	1	1	1	1		1	1	1	1	13
Spa	0	0	0	0	0	0	0	0	0	0		0	0	0	0
Swe	0	1	1	0	1	1	1	1	0	0	1		0	1	8
UK	1	1	1	0	0	1	1	1	0	0	1	0		0	7
US	0	0	0	0	0	0	0	0	0	0	1	0	0		1
TOTAL	4	7	10	1	3	7	8	11	2	0	13	2	4	3	75

Appendix Table 3: Second Iteration

<table>
<tr><th></th><th>Net Strong Dominance (With least dominated by)</th><th>Net Strong Dominance (With most dominated)</th><th>Strong Dominance</th><th>Least Strong Dominated</th><th>Net Weak Dominance (With least dominated by)</th><th>Net Weak Dominance (With most dominated)</th><th>Weak Dominance</th><th>Least Weak Dominated</th></tr>
<tr><td>1</td><td>Norway</td><td>Norway</td><td>Norway</td><td>Norway</td><td>Norway</td><td>Norway</td><td>Norway</td><td>Norway</td></tr>
<tr><td>2</td><td>Denmark</td><td>Denmark</td><td>Denmark</td><td>Denmark</td><td>Denmark</td><td>Denmark</td><td>Denmark</td><td>Denmark</td></tr>
<tr><td>3</td><td rowspan="2">Sweden and Netherlands (Tied)</td><td rowspan="2">Sweden and Netherlands (Tied)</td><td rowspan="2">Sweden and Netherlands (Tied)</td><td rowspan="2">Sweden and Netherlands (Tied)</td><td rowspan="2">Sweden and Netherlands (Tied)</td><td rowspan="2">Sweden and Netherlands (Tied)</td><td rowspan="2">Sweden and Netherlands (Tied)</td><td rowspan="2">Sweden and Netherlands (Tied)</td></tr>
<tr><td>4</td></tr>
<tr><td>5</td><td rowspan="2">Australia and UK (Tied)</td><td rowspan="2">Australia and UK (Tied)</td><td rowspan="2">Australia and UK (Tied)</td><td rowspan="2">Australia and UK (Tied)</td><td rowspan="2">Australia and UK (Tied)</td><td rowspan="2">Australia and UK (Tied)</td><td rowspan="2">Australia and UK (Tied)</td><td>Finland</td></tr>
<tr><td>6</td><td>United States</td></tr>
<tr><td>7</td><td>Finland</td><td>Finland</td><td rowspan="2">Belgium and France (Tied)</td><td>Finland</td><td>Finland</td><td>Finland</td><td rowspan="2">Belgium and France (Tied)</td><td rowspan="2">Australia and UK (Tied)</td></tr>
<tr><td>8</td><td>United States</td><td rowspan="2">Belgium and France (Tied)</td><td>United States</td><td>United States</td><td rowspan="2">Belgium and France (Tied)</td></tr>
<tr><td>9</td><td rowspan="2">Belgium and France (Tied)</td><td>Germany</td><td rowspan="2">Belgium and France (Tied)</td><td rowspan="2">Belgium and France (Tied)</td><td>Finland</td><td rowspan="2">Belgium and France (Tied)</td></tr>
<tr><td>10</td><td>United States</td><td>Finland</td><td>United States</td><td>Germany</td></tr>
<tr><td>11</td><td>Germany</td><td>Germany</td><td>Canada</td><td>Germany</td><td>Germany</td><td>Germany</td><td>Canada</td><td>Germany</td></tr>
<tr><td>12</td><td>Canada</td><td>Canada</td><td>United States</td><td>Canada</td><td>Canada</td><td>Canada</td><td>United States</td><td>Canada</td></tr>
<tr><td>13</td><td>Italy</td><td>Italy</td><td>Italy</td><td>Italy</td><td>Italy</td><td>Italy</td><td>Italy</td><td>Italy</td></tr>
<tr><td>14</td><td>Spain</td><td>Spain</td><td>Spain</td><td>Spain</td><td>Spain</td><td>Spain</td><td>Spain</td><td>Spain</td></tr>
</table>

Appendix Table 4: Third Degree Dominance Patterns

Strong Dominance

	Aus	Bel	Can	Den	Fin	Fra	Ger	Ita	Net	Nor	Spa	Swe	UK	US	TOTAL
Aus		1	1	0	0	1	1	1	0	0	1	0	0	0	6
Bel	0		1	0	0	0	1	1	0	0	1	0	0	0	4
Can	0	0		0	0	0	0	1	0	0	1	0	0	0	2
Den	1	1	1		1	1	1	1	1	0	1	1	1	1	12
Fin	0	0	1	0		0	0	1	0	0	1	0	0	1	4
Fra	0	0	1	0	0		1	1	0	0	1	0	0	0	4
Ger	0	0	1	0	0	0		1	0	0	1	0	0	0	3
Ita	0	0	0	0	0	0	0		0	0	1	0	0	0	1
Net	1	1	1	0	1	1	1	1		0	1	0	1	1	10
Nor	1	1	1	1	1	1	1	1	1		1	1	1	1	13
Spa	0	0	0	0	0	0	0	0	0	0		0	0	0	0
Swe	1	1	1	0	1	1	1	1	0	0	1		1	1	10
UK	0	1	1	0	0	1	1	1	0	0	1	0		0	6
US	0	0	0	0	0	0	0	1	0	0	1	0	0		2
TOTAL	4	6	10	1	4	6	8	12	2	0	13	2	4	5	77

Weak Dominance

	Aus	Bel	Can	Den	Fin	Fra	Ger	Ita	Net	Nor	Spa	Swe	UK	US	TOTAL
Aus		1	1	0	0	1	1	1	0	0	1	0	1	0	7
Bel	0		1	0	0	1	1	1	0	0	1	0	0	0	5
Can	0	0		0	0	0	0	1	0	0	1	0	0	0	2
Den	1	1	1		1	1	1	1	1	0	1	1	1	1	12
Fin	0	0	1	0		0	0	1	0	0	1	0	0	1	4
Fra	0	1	1	0	0		1	1	0	0	1	0	0	0	5
Ger	0	0	1	0	0	0		1	0	0	1	0	0	0	3
Ita	0	0	0	0	0	0	0		0	0	1	0	0	0	1
Net	1	1	1	0	1	1	1	1		0	1	1	1	1	11
Nor	1	1	1	1	1	1	1	1	1		1	1	1	1	13
Spa	0	0	0	0	0	0	0	0	0	0		0	0	0	0
Swe	1	1	1	0	1	1	1	1	1	0	1		1	1	11
UK	1	1	1	0	0	1	1	1	0	0	1	0		0	7
US	0	0	0	0	0	0	0	1	0	0	1	0	0		2
TOTAL	5	7	10	1	4	7	8	12	3	0	13	3	5	5	83

Appendix Table 5: Third Iteration

<table>
<tr><th></th><th>Net Strong Dominance (With least dominated by)</th><th>Net Strong Dominance (With most dominated)</th><th>Strong Dominance</th><th>Least Strong Dominated</th><th>Net Weak Dominance (With least dominated by)</th><th>Net Weak Dominance (With most dominated)</th><th>Weak Dominance</th><th>Least Weak Dominated</th></tr>
<tr><td>1</td><td>Norway</td><td>Norway</td><td>Norway</td><td>Norway</td><td>Norway</td><td>Norway</td><td>Norway</td><td>Norway</td></tr>
<tr><td>2</td><td>Denmark</td><td>Denmark</td><td>Denmark</td><td>Denmark</td><td>Denmark</td><td>Denmark</td><td>Denmark</td><td>Denmark</td></tr>
<tr><td>3</td><td rowspan="2">Sweden and Netherlands (Tied)</td><td rowspan="2">Sweden and Netherlands (Tied)</td><td rowspan="2">Sweden and Netherlands (Tied)</td><td rowspan="2">Sweden and Netherlands (Tied)</td><td rowspan="2">Sweden and Netherlands (Tied)</td><td rowspan="2">Sweden and Netherlands (Tied)</td><td rowspan="2">Sweden and Netherlands (Tied)</td><td rowspan="2">Sweden and Netherlands (Tied)</td></tr>
<tr><td>4</td></tr>
<tr><td>5</td><td rowspan="2">Australia and UK (Tied)</td><td rowspan="2">Australia and UK (Tied)</td><td rowspan="2">Australia and UK (Tied)</td><td rowspan="2">Australia and UK (Tied)</td><td rowspan="2">Australia and UK (Tied)</td><td rowspan="2">Australia and UK (Tied)</td><td rowspan="2">Australia and UK (Tied)</td><td>Finland</td></tr>
<tr><td>6</td><td rowspan="2">Australia and UK (Tied)</td></tr>
<tr><td>7</td><td>Finland</td><td>Finland</td><td>Finland</td><td>Finland</td><td>Finland</td><td>Finland</td><td rowspan="2">Belgium and France (Tied)</td></tr>
<tr><td>8</td><td rowspan="2">Belgium and France (Tied)</td><td rowspan="2">Belgium and France (Tied)</td><td rowspan="2">Belgium and France (Tied)</td><td>United States</td><td rowspan="2">Belgium and France (Tied)</td><td rowspan="2">Belgium and France (Tied)</td><td>United States</td></tr>
<tr><td>9</td><td rowspan="2">Belgium and France (Tied)</td><td>Finland</td><td rowspan="2">Belgium and France (Tied)</td></tr>
<tr><td>10</td><td>United States</td><td>United States</td><td>Germany</td><td>United States</td><td>United States</td><td>Germany</td></tr>
<tr><td>11</td><td>Germany</td><td>Germany</td><td>United States</td><td>Germany</td><td>Germany</td><td>Germany</td><td>United States</td><td>Germany</td></tr>
<tr><td>12</td><td>Canada</td><td>Canada</td><td>Canada</td><td>Canada</td><td>Canada</td><td>Canada</td><td>Canada</td><td>Canada</td></tr>
<tr><td>13</td><td>Italy</td><td>Italy</td><td>Italy</td><td>Italy</td><td>Italy</td><td>Italy</td><td>Italy</td><td>Italy</td></tr>
<tr><td>14</td><td>Spain</td><td>Spain</td><td>Spain</td><td>Spain</td><td>Spain</td><td>Spain</td><td>Spain</td><td>Spain</td></tr>
</table>

Appendix Table 6: Fourth Degree Dominance Patterns

Strong Dominance

	Aus	Bel	Can	Den	Fin	Fra	Ger	Ita	Net	Nor	Spa	Swe	UK	US	TOTAL
Aus		1	1	0	0	1	1	1	0	0	1	0	0	1	7
Bel	0		1	0	0	0	1	1	0	0	1	0	0	0	4
Can	0	0		0	0	0	0	1	0	0	1	0	0	0	2
Den	1	1	1		1	1	1	1	1	0	1	1	1	1	12
Fin	0	0	1	0		0	1	1	0	0	1	0	0	1	5
Fra	0	0	1	0	0		1	1	0	0	1	0	0	0	4
Ger	0	0	1	0	0	0		1	0	0	1	0	0	0	3
Ita	0	0	0	0	0	0	0		0	0	1	0	0	0	1
Net	1	1	1	0	1	1	1	1		0	1	0	1	1	10
Nor	1	1	1	1	1	1	1	1	1		1	1	1	1	13
Spa	0	0	0	0	0	0	0	0	0	0		0	0	0	0
Swe	1	1	1	0	1	1	1	1	0	0	1		1	1	10
UK	0	1	1	0	0	1	1	1	0	0	1	0		1	7
US	0	0	1	0	0	0	0	1	0	0	1	0	0		3
TOTAL	4	6	11	1	4	6	9	12	2	0	13	2	4	7	81

Weak Dominance

	Aus	Bel	Can	Den	Fin	Fra	Ger	Ita	Net	Nor	Spa	Swe	UK	US	TOTAL
Aus		1	1	0	0	1	1	1	0	0	1	0	1	1	8
Bel	0		1	0	0	1	1	1	0	0	1	0	0	0	5
Can	0	0		0	0	0	0	1	0	0	1	0	0	0	2
Den	1	1	1		1	1	1	1	1	0	1	1	1	1	12
Fin	0	0	1	0		0	1	1	0	0	1	0	0	1	5
Fra	0	1	1	0	0		1	1	0	0	1	0	0	0	5
Ger	0	0	1	0	0	0		1	0	0	1	0	0	0	3
Ita	0	0	0	0	0	0	0		0	0	1	0	0	0	1
Net	1	1	1	0	1	1	1	1		0	1	1	1	1	11
Nor	1	1	1	1	1	1	1	1	1		1	1	1	1	13
Spa	0	0	0	0	0	0	0	0	0	0		0	0	0	0
Swe	1	1	1	0	1	1	1	1	1	0	1		1	1	11
UK	1	1	1	0	0	1	1	1	0	0	1	0		1	8
US	0	0	1	0	0	0	0	1	0	0	1	0	0		3
TOTAL	5	7	11	1	4	7	9	12	3	0	13	3	5	7	87

Appendix Table 7: Fourth Iteration

	Net Strong Dominance (With least dominated by)	Net Strong Dominance (With most dominated)	Strong Dominance	Least Strong Dominated	Net Weak Dominance (With least dominated by)	Net Weak Dominance (With most dominated)	Weak Dominance	Least Weak Dominated
1	Norway	Norway	Norway	Norway	Norway	Norway	Norway	Norway
2	Denmark	Denmark	Denmark	Denmark	Denmark	Denmark	Denmark	Denmark
3	Sweden and Netherlands (Tied)	Sweden and Netherlands (Tied)	Sweden and Netherlands (Tied)	Sweden and Netherlands (Tied)	Sweden and Netherlands (Tied)	Sweden and Netherlands (Tied)	Sweden and Netherlands (Tied)	Sweden and Netherlands (Tied)
4								
5	Australia and UK (Tied)	Australia and UK (Tied)	Australia and UK (Tied)	Australia and UK (Tied)	Australia and UK (Tied)	Australia and UK (Tied)	Australia and UK (Tied)	Finland
6								Australia and UK (Tied)
7	Finland	Finland	Finland	Finland	Finland	Finland	Finland	
8	Belgium and France (Tied)	Belgium and France (Tied)	Belgium and France (Tied)	Belgium and France (Tied)	Belgium and France (Tied)	Belgium and France (Tied)	Belgium and France (Tied)	Belgium and France (Tied)
9								
10	United States	United States	United States	United States	United States	United States	United States	United States
11	Germany	Germany	Germany	Germany	Germany	Germany	Germany	Germany
12	Canada	Canada	Canada	Canada	Canada	Canada	Canada	Canada
13	Italy	Italy	Italy	Italy	Italy	Italy	Italy	Italy
14	Spain	Spain	Spain	Spain	Spain	Spain	Spain	Spain

Appendix Table 8: Fifth Degree Dominance Patterns

Strong Dominance

	Aus	Bel	Can	Den	Fin	Fra	Ger	Ita	Net	Nor	Spa	Swe	UK	US	TOTAL
Aus		1	1	0	0	1	1	1	0	0	1	0	0	1	7
Bel	0		1	0	0	0	1	1	0	0	1	0	0	1	5
Can	0	0		0	0	0	0	1	0	0	1	0	0	0	2
Den	1	1	1		1	1	1	1	1	0	1	1	1	1	12
Fin	0	1	1	0		1	1	1	0	0	1	0	0	1	7
Fra	0	0	1	0	0		1	1	0	0	1	0	0	1	5
Ger	0	0	1	0	0	0		1	0	0	1	0	0	0	3
Ita	0	0	0	0	0	0	0		0	0	1	0	0	0	1
Net	1	1	1	0	1	1	1	1		0	1	0	1	1	10
Nor	1	1	1	1	1	1	1	1	1		1	1	1	1	13
Spa	0	0	0	0	0	0	0	0	0	0		0	0	0	0
Swe	1	1	1	0	1	1	1	1	0	0	1		1	1	10
UK	0	1	1	0	0	1	1	1	0	0	1	0		1	7
US	0	0	1	0	0	0	1	1	0	0	1	0	0		4
TOTAL	4	7	11	1	4	7	10	12	2	0	13	2	4	9	86

Weak Dominance

	Aus	Bel	Can	Den	Fin	Fra	Ger	Ita	Net	Nor	Spa	Swe	UK	US	TOTAL
Aus		1	1	0	0	1	1	1	0	0	1	0	1	1	8
Bel	0		1	0	0	1	1	1	0	0	1	0	0	1	6
Can	0	0		0	0	0	0	1	0	0	1	0	0	0	2
Den	1	1	1		1	1	1	1	1	0	1	1	1	1	12
Fin	0	1	1	0		1	1	1	0	0	1	0	0	1	7
Fra	0	1	1	0	0		1	1	0	0	1	0	0	1	6
Ger	0	0	1	0	0	0		1	0	0	1	0	0	0	3
Ita	0	0	0	0	0	0	0		0	0	1	0	0	0	1
Net	1	1	1	0	1	1	1	1		0	1	1	1	1	11
Nor	1	1	1	1	1	1	1	1	1		1	1	1	1	13
Spa	0	0	0	0	0	0	0	0	0	0		0	0	0	0
Swe	1	1	1	0	1	1	1	1	1	0	1		1	1	11
UK	1	1	1	0	0	1	1	1	0	0	1	0		1	8
US	0	0	1	0	0	0	1	1	0	0	1	0	0		4
TOTAL	5	8	11	1	4	8	10	12	3	0	13	3	5	9	92

Appendix Table 9: Fifth Iteration

<table>
<tr><th></th><th>Net Strong Dominance (With least dominated by)</th><th>Net Strong Dominance (With most dominated)</th><th>Strong Dominance</th><th>Least Strong Dominated</th><th>Net Weak Dominance (With least dominated by)</th><th>Net Weak Dominance (With most dominated)</th><th>Weak Dominance</th><th>Least Weak Dominated</th></tr>
<tr><td>1</td><td>Norway</td><td>Norway</td><td>Norway</td><td>Norway</td><td>Norway</td><td>Norway</td><td>Norway</td><td>Norway</td></tr>
<tr><td>2</td><td>Denmark</td><td>Denmark</td><td>Denmark</td><td>Denmark</td><td>Denmark</td><td>Denmark</td><td>Denmark</td><td>Denmark</td></tr>
<tr><td>3</td><td rowspan="2">Sweden and Netherlands (Tied)</td><td rowspan="2">Sweden and Netherlands (Tied)</td><td rowspan="2">Sweden and Netherlands (Tied)</td><td rowspan="2">Sweden and Netherlands (Tied)</td><td rowspan="2">Sweden and Netherlands (Tied)</td><td rowspan="2">Sweden and Netherlands (Tied)</td><td rowspan="2">Sweden and Netherlands (Tied)</td><td rowspan="2">Sweden and Netherlands (Tied)</td></tr>
<tr><td>4</td></tr>
<tr><td>5</td><td rowspan="3">Australia, UK, and Finland (Tied)</td><td rowspan="3">Australia, UK, and Finland (Tied)</td><td rowspan="3">Australia, UK, and Finland (Tied)</td><td rowspan="3">Australia, UK, and Finland (Tied)</td><td>Finland</td><td rowspan="2">Australia and UK (Tied)</td><td rowspan="2">Australia and UK (Tied)</td><td>Finland</td></tr>
<tr><td>6</td><td rowspan="2">Australia and UK (Tied)</td><td rowspan="2">Australia and UK (Tied)</td></tr>
<tr><td>7</td><td>Finland</td><td>Finland</td></tr>
<tr><td>8</td><td rowspan="2">Belgium and France (Tied)</td><td rowspan="2">Belgium and France (Tied)</td><td rowspan="2">Belgium and France (Tied)</td><td rowspan="2">Belgium and France (Tied)</td><td rowspan="2">Belgium and France (Tied)</td><td rowspan="2">Belgium and France (Tied)</td><td rowspan="2">Belgium and France (Tied)</td><td rowspan="2">Belgium and France (Tied)</td></tr>
<tr><td>9</td></tr>
<tr><td>10</td><td>United States</td><td>United States</td><td>United States</td><td>United States</td><td>United States</td><td>United States</td><td>United States</td><td>United States</td></tr>
<tr><td>11</td><td>Germany</td><td>Germany</td><td>Germany</td><td>Germany</td><td>Germany</td><td>Germany</td><td>Germany</td><td>Germany</td></tr>
<tr><td>12</td><td>Canada</td><td>Canada</td><td>Canada</td><td>Canada</td><td>Canada</td><td>Canada</td><td>Canada</td><td>Canada</td></tr>
<tr><td>13</td><td>Italy</td><td>Italy</td><td>Italy</td><td>Italy</td><td>Italy</td><td>Italy</td><td>Italy</td><td>Italy</td></tr>
<tr><td>14</td><td>Spain</td><td>Spain</td><td>Spain</td><td>Spain</td><td>Spain</td><td>Spain</td><td>Spain</td><td>Spain</td></tr>
</table>

Appendix Table 10: Sixth Degree Dominance Patterns

Strong Dominance

	Aus	Bel	Can	Den	Fin	Fra	Ger	Ita	Net	Nor	Spa	Swe	UK	US	TOTAL
Aus		1	1	0	0	1	1	1	0	0	1	0	0	1	7
Bel	0		1	0	0	0	1	1	0	0	1	0	0	1	5
Can	0	0		0	0	0	0	1	0	0	1	0	0	0	2
Den	1	1	1		1	1	1	1	1	0	1	1	1	1	12
Fin	0	1	1	0		1	1	1	0	0	1	0	0	1	7
Fra	0	0	1	0	0		1	1	0	0	1	0	0	1	5
Ger	0	0	1	0	0	0		1	0	0	1	0	0	0	3
Ita	0	0	0	0	0	0	0		0	0	1	0	0	0	1
Net	1	1	1	0	1	1	1	1		0	1	0	1	1	10
Nor	1	1	1	1	1	1	1	1	1		1	1	1	1	13
Spa	0	0	0	0	0	0	0	0	0	0		0	0	0	0
Swe	1	1	1	0	1	1	1	1	0	0	1		1	1	10
UK	0	1	1	0	0	1	1	1	0	0	1	0		1	7
US	0	0	1	0	0	0	1	1	0	0	1	0	0		4
TOTAL	4	7	11	1	4	7	10	12	2	0	13	2	4	9	86

Weak Dominance

	Aus	Bel	Can	Den	Fin	Fra	Ger	Ita	Net	Nor	Spa	Swe	UK	US	TOTAL
Aus		1	1	0	0	1	1	1	0	0	1	0	0	1	7
Bel	0		1	0	0	1	1	1	0	0	1	0	0	1	6
Can	0	0		0	0	0	0	1	0	0	1	0	0	0	2
Den	1	1	1		1	1	1	1	1	0	1	1	1	1	12
Fin	0	1	1	0		1	1	1	0	0	1	0	0	1	7
Fra	0	1	1	0	0		1	1	0	0	1	0	0	1	6
Ger	0	0	1	0	0	0		1	0	0	1	0	0	0	3
Ita	0	0	0	0	0	0	0		0	0	1	0	0	0	1
Net	1	1	1	0	1	1	1	1		0	1	1	1	1	11
Nor	1	1	1	1	1	1	1	1	1		1	1	1	1	13
Spa	0	0	0	0	0	0	0	0	0	0		0	0	0	0
Swe	1	1	1	0	1	1	1	1	1	0	1		1	1	11
UK	0	1	1	0	0	1	1	1	0	0	1	0		1	7
US	0	0	1	0	0	0	1	1	0	0	1	0	0		4

Upon the sixth iteration we reach the equilibrium:

Appendix Table 11: Sixth Iteration

<table>
<tr><td>1</td><td>Norway</td></tr>
<tr><td>2</td><td>Denmark</td></tr>
<tr><td>3</td><td rowspan="2">Sweden and Netherlands (Tied)</td></tr>
<tr><td>4</td></tr>
<tr><td>5</td><td rowspan="3">Australia, UK, and Finland (Tied)</td></tr>
<tr><td>6</td></tr>
<tr><td>7</td></tr>
<tr><td>8</td><td rowspan="2">Belgium and France (Tied)</td></tr>
<tr><td>9</td></tr>
<tr><td>10</td><td>United States</td></tr>
<tr><td>11</td><td>Germany</td></tr>
<tr><td>12</td><td>Canada</td></tr>
<tr><td>13</td><td>Italy</td></tr>
<tr><td>14</td><td>Spain</td></tr>
</table>

This is therefore what we will consider *the most robust* dominance rankings possible.

APPENDIX C – Comparison of Weighting Methods by Preferred Qualities

	Understandable Procedures (Transparency)	***Objective Determination of Societal Valuations***	***Comparable Index Values Across Regions and Time***
Constrained DEA	Constrained DEA is somewhat less transparent than explicit weighting. Although there is an identifiable process for the allocation of weights, this process is statistically involved. Constrained DEA, therefore, is not as transparent as other methods.	Constrained DEA is slightly less objective than DEA in its determination of weights. Although the base methodology remains unchanged, the constraints on each weight are determined by experts. This is therefore not completely objective.	Constrained DEA is not completely comparable across regions and time. The weights vary across countries and time in order to measure the 'efficiency' of each performance. As such, comparisons across countries and time will not be based upon a single set of weights.
Alternative 1	Alternative 1 is an explicit weighting scheme. The process for determining weights is easily communicated.	Alternative 1 is an explicit weighting scheme based on expert weighting. This therefore does not satisfy this quality.	The explicit weights allow for comparisons across both countries and time.
DEA	DEA is somewhat less transparent than explicit weighting. Although there is an identifiable process for the allocation of weights, this process is statistically involved. DEA, therefore, is not as transparent as other methods.	Data envelopment analysis is objective in its choice of weights. No preference is chosen between the different components being considered. Optimal performance in a given domain is considered efficient. Other countries are given weights that optimize their relative performance.	DEA is not completely comparable across regions and time. The weights vary across countries and time in order to measure the 'efficiency' of each performance. As such, comparisons across countries and time will not be based upon a single set of weights.
Alternative 3	Alternative 3 is an explicit weighting scheme. The process for determining weights is easily communicated.	Alternative 3 is an explicit weighting scheme based on expert weighting. This therefore does not satisfy this quality.	The explicit weights allow for comparisons across both countries and time.

	Understandable Procedures (Transparency)	*Objective Determination of Societal Valuations*	*Comparable Index Values Across Regions and Time*
Common (Avg.) Weights	DEA with common weights based on the average optimal weights is less transparent than explicit weighting or constrained DEA and DEA. Although there is an identifiable process for the allocation of weights, this process is statistically involved. The process is then lengthened to produce common weights. Common (avg.) weights DEA, therefore, is not as transparent as other methods.	The common weights based on average optimal weights solution is objective in its determination of weights. Indeed, this method employs DEA and then simply averages the obtained optimal weights to create a set of common weights to be applied to each country.	The common weights based on the average optimal weights solution allows for comparisons across countries on the same weights; however, weights will change from year to year, which will lower comparability across time.
Factor Analysis	Factor analysis the least transparent among considered options. The process is statistically heavy, involving many steps, and is therefore not easily communicated to the public.	Factor analysis objectively determines the weights used in an indicator. Unfortunately, these weights are not generated from optimal valuations yet rather patterns of variance within the data. Therefore, factor analysis has a bias against possibly important variables that vary little.	The weights applied to all countries are the same and thus countries are comparable in a given year. Unfortunately, the base data in the factor analysis will change from year to year, which will make comparisons across time difficult.
Equal Weights	Equal weighting is an explicit weighting scheme. The process for determining weights is easily communicated.	Equal weighting is an explicit weighting scheme. This therefore does not satisfy this quality; however, Hagerty and Land (2007) indicate that is objectively the best option where no other options are optimal.	The explicit weights allow for comparisons across both countries and time.

	Understandable Procedures (Transparency)	***Objective Determination of Societal Valuations***	***Comparable Index Values Across Regions and Time***
Compromise Solution	DEA with the compromise solution weights is less transparent than explicit weighting or constrained DEA and DEA. Although there is an identifiable process for the allocation of weights, this process is statistically involved. The process is then lengthened to produce common weights. The compromise solution, therefore, is not as transparent as other methods.	The compromise solution is objective in its determination of weights. Indeed, this method employs DEA and then determines the weights closest to the set of optimal weights generated by DEA. This creates common weights to be applied to each country.	The compromise solution allows for comparisons across countries on the same weights; however, compromise weights will change from year to year, which will lower comparability across time.
Alternative 2	Alternative 2 is an explicit weighting scheme. The process for determining weights is easily communicated.	Alternative 2 is an explicit weighting scheme based on expert weighting. This therefore does not satisfy this quality.	The explicit weights allow for comparisons across both countries and time.
User-Weighting	The user weights the variables and is therefore familiar with the weighting process.	User-weighting objectively determines the optimal weights *for a particular user*.	User-weighting allows for comparisons across countries and time; however, these comparisons will be different for each user.
Survey Weighting	Survey weights are easily communicated to the public as they are simply derived from a public survey.	Survey weighting, when collected from a representative sample, objectively determines the relative societal valuations of each component.	The weights derived from survey weighting will be explicit. These weights will therefore allow for comparisons across both and time.

	Understandable Procedures (Transparency)	***Objective Determination of Societal Valuations***	***Comparable Index Values Across Regions and Time***
Regression Analysis	Regression analysis is slightly more statistically involved than other methods of weighting and becomes increasingly complex as more variables are added and functional forms are adapted. Regression analysis is therefore less transparent than other methods considered.	Regression analysis does not objectively determine the weights for a composite indicator. Although the weights are endogenously determined, the optimal dependent variable must be determined. Changing this variable will change the results. As such, regression analysis is not completely objective.	The data for the IEWB and the data for any potential dependent variable will change year by year. As such, although comparisons across countries may be possible, the weights will change each year, thereby making limiting comparability across time.
Consumption-Equivalent Measure	The overall nature of the consumption-equivalent measure is simple to explain: each country is ranked according to the value each component has equivalent to consumption. Unfortunately, the details of each conversion are slightly more complicated and require the discussion of economic theory, among other considerations. This is therefore less transparent than other methods.	The consumption-equivalent measure is partially objective in the determination of weights. Weights are determined by a utility function that describes the relationship of each component to consumption; however, these are dependent upon coefficients that the researchers must often base on theory (where household micro-data is not available). As such, this method is not completely objective.	Assuming that the correct utility function has been determined, this function should not change year to year. As such, this function will produce comparable values across all countries and across time.

References

Canadian Council on Learning. "The Methodology of the Composite Learning Index." 2010. 28 July 2011. Available: http://www.cli-ica.ca/en/about/about-cli/methodology.aspx

Charnes, A., W.W. Cooper, and E. Rhodes (1978), Measuring efficiency of decision making units. *European Journal of Operational Research* 2, 429-444.

Cherchye, L., W. Moesen, N. Rogge, T. Van Puyenbroeck, M. Saisana, A. Saltelli, R. Liska, and S. Tarantola (2008a), Creating composite indicators with DEA and robustness analysis: the case of the Technology Achievement Index, *Journal of the Operational Research Society* 59, 239-251.

Cherchye, L., W. Moesen and T. Van Puyenbroeck (2004), Legitimately diverse, yet comparable: on synthesizing social inclusion performance in the EU, *Journal of Common Market Studies* 42, 919-955.

Cherchye, L., E. Ooghe, and T. Van Puyenbroeck (2008b), Robust human development rankings, *The Journal of Economic Inequality* 6, No. 4, 287-321.

Decancq, Koen and Lugo, Maria Ana (2008), Setting weights in multidimensional indices of well-being and deprivation, OPHI Working Paper 18, December.

Doyle, J. (1995) Multiattribute choice for the lazy decision maker: let the alternatives decide. *Organizational Behavior and Human Processes* 62, 87-100.

Hagerty, Michael R. and Kenneth C. Land (2007), Constructing Summary Indices of Quality of Life: A Model for the Effect of Heterogeneous Importance Weights, *Sociological Methods Research* 35, 455-496.

Jones, Charles and Peter Klenow (2010) Beyond GDP? Welfare Across Countries and Time, NBER Working Paper 16352, September.

Kao, C. and H-T. Hung (2005), Data envelopment analysis with common weights: the compromise solution approach, *Journal of the Operational Research Society* 56, 1196-1203.

Maggino, Filomena and Bruno Zumbo (2011), Measuring the Quality of Life and the Construction of Social Indicators. To be published as a chapter in Land, Kenneth C. (ed.), *Handbook of Social Indicators and Quality-of-Life Research.* New York: Springer, 2011, forthcoming.

Makui, A., A. Alinezhad, R. Kiani Mavi, and M. Zohrehbandian (2008), A Goal Programming Method for Finding Common Weights in DEA with an Improved Discriminating Power for Efficiency, *Journal of Industrial and Systems Engineering* 1, No. 4, 293-303.

Melyn W. and W. Moesen (1991), Towards a synthetic indicator of macroeconomic performance: unequal weighting when limited information is available, Public Economic Research Paper 17, CES, KU Leuven.

Nicoletti G., Scarpetta S. and Boylaud O. (2000), Summary indicators of product market regulation with an extension to employment protection legislation, OECD, Economics department working papers, No. 226, ECO/WKP(99)18. Available: http://www.oecd.org/dataoecd/21/13/1880867.pdf

OECD. "What's Your Better Life Index?" n.d., 5 Aug 2011. Available at: http://www.oecdbetterlifeindex.org/about/better-life-initiative/

OECD. "Handbook on Constructing Composite Indicators: Methodology and User Guide," OECD: 2008. Available: http://www.oecd.org/dataoecd/37/42/42495745.pdf

Osberg, Lars and Andrew Sharpe (2002) An Index of Economic Well-being for Selected OECD Countries, *Review of Income and Wealth*, Vol. 48, No. 1, September, pp. 291-316.

Osberg, Lars and Andrew Sharpe (2005) How Should We Measure the 'Economic' Aspects of Well-Being?, *Review of Income and Wealth*, Vol. 51, No. 2, June, pp. 311-336.

Osberg, Lars and Andrew Sharpe (2009a) New Estimates of the Index of Economic Well-being for Selected OECD Countries, 1980-2007, CSLS Research Report No. 2009-11, December. Available: http://www.csls.ca/iwb/oecd.asp

Osberg, Lars and Andrew Sharpe (2009b) New Estimates of the Index of Economic Well-being for Canada and the Provinces, 1981-2008, CSLS Research Report No. 2009-10, December.

Osberg, Lars and Andrew Sharpe (2011) Moving from a GDP-based to a Well-being Based Metric of Economic Performance and Social Progress: Results of the Index of Economic Well-Being for OECD Countries, 1980-2009, Paper presented to the International Statistical Institute Conference, August 21-26, Dublin, Ireland.

Rummel, R. J. (1967), Understanding factor analysis, *Journal of Conflict Resolution* 11, No. 4, 444-480.

Saisana, Michaela. 2007 Composite Learning Index: Robustness Issue and Critical Assessment. European Communities, 2008. Available: http://publications.jrc.ec.europa.eu/repository/bitstream/111111111/4973/1/eur23274_fulldocument.pdf

Saisana, Michaela, Béatrice d'Hombres, and Andrea Saltelli (2011). Rickety numbers: Volatility of university rankings and policy implications, *Research Policy* 40, 165-177.

Salzman, Julia and Andrew Sharpe (2003), Methodological Choices Encountered in the Construction of Composite Indices of Economic and Social Well-Being, Paper presented to the Canadian Economics Association conference, 31 May 2003. Available: http://www.csls.ca/events/cea2003/salzman-typol-cea2003.pdf