Technical Annex for the 2006-2016 daset

The Calculation of the GEI Scores 2006-2016

The GEI scores for all countries are calculated according to the following nine points. Note that we calculate the GEI scores for all 102 countries in the 2006-2016 time period, which results in 619 observations.

1. ***The selection of indicators:*** We start with indicators that come directly from the original sources for each country involved in the analysis. The variables can be at the individual level (personal or business) and come from the Global Entrepreneurship Monitor (GEM) Adult Population Survey, or at the institutional/environmental level and come from various other sources. Altogether we use 16 individual and 20 institutional indicators, some of which are themselves complex. Individual data are calculated from 2006 to 2016, using the two-year moving average principle. Where the proper data are lacking, single-year value is applied. For the institutional variables we applied single-year data. This resulted in 619total observations for 102 countries.

2. **The calculation of variables**: Most cases the indicators are also used as variables. However, two cases in the individual variables and eight cases in the institutional variables contain two or three indicators. The calculation of these variables is the following:

$$Carrier\ Status_{i,l} = Carrier_{i,l} * Status_{i,l} \quad (1a)$$

for all l= 1, 2 the number individual indicators
i= 1 ... k, the number of countries

$$Informal\ Investment\ Mean_{i,l} = Business\ Angel_{i,l} * Informal\ Investment_{i,l} \quad (1b)$$

for all l= 1, 2 the number individual indicators
i= 1 ... k, the number of countries

$$Freedom\ and\ Property_{i,l} = Economic\ Freedom_{i,l} * Property\ Rights_{i,l} \quad (1c)$$

for all l= 1, 2 the number individual indicators
i= 1 ... k, the number of countries

$$Education_{i,l} = Tertiary\ Education_{i,l} * Quality\ of\ Education_{i,l} \quad (1d)$$

for all l= 1, 2 the number individual indicators
i= 1 ... k, the number of countries

$$Connectivity_{i,l} = Urbanization_{i,l} * Infrastructure_{i,l} \quad (1e)$$

for all l= 1, 2 the number individual indicators
i= 1 ... k, the number of countries

$$Taxgovern_{i,l} = Taxation_{i,l} * Good\ Governance_{i,l} \quad (1f)$$

for all l= 1, 2 the number individual indicators
i= 1 ... k, the number of countries

$$Labor\ Marekt_{i,l} = Labor\ Freedom_{i,l} * Staff\ Training_{i,l} \quad (1g)$$

for all l= 1, 2 the number individual indicators
i= 1 … k, the number of countries

$$Compregulation_{i,l} = Regulation_{i,l} * Market\ Dominance_{i,l} \quad (1h)$$

for all l= 1, 2 the number individual indicators
i= 1 … k, the number of countries

$$Science_{i,l} = Gerd_{i,l} * Scientific\ Institutions_{i,l} * Availability\ of\ Scientists_{i,l} \quad (1i)$$

for all l= 1, 2, 3 the number individual indicators
i= 1 … k, the number of countries

$$Finance\ and\ Strategy_{i,l} = Venture\ Capital_{i,l} * Busienss\ Strategy_{i,l} \quad (1j)$$

for all l= 1, 2 the number individual indicators
i= 1 … k, the number of countries

3. ***The construction of the pillars:*** We calculate all pillars from the variables using the interaction variable method; that is, by multiplying the individual variable with the proper institutional variable. The notion behind this technique goes back to Baumol's (1996) idea that the value of entrepreneurship depends on both the individual effort and the institutional context. This multiplication result for all the 674 observations is.

$$z_{i,j} = IND_{i,j} * INS_{i,j} \quad (2)$$

for all j = 1 ... k, the number of individual and institutional variables
$IND_{i,j}$ is the original score value for country i and individual variable j
$INS_{i,j}$ is the original score value for country i and institutional variable j
$z_{i,j}$ is the original pillar value for country i and pillar j

4. ***Normalization:*** pillar values were first normalized to a range from 0 to 1, using the distance method, as shown in equation 2:

$$x_{i,j} = \frac{z_{i,j}}{\max z_{i,j}} \quad (3)$$

for all j = 1 ... k, the number of pillars
where $x_{i,j}$ is the normalized score value for country i and pillar j
$z_{i,j}$ is the original pillar value for country i and pillar j
$max\ z_{i,j}$ is the maximum value for pillar j

This normalization technique compares a given country's performance to the best performing country. Hence, it provides a proper benchmark for evaluating the performance of a particular country in a certain pillar according to the best available practice. The disadvantage of the "min-max" methodology is that it assigns value one to the best country and zero to the worst one, which could exaggerate small differences. Another popular normalization method, the "z" score approach, guarantees normal distribution but results in a variable range of scores. Since we require that country scores be strictly in the [0,1] range, we cannot use this approach.

5. ***Capping:*** All index-building is based on a benchmarking principle. We selected the 95th percentile score adjustment, meaning that any observed values higher than the 95th percentile are lowered to the 95th percentile. For the 102 countries, we used the benchmark values from the full dataset, which contains all the countries and all the years, which results in 619 observations during the 2006-2016 time period. This capping method has two advantages. First, it makes it possible to get rid of the outliers. Second, it provides a reasonable and reachable benchmark for the other countries. Without capping the best country benchmark value, a certain pillar could be extremely high, which would result in unreasonably lower normalized scores for the other countries.

6. ***Average pillar adjustment:*** The different averages of the normalized values of the indicators imply that reaching the same indicator values requires a different effort and resources. Since we want to apply GEI for public policy purposes, the additional resources to achieve the same marginal improvement of the pillar values should be the same for all pillars. However, the marginal effects could differ, depending on the level of the pillar values. Country variations in the marginal effects are also possible. Calculating all the marginal effects for all the countries would be a cumbersome task, so we suggest a simpler solution: to equalize the marginal effects of the components only on the average pillar values of all countries. This technique reduces but does not eliminate the distortion in calculating the marginal effects. Equation 3 shows the calculation of the average value of pillar j:

$$\bar{\mathrm{x}}_{\mathrm{j}} = \frac{\sum_{\mathrm{i}=1}^{\mathrm{n}} \mathrm{x}_{\mathrm{i,j}}}{\mathrm{n}} \text{ for all j} \qquad (4)$$

We want to transform the $x_{i,j}$ values such that the potential values to be in the [0,1] range.

$$y_{i,j} = x_{i,j}^{k} \qquad (5)$$

where k is the "strength of adjustment", the k-th moment of x_j is exactly the needed average, $\bar{y}_j$

We have to find the root of the following equation for k:

$$\sum_{i=1}^{n} x_{i,j}^{k} - n\bar{y}_j = 0 \qquad (6)$$

It is easy to see based on previous conditions and derivatives that the function is decreasing and convex which means it can be quickly solved using the well-known Newton-Raphson method with an initial guess of 0. After obtaining k, the computations are straightforward. Note that if

$$\bar{x}_j < \bar{y}_j \qquad k < 1$$

$$\bar{x}_j = \bar{y}_j \qquad k = 1$$

$$\bar{x}_j > \bar{y}_j \qquad k > 1$$

that is k be thought of as the strength (and direction) of adjustment.

The average marginal rate of compensation (AMRC) for any two average pillars *i* and *j* is the same:

$$AMRC_{i,j} = \frac{d\bar{y}_i}{d\bar{y}_j} \qquad (7)$$

The adjusted pillar values are calculated for the entire 2006-2016 time period, including all 619 observations. Table 2 shows the average pillar values before and after the equalization procedure:

Table 2: GEI Pillar Averages before and after Average Pillar Adjustment

Pillar Name	Average Pillar Values before Equalization	Average Pillar Values after Equalization
Opportunity Perception	0,40	0,47
Startup Skills	0,49	0,47
Risk Acceptance	0,62	0,47
Networking	0,52	0,47
Cultural Support	0,60	0,47
Opportunity Startup	0,49	0,47
Technology Absorption	0,41	0,47
Human Capital	0,40	0,47
Competition	0,52	0,47
Product Innovation	0,59	0,47
Process Innovation	0,33	0,47
High Growth	0,40	0,47
Internationalization	0,49	0,47
Risk Capital	0,29	0,47
Average	0,47	0,47

Source: Author calculations

7. ***Penalizing:*** After these transformations, the penalty for bottleneck methodology was used to create pillar-adjusted PFB values. A bottleneck is defined as the worst performing link or a binding constraint in a particular country's system of entrepreneurship. Here, bottleneck means a shortage or the lowest level of a particular pillar, relative to other pillars. This notion of a bottleneck is important for policy purposes. Our model suggests that pillars interact; if they are out of balance, entrepreneurship is inhibited. The pillar values should be adjusted in a way that takes into account this notion of balance. After equalizing the pillar averages, the value of each pillar of a country is penalized by linking it to the score of the pillar with the weakest scores in that country. This simulates the notion of a bottleneck; if the weakest pillar were improved, the whole GEI would show a significant improvement.

We define our penalty function as:

$$h_{(i),j} = min\, y_{(i),j} + \mathrm{a}(1 - e^{-\mathrm{b}(y_{(i)j} - min\, y_{(i),j})}) \qquad (8)$$

where $h_{i,j}$ is the modified, post-penalty value of pillar j in country i
$y_{i,j}$ is the normalized value of index component j in country i
y_{min} is the lowest value of $y_{i,j}$ for country i
i = 1, 2,……n = the number of countries
j = 1, 2,…….m = the number of pillars
0 ≤a, b ≤ 1 are the penalty parameters, the basic setup is a = b = 1

The penalty function also reflects compensation for the loss of one pillar with a gain in another pillar. Let us define the Marginal Rate of Compensation (MRC) as follows:

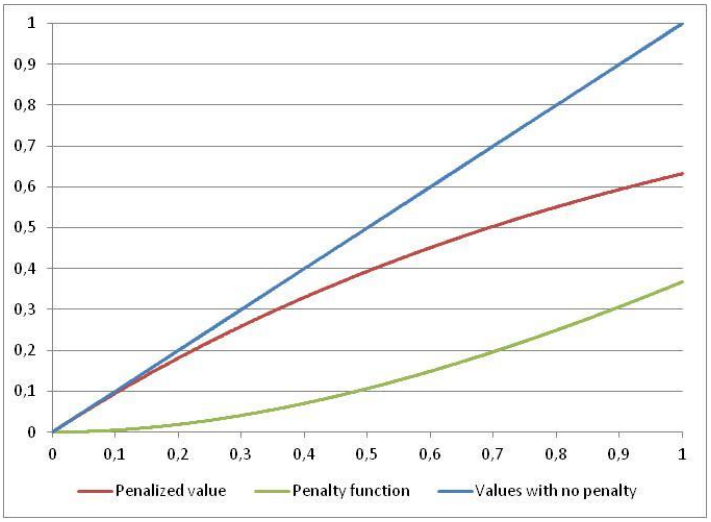
$$MRC_{i,j} = \frac{dy_i}{dy_j} \qquad (9)$$

Full compensability means that a loss in one pillar can be compensated for by the same increase in another pillar. However, this is not realistic. The MRC is the same concept as the Marginal Rate of Substitution for goods and the Marginal Rate of Technical Substitution for inputs (Casadio-Tarabusi and Palazzi 2012), which are reflected in the law of diminishing return. Therefore, the effect of the change of the penalty is not proportional, which means that higher compensation is necessary for the loss in one pillar if the difference between another pillar value and the particular pillar is higher than the lower differences between the pillars. The required positive value of the second derivative means that the pillars are only partially compensable with each other, meaning that the penalty increases at an increasing rate:

$$\frac{dMRC_{i,j}}{dy_j} > 0 \qquad (10)$$

Figure 1 pictures the size of the penalty when the minimum pillar value is 0.

Figure 1: Penalty Function, Penalized Values, and Pillar Values with No Penalty (y_{min} = 0; a, b = 1)



Note that we do not know the size of the penalty. To the best of our knowledge, no objective method exists to define the penalty function exactly. According to Figure 1, the maximum penalty is 0,368, which is about a one-third loss of the original value, which looks reasonable. Larger penalty values rearrange the ranking of the countries considerably. As a result, the average decrease in the GEI points is 9,8%, from 47,1 to 42,5.

Table 3 shows the change of the pillar scores and the size of penalty using the 2015 scores of one country, Ireland, as an example.

Table 3: Average Adjusted Pillar Scores, Penalized Scores, and Penalty in the Case of Ireland, 2015 Values

Pillar Name	Average Adjusted Scores	Penalized Scores	Penalty Scores	Penalty %
Opportunity Perception	0,664	0,630	0,034	5,1%
Startup Skills	0,904	0,792	0,112	12,4%
Risk Acceptance	0,738	0,684	0,054	7,3%
Networking	0,391	0,391	0,000	0,0%
Cultural Support	0,743	0,688	0,055	7,5%
Opportunity Startup	0,907	0,794	0,113	12,4%
Technology Absorption	0,801	0,728	0,074	9,2%
Human Capital	0,926	0,805	0,121	13,0%
Competition	0,920	0,802	0,118	12,9%
Product Innovation	0,842	0,754	0,088	10,4%
Process Innovation	0,756	0,697	0,059	7,8%
High Growth	0,833	0,748	0,085	10,2%
Internationalization	0,827	0,745	0,083	10,0%
Risk Capital	0,630	0,604	0,026	4,2%
Average	0,777	0,704	0,073	8,7%

Note: Red indicates the bottleneck pillar

Ireland's bottleneck pillar is Networking, with a 0,391 average adjusted pillar score. The bottleneck pillar is not penalized. The second lowest pillar is Opportunity Perception, with a 0,664 score. From the system perspective, Ireland cannot fully capitalize on the higher Opportunity Perception performance because the bottleneck pillar is holding it back. The size of the penalty is 0,034, around 5,1%. Ireland's best pillar is Human Capital (0,926). Since the difference between the bottleneck Networking pillar and Human Capital is larger than between Networking and Opportunity Perception, the size of penalty is higher in both absolute value and proportionally, resulting in a 0,121 (13%) penalty. All other penalties are between these two extremes. On average, Ireland is losing 8,7% of its entrepreneurial resources because of the imbalances in its system of entrepreneurship.

8. **Sub-index calculation** The pillars are the basic building blocks of the sub-index: entrepreneurial attitudes, entrepreneurial abilities, and entrepreneurial aspirations.

The value of a sub-index for any country is the arithmetic average of its PFB-adjusted pillars for that sub-index, multiplied by 100. The maximum value of the sub-indices is 100 and the potential minimum is 0, both of which reflect a country's relative position in a particular sub-index.

$$ATT_i = 100 \sum_{j=1}^{5} \frac{h_j}{5} \qquad (11a)$$

$$ABT_i = 100 \sum_{j=6}^{9} \frac{h_j}{4} \qquad (11b)$$

$$ASP_i = 100 \sum_{j=10}^{14} \frac{h_j}{5} \qquad (11c)$$

where $h_{i,j}$ is the modified, post-penalty value of pillar j in country i
i = 1, 2,……n = the number of countries
j = 1, 2,…….14 = the number of pillars

9. **The Global Entrepreneurship Index calculation** The super-index, the Global Entrepreneurship Index, is simply the average of the three sub-indices. Since 100 represents the theoretically available limit, the GEI points can also be interpreted as a measure of the efficiency of the entrepreneurship resources

$$GEI_i = \frac{1}{3}(ATT_i + ABT_i + ASP_i) \qquad (12)$$

where i = 1, 2,……n = the number of countries

The calculation of the Individual and Institutional variables scores

In the GEI publication we also provide data about the Individual and Institutional variables scores independently. These scores are calculated the same way as the GEI average pillar scores, repeating points 3-5.

First, we normalize the variable scores for all the fourteen individual and institutional variables:

$$IND(norm)_{i,j} = \frac{IND_{i,j}}{\max IND_{i,j}} \qquad (13a)$$

$$INS(norm)_{i,j} = \frac{INS_{i,j}}{\max INS_{i,j}} \qquad (13B)$$

for all j = 1 ... k, the number of pillars (equal to the number of individual and institutional variables)
where $IND(norm)_{i,j}$ is the normalized INDIVIDUAL score value for country i and pillar j
where $INS(norm)_{i,j}$ is the normalized INSTITUTIONAL score value for country i and pillar j
$IND_{i,j}$ is the original INDIVIDUAL variable value for country i and pillar j
$INS_{i,j}$ is the original INSTITUTIONAL variable value for country i and pillar j
$max\ IND_{i,j}$ is the maximum value for INDIVIDUAL variable value in the case of pillar j
$max\ INS_{i,j}$ is the maximum value for INDSTITUTIONAL variable value in the case of pillar j

Second, we cap all the normalized individual and institutional variables. Similar to the GEI index score calculation we selected the 95th percentile score adjustment.

Third, we calculate the average adjusted individual and institutional variable pillar scores to equalize their average values. Equation 13a and 13b shows the calculation of the average value of the fourteen normalized and capped individual and institutional variables, respectively:

$$\overline{\mathrm{IND_j}} = \frac{\sum_{i=1}^{n} \mathrm{IND(norm)_{i,j}}}{n} \text{ for all j} \qquad (14a)$$

$$\overline{\mathrm{INS_j}} = \frac{\sum_{i=1}^{n} \mathrm{INS(norm)_{i,j}}}{n} \text{ for all j} \qquad (14b)$$

We want to transform the $INS(norm)_{i,j}$ and the $INS(norm)_{i,j}$ values such that the potential values to be in the [0,1] range.

$$IND_{i,j} = IND(norm)_{i,j}^{k} \qquad (15a)$$

$$INS_{i,j} = INS(norm)_{i,j}^{k} \qquad (15b)$$

where k is the "strength of adjustment", the k-th moment of IND_j and INS_j is exactly the needed average, $\overline{IND_j}$ and $\overline{INS_j}$

We have to find the root of the following equation for k:

$$\sum_{i=1}^{n} IND_{i,j}^{k} - n\overline{IND_j} = 0 \qquad \text{(16a)}$$

$$\sum_{i=1}^{n} INS_{i,j}^{k} - n\overline{INS_j} = 0 \qquad \text{(16a)}$$

Finally the individual and the institutional variables score for country *i* are calculated as the average of the average adjusted individual and institutional variables scores multiplied by 100

$$IND_i = \frac{100}{14}\sum_{j=1}^{14} IND_j \qquad (17a)$$

$$INS_i = \frac{100}{14}\sum_{j=1}^{14} INS_j \qquad (17b)$$