Country level efficiency and National Systems of Entrepreneurship: A Data Envelopment Analysis approach

Esteban Lafuente
Department of Management, Universitat Politècnica de Catalunya (Barcelona Tech)
EPSEB, Av. Gregorio Marañón, 44-50, E-08028 Barcelona, Spain
E-mail: esteban.lafuente@upc.edu

László Szerb
Faculty of Business and Economics, University of Pécs
Pécs, Rákóczi 80, H-7622, Hungary
E-mail: szerb@ktk.pte.hu

Zoltan J. Acs
Department of Management
London School of Economics and Political Science
Houghton Street, London WC2A 2AE
E-mail: Z.J.Acs@lse.ac.uk

Abstract
This paper tests the efficiency hypothesis of the knowledge spillover theory of entrepreneurship. Using a comprehensive database for 63 countries for 2012, we employ Data Envelopment Analysis to directly test how countries capitalize on their available entrepreneurial resources. Results support the efficiency hypothesis of knowledge spillover entrepreneurship. We find that innovation-driven economies make a more efficient use of their resources, and that the accumulation of market potential by existing incumbent businesses explains country-level inefficiency. Regardless of the stage of development, knowledge formation is a response to market opportunities and a healthy national system of entrepreneurship is associated with knowledge spillovers that are a prerequisite for higher levels of efficiency. Public policies promoting economic growth should consider national systems of entrepreneurship as a critical priority, so that entrepreneurs can effectively allocate resources in the economy.

Keywords: Knowledge spillover theory; GEDI; GEM; Efficiency; Data Envelopment Analysis; clusters

JEL classification: C4; O10; L26; M13

1. Introduction

Productivity is not only heterogeneous across countries, but also in terms of the factors explaining productivity differences between and within territories over time (Barro, 1991). A natural presumption is that technology plays a decisive role in shaping territorial productivity. However, when we look at productivity among rich and poor countries the picture gets less clear. It is not obvious that the answer is just technology. The most significant reason against blaming the gap in productivity growth on technology is that most developing countries have access to advanced technology. For example, data from the World Bank[1] reveal that the deepening of the cellular technology has grown in most countries, thus cell phone devices are available today, regardless of the stage of development of the country. Nevertheless, the use of advanced technologies in developing countries is hampered by the limited capacity of these economies to create support structures to efficiently use technological devices or tools (e.g., cell tower networks or bandwidth capacity).

In this context, at the country level we argue that productivity differences do not result exclusively from technology gaps, but also from differences in efficiency (Färe et al., 1994; Boussemart et al., 2003; Mahlberg and Sahoo, 2011). From an economic perspective, efficiency—in terms of input usage or output production—is related to the coefficient of resource utilization introduced by Debreu (1951) and further developed by Farrell (1957), and is represented by a distance function which captures efficiency differences that originate in factors other than differences in technology.

Efficiency is a key concept in economics. For example, in the field of economic growth productivity changes can be decomposed into technology and efficiency:

[1] Data were obtained from the World Bank (http://data.worldbank.org/indicator/IT.CEL.SETS.P2)

efficiency measures how effectively given technology and factors of production are actually used in an economy. The link between economic theory and efficiency measures based on distance functions now seems more evident: irrespective of the amount and quality of production factors, if available input factors are not combined efficiently a country will be off of the production possibilities frontier. While a large literature now exists on distance functions (see *e.g.*, Cooper et al., 2011), the analysis of the impact of entrepreneurship in shaping territorial efficiency remains, to the best of our knowledge, empirically untested. This paper seeks to gain a deeper understanding of efficiency differences at country level by connecting knowledge diffusion and entrepreneurship in endogenous growth models (Braunerhjelm et al., 2010) and the knowledge spillover theory of entrepreneurship (Acs et al., 2009).

Three core conjectures derive from the knowledge spillover theory of entrepreneurship. First, the knowledge hypothesis states that, *ceteris paribus*, entrepreneurial activity will tend to be greater in contexts where investment in knowledge are relatively high, since new firms will be started from knowledge that has spilled over from the source producing that new knowledge (Audretsch et al., 2006). Second, the commercialization *efficiency hypothesis* predicts that the more efficiently incumbents exploit knowledge flows, the smaller the effect of new knowledge on entrepreneurship (Acs et al., 2009). Finally, entrepreneurial activities would likely decrease in contexts characterized by higher regulations, complex administrative barriers and governmental intervention (Pekka et al., 2013).

Empirical analysis provides strong support for the knowledge hypothesis (Anselin et al., 1997), while the commercialization efficiency hypothesis has yet to be tested directly, existing evidence is inconclusive. Audretsch et al. (2006) suggest that a region’s investment in physical capital ‘represents the pursuit of economic opportunities within incumbent firms rather than in start-ups', but the authors find no statistically significant relationship between knowledge spillovers and capital investment. In contrast, arguing that patents indicate incumbents' effort to monopolize the knowledge that would otherwise seed new firms, Acs et al. (2009) find that the rate of self-employment is lower in countries where number of patents is greater. The ambiguity of the results concerning the efficiency hypothesis likely reflects the difficulty of measuring the firm's commercialization efficiency (Sanandaji and Leeson, 2013).[2]

The purpose of this paper is twofold. First, we scrutinize the effects of national systems of entrepreneurship on country-level efficiency. Second, we analyze the relationship between efficiency and certain variables related to the regulatory environment to create and run a business and to the social capital networks. One aspect of this story is that in middle income countries large corporations usually have controlling owners, who are usually very wealthy families. These ownership structures, jointly with high economic entrenchment create inefficiency in the economy: the middle income trap (Morck et al., 2004). In these countries a large number of relatively efficient businesses accumulate market potential, and performance of new businesses does not differ from that of incumbent ones which exploit knowledge spillovers. On contrary, if businesses in the economy are inefficient at exploiting knowledge entrepreneurial activity should be present.

The empirical application an international sample of 63 countries for 2012 and we use input data from the Global Entrepreneurship and Development Index (GEDI)—which captures the multidimensional nature of the country's entrepreneurship ecosystem—and macroeconomic data from the World Bank databases. We use a Data Envelopment Analysis (DEA) frontier method (Cooper et al., 2011) to directly test the

[2] Also see Plummer and Acs (2014) who test the localization hypothesis and localized competition at the local level for U.S. counties.

efficiency hypothesis. DEA is a complex benchmarking non-parametric technique that, through linear programming, yields a production possibilities frontier that approximates the technology of the analyzed units. The flexible nature of DEA models is especially appealing for applications in diverse and heterogeneous contexts (Grifell-Tatjé and Lovell, 1999; Epure and Lafuente, 2015). The second stage proposes a cluster analysis that introduces country-specific factors unconnected to the DEA model that might explain performance differences across the analyzed countries.

The results indicate that a specification that includes the national system of entrepreneurship to model the country's technology significantly contributes to explain efficiency differences. The findings give support to the efficiency hypothesis of the knowledge spillover theory of entrepreneurship. Among the analyzed countries, we find that average inefficiency is 61.68%—which represents the average output expansion that can be achieved to reach the efficiency frontier—and that inefficiency is greater in less developed countries. Although inefficiency widely varies across countries, knowledge investments and friendly environmental conditions to do business are conducive to efficiency, irrespective of the country's stages of development.

The following section presents the theoretical underpinning. Section 3 describes the data and the methodological approach. Section 4 presents the empirical findings, and Section 5 provides the discussion and concluding remarks.

2. Theoretical underpinning and hypothesis formulation

The more recent advance—endogenous growth theory—has been based on the emergence of research and development based models of growth, in the seminal papers of Romer (1990) and Aghion and Howitt (1992). These economic models explicitly aim to explain the role of technological progress in the growth process. R&D based models view technology as the primary determinant of growth and treats it as an endogenous variable. These models add the stock of ideas to the traditional inputs of physical capital and labor. For example, Romer (1990) assumes a knowledge production function in which new knowledge is linear in the existing stock of knowledge, holding the amount of research labor constant. The idea is expressed in the simple model where the growth rate is proportional to $\AA / A = \delta H_A$ where δ denotes the average research productivity, A is the stock of knowledge and H is the number of knowledge workers (R&D). Because, in the Romer's model, long-run per capita growth is driven by technological progress, knowledge growth will increase long-run growth in the economy.

The Romer model (1990) gives us a starting point to frame investigation of sustainable rate of technological progress according to the national knowledge production function:

$$\AA = \delta H_A^{\gamma} A^{\phi} \tag{1}$$

where, ϕ is the elasticity of research productivity of research workers, and γ measures the elasticity of inter-temporal knowledge spillover from the past on current research efforts (standing on the shoulders of giants). Romer assumed a particular form of the knowledge production function. The key restrictions made by Romer in his model are $\phi = 1$ and $\gamma = 1$, which makes $\AA$ linear in A and hence generates growth in the stock of knowledge ($\AA/A$) that depends on L_A unit homogeneously:

$$\AA / A = \delta L_A \tag{2}$$

That is, the growth rate of the stock of knowledge depends positively on the amount of labor devoted to R&D. This key result has important policy implications:

Policies in a country which permanently increase the amount of labor devoted to research have a permanent long run effect on the growth rate of the economy.

The model proposed by Romer captures two important relationships. First, long-run knowledge productions function where the flow of new knowledge depends positively on the existing stock of knowledge *A*, and the number of R&D workers *L*. Second, underlying the Romer's model is the assumption of a long-run positive relationship between total factor productivity and the stock of knowledge in the focal national context. The results indicate the presence of strong inter-temporal knowledge spillovers. The elasticity of new knowledge with respect to existing stocks of knowledge ϕ is at least as large as unity. 'However, the long-run impact of the knowledge stock on TFP is small: doubling the stock of knowledge is estimated to increase TFP by only 10 percent in the long run' (Abdih and Joutz, 2006, p. 244). The focus of the transmission mechanism between knowledge and TFP is needed to explain the parameter γ above.

Productivity not only differs between countries and it also changes within countries over time. A natural presumption is that technology plays a decisive role in this as we saw above. However, when we look at productivity among rich and poor countries the picture gets less clear. It is not obvious that the answer is just technology. But if differences in technology do not explain differences in productivity what does?

The most significant reason against blaming the gap in productivity growth on technology is that most developing countries access advanced technology (e.g., cell phones). Nevertheless, although advanced technologies are available in most developing economies, these countries lack appropriate support structures that allows at efficiently using technological devices or tools (e.g., cell tower networks, bandwidth capacity).

We argue that the other source of productivity differences come from *efficiency*. Efficiency is an umbrella concept used to capture anything that accounts for productivity differences that originate in factors other than differences in technology.

$$P = T \times E \tag{3}$$

where P is a measure of productivity, T is a measure of technology, and E is a measure of efficiency. Country-level data shows wide differences in the level of both technology and productivity. To what extend are the differences due to differences in technology and the differences in efficiency? Let's propose the case of two hypothetical countries (Z and W) where country Z is G years behind country W technologically. Mathematically: $T_{2012,Z} = T_{2012-G,W}$. Let g be the growth rate of technology in country W we can write:

$$T_{2012,z} / T_{2012,\mathrm{w}} = \ 1 + g \ ^{-G} \tag{4}$$

If the growth rate of technology in the country W is 0.54% and country Z is ten years behind the country W, then country Z has technology equivalent to 95% of that in country W. To see the differences in efficiency between two countries by going back to our equation above: $P_Z / P_W = (T_Z / T_W) \times (E_Z / E_W)$.

If for example the level of technology between country Z and country W is 0.31 percent then the left side of the equation is 0.31. The first term on the right side can be calculated from the above equation. If country Z has technology equal to 95% of country W level then efficiency in country Z equals 33% of country W level $(0.95 \times 0.33 = 0.31)$. The point for us is that unless the gap in technology is extremely large the differences in productivity will result from efficiency differences. As we increase the number of years in the technology gap the efficiency gap would continue to remain larger.

So what accounts for the large differences in efficiency between countries? These efficiency differences are about how the production factors and technology are combined. In our view efficiency differences come from differences in institutions as they set the rules of the game and from entrepreneurship that responds to these incentives, $E = I \times C^{*}$, where *E* is efficiency, *I* is institutions and *C** is entrepreneurship by individuals. We now turn to developing a methodology for measuring institutions and agency as they may affect productivity across countries from a systems perspective where $C = T \times NSE$, where *NSE* measures the national system of entrepreneurship.

The national system of entrepreneurship (NSE) refers to the combined effect of individual entrepreneurial initiatives and the context in which these initiatives operate. By definition, the 'National System of Entrepreneurship is the dynamic, institutionally embedded interaction between entrepreneurial attitudes, abilities, and aspirations by individuals, which drives the allocation of resources through the creation and operation of new ventures' (Acs et al., 2014, p. 479).

The analysis of the NSE permits to capture various inter-connected effects related to territorial economic performance. First, the NSE depicts the territory's capacity to mobilize available resources—in the form of interactions between individuals' attitudes, aspirations, and abilities—to the market through new business formation processes. Second, the NSE portrays the interactions between entrepreneurial human capital and accumulated knowledge and the multifaceted economic, social, and institutional contexts in which individuals develop their entrepreneurial activity. Finally, the NSE contributes to understand how entrepreneurial activity fuels territorial economic productivity through the efficient allocation of resources in the economy.

The relevance of the national systems of entrepreneurship flows from the recognition that entrepreneurship is a vital component present in any economy to a larger of lesser extent. Therefore, the systematic analysis of countries' efficiency including variables that account for the effects of entrepreneurial activity—i.e., through the national systems of entrepreneurship—helps not only to enhance the analysis of the factors that contribute to explain economic performance, but also to provide policy makers with valuable information on the economic contribution of entrepreneurship.

Based on the deductions resulting from the theoretical arguments that underpin this study we hypothesize:

H1: *The inclusion of the national system of entrepreneurship for modeling the country's technology contributes to explain efficiency differences across countries, relative to model specifications that do not incorporate national systems of entrepreneurship in the country's production function.*

3. Data and Method

3.1 Data

The data used to carry out this study come from several sources. First, data on the macroeconomic figures of the analyzed countries were obtained from the World Bank databases. Second, variables related to the country's demographic, educational and economic conditions, as well as to the entrepreneurial activity used to estimate the Global Entrepreneurship and Development Index (GEDI) were obtained from different sources, including the Global Entrepreneurship Monitor (GEM) adult population surveys, the Global Competitiveness Index (GCI), and the Doing Business Index. The GEDI scores were computed for 66 countries for 2012. Due to the lack of reliable information, Ethiopia, Taiwan, and Egypt were excluded from the analysis. Thus, the final sample comprises information for 63 countries.

It is worth noting that the representativeness of the sample is ensured insofar as it includes 30 European countries (Austria, Belgium, Bosnia and Herzegovina, Croatia, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Ireland, Israel, Italy, Latvia, Lithuania, Macedonia, Netherlands, Norway, Poland, Portugal, Romania, Russia, Slovak Republic, Slovenia, Spain, Sweden, Switzerland, Turkey, and United Kingdom), 14 American countries, including both North America and Latin America and the Caribbean islands (Argentina, Barbados, Brazil, Chile, Colombia, Costa Rica, Equator, El Salvador, Mexico, Panama, Peru, Trinidad & Tobago, United States, and Uruguay), eight Asian countries (China, Iran, Japan, the Republic of Korea, Malaysia, Pakistan, Singapore, Thailand), and 11 African countries (Algeria, Angola, Botswana, Ghana, Malawi, Namibia, Nigeria, South Africa, Tunisia, Uganda, and Zambia).

3.2 Efficiency Analysis

When dealing with multiple inputs yielding multiple outputs, efficiency literature usually makes use of Data Envelopment Analysis (hereafter DEA) frontier methods (Cooper et al., 2011). DEA is a non-parametric technique that, through linear programming, approximates the true but unknown technology without imposing any restriction on the sample distribution. The fundamental technological assumption of DEA is that any production unit (in our case, country) (*i*) uses $\mathbf{x} = (x_1, \ldots, x_J) \in R_+^J$ inputs to produce $\mathbf{y} = (y_1, \ldots, y_M) \in R_+^M$ outputs, and these sets form the technology (*T*): $T = \{(\mathbf{x}, \mathbf{y}) : \mathbf{x} \text{ can produce } \mathbf{y}\}$. DEA is a complex benchmarking technique that yields a production possibilities set where efficient decision-making units positioned on this surface shape the frontier. For the rest of units DEA computes an inefficiency score indicating the units’ distance to the best practice frontier.

The technology in DEA models has two properties that are worth defining. The first property relates to the returns to scale. In this study the modeled technology exhibits variable returns to scale (VRS) because pure technical efficiency measures (VRS) capture outcomes linked to practices undergone by decision makers in the short term (Chambers and Pope, 1996). The second assumption deals with the measurement orientation (input minimization or output maximization). The proposed DEA model maintains an output orientation. Business managers are often given output targets and told to produce them most efficiently, that is, with minimum inputs (Sengupta, 1987, p. 2290). To the contrary, in the public sector the workforce and assets tend to be fixed and policy-makers seek to produce the maximal possible output given the resources available (Fare et al., 1994, Tone and Sahoo, 2003). The following linear program models the described technology and computes the efficiency score for each country (*i*):

$$
\begin{aligned}
& T\ \ x_i^{'}, y_i^{'}\ \ = \max \theta_i \\
& \text{subject to } \sum_{i=1}^{N} \lambda_i y_{i,m} \geq \theta_i y_{i,m}^{'} \qquad , m = 1, \ldots, M \\
& \qquad \sum_{i=1}^{N} \lambda_i x_{i,j} \leq x_{i,j}^{'} \qquad , j = 1, \ldots, J \\
& \qquad \sum_{i=1}^{N} \lambda_i = 1 \\
& \qquad \lambda_i \geq 0 \qquad , i = 1, \ldots, N
\end{aligned}
\tag{5}
$$

The technology structure in equation (5) describes how countries transform their available resources (**x**: labor, capital and the national system of entrepreneurship) into the maximum possible output (**y**: GDP), uses λ as intensity weights to form the linear combinations of the sampled countries (*N*), and introduces the restriction $\sum_{i=1}^{N} \lambda_i = 1$ to impose variable returns to scale to the technology. The term θ_i is the efficiency score obtained for each country, and for efficient countries $\theta_i = 1$. For inefficient countries $\theta_i > 1$ and $\theta_i - 1$ points to the degree of inefficiency. Figure 1 presents a simplified representation of the distance function. For illustrative purposes, suppose that a fictitious country (*E*) has an inefficiency coefficient of $\theta = 1.25$. Thus, to operate efficiently and reach the frontier (*E**) this country should expand its output by 25%, while keeping its inputs fixed.

----- Insert Figure 1 about here -----

Existing research examines countries' efficiency under the premise that labor and capital generate gross domestic product (Fare et al., 1994; Boussemart et al., 2003; Mahlberg and Sahoo, 2011). In line with these studies the DEA model specification used to compute the world frontier defines an aggregate output (**y**: gross domestic product) that is produced by three inputs (**x**): labor, capital, and the national systems of entrepreneurship. Table 1 presents the descriptive statistics for the input-output set.

The gross domestic product (GDP) for the year 2012 is expressed at 2005 prices in million of PPP International US dollars. Labor is measured as the country's number of employees (expressed in millions of workers). Capital is defined as the gross capital formation, which represents the outlays on additions to the economy's fixed assets (public infrastructures, and commercial and residential buildings) plus net changes in the level of inventories held by firms in the economy[3].

----- Insert Table 1 about here -----

[3] According to the World Bank, gross capital formation consists of outlays on additions to the fixed assets of the economy plus net changes in the level of inventories. Fixed assets include land improvements (fences, ditches, drains, and so on); plant, machinery, and equipment purchases; and the construction of roads, railways, and the like, including schools, offices, hospitals, private residential dwellings, and commercial and industrial buildings. Inventories are stocks of goods held by firms to meet temporary or unexpected fluctuations in production or sales, and 'work in progress.'

The third input, the Global Entrepreneurship and Development Index (GEDI), captures the multidimensional nature of entrepreneurship at the country level. The GEDI index measures the dynamic and institutionally embedded interaction between entrepreneurial attitudes, entrepreneurial abilities and entrepreneurial aspirations by individuals, which drive resource allocation through new business venturing (Acs et al., 2014). The GEDI index, which ranges between zero and 100, is built on 14 pillars which result from 14 individual-level variables properly matched with selected institutional variables related to the country's entrepreneurship ecosystem.

The novelty of the GEDI index lies on the systemic view of countries' entrepreneurship in which the harmonization (configuration) of the analyzed pillars through the penalty for bottleneck (PFB) determines the country's systems of entrepreneurship (Miller 1986, 1996). Through the PFB method the system performance is mainly determined by the weakest element (bottleneck) in the system. The magnitude of the country-specific penalty depends on the absolute difference between each pillar and the weakest pillar. Also, pillars cannot be fully substituted through the PFB method, i.e. a poorly performing pillar can only be partially compensated by a better performing pillar. A detailed description of the structure of the GEDI index (variables and pillars) and the index building methodology are presented in the Appendix 2.

3.3 Second stage analysis

The second stage proposes a supplementary cluster analysis to further scrutinize how country-specific factors—which are unconnected to DEA scores—relate to efficiency. Table 2 presents the descriptive statistics of the variables used to cluster the analyzed countries. The first variable is the country's economic welfare measured by the gross domestic product per capita in 2012 (expressed at 2005 prices in PPP

International US dollars). Second, we account for the quality of the regulatory environment to create and operate a business which is critical for enhancing territorial entrepreneurial activity. Thus, we introduce the values of the doing business index for 2012 developed by the World Bank, with higher values pointing to a more friendly entrepreneurial environment.

The third factor relates to the countries' social capital networks, measured by the social capital index provided by the Legatum Institute (www.prosperity.com). This variable measures the strength of the countries' social cohesion, social engagement, as well as the performance of community and family networks, with higher values indicating greater level of social capital. The last factor is the unemployment rate. This variable has gained relevance in the context of the current economic downturn, as it not only deters the economic activity at the country level, but also sheds some light on the quality of countries' entrepreneurial activity. To enhance estimation accuracy, standardized values for the four variables are introduced in the cluster analysis.

----- Insert Table 2 about here -----

To attain the second stage analysis, we propose a non-hierarchical cluster analysis (*K*-means) using the efficiency scores of the entrepreneurship frontier and the variables in Table 2 as inputs. The cluster analysis is based on the Euclidean distance between vectors of the standardized values of the variables under analysis (Anderberg, 1973; Everitt, 1980). Through this procedure observations are classified according to the similarities of the country-specific dimensions analyzed. The *K*-means cluster analysis requires the establishment of a fixed number of clusters. This represents the main pitfall of non-hierarchical cluster analysis, because in many research fields (including social sciences) cluster analyses are often exploratory.

We adopt two approaches to corroborate the number of clusters and the validity of the analysis. First, we estimate the Calinski and Harabasz (1974) statistic. This index is obtained as $CH(k) = \frac{B(k)/k-1}{W(k)/n-k}$, where *B(k)* and *W(k)* are the between- and within-cluster sums of squares, with *k* clusters. Since the between cluster difference should be high, and the within cluster difference should be low, the largest *CH(k)* value indicates the best clustering. The result of the statistic—pseudo-F value: 277.33—reveals that the number of clusters that maximizes the *CH(k)* index is five. Second, we propose a discriminant analysis to further validate the cluster output, and results in Table 3 confirm that our approach to examine the sampled countries is appropriate.

----- Insert Table 3 about here -----

4. Empirical findings

4.1 Efficiency analysis

This section deals with the efficiency assessment of the analyzed countries. Table 4 presents the summary statistics of the inefficiency measure computed from equation (5), while the country-specific inefficiency scores are presented in Appendix 1.

Prior to reporting the results of our efficiency analysis we have run an additional robustness check to further corroborate that our approach—even if theoretically correct—accurately represents the countries' technology and is not affected by model specification (Nataraja and Johnson, 2011). We adopted the regression-based test by Ruggeiro (2005) to corroborate the impact of the input capturing the national system of entrepreneurship (GEDI index) and the significance of correctly introducing it in the countries' technology. This procedure is based on a variable selection approach in which an initial inefficiency measure—obtained from an input set—is regressed against a set of candidate variables. Variable will be deemed relevant for explaining the analyzed technology if regression coefficients are significant and have the correct sign (positive values for inputs and negative values for outputs).

In our case, we first tested whether the input capturing the national system of entrepreneurship should be included in the efficiency model (equation (5)). More concretely, and similar to Färe et al. (1994), Boussemart et al. (2003) and Mahlberg and Sahoo (2011), we estimated an alternative world economic frontier in which the GDP is produced by labor and capital, and inefficiency scores resulting from this specification are regressed against the candidate input (GEDI index). Following the intuition by Ruggeiro (2005), the result of the OLS regression confirms that the inclusion of the GEDI index in the input set explains inefficiency differences among the sampled countries $(\beta = 0.0178 \text{ and } p - value < 0.001)$. Goodness of fit measures validate this estimation approach (*F-test*: 26.64 and *p-value* < 0.001 – Adj. R2: 0.2956).

To address the threat of collinearity, in the second step we computed the variance inflation factor (VIF). Here, we regressed the inefficiency scores obtained from the model that incorporates the three inputs (labor, capital and GEDI index) against the input values. Although the validity of the regression model (*F-test*: 4.99 and *p-value*< 0.01 – Adj. R2: 0.0611), coefficients for the three input variables are not statistically significant at conventional levels (< 10%). Also, the average VIF value is 7.60 and the only variable for which the VIF value exceeds 10—a generally accepted rule of thumb for assessing collinearity—is capital formation (12.33). The results for this diagnostic test do not raise collinearity concerns, thus confirming that the proposed efficiency model accurately estimates the countries' technology.

To test hypothesis 1 we assessed the influence of introducing the GEDI index in the countries' technology (equation (5)) by examining the DEA model that considers GDP a function of labor and capital and the model that includes the GEDI index in the production function. The direct comparison between the two DEA models reveals that the most significant inefficiency changes resulting from the introduction of the GEDI index in the model are reported for Costa Rica (25.14%), Pakistan (17.57%) and Mexico (11.06%). The Wilcoxon signed-rank test was used to detect differences between the model that considers the GEDI index and the model that assesses economic efficiency. The result supports hypothesis 1. The DEA model that incorporates the GEDI index in the input set attains inefficiency scores significantly different at 1% level from the economic model. This corroborates that the full model considering the national systems of entrepreneurship is not only closer to the real countries' technology, but also enhances estimation and the interpretation of the results. As a result, in what follows we only analyze the scores of the model that considers the GEDI index in the technology.

Results reveal that average inefficiency among the analyzed countries is 61.68%. Figures in Appendix 1 show that six countries are found efficient (Brazil, China, Ireland, Singapore, United Kingdom and United States). Yet, inefficiency widely varies across countries and across stages of development. As expected innovation-driven economies present the best efficiency results (average inefficiency: 21.30%), while inefficiency in factor-driven countries is the highest (113.83%).

----- Insert Table 4 about here -----

European countries show the highest efficiency levels with an average inefficiency of 45.75%. At the country level, the findings indicate that Ireland and the

United Kingdom are efficiently employing their current resources. Additionally, low inefficiency levels are reported for Norway (1.90%), Germany (3.90%), Greece (4.00%) and Italy (9.20%). For interpretation purposes, the result for Germany indicates that, to operate efficiently and reach the world frontier, the country can exploit its available resources to expand its GDP by 3.90%. On contrary, the most inefficient countries in this continent are located in the Baltic area and Eastern Europe (see Appendix 1).

Average inefficiency in North and Latin American countries stands at 62.71%. Besides Brazil and the United States—efficient countries in this continent—Mexico (33.50%), Barbados (34.50%) and Costa Rica (41%) report relatively low inefficiency levels. On contrary, Equator, Peru and Panama present an inefficiency level that exceeds 100%, which implies that an efficient use of resources in these countries would yield more than twice as much output as the countries' actual GDP levels.

China and Singapore lead efficiency results in Asia (average inefficiency: 43.08%), while Thailand (94.60%) and Iran (97.50%) present the highest inefficiency score in this continent. Finally, the highest inefficiency results are found in Africa (average inefficiency: 117.35%), and in this case Angola (12.70%), Nigeria (18.85%) and South Africa (39.10%) are the most efficient countries. It should be noted that the inefficiency dispersion is the greatest in this continent and in the remaining eight African countries inefficiency exceeds 90%, which means that—to operate efficiently and reach the frontier—these countries can exploit their available resources to increase their GDP more than 90%.

4.2 Behavioral path across economies

This section presents the results of the supplementary cluster analysis. Figure 2 illustrates the positioning of the groups of countries according to their inefficiency and

GEDI scores. Overall, the results for both the GEDI and the inefficiency scores are aligned with the path followed by countries based on the analyzed variables.

Results in Figure 2 indicate that five groups emerge from the cluster analysis. Groups 1 and 2 mostly comprise innovation-driven countries with strong national systems of entrepreneurship and low inefficiency levels. Countries in Group 1 show the lowest inefficiency (17.73%), while average inefficiency in Group 2 is 31.73%. Additionally, the result of the Kruskal-Wallis test reveals that inefficiency scores for these two groups are not significantly different. From Figure 2 we note that countries in these two groups benefit from a healthier and more stable economy, a regulatory environment conducive to start and run a business, and stronger social capital networks.

----- Insert Figure 2 about here -----

Group 3 is mainly formed by efficiency-driven economies (64.29%), and seven out of the 14 countries in the group are European former socialist countries. Performing Asian countries are also in this group (Japan, Malaysia, and South Korea). In this Group average inefficiency is 61.70%, and the result of the Kruskal-Wallis test indicates that inefficiency is significantly higher at 1% and 5% level than that reported for countries in Groups 1 and 2, respectively. Also, the values of the GEDI index for countries in this are significantly lower at 1% level than those reported for countries in Groups 1 and 2.

Similar to the results for Group 3, most countries in Group 4 are efficiency-driven economies (88.24%). Also, seven out of the 17 countries are in Latin America, and large emerging economies are in this group (China, Mexico, and Russia). Although the results of the Kruskal-Wallis test show that average inefficiency in this group (66.41%) is not significantly different to that found in Group 3, countries in this group lack efficient national systems of entrepreneurship as their average GEDI index is significantly lower than that reported for countries in Group 3 (Kruskal-Wallis test).

Finally, countries in Group 5 show the poorest results. This group mostly comprises factor-driven economies located in Africa (eight countries). Inefficiency in this group scores the highest (97.72%), and these countries also lag behind in terms of their national systems of entrepreneurship.[4] Countries in this group are characterized by deprived economic conditions and an underdeveloped institutional setting, which contributes to explain both their poor efficiency results and their weak national systems of entrepreneurship.

5. Conclusions and implications

This paper scrutinizes the efficiency hypothesis of the knowledge spillover theory of entrepreneurship. The analysis of the use of available resources by countries is increasingly important in the context of the current economic downturn that affects many economies around the world. Although scholars and policy makers acknowledge the wide array of social and economic advantages resulting from entrepreneurship, the analysis of the relationship between the country's entrepreneurship system and economic efficiency remains unaddressed. In this sense, the debate is open and this study provides evidence that contributes to understand how countries capitalize on their entrepreneurial system.

More concretely, the main contribution of this study relies on the comprehensive efficiency analysis of 63 countries through a non-parametric technique—Data Envelopment Analysis (DEA)—which allows at modeling GDP per head as a function of input variables that can be directly shaped by policy makers. Building on insights

[4] The result of the Kruskal Wallis test confirms that the GEDI index for countries in Group 5 is significantly lower at the 1% level than the value reported for countries in the rest of Groups.

from the knowledge spillover theory of entrepreneurship, we compute a world frontier that incorporates into the model the national system of entrepreneurship as a critical input that contributes to explain efficiency differences across the analyzed economies.

Overall, the findings are consistent with the efficiency hypothesis of the knowledge spillover theory of entrepreneurship. Results indicate that country-level efficiency analyses significantly benefit from the incorporation of variables capturing the countries' entrepreneurial system. Additionally, and although inefficiency widely varies across countries, we find that innovation-driven economies show the best efficiency results, while the group of factor-driven countries are the most inefficient. Regression results support the knowledge commercialization efficiency hypothesis. While Audretsch et al. (2006) report a positive but non-significant effect of incumbent firms on knowledge filter; our results indicate that the accumulation of market potential by existing incumbent businesses explains country-level inefficiency.

We interpret the results of the study in terms of the benefits of national systems of entrepreneurship. Policy makers often allocate fat sums of public money in policies excessively oriented towards the stimulation of employment, capital and knowledge generation in the economy, such as subsidies to support self-employment and human capital formation and investments in research and development. These policies—rooted in the endogenous growth theory—are conducive to growth and they undoubtedly have translated into significant economic outcomes linked to increased levels of employment and education (Braunerhjelm et al., 2010). Nevertheless, the national systems of entrepreneurship have not received appropriate treatment as a country phenomenon.

The results of this study are consistent with the argument that, regardless of the stage of development, knowledge formation is a response to market opportunities, and that a healthy national system of entrepreneurship is associated to spillovers in other economic agents that proves itself a prerequisite for endogenous growth. From a policy perspective, our comprehensive analysis fuels the notion that policy should shift from an excessive focus on capital and labor towards designs that match knowledge and capital formation programs with policies that emphasize the need to enhance the national systems of entrepreneurship. Entrepreneurship support programs would become sterile if entrepreneurs navigate in contexts that do not guarantee the effective exploitation of their knowledge. Thus, policy makers need to turn their attention to the development of appropriate national systems of entrepreneurship; and prioritize policies that seek to improve the way through which the national systems of entrepreneurship channel knowledge to the economy and create economic growth in the long-run.

It must, however, be mentioned a series of limitations to the present study that, in turn, represent avenues for future research. First, the proposed analysis offers a compelling vision of the effects of healthy national systems of entrepreneurship on country-level efficiency. Yet, future research should attempt to introduce into the analysis further measures that permit to capture the knowledge exploitation by incumbent and new businesses as well as to estimate how, in relatively homogeneous entrepreneurial contexts, country-level efficiency is affected by the different types of knowledge exploitation made by entrepreneurs measured by the quality of entrepreneurship. Second, the cross-sectional nature of the study calls for obvious caution when interpreting and generalizing its findings.

Acknowledgement: László Szerb benefited from the financial support of the European Union (TÁMOP project: No. 4.2.2 A–11/1/KONV-2012-0058). Esteban Lafuente acknowledges financial support by the Spanish Ministry of Science and Innovation (ECO2013-48496-C4-4-R).

References

Acs, Z.J., Audretsch, D.B. (1988). Innovation in large and small firms—an empirical analysis. *American Economic Review*, 78, 678-690.

Acs, Z.J., Autio, E., Szerb, L. (2014). National Systems of Entrepreneurship: Measurement issues and policy implications. *Research Policy*, 43, 476-494.

Acs, Z.J., Braunerhjelm, P., Audretsch, D.B., Carlsson, B. (2009). A knowledge spillover theory of entrepreneurship. *Small Business Economics*, 32, 15-30.

Anselin, L., Varga, A., Acs, Z.J. (1997). Local Geographic Spillovers Between University Research and High Technology Innovations, *Journal of Urban Economics*, 42, 422-448.

Anderberg, M.R. (1973). *Cluster analysis for applications*. New York: Academic Press.

Abdih, Y., Joutz, F. (2006). Relating the knowledge production function to total factor productivity: An endogenous growth puzzle, *IMF Staff Papers*, 53(2), 242-271.

Aghion, P., Howitt, P.A. (1992). Model of growth through creative destruction. *Econometrica*, 60 (2), 323-351.

Audretsch, D.B., Keilbach, M.C., Lehmann, E.E. (2006). *Entrepreneurship and Economic Growth*. Oxford University Press, USA.

Baumol, W. (1990). Entrepreneurship: Productive, unproductive and destructive. *Journal of Political Economy*, 98, 893-921.

Barro, R.J., (199l). Economic growth in a cross section of countries, *Quarterly Journal of Economics* CVI, May, 407-444.

Boussemart, J.-P., Briec, W., Kerstens, K., Poutineau, J.-C. (2003). Luenberger and Malmquist Productivity Indices: Theoretical Comparisons and Empirical Illustration. *Bulletin of Economic Research*, 55 (4), 391-405.

Braunerhjelm, P., Acs, Z.J., Audretsch, D., Carlsson, B. (2010). The missing link: knowledge diffusion and entrepreneurship in endogenous growth. *Small Business Economics,* 34 (2), 105-125.

Calinski, R.B., Harabasz, J. (1974). A dendrite method for cluster analysis. *Communications in Statistics*, 3 (1), 1-27.

Chambers, R.G., Pope, R.D. (1996). Aggregate productivity measures. *American Journal of Agricultural Economics*, 78 (5), 1360-1365.

Cooper, W.W., Seiford, L.M., Zhu, J. (2011). *Handbook on data envelopment analysis*. 2nd edition, New York: Springer.

Debreu, G. (1951). The coefficient of resource utilization. *Econometrica*, 19, 273-292.

Djankov, S., La Porta, R., Lopez-de-Silanes, F., Shleifer, A. (2008). The law and economics of self-dealing. *Journal of Financial Economics*, 88, 430–465.

Epure, M., Lafuente, E. (2015). Monitoring bank performance in the presence of risk. *Journal of Productivity Analysis*, in press (DOI: 10.1007/s11123-014-0413-z).

Everitt, B.S. (1980). *Cluster analysis*. 2nd edition, London: Heineman.

Färe, R., Grosskopf, S., Norris, M., Zhang, Z. (1994). Productivity Growth, Technical Progress, and Efficiency Change in Industrialized Countries. *American Economic Review*, 84 (1), 66-81.

Farrell, M.J. (1957). The measurement of productive efficiency. *Journal of the Royal Statistical Society*, 120, 253-281.

Greene, W. (2003). *Econometric analysis*. 5th edition, Upper Saddler River, New Jersey.

Grifell-Tatjé E., Lovell, C.A.K. (1999). Profits and productivity. *Management Science*, 45 (9), 1177-1193.

Epure, M., Lafuente, E. (2015). Monitoring bank performance in the presence of risk. *Journal of Productivity Analysis*, in press (DOI: 10.1007/s11123-014-0413-z).

Mahlberg, B., Sahoo, B.K. (2011). Radial and non-radial decompositions of Luenberger productivity indicator with an illustrative application. *International Journal of Production Economics*, 131 (2), 721-726.

Miller, D. (1986). Configurations of strategy and structure: Towards a synthesis. *Strategic Management Journal*, 7, 233-249.

Miller, D. (1996). Configurations revisited. *Strategic Management Journal*, 17, 505-512.

Morck, R., Wolfenzon, D., Yeung, D. (2004). Corporate Governance, Economic Entrenchment and Growth. National Bureau of Economic Research, Working paper 10692, August 2004.

Nataraja, N.R., Johnson, A.L. (2011). Guidelines for using variable selection techniques in data envelopment analysis. *European Journal of Operational Research*, 215 (3), 662-669.

Parker, S.C. (2009). *The Economics of Entrepreneurship*. Cambridge University Press, London.

Pekka, S., Acs, Z.J., Weubker, R.. (2013). Exploring Country Level Institutional Arrangements on the Rate and Type of Entrepreneurial Activity. *Journal of Business Venturing*, 28 (1), 176-193.

Romer, P. (1990). Endogenous technological change. *Journal of Political Economy*, 98 (October), 71-102.

Ruggiero, J. (2005). Impact assessment of input omission on DEA. *International Journal of Information Technology & Decision Making*, 4 (3), 359-368.

Sanandaji, T., Leson, P.T. (2013). Billionairs. *Industrial and Corporate Change*, 22 (1), 313-337.

Sengupta, J.K. (1987). Efficiency measurement in non-market systems through data envelopment analysis. *International Journal of Systems Science*, 18 (12), 2279-2304.

Simar, L., Wilson, P. (2011). Two-stage DEA: caveat emptor. *Journal Productivity Analysis*, 36 (2), 205-218.

van Praag, M.C. (2007). What is the value of entrepreneurship? A review of recent research. *Small Business Economics*, 29, 351-382.

World Economic Forum (2013). *The Global Competitiveness Report* 2013–2014. Geneva, Switzerland.

List of Figures and Tables

Figure 1. Efficiency analysis based on Data Envelopment Analysis

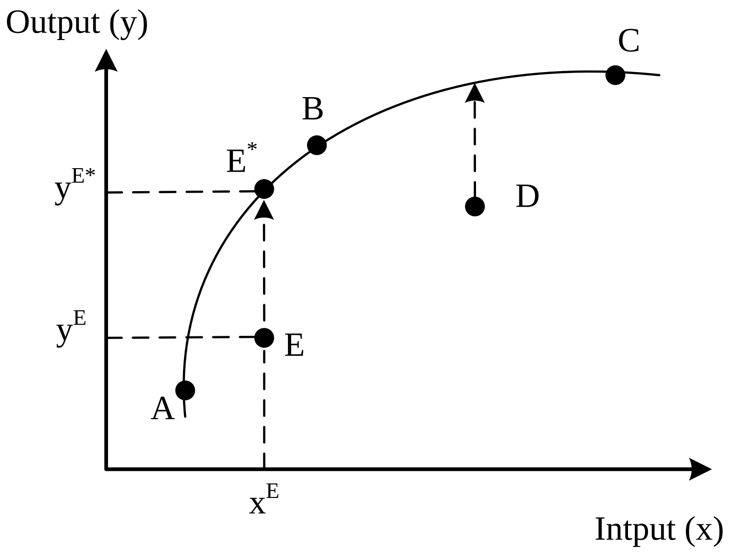
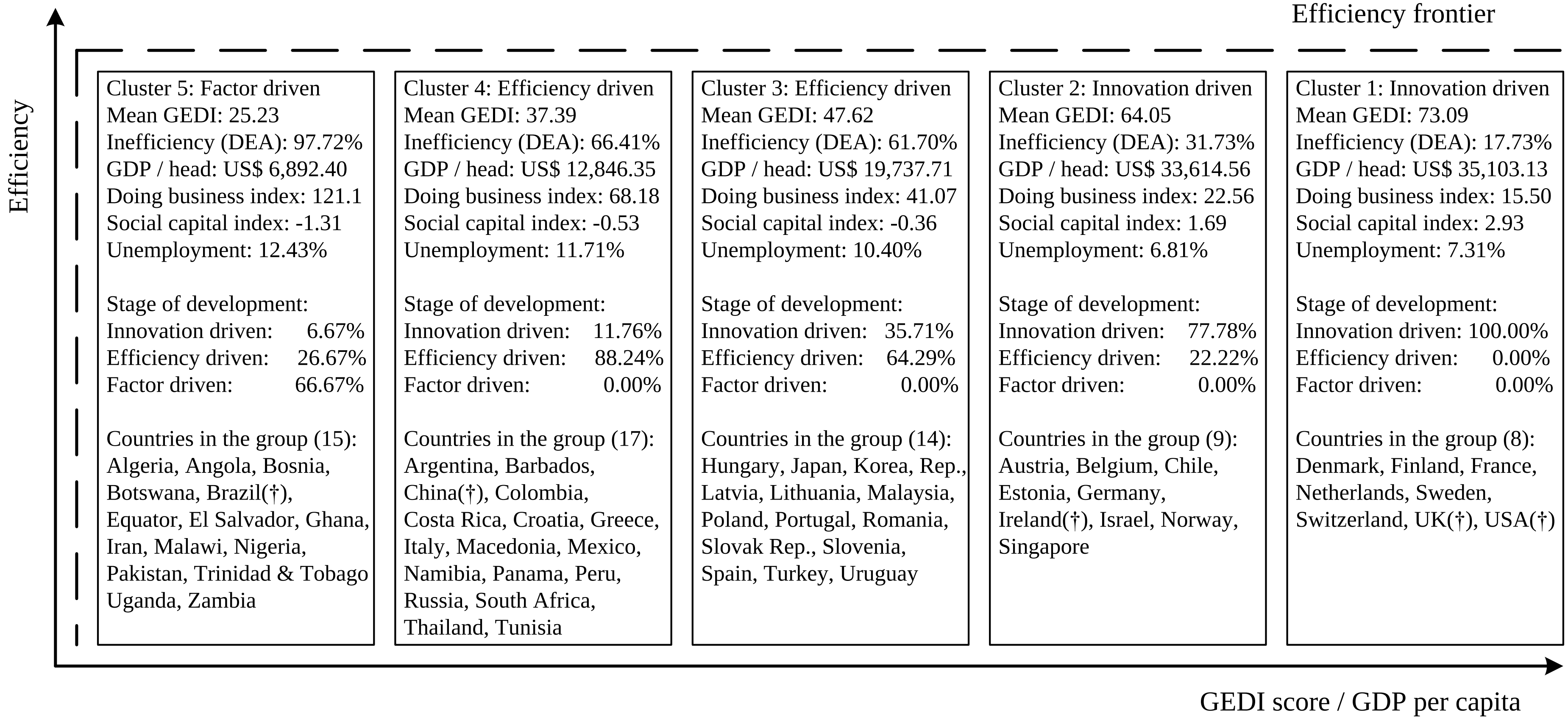


Figure 2. The relationship between the GEDI score and the performance of countries



Data on the stages of economic development were obtained from the World Economic Forum (2013). The reported Gross Domestic Product (GDP) per capita for the year 2012 is expressed at 2005 prices in PPP international US dollars. (†) indicates that the country is efficient.

Table 1. Descriptive statistics for the selected input-output set

	Description	Mean (Std. dev.)	Q1	Median	Q3
Output					
Gross domestic product (GDP)	GDP equals the gross value added by the country producers plus product taxes and minus subsidies not included in the value of the products.	906,663 (2,205,548)	53,607	244,043	636,888
Inputs					
Labor force	Labor force comprises the economically active population: people over 15 years old who supply labor for the production of goods and services.	30.43 (100.79)	2.67	7.20	25.66
Gross capital formation (GCF)	GCF consists of outlays on additions to the fixed assets of the economy plus net changes in the level of inventories.	233,429 (730,409)	11,538	59,776	145,710
GEDI score	Index that measures the country's systems of entrepreneurship	45.1096 (16.7791)	32.7176	43.0896	59.4776

Sample size: 63 countries. Economic and labor figures for the year 2012 were obtained from the World Bank, while the GEDI scores were provided by the International GEM Consortium.

Table 2. Cluster analysis: Descriptive statistics for the selected variables

	Mean	Std. dev.	Q1	Median	Q3
GDP per head (PPP constant 2005 international US$)	18,753.30	12,438.56	9,124.00	15,848.00	27,991.00
Doing business index	61.7937	46.6321	25	51	92
Social capital index	0.0786	1.8230	-1.3740	-0.0650	0.8230
Unemployment rate	0.1033	0.0711	0.0530	0.0790	0.1390

Sample size: 63 countries.

Table 3. Results of the Discriminant Analysis

True groups	Classification according to the discriminant analysis					
	1	2	3	4	5	Observations
Group 1	8 (100.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	8
Group 2	0 (0.00%)	9 (100.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	9
Group 3	0 (0.00%)	0 (0.00%)	14 (100.00%)	0 (0.00%)	0 (0.00%)	14
Group 4	0 (0.00%)	0 (0.00%)	1 (5.88%)	16 (94.12%)	0 (0.00%)	17
Group 5	0 (0.00%)	0 (0.00%)	0 (0.00%)	1 (6.67%)	14 (93.33%)	15
Total	8	9	15	17	14	63

Table 4. Inefficiency scores estimated through Data Envelopment Analysis

	Values
Average inefficiency	61.68%
Standard deviation	54.16%
Bottom quartile (Q1)	18.85%
Median value (Q2)	42.80%
Upper quartile (Q3)	97.20%
Number of efficient countries	6
Total number of countries	63
Innovation-driven countries (N=23)	
Average inefficiency (Std. dev.)	21.30% (20.06%)
Efficiency-driven countries (N=30)	
Average inefficiency (Std. dev.)	75.26% (40.59%)
Factor-driven countries (N=10)	
Average inefficiency (Std. dev.)	113.83% (78.16%)

Appendix 1: Inefficiency score of the analyzed countries

N	Country	Inefficiency score	N	Country	Inefficiency score
European countries			**North and Latin America**		
1	Austria	21.70%	31	Argentina	50.10%
2	Belgium	14.10%	32	Barbados	34.50%
3	Bosnia and Herzegovina	108.80%	33	Brazil	0.00%
4	Croatia	73.50%	34	Chile	83.20%
5	Denmark	28.70%	35	Colombia	71.80%
6	Estonia	121.00%	36	Costa Rica	41.00%
7	Finland	29.00%	37	Ecuador	105.30%
8	France	15.70%	38	El Salvador	48.60%
9	Germany	3.90%	39	Mexico	33.50%
10	Greece	4.00%	40	Panama	135.10%
11	Hungary	49.30%	41	Peru	105.10%
12	Ireland	0.00%	42	Trinidad & Tobago	72.50%
13	Israel	39.70%	43	United States	0.00%
14	Italy	9.20%	44	Uruguay	97.30%
15	Latvia	139.10%			
16	Lithuania	73.80%	**Asian countries**		
17	Macedonia, FYR	166.00%	45	China	0.00%
18	Netherlands	23.70%	46	Iran, Islamic Rep.	97.50%
19	Norway	1.90%	47	Japan	12.30%
20	Poland	42.80%	48	Korea, Rep.	50.90%
21	Portugal	28.20%	49	Malaysia	78.10%
22	Romania	90.90%	50	Pakistan	11.20%
23	Russia	41.70%	51	Singapore	0.00%
24	Slovak Republic	72.40%	52	Thailand	94.60%
25	Slovenia	62.60%			
26	Spain	27.10%	**African countries**		
27	Sweden	21.90%	53	Algeria	156.00%
28	Switzerland	22.80%	54	Angola	12.70%
29	Turkey	39.00%	55	Botswana	174.90%
30	United Kingdom	0.00%	56	Ghana	207.60%
			57	Malawi	90.10%
			58	Namibia	125.00%
			59	Nigeria	18.85%
			60	South Africa	39.10%
			61	Tunisia	97.20%
			62	Uganda	188.50%
			63	Zambia	180.90%

Appendix 2: Global Entrepreneurship and Development Index (GEDI)

Table A1. Structure of the GEDI index

Institutional variable	**Individual variable**	**Pillar**	**Sub-Index**	**GEDI**
Market Agglomeration	Opportunity Recognition	Opportunity Perception	Entrepreneurial attitudes	**Global Entrepreneurship and Development Index (GEDI)**
Tertiary Education	Skill Perception	Start-up Skills		
Business Risk	Risk Acceptance	Non-fear of Failure		
Internet Usage	Know Entrepreneurs	Networking		
Corruption	Career Status	Cultural Support		
Freedom	Opportunity Motivation	Opportunity Startup	Entrepreneurial abilities	
Tech Absorption	Technology Level	Tech Sector		
Staff Training	Educational Level	Quality of Human Resources		
Market Dominance	Competitors	Competition		
Technology Transfer	New Product	Product Innovation	Entrepreneurial aspirations	
GERD	New Tech	Process Innovation		
Business Strategy	Gazelle	High Growth		
Globalization	Export	Internationalization		
Depth of Capital Market	Informal Investment	Risk Capital		

Table A2. Description of the individual variables used to create the GEDI index

Individual variable*	**Description**
Opportunity Recognition	The percentage of the 18-64 aged population recognizing good conditions to start business next 6 months in area he/she lives,
Skill Perception	The percentage of the 18-64 aged population claiming to posses the required knowledge/skills to start business
Risk Acceptance	The percentage of the 18-64 aged population stating that the fear of failure would not prevent starting a business
Know Entrepreneurs	The percentage of the 18-64 aged population knowing someone who started a business in the past 2 years
Carrier	The percentage of the 18-64 aged population saying that people consider starting business as good carrier choice
Status	The percentage of the 18-64 aged population thinking that people attach high status to successful entrepreneurs
Career Status	The status and respect of entrepreneurs calculated as the average of Carrier and Status
Opportunity Motivation	Percentage of the TEA businesses initiated because of opportunity start-up motive
Technology Level	Percentage of the TEA businesses that are active in technology sectors (high or medium)
Educational Level	Percentage of the TEA businesses owner/managers having participated over secondary education
Competitors	Percentage of the TEA businesses started in those markets where not many businesses offer the same product
New Product	Percentage of the TEA businesses offering products that are new to at least some of the customers
New Tech	Percentage of the TEA businesses using new technology that is less than 5 years old average (including 1 year)
Gazelle	Percentage of the TEA businesses having high job expectation average (over 10 more employees and 50% in 5 years)
Export	Percentage of the TEA businesses where at least some customers are outside country (over 1%)
Average informal investment	The mean amount of 3 year informal investment
Business Angel	The percentage of the 18-64 aged population who provided funds for new business in past 3 years excluding stocks & funds, average
Informal Investment	The amount of informal investment calculated as Average informal investment * Business Angel

*All individual variables are from the GEM Adult Population Surveys.

Table A3. Description and source of the GEDI applied institutional variables

Institutional variable	Description	Source of data	Data availability
Domestic Market	Domestic market size that is the sum of gross domestic product plus value of imports of goods and services, minus value of exports of goods and services, normalized on a 1–7 (best) scale data are from the World Economic Forum Competitiveness	World Economic Forum	The Global Competitiveness Report 2013-2014, p. 518
Urbanization	Urbanization that is the percentage of the population living in urban areas, data are from the Population Division of the United Nations, 2011 revision	United Nations	http://esa.un.org/unup/CD-ROM/Urban-Rural-Population.htm
Market Agglomeration	The size of the market: a combined measure of the domestic market size and the urbanization that later measures the potential agglomeration effect. Calculated as domestic market urbanization*	Own calculation	-
Tertiary Education	Gross enrolment ratio in tertiary education, 2012 or latest available data.	UNESCO	http://data.un.org/Data.aspx?d=UNESCO&f=series%3AGER_56
Business Risk	The business climate rate "assesses the overall business environment quality in a country…It reflects whether corporate financial information is available and reliable, whether the legal system provides fair and efficient creditor protection, and whether a country's institutional framework is favorable to intercompany transactions" (http://www.trading-safely.com/). It is a part of the country risk rate. The alphabetical rating is turned to a seven-point Likert scale from 1 (D rating) to 7 (A1 rating). December 30, 2013 data	Coface	http://www.coface.com/Economic-Studies-and-Country-Risks/Rating-table
Internet Usage	The number of Internet users in a particular country per 100 inhabitants, 2013 data	International Telecommunication Union	http://www.itu.int/en/ITU-D/Statistics/Pages/stat/default.aspx
Corruption	The Corruption Perceptions Index (CPI) measures the perceived level of public-sector corruption in a country. "The CPI is a 'survey of surveys', based on 13 different expert and business surveys." (http://www.transparency.org/policy_research/surveys_indices/cpi/2009) Overall performance is measured on a ten-point Likert scale. Data are from 2013.	Transparency International	http://cpi.transparency.org/cpi2013/
Economic Freedom	"Business freedom is a quantitative measure of the ability to start, operate, and close a business that represents the overall burden of regulation, as well as the efficiency of government in the regulatory process. The business freedom score for each country is a number between 0 and 100, with 100 equaling the freest business environment. The score is based on 10 factors, all weighted equally, using data from the World Bank's *Doing Business* study." (http://www.heritage.org/Index/pdf/Index09_Methodology.pdf). Data are from 2012.	Heritage Foundation/ World Bank	http://www.heritage.org/index/explore

Table A3. Continued

Institutional variable	Description	Source of data	Data availability
Tech Absorption	Firm-level technology absorption capability: "Companies in your country are (1 = not able to absorb new technology, 7 = aggressive in absorbing new technology)"	World Economic Forum	The Global Competitiveness Report 2013-2014, p. 511
Staff Training	The extent of staff training: "To what extent do companies in your country invest in training and employee development? (1 = hardly at all; 7 = to a great extent)"	World Economic Forum	The Global Competitiveness Report 2013-2014, p. 467
Market Dominance	Extent of market dominance: "Corporate activity in your country is (1 = dominated by a few business groups, 7 = spread among many firms)"	World Economic Forum	The Global Competitiveness Report 2013-2014, p. 471
Technology Transfer	These are the innovation index points from GCI: a complex measure of innovation, including investment in research and development (R&D) by the private sector, the presence of high-quality scientific research institutions, the collaboration in research between universities and industry, and the protection of intellectual property	World Economic Forum	The Global Competitiveness Report 2013-2014, p. 22
GERD	Gross domestic expenditure on R&D (GERD) as a percentage of GDP, year 2012 or latest available data; Puerto Rico, Dominican Republic, United Arab Emirates, and some African countries are estimated using regional or nearby country data.	UNESCO	http://stats.uis.unesco.org/unesco/TableViewer/tableView.aspx?ReportId=2656
Business Strategy	Refers to the ability of companies to pursue distinctive strategies, which involves differentiated positioning and innovative means of production and service delivery	World Economic Forum	The Global Competitiveness Report 2013-2014, p. 22
Globalization	A part of the Globalization Index measuring the economic dimension of globalization. The variable involves the actual flows of trade, foreign direct investment, portfolio investment, and income payments to foreign nationals, as well as restrictions of hidden import barriers, mean tariff rate, taxes on international trade, and capital account restrictions. Data are from the 2013 report and based on the 2011 survey. http://globalization.kof.ethz.ch/	KOF Swiss Economic Institute	Dreher, Axel, Noel Gaston and Pim Martens (2008), *Measuring Globalisation – Gauging its Consequences* (New York: Springer).
Depth of Capital Market	The depth of capital market is one of the six sub-indices of the Venture Capital and Private Equity Index. This variable is a complex measure of the size and liquidity of the stock market, level of IPO, M&A, and debt and credit market activity. Note that there were some methodological changes over the 2006-2013 time period, so comparison to previous years is not perfect. The dataset is provided by Alexander Groh.* For missing data nearby country data used. For countries having estimated individual data, DCM data are the same way as it is in the case of individual variables (see Table 2 last column)	EMLYON Business School, France and IESE Business School, Barcelona, Spain	Groh, A, H. Liechtenstein and K. Lieser. (2012). The Global Venture Capital and Private Equity Country Attractiveness Index 2012 Annual, http://blog.iese.edu/vcpeindex/about/

*Special thanks for Alexander Groh and his team about the provision of the Depth of Capital Market data.

Estimation of the GEDI index

The GEDI scores for all the countries are calculated according to the following eight points.

1 ***The selection of variables:*** We start with the variables 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) that are coming from the GEM Adult Population Survey or the institutional/environmental level that are coming from various other sources. Individual variables for a particular year is calculated as the two year moving average if a country has two consecutive years individual data, or single year variable if a country participated only in the particular year in the survey. Institutional variables reflect to most recent available data in that particular year. Altogether we use 16 individual and 15 institutional variables (For details see Appendix A).

2 ***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.

$$z_{i,j} = ind_{i,j} \; x \; ins_{i,j} \qquad \text{(A1)}$$

for all *j*=1 ... k, the number of pillars, individual and institutional variables
where $z_{i,j}$ is the original pillar value for the *ith* country and pillar *j*
$ind_{i,j}$ is the original score for the *ith* country and individual variable *j*
$ins_{i,j}$ is the original score for the *ith* country and institutional variable *j*

3 ***Normalization:*** pillars values were first normalized to a range from 0 to 1:

$$x_{i,j} = \frac{z_{i,j}}{\max z_{i,j}} \qquad \text{(A2)}$$

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

4 ***Capping:*** All index building is based on a benchmarking principle. In our case we selected the 95 percentile score adjustment meaning that any observed values higher than the 95 percentile is lowered to the 95 percentile. While we used only 63 country values, the benchmarking calculation is based on all the 425 data points in the whole 2006-2013 time period.

5 ***Average pillar adjustment:*** The different averages of the normalized values of the pillars imply that reaching the same pillar values require different effort and resources. Since we want to apply GEDI for public policy purposes, the additional resources for the same marginal improvement of the indicator values should be the same for all indicators. Therefore, we need a transformation to equate the average values of the components. Equation A3 shows the calculation of the average value of pillar j:

$$\bar{x}_j = \frac{\sum_{i=1}^{n} x_{i,j}}{n}. \tag{A3}$$

We want to transform the $x_{i,j}$ values such that the potential minimum value is 0 and the maximum value is 1:

$$y_{i,j} = x_{i,j}^{k} \tag{A4}$$

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

$$\sum_{i=1}^{n} x_{i,j}^{k} - n\bar{y}_j = 0 \tag{A5}$$

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

$$\bar{x}_j < \bar{y}_j \quad k < 1$$
$$\bar{x}_j = \bar{y}_j \quad k = 1$$
$$\bar{x}_j > \bar{y}_j \quad k > 1$$

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

6 ***Penalizing:*** After these transformations, the PFB methodology was used to create indicator-adjusted PFB values. We define our penalty function following as:

$$h_{(i),j} = min\, y_{(i),j} + \mathrm{a}(1 - e^{-\mathrm{b}(y_{(i)j} - min\, y_{(i),j})}) \tag{A6}$$

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

7 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 a 100. The maximum value of the sub-indices is 100 and the potential minimum is 0, both of which reflect the relative position of a country in a particular sub-index.

$$ATT_i = 100 \sum_{j=1}^{5} h_j \qquad \text{(A7a)}$$

$$ABT_i = 100 \sum_{j=6}^{9} h_j \qquad \text{(A7b)}$$

$$ASP_i = 100 \sum_{j=10}^{14} h_j \qquad \text{(A7c)}$$

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

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

$$GEDI_i = \frac{1}{3}(ATT_i + ABT_i + ASP_i) \qquad \text{(A8)}$$

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