Evaluating countries' innovation potential: an international perspective

Panagiotis Tziogkidis[a*], Dionisis Philippas[b], Alexandros Leontitsis[c], and Robin C. Sickles[d]

[a] Plymouth Business School, Cookworthy Building, Drake Circus, PL4 8AA, Plymouth, UK, panagiotis.tziogkidis@plymouth.ac.uk.

[b] ESSCA School of Management, 55 Quai Alphonse Le Gallo, 92513, Boulogne, Paris, France, dionisis.philippas@essca.fr.

[c] Dubai Data Establishment, PO Box 555995, Dubai, UAE, alexandros.leontitsis@smartdubai.ae.

[d] Department of Economics, Rice University, 6100 Main Street, Houston, Texas 77005 USA, rsickles@rice.edu.

Abstract

The paper proposes a novel two-step approach that evaluates countries' innovation efficiency and their responsiveness to expansions in their innovation inputs, while addressing shortcomings associated with composite indicators. Based on our evaluations, we propose innovation policies tailored to take into account the diverse economic environments of the many countries in our study. Applying multidirectional efficiency analysis on data from the Global Innovation Index, we obtain separate efficiency scores for each innovation input and output. We then estimate different sensitivities for each country, by applying partial least squares on explanatory and response matrices which are determined by the nearest neighbours of the country under consideration. The findings reveal substantial asymmetries with respect to innovation efficiencies and sensitivities, which is indicative of the diversity of national innovation systems. Considering these two dimensions in combination, we outline three policy directions that can be followed, offering a platform for better-informed decision-making.

Keywords: multi-directional efficiency analysis; data mining; partial least squares regression; nearest neighbours; Global Innovation Index

JEL classification: C44, O30, O38, O57

[*] *Corresponding author*

1. Introduction

Innovation is successful when new knowledge is developed, adopted and disseminated, both within a country, as well as across borders (Cassiman and Veugelers, 2002). An important dimension that is often overlooked is innovation efficiency (Cruz-Cázares et al., 2013), which has received more attention since the financial crisis (OECD, 2012). Niosi (2002) attributes innovation inefficiency to a number of factors, including bounded rationality, transactions costs and lock-in situations. Despite the potential benefits of successful innovation, misallocating resources in search of innovative business practices and processes is detrimental to firms, since a large share of innovation investments are financed internally (Hall and Lerner, 2010). After the world recession in the late 2000's many firms cut back on their R&D investments since they viewed these as discretionary expenditures. During this period firms also faced substantial difficulties in obtaining external financing. Innovation was mainly concentrated in established and well-resourced firms. All of these factors have contributed to the declining investments that have been observed in R&D projects (Archibugi et al., 2013) and as the particular circumstances and the institutional constraints to investments in innovation are not well-understood, identifying how innovation resources can be optimally allocated is an important policy concern for governments and an important competitive insight for key industrial sectors and for firms therein. Public financing can help firms to sustain innovation investments, as Paunov (2012) shows for countries in Latin America, while it also has an active role in supporting risky innovations (Mazzucato and Semieniuk, 2017). This creates an added imperative for the efficient use of scarce public resources, which are often misallocated when outcomes are erroneously viewed as easily observable or when political expediency obscures the difficulties in obtaining outcomes that are not easily measured. Despite the fact that various approaches for measuring innovation efficiency have been proposed, two important elements are often missing, at least in combination: (1) accounting for the diversity of national innovation systems (NIS), which makes benchmarking or ranking winners and losers problematic; (2) evaluating the responsiveness of innovation outputs to innovation-related investments. This paper considers both of these important elements.

Policymakers often use composite indicators such as the Global Innovation Index, the European Innovation Scoreboard, the Innovation Capability Index, or the Global Competitiveness Index, to measure countries' innovation performance, where individual innovation indicators are aggregated into a single score. The most comprehensive composite innovation index in terms of country coverage and the spectrum of indicators used is the Global Innovation Index (GII), where the analysis is complemented with the evaluation of countries' innovation efficiency ratio (IER). Measuring innovation efficiency through composite indicators comes with limitations arising from the use of simple arithmetic or geometric averages (Fusco, 2015); most importantly, the diversity of NIS is not adequately captured, as we explain in detail in later sections. Data envelopment analysis (DEA) has been used to address this issue by measuring the innovation efficiency of countries or regions (see Zabala-Iturriagagoitia et al., 2007; Chen and Guan, 2012; Carayannis et al., 2016; Han et al., 2016), and to evaluate the efficiency of innovation systems (Nasierowski and Arcelus, 2003; Guan and Chen, 2012). It also has been used to calculate weights in the construction of composite indicators (see Cherchye et al., 2004; Despotis, 2005; Kao et al., 2008). In fact, the GII report includes a robustness check that compares the IER-based rankings with country rankings based on the constant returns to scale DEA model of Charnes et al. (1978).[1] However, even if weights are determined through optimizing models, such as DEA, instead of using simple arithmetic or geometric averages, such models often do not address the added problem of compensability. As an example, the DEA model used in the robustness exercise of the GII is not free from this issue.

Our paper proposes a novel two-step framework to tailor innovation policies suitable for each country while providing modifications that address the problems we have pointed out above. The analysis is based on the relationship between a country's sensitivity to innovation-related investments and the associated resource misallocation. To account for the latter, we obtain innovation efficiency scores using the multi-directional efficiency analysis (MEA) approach (Bogetoft and Hougaard, 1999; Asmild et al., 2003), explained in detail in section 3. Since MEA is a directional model of efficiency and given the fact that it imposes a different directional vector for each country, it is better suited to deal with the aforementioned compensability issue (Fusco, 2015).[2] Moreover, the fact that separate efficiency scores are calculated for each innovation input and output variable, makes MEA more effective in capturing the diversity of national innovation systems (NIS).

In order to evaluate the responsiveness (*sensitivity*) of innovation outputs to changes in innovation inputs for each country, we propose an iterative multivariate regression approach where the sample of countries varies according to a clustering algorithm, outlined in detail in section 3. In particular, we apply partial least squares regression (PLS) (Wold et al., 1984, Höskuldsson, 1998) on the innovation input (explanatory) and output-related (response) variables, since PLS regression is designed for multivariate systems.[3] PLS is one of many dimension reduction methods used in data mining and is

---

[1] The DEA problem in this comparison is modified somewhat to ensure that each input has a neither a non-negligible nor a dominant weight, the former weight restriction imposed in order to "*…preclude the possibility of a country achieving a perfect score by assigning a zero weight to weak pillars*." For more on this see GII (2016; pp.72).

[2] The issue of compensability is of concern in applications where optimising algorithms are used to calculate alternative scores for composite indicators, compared to the ones resulting from simple averaging and linear aggregation. The problem of perfect compensability also extends to the benefit of the doubt (BoD) approach, which calculates weights that reflect trade-offs between variables. Including weight restrictions (as in the robustness exercise in the GII) deals with the issue partially. However, it does not completely offset it as the resulting weights still represent common trade-offs for all units in the restricted space. Fusco (2015) introduces a directional BoD model which includes a "directional penalty", which simply replaces the radial DEA efficiency measure with a directional one. MEA is one of the possible approaches for choosing the directional vector.

[3] It is important to note that PLS regression, used in this paper, and PLS path modelling, are two distinct branches that share similar origins but are associated with substantially different algorithms. They should not be confused. In this paper, we use the abbreviation PLS to refer to the former.

much like principle components analysis (PCA) in that it can deal with systems that may suffer from collinearity. PLS has the additional benefit that it can account for latent factors (Wold et al., 2001), such as the ones underpinning the transformation of innovation inputs into outputs. To obtain country-specific sensitivities, we run one PLS regression for each country in the sample, while varying the nearest neighbours (*peer countries*) of the reference country. Given that we define the neighbourhood in terms of three economic variables (R&D as a percentage of GDP, FDI inflows as a percentage of GDP and trade openness) our contribution translates to incorporating the notion of economic proximity in innovation policy evaluation, which extends beyond geographical or income boundaries. Finally, using the resulting relationships between the sensitivities and the innovation efficiency scores, our analysis can point out the optimal policy direction that a country should follow: *innovation facilitating*, *innovation-improving* or *hybrid*.

To anticipate empirical results discussed in section 4, we find substantial asymmetries in innovation efficiencies. For example, high-income countries in Europe are more innovation-efficient on average in most aspects of their innovation inputs and outputs than high-income countries in other parts of the world. Also, low-income countries are less innovation-efficient in generating innovation outputs related to creative goods and services and intangible assets. The estimated sensitivities show that countries are more responsive to investments in human capital and research, whereas the lowest average responsiveness is found for the composite index that captures the state of knowledge linkages and transfers in the business community. Moreover, we observe substantial diversity in the sensitivities that different countries exhibit to changes in the various innovation inputs. These do not seem to be related to income or geography. Thus, jointly considering innovation efficiency and output sensitivities to inputs, adds a new dimension to the empirical analysis of innovation. Finally, our results indicate that the optimal innovation policy that a country should follow can differ with respect to the spectrum of input-output combinations, suggesting that a more tailored approach is necessary in order to identify the most promising areas for innovation investments and avoid misallocating resources.

The remainder of the paper is organized as follows. Section 2 outlines how the Global Innovation Index is constructed and discusses in more detail the issues with its use that we have alluded to above. Section 3 outlines the proposed two-step framework we employ in our analysis and discusses other methodological and technical aspects. Section 4 describes the data, presents the empirical results, and discusses implications for policy design. Section 5 concludes. More detailed results of our analyses are provided in the accompanying Supplement.

2. Composite indicators of innovation-related activities

The Global Innovation Index (GII) (2016) assesses the innovation performance of 128 countries through the aggregation of 82 indicators. The aim of the GII is to facilitate policy-making by identifying a country’s relative ranking vis-à-vis other countries in terms of these different indicators.

The indicators are aggregated into groupings that are referred to as "pillars". The first level of aggregation combines the 82 indicators into sub-pillars and these are subsequently aggregated into the five input pillars (*Institutions, Human Capital & Research, Infrastructure, Market Sophistication, Business Sophistication*) and the two output pillars (*Knowledge & Technology* and, *Creative Outputs*) that we utilize in our analyses. The input pillars are averaged to derive the Innovation Input Sub-Index, while averaging the output pillars generates the Innovation Output Sub-Index. Finally, the GII score is computed as the simple average of the Innovation Input and Output Sub-Indices, while the Innovation Efficiency Ratio (IER) is determined by the ratio of the Innovation Output Sub-Index to the Innovation Input Sub-Index. Country rankings based on the GII score differ relative to those based on the IER, as the GII focuses on the magnitude of innovation activities while the IER evaluates the extent to which the available innovation resources have been productively utilized.

A common issue with composite indicators that follows from this process of linear aggregation is that the resulting weights attached to the indicators are compensatory (scaling coefficients), in that low values in one indicator would be perfectly offset by high values in another, without affecting the associated scores. In our context this would suggest, for example, that an increase in *Creative Outputs* could compensate for a decrease in the *Knowledge & Technology* output, leaving the associated metrics (the Innovation Output Sub-Index, and therefore the GII and the IER scores) unaffected. This is arguably somewhat counterintuitive. The problem of compensability also extends to multicriteria models, if not treated appropriately, as well as the standard DEA-based benefit of the doubt (BoD) approach (Cherchye et al., 2004), which is used in the robustness exercise of the GII. In the latter case, directional distance functions have been used to address the issue of compensability (Fusco, 2015; Vidoli et al., 2015).

A related issue, shared by simple averaging and conventional BoD and DEA models, is that they do not fully account for the diversity of national innovation systems (NIS). This is clear for the case of simple averaging since weights are given *a priori*. Regarding DEA-based methods, despite the fact that they can signal the importance of particular variables by assessment of their relative weights generated from an optimization problem, their effectiveness is reduced when radial measures of efficiency are used. This is because innovation inefficiencies determined by conventional BoD and DEA models would indicate equiproportional improvements in inputs and/or outputs, as the calculated weights are consistent with this interpretation. However, the well-documented diversity of NIS is multifaceted and can be influenced by various factors such as capabilities, resources, the environment or the pace of technological change (see Freeman, 1987; Romer, 1990; Grupp and Schubert, 2010; Watkins et al., 2015). Therefore, there are asymmetries in countries' preferences and priorities with respect to innovation that need to be taken into account.

The sources of these asymmetries can be evaluated by examining a country's potential improvements in each dimension considered. We carry out such an analysis using the *multi-directional efficiency analysis* (MEA) model, which estimates directional vectors of potential efficiency improvements with respect to each innovation input and output. MEA, being a directional measure of efficiency, is well-suited for dealing with the compensability issue, while it is more effective compared to other approaches in capturing the diversity of NIS since it allows each country to optimise in each input-output dimension separately.

Another consideration relates to how to plausibly evaluate a country's responsiveness to innovation-related investments, also taking into account the dynamism of the environment. Some countries may be exposed to a substantially less conducive environment to innovation and their ability to innovate or the feasibility of achieving certain targets of innovation activity may be overstated, if the influences of the environment are not taken into account. For example, the GII report classifies countries as 'underperformers', 'achievers' and 'leaders' (Figure 4, GII 2016, p.32) without considering the proximity of their economic environments, which is influential in the development and adoption of innovations. Traditionally, geographical proximity has been the leading factor given that endowments for innovation tend to co-exist in areas with high concentration of good quality physical and human capital, but this is neither a necessary nor a sufficient condition for innovation (Boschma, 2005). There is an extensive literature that supports the roles of foreign direct investments, R&D expenditure and trade as drivers of innovation (see Acemoglu, 2002; OECD, 2005; Onodera, 2008; Mahroum and Al-Saleh, 2013). Given that many of these factors are not adequately emphasized in the GII framework, a more inclusive role for them is provided in this paper.

To account for the role of the environment and obtain country-specific estimates of sensitivity to innovation investments, we propose a novel approach that utilises the nearest neighbours methodology within a multivariate regression framework. The nearest neighbours of each country are determined by its *economic proximity* with other countries, which we define with respect to three dimensions conducive to innovation.[4] The first dimension is investment on R&D within a country, which is a well-documented prerequisite for the development of innovations (OECD, 2015). The second dimension relates to knowledge and technology transfers and spillovers, which are facilitated through foreign direct investments (FDI). FDI contributes to innovation-related activities by financing investments through capital transfers, while developing countries benefit when investments are directed to them from developed countries (OECD, 2002), while He and Maskus (2012) show in their theoretical model that reverse spillovers are also possible. The third dimension is trade openness, which is commonly used as a measure of economic distance (Glass et al., 2016). Trade expands the potential market size and provides incentives to innovate due to product market integration and intensified competition (Grossman and Helpman; 1990 and 1994). A reduction in trade barriers,

[4] We decided to include only supply-based measures given that we are assessing a transformation relationship of innovation inputs to outputs and given the fact that policymakers can exert some control over them. Demand-based factors can be directly or indirectly related to outcomes of the innovation process and we wish to avoid possible endogeneity issues that may be introduced were these factors included in a more structurally focused study which we do not undertake in this paper.

therefore, promotes process innovations as firms need to improve their productive and cost efficiency in order to survive global competition and bear the costs of exporting (Atkeson and Burstein, 2010; Desmet and Parente, 2010; Long et al., 2011). The empirical evidence also shows that such a reduction in trade barriers reallocates skilled labour towards technologically advanced firms (Bloom et al., 2016).

3. Methodology

This section presents the two-step framework of our study. The first step uses the multi-directional efficiency analysis (MEA) approach to obtain non-radial efficiency scores with respect to each input-output dimension. In the second step, we introduce a new approach to estimate country-specific sensitivities of innovation outputs to changes in innovation inputs.

3.1 Multi-directional efficiency analysis

We measure innovation efficiency using the *multi-directional efficiency analysis* (MEA) model (Bogetoft and Hougaard, 1999; Asmild et al., 2003), which builds on the framework of data envelopment analysis (DEA). MEA is a directional efficiency measurement approach that assesses countries' *potential improvements* in each dimension and, therefore, it is possible to obtain separate efficiency scores for each innovation input and output.[5] The resulting efficiency scores indicate areas of improvement in each dimension, therefore revealing asymmetric patterns of innovation efficiency across countries. Such asymmetries are indicative of the diversity of national innovation systems (NIS) and of the differences in national priorities. The choice of a non-parametric technique also finds support in Niosi (2002), who formalised the concept of X-inefficiency for NIS. Given the above and the fact that MEA can deal with the issue of compensability, as previously explained, it is a suitable model to be used in this context.

Consider a set of n decision making units (DMUs) or countries where $i = 1, \dots n,$ that use p inputs $(j = 1, \dots p)$ to produce q $(r = 1, \dots q)$ outputs. Let DMU k have a production plan $(\boldsymbol{x}_k, \boldsymbol{y}_k)$, where $\boldsymbol{x}_k = \left(x_{k,1}, \dots x_{k,p}\right)$ and $\boldsymbol{y}_k = \left(y_{k,1}, \dots y_{k,q}\right)$. In the first step, we calculate potential improvements in

[5] To examine the sensitivity of MEA scores to sampling variations we implemented the *m/n* bootstrap for the case of directional distance functions, in the spirit of Simar et al. (2012). Given the fact that the computational costs for implementing the *m/n* bootstrap on MEA are significantly greater than in the case of directional distance functions (greater by $2 \cdot (p + q)$ times, where p and q are the number of inputs and outputs, respectively), and given the fact that bootstrap methodologies have not been developed for MEA, we only performed the necessary computations for a limited number of block sizes. In particular, we determined the optimal block size along the lines of Politis et al. (2001), but only within a limited range of block sizes, for which we consulted the simulation results of Kneip et al. (2008). Our results indicate that the resulting confidence intervals are reasonably narrow, while the rank correlations between the MEA efficiency scores and the bias-corrected ones are very high (above 0.98). The results of this exercise can be found in the accompanying Supplement.

inputs and outputs for each DMU. We start by defining the *ideal reference point* for the k^{th} DMU, denoted as $(\boldsymbol{x}_k^*, \boldsymbol{y}_k^*) = \left(x_{k,1}^*, \dots x_{k,p}^*, y_{k,1}^*, \dots y_{k,q}^*\right)$. To determine the coordinates of the ideal reference point, we use as many linear programs as dimensions. Consistent with the literature on composite indicators (see Kao et al., 2008; Guen and Chan, 2012), we assume constant returns to scale (CRS). Moreover, we use a non-oriented model given that the innovation outputs are in principle controllable, while innovation inputs require significant investments and effort to be developed and sustained to the desirable level. Instead of interpreting inefficiencies as required input contractions or output expansions, we use them to identify resource misallocations in the innovation process. Taking also into account that multiple indicators are included in the construction of the input and output pillars in the GII framework, the calculated inefficiencies are likely to persist. In this context, innovation inefficiency could be regarded as a form of risk. For the j^{th} input of DMU k, we employ the following linear program:

$$\begin{aligned}
&\min_{\lambda, x_{k,j}^*}\left\{x_{k,j}^*\right\} \quad \text{s.t.} \\
&\sum_{i=1}^{n} \lambda_i x_{i,j} \leq x_{k,j}^* \\
&\sum_{i=1}^{n} \lambda_i x_{i,-j} \leq x_{k,-j} \quad \text{for} \quad -j = 1, \dots j-1, j+1, \dots p \\
&\sum_{i=1}^{n} \lambda_i y_{i,r} \geq y_{k,r} \quad \text{for} \quad r = 1, \dots q \\
&\lambda_i \geq 0
\end{aligned} \tag{1}$$

For the r^{th} output of DMU k, we have:

$$\begin{aligned}
&\max_{\lambda, y_{k,j}^*}\left\{y_{k,r}^*\right\} \quad \text{s.t.} \\
&\sum_{i=1}^{n} \lambda_i x_{i,j} \leq x_{k,j} \quad \text{for} \quad j = 1, \dots p \\
&\sum_{i=1}^{n} \lambda_i y_{i,r} \geq y_{k,r}^* \\
&\sum_{i=1}^{n} \lambda_i y_{i,-r} \geq y_{k,-r} \quad \text{for} \quad -r = 1, \dots r-1, r+1, \dots m \\
&\lambda_i \geq 0
\end{aligned} \tag{2}$$

The linear programs above identify the maximum potential improvements for all inputs and outputs, consecutively defining each coordinate of the ideal reference point for the DMU k. The ideal reference point may lie outside the feasible set, but it is only used to indicate the direction of improvement for the inputs of each DMU separately. If $(\boldsymbol{x}_k^*, \boldsymbol{y}_k^*) = \left(x_{k,1}^*, \dots x_{k,p}^*, y_{k,1}^*, \dots y_{k,q}^*\right) = (\boldsymbol{x}_k, \boldsymbol{y}_k)$, the DMU k utilises its inputs efficiently, while it produces the efficient level of outputs and, therefore, there is no scope for further improvement. However, if $(\boldsymbol{x}_k^*, \boldsymbol{y}_k^*) \neq (\boldsymbol{x}_k, \boldsymbol{y}_k)$, the DMU k should improve in the direction of the ideal reference point $(\boldsymbol{x}_k^*, \boldsymbol{y}_k^*)$.

Denote now the distance of the j^{th} input and r^{th} output of the ideal reference point from the corresponding observed ones as $d_{k,j} = x_{k,j} - x_{k,j}^*$ and $\delta_{k,j} = y_{k,r}^* - y_{k,r}$, respectively. In the second step, the proportion β of the distances ($d_{k,j}$ and $\delta_{k,j}$) from the ideal reference point are identified. If $\beta = 0$, DMU k is by definition efficient, whereas if $\beta > 0$, there are potential improvements in all directions. We estimate β with the following linear program:

$$\max_{\lambda,\beta}(\beta) \quad \text{s.t.}$$

$$\sum_{i=1}^{n} \lambda_i x_{i,j} \leq x_{k,j} - \beta d_{k,j} \qquad \text{for} \quad j = 1, \dots p \tag{3}$$

$$\sum_{i=1}^{n} \lambda_i y_{i,r} \geq y_{k,r} + \beta \delta_{k,r} \qquad \text{for} \quad r = 1, \dots q$$

$$\lambda_i \geq 0$$

The target level of inputs and outputs for DMU k, denoted as $\boldsymbol{x}_k^T = \left(x_{k,1}^T, \dots x_{k,p}^T\right)$ and $\boldsymbol{y}_k^T = \left(y_{k,1}^T, \dots t_{k,p}^T\right)$ respectively, are computed as:

$$\begin{aligned} x_{k,j}^T &= x_{k,j} - \beta_k d_{k,j} \\ y_{k,r}^T &= y_{k,r} + \beta_k \delta_{k,r} \end{aligned} \tag{4}$$

Following Asmild and Matthews (2012), we determine the relative efficiency scores for each input as the ratio of the target inputs over the actual ones, denoted as $\theta_{k,j} = \left(x_{k,j}^T / x_{k,j}\right)$ for each DMU. Similarly, the relative efficiency scores for each output are determined as $\theta_{k,r} = \left(y_{k,r} / y_{k,r}^T\right)$. Finally, we derive an aggregate measure of MEA efficiency using the following aggregation proposed by Tone (2001):

$$\rho_k = \left(1 - \frac{1}{p} \sum_{j=1}^{p} \frac{d_{k,j}\beta_k}{x_{k,j}}\right) \Bigg/ \left(1 + \frac{1}{q} \sum_{r=1}^{q} \frac{\delta_{k,r}\beta_k}{y_{k,r}}\right) \tag{5}$$

3.2 Sensitivity conditional on economic proximity

The second step of our framework estimates country-specific sensitivities of innovation outputs to changes in innovation inputs. This is operationalised by implementing a multivariate regression framework (PLS regression) where the logs of innovation outputs are regressed on the logs of innovation inputs, while conditioning on each country's nearest neighbours (or peer countries).[6] Conceptually, our approach is closer to the conditional efficiency literature (see Bădin et al., 2010;

[6] It is important to note that the estimated sensitivities are the expected responses of outputs to inputs within each country's set of nearest neighbours. Therefore, these partial derivates cannot be used for the characterisation of returns to scale and they should not be confused with scale elasticities.

Daraio and Simar, 2014), but with the difference that they examine the influence of environmental factors on efficiency, instead. Methodologically, the closest approach to ours is that of Guen and Chan (2012), but with the fundamental difference that they apply a PLS regression on efficiency scores to address a different research question. Therefore, our paper differs both in terms of research objectives and the approach it implements to address them.

PLS is a multivariate regression approach that can be used to model the relationship between a response matrix and an explanatory matrix. It belongs to the same family of models as principal components regression, canonical correlation, and ridge regression, and it is more suitable when the column-wise (variables) correlations are high or when one does not want to impose assumptions on the distribution of the error term. PLS regression has been commonly used in data-driven applications in chemometrics (e.g. Tenenhaus et al. 2005), data mining (e.g. Gersende and Lambert-Lacroix, 2005), machine learning (e.g. Khedher et al., 2015), or other data analytics applications (e.g. Yang et al., 2011). We choose to use PLS over alternative suitable candidates, as it can estimate different sensitivities for each input-output combination and, therefore, allows us to evaluate whether investments in certain innovation inputs would find greater response in *Knowledge & Technology* outputs or *Creative Outputs*.[7]

Following Wold et al. (2001) and in line with the notation in subsection 3.1, denote with $\boldsymbol{X}$ the $n \times p$ explanatory matrix containing the logs of innovation inputs and with $\boldsymbol{Y}$ the $n \times q$ response matrix of the logs of innovation outputs. The $\boldsymbol{X}$ and $\boldsymbol{Y}$ matrices are decomposed into orthogonal components as:

$$\boldsymbol{X} = \boldsymbol{T}\boldsymbol{P}' + \boldsymbol{E} \quad \text{and} \quad \boldsymbol{Y} = \boldsymbol{U}\boldsymbol{C}' + \boldsymbol{G} \tag{6}$$

where $\boldsymbol{T}$ and $\boldsymbol{U}$ are score matrices, $\boldsymbol{P}$ and $\boldsymbol{U}$ are factor loading matrices, while $\boldsymbol{E}$ and $\boldsymbol{G}$ are residual matrices, reflecting unexplained variability. PLS regression establishes a linear model that maximises the covariance between the components of $\boldsymbol{Y}$ and $\boldsymbol{X}$:

$$\boldsymbol{Y} = \boldsymbol{T}\boldsymbol{C}' + \boldsymbol{F} = \boldsymbol{X}\boldsymbol{B} + \boldsymbol{F} \tag{7}$$

where $\boldsymbol{F}$ is a matrix of Y-residuals between observed and estimated responses and $\boldsymbol{B}$ is the matrix of PLS regression coefficients. Let β_{ij} be the estimated PLS coefficient of the correspondence between the i^{th} innovation input and the j^{th} innovation output. Since $\boldsymbol{X}$ and $\boldsymbol{Y}$ are expressed in logs, the estimated coefficients reflect the sensitivity (responsiveness) of innovation outputs to changes in innovation inputs. Hence, the greater (smaller) the value of β_{ij}, the greater (smaller) the responsiveness of innovation output j to investments in innovation input i.

[7] Multivariate least squares would only generate as many estimates as explanatory variables. Similarly, when the explanatory variables are common for all response variables, as in our case, the estimated coefficients from a seemingly unrelated regressions system are as many as the (common) explanatory variables.

In our paper, we adjust this framework to account for the influences of the environment, defined through a multidimensional space of user-set variables (here *economic proximity*). To do so, we first identify the *nearest neighbours* (*peers*) of each country, defined as those γ countries that exhibit the smallest squared Euclidian distance from the reference country, in terms of their environmental variables. To determine the *optimal* number of nearest neighbours (denoted as γ^*), we try different values for γ, ranging from $p+1$ to $n-p-1$. Each time, we construct for a reference country k, a $\gamma \times p$ explanatory matrix X_γ^k and a $\gamma \times q$ response matrix Y_γ^k, which contain the innovation inputs and outputs of the *nearest neighbours* of the reference country. We apply the PLS regression of Y_γ^k on X_γ^k to obtain a $\gamma \times q$ matrix of residuals and calculate their element-wise sum of squares $SSE_\gamma^k = \sum_{i=1}^{n} \sum_{j=1}^{q} e_{ij}^2$. We repeat this process for every country and calculate the aggregate sum of squares as $SSE_\gamma = \sum_{i=1}^{n} SSE_\gamma^i$. The optimal number of *nearest neighbours* γ^* is the one that returns the minimum SSE_γ. Our approach generates sensitivity estimates for each country and therefore the total number of estimated coefficients is $n \times p$ for each output, which are all determined from the n PLS regressions corresponding to the γ^* number of neighbours.

4. Empirical analysis and policy implications

This section presents the empirical findings and discusses the policy implications arising from our research hypotheses. We start by providing an overview of the data along with some first insights. We then analyse the MEA innovation efficiency scores for each region and income group. Next, we discuss the results arising from the PLS regression and propose policy directions that consider countries' MEA innovation efficiencies and sensitivities (PLS coefficients), simultaneously.

4.1 Data

We obtain the input-output data from the GII (2016) report and we use its five innovation input pillars as the input variables and its two innovation output pillars as the output variables. The data from the GII are normalised indices which have undergone a four-step process to ensure coherence (GII, 2016; p.61). In the first step, the conceptual consistency is examined, where candidate indicators are selected and innovation pillars are defined, based on the relevant literature. The second step involves checking whether the data, derived from a wide range of sources, conforms to requirements for availability and coverage, while also employing an appropriate treatment where needed. The next step involves determining the weights applied to each indicator, as well as grouping indicators into sub-pillars and subsequently into pillars and sub-indices. Moreover, the statistical coherence of the GII is assessed through principal component and reliability item analysis. In the last step, the overall results are reviewed qualitatively to assess their consistency with other evidence and research, concluding, though, that the GII framework is open for future development (GII, 2016, p. 64).

The input and output pillars that we use for our efficiency computations reflect different aggregated dimensions of innovation. Although it would be possible to reconstruct the pillars using alternative weighting strategies challenging the conceptual framework of the GII is beyond the scope of this study. Our aim is to show how the empirical findings in the GII report can be more informative by incorporating innovation efficiency and output responses in the analysis. We show later in the paper that these two dimensions are independent with each other and with the GII scores, suggesting that further disaggregation would not be necessary to achieve our goals.

The logarithms of the innovation inputs and outputs are then used in our PLS framework. The nearest neighbours (*peers*) are determined for each country through three economic variables (*economic proximity*). The first variable we use is *R&D expenditure* (% GPD), in order to account for the level of R&D investments in a country. The second variable is *FDI net inflows* (% GDP), emphasising the benefits to the host country from inward investments. Finally, to account for trade, we use *trade openness* which is defined as the ratio of the sum of a country's imports and exports to its GDP. The data for the economic variables are obtained from the World Bank and we use the latest available observations. Our methodological framework is flexible for policy experimentation, in that a different set of qualitative or quantitative variables or dimensions of proximity can be accommodated to suit policy objectives.

Table 1 provides a summary of the variables used in our analysis. There is substantial variability in the input-output variables, which largely relates to geography given the asymmetries observed in the regional averages. The *Human Capital & Research* pillar appears with a relatively low value compared to the other input pillars, mostly in low-income regions, such as Sub-Saharan Africa. Moreover, only few countries have environments conducive to innovation, reflected in the positively skewed and leptokurtic distributions of the three variables that we use to define economic proximity. The substantial variability that countries exhibit with respect to these three variables is also indicative of the heterogeneity of their economic environments, further justifying the use of a clustering algorithm. We also find high correlations between innovation inputs and outputs which may introduce an element of double counting under equal weighting, further supporting the use of optimisation-based approaches such as MEA. The high correlations also further justify the use of regression approaches such as PLS, which can deal with multicollinearity. Finally, given that R&D (% GDP) and FDI inflows (% GDP) are indicators within the *Human Capital & Research* and *Business Sophistication* pillars, respectively,[8] we test for endogeneity in these dimensions using the DWH (Durbin–Wu–Hausman) augmented regression test (Davidson and MacKinnon, 1993). The case of endogeneity is rejected for both environmental variables and with respect to both outputs.

[8] In particular, "Government Expenditure on R&D (% GDP)" is one of the three indicators within one of the three sub-pillars of the *Human Capital & Research* input pillar. Also, "FDI net inflows (% GDP)" is one of the four indicators within one of the three sub-pillars that comprise the *Business Sophistication* pillar.

Table 1. Summary statistics

	Institutions	Human Capital & Research	Infrastructure	Market Sophistication	Business Sophistication	Knowledge & Technology	Creative Outputs	R&D (% GDP)	FDI inflows (% GDP)	Trade Openness
Descriptive Statistics										
Mean	63.10	33.91	43.36	45.37	33.55	27.59	31.63	0.89	5.36	0.88
Standard deviation	16.35	15.48	13.09	11.63	10.74	12.64	13.99	0.95	10.17	0.63
Skewness	0.09	0.41	-0.02	0.95	0.60	0.85	0.43	1.60	4.21	2.58
Kurtosis	-0.56	-0.65	-0.88	1.05	-0.17	0.35	-0.10	2.08	20.86	9.52
Regional Averages										
Central and Southern Asia	49.46	24.83	35.01	40.62	25.11	19.93	20.73	0.33	3.81	0.55
Europe	76.01	46.74	52.61	49.81	40.39	37.57	43.62	1.44	6.58	1.17
Latin America and the Caribbean	52.93	26.28	40.14	42.29	30.76	18.09	26.10	0.29	4.30	0.54
Northern America	88.70	54.95	62.00	80.10	49.45	48.70	49.35	2.17	2.85	0.47
Northern Africa and Western Asia	60.43	32.01	44.39	42.11	27.12	24.22	28.73	0.60	4.21	0.81
South East Asia and Oceania	69.70	42.99	50.88	56.93	41.50	36.06	37.48	1.33	8.06	1.24
Sub-Saharan Africa	52.42	17.98	28.21	35.92	27.56	18.41	18.98	0.44	4.46	0.69
Correlations										
Institutions	1									
Human Capital & Research	0.755	1								
Infrastructure	0.817	0.851	1							
Market Sophistication	0.695	0.732	0.701	1						
Business Sophistication	0.727	0.762	0.708	0.697	1					
Knowledge & Technology	0.712	0.810	0.740	0.701	0.814	1				
Creative Outputs	0.783	0.765	0.786	0.675	0.739	0.805	1			
R&D (% GDP)	0.635	0.799	0.672	0.650	0.786	0.813	0.676	1		
FDI inflows (% GDP)	0.245	0.148	0.146	0.224	0.328	0.272	0.252	0.003	1	
Trade Openness	0.448	0.299	0.321	0.230	0.418	0.349	0.408	0.124	0.618	1

Notes: The first section of the table present the mean, the standard deviation, skewness and kurtosis for the input and output sub-indices as well as the environmental variables used in this study. The second and third sections report the respective regional averages and correlation coefficients.

4.2 Patterns of innovation efficiency

Table 2 summarises our findings on innovation efficiency per income group and within each region. We use a colour map to reflect the quartile of innovation efficiency that each group corresponds to; the darker the colour shading, the lower the quartile. Our findings reveal substantial differences across regions and income groups. Considering the two innovation outputs, we find that, in principle, the respective efficiency scores are high for Europe, Northern America as well as South-Eastern Asia and Oceania (SEAO), which mainly include high-income countries. A balanced score, but of a lower magnitude, is observed in Northern Africa and Western Asia (NAWA), as well as Central and Southern Africa (CSA). On the contrary, Sub-Saharan Africa (SSF), the region that comprises low-income countries (except for Nepal), appears considerably less efficient in *Creative Outputs* compared to *Knowledge & Technology*, while the opposite is true for Latin America and the Caribbean (LCN).[9]

Regarding innovation inputs, we find that most countries exhibit a relatively low efficiency score for the *Human Capital & Research* pillar, which is indicative of the spread of performance in a dimension with cited importance for innovation and economic growth. Although this underperformance appears across income groups and regions, countries in Europe, Northern America and SEAO seem to be less affected, in principle. Moreover, they also maintain a balanced performance across innovation inputs and outputs. Similar underperformance is also observed for the *Market Sophistication* pillar, which captures the credit, investment and competitive conditions in a country, and where high-income countries perform significantly better than others on average.

Comparing countries of the same income group across different regions reveals considerable variability in results. For example, European high-income groups perform better than their counterparts in other regions, based on the aggregated efficiency score (ρ_k), while the differences are not uniform with respect to each pillar. Similarly, high-income countries in NAWA, which are mostly oil-rich, perform at the bottom quartiles of innovation efficiency, except for the innovation-efficient Kuwait which is dominated by financial services and has a relatively open market. On the contrary, LCN is the region with the most cases of countries performing at the lower quartile levels, in all income groups. The substantial asymmetries in innovation efficiency that we find across income and geographic groups, suggest that different countries have different approaches and priorities when allocating innovation-related resources. On the one hand, inefficiencies may be associated with disproportionately large endowments that certain countries have and which should be used more productively. On the other hand, these findings may be associated with structural weaknesses of the countries such as high transactions costs, complicated bureaucratic procedures, or market frictions, among others, which policy should aim to remove (Niosi, 2002).

[9] The interested reader may refer to Figure S3 in the accompanying Supplement for a visual aggregation of the results per region and for each income group.

Table 2. Innovation efficiency colour map

	Institutions	*Human Capital & Research*	*Infrastructure*	*Market Sophistication*	*Business Sophistication*	*Knowledge & Technology*	*Creative*	*Eff(ρ_k)*
Central and Southern Asia								
Low	0.83	0.79	0.87	0.85	0.84	0.85	0.79	0.68
Lower - Middle	0.89	0.84	0.87	0.85	0.90	0.82	0.86	0.73
Upper - Middle	0.85	0.77	0.83	0.85	0.87	0.88	0.83	0.72
Europe								
Lower - Middle	0.98	0.95	0.99	0.97	0.99	0.99	0.98	0.97
Upper - Middle	0.84	0.77	0.86	0.85	0.87	0.86	0.78	0.69
High	0.93	0.88	0.93	0.92	0.94	0.93	0.93	0.86
Latin America and Caribbean								
Lower - Middle	0.85	0.76	0.85	0.77	0.81	0.73	0.85	0.64
Upper - Middle	0.88	0.81	0.86	0.83	0.87	0.80	0.87	0.71
High	0.85	0.79	0.84	0.85	0.88	0.84	0.86	0.72
Northern America								
High	0.92	0.87	0.91	0.86	0.92	0.93	0.90	0.82
Northern Africa and Western Asia								
Lower - Middle	0.90	0.85	0.88	0.88	0.93	0.93	0.86	0.79
Upper - Middle	0.88	0.84	0.88	0.87	0.91	0.90	0.89	0.79
High	0.87	0.82	0.86	0.86	0.90	0.85	0.88	0.76
South East Asia and Oceania								
Lower - Middle	0.94	0.91	0.93	0.89	0.93	0.95	0.94	0.87
Upper - Middle	0.93	0.90	0.93	0.90	0.91	0.93	0.91	0.85
High	0.88	0.80	0.88	0.85	0.89	0.87	0.85	0.74
Sub-Saharan Africa								
Low	0.85	0.82	0.90	0.86	0.85	0.85	0.73	0.65
Lower - Middle	0.94	0.91	0.94	0.91	0.93	0.94	0.93	0.87
Upper - Middle	0.80	0.76	0.85	0.80	0.88	0.78	0.83	0.65

Notes: The table presents the average innovation efficiency scores for each income group within each region, as calculated by MEA. Columns 2 through 6 present the average relative efficiency scores for each input. Columns 7 and 8 present the average relative efficiencies for each innovation output. Column *Eff(ρ_k)* presents the average aggregated MEA efficiency scores, as calculated with the alternative ranking measure ρ_k. A dark grey color is applied to denote the bottom quartile, the patterned grey shading corresponds to the lower-middle quartile, and the patterned light grey shading is used for the upper-middle quartile, while light grey reflects the top quartile. Income groups follow the United Nations 2016 classification. Analytical results for each country can be found in the accompanying Supplement, in Table S1.

4.3 Identifying the optimal policy direction

The findings for the second step of our framework are summarised in Table 3, while results are graphically exhibited in Figures 1 and 2 below, for the *Knowledge & Technology* and *Creative Outputs*, respectively. Table 3 shows the averages of the estimated sensitivities and calculates rank correlations. With regards to the figures, the first column plots the estimated PLS coefficients (sensitivities) against innovation inputs for the innovation output under consideration. The second column plots the estimated sensitivities against the respective MEA efficiency scores. Taking into account the interpretation we attach to pillar inefficiencies, the second column can therefore provide insights on the responsiveness of countries to innovation-related investments, given the potential of resource misallocation. Moreover, to evaluate the role of the "magnitude" of innovation in our framework, we vary the marker sizes according to countries' GII score.[10] Finally, the colouring in the shaded areas corresponds to different quartiles of efficiency and it is used to aid in the identification of suitable innovation policies for each country.

All sensitivities have positive signs, confirming the assumption that innovation inputs contribute to the expansion of innovation outputs. Looking at the median sensitivities in Table 3, we find that *Human Capital & Research* exhibits the highest values in both innovation outputs (0.30 and 0.35 respectively). A similar picture is observed for the *Infrastructure* pillar, which exhibits the second highest median values (0.20 and 0.23, respectively), while *Business Sophistication* is associated with the lowest ones (0.16 and 0.14, respectively). Therefore, our results indicate that an increase in *Human Capital & Research* or *Infrastructure* is expected to yield higher returns compared to other input pillars. We also find that the sensitivities of the *Creative Outputs* pillar are higher than those of the *Knowledge & Technology* pillar, suggesting a greater response of the former to changes in innovation inputs.

We do not observe any patterns in the relationship between innovation inputs and sensitivities, given the respective low rank correlations in Table 3.[11] This implies that the extent to which further investments in innovation are expected to generate the desirable returns, does not depend on how well-resourced a country currently is. On the contrary, there is a close link between input endowments and the GII scores, suggesting that well-resourced economies are ranked highly in the GII report. Some exceptions are observed, though, where GII rankings are not necessarily in line with rankings based on input pillars. For example, Kuwait (KW) ranks in the middle-five countries in the GII report, despite exhibiting one of the smallest input values in the *Business Sophistication* pillar, while the opposite is observed for Niger (NE), further adding to the diversity of innovation systems. The high

[10] Since the GII score is the average of the Innovation Input Sub-Index and the Innovation Output Sub-Index, greater values in innovation inputs and outputs are associated with a higher score for the index by definition.

[11] Analytical results can be found in the accompanying Supplement.

rank correlation between GII scores and input pillars, thus, highlights that, if policy aimed at achieving high GII scores, only countries with economies strong enough to invest heavily in innovation inputs would be able to achieve such a goal. However, such investments in innovation inputs may not generate the desirable returns, given our earlier findings on sensitivity for these countries. the increase in innovation outputs may not be the desirable one, given our earlier findings on sensitivity. Taking also into account the low rank correlation between sensitivities and GII scores, we confirm that it is not necessary for highly-ranked countries in the GII report to be as responsive to innovation investments. This reveals further considerations when using the GII scores for performance assessment or policy-making. Similar observations can be made when considering the rank correlations between innovation efficiency and the GII scores. Finally, the low rank correlations between innovation efficiency and sensitivity suggest that they offer additional and independent insights to the GII index.

Table 3. Summary of findings

	Institutions	Human Capital & Research	Infrastructure	Market Sophistication	Business Sophistication
Median Sensitivity (K&T)	0.14	0.30	0.23	0.15	0.12
Median Sensitivity (Cr)	0.21	0.43	0.31	0.22	0.13
Rank Corr: Sens (K&T) with Inputs	0.20	0.29	0.02	0.13	0.40
Rank Corr: Sens (Cr) with Inputs	-0.01	0.05	0.01	0.04	0.43
Rank Corr: Inputs with GII score	0.88	0.88	0.90	0.77	0.78
Rank Corr: Sens (K&T) with GII	0.16	0.20	-0.06	0.17	0.27
Rank Corr: Sens (Cr) with GII	-0.04	0.03	-0.07	0.03	0.32
Rank Corr: In.Eff with GII score	0.31	0.23	0.27	0.36	0.30
Rank Corr: In.Eff with Sens (K&T)	0.08	0.04	-0.09	0.03	0.06
Rank Corr: In.Eff with Sens (Cr)	0.00	0.00	-0.10	-0.02	0.08

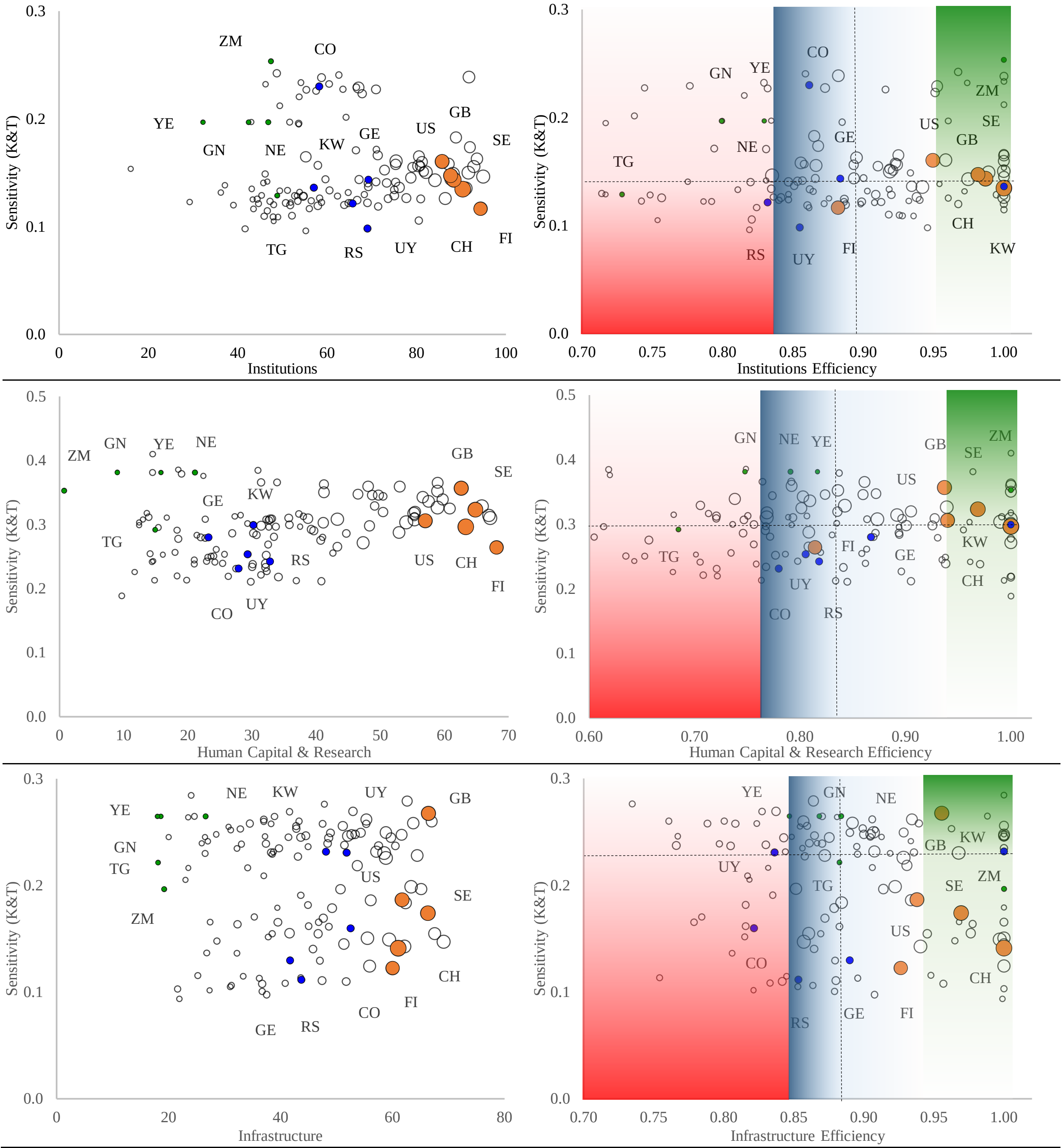
Notes: The table summarises the findings of our proposed framework. The first two rows exhibit the median values of the sensitivities for the *Knowledge & Technology* (K&T) and *Creative outputs* (Cr), against changes in the five innovation inputs (columns 2 to 6). Rows 3 and 4 present the rank correlations between the innovation inputs and the respective estimated sensitivities for the two innovation outputs. Rows 5 to 7 present the rank correlations of the GII scores with the innovation inputs and the estimated sensitivities for the two innovation outputs. The last three rows report the rank correlations between innovation efficiency and the GII scores as well as the respective estimated sensitivities.

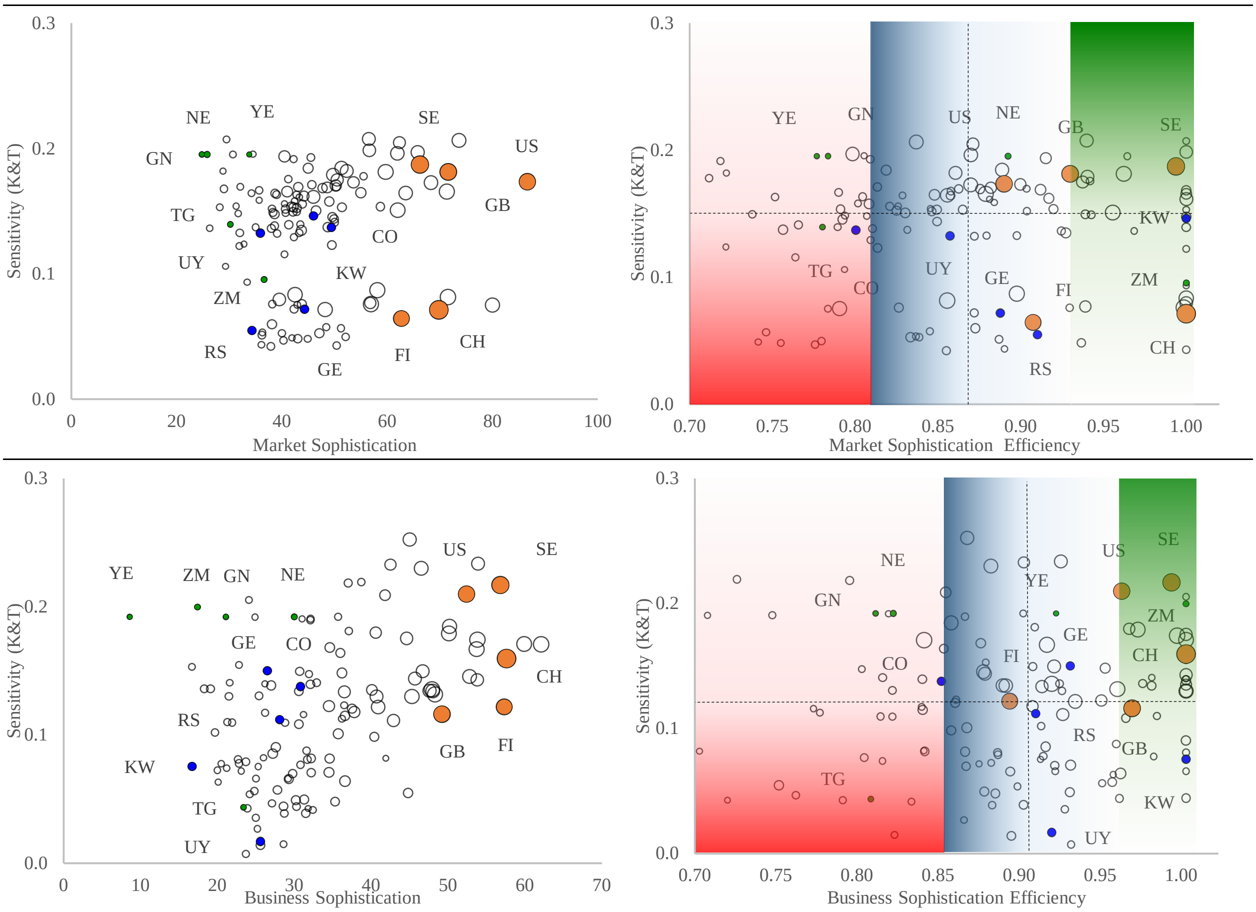
We now consider innovation efficiency and sensitivity in combination to propose tailored innovation policies for each country. In the figures below, we use colour shading to identify three directions. The red shaded areas include countries ranked at the bottom quartile of the innovation efficiency in the respective input, implying a considerable potential for resource misallocation. The colour scaling changes from dark to light red at the point that corresponds to the median sensitivity in the innovation input-output combination under consideration. The darker the red colour, the lower the responsiveness to innovation inputs. Given also the substantial level of inefficiency, an in increase in innovation inputs would probably mean that the used resources would generate less than expected innovation outputs. Even if countries are associated with high sensitivity, the possibility of substantial resource misallocation cannot be disregarded. Hence, policy makers should prioritize improving inefficiencies for countries in the red-shaded area by designing *innovation-improving* policies, as we call in this paper. The exact nature of these policies will depend on the structural characteristics of the countries concerned. For example, policymakers could adjust patent length to achieve an optimal balance between size and frequency of innovation (Horowitz and Lai, 1996). Similarly, Anderlini et al. (2013) show that relaxing legal system rigidities for countries at intermediate stages of technological development can increase the amount of innovations.

Countries in the top quartile are depicted with the green-shaded area in each input pillar. The median sensitivities are used again as the reference point for changing colour grading, so that darker green is associated with greater responsiveness to innovation input expansions. Countries in the dark green area would therefore benefit from policies that promote and support innovation-related investments, which we henceforth refer to as *innovation-facilitating* policies. Examples include initiatives such as grants for basic research (Salter and Martin, 2001), R&D subsidies (Almus and Czarnitzki, 2003), or R&D tax credits (Wilson, 2009), among others.

Finally, countries between the first and fourth quartiles are depicted with the blue-shaded areas. In this case, the colour scaling changes with respect to the median efficiency of the respective pillar, to signify the different nature of policy considerations in this case. The inefficiencies found in these countries are not alarming, but they cannot be disregarded either. Therefore, the lighter the blue shading, the less is the resource misallocation. In this case, a balanced mix of the mentioned policies is more appropriate; we refer to this combination of policies as *hybrid*. The policy mix depends on the position of each country in the quartiles formed by the crossing dotted lines that correspond to the median efficiencies and sensitivities. Thus, countries with relatively high (low) levels in both sensitivity and innovation efficiency should focus more on *innovation facilitating* (*innovation improving*) policies.

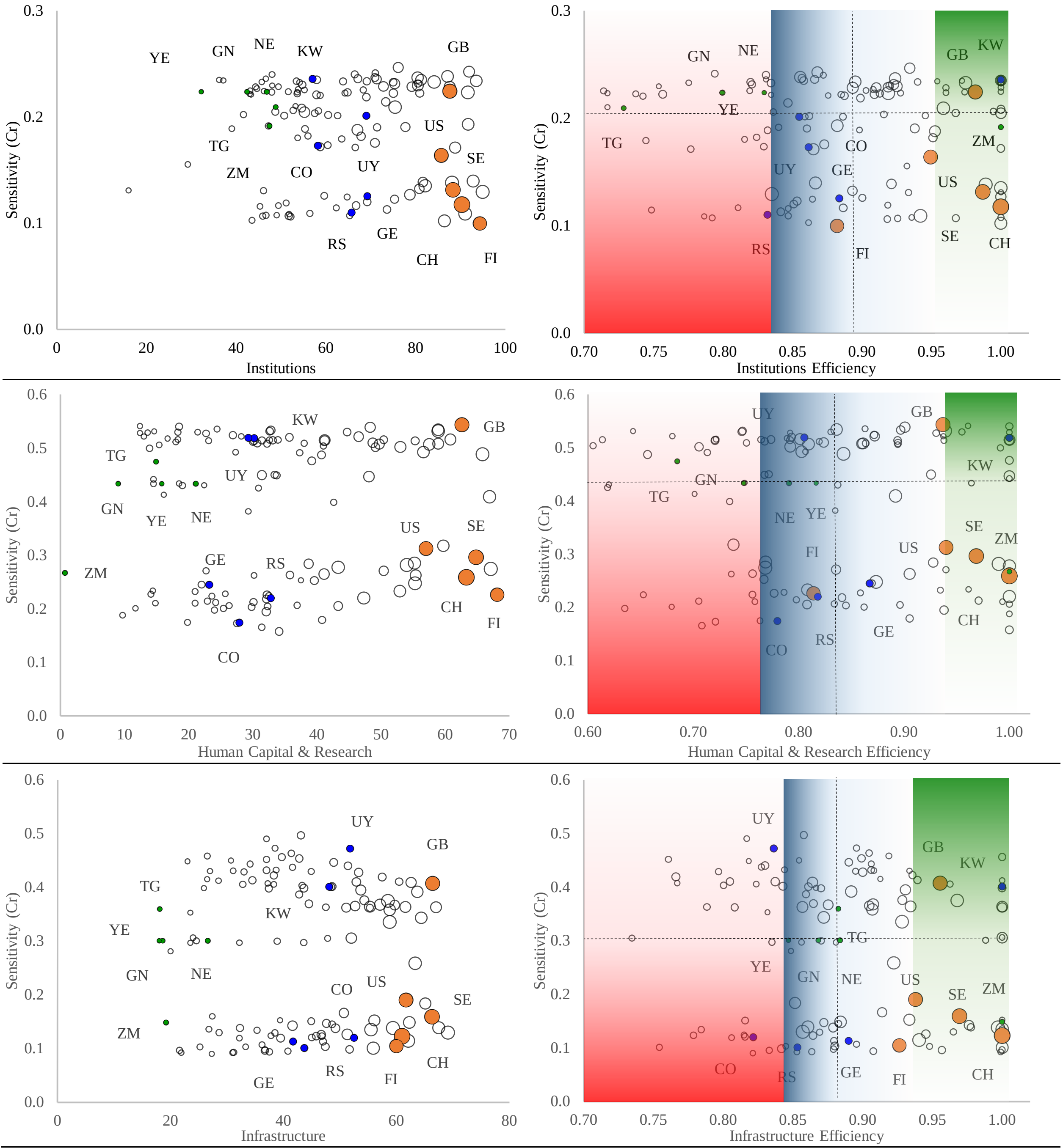
Figure 1. Sensitivities and MEA scores *Knowledge & Technology Outputs*

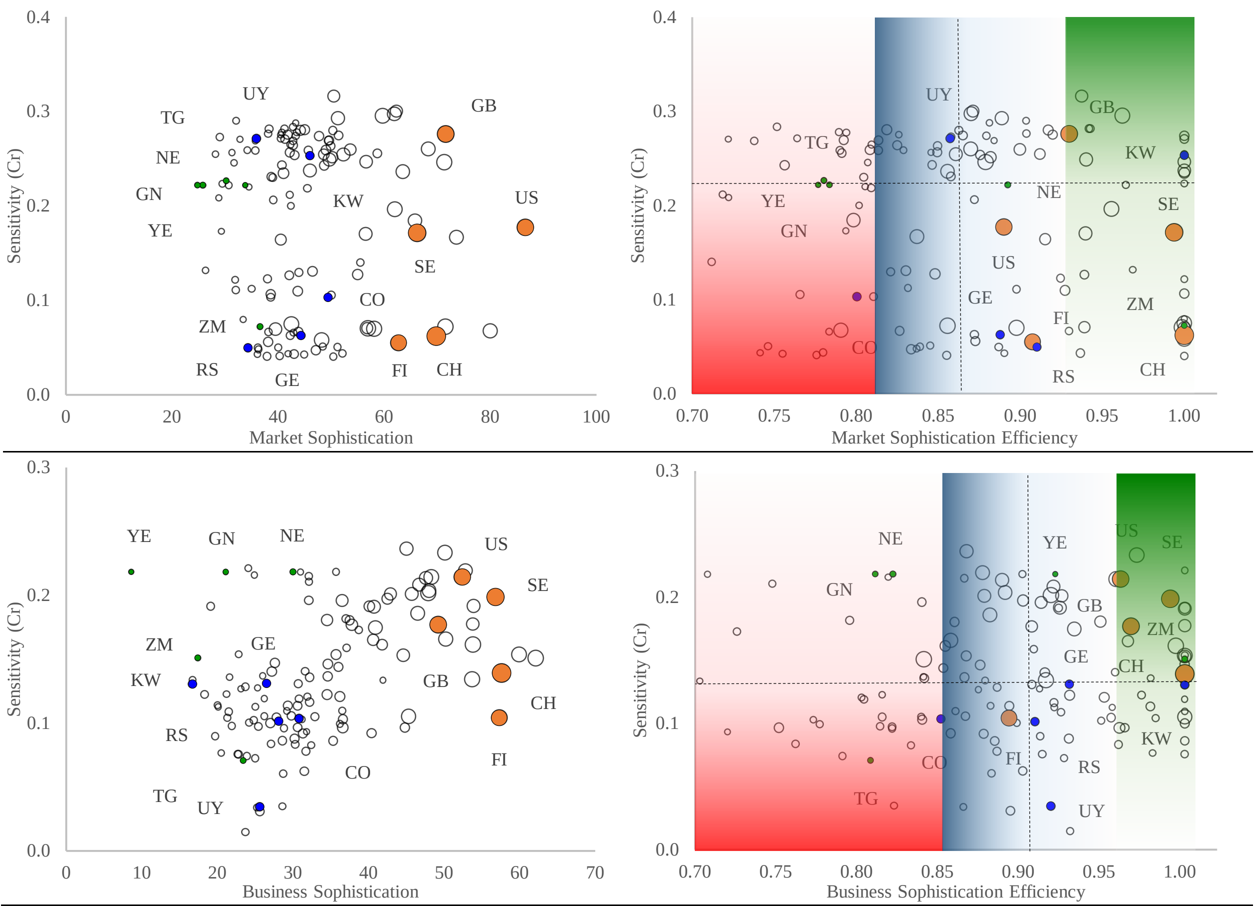




Notes: The first column of the figure plots the input pillar scores (vertical axis) against the sensitivity of *Knowledge & Technology Outputs* for the respective input (vertical axis). The second column presents the scatterplots of the sensitivities against the respective MEA efficiency scores (horizontal axis). The size of markers is proportional to countries' GII scores. The shaded areas provide an indication of the proposed innovation policy that a country should follow. The red-shaded areas correspond to the bottom MEA efficiency quartile for an innovation input, the green-shaded areas correspond to the top quartile, while the blue-shaded areas include the second and third MEA efficiency quartiles. The colour grading in the red and green shaded areas changes at the median sensitivity of the respective pillar, while for the blue areas it changes at the median of the respective MEA scores. The horizontal dotted lines correspond to the median sensitivity, while the vertical ones correspond to the median efficiency. Finally, we highlight as example countries the top five (orange marker), bottom five (green marker) and middle five (blue marker) countries in the GII rankings. The full list of country abbreviations can be found in the accompanying Supplement.

Figure 2. Marginal contributions and MEA scores for Creative Outputs





Note: The figure exhibits the same information as in Figure 1, with the difference that the sensitivities are estimated with respect to *Creative Outputs*. Please refer to the notes in Figure 1 for further details.

In line with our findings on rank correlations, we observe that countries with a high GII score (the sizeable markers) are not necessarily located in the *innovation-facilitating* zone, meaning that an expansion of innovation inputs does not guarantee an efficient or productive use of resources. Similarly, there are examples of countries with a low GII score that are not necessarily in the *innovation-improving* area. For example, out of the top five countries in the GII rankings, only Great Britain (GB) appears consistently in the darker green area (with one marginal exception). Switzerland (CH) and Sweden (SE) also appear consistently within the green region, but the fact that they exhibit mostly below-median sensitivities indicates that they have reached a near-optimal level of innovation activity in the respective dimensions. Hence, although there are no concerns for resource misallocation, further input expansions will be subject to diminishing returns for these countries. Finland (FE) appears in the lower quadrants of the *hybrid* region and should focus on rectifying inefficiencies to improve its innovation performance.

Similar discussion can be made for middle and low-ranked countries in the GII report. Zambia (ZM) and Kuwait (KW) are examples among several low and medium-ranked countries located at the top quartile of efficiency, and are therefore better suited for *innovation-facilitating* policies. Out of the five countries in the bottom of the GII ranking, only Togo (TG) and Guinea (GN) are located in the red area in all input pillars, except for *Infrastructure*. Therefore, *innovation-improving* policies need to be prioritised over any other innovation-related effort for these cases. On the contrary, Niger (NE) is more frequently located in the *hybrid* section, and, in fact, in the top-right quadrant for *Infrastructure* and *Market Sophistication* with respect to *Knowledge & Technology* outputs. Hence, providing more resources in this direction, coupled with monitoring systems to ensure a more efficient resource utilisation, will yield positive innovation outcomes.

The high rank correlations between innovation efficiency scores imply that most countries are located in a similar region across input-output combinations.[12] However, we find that countries are often located in different innovation policy areas across the ten possible efficiency-sensitivity combinations. Therefore, policy should tailor its approach for each country, based on its location on the mapping proposed above. For example, Yemen (YE) would need, in principle, to follow *innovation-improving* policies, especially with respect to *Market Sophistication*, whereas it has a good potential in the direction of *Business Sophistication*. Examining the components of the respective pillars, Yemen should focus on promoting knowledge-intensive employment and knowledge transfers, as well as improving innovation linkages, while it should address frictions, transactions costs and other sources of inefficiency in the process of credit creation, in its investment environment, as well as in maintaining a healthy level of competition in the market.

Our findings, apart from proposing a reconsideration and expansion of the current framework that is used to assess countries' rankings, also carries important policy implications that even challenge statements found in the GII report. The first such statement relates to the claimed innovation divide between high and low-income countries (GII, 2016, p. xxiv). We show that the misallocation of resources is also possible in highly ranked countries, and that even lower-income countries can strike a balance between innovation efficiency and sensitivity. Therefore, the innovation divide may appear overstated when considering more dimensions in the analysis. The second such statement characterises an entire region (Latin America and the Caribbean) as having an "untapped innovation potential" (GII, 2016, p. xxvi), but without providing supporting evidence. Our results show that most countries in *Latin America and the Caribbean* require *innovation-improving* policies, while only few cases can be found where *hybrid* policies could actually be implemented with an emphasis on *innovation-facilitating* ones. Finally, the responsiveness of LCN to innovation inputs is not always great, implying that not only resources may be misallocated but also that the wished outcomes are not necessarily feasible to achieve.

[12] To conserve space, we report the rank correlations of innovation efficiency scores in Table S6 of the Supplement.

5. Conclusion

This paper uses a two-step framework to propose innovation policies which best suit each country, by taking into account their diverse economic environment. These policies are identified by jointly considering countries' innovation efficiency and responsiveness to innovation-related investments. Our approach deals with two significant shortcomings of composite indicators, namely the fact that the diversity of national innovation systems (NIS) is not taken into account due to user-imposed weights and the counter-intuitive compensability property of the resulting composite indicators. Using data from the 2016 Global Innovation Index (GII) report, we highlight the diversity of NIS and propose tailored innovation policy directions, influenced by the extent to which the economic environment of each country is conducive to innovation. Our contribution is therefore twofold; not only do we propose a novel and flexible methodology that can be applied in various contexts, but we also use it to design an innovation policy toolbox that assesses the needs and comparative advantages of each country more accurately compared to current practices.

Two research objectives are evaluated regarding countries' innovation efficiency and sensitivity. The first one relates to the diversity of NIS, which, as expected, is confirmed. Using the MEA model, we obtain separate efficiency scores for each innovation pillar and identify substantial asymmetries. Evidence such as the fact that low-income countries appear significantly more inefficient with respect to *Creative Outputs* compared to *Knowledge & Technology* confirms that national priorities for innovation vary widely. Moreover, the fact that high-income countries in Europe perform better than their counterparts elsewhere, is also indicative of the role that local characteristics play on innovation, but also of the importance of market openness in facilitating knowledge transfers.

The second objective concerns the sensitivity that countries exhibit to changes in their innovation inputs, in relation to their peers. To estimate this, we propose a novel approach that runs separate PLS regressions for each country, whereby the response and explanatory matrices correspond to the reference country's nearest neighbours (peers). Economic proximity is determined in our study by three economic variables, which could be easily extended to include more quantitative or qualitative variables. Hence, we offer a platform for the designing of targeted and feasible innovation policies, which is also flexible for policy experimentation. The most responsive input factor appears to be, on average, *Human Capital & Research* and to a smaller extent *Infrastructure*, whereas *Business Sophistication* seems to be the innovation input that influences innovation outputs the least. However, this pattern does not apply universally, due to the diversity that is also manifested in the sensitivity estimates.

The rankings in innovation efficiency and the estimated sensitivities have little relevance with the GII rankings, which means that we add two independent and intuitive dimensions to the empirical analysis of innovation. Using these two dimensions in combination, we propose three policies on the basis of the potential for resource misallocation (proxied by the innovation inefficiencies) and the expected response to further investments in innovation inputs: *innovation-facilitating*, *innovation-*

improving, and *hybrid*. We observe that certain countries require a mixed approach, whereby *innovation-facilitating* policies may be suitable in one innovation input-output dimension, whereas *hybrid* or *innovation-facilitating* ones may be more suitable in another (or vice versa).

Our findings carry implications which should be of interest to policy makers who use composite indicators to inform their actions. First, the evaluation of countries' ability to innovate should not be based on simplistic approaches and should consider the influences of the environment which may be unique to each country. Characterising countries as 'achievers', 'followers' and underperformers' without accounting for their environment can lead to potentially falsely identified innovation gaps between countries. Second, taking into account both innovation efficiency and sensitivity is necessary to avoid mistakes in the evaluation of countries' potential to innovate and hence mitigate the risk of resource misallocation. Finally, given the diversity of national innovation systems, it is imperative that innovation policy should be tailored for each country, instead of adopting a one-size-fits-all approach that composite indicators tend to promote.

Acknowledgements

The authors would like to thank the seminar attendees of the Centre for Productivity and Performance at Loughborough University and the participants of the NAPW X conference for useful comments on an earlier draft of the paper. Any remaining errors are the authors' responsibility.

References

Acemoglu D., 2002. Directed technical change. Review of Economic Studies 69, 781–809.

Almus M., Czarnitzki D., 2003. The effects of public R&D subsidies on firms' innovation activities: the case of Eastern Germany. Journal of Business and Economic Statistics 21, 226–236.

Anderlini L., Felli L., Immordino G., Riboni A., 2013. Legal institutions, innovation, and growth. International Economic Review 54, 937–956.

Archibugi D., Filippetti A., Frenz M., 2013. Economic crisis and innovation: Is destruction prevailing over accumulation? Research Policy 42, 303–314.

Asmild M., Hougaard L.J., Kronborg D., Kvist K.H., 2003. Measuring inefficiency via potential improvements. Journal of Productivity Analysis 19, 59–76.

Asmild M., Matthews K., 2012. Multi-directional efficiency analysis of efficiency patterns in Chinese banks 1997-2008. European Journal of Operational Research 219, 434–441.

Atkeson A., Burstein T.A., 2010. Firm dynamics and international trade. Journal of Political Economy 118, 433–484.

Bădin L., Daraio C., Simar L. 2010. Optimal bandwidth selection for conditional efficiency measures: a data-driven approach. European Journal of Operational Research 201, 633–640.

Bloom N., Draca M., Reenen V.J., 2016. Trade induced technical change? The impact of Chinese imports on innovation, IT and productivity. Review of Economic Studies 83, 87–117.

Bogetoft P., Hougaard J.L., 1999. Efficiency evaluations based on potential (non-proportional) improvements. Journal of Productivity Analysis 12, 231–245.

Boschma A.R., 2005. Proximity and innovation: A critical assessment. Regional Studies 39, 61–74.

Cassiman B., Veugelers R., 2002. R&D cooperation and spillovers: Some empirical evidence from Belgium, American Economic Review 92, 1169–1184.

Carayannis G.E., Grigoroudis E., Goletsis Y., 2016. A multilevel and multistage efficiency evaluation of innovation systems: a multi-objective DEA approach. Expert Systems with Applications 62, 64–80.

Charnes A., Cooper W.W., Rhodes E., 1978. Measuring the efficiency of decision making units. European Journal of Operational Research 2, 429–444.

Chen K., Guan J., 2012. Measuring the efficiency of China's regional innovation systems: Application of network data envelopment analysis (DEA). Regional Studies 46, 355–377.

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

Cruz-Cázares C., Bayona-Sáez C., García-Marco T., 2013. You can’t manage right what you can’t measure well: Technological innovation efficiency, Research Policy 42, 1239–1250.

Davidson R., MacKinnon J.G., 1993. Estimation and Inference in Econometrics. New York: Oxford University Press.

Daraio C., Simar L., 2014. Directional distances and their robust versions: Computational and testing issues. European Journal of Operational Research 237, 358–369.

Despotis D.K., 2005. A reassessment of the human development index via data envelopment analysis. Journal of Operational Research Society 56, 969–980.

Desmet K., Parente L.S., 2010. Bigger is better: Market size, demand elasticity and innovation. International Economic Review 51(2), 319-333.

Freeman C., 1987. Technology, policy, and economic performance: Lessons from Japan. New York: Pinter Publishers.

Fusco E., 2015. Enhancing non-compensatory composite indicators: a directional proposal. European Journal of Operational Research 242, 620–630.

Glass A.J., Kenjegalieva K., Sickles R.C., 2016. A spatial autoregressive stochastic frontier model for panel data with asymmetric efficiency spillovers. Journal of Econometrics, 190, 289–300.

Grossman M.G., Helpman E., 1990. Trade, innovation and growth. American Economic Review 80, 86–91.

Grossman M. G., Helpman E., 1994. Innovation and growth in the global economy. MIT Press: Cambridge.

Grupp H., Schubert T., 2010. Review and new evidence on composite innovation indicators for evaluating national performance. Research Policy 39, 67–78.

Guan J., Chen K. 2012. Modeling the relative efficiency of national innovation systems. Research Policy 41, 102–115.

Hall B.H., Lerner J., 2010. The financing of R&D and innovation. In: Handbook of the Economics of Innovation. North-Holland, Amsterdam, pp. 609–639.

Han U., Asmild M., Kunc M., 2016. Regional R&D efficiency in Korea from static and dynamic perspectives. Regional Studies 50, 1170–1184.

He Y., Maskus E.K., 2012. Southern innovation and reverse knowledge spillovers: a dynamic FDI model. International Economic Review 53, 279-302.

Horowitz A.W., Lai E.L.C., 1996. Patent length and the rate of innovation. International Economic Review 37, 785–801.

Höskuldsson A., 1988. PLS regression methods. Journal of Chemometrics, 2(3), 211–228.

Kao C., Wu W., Hsieh W., Wang T., Lin C., Chen L., 2008. Measuring the national competitiveness of Southeast Asian countries. European Journal of Operational Research 187, 613–628.

Khedher L., Ramírez J., Górriz J.M., Brahim A., Segovia F., 2015. Early diagnosis of Alzheimer's disease based on partial least squares, principal component analysis and support vector machine using segmented MRI images. Neurocomputing 151, 139–150.

Kneip A., Simar L., Wilson P.W., 2008. Asymptotics and consistent bootstraps for DEA estimators in nonparametric frontier models. Econometric Theory 24, 1663–1697.

Long V.N., Raff H., Stähler F., 2011. Innovation and trade with heterogeneous firms. Journal of International Economics 84, 149–159.

Mahroum S., Al-Saleh Y., 2013. Towards a functional framework for measuring national innovation efficacy. Technovation 33, 320–332.

Mazzucato M., Semieniuk G., 2017. Public financing of innovation: new questions. Oxford Review of Economic Policy 33, 24–48.

Nasierowski W., Arcelus F.J., 2003. On the efficiency of national innovation systems. Socio-Economic Planning Sciences 37, 215–234.

Niosi J., 2002. National systems of innovations are "x-efficient" and (x-effective): Why some are slow learners. Research Policy 31, 291–302.

OECD, 2002. Foreign direct investment for development: Maximizing benefits, minimizing costs. OECD Publishing.

OECD, 2005. Oslo manual: Guidelines for collecting and interpreting innovation data, 3rd edition. OECD Publishing.

OECD, 2012. OECD Science, Technology and Industry Outlook 2012, OECD Publishing.

OECD, 2015. Frascati Manual 2015: Guidelines for collecting and reporting data on research and experimental development: The measurement of scientific, technological and innovation activities.

OECD Publishing.

Onodera O., 2008. Trade and innovation project: a synthesis paper. OECD Trade Policy Working Paper No. 72.

Paunov C., 2012. The global crisis and firms' investments in innovation. Research Policy 41, 24–35.

Politis D.N., Romano J.P., Wolf M., 2001. On the asymptotic theory of subsampling. Statistica Sinica 11, 1105–1124.

Romer P.M., 1990. Endogenous technological change. Journal of Political Economy 98, S71–S102.

Salter A.J., Martin B.R., 2001. The economic benefits of publicly funded basic research: a critical review. Research Policy 30, 509–532.

Simar L., Vanhems A., Wilson P.W., 2012. Statistical inference for DEA estimators of directional distances. European Journal of Operational Research 220, 853–864.

Tenenhaus M., Vinzi V.E., 2005. PLS regression, PLS path modeling and generalized Procrustean analysis: a combined approach for multiblock analysis. Journal of Chemometrics 19, 145–153.

The Global Innovation Index: Winning with Global Innovation, 2016. INSEAD and WIPO. Retrieved from https://www.globalinnovationindex.org/gii-2016-report.

Tone K., 2001. A slacks-based measure of efficiency in data envelopment analysis. European Journal of Operational Research 130, 498–509.

Vidoli F., Fusco E., Mazziotta C., 2015. Non-compensability in composite indicators: a robust directional frontier method. Social Indicators Research 122, 635–652.

Watkins A., Papaioannou T., Mugwagwa J., Kale D., 2015. National innovation systems and the intermediary role of industry associations in building institutional capacities for innovation in developing countries: a critical review of the literature. Research Policy 44, 1407–1418.

Wilson D.J., 2009. Beggar thy neighbor? The in-state, out-of-state, and aggregate effects of R&D tax credits. Review of Economics and Statistics 91, 431–436.

Wold S., Ruhe A., Wold H., Dunn III W. J. (1984). The collinearity problem in linear regression. The partial least squares (PLS) approach to generalized inverses. SIAM Journal on Scientific and Statistical Computing, 5, 735–743.

Wold S., Sjostrom M., Eriksson L., 2001. PLS-regression: a basic tool of chemometrics. Chemometrics and Intelligent Laboratory Systems 58, 109–130.

Yang Z., You W., Ji G., 2011. Using partial least squares and support vector machines for bankruptcy prediction. Expert Systems with Applications 38, 8336–8342.

Zabala-Iturriagagoitia M.J., Voigt P., Gutiérrez-Gracia A., Jiménez-Sáez F., 2007. Regional innovation systems: How to assess performance. Regional Studies 41, 661–672.