Offre pour une position d'assistant(e) de recherche au CERDI

Contrat de 3 mois.

Supervision par Luc Leruth, chercheur associé au CERDI, ancien Directeur de division du Département des finances publiques du Fonds Monétaire International.

CV : https://iset-pi.ge/storage/media/other/2021-04-09/d4750e80-9913-11eb-823a-8dc9808f2d74.pdf

Personne contact et référente de cette offre : Grégoire Rota-Graziosi, professeur au CERDI.

SI vous êtes intéressé(e), vous pouvez envoyer un CV et une lettre présentant votre intérêt/stratégie pour traiter la question à johan.guiot@uca.fr

The technical work – research assistant

The analysis

The book will require a thorough analysis of the most recent PEFA scores that are publicly available on the PEFA / World Bank site.

We plan to use a cluster analysis to identify groups of countries that are similar with respect to the priorities identified above but unlike most previous research, we do not transform the ordinal PEFA scale into a cardinal one. PEFA indicators are measured on an ordinal scale with five or seven levels (A - B+ - B - C+ - C - D+ - D). Most researchers, for reasons of convenience, assimilate ordinal scales to cardinal ones (a cardinal scale is proposed by the PEFA Secretariat). In many cases, this is a practical assumption without major consequences on the results. The implicit assumption underlying such a transformation, however, would be that there is the same distance between D and D+ as between A and B+, which would not be correct. We adopt the distance measure of Podani (1999), who adapted the Gower's (1971) distance, to allow for the treatment of ordinal qualitative variables. The method measures the minimum number of moves that one individual needs to perform on the ordinal scale to reach the position of a second individual.

The results of the hierarchical algorithm are usually visualized on a dendrogram and the process is finalized by deciding where to cut the dendrogram. One rule is to cut the tree at the median distance. There are other, more refined techniques, such as the silhouette approach that identifies the number of clusters so that the individuals are, on average, close to the other members of their cluster (see Charrad et al, 2014 or Kaufman and Rousseeuw, 1990). Technical details are presented below. Note that, irrespective of the methodology used to define the number of clusters, the idea is to group countries where the PFM systems perform in a similar way. That may mean: good, or bad, or present similar characteristics in terms of PFM functions. We have assigned green to the group with better performance and observe that the green groups always contain Ethiopia and Rwanda, the two-better performer in the region.

The software to determine the clusters is called ‘R’. The dendograms can be extracted using a simple program (available from the authors). It may be necessary to also program a way to color maps where countries will have the color of the cluster to which they belong.

Qualifications of the research assistant

An MA in economics or related discipline, in exceptional cases a student in his last year. Strong quantitative skills (stats and maths) including programming. Fluent in English. An interest in international economics and global issues.

Duration

About 3 months with a possibility to undertake a research project under the guidance of the authors, perhaps a PhD (to be discussed).

Background information

Example of expected research outcomes in the case of Sub-Saharan Africa

The worst performers tend to be the same across priorities, but the clusters are not 'stable' for medium performers and only two countries belong to all four clusters of best performers (Ethiopia and Rwanda). This has important implications for the design of a TA program. Unsurprisingly, a lot of reforms are needed (starting with Priority 1) for the poorest performers and the better performers also need additional reforms. More interestingly, the lack of stability across clusters for medium performers suggests that PFM performance is very uneven: one cluster will contain a set of countries for one priority (they do well according to one or two relevant PEFA indicators) but these same countries will not be in the same cluster for another priority. This suggests that future reforms must be designed with a better sense of sequencing and clearer objectives.

The inherited PFM system is not relevant to its current performance. Good or bad performance of a PFM system does not depend on its colonial origin. The analysis never produced a cluster where a substantial majority of countries had a PFM system from the same origin, irrespective of the height at which the clusters were selected in the dendrogram. This is a significant outcome that suggests that when properly operated and implemented, all PFM systems can deliver. Conversely, all systems are sensitive to poor implementation and a high or low level of corruption (or inefficiency) cannot be associated to a specific system.

The complexity of the systems adopted in SSA after the independence explains their poor performance. These systems (typically, the system used by the colonial power at home rather than the system implemented by the colonial power in its colonies) were not suited to the countries' specifics (Bouley et al., 2002). That complexity forced the authorities to rely on exceptional procedures that acted as shortcuts, something that is also apparent from the poor performance of PEFA Pillar 5. By routinely using these exceptional procedures, the authorities weakened financial accountability (Priority 1).

Resources rich countries have weaker PFM performance. Eventually, this affects macroeconomic stability and most of these countries end up in a cluster of poor performers for Priority 2a. This reinforces the widely held view that the volatility of prices for natural resources, or the contracting and large role of public enterprises distort PFM.

Technical issues on the ordinal scale and Podani's distance

Once an appropriate measure for ordinal variable has been adopted, it is reasonable to consider that two individuals reaching, for most indicators, identical levels on the ordinal scale should be more similar than two individuals reaching different levels on several indicators. Podani (1999) suggested an adaptation of Gower's distance for dealing with ordinal data. The idea is to measure the minimum number of moves that one individual needs to perform on the scale in order to reach the position of a second individual, while taking into account the effective range of each variable. Let us give more details. Let x_{ij} denote the observed value for the *i*th individual (*i = 1,...,n*) on the *j*th feature (*j = 1,...,p*). Gower (1971) started by defining a similarity between the individual s*i* and *k* by a weighted sum of similarities computed separately on each variable, i.e.

$$S_G(i,k) = \frac{\sum_{j=1}^{p} w_{ikj} s_{ikj}}{\sum_{j=1} w_{ikj}}$$

where w_{ikj}=0 if the individuals *i*and *k* cannot be compared because x_{ij} orx_{kj} is missing. Gower described how to compute the similarities s_{ikj} between the individuals *i* and *k* using the *j*th variable for all types of variables except for ordinal variables. Podani (1999) completed the definition by adding the ordinal case. First, the weights w_{ikj} are simply given by 1 when both values x_{ij} or x_{kj} are known and 0 otherwise. Then, after replacing all x_{ij} values by their ranks r_{ij} (taking into account the possible "ties", i.e. averaging the ranks when there are equal observed values for several individuals), the similarities s_{ikj} are defined as follows:

$$s_{ikj} = \begin{cases} 1 & \text{if } r_{ij} = r_{kj} \\ 1 - \frac{|r_{ij} - r_{kj}| - (T_{ij}-1)/2 - (T_{kj}-1)/2}{\max\{r_j\} - \min\{r_j\} - (T_{j,\max}-1)/2 - (T_{j,\min}-1)/2} & \text{otherwise} \end{cases}$$

where

T_{ij}= number of individuals having the same rank as the *i*th one on variable *j* (counting the *i*th one as well)

$T_{j,max}$ = number of individuals having the maximal rank (equal to max{r_j}) for variable *j*

$T_{j,min}$= number of individuals having the minimal rank (equal to min{r_j}) for variable *j*

As explained in Podani (1999), the numerator corresponds to the number of "swaps" between neighboring values on the scale of the *j*th variable in order to go from position $x_{ij}$to position x_{kj}. The denominator, when there are no missing values, simplifies to $n - T_{j,max} - T_{j,min} + 1$, which corresponds to the number of individuals lying between the two extreme ranks. Corresponding dissimilarities (usually referred to as Gower distances) are obtained by computing the complement of the similarity with respect to 1. Note that the number of modalities on the ordinal scales has an impact on the number of different values possibly taken by the distance. Working with qualitative variables having a reduced number of modalities diminishes the number of ways in which several countries may differ.

Gower distance, with the adaptation for ordinal data suggested by Podani, can be computed in the statistical software R with the gowdis function of the library FD (see Charrad et al., 2014).