Non-continuous inputs-outputs in DEA for the estimation of knowledge generation and innovation efficiency: the case of CIS data

George T. Koutsouradis and Kostas D. Tsekouras*

Department of Economics, University of Patras
University Campus Rio, 26504 Patras, Greece

Abstract

Estimation of the efficiency of knowledge generation process, in the context of knowledge production functions (Hausman et al, 1984), is often hindered by the non-continuity of the input and output variables which are considered in frontiers parametric and non-parametric approaches. Especially, using Community Innovation Survey (CIS) microdata, crucial information, such as patents, other forms of Intellectual Property Rights and the introduction of several types of innovation is of binary type and therefore does not allow for the estimation of knowledge generation and innovation efficiency (Griliches, 1990).

In the paper at hand, we develop an analytical methodological context which facilitates the inclusion of knowledge and innovation discrete input and output variables in a Data Envelopment Analysis (DEA) framework which is used to evaluate the efficiency of resources devoted to knowledge and innovation processes (KIP). We take advantage and further develop the analytical approach of Banker and Morey (1986) and Kamakura (1988). In this vein, we introduce the necessary handling of information provided by the CIS microdata in non-parametric knowledge and innovation frontiers estimation. More specifically, we first portrait the problems which arise by the mishandling of discrete knowledge and innovation inputs and outputs variables in the efficiency estimation of KIP. Then, we showcase the appropriate handling of non-continuous IPR (i) as output variable in a knowledge generation process, (ii) as an input in a innovation production setting and (iii) where non-continuous variables coexist both in the input and the output side of the frontier which encapsulates the KIP.

We employ an input oriented Variable Returns to Scale non-parametric frontier of DEA type. The main methodological idea is developed on the imposition of legitimate and reasonable hierarchical structures on the non-continuous knowledge and innovation inputs and outputs. Starting from the multiplier form of the corresponding linear programming problem we develop the corresponding envelopment form of the same DEA setting and then we introduce the set of the necessary descriptor binary variables. Therefore, the envelopment form of DEA

is substantially modified including additional linear restrictions. The number of these additional restrictions is dependent on the scaling of the hierarchical structure mentioned above

The introduced methodological approach is tested employing the 2012-2014 wave of CIS microdata for three Southern European Countries, namely Greece, Portugal, and Spain.

References

Banker, R., and Morey, R. (1986). "The Use of Categorical Variables in Data Envelopment Analysis". *Management Science,* **32**, 1613-1627.

Hausman, J. A., B. H. Hall and Z. Griliches (1984). "Econometric Models for Count Data with an Application to the Patents-R&D Relationship". *Econometrica*, **52**, 909-938.

Griliches, Z. (1990). "Patent Statistics as Economic Indicators: A Survey". *Journal of Economic Literature,* **28**, 1661-1707.

Kamakura, W. (1988). A Note on "The Use of Categorical Variables in Data Envelopment Analysis". *Management Science,* **34**, 1273-1276.

Acknowledgements

The research work was supported by the Hellenic Foundation for Research and Innovation (H.F.R.I.) under the "First Call for H.F.R.I. Research Projects to support Faculty members and Researchers and the procurement of high-cost research equipment grant" (Project Number: HFRI-FM17-0320, SPILEF).

* Corresponding Author: tsekour@upatras.gr