NON-CONTINUOUS INPUTS-OUTPUTS IN DEA FOR THE ESTIMATION OF KNOWLEDGE GENERATION AND INNOVATION EFFICIENCY: THE CASE OF CIS

George T. Koutsouradis and Kostas Tsekouras,

Department of Economics, University of Patras, Greece

The research project was supported by the Hellenic Foundation for Research and Innovation (H.F.R.I) under the "1st Call for H.F.R.I. Research Projects to support Faculty Members & Researchers and the Procurement of High-and the procurement of high-cost research equipment grant" (Project Number: HFRI-FM17-0320)

OUTLINE

1. Motivation and Objective
2. Innovation Efficiency
3. Data Envelopment Analysis
4. Treating Binary Inputs
5. Conceptual approach
6. Results
   a) Greece & Portugal
   b) Spain
7. Conclusions
8. References

MOTIVATION AND OBJECTIVE

- Many researchers in Europe use CIS data. Brouwer et al (1999);Catozzella et al (2008);Conte et al (2005);Damijan et al (2017)
- In order to calculate innovation efficiency, we use a multi-input multi output framework. In this case crucial information, such as patents, other forms of Intellectual Property Rights and the introduction of several types of innovation is of binary type.
- The most prominent approach to estimate innovation efficiency is grounded on frontiers methodology.
- DEA is the most popular branch of frontier methodology in innovation efficiency estimation due to its nonparametric characteristics not only allow for multiple inputs and outputs to be used regardless of measurement units, but also do not require prior information on the basic functional form and weight.
- On the contrary the main advantage of a parametric approach (SFA) is that it considers the existence of statistical noise in the data. Its main disadvantage, on the other hand, is the use of a specific functional form that presumably approximates the underlying technology; this however may impose unnecessary structure in the data. (Tsekouras et al 2003)
- However, using non continuous input output variables in DEA and especially in the case of innovation efficiency is not applicable.

MOTIVATION AND OBJECTIVE

- This work aims to introduce an approach on how to handle the non-continuous variables in CIS data by implementation of Data Envelopment Analysis (DEA) model by Banker and Morey (1986).
- Moreover we develop two different approaches of the innovation process with the use of non-continuous data.
- We develop theoretical arguments on which we built estimation techniques of innovation efficiency
- The implementation of modified DEA models is illustrated by the innovation efficiency evaluation of the 3 southern European countries Greece, Portugal and Spain from 2012-2014 wave of CIS microdata.

THEORETICAL ARGUMENTS

- Diversification of knowledge sources.
- Mediating role of Innovation Property Rights (IPRs) in the process of calculating innovation efficiency.
- IPRs are handled as an intermediate factor, acting both as an input and output variable. Therefore, "hard" innovation inputs exhibit direct and indirect effects.
- Innovation efficiency is differentiated in "Moderate Innovator" countries with respect to "Catching Up" countries.

INNOVATION AND EFFICIENCY

- Overall, innovation should not be regarded as a single event, but rather as a continuous and cumulative process. Innovation is not a linear process in which inputs automatically transform into outputs, innovation performance should not be measured as performance, but rather as efficiency, including its input and output altogether.
- In general efficiency is defined as the ratio of outputs over inputs. In case of innovation efficiency is the ability to transform innovation inputs into innovation outputs.
- Innovation efficiency is improved when with the same amount of innovation inputs more innovation outputs are generated (output-orientation) or when less innovation inputs are needed for generating the same amount of innovation outputs (input-orientation).
- R&D plays a significant role in increasing the probability of product innovation, while technological acquisition (TA) increases the likelihood of process innovation. Catozzella et al (2008). Process innovation is much more related to TA , both through the "embodied technical change" acquired by investment in new machinery and equipment and through the purchasing of external technology incorporated in licences, consultancies and know-how. Salter, (1960);Freeman, (1982)

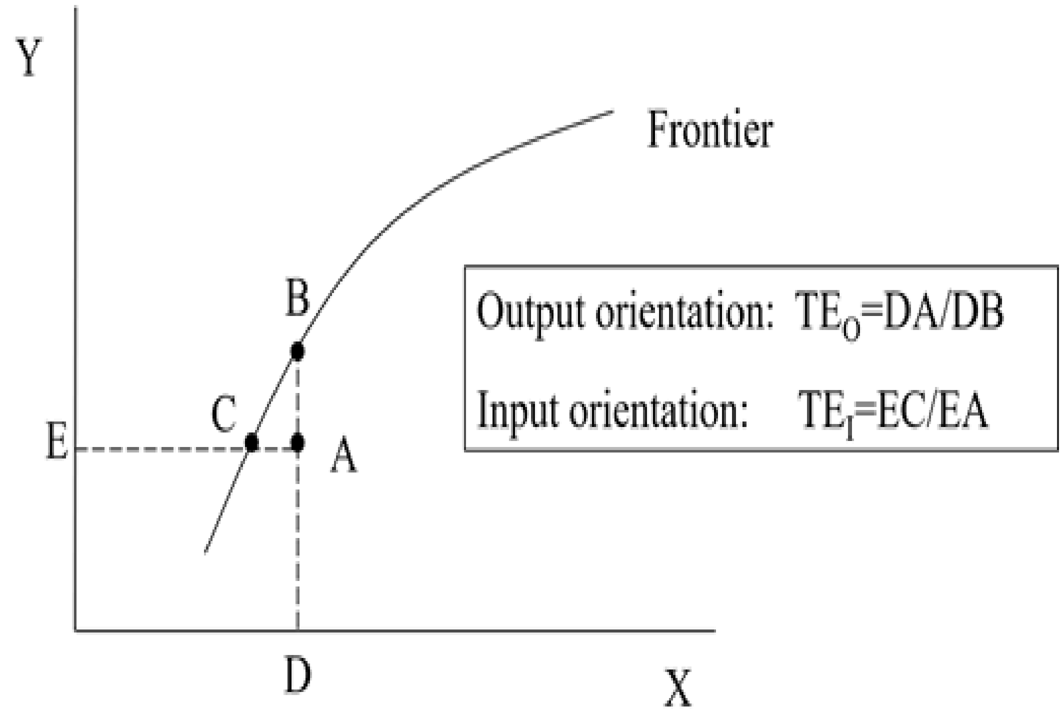
INNOVATION AND EFFICIENCY

- Large firms are characterised by a higher degree of diversification that helps them to deal with the uncertainty of R&D investment. (Cohen and Klepper, 1996; Mairesse and Mohen, 2002)
- Relevant literature on the role of IPRs in innovation and economic growth tends to focus on the strength of IPR protection, given the trade-offs between innovation and diffusion.
- From development perspectives, recent literature has shifted attention to diverse forms of IPRs in promoting innovation and growth, considering not only regular invention patents, but also utility models and trademarks. (Kim et al, 2012)

DATA ENVELOPMENT ANALYSIS

- DEA can be described as a nonparametric technique based on linear programming to evaluate the efficiency of DMUs working in the same industry.
- There are several approaches possible in DEA, for example one can distinguish between a constant returns to scale (CRS) technology and a variable returns to scale (VRS) technology and between an input-oriented version and an output-oriented version. In this work we shall use variable returns to scale technology and input-orientation.

Technical Efficiency Measurement



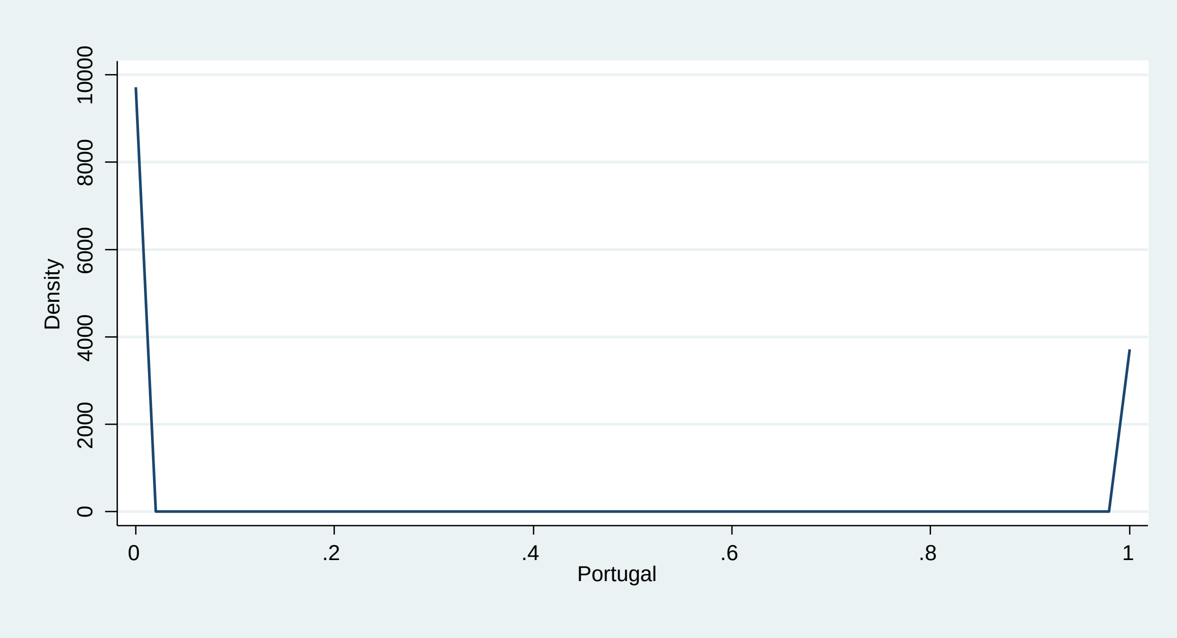
SIMPLE DEA RESULTS

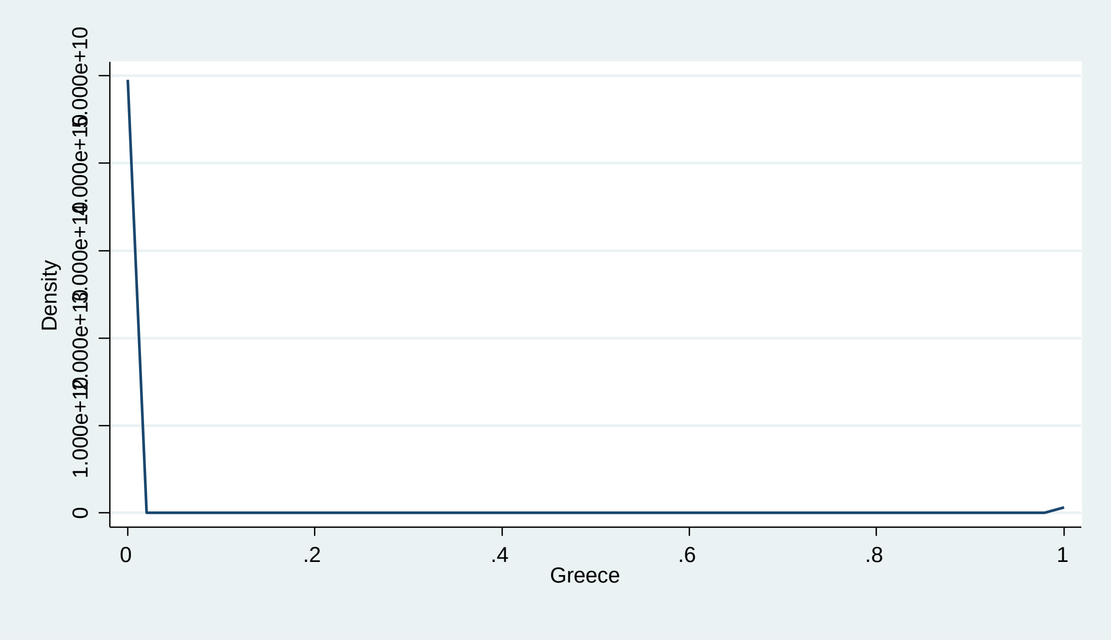
INPUTS

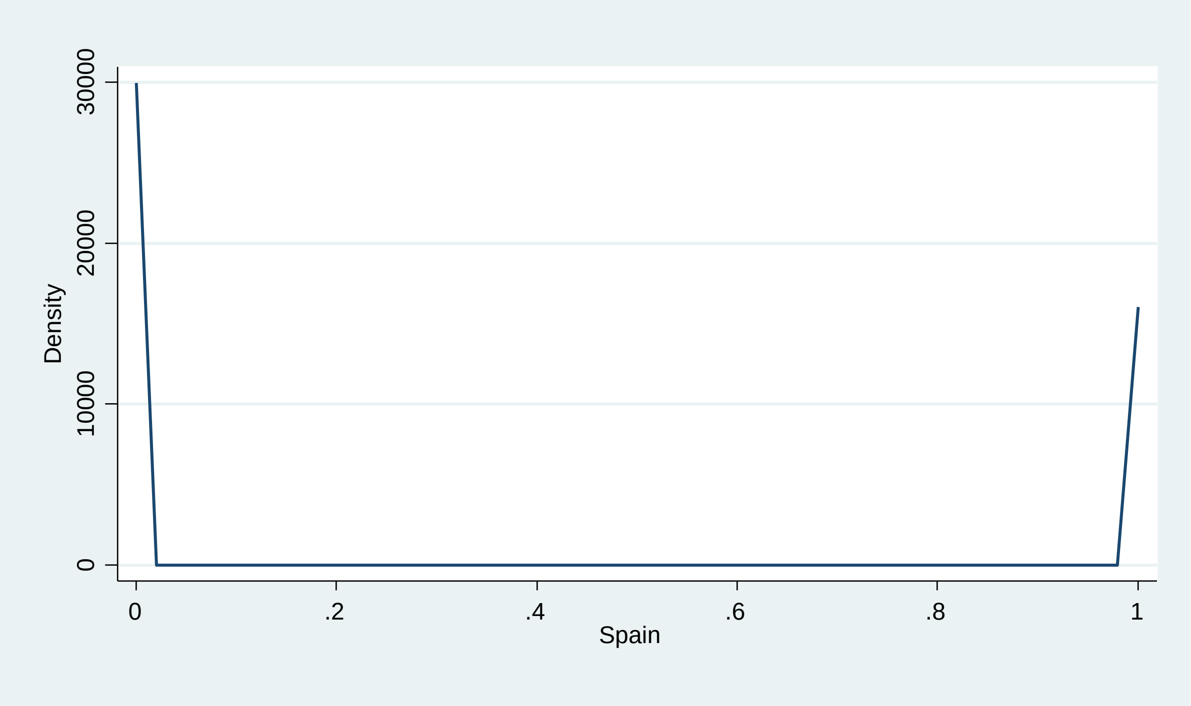
- Expenditures in intramural R&D
- Expenditures in acquisition of machinery
- Expenditures in extramural R&D
- Expenditures in acquisition of external knowledge

OUTPUT

- Innovative sales







CIS DATA BINARY VARIABLES

- 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.
- 87 % of variables in CIS microdata are binary.

TREATING BINARY INPUTS
BANKER AND MOREY(1986)

We define δ new variables $d_{m,j}^{(\delta)}$, where δ + 1 is the number of values the categorical variables can take on

- 4 distinct levels

1. "none"
2. "low"
3. "average"
4. "high"

- Descriptor binary variables

1. $d_{m,j}^{(1)}$
2. $d_{m,j}^{(2)}$
3. $d_{m,j}^{(3)}$

for each of the DMU's

"none" level -> $d_{m,j}^{(1)} = d_{m,j}^{(2)} = d_{m,j}^{(3)} = 0$

"low" level -> $d_{m,j}^{(1)}=1, d_{m,j}^{(2)} = d_{m,j}^{(3)} = 0$

"average" level -> $d_{m,j}^{(1)} = d_{m,j}^{(2)} = 1 , d_{m,j}^{(3)} = 0$

"high" level -> $d_{m,j}^{(1)} = d_{m,j}^{(2)} = d_{m,j}^{(3)} = 1$

Example

Class	d1	d2	d3
"none"	0	0	0
"low"	1	0	0
"medium"	1	1	0
"high"	1	1	1

DATA ENVELOPMENT ANALYSIS

Simple DEA model , input oriented

Efficiency = Min θ

s.t.

$$\sum_{1}^{n} \lambda_j \cdot x_j \leq \theta \cdot x_{r_0}$$

$$\sum_{1}^{n} \lambda_j \cdot y_{r_j} \geq y_{r_0}$$

$$\sum_{1}^{n} \lambda_j = 1$$

$$\lambda_j \geq 0 \; \forall \, j = 1, \dots, r_0, \dots N$$

Modified DEA model , input oriented , with categorical input variables

Efficiency = Min θ

s.t.

$$\sum_{1}^{n} \lambda_j \cdot x_j \leq \theta \cdot x_{r_0}$$

$$\sum_{1}^{n} \lambda_j d_{m_j}^{(\delta)} \leq d_{mr_0}^{(\delta)}$$

$$\sum_{1}^{n} \lambda_j \cdot y_{r_j} \geq y_{r_0}$$

$$\sum_{1}^{n} \lambda_j = 1$$

$$\lambda_j \geq 0 \; \forall \, j = 1, \dots, r_0, \dots N$$

Restriction : There must be at least one continuous input and output in the model

THEORETICAL ARGUMENTS

KNOWLEDGE SOURCES DIVERSIFICATION (KSD)

R&D expenditures are used by the firm as inputs and produce innovative sales.

- The input side of innovation efficiency is dependent on the variety of knowledge sources. Gkypali, A. et al (2017;2018); Leiponen, A. et al (2010)
- Engagement in: intramural R&D, extramural R&D , acquisition of machinery, acquisition of external knowledge, training for innovative activities, market introduction of innovation, are used as R&D-binary variables.
- DMU's which engage in more categories of R&D are considered to employ more distinct knowledge sources.

IPR AS A LATENT MEDIATOR (ILM)

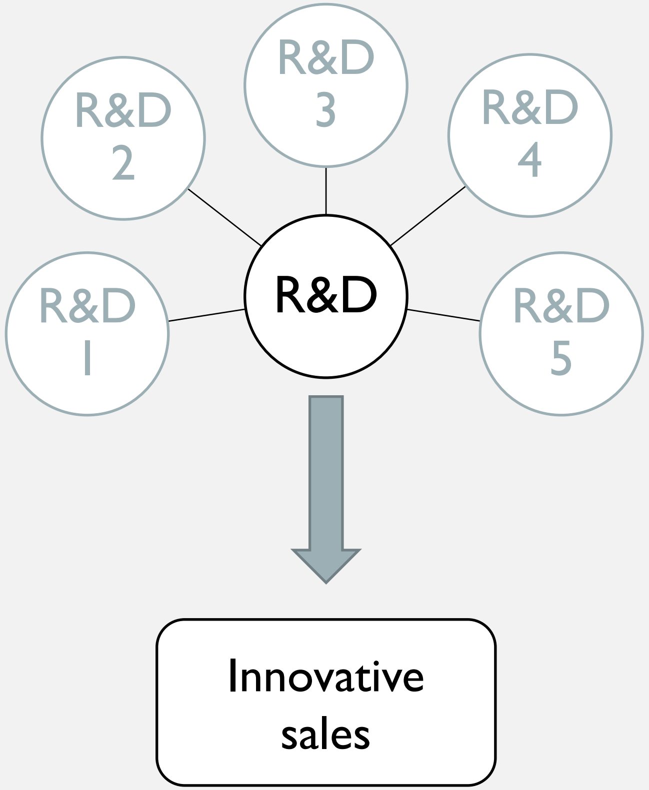
R&D expenditures along with a laten production process produce IPR which are then used as input to produce innovative sales

- IPR are considered as a mediator of the relationship between RD and innovative sales.
- Application for a patent, application for a European utility model, Registered an industrial design right, registered a trademark, licensed out or sell a patent, industrial design right, copyright or trademark to another enterprise, university or research institute are used as IPR-binary variables.
- DMU's which apply for more categories of IPR are considered to be higher on a intellectual property scale.

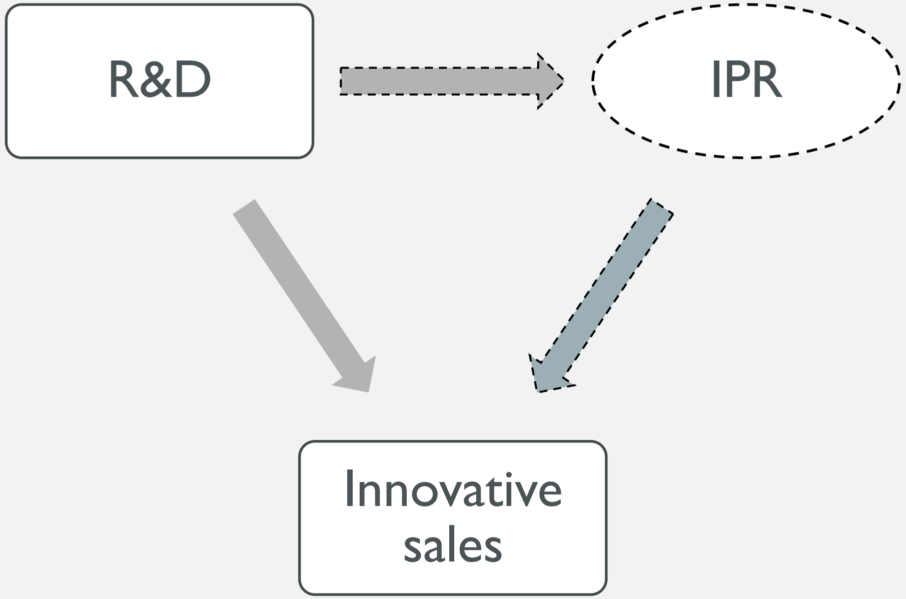
In both cases, as continuous inputs, expenditures in intramural R&D is used in order to capture the product innovation and expenditures of machinery acquisition to capture process innovation.

THEORETICAL ARGUMENTS

KNOWLEDGE SOURCES DIVERSIFICATION



IPR AS A LATTEN MEDIATOR



INPUT / OUTPUT VARIABLES STRUCTURE

Continuous Inputs

1) Expenditures in intramural R&D
2) Expenditures in acquisition of machinery

Categorical Inputs / Intermediate outputs

1) Engagement in intramural R&D
2) Engagement in extramural R&D
3) Engagement in acquisition of machinery
4) Engagement in acquisition of external knowledge
5) Engagement in training for innovative activities

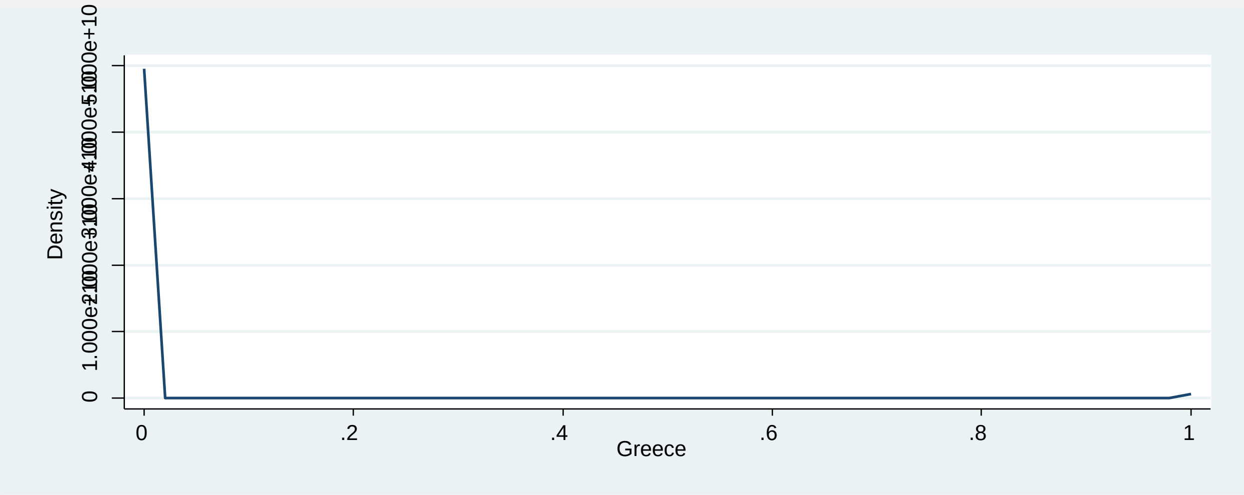
1) Applied for a patent
2) Applied for a European utility model
3) Registered an industrial design right
4) Registered a trademark
5) Licensed out or sell a patent, industrial design right, copyright or trademark to another enterprise, university or research institute

Continuous Outputs

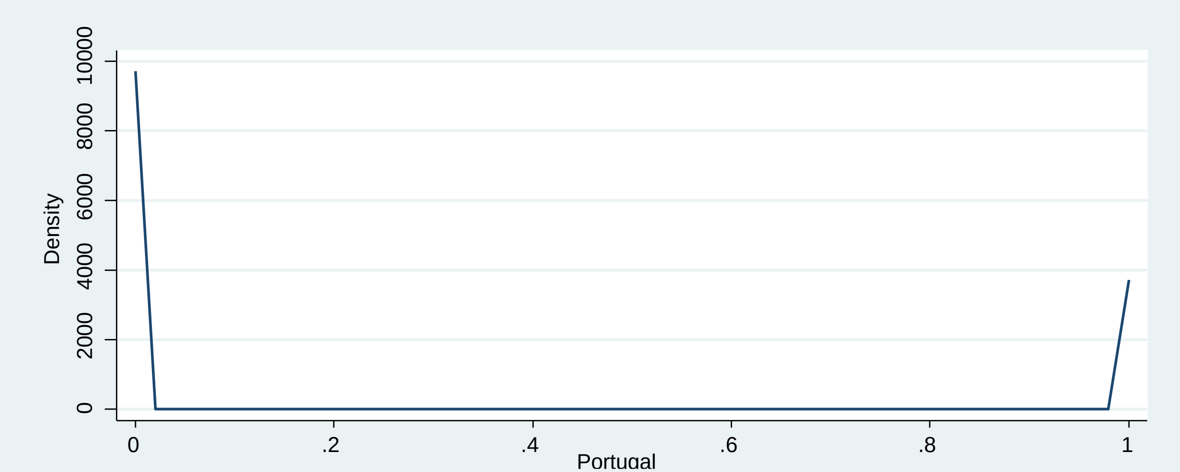
Innovative sales

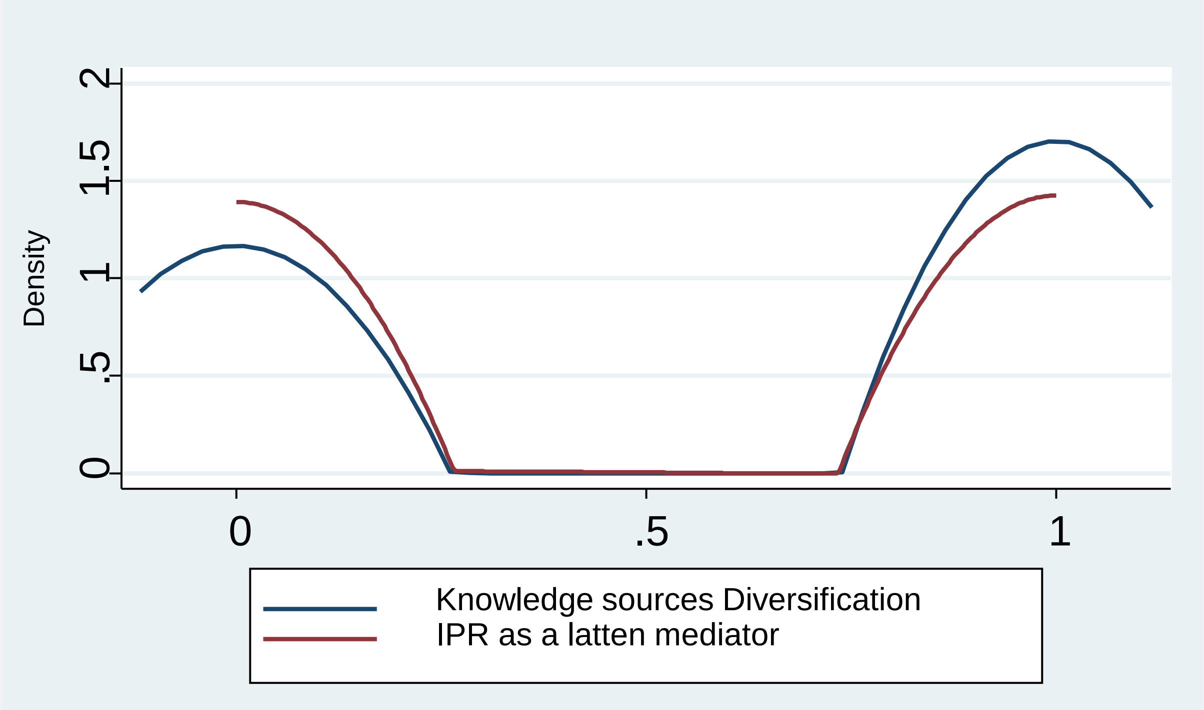
THE CATCHING UP COUNTRIES CASE

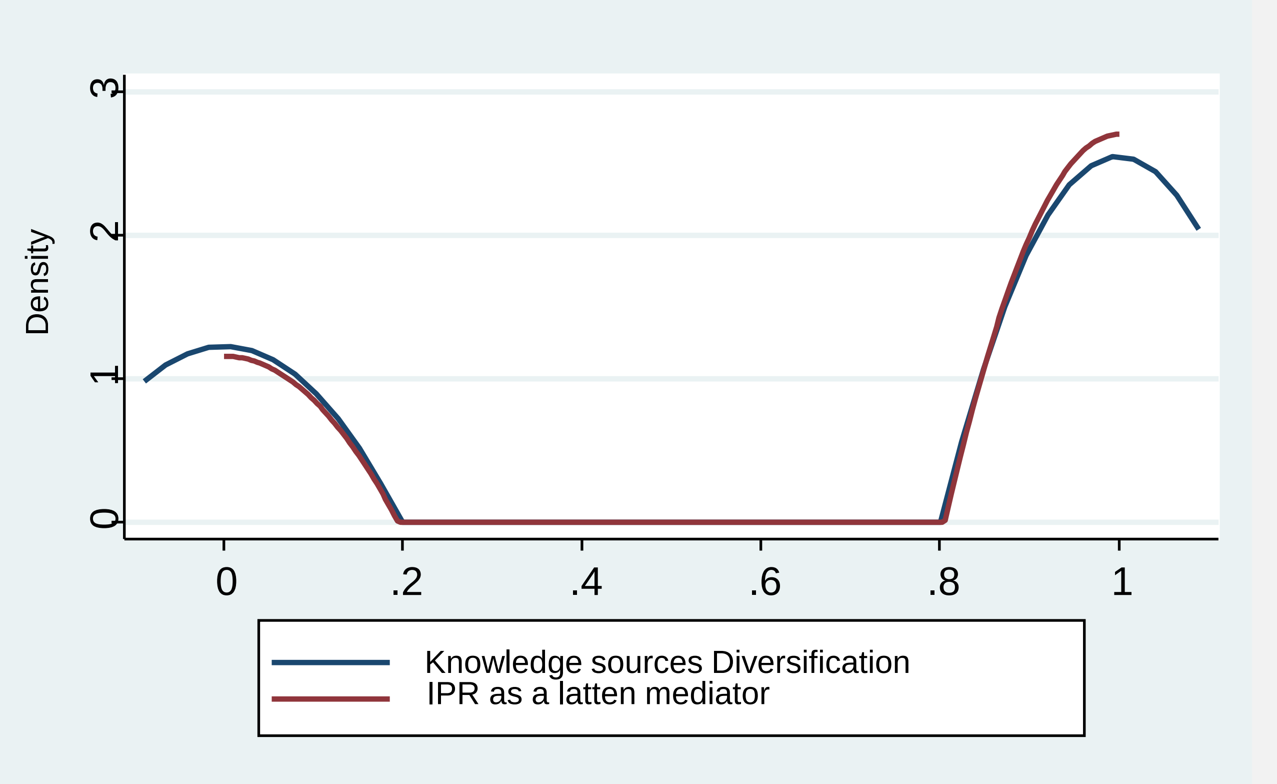
GREECE



PORTUGAL



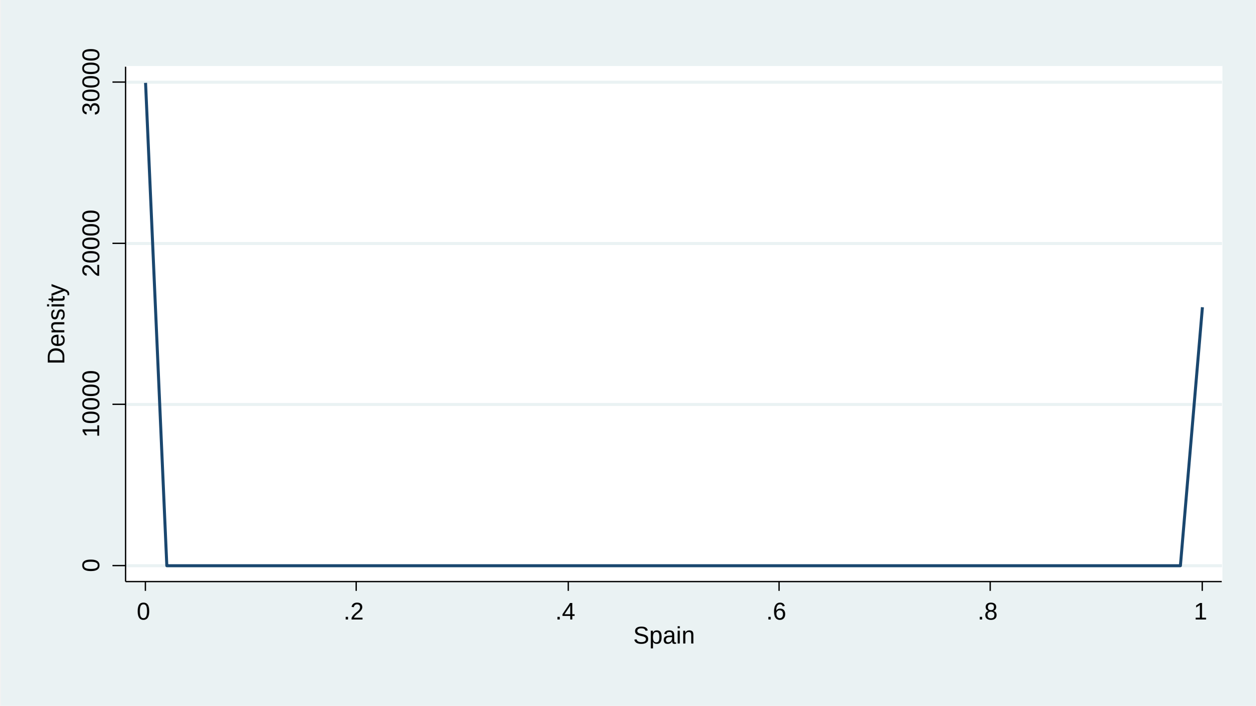


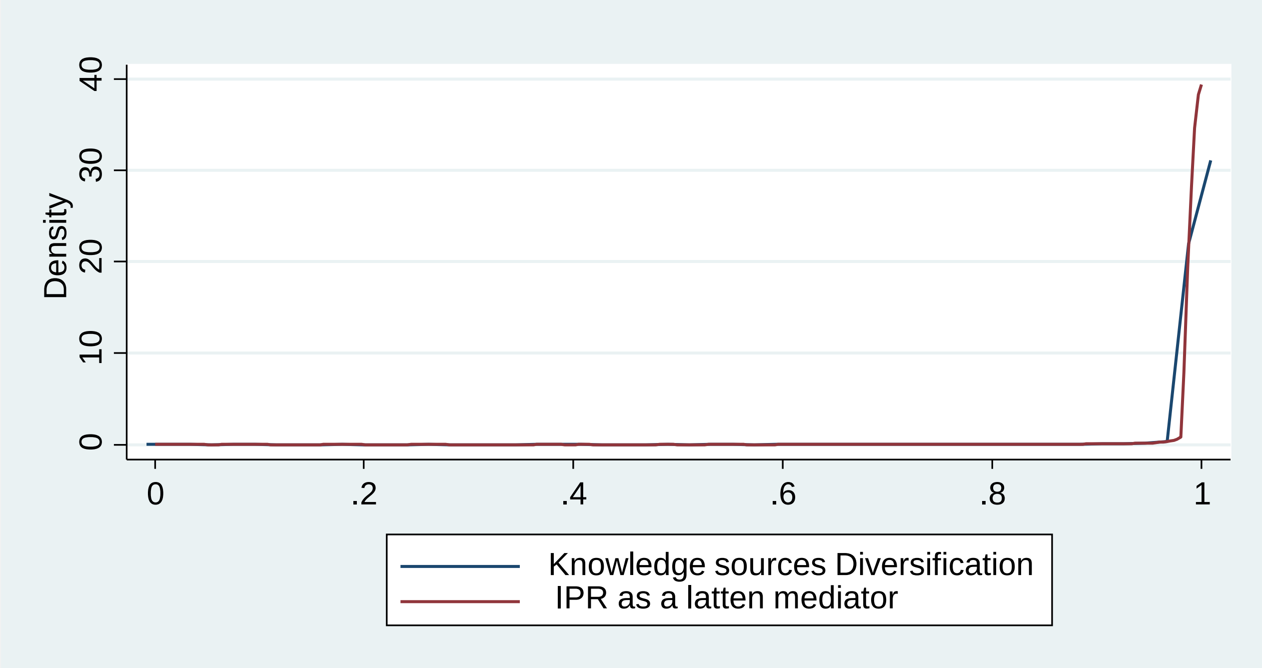


THE CATCHING UP COUNTRIES CASE

- Simple DEA gathers all efficiency scores close to zero.
- Modified DEA along with KSD or ILM approach, creates two efficiency groups, one low efficiency group and one high efficiency group.
- ILM approach inflates low efficiency scores and underestimates high efficiency scores in the case of Greece.
- ILM approach underestimates low efficiency scores and inflates high efficiency scores in the case of Portugal.

THE MODERATE INNOVATOR COUNTRY (SPAIN)





- **Spain** belongs in the "moderate innovators" group. Moderate innovator countries show above average efficiency in Intellectual property.
- **Portugal** and **Greece** belong in the "catching-up countries". Catching-up countries are significantly below average in Intellectual property. This may be because IPR is of less relevance for the innovative activities or there is the potential to generate higher levels of IPR from existing inputs.

CONCLUSIONS AND FURTHER RESEARCH

- We transport innovation efficiency distribution in a more informative way
- Moderate innovation countries need different handling from catching up countries
- Polarization needs further investigation with respect to firm characteristics which are attributed in each one of the two tales of the distribution
- In the case of moderate innovators we need to incorporate binary innovation outcome in the analysis

Average Efficiency	Greece	Portugal	Spain
Simple DEA	0.016	0.0151	0.018
Modified DEA RD	0.593	0.587	0.994
Modified DEA IPR	0.505	0.684	0.994

REFERENCES

- Banker, R., & Morey, R. (1986). The Use of Categorical Variables in Data Envelopment Analysis. *Management Science, 32*(12), 1613-1627.
- Brouwer, E., & Kleinknecht, A. (1999). Innovative output, and a firm's propensity to patent.: An exploration of CIS micro data. *Research policy*, *28*(6), 615-624.
- Catozzella, A., Conte, A., & Vivarelli, M. (2008). The Knowledge production Function Revisited: A Sequential Sample Selection Approach. *Journal of Product Innovation Management*.
- Cohen, W. M., & Klepper, S. (1996). Firm size and the nature of innovation within industries: the case of process and product R&D. *The review of Economics and Statistics*, 232-243.
- Conte, A., & Vivarelli, M. (2005). One or many knowledge production functions? Mapping innovative activity using microdata.

REFERENCES

- Damijan, J., Kostevc, Č., & Rojec, M. (2017). Exporting status and success in innovation: Evidence from CIS micro data for EU countries. *The Journal of International Trade & Economic Development*, *26*(5), 585-611.
- Freeman, C. (1982). The Economics of Industrial Innovation, 2nd edn. London: Pinter.(1974). *The Economics of Industrial Innovation*.
- Gkypali, A., Arvanitis, S., & Tsekouras, K. (2018). Absorptive capacity, exporting activities, innovation openness and innovation performance: A SEM approach towards a unifying framework. *Technological Forecasting and Social Change*, *132*, 143-155.
- Gkypali, A., Filiou, D., & Tsekouras, K. (2017). R&D collaborations: Is diversity enhancing innovation performance?. *Technological Forecasting and Social Change*, *118*, 143-152.
- Guan, J., & Chen, K. (2012). Modelling the relative efficiency of national innovation systems. *Research policy*, *41*(1), 102-115.

REFERENCES

- Hollanders, H. J. G. M., & Esser, F. C. (2007). *Measuring innovation efficiency*. Brussels: European Commission.
- Kim, Y. K., Lee, K., Park, W. G., & Choo, K. (2012). Appropriate intellectual property protection and economic growth in countries at different levels of development. *Research policy*, *41*(2), 358-375.
- Leiponen, A., & Helfat, C. E. (2010). Innovation objectives, knowledge sources, and the benefits of breadth. *Strategic management journal*, *31*(2), 224-236.
- Mairesse, J., & Mohen, P. (2002). The economics of technology and innovation. *American Economic Review*, *92*(2), 226-30.
- Salter, W. E. G. (1960). Productivity and Technical Change (Cambridge, 1960). *Salter Productivity and Technical Change1960*.