



Universiteit
Leiden
The Netherlands

Exploring the Innovative Efficiency of Big Multinational Firms through Conditional Efficiency Analysis

Laurens, P.; Schoen, A.; Toma, P.; Daraio, C.

Citation

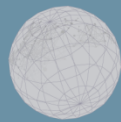
Laurens, P., Schoen, A., Toma, P., & Daraio, C. (2018). Exploring the Innovative Efficiency of Big Multinational Firms through Conditional Efficiency Analysis. *Sti 2018 Conference Proceedings*, 834-842. Retrieved from <https://hdl.handle.net/1887/65296>

Version: Not Applicable (or Unknown)

License: [Leiden University Non-exclusive license](#)

Downloaded from: <https://hdl.handle.net/1887/65296>

Note: To cite this publication please use the final published version (if applicable).



STI 2018 Leiden

23rd International Conference on Science and Technology Indicators
"Science, Technology and Innovation Indicators in Transition"

STI 2018 Conference Proceedings

Proceedings of the 23rd International Conference on Science and Technology Indicators

All papers published in this conference proceedings have been peer reviewed through a peer review process administered by the proceedings Editors. Reviews were conducted by expert referees to the professional and scientific standards expected of a conference proceedings.

Chair of the Conference

Paul Wouters

Scientific Editors

Rodrigo Costas
Thomas Franssen
Alfredo Yegros-Yegros

Layout

Andrea Reyes Elizondo
Suze van der Luijt-Jansen

The articles of this collection can be accessed at <https://hdl.handle.net/1887/64521>

ISBN: 978-90-9031204-0

© of the text: the authors

© 2018 Centre for Science and Technology Studies (CWTS), Leiden University, The Netherlands



This ARTICLE is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License

Exploring the Inventive Efficiency of Multinational Firms through Conditional Efficiency Analysis¹

Patricia Laurens^{*}, Antoine Schoen^{**}, Pierluigi Toma^{***} and Cinzia Daraio^{***}

** patricia.laurens@esiee.fr*

LISIS, Université Paris-Est, CNRS, LISIS, 2 bd Blaise Pascal, Noisy le Grand, 93160 (France)

** Antoine.schoen@esiee.fr*

LISIS, Université Paris-Est, ESIEE, LISIS, 2 bd Blaise Pascal, Noisy le Grand, 93160 (France)

****% toma@diag.uniroma1.it; daraio@diag.uniroma1.it*

DIAG Sapienza University of Rome, Rome (Italy)

Introduction

This paper addresses the issue of the inventive efficiency of multinational firms analysing the Corporate Invention Board (CIB), built in 2009 (and recently updated under the European RISIS project), and applying recently introduced nonparametric efficiency models which include, in a flexible multi-input multi-output Data Envelopment Analysis (DEA) framework, observed factors of heterogeneity and dynamics. Our exploration of the inventive efficiency analysis in R&D intensive companies is carried on a large scale. It includes more than 2000 companies from all over the world and various industrial manufacturing sectors. It combines economic and financial data of the firms, the firm's R&D investments as well as indicators to characterise their inventive production (volume, quality, diversity). It is the first time that such an analysis is carried out with this dataset. Our previous researches analysed the processes of the invention production (Laurens et al., 2015, Alkemade et al. 2015, Laurens et al., 2017).

Related literature

Empirical analyses on firms' R&D efficiency have been developed since the 80's. Griliches (1979) pointed out that all productivity growth, when measured correctly, is related to expenditure on R&D. Griliches and Mairesse (1984) and Griliches (1986), analyzing R&D investment and productivity, found a positive and significant relationship. Firm-level R&D is a driving force for technological innovation and economic growth (Romer, 1986; Lucas, 1988). Several studies have analysed R&D efficiency.

From attempting to measure R&D productivity at the firm or industry levels, Lee and Park (2005) extended R&D productivity measurement to the national level, providing policy implications. This study employed for the first time the data envelopment analysis (DEA) approach to measure R&D productivity and to classify the twenty-seven countries into four clusters based on the output-specialized R&D efficiency. In this field, the DEA method was applied to assess the relative efficiency of R&D activities across countries (Wang and Huang, 2007 ; Wang, 2007 ; Lee and Kang, 2007) or regions (Zhao, 2015).

¹ This work was supported by the European RISIS Project (EU FP7 INFRASTRUCTURE Project 313082)

The effects of entry on incumbent innovation and productivity is instead the topic treated by Aghion et al. (2009) who provided evidence that the threat of technologically advanced entry spurs innovation incentives in sectors close to the technology frontier, where successful innovation allows incumbents to survive the threat, but discourages innovation in laggard sectors, where the threat reduces incumbents' expected rents from innovating.

Cassiman et al. (2010) argued the positive correlation between firm productivity and exports and found that product innovation affects strongly productivity and represents a driver of exports for small non-exporting firms. Pannu et al. (2010) used data envelopment analysis to analyse the relative efficiency and productivity change in Indian pharmaceutical industry during a ten-year period, focusing on the effect of firm size on several performance measures. Results highlighted that firms which have invested in R&D and patents achieve higher efficiency than non-innovative firms.

Empirical analysis

Analytical framework

This paper founds on the innovation production framework and in particular considers the R&D/knowledge generation activity in each multinational firm as a production process.

In the scientific literature (Pakes and Griliches, 1984; Griliches, 1990) this process includes some observable measures, such as R&D expenditures or the number of employees involved in innovative activities.

These measures are intended for producing economically valuable knowledge, which could be thus considered as the output variable. Being the latter unobservable, in line with the literature (Wang & Huang, 2007), the number of priority patents, a quantitative indicator, should be used as a valid proxy for the inventive output.

Our ultimate interest is to study the effect of inputs (Capital, R&D expenditures and number of employees) on the production of inventions of a firm, a process that is also heavily affected by many other variables and policies.

Our model aims at investigating the effects of the levels of firm's internationalization and firm's technological specialization on its invention production, through an efficiency analysis. Maintaining the same traditional input variables as other models well-established in the literature, we introduce two contextual variables, i.e. invention geographical diversity (1-HH index on inventors' countries) and invention technological diversity ((1-HH index on technology fields).

Data Description

This research uses patent data from the Corporate Invention Board (CIB) that combines patent data from the PATSTAT database with data from the ORBIS database about the 2289 worldwide companies with the largest R&D investments (Laurens et al., 2015), as well as economic and financial data and firm's R&D investments from the Industrial R&D Investment (IRI) Scoreboard.

The companies included in the CIB account for 80% of world total private R&D: 730 have their headquarters in Asia, 1002 in Europe and 538 in Northern America.

The corporate patent database was built in several stages. Firm's names were retrieved from the IRI Scoreboard and their subsidiaries identified using the Orbis database to build the firm's consolidated perimeter (including only majority owned subsidiaries). Then, the names of the firms and their subsidiaries have been looked for as potential applicant names in the Patstat database². Checking and cleaning steps are needed to finalise the database. The overall

² A detailed presentation of the building and characterisation of this large firms' patent database is provided in Laurens et al. (2015).

CIB database includes more than 8 million worldwide priority patents applied for by the firms from 1986 to 2009.

Data and indicators were aggregated at the firm's perimeter level. We include in our analysis several data from the IRI Scoreboard either as input or output variables (or as contextual information): the firm's number of employees, the capital expenditure, the net sales and the R&D investments.

From the CIB database, we calculate for each company the following patent indicators as output variables: the total number of priority patents applied for a given period of time (ranging from 2002 to 2008), the number of singleton priority patents, the number of IP5 priority patents. The two last indicators give information on the inventive production providing a patent quality threshold (based on the geographical coverage of the patent protection). The number of singleton patents accounts for low quality patents, the number of IP5 family for higher quality patents (transnational patents with at least one application in one of the 5 largest patent offices in the world). Geographical and technological diversity indexes (using Herfindhal indexes) were calculated. The former is based on the geographical location of the invention production. It uses the country of the inventors as a proxy. The latter uses the distribution of patent (fractional counting) according to the 35 technological fields (WIPO classification). Table 1 summarises the definition and the sources of the data used as variables in the models.

Table 1. Description of the variables used in the analysis

Variables	Definition	Source
Number of employees	Total consolidated average employees	Industrial R&D Investment Scoreboard
Capital Expenditure	Capital expenditure	Industrial R&D Investment Scoreboard
Net sales	Sales, excluding sales taxes and shares of sales of joint ventures & associated companies.	Industrial R&D Investment Scoreboard
R&D expenditures	R&D investment (only investments funded by, and performed for, the companies themselves)	Industrial R&D Investment Scoreboard
Priority patents	Total number of priority patents applied for	CIB Database
Singleton patents	Total number of priority patents applied for only in one patent office	CIB Database
IP5 families patents	Total number of priority transnational patents applied for which the extended family includes a patent filed in one of the IP5 offices	CIB Database
Technological diversity	Invention technological diversity (1-HH index on the 35 technology fields)	CIB Database

Geographical diversity	Invention geographical diversity (1-HH index on inventors' countries)	CIB Database
------------------------	---	--------------

Method

The activity of multinational firms is modelled within an Activity Analysis Framework according to which a set of inputs $X \in R_+^p$ is used to produce a set of outputs $Y \in R_+^q$, with $Z \in R^d$ represents external (contextual) environmental variables, which are neither inputs nor outputs but may affect the performance of the production process.

The set of technically feasible combinations of firms (x, y) (*unconditional production set*) is

$$\Psi = \{(x, y) \in R_+^{p+q} | x \text{ can produce } y\}$$

This can be characterized in a probabilistic way (Daraio and Simar, 2007) by $\Psi = (x, y) | H_{X,Y}(x, y) > 0$ where $H_{X,Y}(x, y) = \text{Prob}(X \leq x, Y \geq y)$. So Ψ is the support of the joint random variable (X, Y) . The unconditional (marginal) output-oriented Farrell–Debreu technical efficiency of a firm (x, y) is defined as:

$$\lambda(x, y) = \sup \lambda | (x, \lambda y) \in \Psi = \sup \lambda | S_{Y|X}(\lambda y | x) > 0$$

where $S_{Y|X}(y|x) = \text{Prob}(Y \geq y | X \leq x)$ is the nonstandard conditional survival function of Y given that $X \leq x$.

Mastromarco and Simar (2015) considered in this framework the time dimension, analysing a panel of data $(x_{i,t}, y_{i,t}, z_{i,t})$ for $i = 1, \dots, n$ and $t = 1, \dots, s$. The time dimension T is introduced as an additional conditioning variable and, for each time period t , the attainable set $\Psi_t^z \subset R_+^{p+q}$ is defined as the support of the conditional probability:

$$H_{X,Y|Z}^t(x, y, z) = \text{Prob}(X \leq x, Y \geq y | Z = z, T = t)$$

Accordingly, the conditional output-oriented technical efficiency of a production plan $(x, y) \in \Psi_t^z$, at time t facing conditions z , is defined in (Daraio and Simar, 2005) as:

$$\lambda_t(x, y|z) = \sup \lambda | (x, \lambda y) \in \Psi_t^z = \sup \lambda | S_{Y|X,Z}^t(\lambda y | x, z) > 0$$

where $S_{Y|X,Z}^t(\lambda y | x, z) = \text{Prob}(Y \geq y | X \leq x, Z = z, T = t)$.

Assuming that the true attainable sets are convex and under free disposability of inputs and outputs, the DEA estimators can be written as (Daraio and Simar, 2007) similarly, at time t and facing the conditions $Z = z$,

$$\Psi_{t,DEA}^z = \{(x, y) \in R_+^p \times R_+^q | \leq \sum_{j \in J(z,t)} \gamma_j y_j; x \geq \sum_{j \in J(z,t)} \gamma_j x_j; \gamma \geq 0 \text{ s.t. } \sum_{j \in J(z,t)} \gamma_j = 1\}$$

where $J(z, t) = \{j = (i, v) | z - h_z < z_{i,v} < z + h_z; t - h_t < v < t + h_t\}$; h_z and h_t are bandwidths of appropriate size selected by data-driven methods. These $J(z, t)$ describe the localizing procedure to estimate the conditional DEA estimates and they determine the data points in a neighbourhood of (z, t) that will be used to compute the local DEA estimate.

External (environmental or contextual) variables may affect the *efficient boundary* of the production set (i.e. the best practice frontier) and/or the *distribution of inefficiency* across the firms (i.e. how far a firm is from the best practice frontier).

The effects of contextual variables on the boundary and on the distribution of the inefficiencies can be assessed using the approach developed by Bădin et al. (2012). The effect on the boundary can be detected by analysing the ratios between *conditional* to *unconditional* efficiency measures:

$$R_O(x, y|z, t) = \frac{\lambda_t(x, y|z)}{\lambda(x, y)}$$

The focus of our study is, in particular, the effects of T and Z on these ratios. For the output orientation, the conditional efficient boundary is by construction below the unconditional one, while ratio is equal to 1, if and only if there is no shift of the efficient boundary of the two attainable sets, at time t and with conditions z . In practice, we use nonparametric estimators of the efficiency scores and we explore the effect of T and Z by looking at the behaviour of $R_O(x, y|z, t)$ and $R_{O,\alpha}(x, y|z, t)$ as a function of T and Z .

Examining the behaviour of the above ratio, we can investigate if the environmental variables (z) affect the production process. According to Bădin et al. (2012), if the ratios as a function of z show an *increasing* trend, this points to a *positive* effect of z on the production process. On the contrary a *decreasing* trend of the ratios shows a *negative* effect of z on the production process while a *flat* trend displays *no effect* of z on the production process.

To analyse the effect of our external variables we use a nonparametric model for the second stage regression Bădin et al. (2012) that does not rely on the very restrictive *separability condition* between the input/output space and the space of external factors. In practice, the separability condition assumes that the external factors do not influence the efficient frontier, but only the distance of firms from the efficient frontier.

Results

Several models were developed to explore the inventive activity of large multinational firms included in the CIB and explain the effects of R&D expenditures. They are summarised in Table 2.

The first model, named *Basic Model*, allows us to analyze the effects of R&D expenditures on the classic production process activity. In this case the selected inputs are the Labour variable, measured by the number of workers, and the Capital variable represented by the Capital Expenditure. The selected output is represented by Net Sales. Considering R&D expenditure as an environmental variable, the model evaluates if R&D affects the efficiency performances of DMUs (Decision Making Unit), and in which measure. This dynamic model allows us to consider the effects of time on efficiency and to evaluate time dynamics. In order to conduct this analysis the time span has been divided into 3 periods: 2002/2004, 2005/2006 and 2007/2008.

The second model, named *Invention Quantity Model*, assesses the effects of business size, identified by the net sales on the R&D production process, using data of 896 companies for the 2002-2008

period. The selected input variables are the number of workers, the capital expenditure and the R&D expenditure, while the total number of priority patents represents the output.

Following the same approach, the third model (*Quality model*) evaluates the effect of size on the Quality of Invention. Using the same inputs as the Quantity Model, the *Invention Quality Model* uses the number of IP5 families patents, as output variable, since it represents a proxy for high quality inventions. In this case, size (measured by Net Sales) together with the low quality inventions (measured by the number of singleton patents) are used as environmental variables. The analysis aims at assessing the influence of size on high quality patent production process, but also accounting for the low quality patent production. According to data availability, this model covers the 2002-2005 time period and used data from 896 companies.

The last model, named *QualQuant*, aims at investigating the effects of internationalisation and specialisation level on the quality and quantity of invention production processes. Maintaining the same traditional input variables as the other models, the *QualQuant Invention Model* uses invention geographical diversity (1-HH index on inventors' countries) and invention technological diversity ((1-HH index on technology fields) as environmental variables, while the number of priority patents and the number IP5 families patents represents the output variables.

Table 2. Models of Invention production

Model	Basic	Quantity	Quality	QualQuant
Input	Number of employees and Capital expenditure	Number of employees, Capital and R&D expenditures	Number of employees, Capital and R&D expenditures	Number of employees, Capital and R&D expenditures
Output	Sales	Priority patents	IP5 families patents	Priority patents, and IP5 families patents
Contextual-environmental factors	R&D expenditure	Sales	Sales and Singleton patents	Geographical diversity, Technological diversity

Due to space limits, we report Figures 1 and 2 as illustrative examples of the explorative results obtained. The figures refer to the QualQuant model. It appears that the geographical diversity as a function of z (Figure 2, top panel) has a slightly decreasing trend pointing to a slightly negative impact of it on the performance. On the contrary, the technological diversity (Figure 2, bottom panel) seems to have a positive impact (increasing trend of the ratios as a function of z).

Figure 1. Effect of geographical diversity (GEO) and technological diversity (TECHNO) on the ratios.

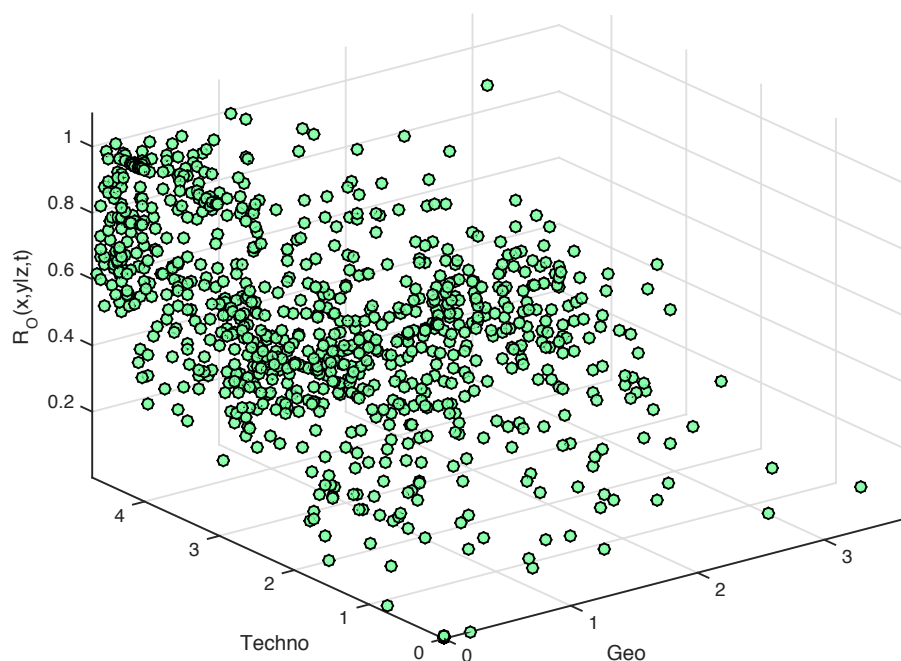
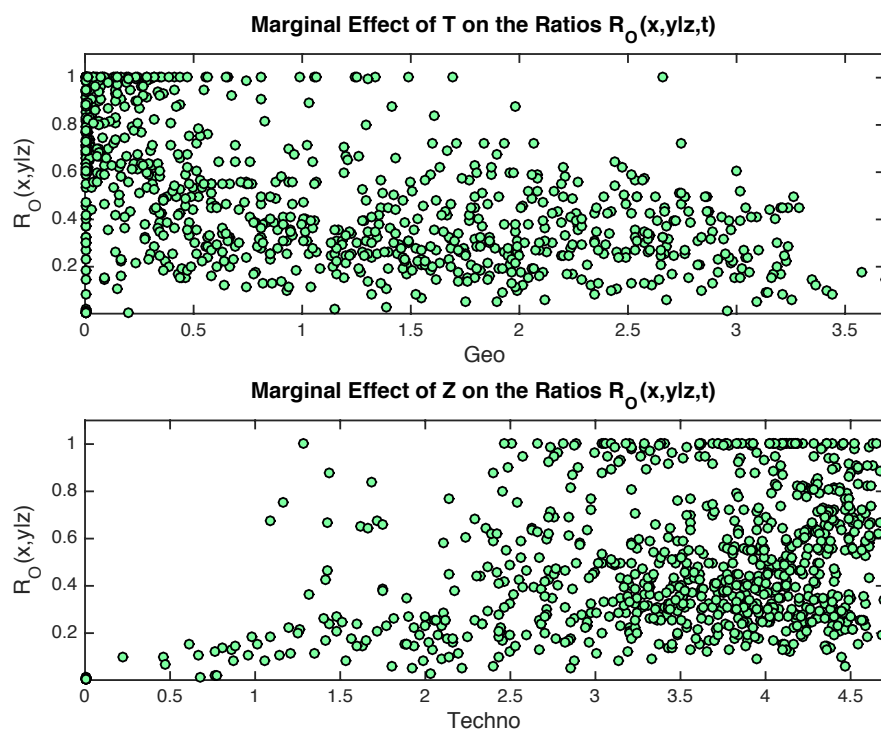


Figure 2. Marginal effects of geographical diversity (GEO) and technological diversity (TECHNO) on the ratios. Top panel: Marginal effect of GEO; bottom panel: marginal effect of TECHNO.



Discussion and conclusions

Thanks to the availability of CIB data, we can explore the inventive production of large and R&D intensive companies by applying recent nonparametric econometric methods. In particular, we applied conditional models of efficiency which are multidimensional, do not rely on strict parametric relationships between inputs and output and do not assume the strict *separability condition* allowing the disentangling of the impact of z variables on the efficient frontier of the best practice and on the distance of each firm from the best practice frontier.

The analysis carried out on the different models of innovation has shown that partial models that consider only one aspect of the production process at a time can mask important non-linear interactions and complementarities. In fact, there may be trade-offs between quantity and quality and these can escape from one-dimensional analyses based on linear innovation models.

In this work we modelled the inventive process with robust, multidimensional non-parametric methods that allowed us to shed some light on the complexity of the inventive performance of large multinational firms.

Further investigations on the models developed in this paper will include the application of a robust two stage nonparametric approach to estimate the “managerial” efficiency of the firms analysed. This will allow the comparison between heterogeneous firms, taking out (or cleaning) the effect of heterogeneity factors (the external environmental variables z).

Selected References

Aghion, P., Blundell, R., Griffith, R., Howitt, P., & Prantl, S. (2009). The effects of entry on incumbent innovation and productivity. *The Review of Economics and Statistics*, 91(1), 20-32.

Alkemade, F., Heimeriks, G., Schoen, A., Villard, L., Laurens, P. (2015). National and Sectoral characteristics of the internationalization of the inventive activity of multinational corporations, *Research Policy* 44 (2015) 1763–1772.

Bădin, L., Daraio, C., & Simar, L. (2012). How to measure the impact of environmental factors in a nonparametric production model. *European Journal of Operational Research*, 223(3), 818-833.

Cassiman, B., Golovko, E., & Martínez-Ros, E. (2010). Innovation, exports and productivity. *International Journal of Industrial Organization*, 28(4), 372-376.

Daraio C., Simar L. (2007), *Advanced Robust and Nonparametric Methods in Efficiency Analysis. Methodology and Applications*, Springer, New York (USA).

European Commission (2018). *The EU Industrial R&D Investment Scoreboard*. <http://iri.jrc.ec.europa.eu/scoreboard.html>.

Guan, J., & Chen, K. (2010). Modeling macro-R&D production frontier performance: an application to Chinese province-level R&D. *Scientometrics*, 82(1), 165-173.

Laurens, P., Le Bas, C., Schoen, A., Villard, L. Larédo, P. (2015). The rate and motives of the internationalisation of large firm R&D (1994-2005): Towards a turning point?. *Research Policy* 44, 765-776.

Lee, H. Y., & Park, Y. T. (2005). An international comparison of R&D efficiency: DEA approach. *Asian Journal of Technology Innovation*, 13(2), 207-222.

Mastromarco, C., & Simar, L. (2015). Effect of FDI and time on catching up: New insights from a conditional nonparametric frontier analysis. *Journal of Applied Econometrics*, 30(5), 826-847.

Simar, L., & Wilson, P. W. (2007). Estimation and inference in two-stage, semi-parametric models of production processes. *Journal of econometrics*, 136(1), 31-64.

Wang, E. C. (2007). R&D efficiency and economic performance: A cross-country analysis using the stochastic frontier approach. *Journal of Policy Modeling*, 29(2), 345-360.

Wang, E. C., & Huang, W. (2007). Relative efficiency of R&D activities: A cross-country study accounting for environmental factors in the DEA approach. *Research Policy*, 36(2), 260-273.