Journal of Social and Economic Statistics

Vol. 4, No. 2. Winter 2015

LOGISTIC REGRESSION MODELS: REASSESSING THE DETERMINANTS OF PRODUCT INNOVATION USING ROMANIAN SURVEY DATA[*]

Elena DRUICA[a], Roxana ADAM[b]

Abstract

The aim of the paper is to examine the factors that increase the firm's propensity towards product innovation, with the goal of identifying the profile of Romanian product-innovative firms. The focus of our analysis is on the obstacles and hampering factors in achieving innovation in enterprises. The methodology makes use of a logistic regression model analysing the relationship between innovation on the one hand and hampering and obstacles on the other hand. We used the Community Innovation Survey (CIS) data, to ensure the representativeness of the study at the national level. The results of the two logistic regressions conducted based on CIS 2012 and CIS 2010 micro data suggest significant relationship between product innovation and: firm size; the strong competition on product quality, reputation or brand; lack of finance; innovation costs and no demand for innovations. Both models have a good prediction power.

Keywords: product innovative enterprises, logistic regression, obstacle factors, hampering factors, Community Innovation Survey (CIS).
JEL Classification: C01, C50, C51, O00, O30

Authors' Affiliation

[a] - University of Bucharest, Department of Economic and Administrative Sciences, elena.druica@faa.unibuc.ro

[b] – The Association for Strategic Decision Analysis

[*] This research has been conducted under the "Partnership in priority domains" Program - PN II, supported by MEN – UEFISCDI, Project nr. PN-II-PT-PCCA-2013-4-1802.

1. Introduction

Nowadays, the capabilities of innovating and successful innovation become important determinants of the global competitiveness and national progress for most of the countries. According to OECD (2007, p.6), since the Industrial Revolution, much of the rise in living standards is due to innovation, and evidence suggests that innovative effort is on the rise as a share of economic activity.

Innovation has been presented as a driver of economic growth in different theoretical and empirical studies (e.g. Smith, 1776; Schumpeter, 1934; Solow, 1957). Many researchers focused their studies on the relationship between innovation and development (e.g. Allen, 1967; Rothwell and Robertson, 1973; Mansfield, 1980; Jaffe, 1986; Fagerberg, 1994; Lehtoranta, 2005; etc.), that innovation is an important driver of economic growth.

O'Sullivan and Dooley (2008, p. 58) discuss that the organizations spend a significant amount of turnover on innovation every year, that may vary from 0.5% of turnover for organizations that operate in stable marketplaces, to more than 20% of turnover for organizations in emerging or turbulent marketplace. According to OECD report (2005), since the mid of 1990s in most OECD countries investment in knowledge has grown more rapidly than investment in machinery and equipment, and it exceeds it in countries such as Finland and the United States.

The EU Industrial R&D Investment Scoreboard (2014) reports the world top 2500 R&D investors that have continued to increase their investment in R&D (4.9%) well above the growth of net sales (2.7%). The same report shows on the top of the R&D investors, companies such as: Volkswagen from Germany (23.4%, up to €11.7bn), Samsung Electronics (South Korea) Microsoft (US) and Intel (US). Other top positions are occupied by Novartis and Roche from Switzerland, Toyota (Japan), Johnson & Johnson and Google from the US and Daimler (Germany).

Official overviews outline a general picture were the developed countries and large companies are the main innovative actors. Although research and development activities are correlated with innovation, the EU Industrial R&D Investment Scoreboard (2014) shows for the 633 EU companies, among the top world 2500 R&D investors, an annual R&D investment growth rate of 2.6%, well below the world average, accompanied by a decrease in sales (-2.0%) and operating profits (-6.6%). Even so, the European Commission finds innovation as a crucial tool for future macroeconomic development: in the context of Horizon 2020, it is expected that investing 3% of EU GDP in R&D by 2020 could create 3.7 million jobs and increase annual GDP by €795 billion by 2025.

A large body of literature (e.g. Huang, C., Arundel, A. and Hollanders, H., 2010) focuses on measuring innovation through various indicators such as: number of patents, R&D expenditures, types of innovations etc. In what follows, we set as the working definition for "innovator", the official statistics approach: an enterprise which during the reference period has introduced a new or significantly improved product over what was previously used or sold by the enterprise, no matter if has been originally developed or used by other enterprises.

Smith (2005, p. 149) points out that the key problems in innovation indicators concerns the underlying conceptualization of the object being measured, the meaning of the measurement concept and the general feasibility of different types of measurement. The author considers that the problems of commensurability are not necessarily insoluble, but a main point arising from recent work is the need for care in distinguishing between what can and what cannot be measured in innovation.

Departing from the literature evidence, the aim of the paper is to examine the factors that increase the enterprises probability towards product innovation, in order to identify a profile of Romanian product innovative enterprises. The focus of our analysis is on obstacle and hampering factors in achieving innovation in enterprises. Logistic regression models have been used as a theoretical support in order to describe the relationship between the dichotomous dependent variable (the status of an enterprise: being or not being product innovative enterprise) and two sets of factors as independent variables: hampering factors and obstacles as factors.

2. Literature review

Adam Smith is known for his studies on human nature end economic development, but in his work, the author also refers to what we call nowadays innovation. Should we reflect on his work *"An Inquiry into the Nature and Causes of the Wealth of Nations",* one of the hypothetical ideas related to a positive relationship between innovation and economic development it is noticed through his statement *"All the improvements in machinery, however, have by no means been the inventions of those who had occasion to use the machines. Many improvements have been made by the ingenuity of the makers of the machines, when to make them became the business of a peculiar trade"* (Smith, 1776). Passing chronologically, we find Joseph Schumpeter (1934) and Robert Solow (1957) as the two main pioneers who have underlined in their works that innovation is an important driver of economic growth. All these are indicated that innovation has been considered, one way or another, a driver of economic growth for more than two hundred years. Although the

importance of innovation was discussed and empirically proven (e.g. Mansfield, 1962; Nadiri, 1993) during the last decades, researchers haven't found yet the most appropriate measure for it.

Canepa and Stoneman (2007) studied the financial factors as constraints to innovation in the UK and found these factors to impact the innovative activity. A recent research (Adam, 2015) shows that high turnover Romanian enterprises are more likely to introduce innovative products on the market. The same study points out that firm size (number of employees) has an influence on the product innovation, and most of the product innovative companies are large enterprises. These results are in line with the Schumpeterian view according to which, a larger firm increases the probability of carrying out innovation activities (Pires-Alves and Rocha, 2008), given that large firms with monopolistic power are seen as more motivated to spend resources helping them to cope with the inherent risk of innovation.

Using bivariate probit estimations, Rouvinen (2002) suggests that one of the most significantly driver of the product and process innovations for Finnish manufacturing is the cooperation with non-academic outside partners. Hall, Lotti, and Mairesse (2009) have shown, in the case of Italy, that firm size, R&D intensity and investment in equipment increase the probability of having process and product innovation in enterprises.

Molero and Garcia (2008) used a logistic regression model to analyse the factors affecting product and process innovation at sector activity level, and considered the following factors: enterprise size, non-R&D innovative expenditures; own R&D; product innovation on inner effort; human resources; public funds; firm size; external knowledge integration; EU funds attraction; basic research and cooperation, and process innovation on external sources. Tiwari et al. (2007) argued that the following hampering factors: firm size, market share, cooperative arrangement, financial constraints, market uncertainty and institutional bottlenecks, regulations and organizational rigidities affect a firm's decision on R&D investment. Mohnen et al. (2008) examine the importance of financial aspects on innovation process in the Netherlands enterprises.

Despite the wide interest and the large body of the contributions on innovation drivers and obstacles, we are far from consent in what concerns the relevant determinants that should be considered. In this respect, data availability is one of the most important constraints. The present paper aims to analyse the influence of hampering and obstacle factors measured thought the 2012 and 2010 Community Innovation Survey (CIS) at Romanian level. The underlying assumption in the factors selection process was that the enterprises are aware of the problems faced in the product innovative process. Consequently, for each of the hampering and obstacle factors we asked the enterprises to rank the levels of influence of the

factors found to have an impact on product innovation. Our logistic regression model examines the relationship between enterprises status (product innovative or non-innovative) and the hampering and obstacle factors by the ranked level of intensity (enterprises effective responses).

3. Methodology

In the previous chapter it was noticed that a variety of variables were considered by scholars in an attempt to analyse different types of innovations, and logistic regression models have successfully been used for this purpose (e.g. Sternberg and Olaf, 2001; Reichstein and Salter, 2006; Vega - Jurado, et al., 2008; Buckley and Chatterjee, 2012; etc.).

Buckley and Chatterjee (2012) analysed product innovation through a logistic regression model by using cross sectional Australian data at firm level, from the official statistic survey. Three different periods were taken into account: 2005 – 2006, 2006 – 2007 and 2005 – 2007. The independent variables chosen in the models were constrained by the common available data across all the three time periods, such as: business plan, compared competitors performance, firm size, market share, operated locally, state-wide, nationally etc. The findings show that is more likely to innovate for firms with a large market share, facing stronger competition, participating in a franchising agreement, in contrast, the preoccupation for financial aspects have no positive influence on innovation.

Our selected independent variables (hampering and obstacle factors) capture the main concerns of scholars regarding the factors affecting innovation process, such as: competition, financial aspects, costs and firm characteristics. The list of the independent variables considered in this research is presented in Appendix 1, as well as the codes in R.

The dependent variable is the probability of an enterprise to introduce new or significantly improved product on the market (product innovation) and, in contrast to the Buckley and Chatterjee (2012) it wasn't investigated in consecutive periods. Since the section of obstacles factors was investigated in 2012 and the hampering factors were studied in 2010, no longitudinal analyse was possible.

To select the variables that will be introduced in the model, we used the Akaike Information Criterion (AIC) and the first logistical regression model, based on the selected CIS 2012 data (see Appendix 1), was:

$$Y_i = \beta_0 + \beta_1 \times \text{EMP12F} + \beta_2 \times \text{OBSPR} + \beta_3 \times \text{OBSQL} + \beta_4 \times \text{OBSCP} + \beta_5 \times \text{OBSFIN} + \varepsilon_i \qquad (1)$$

For the CIS 2010 data, according to the same AIC, the recommended variables are the following:

$$Y_i = \beta_0 + \beta_1 \times EMP10F + \beta_2 \times TURN08F + \beta_3 \times HFENT + \beta_4 \times HFOUT + \beta_5 \times HCOS +$$
$$+ \beta_6 \times HINF + \beta_7 \times HPAR + \beta_8 \times HDOM + \beta_9 \times HDEM + \beta_{10} \times HPRIO + \beta_{11} \times HMAR + \varepsilon_i \quad (2)$$

By conducting the two logistic models we pursue to identify which of the following are factors that influence the probability of an enterprise to introduce new or significantly improved product on the market (product innovation): cost, knowledge, market, financial factors and not to innovate reason, expressed as hampering and obstacle factors within the Community Innovation Survey.

4. Data description

We conducted our analysis based on the Community Innovation Survey (CIS) 2012 and 2010, micro data provided by the Romanian National Institute of Statistics. Each of the two statistical waves provides its specific angles for investigating the determinants of product innovation in Romanian enterprises. For the 2012 wave, we focus the analysis on factors as obstacles related to meeting enterprise's goals, while for the 2010 wave, we consider factors hampering innovation activities.

The variable used were encoded according to Eurostat, using the same corresponding codes from the harmonised survey questionnaire of CIS (see Appendix 1).

The sample for both waves is representative for the entire population of the Romanian firms. The sample is drown from the national Statistical Business Register (REGIS) and includes all the active enterprises with more than 9 employees. According to the CIS methodology, the sampling is random with known probability of selection, reported to the total population according to NACE Rev. 2, enterprises belonging to sections A to M (all enterprises, organizations and institutions whose business activity is the production of goods and services to the market) (INS, metadata, CIS).

For the descriptive statistics and statistical modelling we have used the unweight data. Appendix 2 shows descriptive statistics for the variables of the two statistical waves used in the data analysis.

5. Data analysis

The analysis was conducted in R; to assess the prediction performance of the models, the dataset was split in a train set and a test set. We used the *caTools package* to split randomly

the samples of the two statistical waves, such that the train set contains 70% of the data, and the test set 30% of the total number of observations.

We fitted logistic regression models on the train data (70% enterprises randomly chosen from the main sample of CIS 2012 and 2010) in order to analyse the situation of product innovative enterprises (dependent variable = inpdgd). The *stepwise procedure* from the *nnet package* in R helped to decide the set of the relevant predictors for each of the two models: obstacles factors influencing the product innovative enterprises, and hampering factors influencing the product innovative enterprises. The *stepwise procedure* uses the Akaike information criterion to choose the best model fit by selecting the most significantly variable for the two regression models (Kadane and Lazar, 2004). The results indicated for the first model the following predictors: strong price competition (obspr), innovations by competitors (obscp), lack of adequate finance (obsfin), strong competition on product quality, reputation or brand (obsql), and type of enterprise by the average number of employees in 2012 (emp12f). Then, we predict the first model on the test set. Table 2 shows a high level of accuracy.

Table 1: Model 1: Obstacles factors influencing the product innovative enterprises (train set from the CIS 2012 sample)

Covariates of the model	**Odds Ratio**	**p-value**
(Intercept)	0.01811629	< 2e-16 ***
EMP12F (ref. SMALL)		
MEDIUM	3.05950816	7.24e-05 ***
LARGE	4.32078351	1.80e-05 ***
OBSPR (ref. MEDIUM)		
HIGH	0.64007041	0.37427
LOW	1.86361751	0.02576 *
NOT RELEVANT	0.63566188	0.53438
OBSQL (ref. MEDIUM)		
HIGH	2.04528874	0.03145 *
LOW	0.55091975	0.05990 .
NOT RELEVANT	1.40625838	0.58551
OBSCP (ref. MEDIUM)		
HIGH	0.68096625	0.18201
LOW	1.72272479	0.13733
NOT RELEVANT	0.29071707	0.00444 **
OBSFIN (ref. MEDIUM)		
HIGH	1.43763481	0.22721

LOW	0.89365261	0.75389
NOT RELEVANT	2.450513	0.01158 *

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

Table 2: The accuracy of the logistic regression model 1 (test set from the CIS 2012 sample)

	Threshold		
	0.2	0.3	0.4
ACC	0.936221	0.949459	0.954272

For the second model the following predictors were indicated: type of enterprise by the average number of employees in 2010 (emp10), enterprise's total turnover for 2008 by quintiles (turn08), lack of funds within your enterprise or group (hfent), lack of finance from sources outside the enterprise (hfout), innovation costs too high (hcos), lack of information on markets (hinf), difficulty in finding cooperation partners for innovations (hpar), market dominated by established enterprises (hdom), uncertain demand for innovative goods (hdem), no need due to prior innovations of the enterprise (hprior) and no need because of no demand for innovations (hmar).

Table 3: Model 2: Hampering factors influencing the product innovative enterprises (train set from the CIS 2010 sample)

Covariates of the model	**Odds Ratio**	**p-value**
(Intercept)	0.05745444	< 2e-16 ***
EMP10F (ref. SMALL)		
MEDIUM	1.87332058	2.44e-08 ***
LARGE	3.74419184	< 2e-16 ***
TURN08F (ref. Q3)		
Q0	0.88329506	0.718482
Q1	0.76162342	0.109217
Q2	0.71325535	0.032519 *
Q4	1.37079349	0.017936 *
Q5	1.43487156	0.009848 **
HFENTN (ref. Medium)		
High	0.77627242	0.067081 .
HINF (ref. Medium)		
High	1.94170648	0.000582 ***
Low	0.96698531	0.788895
Factor not experienced	1.03057527	0.862335
HPAR (ref. Medium)		
High	0.7127662	0.056761 .
Low	0.94858737	0.699577
Factor not experienced	1.18811944	0.280306
HDOM (ref. Medium)		
High	1.242416	0.126225
Low	0.96792711	0.809484

Low	0.82835022	0.235519
Factor not experienced	0.34639851	1.89e-07 ***
HFOUT (ref. Medium)		
High	1.57687383	0.003304 **
Low	1.15695192	0.367487
Factor not experienced	1.83980786	0.000319 ***
HCOS (ref. Medium)		
High	0.63848083	0.001040 **
Low	1.0274279	0.85841
Factor not experienced	0.73934982	0.109522
Factor not experienced	0.6866063	0.046162 *
HDEM (ref. Medium)		
High	1.22605202	0.154808
Low	0.87122499	0.317564
Factor not experienced	0.56144187	0.001664 **
HPRIOR (ref. Medium)		
High	0.70499573	0.163216
Low	1.37277663	0.062051 .
Factor not experienced	0.82791118	0.34735
HMAR (ref. Medium)		
High	0.5582566	0.019705 *
Low	2.13562852	1.15e-05 ***
Factor not experienced	2.78117184	4.98e-07 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

Table 4: The accuracy of the logistic regression model 2 (test set from the CIS 2010 sample)

	Threshold		
	0.2	0.3	0.4
ACC	0.811533	0.86726	0.880805

Similar to the first model, the second predict model was conducted on the test set. Comparing the second regression model with the associated baseline model, the data shows a higher level of accuracy for all the tree thresholds tested. For a 0.4 threshold the accuracy (ACC) is 0.88.

6. Discussions

A first observation deriving from both models is related to one of the Schumpeterian hypothesis; in this case it was identified a significant relationship between firm size and product innovative activity at the enterprise level. For example, the results of the first model indicate an increase for *the medium enterprises* of the logarithm by 1.1183 within the relative probability of being innovative vs. non-innovative. The *odds ratio* (OR) reveals that the

probability of a medium enterprise to introduce innovative products on the market is 3 times higher by comparing to a small enterprise. Model 1 shows also that a large enterprise is 4 times more likely to be a product innovative enterprise, compared with a small enterprise.

Model 1 indicates significant statistical relationship between the dependent variable and the following predictors: large enterprises (EMP12F - LARGE), medium enterprises (EMP12F - MEDIUM), low importance of strong price competition (OBSPR - LOW), low and high importance of strong competition on product quality, reputation or brand (OBSQL - LOW and OBSQL - HIGH), no relevant importance of innovations by competitors (OBSCP - NOT RELEVANT) and no relevant importance of lack of adequate finance (OBSFIN - NOT RELEVANT).

The significant predictors for *model 2* are: large enterprises (EMP10F - Large); medium enterprises (EMP10F - Medium); second, fourth and fifth quintile of enterprise's total turnover for 2008 (TURN08F - Q2, TURN08F - Q4 and TURN08F - Q5); factor not experienced and high degree of importance for lack of funds within the enterprise or group (HFENT - Factor not experienced and High); factor not experienced and high importance for lack of finance from sources outside the enterprise (HFOUT - Factor not experienced and High); high importance for innovation costs too high (HCOS - High); high importance for lack of information on markets (HINF - High); high difficulty in finding cooperation partners for innovation (HPAR - High); factor not experienced regarding the market dominated by established enterprises (HDOM - Factor not experienced); factor not experienced regarding the uncertainty of demand for innovative goods or services (HDEM - Factor not experienced); low importance regarding the factor *no need due to prior innovations by the enterprise* (HPRIOR – Low) and the factor *no need because of no demand for innovations* (HMAR – High, Low and Factor not experienced).

The results of the hampering factors analysis show a high influence of *costs* and of *no need of product innovation because of no demand for innovations* in preventing a firm from innovating. For a firm that gives greater weight to *lack of finance from sources outside the enterprise* and to *the lack of information on markets,* is more probably to introduce new or significantly improved goods. Also, the probability of a company to innovative goods is lower in the presence of *no-demand-for-innovation* and *lack-of-external-financing*.

By summarising the results for the first model (CIS 2012), we find that the likelihood of being a product innovative enterprise is higher among medium and large enterprises, when compared with the small ones. The model also shows that for enterprises that are not concerned with price competition, there is a higher likelihood to innovate products. For the enterprises highly competitive in product quality, reputation or brand, the probability to

innovate products also increases. Those firms that are considering the innovations of the competitors as irrelevant are less likely to be product-innovative. The companies were the *lack-of-finance-adequacy* factor has *no relevant* importance, the probability to innovate products is two times higher when compared to those were the factor has a *medium* importance.

Comparing to the two models (CIS 2010 and CIS 2012) we conclude that both of them supports the idea that firm size and financial constraints influence product innovations at enterprise level. A closer look on the analysis conducted on CIS 2010 data completes this portrait with the importance played by enterprise turnover. Therefore, for the enterprises with a turnover situated in the last two quintiles is more probably to innovate, than for those with a turnover situated in the first two quintiles. The overall picture of the enterprise characteristics confirms what we were expected – developed enterprises are more likely to introduce new or significantly improved goods.

The accuracy of the both models is very high, which means that a prediction based on the models is involving a low rate of risk. In other words, having the necessary information about several characteristics of the enterprise, its financial situation, its market and competitors, our model will predict with a high rate of success if an enterprise will introduce new or significantly improved goods.

As the false positive and false negative values of the predicted conditions are low, the manifested errors can be neglected, in addition it can be said that the both models (especially the first one) are having a low false discovery rate and false omission rate.

Although the two models contribute towards increased awareness regarding the factors that influence product innovation, the limitations of the models should be considered. The use of logistic regression models in this study is a consequence of data characteristics and organization. The difficulty of interpreting the results is, however, one of the limitations of any logistic analysis. Another standard limitation of such a model lies in the implicit assumption of a linear relationship between variables. At this stage, our results are in line with previous findings in the literature, however future research may use regression trees for the obstacle factor analysis and structural equation modelling for hampering factors, in order to provide a more intuitive interpretation of the results.

7. Conclusions

In this article we used logistic regression models to explain the likelihood for an enterprise to be product innovative based on two sets of factors. We introduced new sets of

independent variables that summarize the factors analysed in previous contributions: enterprise characteristics, costs, knowledge, market and reasons not to innovate (Rouvinen, 2002; Canepa and Stoneman, 2007; Tiwari et al., 2007; Mohnen et al., 2008; Pires-Alves and Rocha, 2008; Hall, Lotti, and Mairesse, 2009).

The literature review acknowledges the difficulty of selecting the appropriate innovation drivers when analysing the innovation process at enterprise level: in our study, we choose to use factors related to costs, knowledge, market, and reasons not to innovate.

Our findings confirm previous results obtained by other researchers regarding the influence of firm size, financial constraints, market and competition, knowledge on product innovation, by using new sets of factors that are covering the general concerns encountered in literature, related to the forehead mentioned aspects. Summarizing, it is more likely to create new or significantly improved goods for the enterprises with a large number of employees, for those with low financial constraints, active on a strong competition market and for those with access to information on markets. The high level of the model's accuracy encourages us to consider it a proper model for future debates on policies related to support enterprises to innovate. Last, but not least, the models are built on data available through the official statistics, which makes the current analyse easy to replicate.

Disclaimer:
NIS Romania has no responsibility for the results and conclusions of the research.

References:

- Adam, R. (2015). The Romanian Product Innovative Enterprises - A Statistical Analysis Using "R". *Global Economic Observer Journal*, 3(1), pp. 101 - 107.
- Allen, J. A. (1967). *Studies in Innovation in the Steel and Chemical Industries*. Manchester: University of Manchester Press.
- Buckley C. and Chatterjee, I. (2012). *Drivers of Product Innovation by Australian SMEs: The Importance of Managerial Attitude*. Department of Economics, University of Western Australia, WA 6009.
- Canepa, A. and Stoneman, P. (2007). Financial constraints to innovation in the UK: evidence from CIS2 and CIS3. *Oxford University Press*, 60 (4), pp. 711 - 730.
- Fagerberg, J. (1994). Technology and international differences in growth rates. *Journal of Economic Literature*, 32, pp. 1147–1175.
- Hall, B.H., Lotti, F. and Mairesse, J. (2009). Innovation and productivity in SMEs: empirical evidence for Italy. *Small Business Economics*, 33(1), pp. 13-33.

- Huang, C., Arundel, A. and Hollanders, H. (2010). *How Firms Innovate: R&D, Non-R&D, And Technology Adoption*. UNU-MERIT Working Papers, ISSN 1871-9872.
- Jaffe, A.B. (1986). Technological opportunity and spillover of R&D: Evidence from firms' patents, profits and market value. *American Economic Review*, Vol. 76, pp. 984 - 1001.
- Kadane, J. B. and Lazar, N.A. (2004). Methods and Criteria for Model Selection. *Journal of the American Statistical Association*, 99, pp. 279–290.
- Lehtoranta, O. (2005). *Determinants of Innovation and the Economic Growth of Innovators*, [Online]. (http://www.vtt.fi/inf/pdf/workingpapers/2005/W25.pdf). VTT Working Papers 25. (Accessed: 26.10.2015)
- Mansfield, E. (1962). Entry, Gibrat's law, innovation and the growth of firms. *American Economic Review*, 52, pp. 1023 - 1051.
- Mansfield, E. (1980). Basic Research and Productivity Increase in Manufacturing. *American Economic Review*, 70, pp. 863 - 73.
- Mohnen, P., Palm, F. C., Schim van der Loeff, S. and Tiwari. A. (2008). Financial Constraints and Other Obstacles: are they a Threat to Innovation Activity?. *De Economist*, 156(2), pp. 201 - 214.
- Moler, J. and Garcia, A., (2008). *Factors affecting innovation revisited*, WP05/08.
- O'Sullivan, D. and Dooley, L. (2008). *Applying Innovation*. SAGE Publications.
- Pires-Alves, C. and Rocha, F. (2008). *Testing the Schumpeterian Hypotheses for the Brazilian Manufacturing Industry, Brazilian Association of Graduate Programs in Economics,* Working Paper nr 200807091629490.
- Reichstein, T and Salter, A. (2006). Investigating the sources of process innovation among UK manufacturing firms. *Industrial and Corporate Change*, 15 (4), pp. 653 - 682.
- Rothwell, R., and A. B. Robertson (1973). The Role of Communications in Technological Innovation, Research Policy, 2, pp. 204 - 225.
- Rouvinen, P. (2002). Characteristics of product and process innovators: some evidence from the Finnish innovation survey. *Applied Economics Letters*, 9(9), pp. 575 - 580.
- Schumpeter, J.A. (1934). *The Theory of Economic Development, Cambridge*. MA, Harvard University Press.
- Smith, A. (1776). *An Inquiry into the Nature and Causes of the Wealth of Nations*. Modern Library edition, NY, NY, 1937.

- Smith, K. 2005. Measuring innovation. In: Fagerberg, J., Mowery, D. and Nelson, R., eds. *The Oxford Handbook of Innovation*. Oxford University Press, Ch. 6, pp. 148-177.
- Solow, R. M. (1957). Technical Change and the Aggregate Production Function. *Review of Economics and Statistics,* 39, pp. 312 - 320.
- Sternberg, R and Arndt, O. (2001). The Firm or the Region: What Determines the Innovation Behavior of European Firms?. *Economic Geography,* 77(4), pp. 364 - 382.
- Tiwari, A., Mohnen, P., Palm, F. and Schim van der Loeff, S. (2007). *Financial Constraint and R&D Investment: Evidence from CIS*. UNU-MERIT.
- Vega-Jurado, J. (2008). The effect of external and internal factors on firms product innovation. *Research Policy*, 37(4), pp. 616 - 632.
- ***Eurostat. *CIS 2012 Survey Questionnaire*. http://ec.europa.eu/eurostat/documents/203647/203701/Harmonised+survey+questionnaire+2012/164dfdfd-7f97-4b98-b7b5-80d4e32e73ee (Accessed: 29.08.2015).
- ***Eurostat. *CIS 2010 Survey Questionnaire*. http://ec.europa.eu/eurostat/documents/203647/203701/CIS_Survey_form_2010.pdf/b9f2c70e-0c46-4f82-abeb-c7661f1f2166 (Accessed: 29.08.2015).
- *** European Commission. *The 2014 EU Industrial R&D Investment Scoreboard*. http://iri.jrc.ec.europa.eu/scoreboard14.html (Accessed: 30.08.2015).
- ***European Commission. http://ec.europa.eu/research/innovation-union/index_en.cfm?pg= why (Accessed: 30.08.2015).
- *** INS, *metadata*, CIS. colectaredate.insse.ro/metadata (Accessed: 29.08.2015).
- *** OECD. *Science, Technology and Industry Scoreboard 2005*. http://www.oecd-ilibrary.org/science-and-technology/oecd-science-technology-and-industry-scoreboard-2005_sti_scoreboard-2005-en

Appendix 1: The codes of the CIS data

a. CIS 2012

Codes	Variable	Values
inpdgd	New or significantly improved goods (exclude the simple resale of new goods and changes of a solely aesthetic nature)	Yes = 1 No = 0

Codes	Variable	Values	
emp12f	Enterprise's average number of employees in 2012	Small	10 – 49
		Medium	50 – 249
		Large	> 250

Codes	Variable	Values	
turn10f	Enterprise's total turnover for 2010	Q1	< 2611136 RON
		Q2	(2611136 – 6624714] RON
		Q3	(6624714 – 16030753] RON
		Q4	(16030753 – 43830079] RON
		Q5	> 43830079 RON

Codes	Variable	Degree of Importance			
		High	Medium	Low	Not relevant
obspr	Strong price competition	3	2	1	0
obsql	Strong competition on product quality, reputation or brand	3	2	1	0
obslde	Lack of demand	3	2	1	0
obscp	Innovations by competitors	3	2	1	0
obsdmk	Dominant market share held by competitors	3	2	1	0
obsprs	Lack of qualified personnel	3	2	1	0
obsfin	Lack of adequate finance	3	2	1	0
obsamk	High cost of access to new markets	3	2	1	0
obsreg	High cost of meeting government regulations or legal requirements	3	2	1	0

b. CIS 2010

Codes	Variable	Values
inpdgd	New or significantly improved goods (exclude the simple resale of new goods and changes of a solely aesthetic nature)	Yes = 1 No = 0

Codes	Variable	Values	
emp10f	Enterprise's average number of employees in 2010	Small	10 – 49
		Medium	50 – 249
		Large	> 250

Codes	Variable	Values

turn08f	Enterprise's total turnover for 2008	Q0	without turnover
		Q1	< 2188339 RON
		Q2	(2188339 – 5901739] RON
		Q3	(5901739 – 13296909] RON
		Q4	(13296909 – 36578005] RON
		Q5	> 36578005 RON

Codes	Variable	Degree of Importance			
		High	Medium	Low	Factor not experienced
hfent	Lack of funds within your enterprise or group	3	2	1	0
hfout	Lack of finance from sources outside your enterprise	3	2	1	0
hcos	Innovation costs too high	3	2	1	0
hper	Lack of qualified personnel	3	2	1	0
htec	Lack of information on technology	3	2	1	0
hinf	Lack of information on markets	3	2	1	0
hpar	Difficulty in finding cooperation partners for innovation	3	2	1	0
hdom	Market dominated by established enterprises	3	2	1	0
hdem	Uncertain demand for innovative goods or services	3	2	1	0
hprior	No need due to prior innovations by your enterprise	3	2	1	0
hmar	No need because of no demand for innovations	3	2	1	0

Appendix 2: Descriptive statistics

Categ.	**Freq.**	**Freq. Cum**	**Freq. Relat.**
CIS2012			
inpdgd			
Yes	128	2770	0.046
No	2642	2642	0.954
emp12f			
Large	337	337	0.122
Small	1225	1562	0.442
Medium	1208	2770	0.436
turn10f			
Q1	554	554	0.200
Q2	554	1108	0.200
Q3	553	1661	0.200
Q4	554	2215	0.200
Q5	555	2770	0.200

Categ.	**Freq.**	**Freq. Cum**	**Freq. Relat.**
CIS2010			
inpdgd			
Yes	965	965	0.112
No	7660	8625	0.888
emp10f			
Large	1045	1045	0.121
Small	4113	5158	0.477
Medium	3467	8625	0.402
turn08f			
Q0	185	185	0.021
Q1	1688	1873	0.196
Q2	1688	3561	0.196
Q3	1688	5249	0.196
Q4	1688	6937	0.196

obspr			
Medium	1049	1049	0.379
Not relevant	265	1314	0.096
High	241	1555	0.087
Low	1215	2770	0.439
obsql			
Medium	1153	1153	0.416
Not relevant	364	1517	0.131
High	401	1918	0.145
Low	852	2770	0.308
obslde			
Medium	1207	1207	0.436
Not relevant	313	1520	0.113
High	402	1922	0.145
Low	848	2770	0.306
obscp			
Medium	980	980	0.354
Not relevant	743	1723	0.268
High	790	2513	0.285
Low	257	2770	0.093
obsdmk			
Medium	1151	1151	0.416
Not relevant	451	1602	0.163
High	569	2171	0.205
Low	599	2770	0.216
obsprs			
Medium	978	978	0.353
Not relevant	684	1662	0.247
High	885	2547	0.319
Low	223	2770	0.081
obsfin			
Medium	1107	1107	0.400
Not relevant	546	1653	0.197
High	621	2274	0.224

Q5	1688	8625	0.196
hfent			
Factor not exp.	2642	2642	0.306
Medium	2133	4775	0.247
High	2603	7378	0.302
Low	1247	8625	0.145
hfout			
Factor not exp.	3420	3420	0.397
Medium	1939	5359	0.225
High	1894	7253	0.220
Low	1372	8625	0.159
hcos			
Factor not exp.	3075	3075	0.357
Medium	2076	5151	0.241
High	2263	7414	0.262
Low	1211	8625	0.140
hper			
Factor not exp.	3139	3139	0.364
Medium	2309	5448	0.268
High	716	6164	0.083
Low	2461	8625	0.285
htec			
Factor not exp.	3352	3352	0.389
Medium	2164	5516	0.251
High	425	5941	0.049
Low	2684	8625	0.311
hinf			
Factor not exp.	3410	3410	0.395
Medium	2133	5543	0.247
High	478	6021	0.055
Low	2604	8625	0.302
hpar			
Factor not exp.	3830	3830	0.444
Medium	1978	5808	0.229
High	885	6693	0.103

Low	496	2770	0.179
obsamk			
Medium	1151	1151	0.416
Not relevant	533	1684	0.192
High	523	2207	0.189
Low	563	2770	0.203
obsreg			
Medium	1036	1036	0.374
Not relevant	661	1697	0.239
High	604	2301	0.218
Low	469	2770	0.169

Low	1932	8625	0.224
hdom			
Factor not exp.	3384	3384	0.392
Medium	2077	5461	0.241
High	1341	6802	0.155
Low	1823	8625	0.211
hdem			
Factor not exp.	3443	3443	0.399
Medium	2275	5718	0.264
High	1256	6974	0.146
Low	1651	8625	0.191
hprior			
Factor not exp.	4507	4507	0.523
Medium	1521	6028	0.176
High	706	6734	0.082
Low	1891	8625	0.219
hmar			
Factor not exp.	4195	4195	0.486
Medium	1574	5769	0.182
High	1064	6833	0.123
Low	1792	8625	0.208