ACTA UNIVERSITATIS LODZIENSIS
FOLIA OECONOMICA 194, 2005

*Marek Szajt**

IDENTIFICATION OF PATENT-ACTIVITY LEVEL WITH THE USAGE OF DISCRIMINANT ANALYSIS

Abstract

The subject, which is more and more frequently discussed in economic literature, is innovation. A lot of elaborations refer to its description and importance in modern economies. In this paper an attempt is made to separate particular country groups in Europe on the basis of patent activity. The division has been made with the usage of statistical methods – mainly discriminant function. The analysis presented in the paper allows characterizing particular participants and drawing one's attention to the differences in innovative policy conducted in different countries.

Key words: innovation, patent, discriminant function.

I. INTRODUCTION

Nowadays, a significant influence of innovation on economic growth is emphasized more and more often. Along with the forming of general economic theories, a lot of different definitions connected with innovations have been created. Today, an **innovation is usually understood as a relatively new production application of scientific or technical information** (Kot et al., 1993). This definition, although very simple, is clear and equivalent to many terms that have emerged in recent years.

In connection with new problems that arise, we also encounter difficulties connected with their precise quantification. First of all, as the main source of innovation, patents are adopted. Persons or institutions from abroad may submit the patents on the territory of a given country both by its residents and non-residents. The number of patents submitted by the residents reflects the activity of a given country in the sphere of research and

* Ph.D., student, Econometrics and Statistics Department, Technical University in Częstochowa.

development ($R + D$). In order to obtain a better comparability of data concerning the number of patents, the data is quantified per area units or the number of inhabitants.

Innovative activity measured as a number of patents per one thousand of inhabitants is influenced by various factors. Today, as the main determinants of innovative activity, the following factors are enumerated: first of all, expenditures for the $R + D$ sector, employment in this sector and the level of economic growth.

In this paper, an empirical analysis of the abovementioned problems was presented on the basis of information published by Główny Urząd Statystyczny (GUS – the Central Statistical Office) and Organisation of Economic Cooperation and Development (OECD).

II. CHARACTERISTICS OF INNOVATIVE ACTIVITY DETERMINANTS

It would be difficult to overlook the ever-increasing expenditures for research and development activity, both in Poland and in other countries. Despite the fact that Poland spends more money on R + D than results from the tendency that can be observed in the OECD countries (Żółkiewski, 1999), these expenses, due to a low level of national wealth do not have measurable effects.

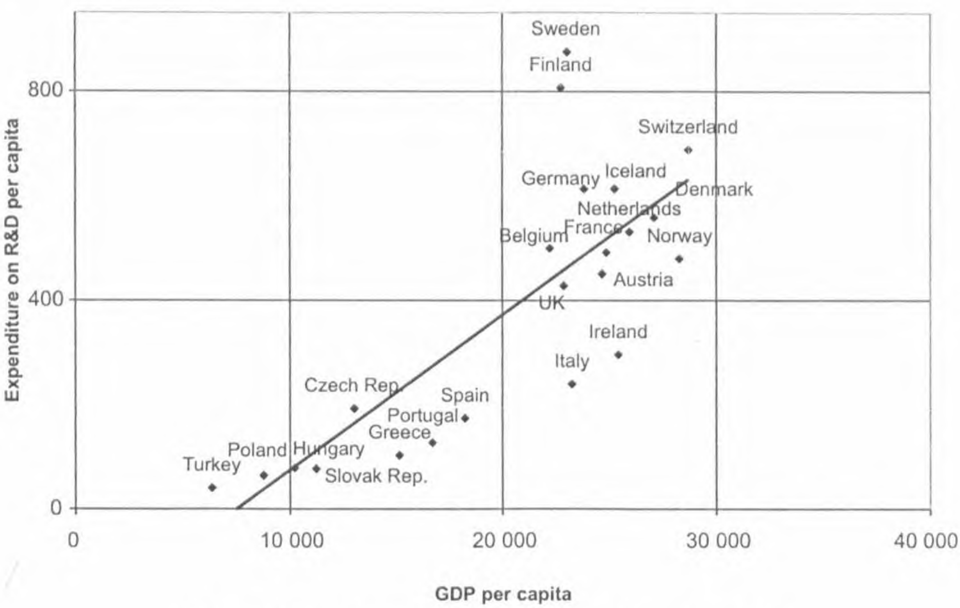

Chart 1. Expenditure for R&D activity and GNP (OECD)
Source: Author's calculations on the basis of GUS data

The straight line in the diagram is a form regression line:

$$y = -212.916 - 0.029x \qquad (1)$$

estimated for the data from the year 1998 concerning expenditures for R + D per capita (Y) in relation to PKB per capita (X) in the OECD countries. Even though fitting of the equation calculated with the determination coefficient $R^2 = 0{,}67$ is not convincing, the value of the parameter present at the independent variable significantly differs from zero. The countries that lie above this line are characterized by a tendency to invest that is higher than the OECD average. The countries lying below the line bear expenses that are lower than expected. Despite this, a certain flaw in our system is an unquestionable domination of budget resources in the financing of R + D activity. A higher share of the enterprise sector in the financing of R + D activity is expected in the future.

Poland has insufficiently developed structures of financing innovation by the enterprise sector, because of which the state budget is excessively burdened. Innovations are becoming increasingly dependent on effective interactions between the scientific base and the business sector.

The vertical lines in the diagram denote a division of countries from the point of view of society wealth calculated with the volume of GNP per 1000 inhabitants, calculated in actual prices according to purchasing power parity. In our case, we distinguished three groups of countries: countries characterized by the level of PKB/1000 inhabitants amounting to less than 10 000, countries characterized by the value between 10 000 and 20 000 and those characterized by the value over 20 000. These groups somehow determine the country's capability to absorb advanced technologies which are expensive because of their character, especially at the moment of their implementation.

The horizontal line divides countries into two groups from the point of view of the innovation financing volume. Below this line, there are countries in which expenses for R + D activity per inhabitant are lower than the average for a given group. It is a certain determinant of the capabilities of a given country and its scientists to create new technologies and inventions, which is usually connected with substantial cost of research and experiments. These costs, besides marketing costs, are the most important ones in the case of creation and introduction of a new product to the market.

The third equally important factor that has an influence on innovative activity is human capital, which we will present as a number of employees in the R + D sector. In order to obtain better comparability, this number was presented on full time basis. Above all research and development employees are taken into consideration because of their highest actual contribution to the creation of new inventions. Additionally, this number is converted into the capability of a given country from the point of view of employment with reference to 1000 employed persons.

In order to investigate the innovative activity, we can use the data concerning the investigated country or a group of countries. Unfortunately, there exist a lot of difficulties connected with the availability of complete data concerning individual countries and with comparability of the data. In this case, the only option is to use the methods of spatial and time analysis or to use analyses based on discrete programming. In this paper, we will present possibilities of use of discriminative analysis as a method for obtaining a division of countries according to the innovation activity criterion.

In the paper presented, a discriminative function estimated for the European members of OECD on the basis of the possessed information from the years 1995–1999 will be used. The sample consists of two p-dimensional normal distributions with expected values vectors $\mathbf{x}_1$ and $\mathbf{x}_2$ the same covariance matrix $\mathbf{S}$. The discriminative function will be $\mathbf{a}'\mathbf{x}$, where as the vector $\mathbf{a}$, we will adopt a vector that maximizes the expression (Morrison, 1990):

$$t^2(\mathbf{a}) = \frac{[\mathbf{a}'(\overline{x}_1 - \overline{x}_2)]^2 \frac{N_1 N_2}{N_1 + N_2}}{\mathbf{a}'\mathbf{S}\mathbf{a}} \tag{2}$$

where $\mathbf{a}'\mathbf{S}\mathbf{a} = 1$

Vector $\mathbf{a}$ is solution of the homogeneous system of equations:

$$[\frac{N_1 N_2}{N_1 + N_2}(\overline{x}_1 - \overline{x}_2)(\overline{x}_1 - \overline{x}_2)' - \lambda\mathbf{S}]\mathbf{a} = 0 \tag{3}$$

where:

$$\lambda = \max_{\mathbf{a}} {}^2(\mathbf{a}) = \frac{N_1 N_2}{N_1 + N_2}(\overline{x}_1 - \overline{x}_2)'\mathbf{S}^{-1}(\overline{x}_1 - \overline{x}_2) = T^2 \tag{4}$$

the matrix rank of this system equals $p-1$, which determines the following form of the linear discriminative function:

$$y = (\overline{x}_1 - \overline{x}_2)'\mathbf{S}^{-1}x. \tag{5}$$

Because of the comparable variance of the observed variables, we can move on to the estimation of the discriminative function.

The discriminative point in our investigation is:

$$population\ A < (\overline{x}_1 - \overline{x}_2)'\mathbf{S}^{-1}(\overline{x}_1 - \overline{x}_2) \leqslant population\ B.$$

As a criterion, we will use the Anderson classification statistic:

$$W = (\overline{x}_1 - \overline{x}_2)'\mathbf{S}^{-1}x - 0.5(\overline{x}_1 - \overline{x}_2)'\mathbf{S}^{-1}(\overline{x}_1 + \overline{x}_2) \tag{6}$$

where: x belongs to population 1 (having low innovative activity) when $W < 0$, and to population 2 (having high innovative activity) when $W > 0$.

As the criterion, we will adopt innovative activity measured as a number of patents submitted by the residents per 1000 inhabitants. We treat the activity as too low (0), when it amounts to less than 0.1 patent per 1000 employed persons, and sufficient (1) when it exceeds the value of 0.1 patent employed person. The feature investigated will depend on three features:

X_{1i} – gross expenditures for the R + D activity per 1000 inhabitants according to the purchasing power parity in $ in actual prices from 1999 for a given country i,

X_{2i} – number of research and development employees per 1000 employed persons for a given country i,

X_{3i} – GNP according to the purchasing power parity in $ per one inhabitant in actual prices from 1999 for a given country i.

Even though in the contents we have taken the year 1999 as the base year, the analyses were carried out on the basis of data from different years, while the data from the year 1999 was presented as the most up-to-date and thus giving the clearest results.

III. RESULTS OF ESTIMATION

In the first phase of the calculation, it turned out that the value of the parameter present at the variable responsible for wealth of the society – X_3 – is insignificant in comparison with the rest of the variables. None of the sample equations for the years from the period 1995–1999 did not confirm its significance. Therefore, in further samples, variables X_1 and X_2 were classified for the equation.

As a result of the estimation, the following discriminative function equation was obtained:

$$y_t = 2.166x_{1t} + 2.711x_{2t} \tag{7}$$

with the discrimination point of the value of 3.389. This equation was assessed through the replacement of individual values of explanatory variables with value 0 for $X_{ji} < 0.1$; 1 for $1 < X_{ji} < 0.4$; 2 for $X_{ji} > 0.4$. In the case of real data, the value of discriminative function parameters amounts to:

$$y_t = 1.388x_{1t} + 8.526x_{2t} \tag{8}$$

at the discriminative point = 4.20
The matching measured with accuracy coefficient (Aczel, 2000) is higher in the case of function (8) and amounts to 0.9524, whereas the proportional chances criterion = 0.528.

Table 1. Values of discriminative functions for selected countries

Country	Function value for equation 7	Function value for equation 8
Czech Republic	2.166	2.141
Belgium	4.877	4.011
Italy	2.166	2.879
Iceland	4.877	4.620
Accuracy coefficient	0.901	0.9524

Source: Author's own calculations.

Table 1 presents the values for the best (from the point of view of the discriminative function value) countries. As we can see, only in the case of Italy the discriminative function value does not agree with the assumptions of the equation (Italy belonged to countries of sufficient innovative activity).

Table 2. Values of discriminative function at simulative values of explanatory variables

No.	X_1	X_2	Function value	No.	X_1	X_2	Function value
1	0	0.1	0.853	11	0.2	0.3	**4.836**
2	0	0.2	1.705	12	0.2	0.4	**5.688**
3	0	0.3	2.558	13	0.3	0.1	**4.269**
4	0	0.4	3.411	14	0.3	0.2	**5.122**
5	0.1	0.1	1.991	15	0.3	0.3	**5.974**
6	0.1	0.2	2.844	16	0.3	0.4	**6.827**
7	0.1	0.3	3.697	17	0.4	0.1	**5.408**
8	0.1	0.4	**4.549**	18	0.4	0.2	**6.260**
9	0.2	0.1	3.130	19	0.4	0.3	**7.113**
10	0.2	0.2	3.983	20	0.4	0.4	**7.966**

Source: Author's own calculations.

It is necessary to stress that the data for Italy – especially the data concerning domestic patents – is incomplete, which may result in a lack of consistency in their presentation. The next table presents simulative values corresponding to individual changes in the value of explanatory changes. In the simulations, values similar to those real were taken into consideration.

The above calculations (Tab. 2) prove that only an appropriate number of research and development employees working on the basis of firm financial basis guarantees a high level of patent activity. It is worth emphasizing that from observation 8 it follows that, despite the correct value of the discriminative function at relatively low expenditures we may talk about a very low efficiency of the R + D personnel in this case. With correct financing, one fourth of this personnel can generate a similar number of patents. It is worth emphasizing the increasing efficiency of researchers in the case of an increase in expenditures, which is not connected with an increase in salary, but rather with an increase in the ability for purchasing research equipment and financing of experiments that are often expensive.

IV. CONCLUSIONS

The presented way of identification of patent activity is characterized by a few advantages:

– it has a better match in comparison with regression functions and arbitrary classification,

– omits outcome of time effects, which results in simplicity of calculations,

– allows adopting of data of lower accuracy,

– quickly reacts to distinct changes of explanatory variables,

– facilitates planning and provides information about the existing flaws in innovative policy.

The presented method of patent activity identification based on linear discriminative function may serve as one of the elements of innovation analysis. Once again, a simple statistical instrument allows the drawing of concrete and, I hope accurate economic conclusions.

REFERENCES

Aczel. A.D. (2000), *Statystyka w zarządzaniu*, PWN, Warszawa.

Jajuga K. (1993), *Statystyczna analiza wielowymiarowa*, PWN, Warszawa.

Kot S.M., Karska A., Zając K. (1993), *Matematyczne modele procesów dyfuzji innowacji*, PWN, Warszawa.

Morrison D.F. (1990), *Wielowymiarowa analiza statystyczna*, PWN, Warszawa.

Żółkiewski Z. (ed.), (1999), *Rachunek satelitarny nauki 1996–1997*, GUS, Warszawa.

Marek Szajt

IDENTYFIKACJA POZIOMU AKTYWNOŚCI PATENTOWEJ Z WYKORZYSTANIEM ANALIZY DYSKRYMINACYJNEJ

Streszczenie

Tematem coraz częściej poruszanym w literaturze ekonomicznej są innowacje. Wiele opracowań dotyczy ich opisu czy zwrócenia uwagi na ich rolę we współczesnych gospodarkach. W pracy podjęto próbę wyodrębnienia poszczególnych grup państw w Europie z punktu widzenia aktywności patentowej. Podział dokonany został z wykorzystaniem metod statystycznych – w tym głównie funkcji dyskryminacyjnej. Przedstawiona analiza pozwala charakteryzować poszczególnych uczestników oraz zwrócić uwagę na różnice w prowadzonej polityce innowacyjnej w różnych państwach.