ΤΕΤΡΑΔΙΑ ΑΝΑΛΥΣΗΣ ΔΕΔΟΜΕΝΩΝ, ΤΕΥΧΟΣ 16 *(σσ. 91-103)*
DATA ANALYSIS BULLETIN, ISSUE 16 *(pp. 91-103)*

A model for the total intramural research and development (R&D) expenditure (GERD) by sectors of performance within the EU21

Joel C. Nwaubani, Nikos Kapoulas, Georgia Lepida

Department of Applied Informatics,
University of Macedonia, Greece

Abstract

In this study, we consider and estimate the most accurate association model of the Categorical Data Analysis (CDAS) for the total intramural research and development (R&D) expenditure (GERD) by sectors of performance in 21 countries of the EU. The data used in this study are obtained from Eurostat statistical surveys which are regularly conducted at national level covering R & D performing entities in the private and public sectors estimated on actual base year from 2000-2010. Since the main focus is to have a better understanding of the total intramural research and development (R&D) expenditures within the 21 countries of the European Union (EU) according to the sector of performance and the source of funds, the Analysis of Association table (ANOAS) is given in order to ascertain the percentage of the data which is covered by each model. We estimate the association model to find the model that has the best fit but none proved accepted. Therefore, we proceeded to the multivariate model to find the model with the best fit and in conclusion we find out that the multivariate Row-Column Effects Association Model (RC) of the (M = 7) has the best fit among all.

Keywords: Association models, Log-linear & Log - nonlinear models, R&D, Gross domestic expenditure (% share of GDP), EU21

1. Introduction

Research and development (R&D) expenditures are the main aggregate used for international comparisons for gross domestic expenditure on (GERD). GERD data and their components are compiled on the basis of the OECD methodology. It is a creative work that are performed within a statistical unit or sector of the economy and undertaken on a systematic basis in order to increase the stock of knowledge, including knowledge of man, culture and society, and the use of this stock of knowledge to devise new applications. One of the most importantl goals of the EU during the last decade has been to encouraging and increase the levels of investment in order to provide an incentive to the EU's competitiveness. The Lisbon strategy set the EU an objective of devoting 3 % of its gross domestic product (GDP) to research and development activities. Following close cooperation with the EU Member States, statistics on science, technology and innovation are based on the European Parliament and of the Council concerning the production and development of Community statistics on science and technology. The Decision was implemented by the European Commission Regulation in 2004

The total research and development (GERD) is divided into four sectors of performance: These four sectors are the business enterprise sector, the government sector, the higher education sector, and the private non-profit sector. GERD is often reported in relative terms as a percentage of GDP, to denote the R&D intensity of an economy. Regional R&D intensity is defined as total intramural expenditures regardless of the source of funds performed in the region in a given year, and are usually expressed in relation to GDP or in relation to population (Fairbanks & Michael, 2005).

1.3. Data Processing and the results

We start by imputing and processing of our observations (data), which are the total intramural R&D expenditure (GERD) by sectors of performance in 21 countries of the EU for the period of 2001-2010.

Table 1. Data showing the total intramural R&D expenditure (GERD) by sectors of performance unit mio_eur Millions of euro from 2001-2010

	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
EU-27	1.86	1.87	1.88	1.87	1.83	1.83	1.85	1.85	1.92	2.01	2.00
Euro area	1.84	1.86	1.88	1.87	1.85	1.84	1.87	1.88	1.96	2.06	2.06
Belgium	1.97	2.07	1.94	1.87	1.86	1.83	1.86	1.89	1.97	2.03	1.99
Bulgaria	0.51	0.46	0.48	0.48	0.49	0.46	0.46	0.45	0.47	0.53	0.60
Czech Republic	1.17	1.16	1.15	1.20	1.20	1.35	1.49	1.48	1.41	1.48	1.56
Denmark (1)	2.24	2.39	2.51	2.58	2.48	2.46	2.48	2.58	2.85	3.06	3.06
Germany	2.47	2.47	2.50	2.54	2.50	2.51	2.54	2.53	2.69	2.82	2.82
Estonia	0.60	0.70	0.72	0.77	0.85	0.93	1.13	1.08	1.28	1.43	1.62
Ireland	1.11	1.09	1.09	1.16	1.22	1.24	1.24	1.28	1.45	1.74	1.79
Greece	:	0.58	:	0.57	0.55	0.60	0.59	0.60	:	:	:
Spain	0.91	0.92	0.99	1.05	1.06	1.12	1.20	1.27	1.35	1.39	1.39
France (2)	2.15	2.20	2.24	2.18	2.16	2.11	2.11	2.08	2.12	2.26	2.26
Italy	1.04	1.08	1.12	1.10	1.09	1.09	1.13	1.17	1.21	1.26	1.26
Cyprus	0.25	0.26	0.30	0.35	0.37	0.41	0.43	0.44	0.43	0.49	0.50
Latvia	0.45	0.41	0.42	0.38	0.42	0.56	0.70	0.60	0.62	0.46	0.60
Lithuania	0.59	0.67	0.66	0.67	0.75	0.75	0.79	0.81	0.79	0.83	0.79
Luxembourg	1.65	:	:	1.65	1.63	1.56	1.66	1.58	1.57	1.66	1.63
Hungary (3)	0.81	0.93	1.00	0.94	0.88	0.94	1.01	0.98	1.00	1.17	1.16
Malta (3)	:	:	0.26	0.25	0.53	0.57	0.62	0.58	0.56	0.54	0.63
Netherlands	1.94	1.93	1.88	1.92	1.93	1.90	1.88	1.81	1.77	1.82	1.83
Austria	1.93	2.05	2.12	2.24	2.24	2.46	2.44	2.51	2.67	2.72	2.76
Poland	0.64	0.62	0.56	0.54	0.56	0.57	0.56	0.57	0.60	0.68	0.74
Portugal	0.73	0.77	0.73	0.71	0.75	0.78	0.99	1.17	1.50	1.64	1.59
Romania	0.37	0.39	0.38	0.39	0.39	0.41	0.45	0.52	0.58	0.47	0.47
Slovenia (4)	1.38	1.49	1.47	1.27	1.39	1.44	1.56	1.45	1.65	1.86	2.11
Slovakia	0.65	0.63	0.57	0.57	0.51	0.51	0.49	0.46	0.47	0.48	0.63
Finland	3.35	3.32	3.36	3.44	3.45	3.48	3.48	3.47	3.70	3.92	3.87
Sweden (5)	:	4.13	:	3.80	3.58	3.56	3.68	3.40	3.70	3.61	3.42
United Kingdom	1.81	1.79	1.79	1.75	1.68	1.73	1.75	1.78	1.79	1.86	1.77
Iceland	2.67	2.95	2.95	2.82	:	2.77	2.99	2.68	2.64	3.11	:
Norway	:	1.59	1.66	1.71	1.58	1.52	1.49	1.62	1.61	1.80	1.71
Switzerland	2.53	:	:	:	2.90	:	:	:	2.99	:	:
Croatia	:	:	0.96	0.96	1.05	0.87	0.75	0.80	0.89	0.83	0.73
Turkey	0.46	0.51	0.51	0.47	0.51	0.58	0.57	0.71	0.73	0.85	:
Japan (4)	3.04	3.12	3.17	3.20	3.17	3.32	3.40	3.44	3.45	:	:
United States	2.69	2.71	2.60	2.60	2.53	2.56	2.60	2.66	2.79	:	:

(1) Break in series, 2007.
(2) Break in series, 2000 and 2004.
(3) Break in series, 2004.
(4) Break in series, 2008.
(5) Break in series, 2005.
Source: Eurostat (online data code: t2020_20), OECD

Among the EU Member States, the highest R&D intensities in 2010 were recorded in Finland (3.87 %), Sweden (3.42 %) and Denmark (3.06 %). There were eight Member States that reported R&D expenditure accounting for less than 1 % of their GDP in 2010; Greece also had a ratio of less than 1 % although its latest available data was for 2007. The Member States with the lowest R&D intensity were generally found in the Southern and Eastern Europe (OECD, 2010).

With the help of the Categorical Data Analysis program (CDAS), we were able to ascertain the results below.

Table 2. Results of the models

Models	Likelihood X^2	Likelihood G^2	Degrees of Freedom	Index of Dissimilariy	Maximum Deviation
O	263487.28526	283368.44412	140	0.32982	0.00000000
U	249478.11005	272362.14566	139	0.33127	0.00073327
R	663953516.13756	157974.15344	120	0.21894	0.15704920
C	242391.18767	236482.08352	133	0.26583	0.00090376
R+C	206225025.37529	141006.84843	114	0.16981	0.18858330
RC	15255981094604	76134.98670	114	0.11738	0.51932580

2. Association Models

We consider six of the most commonly used association models. These are:

a. The model of Independence or the null association models which holds that there is no relationship between the variables which are also symbolized by (O). The log-linear model is: Log (F_{ij}) = $\lambda+\lambda_{A(i)}+\lambda_{B(j)}$, where log denotes the natural logarithm, F_{ij} the expected frequencies under the independence model, $\lambda_{A(i)}$ are the rows main effect and $\lambda_{B(j)}$ are the columns main effect (Eliason & Clogg 1990).

b. The Uniform association model, which is symbolized as (U) in log-linear form is: $\log(F_{ij})=\lambda+\lambda_{A(i)}+ \lambda_{B(j)}+\varphi\chi_i y_j$, where φ ιs a single parameter for interaction and χ_i ,y_j are the scores for the row and column variables (i =1,…,I, j =1,…,J) respectively. (Diewert, 1976).

c. The row effects model (R) where linear-by-linear interaction holds give log (F_{ij}) = $\lambda+\lambda_{A(i)}+\lambda_{B(j)}+\varphi\mu_I y_j$ (Goodman, 1979), where y_j are fixed scores for the column variable (j = 1,…, J) and μ_I are unknown scores for the row variable (i =1,…,I).

d. The column effects model (C) is the same as the R model with a change in subscripts: Log(F_{ij}) = $\lambda+\lambda_{A(i)}+\lambda_{B(j)}+\varphi\nu_j x_I$, where x_I are fixed scores for the row variable (i =1,…,I) and ν_j are unknown scores for the column variable (j =1,…,J), (Diewert, 1995).

e. The model that allows both row and column effects in additive form is called the R+C model, (Goodman, 1979). The log-frequency version of the above model is: log(F_{ij}) = $\lambda+\lambda_{A(i)}+$

$\lambda_{B(j)} + \sum_{\kappa=1}^{I-1} \beta_\kappa y_j Z_{A(\kappa)} + \sum_{\kappa=1}^{J-1} \gamma_\kappa \chi_i Z_{B(\kappa)}$, where χ_i, y_j are the scores as defined earlier, and $Z_{A(i)}$, $Z_{B(j)}$ denote to variables (dummy variables) for the row and column levels, respectively.

f. The model, instead of additive row and column effects on the local odds ratios has multiplicative effects called the RC model or model II, (Goodman, 1981b). The log-multiplicative model is: $\log(F_{ij}) = \lambda + \lambda_{A(i)} + \lambda_{B(j)} + \varphi\mu_I\nu_j$, where the row score parameters μ_I and column score parameters ν_j are not known, but those estimated from the data.

We aim at finding out the model (out of the six) that best fit from the other models which we are examining, i.e., the total intramural research and development (R&D) expenditures within the 21 countries of the EU for the period of 2000-2010. For this reason, first we are going to examine the Index of Dissimilarity (L2), which shows that, the lesser the number, the more our model will give the best fit to match the total intramural research and development (R&D) expenditures within the 21 countries of the European Union (EU) for the period of 2000-2010 in each country compared with other models under consideration.

We analyze the six association models describe above, with the help of the CDAS Statiatical Programme [1]. We used the Pearson chi-squared (X^2) statistic, the likelihood-ratio chi-square(G^2) statistic, the index of dissimilarity

$$D = \sum_{ij} \left| f_{ij}/n - F_{ij}/n \right| / 2 \quad \text{where:}$$

f_{ij} the observed frequencies and

F_{ij} the expected frequencies (under the model)

Additionally, we have the following results:

Table 3. Index of Dissimilarity

Model	Index Of Dissimilarity(D)
1, Null Association-Independence Model (O)	0.32982
2. Uniform Association Model (U)	0.33127
3. Row-Effects Association Model (R)	0.21894
4. Column-Effects Association Model (C)	0.26583
5. Row+Column Effects Association Model (R+C)	0.16981
6. Row Column Effects Association Model (RC)	0.11738

At first sight it seems that the Row Column Effects of the Association Model (RC) adjusted better to the percentage of unemployment in the 27 countries of the EU for the years under study, as it is the one that has the lowest index of dissimilarity with D = 0 .11738.

Since we have models with similar lower ratio, we justify the model which gives the bestfit to match both countries and years by calculating the Index BIC (Bayes information criterion) is being used to give the best solution. The formula for this calculation is:

BIC = G^2 – (d.f.) log (n)

Symbols: n = the size of the sample (351074.4000)

Log (n) = Log (351074.4000) = 12.76875

d.f. = degrees of freedom of the models

G^2 = the likelihood-ratio chi-square statistics

When comparing a number of models, then the model with the smallest value of BIC is the best. So we choose the models, those whose INDEX OF DISSIMILARITY are similar and the lowest out of the six models. More precisely, the 3rd, 5th and 6th models.

3^o Model: BIC = G^2 – (d.f.) Log (n) =157974.15344 – (120 *12.76875) = 156441.75

5^o Model: BIC = G^2 – (d.f.) Log(n) = 141006.84843 – (114 * 12.76875) = 139551.21093

6^o Model: BIC = G^2 – (d.f.) Log(n) = 76134.98670 – (114 * 12.76875) = 74679.3492

As we can see from the above calculations, the best model still remains the 6th - the Row Column - effects of the Association Model (RC)

2.1. Analysis of the association models

In continuation, we check the acceptability or satisfactory fit of each model. We used the Pearson chi-squared (X^2) statistic, the likelihood-ratio chi-square (G^2) statistic and the index of dissimilarity. In the case of the X^2 distribution, the Statgraph programme will be of good help.

Initially, we observe that the likelihood-ratio chi-square statistic for the Independence model (O) is G^2 = 283368.44412 with 140 d.f. (the 95th percentile of the reference chi-square distribution is 168.853). It has unacceptable fit and therefore, we reject it.

The Uniform association model is G^2 = 272362.14566 with 139 d.f. (the 95th percentile of the reference chi-square distribution is 167.75). As it could be noticed both statistics are not accepted and does not have a satisfactory fit.

The statistic G^2 for the Row model (R) is reduced dramatically for 157974.15344 with 120 d.f. (The 95th percentile of the reference chi-square distribution is 146.82). The R model likewise is not accepted.

The Column model (C) has G^2 = 236482.08352 with 133 d.f. (the 95th percentile of the reference chi-square distribution is 161.125) shows even the worst fit.

The statistic of the model R+C, that takes into account the effects for both Countries and Years in additive form, is G^2 = 141006.84843 with114 d.f. (the 95th percentile of the chi-square distribution is 140.143) has equally unacceptable fit.

Finally, the model RC, that is log multiplicative but not log-linear, the G^2 Statistics is 76134.98670 with 114 d.f.(the 95th percentile of the reference chi-square distribution is 140.143, the same as the previous model because they have the same d.f, but the model of multiplicative association effects for both Countries and Years are seen to remain unacceptable. However, we have to realise and in which degree or level of effects it has on each model. In order to verify this we will have to construct the Analysis of association (ANOAS) table.

3. Creation of an Association Table (ANOAS)

The ANOAS table was given by Goodman (1981a). In this table the chi-squared are partitioned as sums of square in a two-factor analysis of variance using the likelihood. The ANOAS table partitions the X^2 so that it could be used as 2 factors analysis of variance using the percent of the G^2 (O) statistic for the basic (zero) independence model which measures the total deviation of the variables. In other words, we can find the percentage of baseline chi-squared X^2 distribution, which have effects on each of our models on the phenomenon being studied.

Table 4.

Models	Likelihood-G^2	Degrees of Freedom	Index of Dissimilarity
O	283368.44412	140	0.32982
U	272362.14566	139	0.33127
R	157974.15344	120	0.21894
C	236482.08352	133	0.26583
R+C	141006.84843	114	0.16981
RC	76134.98670	114	0.11738

The analysis of association table has the following differences of our models: O-U is the total effects model, U-R are the row effects model, R-RC are the column effects model that gives the effect of row, while RC are the residuals of the models.

Table 5. ANOAS

Effects	Models used	G^2	D.F	Percentage
1. General	O-U	11006.2985	1	3.88%
2. Rows	U-R	114387.9922	19	40.37%
3. Columns	R-RC	81839.16674	6	28.89 %
4. Residual	RC	76134.98670	114	26.86%
Total	**O**	**283371.24813**	**140**	**≈ 100.00%**

As shown from the above ANOAS table of the general model, the uniform effects are very weak because the U model accounts for almost 4% of the baseline chi-squared value. The columns effects model covers a reasonable percentage of over 40% of the baseline chi-squared X^2 distribution value. The row effects model are very weak because the R model accounts for roughly 29% of the baseline chi-squared value. Finally, the residual model RC covers only 27%.

We see therefore that on the rate of 27%, the variation which is attributed to the non-independence has been measured from the model of RC. This rate is quite satisfactory and we can say that the correlation on the total intramural R&D expenditure (GERD) by sectors of performance in 21 countries (Belgium, Czech Republic, Denmark, Germany, Estonia, Ireland, Greece, Spain, France, Italy, Cyprus, Latvia, Lithuania, Hungary, Netherlands, Poland, Portugal, Slovenia, Slovakia, Finland, and United Kingdom,) as seen from the data taken from the Eurostat, depended slightly positive on both the countries and of the years (2001-2010).

Again, we can also say that percentage (283368.44 –76134.986) / 283368.44 = 73% of the data explained by the row-column effects model (on the local odds ratios in a multiplicative way), giving it a satisfactory fit.

Moreover, because the best model (RC) under our study show bad fit (rejected), we supposedly proceeded to examining the multivariate models in order to find a solution to the problemn.

4. The multivariate models

In the RC (M) association model, M represents the dimension fit to be, which is utilized by the PROGRAM RCDIM. As shown below the multivariate model RC (M=7) is the acceptable model with the best fit (Haritou & Nwaubani, 2009).

The results are as follows:

Table 6. Multivariate model

Model	**RC(3)**	**RC(5)**	**RC(6)**	**RC(7)**
X^2	46000131943732.02	74864.38177	604.22194	0.16508
G^2	11745.01935	974.70482	293.27414	-4.92255
d.f.	68	30	14	0
D	0.02923	0.00444	0.00116	0.00001

Model RC (3) multivariate row, column, M=3

Model RC (5) multivariate row, column, M=5

Model RC (6) multivariate row, column, M=6

Model RC (7) multivariate row, column, M=7

.

5. Analysis of the multivariate models

The multivariate row column model RC(7) with M=7, has likelihood-ratio chi-square statistic G^2 -4.92255 with 0 degrees of freedom (d.f.). Moreover, the 95th percentile of the reference chi-square distribution is 0.0000 . We also observe that model M=7 covers {(283371.24813– 0.16508) /283371.24813} = 100% of the total data.

Because the model with the smaller M, if it is satisfactory gives the better explanation of effects between the rows and columns, we will prefer model M=7 that has a perfect fit.

6. Estimation of the multivariate model

The expected frequencies under the independent and row columns effects models RC(M=7) for the total number for the total intramural R&D expenditure (GERD) by sectors of performance in 21 countries of the EU of are given below: *Note:* The row columns effects model (on the multivariate model) seems to give better fit, particularly at the end of nominal scale.

Years	Countries	Data	Prices of the Model (0)	Prices of the Model RC(M=7)
1	1	3014.0000	4282.7070	3014.0278
2	1	409.4000	1093.1649	409.3995
3	1	2052.7000	3990.9463	2052.7231
4	1	30129.6000	7258.3310	30129.4755
5	1	5.6000	37.7629	5.6000
6	1	699.2000	355.7020	699.2016
7	1	220.8000	155.4464	220.8143
8	1	2445.4000	3985.6915	2445.3872
9	1	17519.5000	19215.3817	17519.4664

10	1	5511.9000	1327.9587	5511.9708
11	1	2.6000	11.1306	2.6009
12	1	5.1000	25.0558	5.0986
13	1	0.0000	34.8250	0.0156
14	1	109.7000	376.7211	109.7005
15	1	3693.9000	890.1625	3693.9618
16	1	424.1000	522.6612	424.1463
17	1	156.2000	379.7785	156.2225
18	1	134.4000	376.1240	134.3999
19	1	102.4000	133.7107	102.4086
20	1	2238.9000	4480.9800	2238.9189
21	1	14980.5000	34921.6581	14980.4963

Immediately, it is seen how much it fit better the values of the model RC (M=7) to the data.

7. Logarithm of the row effects model (ROW TAU-S)

The logarithm of the row columns effects (M=7) from the Row TAU-S of ther RCDIM are as follows:

Belgium: $\hat{\tau}_1$ = Log(1.300138) = 0.2625 Czech Republic: $\hat{\tau}_2$ = Log(2.023631) = 0.7049

Denmark: $\hat{\tau}_3$ = Log(1.946169) = 0.6659 Germany: $\hat{\tau}_4$ = Log(0.156678) = -1.8536

Estonia: $\hat{\tau}_5$ = -Log(1.526799)= -0.4232 Ireland: $\hat{\tau}_6$ = -Log(2.597227) = -0.9544

Greece: $\hat{\tau}_7$ = -Log(2.987613) = -1.0945 Spain: $\hat{\tau}_8$ = Log(1.317316) = 0.2756

France: $\hat{\tau}_9$ = Log (2.184811) = 0.7815 Italy: $\hat{\tau}_{10}$ = -Log(1.536173) = -0.4293

Cyprus: $\hat{\tau}_{11}$ = -Log(3.423924) = -1.2308 Latvia: $\hat{\tau}_{12}$ = -Log (2.611007) = -0.9597

Lithuania: $\hat{\tau}_{13}$ = - Log(2.358227) = -0.8579 Hungary: $\hat{\tau}_{14}$ = Log(0.221687) = - 1.5065

Netherlands: $\hat{\tau}_{15}$ = -Log(1.910374) = -0.6473 Poland: $\hat{\tau}_{16}$ = - Log (0.90876) = 0.0957

Portugal: $\hat{\tau}_{17}$ = -Log(1.124888) = -0.1177 Slovenia: $\hat{\tau}_{18}$ = Log(0.973098) = -0.0273

Slovakia: $\hat{\tau}_{19}$ = -Log(1.408758) = 0.3427 Finland : $\hat{\tau}_{20}$ = Log(2.050413) = 0.7180

United Kingdom: $\hat{\tau}_{21}$ = Log(5.521336) = -2.25560

8. Comparison of the the total intramural R&D expenditure by sectors of performance in some EU countries

For example, if we want to compare Greece with Cyprus we observe that $\hat{\tau}_7 - \hat{\tau}_{11}$ = 0.1363, exp(0.1363) =1.1460 . This means that Cyprus has 1.1460 intramural research and development expenditure graeter than Greece.

In the case of the Mediterranean countries, Italy and Spain, we find out that $\hat{\tau}_{10} - \hat{\tau}_8$ = -0.7049, exp - (0.7049) = 0.4941. it shows that Italy has 0.4941 lesser than Spain on the total intramural research and development expenditure.

Moving towards the north we observe that the difference of the total intramural expenditure on research and development performance between Poland and Slovenia $\hat{\tau}_{16} - \hat{\tau}_{18}$ = 0.123, exp(0.123) = 1.1309, this means that Slovenia has 1.1309 greater expenditure in research and development performance than Poland.

Even among the advanced European countries there are variations in the number. More specifically between Germany and France $\hat{\tau}_4 - \hat{\tau}_9$ = -2.6351,exp -(2.6351) = 0.0717, here we observe that Germany has 0.0717 lesser than France on the total intramural research and development expenditure.

Additionally, the difference between Germany the United Kingdom $\hat{\tau}_4 - \hat{\tau}_{21}$ = -3.5622, exp -(3.5622) = 0.0284 which shows that Germany has 0.0284 fewer intramural research and development expenditure than the United Kingdom.

Finally, comparing the total spending in business performance between the Scandinavian countries of Finland and Denmark, we see that $\hat{\tau}_{20} - \hat{\tau}_3$ = 0.0521, exp(0.0521) = 1.0535, it means that Finland has 1.0535 greater than Denmark on the total intramural research and development expenditure.

Summary

Generally, we could see that the multivariate model RC (M=7) gives the best fit among all. However, to be more precise, the total intramural expenditure on research and development in the21 EU countries iwas influenced by several factors. These could be as a result of:

- government spending on research and development
- the standard of living of each country
- various other factors which are difficult to be identified or determined.

Moreover, we can easily see from the above data showing the total number of trips by main mode of transport used in 13 EU that over the years the rate of tourism increased due to improved economic conditions

In order to realise the degree of association which exists between the total intramural expenditure on research and development unemployment in the21 EU countries in these subsequent years, we use θ (Theta) in order to calculate the indicator of innate association – i.e. **φ** (phi)**.**

$$\Theta = 1.00963$$

The parameter of interaction is $\varphi = \mathrm{Log}\theta = \mathrm{Log}\,(1.00963) = 0.009584$

And $\boldsymbol{\varphi}\ ½ = \sqrt{}0.009584 = 0.097897906$ and $\varphi^{1/2} = \sqrt{}0.009584 = 0.097897906$

Therefore, we conclude that there is slightly a positive association between the total intramural expenditure on research and development in the21 EU countries and the years.

9. Conclusions

Statistical findings

Gross domestic expenditure on R&D (GERD) stood at EUR 245 673 million in the EU-21 in 2010 at current prices, (3.8 % increase in the previous year), and 43.5 % higher in the last ten years. In 2008 the level of expenditure on R&D in the EU-21 was 88.5 %, although slightly more than double the level of expenditure in Japan and China.

In order to make figures more comparable, General Expenditure on Research and Development (GERD) is often expressed relative to GDP or in relation to population. The ratio of GERD to GDP, one of five key Europe 2020 strategy indicators, increased marginally in the EU-27 during the period up to 2002 reaching a high of 1.88 %, before declining modestly through to 2005 (1.83 %), and climbing again to 2.01 % by 2009. There was a small decline in 2010 when the ratio fell to 2.00 %. The decrease despite the higher absolute level of R & D expenditure was due to the partial recovery from the financial and economic crisis, as GDP increased at a slightly faster pace than GERD in 2010.

The differences in the relative importance of R&D expenditure between countries are often explained by referring to levels of expenditure within the business enterprise sector. An evaluation of the data for the EU Member States shows that those countries with relatively high shares of business enterprise expenditure on R&D namely: Finland, Sweden, Denmark, Austria and Germany also reported relatively high levels of total GERD. Apart from Germany, these countries also tended to feature near the top-end of the ranking of expenditure by the higher education sector, where the Netherlands also had a relatively high share of R&D expenditure. Government R&D expenditure relative to GDP was highest in Germany, Slovenia, France and Finland.

Analytically, R & D expenditure by source of funds shows that more than half (54.1 %) of the total expenditure in 2009 within the EU-21 was funded by business enterprises, while just over one third (34.9 %) was funded by government, and a further 8.4 % from abroad (foreign funds). In 2008, business funded R&D accounted for 78.2 % of total R&D expenditure in Japan and 67.3 % in the United States (both). By contrast, a majority of the gross expenditure on R&D made in Cyprus, Bulgaria, Poland, Romania, Slovakia and Lithuania in 2010 was funded by the government sector. There were also remarkable differences in the source of R&D funding from abroad, with relatively high shares in excess of 15 % of total GERD in 2009 within Latvia, Lithuania, Malta, Austria, the United Kingdom and Ireland.

References

Clogg, C.C. (1990). *ANOAS program.*

Diewert, W. E., (1976). Exact and Superlative Index Numbers, *Journal of Econometrics 4,* 115–145

Diewert, W. E., (1995). Axiomatic and Economic Approaches to Elementary Price indexes. *NBER working paper 5104*.

Eliason, S. R., & Clogg, C.C. (1990). *Ανάλυση Διατεταγμένων Δεδομένων (CDAS).*

Fairbanks, M.. (2005). Changing the Mind of a Nation: Elements in a Process for Creating Prosperity, In *Culture Matters*, Huntington, editors, *New York: Basic Books*, pp. 270–281.

Goodman, L.A., (1979). Multiplicative Models for the Analysis of Occupational Mobility Tables and Other Kinds of Cross-Classification Tables. *American Journal of Sociology*, 84, 804–819.

Goodman, L.A., (1981a). Association models and the Bivariate Normal for Contingency Tables with Ordered Categories. *Biometrika*, *68,* 347–355

Goodman, L.A., (1981b). Association Models and Canonical Correlation in the Analysis of Cross-Classifications Having Ordered Categories. *Journal of American Statistical Association,* 76(3), 20–34.

Eurostat /JP (2010). *European Data Agency, Eurostat-Online data code (t2020_20), OECD statistics at regional level*.

Haritou, A., & Nwaubani, J C. (2006). *Categorical Data Analysis* (*University Press)*.

Ένα Μοντέλο για το σύνολο των εγχώριων δαπανών R&D (GERD) ανά τομέα εκτέλεσης σε 21 χώρες της Ευρωπαϊκής Ένωσης

Joel C. Nwaubani, Nikos Kapoulas, Georgia Lepida

Τμήμα Εφαρμοσμένης Πληροφορικής,
Πανεπιστήμιο Μακεδονίας

Περίληψη

Στην έρευνα αυτή χρησιμοποιούμε όλα τα γνωστά μοντέλα συσχέτισης της Ανάλυσης Διατεταγμένων Δεδομένων και εξετάζουμε το σύνολο των εγχώριων δαπανών R&D (GERD) ανά τομέα εκτέλεσης σε 21 χώρες της Ευρωπαϊκής Ένωσης. Τα δεδομένα που χρησιμοποιήθηκαν στην παρούσα έρευνα σε ετήσια βάση προέρχονται από στατιστικές έρευνες του Eurostat που διεξάγονται τακτικά σε εθνικό επίπεδο και αφορούν τα έξοδα στην έρευνα και ανάπτυξη στον ιδιωτικό και τον δημόσιο τομέα, υπολογισμένα ανά έτος για τα έτη 2000 – 2010. Δεδομένου ότι ο κύριος στόχος είναι να έχουμε μια καλύτερη κατανόηση του συνόλου των εγχώριων δαπανών στην έρευνα και ανάπτυξη (R&D) εντός των 21 χωρών της Ευρωπαϊκής Ένωσης (ΕΕ), ανάλογα με τον τομέα εκτέλεσης και την προέλευση των κεφαλαίων, η ανάλυση του πίνακας συσχέτισης (ANOAS) δίνεται προκειμένου να εξακριβωθεί το ποσοστό των δεδομένων που καλύπτονται από κάθε μοντέλο. κάνουμε την εκτίμησή και τέλος διαπιστώνουμε από τα συμπεράσματα μας ότι κανένα από τα έξι μοντέλα δεν προσαρμόζεται. Προχωρήσαμε με το πολυμεταβλητό μοντέλο για να βρούμε το μοντέλο με την καλύτερη προσαρμογή και ως εκ τούτου το πολυμεταβλητό μοντέλο επίδρασης των στηλών RC (M = 7) έχει την καλύτερη προσαρμογή μεταξύ όλων.

Λέξεις κλειδιά: Μοντέλα Συσχέτισης, Λογαριθμογραμμικά και μη Λογαριθμογραμμικά Μοντέλα, ΑΕΠ, έρευνα και ανάπτυξη, Ακαθάριστες εγχώριες δαπάνες (σε ποσοστό % του ΑΕΠ), 21 χώρες της ΕΕ.