



Munich Personal RePEc Archive

# **The Application of the Econometric Models with Qualitative Variables in the Analysis of the Non Academic Behaviors at the Level of the Romanian Higher Education System**

Andrei, Tudorel and Teodorescu, Daniel and Iacob, Andreea Iluzia E. S. and Stancu, Stelian

Academy of Economic Studies Bucharest, Institutional Planning Research, Emory University, Atlanta, USA

December 2007

Online at <https://mpra.ub.uni-muenchen.de/6889/>  
MPRA Paper No. 6889, posted 26 Jan 2008 05:45 UTC

**Prof. Tudorel ANDREI, PhD**  
**Academy of Economic Studies, Bucharest**  
**Daniel TEODORESCU, PhD**  
**Institutional Planning & Research, Emory University, Atlanta, SUA**  
**Associate Prof. Andreea Iluzia IACOB, PhD**  
**Academy of Economic Studies, Bucharest**  
**Prof. Stelian STANCU, PhD**  
**Academy of Economic Studies, Bucharest**

## **THE APPLICATION OF THE ECONOMETRIC MODELS WITH QUALITATIVE VARIABLES IN THE ANALYSIS OF THE NON ACADEMIC BEHAVIORS AT THE LEVEL OF THE ROMANIAN HIGHER EDUCATION SYSTEM<sup>1</sup>**

*The purpose of this paper was to apply the econometric models with qualitative variables in order to analyze two non academic behaviors at the level of the Romanian higher education system: cheating on the exams by copying or by direct or intermediary intervention at the professor.*

**Key-words:** *econometrics, Logit Model, survey, sample, fraud.*

### **1. INTRODUCTION**

Romania, country of Eastern Europe, after the decision adopted at Helsinki in December 1999, began the negotiations to adhere at the European Union. In December 2004 it finished the negotiations to adhere at the European Union. Since January first, 2007 it became member of the European Union.

The transition period generated a lot of important changes at the tertiary education system level. The process of reforms in the Romanian higher education system was accelerated after the signature of the treaty of Bologna. During the last 16 years the number of the enrolled students registered a spectacular growth. If during 1990/1991 academic year were enrolled 192810 students, all in the tertiary education public institutions, during 2005/2006 academic year were enrolled

---

<sup>1</sup> Research for this paper was supported by the excellence research project no. CEEX 05-D8-65/10.11.2005

716464 students from which 513678 students followed the courses of the higher education public institutions.

During this period, the ratio between the number of the pupils and the enrolled students decreased. Thus, if for 1990/1991 academic year the ratio was equal with 5,16, for 2005/2006 academic year the ratio was equal with 4,18. During this period the number of the students per 10000 inhabitants increased almost four times.

In order to determine some characteristics of the higher education, in the University Center of Bucharest, during November 2005, was organized a research based on a statistical survey

The objective of the research was the analysis of the academic behavior of the students and the professors, reported at the characteristics of the performances of the education process.

The sample was stratified by field and year of study and included 1025 students enrolled in day study programs at major state universities from Bucharest. The first-year students were not included in the study because of their limited experience in the field. The results were guaranteed with a 2% error.

The structure of the students by year of study was as follows: second year - 38.6%; third year -33.0%; fourth and fifth years - 28.4%. The distribution of the students by field of study was as follows: arts and sciences - 32.3%, technology - 27.6%, economy - 22.2%, medicine - 6.7%, law - 5.5%, agriculture - 4.0% and physical training and sports - 1.8%.

The obtained data at the sample level were used to analyze two non academic behaviors at higher education level: the fraud at the exam by copying or by the direct intervention or by intermediary at the professor.

## **2. THE APPLICATION OF THE LOGIT MODEL IN ORDER TO ANALYZE THE FRAUD AT THE EXAMS BY COPYING THE EXAM**

In order to define the Logit Model it is considered that the exam fraud by copying the exam ( $y$ ) is dependant on the following characteristics: the sex of the person ( $x_1$ ), the level of the corruption in the university determined by the behavior of the professors ( $x_2$ ), the level of professional preparation of a student which is defined by the time spent for study, on average, during the week ( $x_3$ ) and the appreciation of the students concerning the level of preparation of the colleagues ( $x_4$ ), the quality of the education activity at the university level, appreciated by the relevance of the courses ( $x_5$ ) and by the presence of the students at the courses ( $x_6$ ), the manner of spending the spare time outside the university campus taking account of the time spent working off-campus during a week ( $x_7$ ), and the

extraprofessional activities ( $x_8$ ), the tendency to cheat on exams taking account of the practice to cheat on exams during college ( $x_9$ ).

1. **The dependent variable of the logit model** expresses the student exam fraud by copying the exam from a colleague or by using other sources, excluding the situation in which one intervenes directly or by intermediary at the professor. Information was recorded with the answers given at the question: *“Suppose that, for the moment, there is no chance to pass on a certain exam. In this situation would you try to copy the exam from a colleague or by using other hidden sources (cribs, manual etc.)?”*. The answer is defined as a binary variable. Thus, one allots value 1 if the answer of the student was “Yes” and 0 for the answer “No” or “I don’t know”.

2. **The independent variables of the logit model** are:

A. The sex of the person

B. The level of the corruption in the university is determined by the behavior of the professors. The collection of information was carried out analyzing the answers at the following questions: *“Did a professor asked for or accepted money from the students in exchange of a passing grade on exams?”*, *“Did a professor asked for or accepted money from the students to guarantee the promotion at the University admission exam?”* and *“Do you know a professor who asked the students to buy a book or a handbook published by him?”*. Using these three sets of data was defined a new variable, calculated as an average of the values of these three characteristics granted by each student. Using this new variable can be measured this element of the corruption in the university educational system perceived by the students as being determined by the professors.

C. The level of professional preparation of the students, which is measured using the following variables: the time spent for individual study during the week and the appreciation of the students concerning the level of preparation of the colleagues.

D. The quality of the teaching activities in the universities is measured using the following variables: the relevance of the courses for the future profession; the presence at the courses and at the seminars, which is defined as an average of the answers of the students at the following questions: *“During 2004/2005 academic year, how often did you participate at the courses?”* and *“During 2004/2005 academic year, how often did you participate at the seminars/laboratories?”*.

E. The manner of spending time off-campus was measured using the following characteristics: the time spent working off-campus and the manner of spending the spare time which is defined as an average of three variables which characterize the time spent with the friends or the colleagues, the time spent for Internet surfing and the time spent to play video games or in another manner.

F. The tendency to fraud the exam, consequence of the behavior of the current students during college. This variable is defined by the answers at the

question: “On your opinion, which is the proportion of the colleagues who used illegal methods during college to obtain higher grades?”.

Using these variables will be defined the following linear probability model:

$$Y_i = a_0 + a_1x_1 + a_2x_2 + a_3x_3 + a_4x_4 + a_5x_5 + a_6x_6 + a_7x_7 + a_8x_8 + a_9x_9 + \varepsilon_i \quad (1)$$

Where  $\varepsilon_i$  represents the residual variable which quantifies the influence of other factors which can determine the intention of a student to cheat on exams. This variable is normally distributed with zero mean ( $N(0, \sigma^2)$ ).

Using the registered values for the past variables will be estimated the parameters of the linear probability model and will result the following model:

$$\hat{Y}_i = a_0 + a_1x_1 + a_2x_2 + a_3x_3 + a_4x_4 + a_5x_5 + a_6x_6 + a_7x_7 + a_8x_8 + a_9x_9 \quad (2)$$

Where  $\hat{Y}_i$  is estimating the probability that  $Y_i = 1$ . By estimating this model we'll determine:

1. The sign and the estimated value of each parameter, specifying thus if one factor has a positive or negative contribution on exam fraud. Each coefficient shows with how much is modified the probability that  $Y_i = 0$ , if an independent variable is modified by a unit, while the others are maintained constant.
2. The probability that a particular student' profile can or can't determine the exam fraud. For example, if for a particular combination of factors we'll obtain that  $\hat{Y}_i = 0.4$ , then this one is estimating the probability that  $Y_i = 1$ . While the observed values of the variable  $Y_i$  are 0 or 1, the estimated values  $\hat{Y}_i$  are included between the two extreme values.

In the case of the Logit Model, the distribution function which defines the probability  $p_i = P(y_i = 1)$  is the logistic distribution function and is defined:

$$\begin{aligned} p_i &= P(y_i = 1) = P(v_i^* > v_i) = P(\varepsilon_i > v_i - (a_0 + a_1x_{1i} + \dots + a_px_{pi})) \\ &= P(\varepsilon_i < a_0 + a_1x_{1i} + \dots + a_px_{pi} - v_i) = F(\varepsilon_i < a_0 + a_1x_{1i} + \dots + a_px_{pi} - v_i) \quad (3) \\ &= \frac{1}{1 + e^{-[a_0 + a_1x_{1i} + \dots + a_px_{pi} - v_i]}} \end{aligned}$$

The distribution function can be also written in the equivalent form:

$$p_i = P(y_i = 1) = \frac{e^{[a_0 + a_1 x_{1i} + \dots + a_p x_{pi} - v_i]}}{1 + e^{[a_0 + a_1 x_{1i} + \dots + a_p x_{pi} - v_i]}} \quad (4)$$

The proprieties of this function are:

1. In the conditions in which  $a_0 + a_1 x_{1i} + \dots + a_p x_{pi} - v_i > 0$ , the probability  $p_i \in [0, 1]$ . The maximum value of  $p_i$  is 1 and it's obtained for the situation in which is verified the condition  $a_0 + a_1 x_{1i} + \dots + a_p x_{pi} - v_i \rightarrow \infty$ . The value of  $p_i$  is equal with zero if the following condition is verified:  $a_0 + a_1 x_{1i} + \dots + a_p x_{pi} - v_i \rightarrow -\infty$ . Using this function are eliminated the situations in which the estimated values  $\hat{Y}_i = P(Y_i = 0)$  are not included in the interval  $[0, 1]$ . The logistic function is estimating the linear probability model.
2. The distribution function is a symmetric one. From the relation (3) it results that:  $F(-x) = 1 - F(x) = F(x)$ .
3. A transformation of the model defined by the relation (3) allows us to write it in the same manner as the linear probability model defined in the relation (1):

$$\ln\left(\frac{p_i}{1 - p_i}\right) = a_0 + a_1 x_{1i} + \dots + a_p x_{pi} + \varepsilon_i. \quad (5)$$

In these conditions is obtained the following equality:

$$P(Y_i = 1) = \ln\left(\frac{p_i}{1 - p_i}\right) \quad (6)$$

The estimated results of the Logit Model and its characteristics are presented in the following table:

*The estimated results of the Logit Model and its characteristics in order to analyze the fraud at the exams by copying the exam*

*Table 1*

Variable	B	S.E.	Wald	Sig.	Exp (B)
A. The sex of the person ( $x_1$ )	- 0,0261	0,191	1,867	0,172	0,770
B. The level of the corruption in the university educational system ( $x_2$ )	0,203	0,089	5,237	0,022	1,225
C. The level of preparation of the students					
C1. The individual study ( $x_3$ )	- 0,238	0,067	12,630	0,000	0,788
C2. The evaluation of the students concerning the level of preparation of the colleagues ( $x_4$ )	0,183	0,110	2,759	0,097	1,201
D. The quality of the education activities in the universities					
D1. The relevance of the courses ( $x_5$ )	- 0,224	0,087	6,605	0,010	0,799
D2. The presence at the courses ( $x_6$ )	- 0,543	0,164	10,934	0,001	0,581
E. The manner of spending the spare time					
E1. Time spent working-off campus ( $x_7$ )	- 0,133	0,044	9,067	0,003	0,876
E2. Extraprofessional activities ( $x_8$ )	0,268	0,070	14,631	0,000	1,307
F. The tendency to copy the exam during college ( $x_9$ )	0,268	0,094	3,434	0,064	1,190
Constant	0,488	0,749	0,424	0,515	1,628

In this case, the Logit Model is defined by:

$$P(\text{the fraud at the exam by copying}) = \frac{\exp(0,488 - 0,0261x_1 + 0,203x_2 - 0,238x_3 + 0,183x_4 - 0,224x_5 - 0,543x_6 - 0,133x_7 + 0,268x_8 + 0,268x_9)}{1 + \exp(0,488 - 0,0261x_1 + 0,203x_2 - 0,238x_3 + 0,183x_4 - 0,224x_5 - 0,543x_6 - 0,133x_7 + 0,268x_8 + 0,268x_9)}$$

The characteristics which measure the time spent for individual study during the week, the relevance of the courses and the presence of the students at the courses determine a decrease of the probability of cheating on exams. The students

which develop regularly an activity off-campus are less tempted to cheat on exams. On the contrary, extraprofessional activities like: the time spend with the friends, the time spent on Internet surfing or playing video games conduce the fraud at the exam. More the time spent by a student for these activities, larger will be its intention to cheat on exams. The absence of objectivity in the evaluation of the level of preparation of the colleagues determines a growth of the intention towards the fraud at the exams. The sex of the person has a reduced relevance in the decision to cheat on exams.

### 3. THE APPLICATION OF THE LOGIT MODEL IN THE ANALYSIS OF THE FRAUD AT EXAMS BY INTERVENTION AT THE PROFESSOR

The dependent variable is represented in this model by the intention of a student to cheat at the exam by direct intervention or by intermediary at the professor. For this variable one allots value 1 for the students who intervene at the professor in order to fraud the exam and 0 for the ones which do not intervene or do not know how to proceed.

The independent variables of this model are: the sex of the person, the level of the corruption in the higher education system, the number of reported frauds at the direction of the university and the number of punished frauds.

The estimated results of this Logit Model and its characteristics are presented in the following table:

*The estimation results of the Logit Model and its characteristics in order to analyze the fraud at the exam by intervention at the professor*

*Table 2*

Variable	B	S.E.	Wald	Sig.	Exp (B)
The sex of the person ( $x_1$ )	- 0,672	0,359	3,498	0,061	0,511
The level of the corruption in the higher education system ( $x_2$ )	0,298	0,162	3,384	0,066	1,348
Number of reported frauds at the direction of the university ( $x_{10}$ )	0,494	0,217	5,184	0,023	1,639
Number of punished frauds ( $x_{11}$ )	0,582	0,280	4,312	0,038	1,789
Constant	- 2,799	0,807	12,034	0,001	0,061

$$P(\text{the fraud at the exam by intervention at the professor}) = \frac{\exp(-2,799 - 0,672 x_1 + 0,298 x_2 + 0,494 x_{10} + 0,582 x_{11})}{1 + \exp(-2,799 - 0,672 x_1 + 0,298 x_2 + 0,494 x_{10} + 0,582 x_{11})}$$



In this case, this model marks out a different behavior between the boys and girls in the attempt to cheat on exams by direct intervention or by intermediary. The other three variables included in the model have a negative influence on the level of the moral integrity in the university. The rarely report of the cases of corruption at the direction of the university and also the absence of severe measurements to punish the professors who encourage the fraud at the exams represent factors which lead to the reduction of moral integrity.

### REFERENCES

- [1] **Andrei, T. (2004)** – *Statistică și Econometrie*, Ed. Economica, București.
- [2] **Bandura, A. (1986)** - *Social Foundations of Thought and Action: A Social Cognitive Theory*, Prentice-Hall, Englewood Cliffs, NJ.
- [3] **Gourieroux, C. (1984)** - *Econometrie des Variables Qualitatives*, Ed. Economica, Paris.
- [4] **Kramer, J. S. (1991)** – *The Logit Model for the Economists*, Edward Arnold Publishers, London.
- [5] **Maddala, S. (1983)** – *Limited Dependent and Qualitative Variables in Econometrics*, Cambridge University Press, New York.
- [6] **McCabe, D. L., Trevino, L. K. (1997)** - *Individual and Contextual Influences on Academic Dishonesty: A Multi-Campus Investigation*, *Research in Higher Education*, 38(3), 379-396.
- [7] **McCabe, D. L., Trevino, L. K. (1993)** - *Academic Dishonesty: Honour Codes and Other Contextual Influences*, *Journal of Higher Education*, 64, 520-538, 1993.
- [8] **Pulvers, K., Diekhoff, G. M. (1999)** - *The Relationship between Academic Dishonesty and College Classroom Environment*, *Research in Higher Education*, 40, 487-498, 1999.
- [9] **Whitley, B. E. Jr. (1998)** - *Factors Associated with Cheating among College Students: A Review*, *Research in Higher Education*, 39(3), 235-274.