# Econometric Models Used in the Analysis of the Informal Economy at the Regional Level 

Tudorel ANDREI ${ }^{1}$, Bogdan OANCEA ${ }^{2}$, Stelian STANCU ${ }^{3}$, Andreea Iluzia IACOB ${ }^{1}$<br>${ }^{1}$ Academy of Economic Studies, Bucharest, Department of Statistics and Econometrics,<br>${ }^{2}$ Artifex University of Bucharest,<br>${ }^{3}$ Academy of Economic Studies, Bucharest, Department of Economic Cybernetics andreitudorel@yahoo.com, oanceab@ie.ase.ro, aiacob@ase.ro, stelian_stancu@yahoo.com

The disintegration of planned economies and the first elements of the market economy have resulted in a new attitude from some companies to transfer their activities to the informal economy. Assessing the probability that a company decides to conduct an economic activity can be done based on the application of the logit and probit econometric models. A series of data for different latent variables should be known for their implementation. This paper presents a number of problems related to the application of these models at the level of economic regions. Binary models are recommended because of a number of variables such as: the size of the informal economy at the level of a company, the size of unpaid taxes and duties due to their activities in the informal economy, which are not known in official statistics. These variables can be substituted to some extent by the latent variables for which data series are known.
Keywords: informal economy, econometric models, logit, probit

1Introduction
The disintegration of planned economies and the first elements of the market economy have resulted in a new attitude from some companies to transfer their activities from official area to the unofficial area [6], [7]. Thus, by acceptance of the risks arising from this, some companies avoid payment of taxes or duties to the state or avoid a series of bureaucratic barriers. Under these conditions, the analysis of the informal economy in Romania has become an important subject for economic theory and practice. Several studies have shown that taking measures to reduce the financial frauds and the informal economy represent serious problems for improving the economic environment of an economy during transition. These measures will stimulate the attraction of foreign investors with a high potential to support the economic development of Romania over the medium and long term. In the economic theory there is a large variety of methods used to estimate the size and the dynamics of the informal economy. Among the most important ones we can mention: monetary approach to the informal economy; the method of implicit labor offer applied to our
country by the data obtained by AMIGO; the method of energy consumption etc. The results obtained by applying various methods are different, especially in the assessment of the size of the informal economy for the economies during transition [1]. For example, in Romania [3], this dimension, as a share of Gross Domestic Product is $20 \%$ if it assessed by the energy consumption and over $45 \%$ if the assessment is carried out using monetary method. The National Statistics Institute assesses the informal economy by the methodology of National Accounts, and its value is approximately 25$28 \%$ of the formal economy based on GDP.

## 2 Informal Economy in Europe

Informal economy in the European countries is very heterogeneous. Differences are significant if we consider its size, but also the causes that generate it. In this economic space we meet countries with the lowest level of the informal economy in the world (Switzerland 8\%, Austria 10.2\% etc.), but also countries that have a high level of informal economy (Moldavia 45.1\%).
Countries recently admitted in the European Union are heterogeneous regarding informal
economy. The table 1 shows the size of the informal economy as a share of the formal economy in several countries recently admitted into the European Union. The values in this graphic emphasize a high level of informal economy in Bulgaria and Romania, but also countries with a low level, such as Slovakia and the Czech Republic. For example, Slovakia has a level of informal economy by nearly two times lower than in Bulgaria. Moderate values of the informal economy share of the GDP are met in Hungary.

An important factor influencing the development of the informal economy in the European countries in the process of transition was the work on the black market [4] due to lack of segmentation of the economy and the continuous changing of the legislation. Economic changes in the Eastern Europe have generated throughout the transition period a consistent migration of people from the East to the West part of the continent. This process has contributed directly to the increase in the informal economy in Western countries [5].

Table 1. The share of the informal economy in GDP for Eastern European countries

| Country | Share (\%) |
| :---: | :---: |
| Bulgaria | 36.9 |
| Czech Republic | 19.1 |
| Hungary | 25.1 |
| Romania | 34.4 |
| Slovakia | 18.9 |
| Slovenia | 27.1 |
| Poland | 27.6 |

For example, in Spain, the informal economy has increased from $16.1 \%$ in 1989 to no less than $22.3 \%$ in 2003. The chart below shows the size of the informal economy for various
countries in Western Europe. The highest rates of informal economy, situated at the level of Hungary and Poland, is met in Greece for the year 2003. (figure 2).


Fig. 2. The share of the informal economy in GDP for some Western European countries

## 3 The method

Models for which the dependent variable takes only discrete values are called models of discrete choice. In relation to the number of values taken by the dependent variable there are two types of discrete choice models:

1. Binary models - the dependent variable is of binary type. It has two values, symbolized as a rule, with 0 and 1 and is used to indicate whether an event occurred or whether a statistical unit has or not a particular property. These models are used in various areas such as: financial frauds analysis, medicine, economics, sociology, psychology

These models try to estimate the probability $p_{i}$ that the binary variable is equal to 1 .
We present a summary of two examples of the use of binary models and the manner of defining the dependent variable in such a model.
A. In the area of the financial frauds or decision for a company or a person to operate in the area of the informal economy. In this case, the $Y$ variable characterizes the state of the company or person to conduct activities in the informal economy. The values of this variable are defined as:

$$
y_{i}=\left\{\begin{array}{l}
1, \text { the company / person operates in the formal economy } \\
0, \text { the company / person operates in the informal economy }
\end{array}\right.
$$

In this situation, state institutions may not observe all the people and economic agents operating in the informal economy. Observation cannot be a comprehensive one for various reasons: the operational costs of these institutions are enormous in this case; the ways in which companies elude the formal economy can be diversified according to the methods and ability to control of the state institutions etc.
For these reasons, we will apply econometric models to estimate the probability that a company may decide to operate in the

$$
y_{i}=\left\{\begin{array}{l}
1, \text { the customer i buys the product } \\
0, \text { the customer i does not buy the product }
\end{array}\right.
$$

When the product is sold the option of the customer to purchase the product is not known, but other information may be available that allow the estimation of the probability to buy, such as income, age, sex, habit of consuming a similar product, number of persons of a family etc.
From the binary models we can mention: Logit model, Probit model and Weibull model.
2. Models with multiple responses are those models for which the dependent variable has more than two discrete values. In this case, the dependent variable may be an ordinal or nominal variable.
informal area according to data recorded by different sets of data defined by the latent variables. These are observable and correlated with the $Y$ variable.
B. For a better understanding of the accepted methodology we the present another example from the field of marketing.
In the marketing domain, the dependent variable characterizes the attitude of a customer to buy or not to buy a product. We define the $y_{i}$ variable in this case by:

## 4 Binary models for the analysis of financial frauds

First we present a simple binary variable used for the analysis of financial fraud of a company at the level of an economic development region. In this case, a series of economic information may be available in an economic region.
For defining the binary model we consider the following variables:

1. The dependent variable of the model is denoted by $Y$ and it is used to determine if a company decides to pursue an economic activity in the informal economy. In this respect, the list of companies that committed economic illegalities during the last financial year is known. Data can be available from

Local Public Administrations.
2. The explanatory variables of the model are:
a. The number of employees at the beginning of the year (X1);
b. The number of employees at the end of the year (X2);
c. Annual Turnover (X3);
d. The number of controls that the company was subject to by the representatives of AFPL (Local Public Administrations) or Financial Guard (X4);
e. The weight of the export-oriented production (X5);
f. European funding attracted by the company (X6);
g. The amount of fines paid by the company during the year (X7);
h. The amount of VAT paid by the company (X8);
i. The amount of VAT refunded to the company (X9).
Using the above variables we define the linear probability model:

$$
Y_{i}=a_{0}+\sum_{i=1}^{9} a_{i} X_{i}+\varepsilon_{i},
$$

where $\varepsilon_{i}$ is a residual variable that measures the influence of other factors that may determine the company to make economic frauds. This is a variable with the property that follows a normal repartition with zero mean $\left[\mathrm{N}\left(0, \sigma^{2}\right)\right]$.
On the basis of the values recorded for the above mentioned variables we estimate the linear probability model's parameters:

$$
\hat{Y}_{i}=a_{0}+\sum_{i=1}^{9} \hat{a}_{i} X_{i}
$$

where $\widehat{Y}_{i}$ estimates the probability that $Y_{i}=1$. After the estimation we compute the following:

1. The sign and the actual value of each parameter, specifying whether such a factor has a positive or a negative contribution in fraud examination. Each coefficient indicates by what extent the probability that $Y_{i}=0$, changes if an independent variable is amended by a unit, while other variables are constant;
2. The probability that a company with a specific profile can make financial fraud by transferring certain activities in the informal area. For example, if a particular combination of factors gives $\hat{Y}_{i}=0,4$, then it estimates the probability for $Y_{i}=1$. While the observed values of the variable $Y_{i}$ are 0 or 1, the estimated values $\widehat{Y}_{i}$, based on regression model, are between those two extremes.

## 5 Problems and consequences of the binary specification

If the analysis of the feature $Y$ is achieved by a regression model based on the explanatory variables $X_{1}, X_{2}, \ldots, X_{p}$, then it is represented by the relationship given below:

$$
y_{i}=a_{0}+a_{1} x_{1 i}+\cdots+a_{p} x_{p i}+\varepsilon_{i} .
$$

The estimation of the parameters and their interpretation raises special problems in some particular situations. Two situations are eloquent in this regard:

1. The first case is when the dependent variable is likely different in relation to one or more explanatory variables. It is the case of the present model used to analyze the financial fraud;
2. The second situation is one in which the values of the explained variable are not directly observable, in which case we define the regression models with latent variables.

## Problems of the binary specification

We present some issues where the first dependent variable is likely different in relation to one or more explanatory variables. For example, if the variable $Y$ is a qualitative variable of binary type and one or more explanatory variables are quantitative, then the equality defined by the relationship (1) has no meaning. For a better understanding of this situation we consider a regression model with one explanatory quantitative variable and a binary dependent variable. We show in this example which is the consequence of changing the way the binary variable is defined on the estimators of the regression model.


Fig. 3. The data series and the series adjusted by the linear model

Thus, we consider the simple regression model used to analyze the dependence between variables $Y$ and $x$ :

$$
y_{i}=b+a x_{i}+\varepsilon_{i} \quad i=1, \cdots, n
$$

We denote in the above model with $a$ and $b$ the two parameters of the regression model and define the variables as follows:

1. The binary variable $Y$ indicates whether a person/company carries an illegal activity. This is defined by:
$y_{i}=\left\{\begin{array}{l}1, \text { the company operates in the formal economy } \\ 0, \text { the company operates in the informal economy }\end{array}\right.$
2. The value of the feature $X$ is the income of the individual $i$ in Euro made over a period of time;
3. $\varepsilon_{i}$ is the specification error of the model.

We present data series that are represented by the graphic in Figure 3 to show that this linear model, whose parameters are estimated by the least squares method, is not appropriate under the circumstances in which the two variables are likely different. As one can see in Figure 3, the data series $\left(x_{i}, y_{i}\right)_{i=\overline{1,50}}$ , with $y_{i} \in\{0,1\}$ is represented by two parallel lines, one for $y_{i}=0$ and the other one for $y_{i}=1$. Under these circumstances, it is virtually impossible to find a line to be estimated on the basis of the points of coordinates $\left(x_{i}, y_{i}\right)_{i=1,50}$. For the data series
that is represented in Figure 3 we estimated the parameters by the least squares method and we obtained:

$$
\hat{y}_{i}=\underset{(0.1509)}{0.6112}-\underset{(0.051)}{0.0015} x_{i}
$$

The value of $t$-Student statistics show that the slope of the regression line does not differ significantly from zero. Moreover, if we change the encoding for the binary variable, then the two parameters change too. For example, if value 1 is replaced with value 10 , then the intercept becomes 10b. This result highlights the fact that the values of the two parameters cannot be interpreted in economic terms.
We formulate the following conclusions for the case of linear regression model with the dependent variable of binary type:

1. One cannot compute a regression line to pass near the points of coordinates $\left(x_{i}, y_{i}\right)_{i=\overline{1, n}}$. This statement is verified by the graphic in figure 3 which shows a clear discrepancy between the actual and estimated values of the variable $Y$. Under these circumstances, $R^{2}$ or $\bar{R}^{2}$ are not appropriate statistical measures to characterize the quality of the estimates;
2. The intercept of the regression model is modified by changing the encoding of the binary variable. This situation highlights the impossibility of economic interpretation of the parameters of the regression model;
3. While the values of the variable $Y$ are
situated in the $\{0,1\}$ set, the estimated values are not limited in all cases to the two values. There may be situations in which the estimated values satisfy one of the following two inequalities: $\hat{y}_{i}<0$ or $\hat{y}_{i}>1$;
4. Because the error $\left(\varepsilon_{i}\right)$ may take only two values, it will follow a discrete distribution, so that the assumption of the normal distribution of the errors is not verified.
If we note with $p_{i}$ the probability that the
value of $Y$ is equal to 1 , then the following equality: $\varepsilon_{i}=1-\left(b+a x_{i}\right)$ is verified with the same probability. Because the value of $Y$ is equal to 0 with the probability $1-p_{i}$, then the value $\varepsilon_{i}=-\left(b+a x_{i}\right)$ is verified with the same probability. In these circumstances, the mean value and the variance of the residual variable are:
The mean of the variable is:

$$
E \varepsilon_{i}=\left(1-b-a x_{i}\right) p_{i}+\left(-b-a x_{i}\right)\left(1-p_{i}\right)=p_{i}+\left(-b-a x_{i}\right)
$$

From the restriction that the mean of the residual variable is equal to zero we obtain the restriction on the parameters of the linear regression model $p_{i}=b+a x_{i} \in[0,1]$. For this
reason the model is also called linear probability model.
The variance of the residual variable is:

$$
V \varepsilon_{i}=E\left(\varepsilon_{i}^{2}\right)=\left(1-b-a x_{i}\right)^{2} p_{i}+\left(-b-a x_{i}\right)^{2}\left(1-p_{i}\right)=p_{i}\left(1-p_{i}\right),
$$

To determine this relationship we take into account the restriction $p_{i}=b+a x_{i}$.
One can observe that the variance differs from an observation to another. The phenomenon of heteroskedasticity can be noticed at the model's level, so that the least squares method used to estimate the parameters do not lead to convenient results. In the same time, we cannot directly apply the generalized least squares method, since probabilities $p_{i}$ depend on parameters to be estimated.

## Latent variables

There are situations in which the values of the variable $Y$ are not directly observable. In such conditions, in order to define the econometric model, we introduce the concept of latent variable. It is a continuous variable that is not directly observable, but is representative for the phenomenon under study. For example, in the analysis of the scale of the financial fraud of companies in a development region it is difficult to directly assess the total amount of frauds, but we know the total number of frauds and the total fines paid by companies for the discovered financial fraud.

Thus, for each company we record the endogenous binary variable value ( $y_{i}$ ) by the rule below:
$y_{i}=\left\{\begin{array}{l}1, \text { the company paid a penalty } \\ 0, \text { the company haven't paid a penalty }\end{array}\right.$
The value taken by this observable variable depends on a set of exogenous variables that are denoted by $X_{1}, \ldots, X_{p}$. We denote by $v_{i}$ the maximum level of VAT that company can pay, so they could operate in the formal economy and by $v_{i}^{*}$ the latent variable that quantifies the business income. This feature is random because two companies that have the same characteristics $X_{1}, \ldots, X_{p}$ and have equal values for $v_{i}^{*}$ can be in different situations in relation with the area of formal or informal economy.
To define the variable $v_{i}^{*}$ we use the following regression model:
$v_{i}^{*}=a_{0}+a_{1} x_{1 i}+\cdots a_{p} x_{p i}+\varepsilon_{i}$.
The explanatory observable variable is one of dichotomic type. This is defined on the basis of the relationship given below:

$$
y_{i}=\left\{\begin{array}{l}
1, \text { if } v_{i}^{*}>v_{i} \\
0, \text { if } v_{i}^{*} \leq v_{i}
\end{array}\right.
$$

We compute the probability that a company have an income which is less than the level of the payable VAT ( $p_{i}=P\left(y_{i}=1\right)$ ), taking
into account that the residual variable $\varepsilon_{i}$ has a certain distribution. We successively obtain the following relations:

$$
\begin{align*}
p_{i} & =P\left(v_{i}^{*}>v_{i}\right)=P\left(a_{0}+a_{1} x_{1 i}+\cdots a_{p} x_{p}+\mathcal{q}_{i i}>v_{i}\right)= \\
& =P\left(\varepsilon_{i}>v_{i}-\left(a_{0}+a_{1} x_{1 i}+\cdots a_{p} x_{p i}\right)\right)  \tag{3}\\
& =P\left(\frac{\varepsilon_{i}}{\sigma}<-\frac{v_{i}-\left(a_{0}+a_{1} x_{1 i}+\cdots a_{p} x_{p i}\right)}{\sigma}\right)=F\left(-\frac{v_{i}-\left(a_{0}+a_{1} x_{1 i}+\cdots a_{p} x_{p i}\right)}{\sigma}\right)
\end{align*}
$$

Under these conditions, the probability $p_{i}$ depends on the residual variable distribution. Therefore, in order to define the binary model, we distinguish the following cases:

1. If the residual variable has a normal distribution, then we use a Probit model; 2. However, if the distribution is one of a logistic type, then we use a Logit model; 3. The distribution function may be a Weibull type.

## 6 Probit and Logit models Probit model definition

For the Probit model, the error repartition function is $N(0,1)$. Under these conditions we define the probability $p_{i}=P\left(y_{i}=1\right)$ by:

$$
p_{i}=\int_{-\infty}^{v_{i}^{*}} \frac{1}{\sqrt{2 \pi}} e^{-t^{2} / 2} d t \text {, (4) }
$$

where $v_{i}^{*}=a_{0}+a_{1} x_{1 i}+\cdots a_{p} x_{p i}$.

If we denote by $F$ the repartition function of $N(0,1)$, then, the Probit model has the equivalent form:

$$
\begin{equation*}
v_{i}^{*}=F^{-1}\left(p_{i}\right)=a_{0}+a_{1} x_{1 i}+\cdots a_{p} x_{p i} . \tag{5}
\end{equation*}
$$

If the variable $\varepsilon_{i}, i=\overline{1, n}$ is an independent one and the variable $y_{i}$ has only two values, then the verosimility function is defined by:

$$
L\left(y, a_{0}, \ldots, a_{p}\right)=\prod_{i=1}^{n} p_{i}^{y_{i}}\left(1-p_{i}\right)^{1-y_{i}} .
$$

After the estimation of the Probit model parameters by the maximum verosimility method we obtain the form given by equation (5).

## Logit model definition

In the case of the Logit model, the repartition function that defines the probability $p_{i}=P\left(y_{i}=1\right)$ is of logistic type:

$$
\begin{align*}
p_{i} & =P\left(y_{i}=1\right)=P\left(v_{i}^{*}>v_{i}\right)=P\left(\varepsilon_{i}>v_{i}-\left(a_{0}+a_{1} x_{1 i}+\cdots a_{p} x_{p i}\right)\right) \\
& =P\left(\varepsilon_{i}<a_{0}+a_{1} x_{1 i}+\cdots a_{p} x_{p i}-v_{i}\right)=F\left(\varepsilon_{i}<a_{0}+a_{1} x_{1 i}+\cdots a_{p} x_{p i}-v_{i}\right)(6  \tag{6}\\
& =\frac{1}{1+e^{-\left[a_{0}+a_{1} x_{i j}+\cdots a_{p} x_{p i}-v_{i}\right]}}
\end{align*}
$$

The repartition function can be written under the equivalent form:

$$
p_{i}=P\left(y_{i}=1\right)=\frac{e^{\left[a_{0}+a_{1} x_{1 i}+\cdots a_{p} x_{p i}-v_{i}\right]}}{1+e^{\left[a_{0}+a_{1} x_{1 i}+\cdots a_{p} x_{p i}-v_{i}\right]}}
$$

The properties of this function are:

1. If $a_{0}+a_{1} x_{1 i}+\cdots a_{p} x_{p i}-v_{i}>0$, the probability $p_{i} \in[0,1]$. Obviously, the maximum value of $p_{i}$ is equal to 1 and it can be obtained when the following condition stands: $a_{0}+a_{1} x_{1 i}+\cdots a_{p} x_{p i}-v_{i} \rightarrow \infty$. The value is equal to zero if $a_{0}+a_{1} x_{1 i}+\cdots a_{p} x_{p i}-v_{i} \rightarrow-\infty$. Using
this function we can eliminate the situations when the estimated values $\widehat{Y}_{i}=P\left(Y_{i}=0\right)$ are outside the [0,1]. More, the logistic function estimates the linear probability model. The two results obtained from a theoretical point of view are obvious for the practical side too. For example, if the income of all households is much greater then the necessary funds used to purchase a house, then the probability that all households have their own house tends toward 1. The reparation function is symmetrical. From (6) it results that:

$$
F(-x)=1-F(x)=F(x)
$$

2. The transformation of the model defined by (6) allows us to formulate this model under the form of the linear probability model (2):

$$
\begin{equation*}
\ln \left(\frac{p_{i}}{1-p_{i}}\right)=a_{0}+a_{1} x_{1 i}+\cdots a_{p} x_{p i}+\varepsilon_{i} \tag{8}
\end{equation*}
$$

Under these conditions we have:

$$
P\left(Y_{i}=1\right)=\ln \left(\frac{p_{i}}{1-p_{i}}\right)
$$

## Results interpretation

The estimation of the parameters of the Probit or Logit model can be carried out using algorithms to maximize a logverosimility function. Although the results obtained in the two cases (Probit and Logit) are relatively close, however, estimated coefficients are not directly comparable. Contrary to the situation in which the parameters of linear regression models are estimated by the least square method, when they are interpreted in terms of economic marginal propensity or elasticity, the values of the parameters of the model with linear probability have no direct interpretation. For the binomial model, two aspects are important: first, whether a parameter differs significantly from zero and the second is the sign of each parameter that indicates whether an explanatory characteristic act positively or negatively on the probability $p_{i}$.
However, it is possible to calculate for each explanatory variable the marginal rate to know the sensitivity of its variation on the probability $p_{i}$.
To test some aspects related to the binary model we take into account the following two possibilities:

1. Separately testing the signification of each explicative variable. We estimate the value of each parameter and then we calculate the mean square deviation and the $t$ statistics. For the binary model the $t$-statistics doesn't follow a Student repartition as for the general linear model, but a normal distribution;
2. Testing the signification of all explanatory variables. The elements of this test are presented below:
a. Test hypotheses are:

$$
\begin{aligned}
& H_{0}: a_{1}=\cdots a_{p}=0 \\
& H_{1}: \exists a_{i} \neq 0
\end{aligned}
$$

b. Test statistics is:

$$
L R=-2\left(\ln \left(L_{R}\right)-\ln \left(L_{U}\right)\right),
$$

where $L_{R}$ is the value of the log-verosimility function under the restrictions imposed by the null hypothesis, and $L_{U}$ is the value of the same function for the unrestricted model. According to the null hypothesis $L R$ statistics has $\chi^{2}$ repartition with $p$ degrees of freedom.
c. The decision for the $L R$ test is taken by comparing the value of this statistics with the value of the $\chi^{2}$ repartition from the table for a specified significance level and $p$ degrees of freedom. If the value of the $L R$ statistics is greater than the value read from the table then the $\mathrm{H}_{0}$ hypothesis is rejected.
Given the endogenous variable that has only two values, coded with 0 or 1 , the $\mathrm{R}^{2}$ coefficient of determination can't be interpreted in terms of adjusting the model. Therefore, in order to characterize the quality of the estimated model we use the pseudo- $\mathrm{R}^{2}$ statistics which is calculated using the following relationship:

$$
R^{2}=1-\frac{\log \left(L_{U}\right)}{\log \left(L_{R}\right)}
$$

## 7 Conclusions

Binary models are recommended as often a number of variables, such as the size of the informal economy at the level of a company and the size of unpaid taxes and duties due to their activities in the informal economy, are not known in official statistics. These variables can be substituted to some extent by the latent variables, for which data series are known and are correlated with variables that characterize certain aspects of the informal economy.
By means of the Logit or Probit models we can estimate the probability that a company or a person accept the risk of operating in the informal economy with all the advantages and disadvantages of this decision.
For the estimation of the parameters of these
models SPSS or EVIEWS procedures can be used [2].

## Acknowledgements.

The work related to this paper was supported by the PN II projects, financing project no. 862/13.01.2009, CNCSIS Code: 1793, financing project no. 763/13.01.2009, CNCSIS Code: 1814, financing project no. 953/13.01.2009, CNCSIS Code: 1857 and CNMP-PNCDI II projects, financing project no. 92082/01.10.2008 and financing project no. 91054/18.09.2007.

## References

[1] L.-L. Albu, "Tax Evasion and the Size of Underground Economy: a Theoretical and Empirical Investigation", Romanian Journal of Economic Forecasting, vol. 1-2, no. 1-2, pp. 16-31, 2001.
[2] T. Andrei and R. Bourbonnais, Econometrie, Bucharest: Economica Publishing House, 2008.
[3] R. French, M. Balaita and M. Ticsa, "Estimating the Size and Policy Implications of the Underground Economy in Romania", US Department of the Treasury, Office of Technical Assistance, Bucharest, August 1999.
[4] R. Hussmanns, "International Standards on the Measurement of Economic Activity, Employment, Unemployment and Underemployment", Labour Statistics for a Market Economy: Challenges and Solutions in the Transition Countries of Central and Eastern Europe and the Former Soviet Union, Budapest: Central European University Press in association with the International Labour Office, 1994.
[5] R. Hussmanns, "Statistical Definition of Informal Employment: Guidelines Endorsed by the Seventeenth International Conference of Labour Statisticians (2003)", Geneva: International Labour Office, Bureau of Statistics, 2004.
[6] S. Kim, "Expansion of Markets and the Geographic Distribution of Economic Activities: The Trends in US Regional Manufacturing Structure. 1860 - 1987". Quarterly Journal of Economics, no. 70, pp. 881-908, 1995.
[7] F. Schneider, "Shadow Economies Around the World: What Do We Really Know?" European Journal of Political Economy, vol. 21, no. 3, pp. 598-642, 2005.


Tudorel ANDREI has graduated the Faculty of Planning and Economic Cybernetics, Bucharest University of Economics in 1989 and The Faculty of Mathematics, University of Bucharest in 1997. He holds a PhD diploma in "Statistics" from 1996 and has a remarkable didactic and research activity since 1991 when he joined the staff of Bucharest Academy of Economic Studies. Currently, he is full Professor of Statistics and Econometrics within the Department of Statistics and Econometrics. He is the author of more than 15 books and an impressive number of articles. He is also project director or member in many important research projects. He has developed abroad teaching activity as Invited Professor at the Paris - Dauphine IX University, France and made various research stages at Paris Dauphine IX University, France, Emory University, Atlanta, USA etc. His research interests include econometrics, macro econometrics, micro econometrics, statistics, public administration, economy of corruption, microeconomics etc.


Bogdan OANCEA has graduated the Faculty of Physics, University of Bucharest in 1991 and the Faculty of Cybernetics, Statistics and Economic Informatics, Bucharest University of Economics in 1998. He holds a PhD diploma in "Cybernetics and Statistics" from 2001 and has a remarkable didactic and research activity since 1991. Currently, he is full Professor of Economic Informatics at the Artifex University. He is the author of 7 books in the field of computer science and an impressive number of articles. He was
also project director or member in many important research projects. His research interests include parallel algorithms for matrix computations, parallel algorithms in econometrics, operating systems, GPU computing.


Stelian STANCU has graduated the Faculty of Planning and Economic Cybernetics, Bucharest University of Economics in 1991 and the Faculty of Mathematics, University of Bucharest in 1998. He holds a PhD diploma in "Economic Cybernetics and Prediction" from 1996 and has a remarkable didactic and research activity since 1991 when he joined the staff of Bucharest Academy of Economic Studies. Currently, he is full Professor of Microeconomics, Quantitative Economics, Economics of Information, Methods and techniques for project prediction, Processes and business modeling within the Department of Economic Cybernetics. He is the author and coauthor of more than 20 books and an impressive number of articles. He is also project director or member in many important research projects. He made various research stages at Paris - Dauphine IX University, France, Emory University, Atlanta, USA. His research interests include econometrics, statistics, microeconomics, macroeconomics etc.


Andreea Iluzia IACOB has graduated the Faculty of Cybernetics, Statistics and Economic Informatics, Bucharest University of Economics in 1994. She holds a Master Diploma in "Quantitative Economics" from 1995 and a PhD Diploma in "Cybernetics and Economic Statistics" from 2003 and has a remarkable didactic and research activity since 1995 when she joined the staff of Bucharest Academy of Economic Studies. Currently, she is full Professor of Econometrics within the Department of Statistics and Econometrics. She is the author and coauthor of more than 10 books and an impressive number of articles. She is also project director or member in many important research projects. She has developed abroad teaching activity as Invited Professor at the University of Alcala de Henares (Madrid), Spain and at the Paris - Dauphine IX University, France and made various research stages at Paris - Dauphine IX University, France, Her research interests include econometrics, statistics and economic statistics, macroeconomics, macro econometrics etc.

