A Model to Estimate Informal Economy at Regional Level: Theoretical and Empirical Investigation

Lucian-Liviu Albu
Institute for Economic Forecasting, Romanian Academy
E-mail: albul@ipe.ro

- Abstract -

Many problems emerge since it is widely believed that high tax rates and ineffective tax collection by government are the main causes contributing to the rise of the informal economy. Already the economists have established a relationship between tax rates and tax evasion or size of the informal economy. The higher is the level of taxation, the greater incentive is to participate in informal economic activities and escape taxes.

At the macroeconomic level, there is a number of so-called indirect methods used to estimate the size and dynamics of informal economy, reported in literature as "Monetary Approach", "Implicit Labour Supply Method", "National Accountancy", "Energy Consumption Method", etc. Unfortunately, many times there are huge differences among the estimated shares of informal or underground economy obtained by various methods. For instance, in case of Romania the figures are between about 20% of GDP, obtained on the base of the energy consumption method and more than 45% computed by using the monetary approach. Also, the figures reported by the National Institute for Statistics (NIS), based on national accounts methodology, increased (mainly due to changes in methodology) from about 5% in 1992, to 18% in 1997 and to 20-22% after 2000. Adding to these figures about 7% of GDP, representing the estimated level for self-consumption in case of a rural household, legal non-registered but informal, resulted that last years the informal economy is responsible of 27-29% of national economy.

In this article, coming from certain general accepted finding of the theory in matter of modelling underground economy, we concentrate on evaluating analytically the limit-values of certain important parameters involved in models used to estimate the size of underground economy and to explain the mechanisms of its dynamics. Then we shall simulate some exercises on available data. The second goal of the paper is to report some conclusions of our investigation based on data supplied by special surveys organised in Romania. Also, in order to see since certain hypotheses (referring to the complex transmission mechanism from the tax policy decisions to the effective implication of agents into informal economy) are statistically verified and to extend the study from the aggregate level to a deep research inside the population set in regions, we used data supplied by this special large survey, which already were processed and are available in our database.

1. Introduction

There is a vast literature concerning the impact of fiscal policy and tax rate on macroeconomic stabilisation programs. New developments are recorded regarding this subject, from the emergence of supply-side economics. Without ignoring the regulatory role of the government budget, the main idea is that the expansion of tax rate over some thresholds, continuing to have positive effects on the demand-side and, consequently, on the stabilisation process in the short-run, can have negative impact on the supply-side. In the long run, the stabilisation process itself will be affected. The general explanation is that an increasing tax rate may limit the stimulation of private (capitalistic) firms to investment, which is the main factor of development in a free market economic system.

Despite the fact that there are many complications, at macroeconomic level, induced by the structural changes, significant differences between the responses to various taxes, as well as by the business cycles and many other factors, however, the mentioned impact was demonstrated, at least as a general trend in the long run. Generally accepted the idea that an increase of tax rate provokes a migration of activity from the visible sector to the invisible sector, the remaining problem is: what is the intensity of the correlation between the tax rate level and the proportion of this migration. Moreover, there is the problem of how should this intensity change and by what quantity under the impact of other economic or/and non-economic factors. To evaluate some essential parameters of this transfer of activity (migration), and to produce explanations accepted by standard economics are the main goals of the underground economy modelling.

We are concentrated on some theoretical aspects and on empirical study of the impact of various factors on the size and dynamics of informal economy. Coming from the existing literature we try to build a simple model in order to estimate the theoretical limit-values for informal economy based only on a limited number of available macroeconomic data, on the one hand. On the other hand, using data from some old surveys on Romanian economy, we investigate the behaviour of various groups of households in matter of participation in informal activities. Empirical study demonstrated that the main reason of informal economy is the general household's standard of living or more concrete the level of actual income obtained by work in formal (official or visible) sector of the economy. Then, coming from the regression equations estimated on the available data in two surveys organised in Romania in 1996 and 2003, respectively, we tried to extend the model to estimate the size and dynamics of informal economy at regional level. In Romania there are eight geographic regions for which beginning from few years ago the National Institute of Statistics publish in the Territorial Statistics section some essential indicators such as total income of main household categories. According to NSI, the regions are as follows:

1) North - East; 2) South - East; 3) South - Muntenia; 4) South - West Oltenia; 5) West; 6) North - West; 7) Center; 8) Bucharest - Ilfov.

2. Standard theoretical models

There are many studies that approach the psychological effect of the modern State development that has been to dig an important gap between the government and the citizen. Some authors do not hesitate to see in the underground economy a refuge against the will of the State to organise the society, and they affirm that the main "production" of a state is the law. Sometimes this production becomes an inside goal. There is a vicious cycle, then: the state produces many laws, of which some are not complied with in practice, and then the state produces more laws to counteract this practical inefficiency,

2

and then more laws are not complied with in practice, and so on. According to this opinion, in order to be optimal a state must produce only those laws that will have a chance to be complied with in practice. Each excess-output relating to this optimal level will generate inefficiency and perverse effects, as well as increasing of underground economy. By describing the attitudes of taxpayers, many reasons can be evoked - coercive nature of fiscal policy, lack of identification with the administration, aspects relating to the risk behaviour in the case of law infringement detection and the inefficiency of the administrative management.

To describe it in a sentence, the more the fiscal cost is strong; the greater will be the desire to avoid paying for each individual reason. This phenomenon has been made evident by the Laffer curve, according to which the receipts of the State begin by increasing, and then decrease below one certain threshold, as the average taxation rate increases. Gutmann has amended the Laffer curve. He thinks that the increase of the rate of taxation increasingly incites the individuals to escape the fiscal controls, by penetrating in the sphere of the underground activities to the point to distort the Laffer curve (Gutmann, 1977). According to Gutmann, the Laffer curve must be in fact skewed towards the right part of the graph. The weight of the fiscal payments and social contributions, it is to tell the obligatory payments, has constantly increased during the last century, and its rate of growth is now such that the acceleration is perceptible in a single generation. Generally speaking, high taxation rates, few controllers of taxes and relatively weak amend for fiscal fraud contribute to persuade the people to take their chance to be undiscovered. On a theoretical plan, the model of Allingham and Sandmo (1972) strengthen this position. The problem of taxpayer consists indeed in maximisation of the hoped usefulness of the obtained income if he develops a fiscal fraud strategy.

Consider the model of Allingham and Sandmo where a risk-averse taxpayer is allowed to declare less than his actual income, X. Declared or normal income, $Xn > 0$, is taxed by a constant rate, $\theta > 0$, whereas the undeclared income, $X - Xn$, is taxed, if detected, by a higher rate, π. The taxpayer chooses Xn^* to maximise his expected utility:

$$E[U] = (1 - p) U(Y) + p U(Z) \tag{1}$$

where p is the (exogenously given) probability of detection, and

$$Y = X - \theta Xn \tag{2}$$
$$Z = X - \theta Xn - \pi (X - Xn) \tag{3}$$

represent his income in the case of detection and non-detection, respectively. The first-order condition for the maximisation of (1) is

$$dE[U] / dXn = -\theta (1 - p) U'(Y) + (\pi - \theta) p U'(Z) = 0 \tag{4}$$

from which the taxpayer's response to a change in θ may be derived. This is given by

$$dXn^* / d\theta = -D^{-1} (1 - p) U'(Y) \{\theta Xn [R_A(Z) - R_A(Y)] - [\pi / (\pi - \theta)]\} \tag{5}$$

where $D \equiv \theta^2 (1 - p) U''(Y) + (\pi - \theta)^2 p U''(Z) < 0$ is the second-order condition for the maximisation of (1) and $R_A(I) = -U''(I) / U'(I) > 0$ is the Arrow-Pratt absolute risk-aversion measure, evaluated at $I = Y, Z$.

Under decreasing absolute risk-aversion [$R_A(Z) > R_A(Y)$], the sign of (5) is ambiguous, as asserted by Allingham and Sandmo. However, given that the relative risk-aversion is constant ($R_A(I) = c$), implies that

3

$dX_n^* / d\theta \ > (\text{or} <) \ 0 \ \text{if} \ c\theta X_n \ [(Y-Z)/(YZ)] > (\text{or} <) \ \pi / (\pi - \theta)$ (5')

or, by substituting (2) and (3) into (5') and rearranging, that

$dX_n^* / d\theta \ > (\text{or} <) \ 0 \ \text{if} \ c > (\text{or} <) \ \alpha \ [1 + \beta \ (X/X_n)]$ (5'')

where $\alpha \equiv (X - \theta X_n) / \theta(X - X_n)$ and $\beta \equiv (1 - \pi) / (\pi - \theta)$. Clearly, $\alpha > 1$, and $\beta > (\text{or} <) 0$ if $\pi > (\text{or} <) 1$. Hence, $\pi \leq 1$ ensure that $\alpha(1 + \beta) > 1$, so that $dX_n^*/d\theta < 0$ if $c \leq 1$ (notice that $c > 1$ does not necessarily imply that $dX_n^*/d\theta > 0$). However, since β varies inversely with π, the stricter the restriction imposed on π, the higher the ceiling on c allowed for the satisfaction of $dX_n^*/d\theta < 0$ (recall also the Allingham and Sandmo's result that $dX_n^*/d\pi > 0$, so that X/X_n rises as π falls). Still, since all that is known on X/X_n is that it exceeds unity, the upper bound on c which may safely be identified as yielding the desired prediction is just $1+\beta = (1 - \theta) / (\pi - \theta)$, which rises with θ. Referring again to its minimal value, we conclude that if $\pi \leq 1$ and $c \leq 1/\pi$, condition (5'') surely implies that when "the fruits of evasion become sweeter, a rational taxpayer will take a bigger bite".

If the penalty rate is 100% (i.e. the entire undeclared income is confiscated), a negative relationship between declared income and the income tax rate would unambiguously hold if the relative risk-aversion coefficient does not exceed unity. In more realistic cases, where the penalty rate is lower, the desired prediction will also hold for a higher than unity coefficient of relative risk-aversion. For instance, in the U.S., a detected evader is obliged to pay less than 1.5 times his *evaded taxes* for most violations (Pencavel, 1979, p. 122), whereas the ratio of tax payments to adjusted gross income averages less than 20% (ibid., p. 122). This means that the penalty rate on *undeclared income* is less than 30%. Hence, if the penalty scheme in the U.S. were made dependent on undeclared income, keeping the punishment level intact, a negative relationship between declared income and the income tax rate would be predicted for a relative risk-aversion coefficient that does not exceed 3.33.

Since this work, a sizeable literature applying economic analysis to tax evasion has appeared. Many of the theoretical papers on this topic have used models in which a representative taxpayer receives income from a single source and then decides how much of that income to declare. While such studies reveal much about the behaviour of a single underreporting agent, they do not incorporate one important empirical fact about tax evasion, namely that the extent to which one evades taxes is strongly correlated with the source of one's income (Clotfelter, 1983).

In the recent years, several theoretical papers have recognised that opportunities for evasion differ among occupations. Such papers have also emphasised that these differences may affect an agent's labour market behaviour. An agent may, in other words, base his labour supply decisions in part on the ability to evade taxes. Theoretical models that recognise tax evasion to be easier when income is received from certain sources (self-employment, for instance) rather than other sources (corporate employment, more frequently) typically assume that the economy has two sectors: one in which evasion is impossible, presumably because of tax holding and information reporting, and another in which evasion is possible. Such papers are often referred to as studies of the underground economy.

Generally, the models focused on the situation of the underground economy in Western countries. Coming from the fundamental tension in the applied literature, which refers to the interpretation of labour force in the underground sector, Gibson and Kelley (1994) developed a theoretical model that, beside the general problems of underground economy,

focused on the case of informal sector in the developing countries. An important conclusion derived from their model is that the costs in the informal sector must be greater than in formal sector. Then, if the informal sector uses more resources per unit of output, the social surplus will expand as the formal sector replaces the informal sector. The outcome, however, depends upon the existing distribution of income, which further complicates matters.

In case of Romania in transition period, we mention the methodology we used (Albu et al., 1998) in order to estimate the impact of underground sector to the general dynamics of economy, and the relation between underground economy and institutional development (Daianu and Albu, 1997). Some studies seem to confirm the hypothesis of Laffer's curve model (Lemieux et al., 1994), but other authors contest the availability of this model at least at macroeconomic level (Pestieau, 1989). A detailed study on the Laffer's curve model and its implications on the fiscal policy plan can be find in our work achieved in 1995 within CEPREMAP (Albu, 1995).

Also, in literature there are a number of other models trying to estimate the size of underground economy and to simulate its formatting mechanisms. One of them is the model based on the so-called monetary approach, but which generally produced some not very robust estimation and was many times contested by economists (see Isachsen and Strom, 1985, for the case of Norway, and French et al., 1999, for the Romanian economy). Also, we used two generalised models based on the so-called labour force supply approach and the allocation of working time fund approach, respectively. Unfortunately, due to some poor adequate statistical data it is possible to simulate these models on Romanian economy in the period of transition only as an exercise.

3. New hypotheses and estimation models

Now, we are concentrating on developing the previous theoretic model of tax evasion in order to estimate underground economy bases on usually available macroeconomic data. Thus, we consider that at macroeconomic or aggregated level the behaviour of taxpayer is conforming to some supplementary hypotheses, expressed by the following equations. To facilitate understanding, they are grouped in a number of sub-models. Then, the sub-models will be assembled within a model that takes into account the impact of the penalty tax. This type of tax is applied when the fiscal authorities discover subtractions from tax pay duties or illegal activities. It is as a rule higher than the so-called normal tax rate, but smaller than 1 (the case corresponding to the confiscation of the entire non-declared income by the fiscal authorities).

Sub-model S1

In order to capture the impact of a penalty tax rate on the yearly average tax rate, we consider a Laffer's curve model. Thus, we consider two conventional stages of the tax collecting process.

The first stage is referring to the normally paid taxes that are collected from taxpayers. To express the total amount of taxes received by fiscal authorities in this stage we recall the simple expression of a standard Laffer's curve model:

$$X_n(\theta) = (1 - \theta) X \tag{6}$$
$$X_a(\theta) = \theta X \tag{7}$$
$$T_n(\theta) = \theta X_n = \theta (1 - \theta) X \tag{8}$$
$$Y_n(\theta) = (1 - \theta) X_n = (1 - \theta)^2 X \tag{9}$$
$$Y_I(\theta) = X - T_n = (1 - \theta + \theta^2) X \tag{10}$$

where Xn is the declared income, θ - average value of normal tax rate, X – total actual income or total GDP (declared and non-declared), Xa – undeclared income or hidden GDP at the end of normal tax procedure, Tn – total amount of taxes received by normal procedures, Yn – the remaining disposable income from declared income after its normal taxation, and YI – total remaining disposable private income (households and private sector) in case of first stage with no penalty procedure (the entire initially hidden income, Xa, remains non-discovered, and consequently YI=Yn+Xa). A graphical representation can be found in Figure 1, where the indicators are reported as percentage of total income or GDP, X.

   Sub-model S2

   In case of considering the second stage, the fiscal authorities discover a part of the initially hidden activity, which now became visible, and by applying penalty procedures they gather supplementary income for the public budget, as follows:

$$X_d(p) = p\,X_a \tag{11}$$
$$X_i(p) = (1 - p)\,X_a \tag{12}$$
$$T_d(p) = \pi\,X_d = \pi\,p\,X_a \tag{13}$$
$$Y_d(p) = (1 - \pi)\,X_d = p\,(1 - \pi)\,X_a \tag{14}$$

where Xd is the detected part of the initially non-declared income, p – average probability of detection, Xi – total actual income finally non-detected or invisible GDP (non-declared and non-detected), Td – total amount of taxes received by penalty procedures, π – average of penalty tax rate, and Yd – the remaining disposable income from discovered income after penalty taxation. A graphical representation can be found in Figure 2, where the indicators are reported as percentage of total initially hidden income or GDP, Xa.
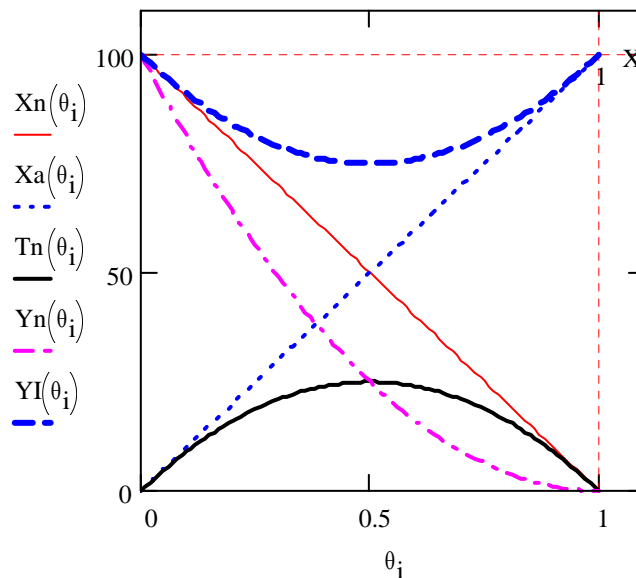

Figure 1. The behaviour of the system under the hypotheses of Sub-model S1

   Sub-model S3

   The main hypothesis of this model is that the probability of detection, p, depends positively both on the initially hidden economy size and on the total amount of taxes

6

finally collected by the state (i.e., the probability of detection becomes endogenous), expressed as shares in total actual income or GDP, as follows:

$$p(xa,t) = xa\, t \qquad\qquad (15)$$
$$xa = Xa / X = \theta \qquad\qquad (16)$$
$$t = T / X = (Tn + Td) / X = \theta(1 - \theta) + \pi\, p\, \theta \qquad\qquad (17)$$

where $T = Tn + Td$.

The probability of detection increases when the share of underground (hidden) economy becomes larger (the so-called omnipresence of informal economy). On the other hand, it is increasing when the share of total collected taxes, t, in GDP is growing (the so-called omnipresence of the state's power).

Based on relations (15)-(17), after some algebraic operations, the probability of detection can be written as following:

$$p(\theta) = [\theta^2(1-\theta)] / (1 - \pi\,\theta^2) = [\theta^2(1-\theta)] / (1 - k\,\theta^3) \qquad\qquad (18)$$

where k is the coefficient of multiplying the normal taxation rate, θ,

and $\pi = k\,\theta$ \qquad\qquad (18')
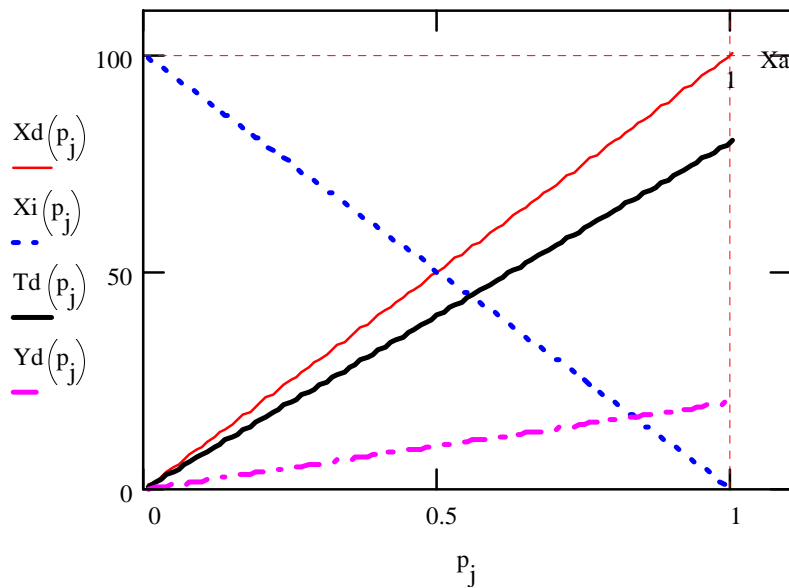
Figure 2. The behaviour of the system under the hypotheses of Sub-model S2

Taking into account that the maximum level of penalty rate, π, is 100% (i.e., the entire discovered income is confiscated), in case of a given level of θ, the maximum level of coefficient k will be equal to $1/\theta$ (its minimum level is indeed 1, i.e. $\pi = \theta$). So, the domain in which the probability of detection, p, could be placed is delimited by the lines $pmin(\theta) = [\theta^2(1-\theta)] / (1-kmin\,\theta^3)$ and $pmax(\theta) = [\theta^2(1-\theta)] / (1-kmax\,\theta^3)$, as it is shown in Figure 3.

Figure 3. The probability function under the hypotheses of Sub-model S3

<u>The general model</u>

Combining the three above-mentioned sub-models a general model, which can be described by the following set of equations, resulted. Thus, based on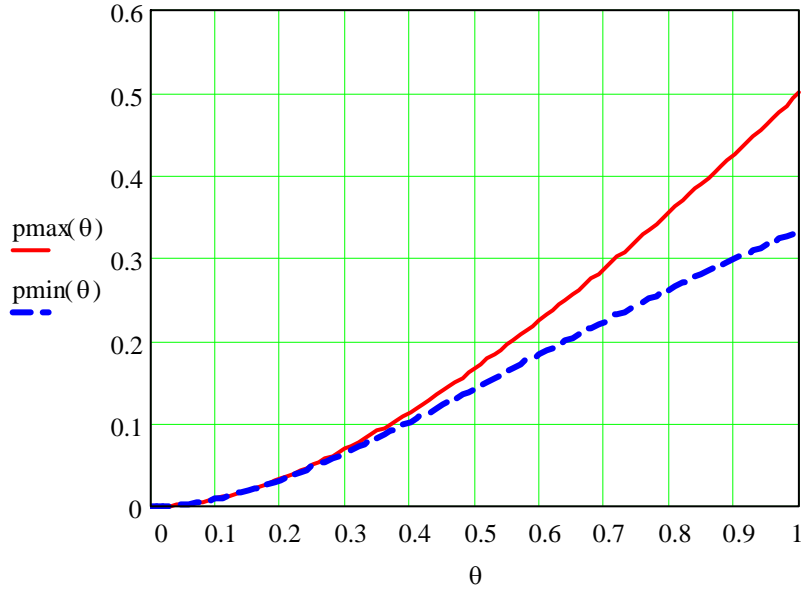 the behaviour of taxpayers, the national income or GDP is divided in three fundamental parts: Xn – declared income, Xd – undeclared income, but then discovered by authorities, and Xi – undeclared income, but then continuing to remain undiscovered by authorities. The latter is called "invisible sector" in order to make difference from the "visible sector", which comprises two parts ($Xv = Xn + Xd$). The official published statistical data are usually referring only to the visible sector. Thus, the yearly or quarterly published GDP means in fact only Xv.

1) Distribution of total income among sectors:

$$X \; = \; Xn + Xa \; = \; Xn + Xd + Xi \; = \; Xv + Xi \tag{19}$$

where $Xn = xnX = (1-\theta)X$; $Xa = \theta X$; $Xd = pXa = \theta pX$; $Xi = \theta(1-p)X$; $Xv = (1-\theta+\theta p)X$.

2) Distribution of income within sectors:

$$Xn \; = \; Tn + Yn \; = \; \theta \, Xn \; + \; (\, 1 - \theta \,) \, Xn \tag{20}$$
$$Xd \; = \; Td + Yd \; = \; \pi \, Xd \; + \; (\, 1 - \pi \,) \, Xd \tag{21}$$
$$Xi \; = \; Yi \tag{22}$$

3) Distribution of total income by destinations:

$$X \; = \; T + Y \tag{23}$$

where

$$T = Tn + Td = \theta \, Xn + \pi \, Xd = \theta \, (\, 1 - \theta \,) \, X + \theta \, \pi \, p \, X = \theta \, (\, 1 - \theta + \pi \, p \,) \, X \tag{24}$$
$$Y = Yn + Yd + Yi = (\, 1 - \theta \,) \, Xn + \theta \, p \, (\, 1 - \pi \,) \, X + \theta \, (\, 1 - p \,) \, X \tag{25}$$

8

The economic system described by the general model demonstrates a very complicated dynamics, essentially varying with parameters as θ and π or, by extension, with k and p. One conclusion is that ***the economic policy and its tools can be oriented either to the fiscal pressure side*** (i.e., θ) ***or to the improvement of the penalty side*** (i.e., π). However, to find an optimal solution between the two sides of policy continues to be an unsolved problem, at least in the case of applied policy. Such complicated dynamics is in a way similar to that deduced from the Allingham and Sandmo model.

In the next section, focusing on obtaining a procedure to evaluate the limits between which the underground economy could be framed, many times we shall use implicitly the so-called *computer assisted* proofs instead of rigorous proofs. Concerning latter, we feel that most economists will not be interested in the means and often very technical details; therefore we sometimes just give sketches.

4. Estimating the limit-values of the underground economy size

Usually, at macroeconomic level, the available statistical data are only the following aggregated indicators: Xv – total visible income or reported GDP, T – total revenues collected from people and economic agents (there are included here mainly those coming from the state budget, local budgets, and state social insurance budget), and *eventually* Td – total amount of taxes collected by penalty procedures, in case of detection (in case of Romania, here there are included: "Tax on profit from illegal commercial activities or against Law on the consumer protection", "Delay increases and penalties for term unpaid revenues", "Judicial fines", etc.).

Considering the case when only statistical data series of Xv and T are available, in the context of our model it is useful to recall the following two fundamental relations: T = θ(1-θ+pπ)X, and Xv = (1-θ+θp)X, respectively. They permit to calculate the statistical reported tax rate (tst), as follows:

$$\text{tst} \;=\; T \,/\, Xv \;=\; [\,\theta\,(\,1 - \theta + p\,\pi\,)\,] \,/\, [\,(\,1 - \theta + \theta\, p\,)\,] \qquad (26)$$

Now, replacing p and π by their definitions (relations (18) and respectively (18')), the actually tax rate can be written as follows:

$$\text{tst}\,(\theta, k) \;=\; \theta \,/\, [\,1 - \theta^{\,3}\,(\,k - 1\,)\,], \quad \text{with } 0 < \text{tst} < 1 \qquad (27)$$

Also, we can write the reverse relation of k, as follows:

$$k\,(\theta, \text{tst}) \;=\; 1 + \{\,(\,1 \,/\, \theta^{\,2}\,)\,[\,(\,1 \,/\, \theta\,) - (\,1 \,/\, \text{tst}\,)\,]\,\} \qquad (28)$$

with $\theta < \text{tst}$ and $k_{max} = 1/\theta$.

Taking into account that tst could theoretically be placed within the interval [0; 1] and k within [1; 1/θ], now we can calculate the new form of k_{max} as a function of θ, as follows:

$$k_{max}\,(\theta) \;=\; [\,1 - \theta\,(\,1 - \theta^{\,2}\,)\,] \,/\, (\,\theta^{\,3}\,) \qquad (29)$$

The most important result obtained was coming from the two theoretic limits of k: a very restrained interval of values in which could be placed the variable θ. Thus, in the case of our model (where θ is only theoretically between 0 and 1), the maximum level of θ is in the case of zero detection $\theta_{max} = \text{tst}$, and its minimum level, respectively, is given by only

one solution (having economic significance) of the three solutions of the following equation:

$$tst\,\theta^3 - tst\,\theta^2 - \theta + tst = 0 \qquad (30)$$

The analytical expression of this remarkable value of θ, noted as "$\theta 1$", was obtained by using the so-called *computer assistance* (here, because its formula is too large to be presented it was omitted).

Now, the main equations that are describing the economic system behaviour and may be used to build up an efficient econometric strategy can be expressed as follows:

$$test = [\,(1-\theta)\,tst\,]\,/\,(1-\theta^2\,tst\,) \qquad (31)$$
$$xvest = (1-\theta)\,/\,(1-\theta^2\,tst\,) \qquad (32)$$
$$pest = [\,(1-\theta)\,\theta\,tst\,]\,/\,(1-\theta^2\,tst\,) \qquad (33)$$

where test is the best estimator of $t=T/X$, xvest – the best estimator of $xv=Xv/X$, pest – the best estimator of $p=Xd/Xa$, and

$$\theta \in [\,\theta 1(tst),\,tst\,] \qquad (34)$$

Alternatively, we should consider also a simplified variant for the probability of detection definition to replace relation (15), as follows:

$$p\,(xa) = xa \qquad (15')$$

In this case, again the resulted maximum level of θ (obtained by making $\pi=\theta$ in the relation of definition for tst) is equal to tst, but its minimum level (obtained by making $\pi=1$ in the relation of definition for tst) will be as follows:

$$\theta 2\,(tst) = [(1+tst) - (1 + 2\,tst - 3\,tst2)^{1/2}]\,/\,(2\,tst) \qquad (35)$$

and

$$\theta \in [\,\theta 2(tst),\,tst\,] \qquad (34')$$

A graphical representation of the area within which parameter θ is forced to vary in the two cases can be found in Figure 4. The two variation areas of θ are between the solid line, tst, and $\theta 1$ (doted line) and between the solid line, tst, and $\theta 2$ (dashed line) respectively. Figure 5 presents the behaviour of t and xv, as function of tst and respectively θ in case of using $\theta 1$ parameter (here, tst_i are the i values of tst and θ_{ij} are the ij values of θ; but $\theta 1(tst_i)$ $< \theta_{ij} < tst_i$, where $i = 1, 2,\ldots, 40$ and $j = 1, 2, \ldots, 20$). Figure 6 presents the behaviour of t and xv, as function of tst and respectively θ in case of using $\theta 2$ parameter (here, tst_i are the i values of tst and θ_{ij} are the ij values of θ; but $\theta 2(tst_i) < \theta_{ij} < tst_i$, where $i = 1, 2,\ldots, 40$ and $j = 1, 2, \ldots, 20$).

In order to apply the model to the Romanian economy for 1989-2004, we used statistical data from National Accounts. Also, as a measure of overall tax ratio we considered, according to the OECD methodology (Blades, 1922; Schneider, 2002), the fiscal ratio together with the social ratio. Estimated values for certain indicators obtained by using $\theta 2$ parameter are presented in Table 1.

Despite the model makes more complex the variation map of certain factors involved in explaining the mechanism of stimulating people to participate in informal activities, many

others remain unknown at regional level. Thus, aside to the tax rate other candidates to explain the complex process of underground sector development seem to be the productivity level expected in this sector (and further specific level of wages), level of income obtained in visible sector, etc.
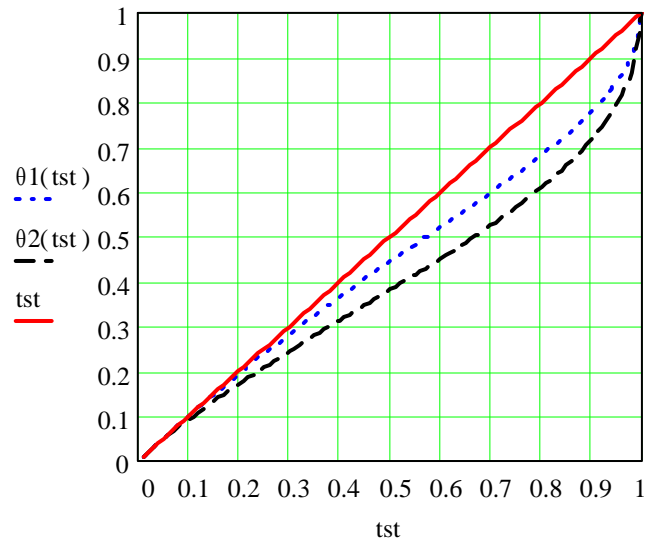

Figure 4. The restricted variation areas for parameter θ: θ ∈ [θ1(tst), tst] and θ ∈ [θ2(tst), tst] respectively
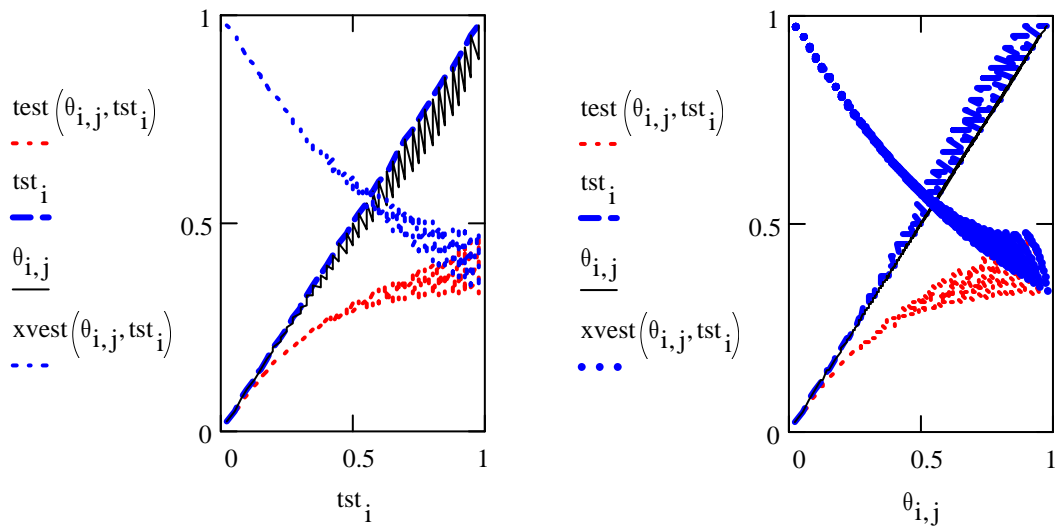

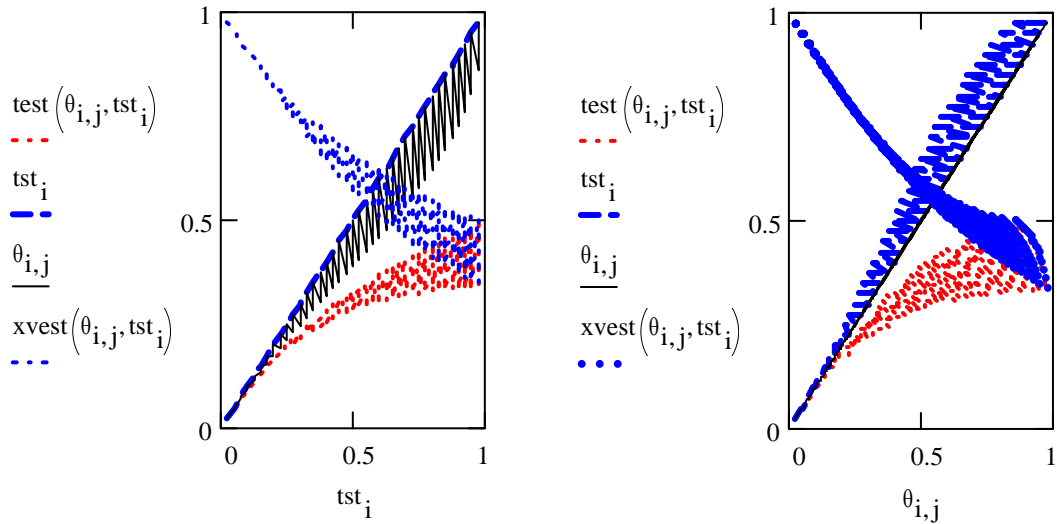**Figure 5. The behaviour of t and xv, as function of tst and θ for θ1 < θ < tst**

Figure 6. The behaviour of t and xv, as function of tst and θ for θ2 < θ < tst

Table 1. Main indicators expressing the informal economy extension in Romania
- as % of X (total GDP) -

	tst	xv	xi	xn	xa	t	y
1989	35.9	67.2-73.6	26.4-32.8	64.1-71.4	28.6-35.9	24.1-26.4	73.6-75.9
1990	35.2	67.8-74.0	26.0-32.2	64.8-71.9	28.1-35.2	23.8-26.0	74.0-76.2
1991	34.0	68.7-74.6	25.4-31.3	66.0-72.7	27.3-34.0	23.4-25.4	74.6-76.6
1992	31.4	70.8-76.1	23.9-29.2	68.6-74.6	25.4-31.4	22.2-23.9	76.1-77.8
1993	28.5	73.2-77.8	22.2-26.8	71.5-76.6	23.4-28.5	20.9-22.2	77.8-79.1
1994	28.7	73.0-77.7	22.3-27.0	71.3-76.5	23.5-28.7	21.0-22.3	77.7-79.0
1995	26.9	74.6-78.8	21.2-25.4	73.1-77.8	22.2-26.9	20.1-21.2	78.8-79.9
1996	26.5	74.9-79.1	20.9-25.1	73.5-78.0	22.0-26.5	19.8-20.9	79.1-80.2
1997	28.0	73.6-78.1	21.9-26.4	72.0-77.0	23.0-28.0	20.6-21.9	78.1-79.4
1998	30.4	71.6-76.7	23.3-28.4	69.6-75.3	24.7-30.4	21.8-23.3	76.7-78.2
1999	32.5	69.9-75.5	24.5-30.1	67.5-73.8	26.2-32.5	22.7-24.5	75.5-77.3
2000	35.0	67.9-74.1	25.9-32.1	65.0-72.0	28.0-35.0	23.8-25.9	74.1-76.2
2001	28.4	73.3-77.9	22.1-26.7	71.6-76.7	23.3-28.4	20.8-22.1	77.9-79.2
2002	28.3	73.4-77.9	22.1-26.6	71.7-76.7	23.3-28.3	20.8-22.1	77.9-79.2
2003	27.5	74.0-78.4	21.6-26.0	72.5-77.3	22.7-27.5	20.4-21.6	78.4-79.6
2004	27.3	74.2-78.6	21.4-25.8	72.7-77.5	22.5-27.3	20.3-21.4	78.6-79.7

5. Estimating the informal economy at regional level

As empirical studies based on surveys demonstrate, one of the most significant determinants of participation in informal activities is the average income per person in household obtained in formal sector. Moreover, we should consider many other factors as stimulating households to involve in the informal sector, such as occupation, region, age, education, etc., as we proceeded in some old studies (Duchene et al., 1998; Albu et al., 2002a and 2002b; Albu and Nicolae, 2002). Based on such special organised surveys certain useful conclusions could be outlined: 1) People in households perceive the high rate

of taxation as the main cause of the underground activity (more than 80% of answers in surveys demonstrate this idea); 2) Separating the main motivations of operating in informal sector in two groups – "subsistence" and "enterprise", respectively, the data in surveys suggest that the subsistence represented a relevant reason for the households' decision to operate in the informal sector; 3) Informal activities supply a "safety valve" within the surviving strategies adopted by the poorest households; 4) Participation in the informal sector seems to not be simply correlated with poverty: in the informal activities are involved poor people (having probably low level of education), but also relatively rich persons. However, their motivations are quite different. The former are practically "forced" to operate in informal sector (the "subsistence" criterion), but the latter are "invited" to participate in it (the "enterprise" criterion). In both cases, at least during the first stages of transition to a free market system in Romania, the environment was propitious due to legislative incoherence, feeble penalty system in the cases of fraudulent activities, and existence of some accompanying elements of proper informal activity, such as corruption, bureaucracy, etc. Moreover, the behaviour related to the informal economy is sometimes fundamentally different between the two groups of population. The most synthetic expression of this idea could be as follows: along with their formal income growth, the households tend to wish to obtain more and more informal income in absolute terms, but in the same time its share in the total income tends to decrease (sharply down until a reasonable average level of formal income is obtained and slowly down in the case of the richest households). Probably, the main reason for which the rich people could be involved in the informal sector is provided by the attempt to avoid taxes and to follow an optimising strategy in this matter.

Based on database resulted from two special surveys organised in Romania (1996 and 2003, respectively) permitted us to estimated the parameters for the correlation between formal income in household and its participation in informal activities. Omitting other details regarding the complicate methodologies and procedures that we used to obtain a set of data on informal income in households, we report here only the final results and the strategy that we use now in order to extent them from households' population in survey to the entire population at regional and national levels.

Certain behavioural regimes were outlined in matter of potential implication in informal sector. Thus, in the case of poor households (obtaining relative low income from their formal activity) there is a large availability to work in the informal sector. On the other hand, in the case of rich households (obtaining relative large income from their work in the formal sector), their availability for informal jobs becomes smaller; however still remain temptation for the richest people to accept informal jobs in order to supplement their income or, perhaps, to avoid taxation. Despite general decreasing trend of the share of desired informal income along with the growth of the basic formal income of household, in absolute terms the desired informal income has an ascending trend.

In similar way to the case of desired informal income, we used the hypothesis of a hyperbolic-type function for $z\%(v)$. This time z means the effective informal income, being different from the desired informal income. Thus, in order to estimate the coefficients we selected as basic regression equation that expressing the share of informal income in the total household's income, $z\%$, as being correlated with the level of the actual formal income, v, in household, as follows:

$$z\% = a / (v + b) + (1 - a/b) + u \qquad (36)$$

where a, b are coefficients to be estimated and u is residual variance.

13

Then, using the estimated values of coefficients we can write, along with changes in the level of formal income, the expected trajectories, as follows (see for details, Albu, 2004):

$$ze\% = a / (v + b) + (1 - a/b) \tag{36'}$$

$$ze(v) = [(b - a) / a] \cdot v + (b^2 / a), \text{ with } ze(0) = (b^2/a) \tag{37}$$

In order to estimate the real level for informal income, according to the available data from survey of 1996, we used two sub-samples, noted as A and B. In the case of the sample A the function of informal income share reflects indirectly the impact of changing the proportion of households operating in the informal sector (or, equivalently: the impact of changing the probability that a household is involved in the informal sector) along with the growth of the formal income per person in household. Consequently, it could be used directly to expand the estimation procedure to the regional and national levels. An impediment remains: it is implicitly supposed the same distribution of the entire population by formal income as in the case of the A sample. On the other hand, within sample A there is the sub-sample B comprising only the households obtaining informal income. In this case, to simply extrapolate the $z\%(v)$ function to the entire set of households' population is not a good solution (it is the case of the so-called hypothesis of a generalized informal economy). Thus, we have to amend the $z\%(v)$ function by multiplying it with the probability function computed by deciles of formal income.

As a first step, we amended the last estimating equation by adding a supplementary equation concerning the probability that a person in a household is involved in the informal economy. This was estimated by regressing within the sample A the proportion of persons in households obtaining effectively informal income in the total number of deciles of formal income in which they are located (the total number of this special category of household is just the sub-sample B):

$$p = a \cdot d + b + u \tag{38}$$

and from which the equation (38) was rewritten as

$$zpe(v) = ze(v) \cdot pe(d) \tag{39}$$

where d are deciles (d=1…10); pe(d)=ad+b is the estimation equation of the probability that a person in a household is involved in the informal economy, p; a and b are the estimated coefficients, and u is residual variance in the equation (38). The estimation procedure (39) is noted as C.

Then, we extended the three estimation procedures, A, B, and C to the regional and national level over the period 2000-2005. In order to conserve estimated values for coefficients in case of extending the model to the regional and national level, all data were expressed in constant prices of September 1996 (as in the original data of the 1996 survey).

Synthetically, the conclusion is that over the period 2000-2005 the informal income decreased in Romania from 22.3-22.8% in the total income of households to 17.2-18.3%, as we can see in Table 2. Under the very improbable hypothesis of a generalised participation in informal activities (in theoretical case when all household are involved in informal activities, as in case of the sub-sample B), the computed share decreased from 33.6% in 2000 to 27.9% in 2005. The main factor of this favourable dynamics of informal income was the formal income growth (+42%, from about 158300 to 225600 Lei/person/month, computed in September 1996 prices).

14

More interesting conclusions could be extracted in the case of analysing by regions the dynamic process of involvement in the informal sector. In Appendix 1 are presented the three matrixes comprising the shares of informal income within the total income in the case of the eight regions of Romania for each year of the period 2000-2005, corresponding to the three estimating methods. In Appendix 2 is presented the contribution of regions to the total informal income at national level, also corresponding to the three methods.

Figures 6 and 7 show the estimated dynamics of the average share of informal income in total income at the national level, based on the two estimation procedures, A and C, over the period 2000-2005 (the year 2000 is denoted as 0 and 2005 as 5), and its relatively strong inverse correlation with the distribution of formal income grouped by regions, respectively (regions are noted as $i=1\ldots8$, and years as $j=0\ldots5$). z%M represents the yearly average share of the informal income in the total income at national level, resulted from the regression equation based on the procedure A (sample A) and zp%M from that based on the procedure C (applying the regression equation on sub-sample B amended by the probability function), respectively.
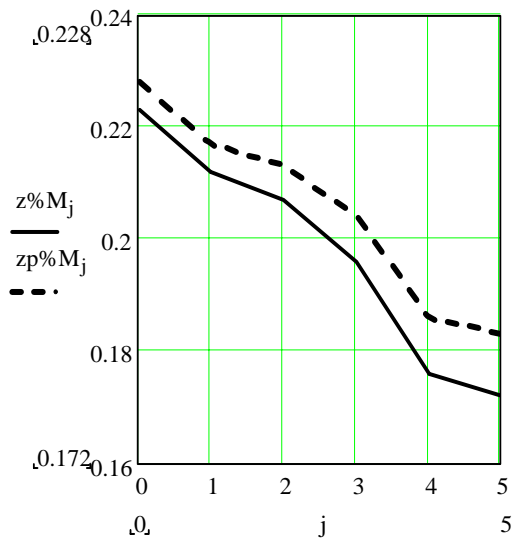

Figure 7.

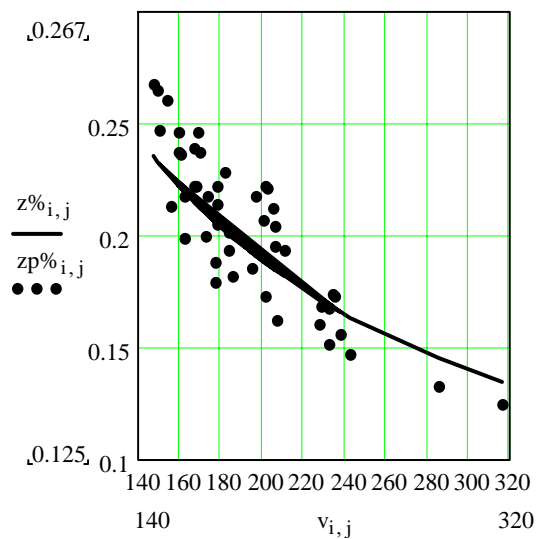**Figure 8.**

Table 2. Average shares of informal income in the total income of households over the period 2000-2005

Years	z%M	zp%M
2000	22.3	22.8
2001	21.2	21.7
2002	20.7	21.3
2003	19.6	20.4
2004	17.6	18.6
2005	17.2	18.3

15

References

Albu, L.-L. (1995): "Underground Economy and Fiscal Policies Modeling", *ACE-Phare Project,* Université de Paris I, Centre d'Etudes Prospectives d'Economie Mathematique Appliquées à la Planification, August.

Albu, L.-L. (2004): "Estimating the Size of Underground Economy in Romania", *Tax Evasion, Underground Economy and Fiscal Policies in Candidate Countries (Case of Romania)*, GDN Project, Prague, June.

Albu, L.-L., Daianu, D., Pauna, B., and Pavelescu, F. (1998): "Endogenous Cycles and Underground Economy in Europe", *MEET- IV conference on the East European Economies in Transition*, University of Leicester – CEES, 20-21 June.

Albu, L.-L., Kim, B.-Y., and Duchene, G. (2002a): An Attempt to Estimate the Size of Informal Economy Based on Household Behaviour Modeling, *Romanian Journal of Economic Forecasting, Vol. 1 (9)*, Bucharest.

Albu, L.-L., Kim, B.-Y., and Duchene, G. (2002b): Households' Activities in Informal Economy: Size and Behavioural Aspects, *Romanian Journal of Economic Forecasting, Vol. 3-4 (11-12)*, Bucharest.

Albu, L.-L. and Nicolae, M. (2002): Use of Households Survey Data to Estimate the Size of the Informal Economy in Romania, *The Informal Economy in the EU Accession Countries: Size, Scope, Trends and Challenges to the Process of EU Enlargement*, Network for Integration of Central and Eastern-European Countries into the European Union, March, Sofia.

Allingham, M.G. and Sandmo, A. (1972): Income Tax Evasion: A Theoretical Analysis, *Journal of Public Economic*s, November, 1(3-4).

Blades, D. (1982): "The Hidden Economy and the National Accounts", *OECD (Occasional Studies)*, Paris, pp. 28-44.

Clotfelter, C.T. (1983): Tax Evasion and Tax Rates: An Analysis of Individual Returns, *Review of Economics and Statistic*s, Vol. 65.

Daianu, D. and Albu, L.-L. (1997): "Institutions, Strain and the Underground Economy". International Conference on the Importance of the Underground Economy in Economic Transition, University of Zagreb. *The Davidson Institute Working Paper Series*, 98.

Duchene, G. (coordonator), Adair, P., Albu, L.-L., Ivan-Ungureanu, C., Neff, R., and Tanase, F. (1998): "Informal economy in Romania", *ACE-Phare Project*, ROSES, Paris, September.

French, R., Balaita, M. and Ticsa M. (1999): "Estimating The Size And Policy Implications Of The Underground Economy In Romania", US Department of the Treasury, Office of Technical Assistance, Bucharest, August.

Gibson, B. and Kelley, B. (1994): A Classical Theory of the Informal Sector, *The Manchester School*, Vol. LXII, 1.

Gutmann, P. M. (1977): The Subterranean Economy*, Financial Analysts Journal*, 34:1, pp. 24-27.

Isachsen, A.J., and Strom, S. (1985): The size and growth of the hidden economy in Norway, *Review of Income and Wealth*, Series 31. No.1.

Lemieux, T., Fortin, B., and Fréchette, P. (1994): "The Effect of Taxes on Labor Supply in the Underground Economy", *The American Economic Revie*w, March, 84(1).

Pencavel, J.H. (1979), A Note on Income Tax Evasion, Labour Supply, and Non-linear Tax Schedules, *Journal of Public Economics*, Vol. 12.

Pestieau, P. (1989): *L' Economie souterrain*e, Hachette, Pluriel.

Schneider, Friedrich (2002): The Size and Development of the Shadow Economies and Shadow Economy Labor Force of 22 Transition and 21 OECD Countries: What Do We Really Know?, *The Informal Economy in the EU Accession Countries: Size, Scope, Trends and Challenges to the Process of EU Enlargement*, Network for Integration of Central and Eastern-European Countries into the European Union, March, Sofia.

*** National Institute of Statistics (1996-2006): Population income, expenditure and consumption, *Romanian Statistical Yearbook 1996-2006*, Bucharest.

Shares of informal income in total income by regions

H1 Estimations under the hypotheses of procedure A
(regression equation on sample A)

Years		**2000**	**2001**	**2002**	**2003**	**2004**	**2005**
Regions	**R1**	0.235	0.233	0.228	0.214	0.187	0.188
	R2	0.222	0.215	0.212	0.202	0.191	0.185
	R3	0.232	0.221	0.222	0.205	0.188	0.184
	R4	0.215	0.209	0.215	0.205	0.184	0.182
	R5	0.219	0.205	0.201	0.192	0.168	0.167
	R6	0.226	0.210	0.201	0.192	0.169	0.171
	R7	0.220	0.206	0.199	0.188	0.171	0.166
	R8	0.206	0.184	0.169	0.163	0.145	0.134
Average		**0.223**	**0.212**	**0.207**	**0.196**	**0.176**	**0.172**

H2 Estimations under the hypotheses of procedure C
(regression equation on sub-sample B amended by the regression
equation of probability sub-sample B in sample A)

Years		**2000**	**2001**	**2002**	**2003**	**2004**	**2005**
Regions	**R1**	0.267	0.265	0.260	0.246	0.221	0.222
	R2	0.246	0.239	0.237	0.228	0.217	0.212
	R3	0.247	0.236	0.237	0.222	0.207	0.204
	R4	0.222	0.217	0.222	0.214	0.195	0.193
	R5	0.217	0.205	0.201	0.194	0.174	0.173
	R6	0.213	0.200	0.193	0.185	0.167	0.168
	R7	0.199	0.188	0.182	0.173	0.160	0.156
	R8	0.179	0.162	0.151	0.147	0.133	0.125
Average		**0.228**	**0.217**	**0.213**	**0.204**	**0.186**	**0.183**

H3 Estimations under the hypothesis of procedure B
(a generalized informal economy based on the equation of regression
used in case of sub-sample B)

Years		**2000**	**2001**	**2002**	**2003**	**2004**	**2005**
Regions	**R1**	0.349	0.347	0.342	0.326	0.296	0.297
	R2	0.336	0.327	0.324	0.313	0.300	0.294
	R3	0.347	0.334	0.335	0.317	0.297	0.293
	R4	0.327	0.321	0.327	0.317	0.293	0.290
	R5	0.332	0.316	0.311	0.302	0.274	0.273
	R6	0.339	0.321	0.312	0.301	0.275	0.277
	R7	0.333	0.317	0.309	0.297	0.278	0.272
	R8	0.317	0.292	0.275	0.269	0.247	0.234
Average		**0.336**	**0.324**	**0.318**	**0.306**	**0.283**	**0.279**

Shares of informal income in total income by years

H1 Estimations under the hypotheses of procedure A
(regression equation on sample A)

Years		**2000**	**2001**	**2002**	**2003**	**2004**	**2005**
Regions	**R1**	0.170	0.169	0.170	0.171	0.171	0.171
	R2	0.131	0.131	0.131	0.131	0.130	0.130
	R3	0.154	0.154	0.154	0.154	0.153	0.153
	R4	0.107	0.107	0.107	0.107	0.106	0.106
	R5	0.091	0.091	0.090	0.090	0.090	0.090
	R6	0.127	0.127	0.127	0.127	0.127	0.127
	R7	0.118	0.118	0.117	0.118	0.118	0.118
	R8	0.103	0.103	0.104	0.104	0.105	0.106
Total		**1.000**	**1.000**	**1.000**	**1.000**	**1.000**	**1.000**

H2 Estimations under the hypotheses of procedure C
(regression equation on sub-sample B amended by the regression
equation of probability sub-sample B in sample A)

Years		**2000**	**2001**	**2002**	**2003**	**2004**	**2005**
Regions	**R1**	0.196	0.194	0.194	0.195	0.197	0.196
	R2	0.146	0.145	0.145	0.145	0.143	0.143
	R3	0.162	0.162	0.161	0.162	0.16	0.160
	R4	0.109	0.108	0.107	0.107	0.106	0.106
	R5	0.088	0.088	0.087	0.086	0.087	0.087
	R6	0.115	0.115	0.116	0.116	0.117	0.116
	R7	0.101	0.102	0.101	0.102	0.101	0.102
	R8	0.084	0.086	0.087	0.087	0.089	0.091
Total		**1.000**	**1.000**	**1.000**	**1.000**	**1.000**	**1.000**

H3 Estimations under the hypothesis of procedure B
(a generalized informal economy based on the equation of regression
used in case of sub-sample B)

Years		**2000**	**2001**	**2002**	**2003**	**2004**	**2005**
Regions	**R1**	0.168	0.167	0.167	0.168	0.169	0.168
	R2	0.131	0.130	0.131	0.130	0.128	0.128
	R3	0.153	0.152	0.152	0.152	0.151	0.150
	R4	0.108	0.107	0.106	0.106	0.105	0.105
	R5	0.091	0.092	0.091	0.090	0.091	0.090
	R6	0.126	0.127	0.128	0.127	0.128	0.127
	R7	0.118	0.119	0.118	0.119	0.118	0.119
	R8	0.104	0.106	0.109	0.108	0.110	0.113
Total		**1.000**	**1.000**	**1.000**	**1.000**	**1.000**	**1.000**