



Munich Personal RePEc Archive

## **An attempt to estimate the size of informal economy based on household behaviour modeling**

Albu, Lucian-Liviu and Kim, Byung-Yeon and Duchene,  
Gerard

31 March 2002

Online at <https://mpra.ub.uni-muenchen.de/13026/>  
MPRA Paper No. 13026, posted 28 Jan 2009 03:24 UTC

# **An Attempt to Estimate the Size of Informal Economy Based on Household Behaviour Modeling**

Lucian-Liviu ALBU  
(Institute for Economic Forecasting)

Byung-Yeon KIM  
(University of Essex)

Gerard DUCHENE  
(University of Paris I, ROSES)

## *Abstract*

Having in view the three main methods for the estimation of the informal economy's size, the study focuses in the first part on data and methodological problems, trying at the same time to outline a few behavioral aspects by using only simple simulation models. Thus, based on a 288-sample and using different types of computations, the authors describe the households' participation in informal activities as varying between 20.4% and 30.6%. In the second part of the article, some more sophisticated dynamic models to simulate very complicated types of household behavior relating to participation in secondary activities and in informal economy are presented.

JEL Classification: C51, H31, O17

Keywords: Informal Economy, Estimation, Household's Behaviour

Note: Paper prepared on the basis of research within two Phare ACE Programmes 1996 and 1998, respectively ("The informal economy in Romania" and "Household informal economy activities in candidate countries: size, determinants and implications for the enlargement - empirical evidence from Romania and Bulgaria"). The views expressed here do not necessarily reflect the European Union's view.

## 1. Introduction

Broadly speaking, there are three methods for estimation of the size of the informal economy: time-series analysis based on cash demand; estimates based on the discrepancy between total incomes and total expenditures at the aggregate level; discrepancies between income and expenditure at the microeconomic level (Smith, 1986, Thomas 1992). Among other things, Smith (1986) points out that the overall accuracy of the national accounts method is lower than the other two due to the inclusion of various errors in both the income and expenditure measures of GDP. The lack of reliable historical data before transition and the possible structural break between pre and post transition suggest that time-series method is not feasible to a large extent. The only remaining alternative is, therefore, to analyse the individual household data. Furthermore, the results of the analysis on the basis of micro-data might provide more significant information for policy-making because they, unlike those using aggregate data, can highlight the main participants in the informal economy and the effects on welfare/behaviour of the households.

The study focuses in a first part on data and methodological problems trying at the same time to outline a few behavioural aspects by using only simple simulation models. In the second part of article, we present some more sophisticated dynamic models, which will simulate very complicated types of household behaviour relating to participation in secondary activities and in informal economy.

## 2. Data and methodological aspects

The so-called Integrated Household Survey (IHS), comprising a sample of approximately 36,000 observation units from about 500 urban and rural research areas, is the main source of information in order to study the household behaviour. Also, in 1996, in Romania a Supplementary Survey on Household Informal Economy Activities was conducted in September (Duchene et al., 1998). The Supplementary Survey, which used a sample size of around 2,600 households, was focused on informal economy activities carried out by households. The survey was divided into 21 sub-sections that contained detailed questions, but indirectly formulated by answering means, about informal economy. It was essential for our work to correlate the two sources of data. The survey asked about the ratio of income from main activity to that from secondary activity. Using the information, we obtained an absolute measure of households' income from the secondary activity.

Based on answers provided by the question in which all members of household compared their two incomes (from main activity and from declared secondary activity, respectively), we computed a composite coefficient ( $k_s$ ) for every household in the sample, in order to characterise the share of the two types of activity.

$$V_s = k_s \cdot V \quad (1)$$

where  $V_s$  is income from second activity,  $k_s$  is the share of income from declared secondary activities in total declared income, and  $V$  is the total declared income.

Also, the income corresponding to the main activity ( $V_b$ ) was obtained as follows:

$$V_b = V - V_s \quad (2)$$

Using (1) and (2), we rewrite the shares of the two components in the total declared income of a household:

$$k_s = V_s / (V_s + V_b) \quad (3)$$

and

$$k_b = V_b / (V_s + V_b), \text{ respectively} \quad (3')$$

where  $k_b$  is the share of income from the main job in the total declared income.

One important result was also obtained by comparing the so-called decent (or desired) income with the actual size of income. So, in order to find the size of informal (or hidden) economy we computed the difference between the two types of income:

$$V_a = H - V \quad (4)$$

where  $V_a$  is hidden (informal) income,  $H$  is decent income (or the maximum level of desired income) and  $V$  is actual total declared income ( $V = V_s + V_b$ ).

The computing outputs obtained by grouping data conforming to the last criterion are presented synthetically in Table 1, where the desired (decent) level of income,  $H$ , was replaced by  $H^*$ , which means that data were adjusted in order to solve some logical incoherence existing in the initial answering data (for instance, in case when a qualitative answer relating to the report between actual income and desired income was  $V < H$ , but from other answers on the precise levels of actual income and respectively desired income resulted an opposite situation,  $V > H$ , the respective household was moved into this last group, moreover denoting the new obtained classification by  $H^*$ ). As we can see from computed data, indeed in case of the richest group ( $V > H^*$ ), where the average of declared actual income is 311.8 thousands Lei/person, there is no informal activity. A different situation is registered in the case of poorest group of households ( $V < H^*$ ), where the difference between the considered decent income and actual declared income is huge one. We could interpret this difference as a measure of “potential supply of informal activity” by people. But in reality this “supply” probably is not covered by a corresponding demand coming from the real economy only in a smaller proportion. To discover the real dimension of this proportion in case of the poorest group ( $V < H$ ) continues to be a challenge for economists and statisticians from everywhere.

**Table 1****Households income grouped by the ratio of V to H, in 1996**

- thou. Lei/person -

	Number of households	Declared income		Desired income H*	Potential informal activity H* - (Vb + Vs)
		Main activity Vb	Second activity Vs		
V<H*	2181	127.0	9.8	368.2	231.5
V=H*	288	263.6	20.1	384.3	100.6
V>H*	92	258.0	53.8	311.8	0.0
Total sample	2561	146.6	12.6	367.7	208.5

**3. Estimation of hidden economy in the case of 288-sample**

In order to capture the size of hidden income, we computed the differences between the two levels of income in the case of households declaring that they are the same, so the group V=H (the 288-sample). However, a problem, which was also solved when we computed the adjusted data, was in the case when the desired (decent) income in Supplementary Survey declared by a household was smaller than the same indicator reported in IHS. So, in Table 2 there are data on both corresponding situations: initial data and adjusted data, respectively (when the difference between declared decent income and actual income was negative, it was replaced with zero, considering that in the case of the 288-sample the decent income must be at least equal to the actual declared income in IHS).

Corresponding to the two considered hypotheses (initial data and respectively adjusted data), as it is shown in Table 2, at the level of the whole 288-sample, the share of probable informal (hidden) activity,  $V_a/H$ , was around 20.4% and respectively 26.2%, on average. In other words, the composition of total income, H, by sources was, in 1996, in the case of initial data: 74.0% main activity, 5.6% second declared activity, and 20.4% informal activity. In the case of adjusted data, the structure of the total households' income, H\*, was: 68.6% main activity, 5.2% second activity, and 26.2% informal activity. More analytic conclusions were provided by a deeper analysis in which the two conventional sectors, SI (households operating only in one activity, main or basic activity), and SII (households operating in more than one activity, main activity and secondary activities), were compared.

So, the last estimated level of informal activity, 26.2% of total income, H\*, may be used as first estimation in order to obtain parameters in a more general regression equation and to find the behaviour of household. A very severe restriction, which occurs when we wish to extrapolate some conclusions in order to capture the general household behaviour, is brought about by the asymmetry between the 288-sample and the whole sample including more than 2500 households. The best general fitting function to estimate the household behaviour seems to be one expressing a complex inverse relation between the average level of income provided by main activity and participation rate in informal activity.

**Table 2**

**Estimated shares of hidden income based on the 288-sample**

- thou. Lei/person -

	Initial data				Adjusted data			
	Declared Income		Hidden Income	Total	Declared Income		Hidden Income	Total
	Main activity Vb	Second activity Vs	Va	H	Main activity Vb	Second activity Vs	Va	H*
SI	162.9	-	35.9	198.8	162.9	-	52.6	215.5
Average	340.1	-	74.9	415.0	340.1	-	109.8	449.9
SII	44.3	15.8	21.1	81.2	44.3	15.8	26.5	86.6
Average	144.3	51.5	68.8	264.6	144.3	51.5	86.2	282.0
Total (SI+SII)	207.2	15.8	57.0	280.0	207.2	15.8	79.1	302.1
Average	263.6	20.1	72.5	356.2	263.6	20.1	100.6	384.3

SI – Households operating only in one activity (main or basic activity, conforming to the definitions included in the Supplementary Survey questionnaire).

SII - Households operating in more than one activity (main activity and secondary activities, conforming to the same definitions comprised in Supplementary Survey).

The first and second rows in each category refer to total household income of the sample (in million Lei) and to the average level of income per person (in thousand Lei), respectively.

However, based on the 288-sample, one may affirm that, at an average level of income provided by the main activity amounting to 263600 lei/person, some people are forced to work in a secondary job and also in informal sector in order to obtain supplementary income for their families (20100 lei/person from secondary job and respectively 72500 lei/person from informal activity). Moreover, at an average monthly income of only 144300 lei/person, people are forced to be relatively more involved both in secondary activities and in informal activities (in order to add supplementary incomes of 51500 lei/person and 68800 lei/person, respectively). So, in the last case, the structure of final income is: 51.2% main activity, 18.3% secondary declared activities, and 30.6% informal activities. However, the situation is quite different when the basic income rises. For instance, at an income obtained from the main activity of 340100 lei/person (average level of SI in Table 2), there were supplementary opportunities for people to work in the informal sector in a smaller proportion, 24.4% of their total income.

Indeed, besides the level of income provided by their main activity, the households' informal activities are probably affected by occupation, region, age, education and many other factors. However, at this stage of our investigation, a few conclusions could be outlined:

- Taxation is perceived by people as the main cause of underground activity.
- Separating the main motivations of operating in informal sector in two groups – “subsistence” one and “enterprise” one, respectively – the Supplementary Survey suggests

that, at least in 1996, the subsistence represented a relevant reason for the households' decision to operate in informal economy, including its underground segment.

- Informal activities supplied a “safety valve” within the surviving strategies adopted by the poorest households.

- Participation in informal economy seems not to be simply correlated with poverty: in informal economy poor people (having probably a small level of instruction) as well as relatively rich persons are involved. But their motivations are quite different. The former are practically “forced” to operate in informal economy (“subsistence” criterion), but the latter are “invited” to participate in it (“enterprise” criterion). In both cases, at least during the last stages of transition, the environment was propitious owing to legislative incoherence, feeble penalty system in cases of fraudulent activities, and existence of some accompanying elements of proper informal activity, such as corruption, bureaucracy, etc. However, the behaviour towards the informal economy is sometimes fundamentally different between the two groups of population. This is why we consider that deeper investigations focused on the behavioural aspects of different groups of population will be further needed in the future.

#### 4. Empirical analysis

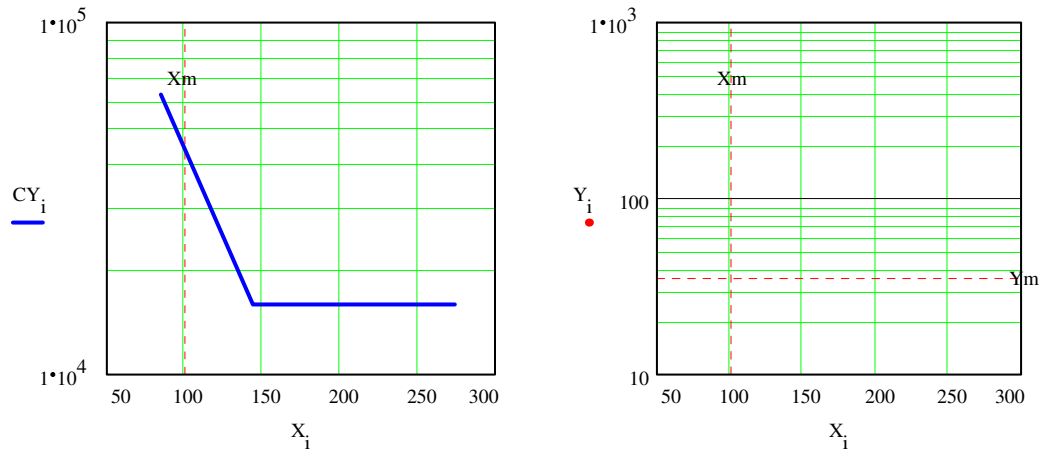
In order to systematically study the households' behaviour, as a first stage of the research, we consider some simple empirical analyses. So, to identify the type of diverse relationships between components of the total income, it may be useful to see the simple graphical representations of such relations based on raw registered data.

First, considering only the average levels, we identified the general types of demand-curves and supply-curves in the case of secondary activities and informal activities, respectively, as they are shown in Figures 1a and 1b and in Figures 2a and 2b, respectively. The samples we used were obtained from the entire SII-sample (931 households) sub-grouped by the criterion of the ratio of  $V$  and  $H^*$ . In the figures, the denotations represent:  $CY$  and  $CZD$  – total demand for  $V_s$  and total desired demand for  $V_a$ , respectively (it is the already mentioned “potential informal activity” in Table 1);  $Y=V_s$ ,  $X=V_b$ , and  $Z=V_a$ ;  $OY$  and  $OZ$  – supply for  $V_s$  and for  $V_a$ , respectively. Also, we mention that on abscise,  $X$  may be interpreted as the level of qualification, on which the demand of real economy is established. So, it must be noted that in the case of informal activity the demand is only that desired by people but not the demand coming from the real economy. In Figures 1b and 2b, the dotted curves represent the real supply and their intersection with the demand curves means even the registered (empiric) data.

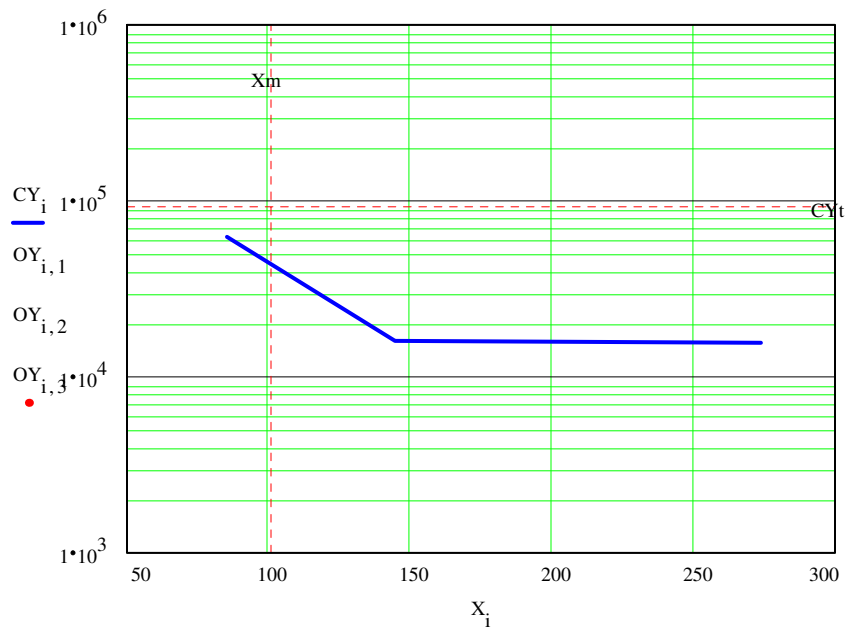
Also, Figure 3a shows the distribution-maps of shares of income components ( $x=V_b/H^*$ ,  $y=V_s/H^*$ , and  $z=V_a/H^*$ ) against the level of income supplied by the main activity, in the case of 288-sample and Figure 3b, some equivalent 3-D graphs. This pictures suggest the existence of certain inverse correlations  $y-X$  and  $z-X$  and of a direct correlation  $x-X$ , respectively. Further, we shall analyse in a more detailed way the relationships between  $y$

and X and between z and X, respectively, by using as descriptive functions the linear ones and the hyperbolic ones.

**Figure 1a**

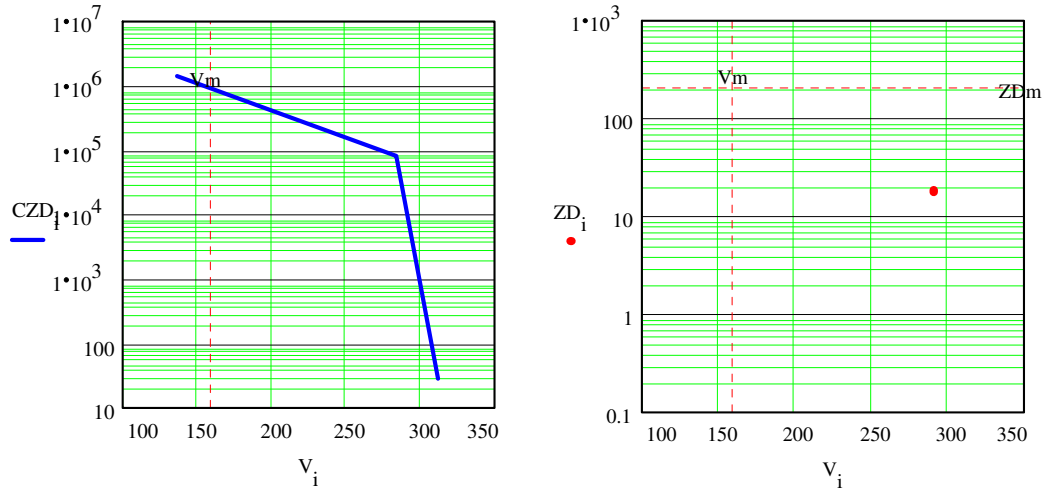


**Figure 1b**

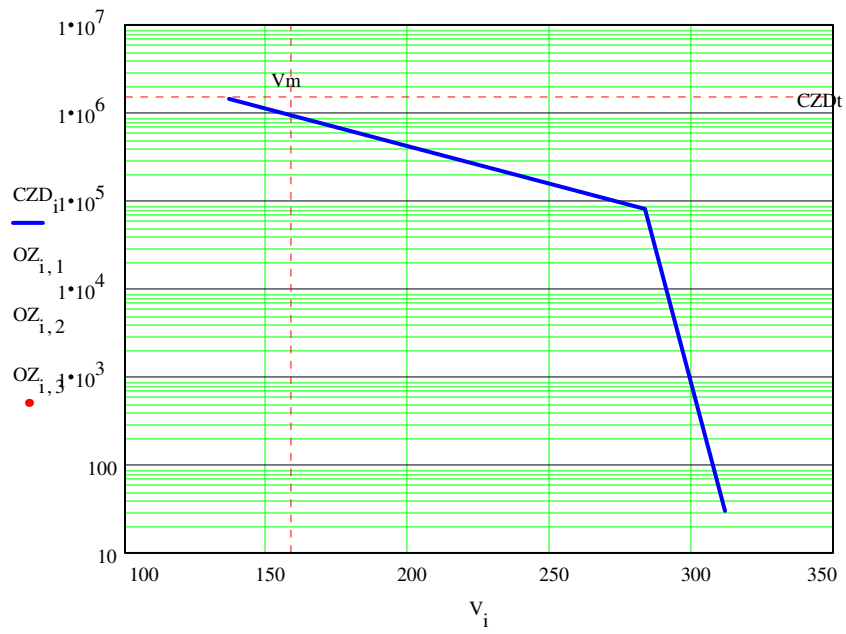




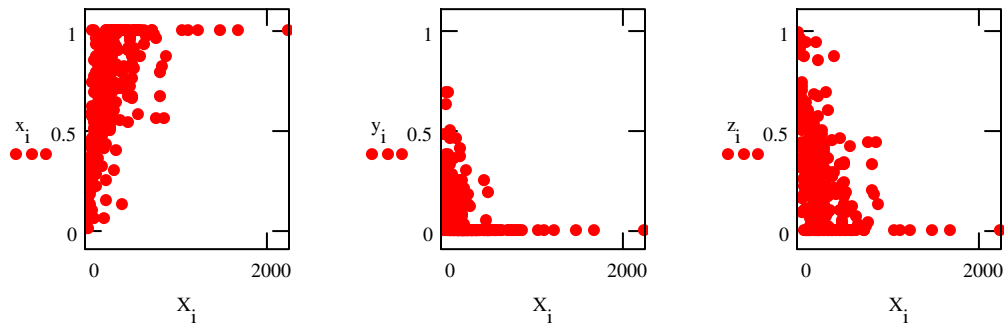
**Figure 2a**



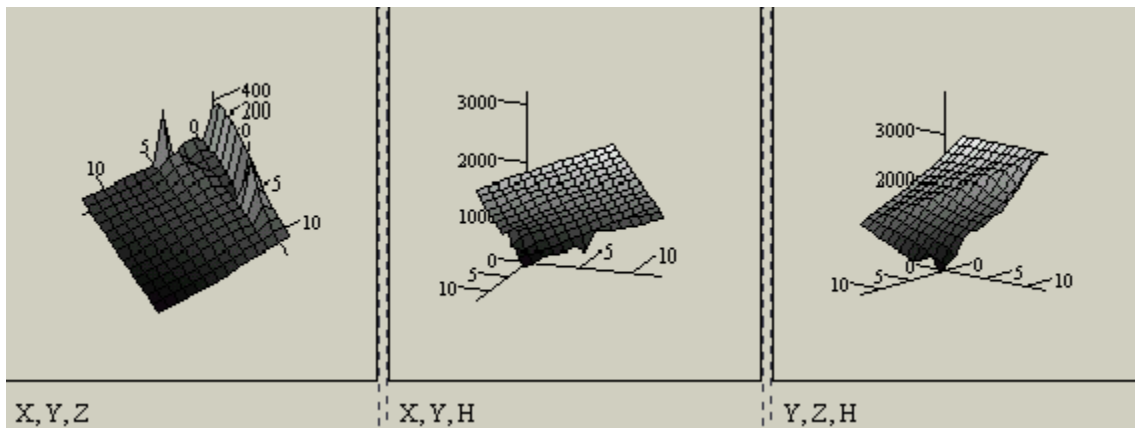
**Figure 2b**



**Figure 3a**



**Figure 3b**



## References

Albu, L., Tarhoaca, C., Croitoru, C., Daianu, D., Dobrescu, D., Doltu, C., Ivan-Ungureanu, C., Lazea, V. (2001), "Underground Economy in Romania", CERPE, USAID, IRIS, Bucharest.

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

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.

Duchene, G., (1999), "Les revenus informels en Roumanie. Estimation par enquête", in *Revue d'études comparatives Est-Ouest*, vol. 30, no. 4, p. 35-64.

Smith, S. (1986), *Britain's Shadow Economy*, Clarendon Press, Oxford.

Thomas, J., (1992), *Informal Economic Activity*, Harvester Wheatsheaf, London.

\*\*\* Institutul National de Statistica (2001), "Veniturile, cheltuielile si consumul populatiei in anul 2000", Bucharest.