The distributional implications of income underreporting in Hungary

Benedek, Dora and Lelkes, Orsolya
Central European University , Ministry of Finance, Hungary, European Centre for Social Welfare Policy and Research

09. September 2009

THE DISTRIBUTIONAL IMPLICATIONS OF INCOME UNDERREPORTING IN HUNGARY*

Dóra Benedek[+]

Orsolya Lelkes[++]

9 September 2009

Abstract

The paper estimates the distributional implications of income tax evasion in Hungary based on a random sample of administrative tax records of 230 thousand individuals. Gross incomes in the administrative tax records are compared with those in a nationally representative household budget survey, assuming that tax-evaders are more likely to report their true incomes in an anonymous interview. Our estimates show that the average rate of underreporting is 11%, which conceals large differences between self-employed (who hide the majority of their incomes) and employees. The estimates are likely to be lower bound, due to measurement error in the income survey. These rates are then used in EUROMOD, a tax-benefit microsimulation model to calculate the fiscal and distributional implications of underreporting, while taking account of all major direct taxes and cash benefits and also their interactions. Tax evasion reduces fiscal revenues from personal income taxes by about 19%. While the occurrence of poverty is not affected, income inequality becomes significantly higher (the Gini coefficient increases by 7%), suggesting that high earners tend to evade proportionately more. Finally, we find that tax evasion largely reduces the progressivity of the tax system.

JEL Codes: C8, D31, H26, H22, H21, I38
Keywords: tax policy, tax evasion, income distribution, self-employed

---

* We are grateful for helpful discussions with Zsombor Cseres-Gergely, István János Tóth, Manos Matsaganis, Maria Flevotomou, Daniela Mantovani, Panos Tsakloglou, Holly Sutherland and participants of the AIM-AP project. We thank Christine Lietz for her efficient help with EUROMOD. Dóra Benedek is thankful for the APEH, the Hungarian tax authority, for the provision of the random sample of tax records for research purposes of the Ministry of Finance. We gratefully acknowledge financial support from the European Commission under the AIM-AP (Accurate Income Measurement for the Assessment of Public Policies) project.

+ Corresponding author, Central European University and Ministry of Finance, Hungary, email: cphbed01@ceu-budapest.edu, address: CEU Economics Department, H-1051, Budapest, Nádor u. 9.
++ European Centre, Vienna and CASE, LSE, London, email: lelkes@euro.centre.org

1 Introduction

Tax evasion obviously interferes with the evaluation of tax policy changes. In addition, where the practice of assessing eligibility to benefits relies on a scrutiny of tax returns, tax evasion will also cause target inefficiency in the form of benefit "leakage" to ineligible recipients. For these reasons, ignoring tax evasion can seriously misjudge the distributive and fiscal effect of changes in social benefits and the tax system. In this light, we aim to explore a procedure to correct income data for tax evasion.

The aim of the paper is to provide preliminary estimates of the size and distribution of income underreporting and thus tax evasion in Hungary. The paper takes advantage of access to a random sample of income tax returns reporting on incomes earned in 2005, containing information on 227.688 individuals, about 5,4% of total taxpayers, provided to one of the authors by the tax authority.

The level of income taxes is high in Hungary, but a large proportion of the population does not pay these in full. The bottom 80% of taxpayers pays altogether only 20% of the personal income tax (Krekó and P. Kiss 2007, p. 26). In order to design a fair income redistributive system, policy makers need to know not only the incomes of the individuals, but also how they actually comply with the tax regulations. Currently, very little is known on the latter. This paper uses administrative tax records, thus provides novel results on the extent and distribution of income tax evasion.

Tax compliance may be low for two reasons in the country. The first is an attitudinal reason: low trust in the political system, little public awareness on the cost of public services and the budget policy as such (e.g. Csontos et al. 1998), and the low perceived quality of public services (Hanousek and Palda 2004).

We find that the average rate of underreporting is 11% on average, including around 65% of the self-employed and 3% of employees. However these estimates are lower bound, due to measurement error in the income survey. Tax evasion reduces government revenues from personal income taxes by about 19% and significantly increases income inequality and reduces the progressivity of the tax system.

The paper is organized as follows. Section 2 briefly reviews the literature on tax evasion in Hungary. Section 3 explains the main features of tax policy, with special focus on entrepreneurs. Section 4 presents the data and section 5 the main results. Last, section 6 concludes.

2 Literature

Why do people evade taxes? Or rather, why do people pay taxes at all? While classical theory (Allingham and Sandmo 1972) was preoccupied with the former question, recently the latter one has received increasing attention (see e.g. the overview of Andreoni et al. 1998). There is a great discrepancy between the standard theoretical model and actual compliance, the former predicting greater non-compliance than observable in the real world. Given that audit probability is rather low in general, Expected Utility Theory (using audit probability, penalty rate and expected return on evading the tax in the model) predicts the expected return on tax evasion between 91 and 98%, implying that *all* taxpayers should hide some income (Dhami

and al-Nowaihi 2007). Psychological factors or moral preferences, including loss aversion, stigma, guilt, shame or sense of duty, might explain why it is not the case.

A growing strand of authors criticize the traditional neoclassical model of tax compliance, where the tax-payer is treated as an isolated and amoral individual and present evidence that tax evasion is an interdependent decision, and is greatly influenced by social norms and social interactions (Fortin et al. 2007; Frey and Torgler 2007). If taxpayers believe tax evasion to be common, tax evasion decreases.

The institutional framework of tax compliance also merits more attention, including the structure, the functioning of the tax authority, its setting within the government, and its dynamic, repetitive interaction with the taxpayers (Andreoni et al 1998). The quality of political institutions (accountability, political stability, government effectiveness, regulatory quality, rule of law, control of corruption) was shown to have positive observable effect on tax morale (Frey and Torgler 2007). Entrepreneurs were found to go underground not to avoid taxes but to reduce the burden of bureaucracy and corruption (Friedman et al 2000).

There are three main methods of measuring tax compliance, which may be called indirect approach, direct approach and the modelling approach (Schneider and Enste 2000; Frey and Schneider 2001). The indirect approaches are mostly based on macroeconomic data. Possible indicators include the discrepancy between national expenditure and income statistics, the discrepancy between official and actual labour force, the currency demand approach and the electricity consumption method. Direct methods, on the other hand, are based on micro datasets. They could include specific surveys or tax auditing data. The modelling approach, developed by Frey, focuses on the causes and effects of the undergrounds economy, and builds on a behavioural model. This requires a large amount of data, which are often not available. In our paper, we are using a direct method.

Direct methods seem to be better suited for Eastern European countries. Hanousek and Palda (2006) express a rather critical view about macro approaches measuring tax evasion in transition economies (e.g. money demand and electricity demand equations), and plea for the use of direct methods. They argue that the pace of change of the parameters of the money demand and electricity demand equations is too variable for these methods, demonstrating with data for the Czech Republic for the period between 1990 and 2000.

The direct approach has been used by Fiorio and D'Amuri (2005) for Italy and by Matsaganis and Flevotomou (2008) for Greece, with similar basic assumptions as in this paper. These analyses were based on comparing incomes reported in administrative tax records and income survey, assuming that people might consider declaring a closer-to-true income to an anonymous interviewer. They focused on the active population only, and measured only income tax evasion, not that of social security contributions or VAT.

Existing evidence on the shadow economy in Hungary

The share of the underground (or shadow) economy[1] is rather high in Hungary in European comparison, ranging between 18 and 25% in 1999-2000 (Schneider and Klinglmair 2004).

[1] *Economic underground (our focus in this paper)* consists of activities that are productive in an economic sense and quite legal, but which are deliberately concealed from public authorities in order to avoid the payment of taxes or social security contributions. It includes underreporting of production, (understating of revenues or

This places the country into a group of high evaders together with other former Communist countries, but also Mediterranean countries such as Greece and Italy.

The calculations of Christie et al. (2005) provide one of the few recent comparative evidence on the extent of tax evasion. Their approach is particularly relevant to us, as they include both Personal Income Taxes (PIT) and Social Security Contributions (SSC) for the estimation of tax evasion. Their fundamental assumption was that tax evasion is uniform across all income groups (as they had no access to individual tax records). Our paper aims to relieve this latter assumption and explores differences across population groups. They found a higher compliance for PIT than for SSC in Hungary: while SSC compliance was 64%, PIT compliance reached 70% in 2002 (see Table 1).

Table 1. Income tax evasion in European countries

Country	SSC Compliance, %	PIT Compliance, %	SSC Theoretical Effective Rate, %	PIT Theoretical Effective Rate, %	Year
Austria	91	75	14,1	19,0	2003
Belgium	69	70	10,9	25,4	2002
Czech Republic	67	77	12,8	12,1	2003
Estonia		56		21,6	2003
France	72	60	3,8	16,5	1999
Germany	84	75	17,8	17,7	2002
Hungary	**64**	**70**	**8,5**	**21,1**	**2002**
Italy	83	62	6,0	22,7	2002
Latvia	53	45	6,8	18,9	2002
Netherlands		73		13,3	1998
Poland		66		18,6	1998
Portugal	66	68	7,7	12,1	2002
Slovakia	69	56	10,2	11,2	2002
UK	65	78	5,4	16,9	2002*

Note: *UK fiscal year: 6 April 2002 - 5 April 2003
Source: Christie et al. 2005

Semjén et al (2008), based on attitudinal survey questions, find that about 15% of all respondents received a share of their income cash-in-hand and 14% received part of their wage incomes as enterprise income. Altogether 26% of the respondents evaded some of their income taxes in 2006 and 2007.

Elek et al (2009) estimate the share of unregistered employment to 16-17% of the labour force, based on the comparison of administrative (pension insurance registry data) and survey data for 2001-2004. They find that in 2003 about 50-60% of those reporting minimum wage underreported their incomes and on average received about one-third of their actual income as "envelope wages".

3 Data

Our estimation of income underreporting is based on two datasets: a random sample of unaudited administrative tax records and the Household Budget Survey (HBS) of the Hungarian Central Statistical Office, both containing data on the 2005 incomes.

---

overstating costs), and also intentionally not registering (whole enterprises or parts of a registered enterprise). For a discussion of the concepts and their definitions, see the1993 SNA (System of National Accounts) and the OECD Handbook for the Measurement of the Non-Observed Economy.

The sample of administrative tax records (also referred to as APEH, reflecting the name of the tax authority) includes 227.688 individuals, about 5,4% of total taxpayers in the country[2]. The data refer to annual incomes from 2005. The sample size falls to 217,530 in the sample used for analysis. We top-coded the dataset by excluding those taxpayers who have incomes (any type of income) above the highest value in the survey data.[3] We also excluded taxpayers with zero taxable income[4]. The tax records include some socio-demographic characteristics of individuals, including age, sex, ZIP code, and for a smaller sub-sample the number of dependants, and occupation.

The HBS dataset includes 24.549 individuals in 9058 households. Income data is collected from household members aged 16 and over, while demographic information is available on all members. The income reference period is the calendar year of 2005.[5]

The results of the analysis crucially depend on the comparability of the two datasets: both in terms of the target population and in terms of income. Initially we thus assessed the comparability as such.

The main differences of the survey data and the administrative data include the following: (1) HBS is based on voluntary participation, while the filing of tax records is a legal obligation for those with taxable incomes; (2) under-sampling of high-income households may be present in the HBS due to non-response, thus underestimating top incomes and the extent of inequality; (3) incomes in the HBS are self-reported, thus recall errors might occur (respondents not remembering correctly).

The datasets, however, prove to be similar in crucial aspects: (1) both include personal incomes with reference to the calendar year of 2005, and incomes are measured on an annual basis (rather than e.g. on a monthly basis); (2) information on gross income is available in both; (3) both datasets include basic demographic information on respondents, including sex, age, region, employment status (employee or self-employed). These features ensure that the two datasets are actually comparable.

We created a comparable reference population in the two datasets, by (1) assuring the representativeness of the tax records sample, and by (2) reconciling the taxpayer population identified in the two datasets.

[2] There were only 4,37 million taxpayers in 2005, 44% of the total population, out of which 4,2 million had tax base taxed under the progressive tax system (APEH 2006). This later is our reference population.

[3] In order to exclude outliers from the Tax Audit data we excluded taxpayers with total income tax base above 26,88 million HUF, wage income above 19,67 million HUF, self-employment income above 24 million HUF and other taxable income above 7,21 million HUF. The number of these excluded observations is not substantial, altogether make up about 0,2% of the sample.

[4] We use a broad definition of taxable income including income subject to the progressive tax scheme + separately taxed self-employment income (tax base of Simplified Business Tax is not included). Note that other separately taxed income, such as capital income is not included in the analysis. That is why we excluded those with zero taxable income in order not to include taxpayers with only capital income in the sample.

[5] According to National Statistical Office high income households are underrepresented and low income households are overrepresented in the HBS, therefore average income is somewhat underestimated based on the HBS (KSH 2004, pp. 29.)

First, we thus reweighted the tax records sample on the basis of aggregate data on the entire population of taxpayers. The weights were calculated on the basis of region and employment status.

Second, the reweighted tax sample had to be reconciled with the HBS by restricting the latter sample to taxpayers. We thus reduced the HBS sample to those who had positive taxable income and declared to have filed a tax report[6]. This reduces the sample size to 9313. This is a good approximation of Hungarian taxpayers, because most social incomes, including pensions, universal family benefits and other cash transfers, were tax exempt and not reported to the tax authorities in 2005. It also implies that pensioners are only included if they have employment incomes (besides pensions). The pension, however, remained tax exempt in this case.

4 Methodology

Our aim is to assess the distributional implications of income tax evasion which is a result of income underreporting to tax authorities. For this, we construct a distribution of „true income", as observed in the income survey, and a synthetic distribution of "reported income", assumed to be revealed to tax authorities, also based on the HBS survey, but corrected for income under-reporting using information derived from the income tax records. The main methodological problem we have to overcome is that there is no single dataset that includes both the "true" and the reported taxable income of individuals. If their joint distribution $f(y_r, y_d, X)$, was observable, we could analyse its various features through summary statistics. y_r is true taxable income, y_d is reported taxable income and X is individual characteristics.

In the real world, such joint distributions are not observable. What we observe are two distributions in separate datasets, often measured for separate populations, and also often with measurement error. There is no trivial way therefore to re-create the joint distribution to connect them. Our aim here is to devise a method that allows us to re-generate one of the distributions in a dataset where only the other is observed.

Assume first that we have two datasets, where true and reported incomes are measured without error and in both datasets we are able to observe some individual characteristics X. This gives us two distributions: the true and declared taxable income: $f(y_r, X)$ and $g(y_d, X)$. Clearly, the only information connecting these distributions is the information contained in the X variables. If we are to successfully connect these distributions, we have to make use of X.

We define both income categories as a function of some individual characteristics, where $Y_d(X)$ is the distribution of declared income, $Y_r(Z)$ is the distribution of true taxable income and X and Z are non-exclusive sets of individual characteristics. We can use the common variables in X and Z to impute the relationship of the two income categories.

We can estimate declared taxable income from the tax returns data as a function of the common set of individual characteristics (x_i):

$$y_{di} = \alpha + \beta x_i + \varepsilon_i \qquad (2)$$

[6] The variable indicating the amount of "tax liabilities" was not a missing value.

Then from (2) we have $\hat{\alpha}$ and $\hat{\beta}$. According to the standard OLS case if there is measurement error, due to omitted variables, then $\hat{\alpha}$ and $\hat{\beta}$ will be biased.

By using $\hat{\alpha}$ and $\hat{\beta}$ the imputed declared income measure in the HBS dataset can be calculated by:

$$\hat{y}_{di} = \hat{\alpha} + \hat{\beta}\, x_{ih} \qquad (3)$$

where x_{ih} is the individual characteristics in the HBS dataset.

In case we are working with a set of indicator variables, estimating (2) in the tax returns data is equivalent to calculating population averages for subgroups defined by this set of variables. This way estimating declared income in the survey dataset (3) becomes

$$y_d = \pi_1 D_1 + \pi_2 D_2 + \ldots + e \qquad (5)$$

where D_i-s are the dummies for the subgroups defined by the individual characteristics (region, agegroup, sex) and $\hat{\pi}_i$ -s will be equal to group average reported incomes.

We would like to estimate the relationship of y_d and y_r (τ, income underreporting):

$$\tau = y_d / y_r \qquad (6)$$

with heterogeneity in τ.

Using (5) we have

$$\hat{\tau}_i = (\hat{\pi}_1 D_1 + \hat{\pi}_2 D_2 + \ldots)/ y_{ri} \qquad (7)$$

or

$$\hat{\tau}_i = \hat{\pi}_1 D_1 / y_{ri} + \hat{\pi}_2 D_2 / y_{ri} + \ldots \qquad (8)$$

If the same D_i-s are defined in the 2 datasets (same set of individual characteristics are observable), $\hat{\tau}$ will be defined for each population subgroup and will be equal to the ratio of average reported income and average true income in the given subgroup.

Based on the above, our estimation strategy is the following. [7] First we defined comparable income categories in the two datasets: gross personal taxable income. It is gross, thus before the payment of taxes. It is personal, rather than household level, which adequately matches the individual based taxation system in Hungary. It is taxable, thus it refers to positive incomes subject to tax. Thus, we only included taxpayers in the sample who had positive total income (in both datasets).[8]

[7] On the features of the Hungarian tax policy see Annex B.

[8] We defined the following variables on the two datasets:

- Wage income: this is part of the tax base of the progressive tax scheme
- Self-employment income: sum of wage income form self-employment (the tax base of the progressive tax scheme) and other income from self-employment (regarded as capital income in the Hungarian tax system).

As pensions and benefits are exempt from tax, we did not include them in our definition of the taxable income. Capital incomes (other than income from self-employment) are not included in the definition of incomes used here[9].

We assume that "true taxable income" is reported in the income survey and "reported taxable income" is revealed to the tax authority. Based on the survey income we construct a synthetic distribution of reported income by adjusting survey income of each individual for underreporting with adjustment factors calculated for each subgroup. We calculate these adjustment factors by comparing the average tax returns income to the average survey income for a given subgroup, separately for wage income, self-employment income and other income. The common set of individual characteristics that we can use for defining these subgroups is very limited: we only have age group, gender and region, but age and gender are usually endogenous with income. Using all the 3 variables we can define 70 (5*2*7) subgroups ("cells") whereas using only region gives only 7 subgroups.

It is important to note that in our analysis we were not concerned about the employment status of the individual (employed or self-employed) but the part of income coming from the different sources. We calculated reported income for the part of income coming from wage and self-employment sources separately and then added up these figures for total declared income. Later on when we refer to "self-employed" (rather than self-employment income), we use our definition of self-employed rather than a self declared status. We regarded those taxpayers self-employed who declared at least 1 forint (HUF) of any kind of self-employment income in 2005.

Our results crucially depend on the assumption that tax evaders have no incentive to conceal their true income when responding to an income survey. Some papers (e.g. Fiorio and D'Amuri (2005)) argue for this assumption whereas others, especially those using the consumption-based approach to estimate income underreporting (e.g. Pissarides and Weber (1989), Lyssiotou et al (2004)) suppose that income in surveys is unreliable. Even these studies use a group of individuals as reference population, for whom they suppose that reported income is real. Normally this reference population is the employed and they calculate income underreporting of the self-employed compared to them.

We assume that survey incomes are "true", at least there are weaker incentives to underreport income in an anonymous survey than to the tax authorities. In case income is underreported in the HBS, our estimate of underreporting will be lower bound estimates.

5 Results

- Other taxable income: other income components that are part of the tax base of the progressive tax scheme, except wage income from self-employment

- Total income is the sum of wage income, self-employment income and other taxable income.

[9] Also note that those income that are not part of the tax base but increases the total tax liability (adoterhet nem viselo jarandosagok) are not part of our total taxable income either. The reason is that the information on these items in the HBS is insufficient. However this does not affect our results substantially as in 2005 pension was completely tax free and GYES was exclusive of wage income, so the sum of these types of income is very limited.

According to our calculations, income underreporting rate is 9-13%, when all three explanatory variables (age, gender and region) are used (see Table 3). With respect to income types, wage income is underreported by some 4%, self-employment income by about 73% and other income by 10%.[10]

The extent of income underreporting is not uniform across income groups, but the highest in the richest two deciles (about 11-21%) and around the minimum wage (second decile, ca. 10%). Thus, the rich benefit the most from income tax evasion, not only in monetary value but also as a proportion of their incomes.

Table 3. Underreporting by level of income under different specifications

Income decile of taxpayers	Mean „true" income	Rate of underreporting			
		(1) explanatory var.s: region, agegr, gender – not topcoded	(2) explanatory var.s: region, agegr, gender – topcoded	(3) explanatory var.: region – not topcoded	(4) explanatory var.: region – topcoded
1 (poorest)	301	-8%	5%	1%	4%
2	694	4%	10%	8%	10%
3	90	1%	7%	5%	7%
4	1090	1%	6%	4%	5%
5	1265	2%	7%	4%	6%
6	1460	5%	9%	7%	8%
7	1719	6%	9%	7%	8%
8	2053	6%	9%	7%	9%
9	2634	11%	13%	11%	12%
10 (richest)	4645	19%	21%	19%	19%
Total	1676	9%	13%	10%	11%

Notes:
Reported Income = Adjustment Factor * True Income
underreporting=(True income-Reported Income)/True Income
Top-coded means that adjustment factors are maximised to 1, not allowing income overreporting.
Mean income by income group is annual gross personal income in thousand forints. "True" income is as observed in the HBS dataset. Income quantiles of taxpayers were generated based on "true" income, excluding those earning zero or negative incomes.

We estimate reported income for all taxpayers by applying our adjustment factors - the estimated rate of income underreporting - to HBS taxable income measures. We find that underreporting is quite different for different population groups. Self-employed[11] tend to underreport most: about two thirds of their incomes are not reported as a tax base to the tax authority. In contrast, employees seem to comply with tax rules at large, with an overall rate of underreporting of 4% (Table 4).

Underreporting is the highest in the highest income region, Central Hungary (including Budapest). This may be explained with the higher share of economic sectors particularly

[10] As mentioned in footnote 5 average income is underestimated in the HBS, therefore our results are lower bound estimates.
[11] We define self-employed as those who earned at least 1 HUF of any kind of self-employment income in the reference period.

prone to tax evasion. Much of the construction industry and the service sector can be found around the capital, in the highest income region. It is followed by a rich (West-Transdanubia) and a poor (South Great Plain) region.

Tax evasion is higher among men and somewhat increases by age, especially for people around the retirement age. This is partly a composition effect: (i) men tend to be over twice as frequently entrepreneurs than women, (ii) there are more self-employed among those aged 60 or over than in younger age groups. Our calculations also show that men underreport both wage income and self-employment income more than women do and the elderly underreport more. Based on a specific survey on Hungary, Tóth (2008) also finds that men tend to underreport wages more than women, while he concludes that envelope wages are more widespread among young and middle-aged people than the old.

Table 4. Underreporting by main source of income, region, age and gender

	Population share	*"True income"*	*Calculated reported income*	*Rate of underreporting*
Employment status				
Wage earners	90%	1603	1542	4%
Self-employed	10%	2325	770	67%
Region				
Central Hungary	31%	2166	1787	17%
Central Transdanubia	11%	1488	1376	7%
West-Transdanubia	12%	1549	1347	13%
South Transdanubia	7%	1396	1274	9%
North Hungary	12%	1431	1355	5%
North Great Plain	13%	1402	1278	9%
South Great Plain	14%	1468	1289	12%
Gender				
Male	50%	1898	1575	17%
Female	50%	1455	1354	7%
Age group				
0-14	0%	1138	230	80%
15-29	18%	1264	1144	10%
30-44 yrs	39%	1739	1496	14%
45-59 yrs	41%	1802	1589	12%
60+ yrs	2%	1686	1313	22%

Notes:
Reported Income = Adjustment Factor * True Income
Rate of underreporting=(True income-Reported Income)/True Income
Here we present results for the case where region, agegroup and gender are used as explanatory variables and adjustment factors are topcoded. Topcoded means that adjustment factors are maximised to 1, not allowing income overreporting.
Mean income by income group is annual gross personal income in thousand forints. "True" income is as observed in the HBS dataset. Income quantiles of taxpayers were generated based on "true" income, excluding those earning zero or negative incomes.

In order to see the potential effects of tax evasion we make use of a microsimulation model that can take into account interactions of the elements of the tax and benefit system. Therefore in the next step we apply the estimated adjustment factors in the tax-benefit microsimulation model (EUROMOD) in order to calculate how income underreporting affects tax revenues, progressivity and the distribution of incomes. The total synthetic reported income for each individual comprises the sum of the synthetic reported wage income and synthetic reported

self-employment income. Income underreporting can therefore be calculated by comparing the survey taxable income to this total synthetic reported income.

Table 6. Adjustment factors used in the microsimulation model by region and type of income

	Employment income	*Self-employment income*
Central Hungary	0,960	0,211
Transdanubia	1,000	0,288
Great Plain and North	0,990	0,342

Notes: The estimated adjustment factors for employment income in Transdanubia was recoded to 1, thus excluding the possibility of over-reporting to the tax authorities.
Since information is available only for three main regions (NUTS1 level) and two main types of income in the microsimulation database (EU-SILC), the number of adjustment factors was limited to six.

Income underreporting as such modifies total household incomes, but the impact is largely dependent on the tax system and also on the system of cash benefits. The use of a tax-benefit model allows us to account for (i) the complexity of the tax system, including the fact that only some income components are subject to tax, while others are not, (ii) a potential interaction between specific cash and tax policies (benefit entitlements may also change as a result of tax evasion).

We aim to estimate the distribution of "net income", calculating net incomes in the following way.

$$Y_{net} = Y_{reported} * (1- t) + Y_{hidden} + Y_{social}$$

Where Y_{net} is total personal disposable income for individual i, $Y_{reported}$ refers to reported income, t is the total tax rate, Y_{hidden} is the income not-reported to the tax authority, and Y_{social} is the amount of social transfers received by individual i. Note that in the Hungarian case $Y_{reported}$ consists of only labour income (and not social benefits), and Y_{social} is non-taxable, with the exception of insurance-based maternity benefits, which are included in labour incomes.

Next we present our calculations based on EUROMOD, the European tax-benefit microsimulation model (see Lietz and Mantovani 2007, Lelkes and Sutherland 2009) [12]. This model includes all major direct tax and cash benefit policies in 2005, thus allows us to consider the complexity and interaction of these. In addition to the policy rules, the model includes a nationally representative household dataset, which enables the user to assess the impact of specific policy measures (or their changes) on the total population. In our case, we do not model policy change as such, but rather a change in one of the basic assumption of microsimulation models, that of full compliance. Using our adjustment factors for income underreporting we are able to assess the impact of tax evasion on the budget, but also on individuals' incomes and their distribution.

Table 7. Fiscal and distributional implications of tax evasion

	Full compliance	*Tax evasion*	*Difference*

[12] This paper uses EUROMOD, the European tax-benefit microsimulation model, as a means for estimating the incidence of income tax evasion. The EUROMOD is based on the EU SILC, therefore using the same dataset, instead of the HBS dataset for estimating underreporting would have been a logical choice. However our preliminary calculations showed that income data is not detailed enough in the EU SILC for our analysis, which hindered the comparison with the tax audit data. Thus, we concluded that using the HBS data for underreporting estimates and using the EU-SILC based microsimulation model for distribution analysis is more appropriate.

Personal income tax receipts (billion HUF, annual)	1.119	902	-19,4%
Poverty line (HUF, monthly)	45.279	46.822	3,4%
Poverty rate (FGT a=0)	14,8	14,6	not sign.
Poverty gap (FGT a=1)	3,4	3,3	not sign
Gini	0,274	0,292	6,8%
S80/S20	4,020	4,320	7,5%
Atkinson e=0.5	0,064	0,076	17,9%
Atkinson e=2	0,235	0,253	7,6%
Theil	0,139	0,173	24,5%
Kakwani	0,259	0,231	-10,8%
Reynolds-Smolensky	0,073	0,053	-26,9%
Suits	0,294	0,254	-13,5%

Notes: full compliance provides estimates of income tax variables assuming incomes are reported to tax authorities as observed in the EU-SILC. Tax evasion provides estimates of the same variables assuming incomes are under-reported to tax authorities by the adjustment factors shown in Table 6. FGT refer to the Foster Greer Thorbecke family of poverty indices.
Income concept: equivalised household income, monthly.
Difference in the poverty rates and the poverty gaps under the two scenarios are not statistically significant.
The actual amount of personal income tax receipts in the state budget was 1.207 billion forints in 2005 (APEH 2006). It includes tax on other types of incomes (e.g. agricultural incomes, intellectual activities) which are not measured accurately in the EU-SILC survey.

The fiscal implications of tax evasion are substantial; the "tax gap", the difference between the taxes households actually owe and what they report, equals about 220 billion forints. Revenues fall short by about 19%. This figure is somewhat higher than the Ministry of Finance estimate, quoted by the World Bank report (World Bank, 2008) which estimates the tax gap to be above 10%. The microsimulation model somewhat underestimates the total budget revenue from income taxes, which was 1.207 billion forints in 2005 (APEH 2006), due to data limits. For specific income types, the number of observations is very small, for some others, there are measurement errors (e.g. property incomes, agricultural incomes, intellectual activities). On the other hand, certain tax rules (especially tax credits) are simplified in the model, as there is no adequate information in the income survey (e.g. on donations to charities). These modelling features, however, are not likely to affect the estimated implications of tax evasion as such, as they are expected to affect the results under both scenarios equally.

Income inequality is significantly higher under tax evasion. The Gini coefficient and the quintile ratio increase by 7-8%. Similarly, the Atkinson index also rises for both alternative values of the inequality aversion parameter[13]. This has two major implications. First, high earners tend to evade proportionately more, and second, progressivity of the tax system is lower under tax evasion than implied by the tax legislation. This is confirmed by the indicators measuring progressivity.

Tax evasion reduces the progressivity of the income tax system to a rather large extent. All indices of progressivity suggest that income tax evasion reduces progressivity. The Kakwani and Suits indices indicate a decline of 11-14%, while the Reynolds-Smolensky an even greater, 27% fall. In these calculations we consider the impact of both personal income taxes and social security contributions, thus all taxes on labour at the employee level. Note that we

[13] Note that the extent of the change is greater when the inequality aversion parameter is smaller (e=0.5), when less weight is attached by the society to redistribution to the poor. (The potential range of e is from 0, which means that the society is indifferent to the redistribution, to infinity, where the society is concerned only with the position of the poorest income group.)

measure progressivity of (equivalised) household incomes, rather than personal incomes. Equivalised household income is a better proxy for individual resources, as individuals live in households and share incomes. The use of (equivalised) household income for measuring tax progressivity is clearly a virtue of tax benefit microsimulation models as normally information on tax payments (e.g. in tax records) cannot be linked to household structures or other sources of household incomes.

By international comparison, the estimated tax evasion in Hungary is rather high, but not outstanding. In Greece, the tax gap with respect to personal income taxes was found to be 25% (Matsaganis and Flevotomou 2008). The tax gap (of all federal taxes, not simply income taxes as in our case) in the US was estimated to be 17% in the early 1990's, and found to be relatively unchanged in the past 20 years (Andreoni et al 1998). Non-compliance was calculated using data from intensive audits conducted on a stratified random sample of tax returns. One third of taxpayers were found to underreport their incomes.

An important implication of our results is that tax evasion is not uniform across social groups therefore tax authorities can target audits to certain groups in order to increase efficiency. We do no assess, however, how taxpayers may react to policy changes. Bakos et al. (2008) analyzed taxpayer behaviour using the 2005 tax reform which changed marginal and average tax rates but kept tax enforcement constant. They assessed the impact of tax changes on taxable income and found that tax elasticity is high for mid- and high-income earners and state that it might come from high responsiveness of income tax evasion to tax rates. As shown in this paper, tax evasion is indeed comparatively high at the top of the income distribution. Therefore it is likely that high income earners would react to lower marginal tax rate by lowering tax evasion which would then increase budget revenues.

6 Conclusions

The paper estimated the incidence of income tax evasion in Hungary on the basis of a random sample of administrative tax records of 230 thousand individuals, not accessible for research so far. Gross incomes in the administrative tax records are compared with those in a nationally representative income survey, HBS, assuming that tax-evaders are more likely to report their true incomes in an anonymous interview. We estimated income underreporting for those who report at least some income to the tax authorities, leaving out those individuals who do not declare any of their income. The method we applied in our paper provided the first micro-data-based estimates for personal income tax evasion in Hungary.

Our estimates show that the average rate of underreporting is 11%, which conceals large differences between the self-employed (who hide the majority of their incomes) and the employees. Men are more likely to hide their incomes than women, but it is due to the composition effect: the majority of self-employed are men. The rate of underreporting was found to be the highest in case of taxpayers at the top of the distribution and those with incomes around the minimum wage level. Due to the parameters of the tax system (progressive tax scheme and the employee tax credit), the similar rate of underreporting, however implies rather different actual tax payments in these two income groups. The progressive tax scheme imposes greater relative burden on those with high incomes, and in parallel, employee tax credit benefits employees below the average wage. In absolute terms, however, the evasion of top earners affects the budget more, as expected.

The estimated rates of underreporting are then used in EUROMOD, the European tax-benefit microsimulation model to calculate the fiscal and distributional implications of underreporting, comparing the scenarios of full compliance to that of tax evasion, while taking account of all major direct taxes and cash benefits and their interactions. Tax evasion reduces fiscal revenues from personal income taxes by about 19%. While poverty does not change under tax evasion, income inequality becomes significantly higher (the Gini coefficient and the quintile ratio increase by 7-8%), showing that high earners tend to evade proportionately more. This implies that the progressivity of the tax system is lower under tax evasion than intended. The effects are rather large, as shown by various indices of progressivity, suggesting a decline of 11-26%. As our estimates are likely to be lower bound (due to potential measurement error in the income survey), the actual effects of tax evasion on the total redistribution are expected to be larger.

In the policy debate, tax evasion is often attributed exclusively to the high level of taxes in Hungary, or to a culture of free-riding by citizens. Policy-makers tend to be concerned primarily with the fiscal loss arising from tax evasion. Our results contribute with a new aspect to this debate: due to the greater ability of high income earners to evade taxes, tax evasion tends to increase income inequality. The reduced progressivity of the income tax scheme is likely to alter social outcomes from the pursued policy goals, and thus may undermine the equity of income redistribution. We have also shown that specific rates of tax evasion vary substantially by social groups and these differences alter income inequality, poverty and tax progressivity. As long as tax evasion prevails, informed policy decisions aiming to promote social welfare need to consider its implications.

Annex A. Descriptive and summary statistics

Table A1. Main characteristics of the taxpayers in administrative and survey datasets

Number of observation

APEH				*HBS*			
regions	*employed*	*entrepr*	*total*	*regions*	*employed*	*entrepr*	*total*
Central Hungary	55097	3465	58562	Central Hungary	2481	378	2859
Central Transdanubia	18632	1279	19911	Central Transdanubia	963	86	1049
West Transdanubia	18017	1266	19283	West Transdanubia	992	123	1115
South Transdanubia	21320	1314	22634	South Transdanubia	630	55	685
North Hungary	18977	1148	20125	North Hungary	1070	79	1149
North Great Plain	24161	1650	25811	North Great Plain	1094	102	1196
South Great Plain	21556	1577	23133	South Great Plain	1142	118	1260
Total	177760	11699	189459	Total	8372	941	9313

Share (%)

APEH	*employed*	*entrepr*	*total*	*HBS*	*employed*	*entrepr*	*total*
Central Hungary	29%	2%	31%	Central Hungary	27%	4%	31%
Central Transdanubia	10%	1%	11%	Central Transdanubia	10%	1%	11%
West Transdanubia	10%	1%	10%	West Transdanubia	11%	1%	12%
South Transdanubia	11%	1%	12%	South Transdanubia	7%	1%	7%
North Hungary	10%	1%	11%	North Hungary	11%	1%	12%
North Great Plain	13%	1%	14%	North Great Plain	12%	1%	13%
South Great Plain	11%	1%	12%	South Great Plain	12%	1%	14%
Total	94%	6%	100%	Total	90%	10%	100%

Number of observation

APEH	*employed*	*entrepr*	*total*	*HBS*	*employed*	*entrepr*	*total*
male	95606	8203	103809	Male	4013	628	4641
female	108201	5520	113721	Female	4359	313	4672
Total	203807	13723	217530	Total	8372	941	9313

Share (%)

APEH	*employed*	*entrepr*	*total*	*HBS*	*employed*	*entrepr*	*total*
male	44%	4%	48%	Male	43%	7%	50%
female	50%	3%	52%	Female	47%	3%	50%
Total	94%	6%	100%	Total	90%	10%	100%

Notes:
APEH: administrative data from the tax authority
HBS: Household Budget Survey

Annex B. Features of the Hungarian tax policy

The income tax system is subject to frequent (mostly annual) changes, including both the tax rates and the tax brackets. Some of this might be explained with the indexation to inflation, but in most cases they reflect changing policy priorities, at times focusing on raising revenues, other times on lowering the tax burden.

Table 2. Personal income tax brackets (in HUF) and rates

2003		*2004*		*2005*		*2006*	
Tax bracket	*Rate*	*Tax bracket*	*Rate*	*Tax bracket*	*Rate*	*Tax bracket*	*Rate*
0-650.000	20%	0-800.000	18%	0-1.500.000	18%	0-1.550.000	18%
650.001- 1.350.000	30%	800.001- 1.500.000	26%				
1.350.001-	40%	1.500.001-	38%	1.500.001-	38%	1.550.001-	36%

In addition to PIT, employers had to pay a total of 32% as social security contributions on their labour incomes in 2004. Employee social security contribution amounted to 13,5%. These rates had been relatively stable and remained the same between 2004 and 2006.

Budget revenue from personal income taxes made up 6,6% of the GDP in 2005 (European Commission 2007). The Hungarian budget however relies heavily on indirect taxes, and received about 53% higher amount from VAT than from PIT in 2005[14].

With respect to the distribution of tax burden in the country: about one third of taxpayers paid the higher marginal tax rate of 38% on some of their total incomes in 2005.

[14] Ministry of Finance, Hungary, Balance sheet of the central government annual budget. Retrieved on 14 January 2009 from www.pm.gov.hu

References

Allingham, M. G. and A. Sandmo (1972). "Income tax evasion: a theoretical analysis." Journal of Public Economics **1**(3-4): 323-338.

Andreoni, J., B. Erard and J. Feinstein (1998). "Tax Compliance." Journal of Economic Literature **36**(2): 818-860.

APEH (2006). Az APEH világa 2006. évkönyv [APEH Bulletin 2006] Budapest, APEH.

Bakos, P., P. Benczur and D. Benedek (2008) The Elasticity of Taxable Income: Estimates and Flat Tax Predictions Using the Hungarian Tax Changes in 2005, MNB Working Papers 2008/7., Hungarian National Bank

Christie, E., M. Holzner and M. Lacko (2005). Taxation and the Shadow Economy in Europe. Vienna, wiiw. The Vienna Institute for International Economic Studies. OeNB Jubiläumsfondsproject No. 10867.

Csontos, L., J. Kornai and I. G. Tóth (1998). "Tax awareness and reform of the welfare state: Hungarian survey results." Economics of Transition **6**(2): 287-312.

Dhami, S. and A. al-Nowaihi (2007). "Why do people pay taxes? Prospect theory versus expected utility theory." Journal of Economic Behavior & Organization **64**: 171-192.

Elek, P., A. Scharle, B. Szabó, P.A. Szabó (2009): A feketefoglalkoztatás mértéke Magyarországon (Employment in the black economy in Hungary), in: Semjén A. – Tóth I.J. (ed.) Shadow economy, MTA-KTI, Budapest. pp. 84-102

European Commission (2007). Taxation trends in the European Union, Data for the Member States and Norway, Luxembourg

Fiorio, C. and F. D'Amuri (2005). "Workers' tax evasion in Italy." Giornale degli Economisti e Annali di Economia **64**(2/3).

Fortin, B., G. Lacroix and M.-C. Villeval (2007). "Tax evasion and social interactions." Journal of Public Economics **91**(11-12): 2089-2112.

Frey, B. and F. Schneider (2001). Informal and Underground Economies. International Encyclopedia of the Social and Behavioral Sciences. N. J. Smelser and P. B. Baltes. Oxford, Elsevier**:** 7441-7446.

Frey, B. S. and B. Torgler (2007). "Tax morale and conditional cooperation." Journal of Comparative Economics **35**(1): 136-159.

Friedman, E., S. Johnson, D. Kaufmann and P. Zoido-Lobaton (2000). "Dodging the grabbing hand: the determinants of unofficial activity in 69 countries." Journal of Public Economics **76**(3): 459-493.

Hanousek, J. and F. Palda (2004). "Quality of government services and the civic duty to pay taxes in the Czech and Slovak Republics, and other transition countries." Kyklos **57**(2): 237-252.

Hanousek, J. and F. Palda (2006). "Problems measuring the underground economy in transition." Economics of Transition **14**(4): 707-718.

Krekó, J. and G. P. Kiss (2007). Adóelkerülés és a magyar adórendszer [Tax evasion and the Hungarian tax system]. MNB-tanulmányok 65. Budapest, Hungarian National Bank.

KSH (2004), Háztartás-statisztikai közlemények [Reports on Household Statistics], No. 25. 2004/I, Budapest, Központi Statisztikai Hivatal

Lelkes, O. and H. Sutherland, Eds. (2009). Tax and Benefit Policies in the Enlarged Europe. Assessing the Impact with Microsimulation Models. Aldershot, Ashgate.

Lietz, C. and D. Mantovani (2007). "A Short Introduction to EUROMOD: An Integrated European Tax-Benefit Model". Micro-simulation in action: Policy analysis in Europe using EUROMOD. O. Bargain, Elsevier Ltd. Research in Labor Economics, 25: 1-26.

Lyssiotou, P., P. Pashardes and T. Stengos (2004), Estimates of the Black Economy Based on Consumer Demand Approaches, Economic Journal 114, 622-640.

Matsaganis, M. and M. Flevotomou (2008). "Distributional Implications of Income Tax Evasion: the Case of Greece." manuscript.

Pissarides, C. and G. Weber (1989), An Expenditure-Based Estimate of Britain's Black Economy, Journal of Public Economics 39, 17-32.

Schneider, F. and D. H. Enste (2000). "Shadow Economies: Sizes, Causes, and Consequences " Journal of Economic Literature **38**(March): 77-114.

Schneider, F. and Klinglmair, R. (2004). Shadow Economies around the World: What do we Know? CESifo Working Paper No. 1167 and IZA Discussion Paper No. 1043. University of Linz.

Semjén, A., I. J. Tóth, M. Medgyesi and Á. Czibik (2008). Tax evasion and corruption: population involvement and acceptance. Survey results, DP-2008/13. Discussion paper of the Institute of Economics, Hungarian Academy of Sciences, Budapest.

Tóth. I. J. (2008). "Be nem jelentett munkavégzés Magyarországon – elterjedtség és társadalmi összetevők [Unregistered employment in Hungary – size and social causes]" Conference presentation, ILO Triparty Conference, December 8., 2008, Budapest.

World Bank (2008). Reducing Undeclared Employment in Hungary, Synthesis Report of the World Bank Study, Draft Report, March 2008