University of Tartu

Faculty of Economics and Business Administration

WHY DO INDIVIDUALS EVADE PAYROLL AND INCOME TAXATION IN ESTONIA?

Kenneth A. Kriz, Jaanika Meriküll,
Alari Paulus, Karsten Staehr

Tartu 2007

ISSN 1406–5967
ISBN 978–9949–11–541–9
Tartu University Press
www.tyk.ee
Order No. 6

WHY DO INDIVIDUALS EVADE PAYROLL AND INCOME TAXATION IN ESTONIA?

Kenneth A. Kriz[1], Jaanika Meriküll[2], Alari Paulus[3], Karsten Staehr[4]

Abstract

This paper employs micro-level data to determine the factors characterizing individuals who evade payroll and income taxation in Estonia. Using logit estimation on three different cross-sectional datasets, we estimate the marginal effects of different individual characteristics on tax evasion. The three datasets give broadly analogous results. Payroll and income tax evasion is most prevalent in small firms and in the construction and agricultural sectors. Evasion is more common among individuals who work part-time, are of non-Estonian ethnicity, have relatively short education, earn a low income and are men. Tax evasion is more frequent among the young and the elderly than among the middle-aged. There are clear regional differences. The overall picture is that the relatively disenfranchised are most likely to evade payroll and income taxation in Estonia.

Keywords: Tax evasion, unreported work, incentives, tax system

JEL classification: H26, H24, D19

---

[1] School of Public Administration, University of Nebraska at Omaha, Omaha, USA.

[2] Faculty of Economics and Business Administration, University of Tartu, Tartu, Estonia.

[3] Institute for Social and Economic Research, University of Essex, Colchester, United Kingdom.

[4] Corresponding author: Research Department, Bank of Estonia, Estonia pst 13, 15095 Tallinn, Estonia. E-mail: karsten.staehr@epbe.ee. Tel.: +372 6680926, Fax: +372 6311240.

1. INTRODUCTION

This paper uses micro-level data to determine the characteristics of individuals who evade payroll and income taxation in Estonia. The extent and the distribution of tax evasion affect the efficiency and distributional characteristics of the tax system. Specific knowledge of evasion patterns is important when assessing the effects of individual taxes and can provide useful background information for the design and reform of taxation, auditing and penalty schemes.

The effects of tax evasion on efficiency and distribution are complex (Andreoni *et al.* 1998, Cowell 1990). Evasion reduces the tax base, which calls for higher tax rates and increases the excess burden of taxation. In certain cases, however, the excess burden is reduced if the evasion is primarily undertaken by individuals whose trades would otherwise have been deterred by the tax, i.e. by the taxpayers who would have borne the excess burden in the absence of evasion. The same ambiguity applies to the equity dimension. Evasion can make the distribution more arbitrary and unequal, but evasion can also affect the distribution positively if it primarily benefits individuals who are socially important, e.g. less advantaged individuals. In many low-income countries untaxed income from the informal sector constitutes an important 'safety net' for disenfranchised persons.

The overall extent of tax evasion in an economy is important, but to assess the welfare economic consequences it is equally important to ascertain *who* evades taxation. The pioneering paper by Allingham & Sandmo (1972) considered a risk-averse rational individual with an exogenous income deciding his or her tax evasion given the tax and penalty rates and the probability of detection. Subsequent papers extended the analysis, e.g. by considering different penalty schemes and endogenizing the labor supply. The literature generally finds few unambiguous results although many models predict that higher income leads to more evasion. It is also clear from this literature that the individual's risk preferences and perception of the auditing and penalty schemes play an important role for evasion decision (Andreoni *et al.* 1998: sec. 6). The individual's risk preferences and perceptions are likely to be related to different characteristics like age, gender, race, and

education as well as factors as the individual's workplace and residence. These variables may therefore help explain the likelihood of an individual engaging in tax evasion.

Another strand of the theoretical literature asserts that morals or social norms help explain an individual's decision to evade taxation. This reasoning is supported by the fact that there is generally less evasion in practice than models with rational individuals would suggest (Andreoni *et al.* 1998: sec. 8). A taxpayer may evade less if norms in society make the individual feel guilt in case of evasion. Social morality and norms may be an economy-wide phenomenon, but it could also differ across social groups or regional areas. This line of reasoning implies — as in the models with a rational taxpayer — that variables like age, gender, race, education, workplace, residence etc. affect the decision to leave income unreported.

Turning now to the empirical literature, much effort has gone into estimating the overall size of the informal sector in individual countries or across countries. Fewer studies have sought to uncover what determines the prevalence or degree of tax evasion among individuals, especially for countries outside the USA. Andreoni *et al.* (1998) provides a survey of microeconometric studies of the determinants of income tax evasion in high-income countries: (i) The estimated relationship between tax rates and evasion varies, but most studies find that higher income is associated with more evasion. (ii) More frequent auditing, more intensive auditing and higher fines are usually found to lead to less evasion, but the effects are small. (iii) Social norms stressing law abidance and perceptions of the tax system being 'fair' and the government well functioning lead to less evasion. (iv) Background variables like gender, age, race, education and family relationship are important predictors of evasion.

Only a small number of studies have used econometric methods to assess the factors determining tax evasion or work in the informal sector in *transition economies*. Gardes & Starzec (2002) use data from an enlarged Labor Force Survey undertaken in Poland in 1995 and estimate the probability of an individual working in the informal sector. They find that the likelihood of informal employ-

ment increases if the individual is otherwise unemployed, a man, has residence in a region with high unemployment, is more than 60 years old, lives in the countryside, has only primary education or is self-employed. They also find evidence of a 'network effect' where familiarity with informal markets increases the individual's probability of participation in such markets.

Kim (2005) also finds — using household survey data from 1996 — that the informal sector functions as a resort for poor families in Romania. Higher income from the formal sector reduces the participation in the informal sector, while perceived poverty increases the participation. Families living in rural areas engage more in informal activities than families living elsewhere, while the coefficients to background variables like age and education are insignificant.

Kolev (1998) finds for Russia that social misfortunes, as e.g. unemployment, lead to an increased likelihood of work in the informal sector, but the earning possibilities in the sector are also important. The study uses data from the mid-1990s when the transition was still in its early stages.

Hanousek & Palda (2004) analyze surveys of individuals in a number of Central European countries (from 2002) and find that tax evasion is more prevalent among individuals who believe that the probability of being audited is small and among individuals who are dissatisfied with the level of government services. Torgler (2003) compares the tax *morale* across transition countries using data from the World Value Survey. Tax morale is higher in Central and Eastern Europe than in the countries emerging from the Soviet Union. Trust in the legal system and in the government is positively correlated with tax morale.[5]

In sum, empirical research for transition countries in the 1990s indicates that informal work and tax evasion were most prevalent

[5] Johnson *et al.* (2000) show that firms in Russia and Ukraine declare less of their output than firms in Poland, Slovakia and Romania. This is partly a result of the more pervasive bureaucratic corruption in Russia and Ukraine.

among disenfranchised individuals (Kolev 1998, Gardes & Starzec 2002, Kim 2005) and in many cases functioned as a social safety net (cf. also the contributions in Neef & Stanculescu (2002)).

This study contributes to the literature in several ways. First, the study considers the prevalence of tax evasion in a post-communist country using data from 2002-04 — instead of data from the mid-1990s as in the studies above. Using the more recent data implies that inferences can be made for a country which has 'graduated' from the transition process. By 2002-04 the transition-induced restructuring was largely completed in Estonia and the economy grew rapidly. Second, the Estonian tax system is unusually 'clean' with essentially everybody facing the same marginal tax rate. This reduces the problems of disentangling the effects of different marginal tax rates and other determining factors. Third, the study addresses the data problems inherent in tax evasion estimations by contrasting the results from three different datasets.

The paper proceeds as follows: Section 2 provides background information and briefly discusses the data sources and the research methodology. The next three sections present the empirical results using three different datasets. Section 3 uses survey data from the Estonian Institute for Economic Research, section 4 uses audit data from the Estonian Tax and Customs Board, and section 5 uses data from the Estonian Labour Force Survey. Section 6 brings together the results from the three datasets and presents an overall picture of the determinants of payroll and income tax evasion in Estonia.

2. BACKGROUND, DATA AND METHODOLOGY

Estonia is a small country in Northern Europe with 1.3 million inhabitants. The country regained independence from the Soviet Union in 1991 and embarked immediately on a comprehensive reform program (Staehr 2004). In spite of impressive growth rates since the mid-1990s, the per capita GDP in 2004 was only little more than half the EU average (Eurostat 2006). Transition has left the income distribution relatively unequal and pockets of poverty

remain among those who have found stable employment, in particular the young and the elderly, individuals with little education and the non-Estonian speaking part of the population (UNDP 2001).

The Estonian payroll and income tax system is relatively simple (Ministry of Finance 2006). The payroll tax paid by the employer amounts to 33% of the gross wage, but 4%-points of the payroll tax has for most Estonians been transferred to private retirement accounts since 2002. The flat rate income tax is payable by the employee but withheld by the employer. In the period 1994–2004 the flat rate was 26% with a relatively small tax-free allowance.[6] Since 2002 employers and employees have paid a modest compulsory fee to the Estonian unemployment insurance fund.[7] The Estonian tax and contribution system implies that all individuals with income above the tax-exempt amount pay the same marginal tax.

The extent of informal sector activities and tax evasion in Estonia appears to have fallen since the mid-1990s. The GDP exhaustiveness adjustments undertaken by Statistics Estonia suggest that the share of the informal sector in the Estonian GDP has fallen from 12% of GDP in 1997 to 8.3% of GDP in 2001 (Leetmaa & Vork 2004). Data for the share of wage earners receiving unreported income show a similar trend. Antila & Ylostalo (2002) estimate that 19% of the working age population received unreported income in 1998, falling to 10% in 2002. Surveys undertaken by the Estonian Institute of Economic Research (EKI) indicate that 19% of all working-age respondents received unreported wage income in 1999, while the share had fallen to 14% in 2004 (EKI 2005: 16).

Renoy *et al.* (2004) compare the extent of unreported work across the eight new EU countries from Central and Eastern Europe, the

[6] The tax rate was reduced to 24% in 2005 and to 23% in 2006. The monthly tax-exempt amount was gradually increased from 300 EEK or 19 EUR per month in 1994 to 1400 EEK or 89 EUR in 2004.

[7] In 2002-05 employers paid 0.5% and employees 1% of the gross wage bill to the unemployment insurance fund.

two candidate countries Romania and Bulgaria, and — in less detail — the 15 'old' EU members, using data from 2001 or a nearby year. The work income going unreported in Estonia is estimated to 8-9% of GDP with a falling trend during the 1990s. This places Estonia as one of the countries in Central and East European with the smallest informal economy. The Estonian informal sector is relatively large compared to that of Northern EU countries, but smaller than in most Southern European countries (Renoy *et al.* 2004: 24–30).[8]

Empirical analyses of tax compliance and evasion are constrained by a lack of reliable data as tax evasion is by definition unrecorded and therefore difficult to measure precisely. Andreoni *et al.* (1998) discuss four possible sources of tax evasion data, namely tax audits, surveys and interviews, tax amnesties and laboratory experiments. Tax audits can be randomized or undertaken based on a suspicion of unreported income. The randomized audits are generally considered the most reliable data source, but such compliance measurement programs are only undertaken in few countries. The advantage of survey data is that the method makes it easy to obtain additional background information on the individual taxpayer e.g. in terms of socioeconomic, demographic and attitudinal factors. The disadvantage of such data is that participants may overstate their compliance. Tax amnesty data provide information about non-compliance, but cannot be extended to the whole population of evaders. Information gathered via experiments miss factors that cannot be replicated outside the laboratory environment.

The problem of poor data quality also prevails in the Estonian case. No data from randomized audits are available, so less reliable data sources must be used. We address the data problems in two ways. First, we undertake a number of robustness tests to ensure that our main results are robust to changes in data and model specifications. Second, we use data from three different sources allowing a juxtaposition of the results.

---

[8] See also Schneider & Enste (2000) for estimates of the size of the shadow economy in a cross-country sample at the beginning of the 1990s.

Our analysis is based on three datasets, all measuring the prevalence of unreported income in different ways and with considerable uncertainty: [9]

(a) A survey of self-reported tax evasion by the Estonian Institute for Economic Research. Survey respondents are asked whether they have received 'envelope wages', the Estonian term for undeclared wage income. This self-reported measure of the incidence of tax evasion is likely to be smaller than the actual evasion.

(b) Compliance data from the audits of individuals by the Estonian Tax and Customs Board. The data comprise unreported labor income but also other forms of tax evasion. The data are subject to a selection bias, as the audited individuals are not chosen randomly, but based on tip-offs or auditing information from previous years.

(c) The Labour Force Survey of Statistics Estonia where respondents self-report their type of employment contract. The survey contains no direct information on tax evasion, but asks individuals about their type of employment contract. Tax evasion is likely to take place if an individual indicates that he or she works according to a verbal contract as such contracts are only legal for short employment spells. The survey data offer a rich set of background information about the respondents.

Two comments are in place here. First, the three datasets in our study contain information on, respectively, income subject to tax evasion, cf. (a) and (b), and income that is merely unregistered, cf. (c). In practice, it is difficult to distinguish between income remaining unregistered in order to evade taxation and income remaining unregistered for other reasons. Tax evasion is, however, likely to be the main objective behind most unregistered activities

[9] We also sought to use a dataset based on the Estonian Household Budget Survey collected by Statistics Estonia. There were many missing observations in the data and it was difficult to establish whether or not individual households had evaded income taxation. After some experimentation we decided to drop the Household Budget Survey from our analysis.

(Schneider & Enste 2000).[10] Second, firms may require that the salary be fully or partly unregistered and the individual can therefore not avoid tax evasion. The individual can, however, leave the firm and seek work elsewhere. Tax evasion is thus an individual choice — at least as long as there are employment possibilities both in the formal and informal sectors of the economy.

Andreoni *et al.* (1998) asserts that a serious shortcoming of most empirical work on tax evasion is that it is only loosely connected with theory; empirical studies can rarely be interpreted as tests of a specific theory. The purpose of this paper is to uncover the factors that characterize individuals who engage in payroll and income tax evasion in Estonia, largely to ascertain how evasion is distributed across income, education, gender, etc. The emphasis is thus not directly on testing any specific theory. Still, the relations between evasion and different explanatory variables can also lend support to different theoretical explanations of evasion behavior as explained in section 6.

3. EMPIRICAL RESULTS USING THE EKI SURVEY OF ENVELOPE WAGES

The Estonian Institute of Economic Research (EKI) has since 1999 conducted annual surveys on individuals receiving envelope wages (EKI 2005: 16). The survey results are chiefly used to estimate the overall size of the informal sector in Estonia. In 2004 the number of respondents was 744 with the sample broadly mirroring the distribution of the Estonian labor force in terms of gender, ethnicity, residence, region, age, level of education and income. Of the questioned individuals 514 (69%) stated that they worked, and 499 of them (97%) answered the question concerning the receipt of envelope wages.

The respondents were asked: 'Did you receive envelope wages in 2004?', and offered three answer possibilities: 'Yes, regularly',

[10] See also the discussion in Tanzi (1999) on the linkages between unreported economic activities and tax evasion.

'Sometimes' or 'Never'. In total 5% of the respondents answered that they received envelope wages regularly, 9% said they received it sometimes, and 86% that they did not at all receive envelope wages in 2004.

The setup of the question above leaves several possibilities for coding the dependent variable. One possibility is to use ordered logistic (or probit) regression, as the answers can be ordered using a logical scale from regular evaders to those who never received envelope wage. This ordering is questionable if there is not a consistent ranking from no evasion to occasional evasion and from occasional evasion to regular evasion. Thus, multinominal logistic regression could also be a suitable method for econometric analysis. We have chosen to use binary regression for our baseline estimations, as this method is the only applicable for the two other datasets. Robustness checks show that the choice of estimation method is of little importance for the results (see below). A dummy variable was constructed taking the value 0 if the respondent did not receive envelope wages during 2004, and 1 if the respondent regularly or sometimes received envelope wages during 2004.

As discussed in section 1, theory suggests that the decision to evade income taxation is determined by a range of factors. However, the structure of the Estonian tax system and the data available in the EKI dataset limit the set of variables used as explanatory variables in the estimations. The flat tax implies that virtually all taxpayers have the same marginal tax rate. The dataset contains no information on auditing probabilities and penalties facing the individual. The dataset does, however, contain a range of variables capturing the characteristics of the individual and his or her employer. The personal characteristics comprise ethnicity, education, gender, age and income as well as a regional dummy for residence. From the employer's side, the sector of production and the size of the firm are included. Among production sectors the field of activities is aggregated to construction, manufacturing, services, trade, agriculture and government.

The results of the logit model based on the EKI data are shown in Table 1. The impact of the explanatory variables on the probability of evasion is expressed as marginal effects. The marginal effects of

continuous and ordered explanatory variables were calculated for the average of the explanatory variable; for binary explanatory variables the marginal effect was calculated as the difference in probabilities for the extremes of the explanatory variable.

Table 1. EKI data: logit estimation of self-reported receipt of envelope wages, 2004

	Parameter estimate	***z*-statistic**	**Variable average**	**Marginal effect**
Northern Estonia (‘1’ Northern, ‘0’ other)	0.52	1.10	0.47	0.043
Central Estonia (‘1’ Central, ‘0’ other)	0.37	0.57	0.08	0.034
North-Eastern Estonia (‘1’ North-Eastern, ‘0’ other)	0.38	0.60	0.14	0.034
Western Estonia (‘1’ Western, ‘0’ other)	–0.09	–0.14	0.10	–0.007
Construction sector (‘1’ construction, ‘0’ other)	1.45**	2.35	0.08	0.191
Manufacturing sector (‘1’ manufacturing, ‘0’ other)	–0.04	–0.06	0.18	–0.003
Service sector (‘1’ services, ‘0’ other)	0.86*	1.67	0.35	0.079
Trading sector (‘1’ trading, ‘0’ other)	0.65	1.02	0.11	0.065
Agricultural sector (‘1’ agriculture, ‘0’ other)	1.47*	1.79	0.04	0.204
Firm size group 2 (‘1’ 5–19 employees, ‘0’ other)	0.36	0.73	0.20	0.033
Firm size group 3 (‘1’ 20–49 employees, ‘0’ other)	0.10	0.20	0.21	0.009
Firm size group 4 (‘1’ 50–249 employees, ‘0’ other)	–0.44	–0.82	0.30	–0.034
Firm size group 5 (‘1’ > 249 employees, ‘0’ other)	–0.90	–1.37	0.18	–0.058
Ethnicity (‘1’ Estonian, ‘0’ other)	–0.27	–0.70	0.71	–0.023
Secondary education (‘1’ secondary, ‘0’ other)	0.15	0.23	0.59	0.012
Tertiary education (‘1’ tertiary, ‘0’ other)	–0.59	–0.85	0.35	–0.045
Gender (‘1’ man, ‘0’ woman)	0.45	1.38	0.46	0.038
Respondent’s age group 2 (‘1’ 30–49, ‘0’ other)	–0.28	–0.80	0.46	–0.023

	Parameter estimate	*z*-statistic	Variable average	Marginal effect
Respondent's age group 3 ('1' 50–64, '0' other)	–1.10**	–2.23	0.24	–0.073
Respondent's age group 4 ('1' 65–74, '0' other)	–0.99	–1.21	0.06	–0.057
Income group 2 ('1' 1001–2000, '0' other) [a)]	1.12	1.31	0.20	0.121
Income group 3 ('1' 2001–3500, '0' other) [a)]	1.33	1.62	0.30	0.138
Income group 4 ('1' > 3500, '0' other) [a)]	1.46*	1.71	0.43	0.135
Constant	–3.67***	–2.99	..	..
Log likelihood	–152.33			
Pseudo R^2	0.130			
Number of observations	466			
Actual share of evaders in sample	0.125			
Predicted share of evaders	0.090			

***, ** or * denotes that the estimated coefficient is significantly different from 0 at, respectively, the 1%, 5%, and 10% level.
Note: Omitted variables are Southern Estonia, Government sector, Firm size 1 (1–4 employees), Non-Estonian ethnicity, Primary education, Woman, Age group 1 (18–29 years), Income group 1 (income less than 1000 EEK per month).
[a)] Monthly total income per household member in EEK, net of paid taxes (€ 1 = 15.65 EEK).

None of the regional dummies are significant. Among the sectoral dummies, construction, services and agriculture are positive and statistically significant, implying more evasion than in the government sector, which is the omitted variable. The probability of receiving envelope wages in the construction sector is 19.1%, in the services sector 7.9% and in the agricultural sector 20.4% higher than in the government sector. The variables depicting the size of the firm are not significant.[11]

[11] The EKI survey does not make it possible to ascertain whether the respondents refer to their work in the formal and/or in the informal sector when they answer questions on firm size, sector and income.

Among the individual personal characteristics only one variable for the respondent's age and one variable for the respondent's income are statistically significant. The group of middle-aged (50-64 years old) evades taxation less frequently than the reference group of young people. There may be a weak tendency that the likelihood of the respondent receiving envelope wages increases with income. Only the group with the highest income is significant and then only marginally at the 10% level. The three groups with the highest income (groups 2–4) exhibit essentially the same marginal effect.

The averages of the explanatory variables for the whole sample and for the sample used for estimation differ somewhat. The estimation sample contains fewer respondents from lower income groups than the original sample — and individuals with low income per household member report information about envelope wages less frequently. The differences between the two samples are, however, small and unlikely to affect the results markedly: 10% of individuals in the whole sample belong to the lowest income group, while the corresponding number is 7% in the estimation sample (see Table 1).

The results in Table 1 are derived using the groupings of the respondents' characteristics used in the original data from EKI. Because of the relatively low number of evaders in the sample, the many different characteristics, and the numerous groups with each characteristic, there will be cases where very few respondents belong to a particular group. The many small groups may also lead to large standard errors and thus be behind the relatively few significant parameters.

To address these concerns, we first sought to cut down on the number of explanatory variables by omitting the regional dummies. A Wald test failed to reject the hypothesis that these variables are jointly insignificant in the regression. Still, the results (not shown) were qualitatively identical to those in Table 1.

Reintroducing the regional dummies, we then estimated an alternative model with firm size, education, age and income as ordered variables. For example, the five dummy variables indicating the size of the firm were converted into one ordered variable taking the

values 1, 2, …, 5 depending on the size of the firm. This approach essentially imposes constraints across the grouped variables in order to addresses the problems discussed above. These constraints may or may not be warranted, but in this case where the parameters to the grouped variables are very imprecisely estimated, Wald tests cannot reject that the imposed constraints are valid. The results are shown in Table 2.

Table 2. EKI data: logit estimation of self-reported receipt of envelope wages with ordered explanatory variables, 2004

	Parameter estimate	***z*-statistic**	**Variable average**	**Marginal effect**
Northern Estonia (‘1’ Northern, ‘0’ other)	0.50	1.07	0.47	0.044
Central Estonia (‘1’ Central, ‘0’ other)	0.38	0.60	0.08	0.037
North-Eastern Estonia (‘1’ North-Eastern, ‘0’ other)	0.31	0.50	0.14	0.030
Western Estonia (‘1’ Western, ‘0’ other)	–0.10	–0.15	0.10	–0.008
Construction sector (‘1’ construction, ‘0’ other)	1.42**	2.36	0.08	0.193
Manufacturing sector (‘1’ manufacturing, ‘0’ other)	–0.03	–0.05	0.18	–0.003
Service sector (‘1’ services, ‘0’ other)	0.83*	1.66	0.35	0.080
Trading sector (‘1’ trading, ‘0’ other)	0.68	1.10	0.11	0.072
Agricultural sector (‘1’ agriculture, ‘0’ other)	1.38*	1.71	0.04	0.193
Firm size (‘1’ 1–4, ‘2’ 5–19, ‘3’ 20–49, ‘4’ 50–249, ‘5’ > 249 employees)	–0.26**	–2.13	3.23	–0.022
Ethnicity (‘1’ Estonian, ‘0’ other)	–0.27	–0.71	0.71	–0.024
Education (‘1’ primary, ‘2’ secondary, ‘3’ tertiary)	–0.49*	–1.71	2.28	–0.042
Gender (‘1’ man, ‘0’ woman)	0.39	1.22	0.46	0.034
Age group (‘1’ 18–29, ‘2’ 30–49, ‘3’ 50–64, ‘4’ 65–74)	–0.47**	–2.40	2.12	–0.040
Income group (‘1’ < 1001, ‘2’ 1001–2000, ‘3’ 2001–3500, ‘4’ > 3500) [a)]	0.26	1.39	3.10	0.022
Constant	–0.92	–0.87	..	..
Log likelihood	–155.79			

	Parameter estimate	*z*-statistic	Variable average	Marginal effect
Pseudo R^2	0.110			
Number of observations	466			
Actual share of evaders in sample	0.125			
Predicted share of evaders	0.095			

***, ** or * denotes that the estimated coefficient is significantly different from 0 at, respectively, the 1%, 5%, and 10% level.
Note: Omitted variables are Southern Estonia, Government sector, Non-Estonian ethnicity, Woman.
[a)] Monthly total income per household member in EEK, net of paid taxes.

The results and the properties of the model in Table 2 with ordered explanatory variables are broadly similar to those of the model in Table 1 with grouped variables. The marginal effects of variables statistically significant in both models are approximately the same size. In the model with ordered variables, there is a statistically significant negative relationship between the size of the firm and the likelihood of evasion. It appears that individuals with higher levels of education evade taxes less often, but the parameter estimate is only significant at the 10%-level. There is a negative and significant relation between age and evasion. Thus, these two models — with grouped or with ordered independent variables — produce essentially similar results from the EKI dataset.

As yet another robustness check of results of the model with grouped variables, we undertook an ordered logit estimation with the dependent variable taking three values: '1' if envelope wage was not received, '2' if envelope wage was received sometimes, '3' if envelope wage was received regularly. The result is shown in Table A.1 in Appendix 1. The results in Tables 1 and A.1 are very similar, indicating that the effect of the explanatory variables on the latent variable (tax evasion) are analogous irrespective of whether evasion is measured as a binary discrete variable or an ordered variable.

Overall, the results from the EKI dataset suggest that firm-side factors are important in explaining the prevalence of envelope

wages, while the importance of the employees' personal characteristics cannot be estimated precisely. This may indicate that the decision to receive all or some of the salary as envelope wages is to a large extent made by the firm and that the employees have little chance to influence that decision irrespective of personal characteristics such as education and gender. This corresponds to the survey respondents' attitude to envelope wages: 45% of the respondents receiving envelope wage were not pleased with the situation; 55% of them said that they would lose their job if they did not accept the envelope wage (EKI 2005).

Logit models are estimated using Maximum Likelihood methods and the reliability of the results, including the consistency of the parameter estimates, hinges on the model not being misspecified (Green 2000: sec. 19.4). The overall quality of our data compelled us to examine whether outliers were affecting results in an unduly manner. The Delta-Beta influence statistics test (Pregibon 1981) indicated that outliers were not of importance for any of the three models based on the EKI dataset.

A related concern was the possibility of heteroscedasticity. An LM test indicated that heteroscedasticity could be related to some of the binary explanatory variables in all three models using the EKI dataset. We undertook additional analyses to assess the importance of the heteroscedasticity problem and concluded that it is unlikely to affect the qualitative results. First, the LM test has very high power and this could result in many 'false alarms'. We reestimated the first model using probit and then undertook LR and Wald tests which generally indicated that the heteroscedasticity problems were unimportant in the reestimated probit model. Second, we calculated the Huber-White and the GLM robust standard errors and they were in all cases essentially identical to the ordinary standard errors presented in Tables 1, 2 and A.1. Third, some experimentation with removal of the binary explanatory variables, which LM tests suggested were responsible for the heteroscedasticity problems, changed results little.

4. EMPIRICAL RESULTS USING AUDITS OF THE ESTONIAN TAX AND CUSTOMS BOARD

The Estonian Tax and Customs Board undertakes regular audits of corporate and individual taxpayers. In this study we use data from the audits of individual taxpayers in 2002. The audits are non-random as individuals were only audited if the tax board had received a tip-off or for other reasons suspected tax evasion. The individuals we selected for auditing based on e.g. income and expenditure records, real property registrations, and criminal records.

A total of 2655 taxpayers were audited in 2002 amounting to 0.3% of all Estonian personal taxpayers.[12] Tax evasion was detected in 66% of the audits. The sample selection explains the high share of evasion among the audit subjects, but also raises some problems for the interpretation of the econometric analysis. Thus, the results of the analysis of the Tax Board data should not be interpreted as pertaining to the whole population but *only* to the group of individuals selected for auditing in 2002.

Table 3 presents the results of logit estimations using audit data from the Tax and Customs Board. The available data allowed us to include only few background variables. Most of the estimated coefficients are significant at the 1% level of significance. Among people selected for tax audits in 2002, the probability of uncovering tax irregularities is higher for individuals living in Northern, Central and North-Eastern Estonia and lower for individuals living in Western Estonia than for the individuals in Southern Estonia. We return in section 6 to possible interpretations of the regional dummies.

[12] By means of comparison, the proportion of audited individuals in the Taxpayer Compliance Measurement Program (TCMP) in the USA was 1.7% in 1995 (Andreoni *et al.* 1998: 820). In the TCMP, however, sampling is stratified random.

Table 3. Tax and Customs Board audit data: logit estimation of detected tax evasion, 2002

	Parameter estimate	***z*-statistic**	**Variable average**	**Marginal effect**
Northern Estonia ('1' Northern, '0' other)	2.47***	14.7	0.32	0.383
Central Estonia ('1' Central, '0' other)	2.45***	8.99	0.10	0.282
North-Eastern Estonia ('1' North-Eastern, '0' other)	0.74**	2.26	0.02	0.121
Western Estonia ('1' Western, '0' other)	–0.66***	–5.77	0.31	–0.136
Gender ('1' man, '0' woman)	0.16	1.55	0.62	0.033
Respondent's age	–0.090***	–2.79	43.27	–0.018
Respondent's age squared	$1.16e^{-3}$***	3.21	2002.6	$2.28e^{-4}$
Income [a)]	$-2.23e^{-3}$***	–3.78	57.86	$-4.40e^{-4}$
Constant	1.78***	2.51	..	..
Log likelihood	–1172.84			
Pseudo R^2	0.240			
Number of observations	2392			
Actual share of evaders in sample	0.65			
Predicted share of evaders	0.73			

***, ** ,* denote that the coefficient estimate is significantly different from 0 at, respectively, the 1%, 5%, and 10% level.
Note: Omitted variables are Southern Estonia, Woman.
[a)] Annual declared gross income per taxpayer, thousands EEK.

The coefficient of the gender variable is statistically insignificant. Higher income among individuals in the sample lowers the probability to be an evader; if the declared annual income increases by 10,000 EEK, the propensity to evade taxation decreases by 0.4%. From the quadratic relation between evasion and age, it can be seen that young people and old people are more prone to evasion than middle-aged persons. The finding that old people evade taxation more frequently than middle-aged persons is a somewhat surprising as it is generally found in advanced econo-

mies that old people are more compliant (Andreoni *et al.* 1998). This could be explained by the disadvantaged and rejected position of many elderly persons in the Estonian labor market, which may be explained by the fundamental changes to the economy resulting from the transition process. The finding broadly supports the results from Poland in Gardes & Starzec (2002).

An obvious concern is that the model using the audit data from the Tax and Customs Board suffers from problems stemming from omitted variables. We have available a very limited number of explanatory variables and the very high *z*-statistics could also indicate that the model is under-parameterized. In lieu of these problems, it is important to examine whether there remains systematic variation in the residuals from the regression shown in Table 3.

The lack of variables and the structure of the estimated model limit the possibilities of specification testing, but it is possible to undertake a Link test. The Link test is based on an auxiliary regression where the predicted values and the *squared* predicted values from the original regression are used as explanatory variables in an auxiliary logit estimation of individual tax evasion. The parameter to the squared predicted value is insignificant (*t*-value = –0.99, p = 0.324), so the Link test fails to reject the hypothesis that the model is specified correctly in this case. The result of the Link test may be taken to signify that the problem of an omitted variables bias will be relatively small.

A Pregibon (1981) Delta-Beta test for outliers indicates that there is only one outlier. Removing the observation generating the outlier makes no difference in terms of the parameter estimates and the significance levels for the explanatory variables. Outliers appear not to be a problem here. LM-tests indicated that heteroscedasticity cannot be ruled out, but as argued before the test might not be reliable as it might pick up other forms of misspecifications (Green 2000: 829–830). Experimentation with different non-linear specifications of the income variable did not alter the results qualitatively. We conclude that any heteroscedasticity problem is unlikely to affect the results substantially.

5. EMPIRICAL RESULTS USING THE ESTONIAN LABOUR FORCE SURVEY

Statistics Estonia regularly carries out a Labour Force Survey using the methodology of the International Labor Organization (Statistics Estonia 2005). This paper uses data from 2004 with 14,645 observations in the sample. Information about evasion of personal income taxation is ascertained indirectly from the following multiple-choice question in the survey: 'Did you work in this enterprise / organization under an employment contract, a contract of agreement, a 'Public Service Act' or according to a verbal contract?'

According to the survey the Estonian employment rate was 56.8% in 2004 among respondents aged 15–74. In total 2.7% of the employed respondents worked under a verbal contract, but 0.3% did not report other characteristics used as explanatory variables and were therefore dropped from the sample. The answer of working 'according to a verbal contract' implies that the work in most cases will remain unreported and taxation evaded. According to Estonian law, employees can only work under an oral contract if the duration of the work is shorter than two weeks. Almost all respondents answering that they worked according to a verbal contract also indicated that they had remained in the position for more than one month. Thus, these individuals presumably break the law — most likely to avoid taxation.[13] Schneider & Enste (2000) argue that the main objective behind most unregistered transactions is tax evasion. Guariglia & Kim (2004) advance the same argument and use unregistered employment as an indicator of evasion of taxation of the income from this employment.

The dependent variable in our analyses of the Labour Force Survey sample is undeclared work; the variable is equal to 1 if the

[13] The Labour Force Survey also asks respondents to state their gross and net wages. This information in combination with the Estonian flat rate income tax system allows us, in principle, to calculate the total evaded tax. About one-third of all employed respondents reported both their gross and net wage, but the calculated evaded amounts were often unreasonable.

respondent worked under a verbal contract; and 0 otherwise. In the sample used for analysis the share of unregistered work is 2.4%. Table 4 shows the results of a logit estimation using the undeclared work variable as the dependent variable. Most explanatory variables are entered as grouped variables, but the size of the firm was included as an ordered variable in order to reduce the otherwise large number of variables and to avoid sensitivity of estimates due to small amounts of observations in individual groups.

Table 4. Labour Force Survey: logit estimation of verbal work contract, 2004

	Parameter estimate	***z*-statistic**	**Variable average**	**Marginal effect**
Northern Estonia (‘1’ Northern, ‘0’ other)	0.74***	2.64	0.44	0.003
Central Estonia (‘1’ Central, ‘0’ other)	0.78**	2.40	0.10	0.004
North-Eastern Estonia (‘1’ North-Eastern, ‘0’ other)	–0.23	–0.48	0.13	–0.001
Western Estonia (‘1’ Western, ‘0’ other)	–0.80*	–1.70	0.10	–0.002
Construction sector (‘1’ construction, ‘0’ other)	3.00***	6.52	0.07	0.052
Manufacturing sector (‘1’ manufacturing, ‘0’ other)	1.28***	2.79	0.30	0.006
Service sector (‘1’ services, ‘0’ other)	0.65	1.41	0.18	0.003
Trading sector (‘1’ trading, ‘0’ other)	1.17**	2.53	0.13	0.007
Agricultural sector (‘1’ agriculture, ‘0’ other)	2.96***	6.32	0.04	0.055
Firm size (‘1’ 1–10, ‘2’ 11–19, ‘3’ 20–49, ‘4’ 50–99, ‘5’ 100–199, ‘6’ 200–499, ‘7’ 500–999, ‘8’ > 1000 employees)	–0.81***	–8.01	3.25	–0.003
Full-time work (‘1’ full-time, ‘0’ part-time)	–1.47***	–4.92	0.91	–0.010

	Parameter estimate	***z*-statistic**	**Variable average**	**Marginal effect**
Ethnicity (‘1’ Estonian, ‘0’ other)	–0.53**	–2.33	0.64	–0.002
Education level low (‘1’ level 2, ‘0’ other) [a)]	–1.25***	–5.50	0.57	–0.005
Education level mid (‘1’ level 3, ‘0’ other) [a)]	–1.43***	–2.97	0.11	–0.003
Education level high (‘1’ levels 4, 5 and 6, ‘0’ other) [a)]	–1.44***	–3.78	0.21	–0.004
Gender (‘1’ man, ‘0’ woman)	1.10***	4.44	0.47	0.004
Respondent’s age	–0.11**	–2.51	42.28	$-3.78e^{-4}$
Respondent’s age squared	$1.29e^{-3}$***	2.66	1950.1	$4.59e^{-6}$
Wage income (monthly, EEK net of paid tax)	$-1.02e^{-4}$**	–1.96	4458.9	$-3.62e^{-7}$
Constant	0.54	0.57	..	..
Log likelihood	–436.68			
Pseudo R^2	0.324			
Number of observations	5725			
Actual share of evaders in sample	0.024			
Predicted share of evaders	0.00357			

***, ** ,* denote that the coefficient estimate is significantly different from 0 at, respectively, the 1%, 5%, and 10% level.

Note: Omitted variables are Southern Estonia, Government sector, Part-time work, Non-Estonian ethnicity, Primary education, Woman.

[a)] Level of education according to ISCED 1997 classification (UNESCO 1997):

Level 1 – Primary education or first stage of basic education; ISCED level 0 is also included as there is only one observation with pre-primary education in the sample.

Level 2 – Lower secondary or second stage of basic education.

Level 3 – (Upper) secondary education.

Level 4 – Post-secondary non-tertiary education.

Level 5 – First stage of tertiary education.

Level 6 – Second stage of tertiary education.

Except for the regional dummy for North-Eastern Estonia and the service sector dummy, all explanatory variables are statistically significant at the 10% level of significance or better. Working in construction, manufacturing, trade or agriculture raises the

probability of being an evader compared to working in government. The two sectors where working under verbal contract is most common are construction and agriculture; in both sectors the probability to be an evader is about 5% higher than for individuals working in the government sector. An important factor from the employer's side is also the size of the firm; more employees working in a firm lower the probability of people working under a verbal contract. Among the geographical regions, the probability to work under verbal contract is higher in Northern and Central parts of Estonia; and lower in Western parts of the country. The omitted variable is as in the previous analysis the Southern region.

Being employed full-time reduces the probability of evasion by 1%.[14] Estonian ethnicity and higher level of education lower the probability to be an evader, and male probability to be an evader is on average 0.4% higher than female. Similarly to Tax and Customs Board audits data there is a quadratic convex relation between evasion and age. Thus, the young and old people are more likely to work under a verbal contract than the middle-aged. The probability to be an evader decreases with increasing net wage income.

We were concerned that the use of dummy (grouped) variables for the education level could give misleading results. We constructed an ordered education variable and redid the estimation, but obtained results essentially analogous to those presented in Table 4.

A potential cause of concern using the Labour Force Survey data is the unbalanced sample with a very small share of individuals with unreported employment (2.4%). An unbalanced sample does not affect the consistency or efficiency of the estimated parameters, but the model selection is complicated by two factors (Cramer 1999). First, the detection of outliers is difficult. In a discrete model, a small value of the estimated probability of the individual observation indicates an outlier. In the case of an unbalanced sample the likelihood of an individual observation to be an outlier is not equal within two outcome sets: the smaller outcome group's

[14] Recall that for binary explanatory variables the marginal effect has been calculated as the difference in probabilities at the binary extremes.

observations are more often picked out as outliers since their predicted probabilities are lower. The Pregibon (1981) Delta-Beta influence statistic test, however, does not indicate any outliers. The second complication stems from assessing the estimated model based on its forecasting properties. Appendix 2 shows that the model's underprediction of evasion is the mainly the result of the unbalanced sample and thus unlikely to stem from misspecification of the mode. We conclude, overall, that the unbalanced sample is not a major concern for the selection and interpretation of the model based on the Labour Force Survey.

6. COMPARISON OF RESULTS ACROSS DATASETS

The purpose of this paper is to determine the characteristics of individuals engaging in payroll and income tax evasion in Estonia. To address a lack of reliable data, we have used data from three different data sources. Each dataset has its own strengths and weaknesses; brought together the results may help provide a broad picture of the prevalence of tax evasion and unreported work in Estonia. Table 5 contrasts the results from the three different datasets. A significant positive relationship is indicated with a (+) and a significant negative relationship is indicated with a (–).

It follows from Table 5 that the estimated share of evasion is around 14% for self-reported receipt of envelope wages; the share of unreported income is 66% in the non-random sample of audits from the Tax and Customs Board; the share of employees who work according to a verbal contract is 2.7% according to the Labour Force Survey. The very different evasion frequencies in the three samples are the result of the different definitions of the dependent variable and different sampling methodologies. The detailed results for each dataset were discussed in sections 3–5.

Table 5 reveals that the qualitative results are surprisingly similar across the three datasets. The signs of the marginal effects are relatively congruent across the three datasets — in spite of the described data limitations, the somewhat different contents of the

dependent variables, the different sampling methodologies, and the numerous methodological and econometric problems.

Table 5. Effects of explanatory variables on tax evasion across different data sources

	EKI Envelope Wages Survey 2004	**Tax and Customs Board Audits 2002**	**Labour Force Survey 2004**
Measurement of evasion	Self-reported receipt of envelope wages	Non-random audits of individual taxpayers	Self-reported work under verbal contract
Share of evaders	14.4%	66%	2.7%
Share of evaders in regression	12.5%	65%	2.4%
Northern Estonia	~	(+)	(+)
Central Estonia	~	(+)	(+)
North-Eastern Estonia	~	(+)	~
Western Estonia	~	(–)	(–)
Construction sector	(+)	..	(+)
Manufacturing sector	~	..	(+)
Service sector	(+)	..	~
Trading sector	~	..	(+)
Agricultural sector	(+)	..	(+)
Firm size	(–)	..	(–)
Full-time work	..	..	(–)
Estonian ethnicity	~	..	(–)
Education level	~	..	(–)
Gender 'man'	~	~	(+)
Age	(–)	(–)	(–)
Age squared	..	(+)	(+)
Income	~	(–)	(–)

(+) indicates a positive and statistically significant relation.
(–) indicates a negative and statistically significant relation.
~ indicates that the variable is not statistically significant in the model.
.. indicates that no information is available on the item in the dataset.

Living in Northern and Central parts of Estonia increases the probability of evasion, while living in the Western part of the country decreases the probability of evasion. Several interpretations of the regional dummies are possible:

- The region dummies could simply be considered control variables proxying for some unobserved effects. The parameter estimates are then of little importance.
- The regional dummies could account for different socio-economic conditions in various regions in Estonia. We have tried to replace the regional dummies with a variable comprising regional unemployment rates but the variable was insignificant. However, when we replaced the regional dummies with regional GDP per capita, the GDP variable attained a positive parameter and became highly significant in the models based on data from the Tax and Customs Board and from the Labour Force Survey. A possible — but rather speculative — interpretation is that whereas tax evasion and unreported work are most prevalent among relatively disenfranchised individuals (e.g., those with low income), then the possibilities for attaining unregistered work might be better in the relatively affluent parts of Estonia than in poorer areas of the country.[15]
- The significant coefficients of the regional dummies may also reflect different social norms across different parts of Estonia. Northern and Central Estonia experienced rapid economic modernization and social fragmentation, while Western Estonia has remained a rural and 'traditional' region. This explanation would be consistent with theories predicting that social norms toward tax evasion could differ across regions or countries.[16]

[15] Rosser *et al.* (2000) provide some evidence (from the early transition period) showing that the transition countries with highest income inequality are those were informal economy activities are most prevalent.

[16] Alm & Torgler (2006) present evidence showing that the tax morale differs markedly between Europe and the USA, but also across European countries.

There is evidence that tax evasion is more frequent in the construction and agricultural sectors and possibly also the services sector. The tax evasion in these sectors may be a result of low detection probabilities given the nature of the business conducted. Smaller firms are more likely both to pay envelope wages and to hire employees on verbal contracts. This result may be explained by an expected higher probability of audits for bigger companies and/or by higher penalties after auditing, as the costs due to probable closure or loss of reputation are higher for larger companies.

In terms of age there is U-shaped relation between age and the prevalence of payroll and income tax evasion. Relatively young and old people are more likely to evade than are middle-aged persons. The relatively high occurence of tax evasion among elderly is unusual in an international context and may be a result of the Estonian transition process, where especially older persons have experienced problems getting jobs in the formal economy. Two datasets point to a negative relation between income and tax evasion, while the EKI dataset yields no significant parameters in this regard. One dataset (the Labour Force Survey) suggests that unreported work is more common among individuals who work part-time, are of non-Estonian ethnicity, have relatively short education, and/or who are men.

Returning to the determinants of tax evasion and unreported work in empirical analyses, we attain only mixed support for the main findings (Andreoni *et al.* 1998, cf. section 1). (i) Tax evasion appears to be a declining function of income in Estonia while it is frequently an increasing function in high-income countries. (ii) Data limitations meant that we could not test how auditing and penalty schemes affect tax evasion in Estonia. (iii) The regional patterns in tax evasion in Estonia may lend support to social norms and customs being of importance. (iv) Individual background variables are important.

Overall, our analyses on data from 2002-04 are broadly in line with the existing literature from the mid-1990s on tax evasion in transition economies. The relatively disenfranchised appear to be most likely to evade payroll and income taxation in Estonia. A

similar pattern is found in earlier studies for other transition countries (Kolev 1998, Gardes & Starzes 2002, Kim 2005). In other words, in spite of almost a decade between the previous studies and the study, the overall functioning of tax evasion and informal employment remains unchanged. The informal economy still functions as a partial safety net for persons who cannot find employment and income opportunities within the formal economy.

These findings have important policy implications. Estonia has during the last decade taken numerous steps to strengthen its revenue collection, improve the auditing system and crack down on tax evasion. These steps have likely helped Estonia retain a position as one of the Central and Eastern European countries with the least tax evasion. Our analysis suggests that a crack down on tax evasion also has the potential to harm the relatively disenfranchised in society. Further strengthening of the auditing and penalty schemes may thus be more successful if accompanied by steps making it easier for disenfranchised persons to gain access to formal sector employment, skills upgrading or social assistance.

The results in this paper are clearly circumscribed by considerable uncertainty stemming from the underlying data being unrepresentative or inaccurate; tax evasion is per definition difficult to measure. We believe that our approach of using three different datasets has provided additional insights and more reliable results. Until randomized audits become available, the results obtained in this paper comprise the most accurate picture of the determinants of tax evasion in Estonia.

ACKNOWLEDGEMENTS

The authors would like to thank Ross Chambers for research assistance and Aurelijus Dabusinskas, Dmitry Kulikov, and Tonu Roolaht for useful comments to earlier versions of the paper. They also benefited from comments by the discussant, Alberto Zanardi, and other conference participants at the 8th INFER Annual Conference, Cork, 2006. The authors alone are responsible for all interpretations and remaining errors. The views expressed are those of the authors and do not represent official views of the institutions in which they work. Marje Josing and Evelin Ahermaa from the Estonian Institute of Economic Research and Jaanus Laane from the Estonian Tax and Customs Board were helpful providing data for our research. This research has been supported by grants from EuroFaculty, Riga, Latvia; EuroCollege at Tartu University, Estonia; and the University of Nebraska at Omaha, U.S.A.

LITERATURE

Allingham, Michael G. & Agnar Sandmo (1972): 'Income Tax Evasion: A Theoretical Analysis', *Journal of Public Economics*, 1(3–4): 323–338.

Alm, James & Benno Torgler (2006): 'Culture Differences and Tax Morale in the United States and in Europe', *Journal of Economic Psychology*, 27(2): 224–246.

Andreoni, James, Brian Erard & Jonathan S. Feinstein (1998): 'Tax Compliance', *Journal of Economic Literature*, 36(2): 818–860.

Antila, Juha & Pekka Ylostalo (2003): 'Working Life Barometer in the Baltic Countries 2002', Finnish Ministry of Labour, Labour Policy Studies. http://www.sm.ee/est/HtmlPages/baromeeter2002/$file/baromeeter2002.pdf.

Cowell, Frank A. (1990): *Cheating the Government: The Economics of Evasion*, Cambridge, M.A.: MIT Press.

Cramer, J. S. (1999): 'Predictive Performance of the Binary Logit Model in Unbalanced Samples', *Journal of the Royal Statistical Society (Series D): The Statistician*, 48(1): 85–94.

EKI (2005): 'Varimajandus Eestis' [The Estonian Shadow Economy], Tallinn, Estonia: Estonian Institute of Economic Research, January. http://www.ki.ee/publikatsioonid/valmis/Varimajandus_Eestis_2004(elanike_hinnangute_alusel).pdf.

Eurostat (2006): 'Tables: GDP per Capita in PPS'. http://epp.eurostat.cec.eu.int/.

Gardes, Francois & Christophe Starzec (2002): 'Polish Households between Transition and Informal Markets', mimeo, http://www.rennes.inra.fr/jma2002/pdf/gardes.pdf.

Greene, William (2000): *Econometric Analysis*, Prentice Hall, 4th ed.

Guariglia, Alessandra & Byung-Yeon Kim (2004): "Earnings uncertainty, precautionary saving, and moonlighting in Russia", *Journal of Population Economics*, 17(2): 289–310.

Hanousek, Jan & Filip 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.

Johnson, Simon, Daniel Kaufmann, John McMillanc & Christopher Woodruff (2000): Why Do Firms Hide? Bribes and Unofficial Activity after Communism, *Journal of Public Economics*, 76(3): 495–520.

Kim, Byung-Yeon (2005): 'Poverty and Informal Economy Participation. Evidence from Romania', *Economics of Transition*, 13(1): 163–185.

Kolev, Alexandre (1998): 'Labour Supply in the Informal Economy in Russia during Transition', CEPR Discussion Paper no. 2024, London: Centre for Economic Policy Research.

Leetmaa, Reelika & Andres Vork (2004): 'Estonia', *European Employment Observatory Review*, Autumn 2004: 83–88, European Commission.

Ministry of Finance (2006): 'Summary of the Taxation System', Estonian Ministry of Finance. http://www.fin.ee/?id=3814

Neef, Rainer & Manuela Stanculescu (2002, eds.): *The Social Impact of Informal Economies in Eastern Europe*, Aldershot: Ashgate Publishing.

Pregibon, Daryl (1981): 'Logistic Regression Diagnostics', *Annals of Statistics*, 9(4): 705–724.

Renoy, Piet, Staffan Ivarsson, Olga van der Wusten-Gritsai & Emco Meijer (2004): *Undeclared Work in an Enlarged Union. An Analysis of Undeclared Work: An In-Depth Study of Specific Items*, European Commission, Directorate-General for Employment and Social Affairs. http://europa.eu.int/comm/employment_social/employment_analysis/work/undecl_work_final_en.pdf

Rosser, Barkley, Marina Rosser & Ehsan Ahmed (2000): 'Income Inequality and the Informal Economy in Transition Economies', *Journal of Comparative Economics*, 28(1): 156–171.

Schneider, Friedrich & Dominik H. Enste (2000): 'Shadow Economies: Size, Causes and Consequences', *Journal of Economic Literature*, 38(1): 77–114.

Staehr, Karsten (2004): 'The Economic Transition in Estonia. Background, Reforms and Results, in Rindzeviciute, Egle (ed.): *Contemporary Change in Estonia*, Huddinge: Sodertorns hogskola, Baltic & East European Graduate School, 37–67.

Statistics Estonia (2005): 'Labour Market 2004', Tallinn: Statistical Office of Estonia.

Tanzi, Vito (1999): 'Uses and Abuses of Estimates of the Underground Economy', *Economic Journal*, 109(June): 338–347.

Torgler, Benno (2003): 'Tax Morale in Transition Countries', *Post-Communist Economies*, 15(3): 357–381.

UNDP (2001): *Estonian Human Development Report 2001*, United Nations Development Program.

UNESCO (1997): 'International Standard Classification of Education' (ISCED), United Nations Educational, Scientific and Cultural Organization. http://www.uis.unesco.org/TEMPLATE/pdf/isced/ISCED_A.pdf.

[All links valid 14 December 2006.]

APPENDIX 1

Table A.1. EKI data: ordered logit estimation of receipt of self-reported envelope wages, 2004

	Para-meter estimate	***z*-statistic**	**Marginal effect: '3' evading regularly**	**Marginal effect: '2' evading sometimes**
Northern Estonia ('1' Northern, '0' other)	0.44	0.95	0.012	0.024
Central Estonia ('1' Central, '0' other)	0.33	0.52	0.010	0.020
North-Eastern Estonia ('1' North-Eastern, '0' other)	0.50	0.79	0.016	0.031
Western Estonia ('1' Western, '0' other)	–0.18	–0.27	–0.005	–0.009
Construction sector ('1' construction, '0' other)	1.64***	2.69	0.091	0.137
Manufacturing sector ('1' manufacturing, '0' other)	0.02	0.04	0.001	0.001
Service sector ('1' services, '0' other)	0.85	1.64	0.027	0.051
Trading sector ('1' trading, '0' other)	0.73	1.15	0.027	0.048
Agricultural sector ('1' agriculture, '0' other)	1.44*	1.76	0.077	0.119
Firm size group 2 ('1' 5–19 employees, '0' other)	0.39	0.78	0.012	0.023
Firm size group 3 ('1' 20–49 employees, '0' other)	0.07	0.14	0.002	0.004
Firm size group 4 ('1' 50–249 employees, '0' other)	–0.50	–0.94	–0.013	–0.025
Firm size group 5 ('1' 249 >employees, '0' other)	–1.03	–1.57	–0.021	–0.044
Ethnicity ('1' Estonian, '0' other)	–0.31	–0.82	–0.009	–0.018
Secondary education ('1' secondary, '0' other)	0.12	0.20	0.003	0.007

	Parameter estimate	***z*-statistic**	**Marginal effect: '3' evading regularly**	**Marginal effect: '2' evading sometimes**
Tertiary education ('1' tertiary, '0' other)	–0.63	–0.92	–0.016	–0.032
Gender ('1' man, '0' woman)	0.45	1.38	0.013	0.025
Respondent's age group 2 ('1' 30–49, '0' other)	–0.39	–1.11	–0.011	–0.021
Respondent's age group 3 ('1' 50–64, '0' other)	–1.14**	–2.31	–0.025	–0.050
Respondent's age group 4 ('1' 65–74, '0' other)	–1.05	–1.29	–0.019	–0.040
Income group 2 ('1' 1001–2000, '0' other) [a)]	1.09	1.28	0.042	0.074
Income group 3 ('1' 2001–3500, '0' other) [a)]	1.25	1.54	0.046	0.082
Income group 4 ('1' > 3500, '0' other) [a)]	1.51*	1.78	0.049	0.090
Cut 1	3.50	1.21	..	..
Cut 2	4.71	1.23	..	..
Log likelihood	–188.81	..	..	..
Pseudo R^2	0.114	..	..	..
Number of observations	466	..	..	..
Actual share of evaders	..	..	0.045	0.079
Predicted share of evaders	..	..	0.028	0.061

***, ** ,* denote that the coefficient estimate is significantly different from 0 at, respectively, the 1%, 5%, and 10% level.

Averages of explanatory variables are the same as reported in Table 1.

Note: Omitted variables are Southern Estonia, Government sector, Firm size 1 (1–4 employees), Non-Estonian ethnicity, Primary education, Woman, Age group 1 (18–29 years), Income group 1 (income less than 1000 EEK per month).

[a)] Monthly total income per household member in EEK, net of paid taxes (€ 1 = 15.65 EEK).

APPENDIX 2

A very large discrepancy between the actual share of unregistered work in the sample and the predicted share can either be the result of model misspecification or the consequence of using an unbalanced sample. With an unbalanced sample, the estimated prediction probabilities are usually found to overpredict for the greater share in the sample and underpredict for the smaller share (Cramer 1999).

The model in Table 4 predicts a share of evaders equal to 0.36% against 2.4% in the sample. To examine whether or not this underprediction is mainly a consequence of the unbalanced sample (and hence innocuous), we construct a *reweighted sample* with equal shares of individuals with undeclared and declared work. The construction of the reweighed balanced sample proceeds as follows: In total 135 observations (the number of individuals with undeclared work) are drawn randomly from the share of individuals with declared work. The 135 randomly drawn individuals with declared work are added to the 135 individuals with undeclared work. The 50/50 balanced sample comprises in total 270 observations. The marginal effects with respect to the explanatory variables should be the same in the original sample and the rebalanced sample.

Table A.2 shows the estimation results with the rebalanced sample. The rebalanced model has adequate prediction properties while the qualitative results are otherwise qualitatively unchanged. We conclude that the discrepancy between the actual and the predicted shares of individuals with unregistered work is mainly stem from the unbalanced sample and, consequently, is not a sign of misspecification.

Table A.2. Labour Force Survey: logit estimation of verbal work contract using 50/50 rebalanced sample, 2004

	Parameter estimate	*z*-statistic	Variable average	Marginal effect	
				50/50 sample	Original sample
Northern Estonia (‘1’ Northern, ‘0’ other)	0.47	0.71	0.49	0.113	0.008
Central Estonia (‘1’ Central, ‘0’ other)	1.54**	2.05	0.14	0.364	0.046
North-Eastern Estonia (‘1’ North-Eastern, ‘0’ other)	–0.51	–0.53	0.10	–0.118	–0.007
Western Estonia (‘1’ Western, ‘0’ other)	1.00	0.97	0.05	0.244	0.026
Construction sector (‘1’ construction, ‘0’ other)	5.10***	4.57	0.16	0.244	0.026
Manufacturing sector (‘1’ manufacturing, ‘0’ other)	1.96**	2.34	0.21	0.451	0.061
Service sector (‘1’ services, ‘0’ other)	0.89	1.07	0.21	0.218	0.019
Trading sector (‘1’ trading, ‘0’ other)	1.97**	2.17	0.14	0.446	0.071
Agricultural sector (‘1’ agriculture, ‘0’ other)	4.39***	4.07	0.12	0.679	0.435
Firm size (‘1’ 1–10, ‘2’ 11–19, ‘3’ 20–49, ‘4’ 50–99, ‘5’ 100–199, ‘6’ 200–499, ‘7’ 500–999, ‘8’ > 1000 empl.)	–1.03***	–5.01	2.47	–0.248	–0.104
Full-time work (‘1’ full-time, ‘0’ part-time)	–0.28	–0.45	0.84	–0.069	–0.005
Ethnicity (‘1’ Estonian, ‘0’ other)	–1.54***	–2.93	0.65	–0.367	–0.034
Education level low (‘1’ level 2, ‘0’ other) [a)]	–0.47	–0.82	0.50	–0.114	–0.008
Education level mid (‘1’ level 3, ‘0’ other) [a)]	–1.98*	–1.90	0.10	–0.353	–0.017
Education level high (‘1’ levels 4, 5 and 6, ‘0’ other) [a)]	–0.57	–0.64	0.14	–0.132	–0.008

	Parameter estimate	*z*-statistic	Variable average	Marginal effect	
				50/50 sample	Original sample
Gender (‘1’ man, ‘0’ woman)	2.22***	4.15	0.59	0.476	0.036
Respondent’s age	–0.21**	–2.03	41.33	–0.050	–0.003
Respondent’s age squared	$2.85e^{-3}$***	2.37	1931.8	$6.89e^{-4}$	$1.94e^{-7}$
Wage income (monthly, EEK net of paid tax)	$-2.70e^{-4}$**	–2.46	3879.4	$-6.54e^{-5}$	$-1.20e^{-5}$
Constant	4.02**	1.99	..		
Log likelihood	–82.96				
Pseudo R^2	0.527				
Number of observations	270				
Actual share of evaders in rebalanced sample	50				
Predicted share of evaders				0.41	0.016

***, ** ,* denote that the coefficient estimate is significantly different from 0 at, respectively, the 1%, 5%, and 10% level.

Note: Omitted variables are Southern Estonia, Government sector, Part-time work, Non-Estonian ethnicity, Primary education, Woman.

[a)] Education levels as in Table 4.

KOKKUVÕTE

Miks indiviidid hoiavad kõrvale üksikisiku- ja sotsiaalmaksust Eestis?

Maksudest kõrvale hoidmise ulatus ja jagunemine erinevate maksumaksjate lõikes mõjutab maksusüsteemi efektiivsust ja maksukoormuse jagunemist. Täpsem teadmine sellest, kes maksudest kõrvale hoiavad, võimaldab hinnata erinevate maksude mõju ja annab vajalikku taustinformatsiooni maksude kavandamisel ja reformimisel, auditeerimisel ning trahvide kehtestamisel. Käesolevas artiklis kasutatakse kolme erinevat individuaalkirjetega andmebaasi, leidmaks tegureid, mis iseloomustavad üksikisiku- ja sotsiaalmaksust kõrvale hoidmist Eestis. Maksudest kõrvale hoidmist on üleminekuriikide andmete põhjal vähe uuritud. Peamiseks takistuseks sellelaadsetes uuringutes on andmete puudumine või nende madal usaldusväärsus. Selle probleemi leevendamiseks kasutataksegi käesolevas artiklis kolme erinevat andmebaasi ning kõrvutatakse saadud tulemusi.

Eesti Konjunktuuriinstituudi (2004), Maksuameti (2002) ja Eesti Statistika tööjõu-uuringu (2004) individuaalkirjetega andmebaasidе põhjal kasutatakse logistilisi mudelid ning hinnatakse erinevate karakteristikute mõju maksudest kõrvale hoidmise tõenäosusele (marginaalsed efektid). Kasutatud andmebaasid on koostamispõhimõtete ja maksudest kõrvalehoidmist iseloomustavate muutujate osas erinevad. Konjunktuuriinstituudi isikküsitlus sisaldab infot ümbrikupalkade saamise, Eesti Statistika tööjõu-uuringu isikküsitlus deklareerimata töö kohta ning erinevalt eelmisest kahest üldkogumile mittelaienevad Maksuameti registri andmed auditeerimise tulemusi. Vaatamata erinevustele annavad nende kolme andmebaasi põhjal tehtud arvutused sarnased tulemused.

Uuringu kohaselt on Eestis üksikisiku- ja sotsiaalmaksust kõrvale hoidmine enam levinud väikestes ettevõtetes ning ehitus- ja põllumajandussektoris. Indiviidide karakteristikute lõikes hoiavad maksudest kõrvale enam osa-ajaga töötajad, mitte-eestlased vähese hariduse ja madala sissetulekuga indiviidid ning mehed. Maksudest kõrvale hoidmine on enam levinud noorte ja vanade, mitte kesk-

ealiste, indiviidide lõikes. Maksudest kõrvale hoidmise tõenäosus sõltub ka indiviidi elukohast regioonide lõikes, nt on see tõenäosus kõrgem Harjumaal ja madalam Läänemaal elavatel indiviididel. Kokkuvõtvalt võib välja tuua, et Eestis on üksikisiku- ja sotsiaalmaksust kõrvale hoidmise tõenäosus suurem tööturul ebasoodsas olukorras olevatel indiviididel.