MODELLING OF HOUSEHOLD BEHAVIOUR IN TERMS OF SHADOW ECONOMY: LABOUR SUPPLY

by

Khmilevska Nataliya

A thesis submitted in partial fulfilment
of the requirements for the degree of

Master of Arts in Economics

Economics Education and Research
Consortium

MA program in Economics

National University of "Kyiv-Mohyla Academy"

2002

Approved by _____
Chairperson of Supervisory Committee

_____

_____

_____

Program Authorized
to Offer Degree _____

Date _____

Economics Education and
Research Consortium
MA program in Economics

National University of "Kyiv-Mohyla Academy"

Abstract

MODELLING OF HOUSEHOLD BEHAVIOUR IN TERMS OF SHADOW ECONOMY: LABOUR SUPPLY

by Khmilevska Nataliya

Chairperson of the Supervisory Committee:Professor Serhiy Korablin,
Institute for Economic Forecasting at the National Academy of Sciences of Ukraine

The model describes the household behaviour in terms of labour allocation into the official and unofficial economic sectors. A decision on the particular labour efforts in both sectors is designed as a constrained optimisation of a worker's utility, derived from the composite consumption good and the hours of leisure left. A government is implicitly involved into the modelling by introducing the marginal tax rates in the regular and irregular sectors. The analysis of tax evasion decisions reveals the major determinants of worker's unlawful behaviour. Finally, the model results suggest that the labour-hours in the unofficial sector are more sensitive to the changes of the gross hourly wage rate and the marginal tax rate in the regular sector than to the enforcement level, designed as a marginal tax rate in the unofficial sector.

ACKNOWLEDGMENTS

The author wishes to thank professor Michael Hemesath, the thesis advisor, for constant attention and encouragement of new ideas. I also want to thank professor Gene Ellis for the inspiration and confidence in my forces. The work owes much to the valuable comments of professor Ronald Johnson. The thesis would not be completed without the support of the project UEPLAC, Tacis and, personally, Petro Morgos. I am also grateful to professors Ghaffar Mughal, Stefan Lutz and Irina Lukyanenko for being considerate and supportive throughout the process of the preparation.

TABLE OF CONTENTS

LIST OF FIGURES

Number *Page*

LIST OF TABLES

GLOSSARY

Household – one person or a group of people, who reside in the same living quarters and perform joint housekeeping (share expenditures for living, eating, etc.).[1]

Able-bodied citizens – male at age 16-59 and female at age 16-54.[1]

Official unemployment rate – ratio of registered unemployed people to able-bodied citizens.[1]

Unemployed by ILO definition – people, who simultaneously satisfy the following three criteria:

(1) They are not employed or do not earn income;
(2) They have been actively looking for a job during the preceding four weeks before the survey or tried to start own business;
(3) They are ready to start working during the nearest two weeks.

Official GDP – part of the total economy, which corresponds to the production volumes calculated by the official statistical bodies according to the reported economic activities and partially accounting for the unreported economic activities.[2]

Total GDP – includes the reported and unreported economic activities in a country.

[1] Definition by State Committee of Statistics of Ukraine

[2] Definition from System of National Accounts

Introduction

Although we often hear a term "shadow economy" applied to the current economic situation in Ukraine and other transition economies, we never assimilate ourselves with the unofficial sector. Nevertheless, the representative household survey used in the thesis suggests that three quarters of able-bodied citizens of Ukraine are involved in tax evasion. This thesis attempts to answer the question how do we decide to allocate labour efforts to the unofficial sector and what should the government implement in order to reduce the scope of the unofficial economic activities.

To begin with, it is necessary to clarify a definition of "shadow economy". Opposite to the meaning of the term relevant to the developed and Latin America countries, "shadow sector" in transition economies relates to tax evasion. Hence, a term of "shadow economy" applied in the thesis pertains to the legal economic activities, deliberately hidden from the state authorities in order to evade or understate due tax payments.

According to various surveys (Kaufmann and Kaliberda [1996], Schneider and Enste [1998], Johnson, Kaufmann and Zoido-Lobaton [1998], Eilat and Zinnes [2000]) conducted by prominent economists in FSU (former Soviet Union) countries, a "shadow" economy share varies from 2 to 261% of a country's GDP and there is no cogent evidence that an expansion of the unofficial sector will halt in the near future. The very existence of "shadow" economy implies deterrent development environment and failure of government economic policy. In addition, the essential funds for social maintenance may become exhausted; moreover, the role of the government as a regulatory body maintaining social system and voluntary tax compliance is undermined. The scope of "shadow economy" is stipulated by the existing incentive structure, so that a large share of "shadow" economy is a manifestation of social discontent. The most transparent consequences of the informal sector activities involve the deficit of budget revenues and excessive tax burden for the economic entities functioning in the formal sector. When a

single country's economy splits into two sectors regulated by the different rules, economic development, new technological implementation, social equality of the citizens are obviously unattainable.

Although it may seem from the term of "shadow economy" that this phenomenon is undesirable for a society, it has its own benefits as well as costs. On one hand, shadow economy creates additional job placements, raises private consumption through enlarged family earnings and greater variety of shadow goods and services. Consequently, it may enhance the living standards of the country's population. On the other hand, the spread of shadow economy impedes future economic development due to:

1) Falling tax revenues to finance consolidated budget.
2) Undermining the government social function to supply the necessary services and maintain the minimum living standards of the citizens.
3) Raising tax burden for the official enterprises, and eventually forcing them to choose among bankruptcy and moving toward the shadow sector.

Specific feature of the phenomenon is its spiral nature. Shortage of the necessary tax revenues to provide services to the legitimate enterprises prompts the government to raise taxes. High taxes combined with the insufficient amount of public services make "unofficialdom" more attractive environment for businesses to develop. Shadow sector becomes even more advantageous owing to unequal conditions for the legitimate and hidden manufactures, when the latter do not pay taxes and can maintain comparatively lower price level, which is not feasible for the lawful businesses.

The benefits of shadow economy make it worthwhile to divert economic activities from the unofficial to legal sectors rather than aiming at eradication of the shadow economic activities. To build up an efficient incentive framework for the hidden businesses, a state should adopt a set of reforms such as transparent and equitable regulation and tax system, lower tax rates and sufficient amount of public services devoted to the legitimate

economic agents, proportional to present contributions and timely paid pension benefits, enforcement of legal contracts, protection from criminal elements, etc.

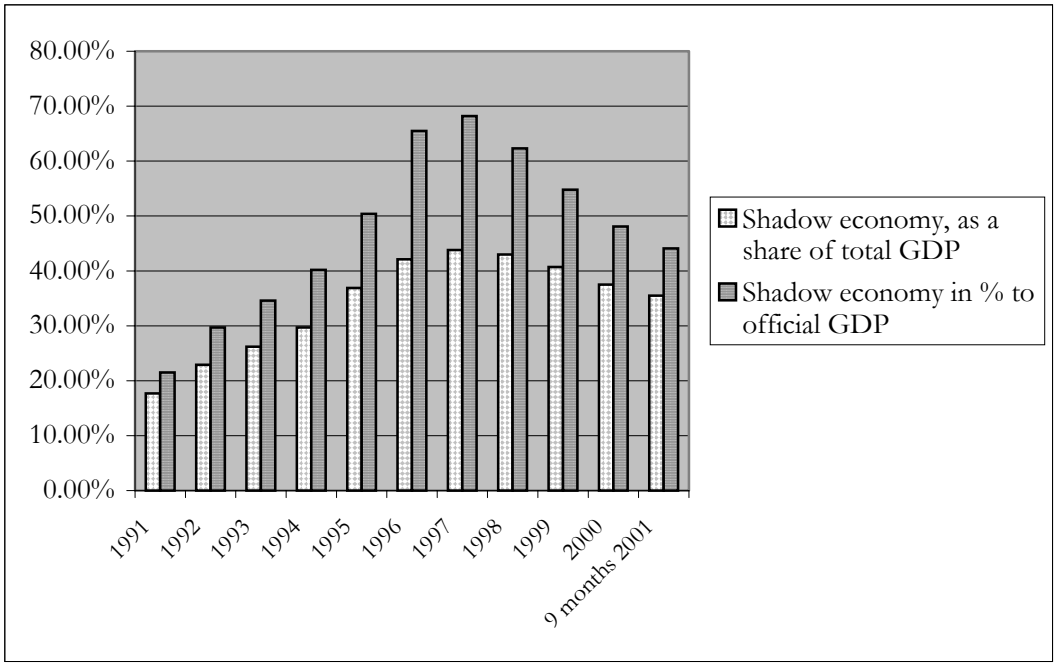
Research of shadow economy can be divided into two major streams: macro and microanalyses. Macro analysis deals with the issue of shadow economy's dynamics and its proper measurement; therefore, macro approach is an ex-post attempt to explain the scope and consequences of the phenomenon. Microanalysis in turn attempts to model the incentive structure of separate entities (households, businesses, government), and consequently, relates to an ex-ante approach.

There are four commonly applied methods to estimate the size of shadow sector: financial method (currency approach - Cagan [1958]), monetary method (Gutmann method), electricity consumption method (Kaufmann and Kaliberda [1996]) and household expenditure method. Based on these methods to assess a shadow economy share, Ministry of Economy and European Integration of Ukraine independently calculates an integral coefficient of shadow economy in Ukraine (Analytical note #2 [2002]). Figure 1 represents development of the shadow economic sector as a share of official GDP and a share of total GDP, where total GDP includes both the legal economy and the shadow sector. Estimates indicate that the size of the unofficial economy in Ukraine for the first nine months of 2001 is of order 44.1% of official GDP and 35.5% of total GDP; thus, these numbers speak themselves for the importance of consideration. Johnson, Kaufmann and Zoido-Lobaton [1998] also evaluate an average size of the shadow sector for transition economies for 1990-1997 and conclude that only Azerbaijan and Georgia surpass the significant size of Ukrainian shadow sector (48.9% to official GDP) with 60.6% and 62.6% respectively.

Essentially, macro analysis suggests that shadow sector largely contributes to total GDP of Ukraine, and therefore, Ukraine occupies one of the leading positions in a rating of greatest shadow economy among the countries of FSU.

Figure 1

Dynamics of the shadow economic sector in Ukraine as a percent from GDP



Source: Ministry of Economy and European Integration, The State Committee of Statistics

Micro-level analysis deals with modelling an incentive framework for firms as demanders for the unreported labour force and households as suppliers of such labour units. Usually, government is implicitly introduced in such framework by setting different taxes and formatting rules of the legal environment.

Research of Johnson, Kaufmann, McMillan and Woodruff [1999] and a country study prepared by the experts of World Bank (A World Bank country study [2000]) suggest that the following factors are the major determinants of Ukrainian businesses' hidden economic activities:

1. Excessive taxation. Ukrainian entrepreneurs reports excessive taxation as one of the major impeding factors to business development and especially business start-up. According to the 1997 Global Competitiveness Report (A World Bank country study [2000]), Ukraine is at the lowest score of 1.58 on 1-to-7-scale basis, which evaluates country's taxation enhancing competitiveness.

2. Burdensome regulation. This factor can be measured in a number of different ways, while among the ones often cited are time required for managers to comply with the regulation and frequency of supervision visits by different state authorities. The questioned sample of managers in Ukraine reports that a quarter of their working time is devoted to the state regulatory matters, which signifies vague and intricate regulation system and explains businessmen desire to escape numerous control procedures by moving into unofficialdom.
3. Corruption. Almost 90% of the firms declared extra payments for public services. Unequal distribution of public services, which is especially important for newly established firms, leads to deterring of new businesses and increases the benefits of the shadow sector.

Concluding with the premises, Ukraine's failures of tax, legal, regulatory and public services systems generate conducive conditions for the shadow economic sector. Favourable position of unofficial enterprises in turn reduces tax revenues, financing of public services and undermines the law standards of the society.

Finally, at the core of business unofficial activities there is a household decision to supply the labour. Looking for maximization of the earnings obtained from the legal and unofficial sectors, an individual encounters the marginal tax rates in both. Thus, in order to affect a personal decision to evade taxes, government can operate with the official tax system and penalties for tax evasion. As a necessary complement to the implicit taxation of the unofficial sector, a probability of being caught and charged with tax evasion is also determined by the government enforcement efforts.

Beside these direct methods to influence labour supply decisions, a government can also affect the qualitative factors of tax compliance by creating favourable economic environment for particular groups of able-bodied citizens. At this level, it is important to investigate the social characteristics of a worker in the unofficial sector, which would contribute to understanding of the necessary alterations in government policy. Moreover,

knowing quantitative magnitudes of such factors would help to assess the volume of the informal activities and its elasticity with respect to the factors' dynamics. Thus, modelling of household behaviour in terms of labour allocation to the two economic sectors is the most essential step to understanding of shadow economy appearance and its dynamics. As an ultimate aim of a research, such a model would provide possible solutions to the problem of shadow economy.

The thesis is organized as follows. Chapter 1 presents the retrospective of research of the household behaviour in the irregular economy and description of the data sample used in our analysis. Chapter 2 formulates the theoretical framework and estimation strategies. The empirical part begins with Chapter 3 discussing the determinants of tax evasion in Ukraine. Chapter 4 describes the estimation results and policy implications for Ukraine. Conclusions summarize the major findings of the current research.

Chapter 1

PROBLEM DESCRIPTION

Literature review

At the initial stages of the problem research, which coincide with the growing attention to the labour economics, theoretical trust prevails over the empirical contents. Earlier models were built on the erroneous methodology, which yielded the biased results for the magnitude of income and substitution effects. The main theoretical underpinnings included the justification of a particular model and estimation strategy applied. The major drawback of these papers was the lack of empirical testing. Only in the middle of 1980s there were several seminal papers that set a break through in the investigated problem. Finally, little if any research on labour supply concerns the transition economies, which are commonly deemed to have the largest shares of shadow economy in the country's total GDP.

Heckman [1993] provides a good summary of the estimation strategies and theoretical models applied to an investigation of the labour supply decisions, drawing a special attention to the problems of selection bias and measurement error. The author divides the previously tested models into 4 major categories:

(1) $E(H \mid w, Y, \varepsilon)$,

where H – hours of work, w – hourly wage rate, Y – non-labour income, ε - individual tastes

(2) $E(H \mid w, Y, H > 0)$

(3) $E(H \mid w, Y) = E(H \mid w, Y, H > 0) * probability(H > 0 \mid w, Y)$

(4) $probability(H > 0 \mid w, Y)$

First equation is a structural labour supply equation, which allows for Slutsky decomposition into income and substitution effects, controlling for the individual tastes. Second equation is commonly tested empirically without

any reference to the individual heterogeneity. Fourth equation describes the participation decisions and the third one specifies the aggregate labour supply choices. The labour supply choices based on equations (3) and (4) reveal greater sensitivity to the variation in wage and non-labour income mostly due to their impact on the employment decisions ("extensive margin" effects). Consequently, elasticities and tax effects computed at the "intensive margin" (equations (1) and (2)) are closer to zero and produce weak estimates and spurious results due to the missing data on the non-workers. Heckman concludes on the selection bias pertinent to the equations (1) and (2) and suggests that much of the elasticity in estimated labour supply comes in "entry and exit" decisions.

Another common problem encountered in estimating labour supply equations is that of a measurement error in hours worked and hourly wage rate. Evidently, the errors associated with the hourly wage rate are positively correlated with the true values of the wage rate and education and age variables. On the contrary, the errors pertinent to the reported working hours are negatively correlated with the true values. Due to the measurement errors of working hours and hourly wage rates, estimates of income and substitution effects are biased toward zero. Heckman suggests correcting for the measurement error with the instrumental variable approach, though excluding age and education dummies from the instruments list.

Hausman and Ruud [1984] investigate the family labour supply with a progressive income tax ("marriage tax") and the effect of deductions introduced for the two-earner couples on household's labour supply. Their model is based on an indirect utility function with the coefficients optimised by maximum-likelihood procedure. Three important points arise in the paper, which are the essential bricks for our study. First of all, this survey emphasizes the interdependent nature of family labour supply decisions. In order to accommodate for the conditional choice of one of the spouses, the authors introduce "virtual non-labour" income into the labour supply equation, which contains the earnings of the other spouse. Depending upon the conditional

choice of the second worker in a family, the two-earner couple faces different marginal tax rates. Thus, the second issue is the non-linearity of the budget sets appearing due to different marginal taxes. Finally, Hausman and Ruud argue for the general reliability of the cross-section data in estimating the uncompensated (Marshallian) labour supply. In other words, the correct estimates of the uncompensated labour supply should account for the individual heterogeneity.

The issue of the interdependence of labour supply decisions within a household is also discussed in Basu, Genicot and Stiglitz [2000]. They develop a model for aggregate labour supply conditional on the changes in labour demand, particularly, the responses of household's labour supply to the adverse shocks in aggregate labour demand. There are two distinguished effects: "added worker effect", when the other members of a household join the labour force in order to compensate for the loss in the family's total income, and "discouraging worker effect" that prevents potential workers from entering the labour market due to subjective perception of the lower chances to find a job. Although the theoretical model developed in the paper does not directly relates to the current thesis, the authors emphasize the necessity to consider marital status and total number of adult members in the household while modelling the household labour supply function. Similar to Hausman and Ruud, the authors argue for distinction of secondary workers' response and possible constraints on the working hours in the primary job market.

The next dimension of the investigated problem concerns the introduction of taxes and enforcement efforts into the modelling of labour supply. The obvious candidates commonly applied as proxies for the enforcement level are penalties for tax evasion and probabilities of being caught and convicted in tax evasion.

Charles Clotfelter [1983] claims that the marginal tax rates have a significant effect on the amount of tax evasion, creating a substitution effect, i.e. encouraging tax incompliance and the hidden economic activities.

Investigating the effectiveness of IRS Taxpayer Compliance Measurement Program in 1969, the author concludes that approximately 20-53% of taxpayers evade different due payments. Professor Clotfelter highlights the corner-stone issue in modelling the shadow economy. He states that the underreported earnings is the basic indicator of tax evasion, which is the approach adopted in the thesis. He also claims for endogeneity of the reported penalties and audit rates, if such are used as proxies for penalties and probabilities of tax evasion in the modelling. These rates are rather the exposition of unlawful behaviour than the expected values for tax evaders. Thus, if such reported variables are included into the model of the unregistered earnings, they lead to simultaneity bias. Moreover, Clotfelter distinguishes among the purposely evaded taxes and "honest" tax errors performed by a taxpayer. However, it is usually impossible to empirically separate the two categories and the expected mean of the latter is assumed to be zero; the author agrees that we cannot do better but to ignore such "honest" mistakes.

Burtless and Hausman [1978] build up an intricate model of non-linear budget sets appearing due to different marginal tax rates and evaluate the consequences of Gary income maintenance experiment in Indiana, 1971-1974. Authors introduce a structural model of the net wage depending on the working efforts and allow for a distribution of preferences for the labour-leisure choices in the population, which has not been previously accounted for. They assume a constant elasticity labour supply function:

$$H = k * w^{\alpha} * Y^{\beta},$$

where H – total number of hours supplied, w – hourly wage rate, Y – non-labour income, and $k_i = \exp(Z_i\delta + \varepsilon_i)$ - parameter describing individual heterogeneous characteristics. The authors claim that progressive income taxation and government transfer programs generate non-convex budget sets with kinks at points, where the income rises enough to place an individual into the next higher tax bracket. Rather than restricting a model to any

particular form of a utility function, Burtless and Hausman integrate the Roy's identity relationship in order to construct an indirect utility function.

The ultimate model for estimation is presented in a form of a probability of observing the actual hours worked depending on the unknown parameters (k, α, β). This function is constructed for the non-convex and convex budget sets, which include only two segments (i.e. one kink point) for simplicity of exposition. However, the rigorous model does not reveal a significant direct effect of the different marginal tax rates on the labour supplied. The authors conclude that wide variation in the after-tax wage rates has little effect on labour supply, though possibly a larger indirect effect through the impact of taxes on the family non-wage income (changes of intercepts).

Another distinguishing paper by Lemieux, Fortin and Frechette [1994] describes the decision of a household to work unofficially as a constrained optimisation problem. The authors develop a model of an individual time allocation to the official and unofficial work, and leisure, subject to the constraints imposed on the consumption volume and leisure-hours. The model incorporates stochastic variables and specific tax rates for the official sector. The household decision is affected by the probability of being caught and having to pay a penalty and the imposed tax rate, which impacts the official wage and further the propensity to engage into the shadow production. The authors conclude that hours in the official sector are not affected by the official wage rates, however, they define the hours of the unofficial activities. Moreover, the earnings are a linear function for the official and concave function for the unofficial activities, suggesting the existence of a boundary on the possible earnings growth in the shadow economy. One of the reasons that explain the phenomena is the limited scope of unofficial economy, which makes the latter easier to detect after a certain level of production.

The structural labour supply model is adopted as a basis for a theoretical framework in the current thesis.

The seminal work of Lacroix and Fortin [1992] introduces the unprecedented tests and assumptions behind a utility function in modelling labour supply decisions. First of all, the model allows for the spillover effect between the regular and unofficial markets and this assumption is justified empirically, suggesting that the two markets are not perfect substitutes. Secondly, the assumptions about the preferences are tested on a basis of a quadratic utility function (see *Theoretical Framework* section). Moreover, the authors estimate constrained and unconstrained versions of working hours in the regular sector, which all together constitute six regimes of a worker's behaviour (including working and non-working strategies for both sectors). Finally, the constructed model yields unambiguously defined effects of the proportional taxes in the regular sector and the penalties and probabilities of tax evasion on the hours worked in unofficial sector. Applying maximum-likelihood estimation to the data set generated by a household survey, the authors provide the following estimates of the elasticities:

- $\varepsilon(hr\_unofficial, w\_official) = -0.79$ for the unconstrained version and (-2.88) for the constrained version;
- $\varepsilon(hr\_official, w\_official) = 0.49$ and 0.25 for the unconstrained model and 1.06 and 0.45 for the constrained model;
- $\varepsilon(hr\_unofficial, [1 - tax\ rate]) = -0.65$ for the unconstrained model and (-2.55) for the constrained version.

Lacroix and Fortin explain the larger elasticities for the constrained version by few individuals working little hours. The authors conclude that the higher marginal tax rate in the official sector induces larger efforts devoted to the unofficial sector, whereas the higher probabilities and penalties entails less hours worked in the unofficial sector.

Friedman et al. [1999] propose an extensive survey on the determinants of shadow economy in 69 countries. The authors question several seemingly indisputable "truths":

a) High taxes are correlated with higher share of shadow activities in economy;

b) Larger share of underground economy lead to weak institutions setting (the work traces institutional variables such as tax rates, over-regulation, legal environment, corruption);
c) High taxation enhances government performance.

The paper covers the application of various indexes and approaches in order to properly estimate the size of the underground sector, such as Multiple-Indicator Multiple Cause approach (MIMIC), cash (currency demand) approach, and measures of business environment. The authors attempt to estimate the net effect of higher direct taxes. On the one hand, larger tax rate would increase the incentives to divert from legal activities; on the other, higher tax revenues would improve the level of law and order and allow for greater provision of the public goods (such as education and infrastructure), which in turn enhance the impetus to produce officially. Therefore, the proposed solution is some "threshold" tax rate that maintains the sufficient budget revenues for a proper state functioning. According to the cross-country estimates the tax rates have a dominant indirect effect, which implies that the higher tax rate rather improves legal environment than distorts the working efforts in the regular economy. The over-regulation, weaker legal environment and higher corruption drive the intensity of the unofficial activities. The empirical findings reveal that the weak institutions rather than high tax rate impel the growth of underground economy and, thus, lead to lower government revenues.

A number of decisive papers are concerned with estimating the size of the shadow economy and determining the factors that define a household decision to get involved in the unofficial activities. However, the works based on an empirical survey for the transition economies, which unconditionally recognized the leaders in the list of informal sector share, are limited.

Guariglia and Kim [2000] in their paper provide an unusual view on the "moonlighting" issue in Russia. They assert that in fact the shadow economy positively affects citizens' welfare, since the effects of moonlighting are of a transitory character. By their nature, people are generally risk-averse; so that

temporary multiple-job holding creates an opportunity for an individual to try his/her efforts in a new possible occupation and obtain necessary skills and information relating to a new job. There are costs of reversal to the former job if the new one does not coincide with a worker's expectation; thus, an unofficial work should be treated a "human capital enhancing activity". These theoretical postulates are confirmed empirically with a household survey based on panel data. Analysing the dynamics of a number of households in Russia, the authors come to the conclusion that after a transitory period of moonlighting, workers return to a single job holding. Moreover, a past experience of moonlighting increases the probability of a job change; therefore, an informal employment serves as a transitory stage in one's career ladder. Another interesting observation is a high correlation between the moonlighting experience and the successive self-employment decision. The authors argue that an involvement into the informal activities contributes to a consecutive starting of own business and long-term benefits to the economy, which may outweigh the short-term costs of significant unofficial production.

In a household survey conducted by UEPLAC (Tacis Programme) in Ukraine during August-October 2000 the authors develop a probit model aiming at estimating a real scale of the shadow economy in Ukraine. The joint work of Novoseletska and Najman [2001] defines the specific features of households that stipulate a propensity to tax evasion. The survey raises an interesting aspect of the shadow economy. Not only the characteristics of households stipulate the level of the unofficial activities, but also the government creates conducive conditions for the underground economy development. Therefore, the model indicates two directions to reduce the unofficial production:

1) Implicit state influence on a household performance (female employment, public education and training programs, etc.).
2) Direct government intervention (taxes and enforcement, transparent legislature, etc.).

Another strong point of the paper is that the authors clearly distinguish among the types of activities that the households conduct (e.g., single formal activities, single informal activities, multiple activities, activities on land plots, etc.) and provide a crosschecking for the true answers.

Data description

The data used in the paper is generously supplied by Tacis UEPLAC project, that, following the request of Cabinet of Ministers of Ukraine #02-105 from 11.16.1999, conducted a household survey in August – October 2000. The data set was generated upon the filled questionnaire by 1017 different households or, respectively, 2315 adults over 16 years old, which possessed the representative characteristics of the total labour force in Ukraine as of the mentioned period. Referring interested readers to the detailed description of the survey procedure in the original report, it is important for the estimation purposes to emphasize the representability of the obtained answers. The applied sampling technique involved four levels of selection, targeting at representative exposition of territorial, administrative, compositional and welfare status of a random household.

Another point, which is commonly arguable, is general reliability of a survey as a research method. Although not guaranteeing true answers provision, a survey is the optimal combination of minimum costs and qualitative output. It can be argued that since the respondents are not punished for the inaccurate answers (especially on specific questions concerning informal performance), the data set generated by this method is not a trustworthy one. On the other hand, in order to solve the problem of validity, the next best trial is a direct supervision of households' behaviour, which is a rather costly procedure. Thus, a technique of household survey is a justified attempt to obtain the plausible overview of forming labour supply in Ukraine.

Whenever it is necessary, the procedure of deriving answers to the specific questions (e.g. unofficial earnings, hours devoted to the unofficial occupation, amount of the unregistered consumption) will be presented.

In order to support the representative validity of the survey, the data set is checked for the following criteria:

1) *Geographical*. According to the data of The State Committee of Statistics of Ukraine [*Quarterly predictions, June 2001*], 69% of households reside in urban area and 31% - in rural. The survey presents 67% of urban dwellers and 33% of rural inhabitants.
2) *Average size*. According to the same release of The State Committee of Statistics of Ukraine [*Quarterly predictions, June 2001*], 49.3% of households consist of 1-2 family members, 22% - of three members, 18.9% - of four members. The survey includes 44% of one-two-member households, 22% of three-member households, and 34% of four-member and more households.
3) *Land plot*. Official statistics [*Quarterly predictions, June 2001*] claims that 62.7% of households possess own land plots and use the own crop as additional income source. 60% of survey respondents possess land plots and 18.5% of households (or 31% of households who possess land plots) sell the produce from own land plot.
4) *Unemployment*. In 2000 the official unemployment rate was 4.2% [*Monitoring of macroeconomic and industrial indicators, September 2001*] and unemployment rate computed by ILO method constituted 11.7% [*Perspective research: Economic Statistics in Ukraine, November 2000*]. The survey suggests 13.8% and 13.8% respectively. Significantly higher official unemployment rate revealed by the survey can be explained by the status self-assessment; hence, an individual present his opinion on the current personal status rather than the official registered status.
5) *Monthly nominal average wage*. As for September 2000 monthly nominal average wage constituted 249.09 hryvna [*Monitoring of macroeconomic and industrial indicators, September 2001*]; the survey offers a number of 253.5

hryvna from all possible wage earnings (regular registered, extra registered, unregistered).

Thus, the survey sample possesses the characteristics similar to those of the labour force composition in Ukraine as of the respective period. This brief analysis suggests that we can make an inference about the total labour force based on the estimates obtained in our sample analysis.

An initial analysis of the sample discover some specific features, which are to be incorporated later into the model assumptions:

1. *Composition of current employment status*. Figure A1 (Appendix A) represents the composition of current employment status and reveals that about one half of the surveyed adults are employed, almost one third (28%) are on retirement, students constitute approximately seven percent, and ten percent belong to the officially unemployed. Thus, from the available pool of potential labour force only 49% of workers are officially employed, which signifies a low level of labour involvement.
2. *Multiple job holdings*. Analysis of households' labour activities suggests that only 21.6% of able-bodied citizens occupy single job, whereas almost 64% of the respondents account for the largest three multiple job holdings (Figure A2, Appendix A). Therefore, specific character of Ukrainian labour market is multiple job holdings.
3. *Unemployment welfare benefits*. Although the initial analysis indicates relatively high level of unemployment, a choice of "not-working" behaviour rather than working in any sector cannot be considered voluntary. According to The State Committee of Statistics of Ukraine [*Monitoring of macroeconomic and industrial indicators, September 2001*], an average level of unemployment welfare benefits paid in August 2000 was 55.26 hryvna, which was significantly lower than the minimum wage rate as of the same period – 118 hryvna. Therefore, being unemployed cannot be considered as a rational choice in the current model.

4. *Awareness of taxes*. The survey suggests that only 15% of the respondents did not file a tax return in 1999 because of unawareness of the procedure. Consequently, for the purposes of the modelling, awareness of taxes can be considered as a model assumption.
5. *Occurrence of wage arrears*. According to the survey data, 27% of regular workers and 5.4% of workers at extra jobs incurred wage arrears. However, the average size of wage arrears was quite different for these categories: 819 hryvna for regular workers and 540 hryvna for extra workers. Thus, in modelling labour-allocation behaviour, an occurrence of wage arrears should be taken into account.

Initial analysis of the data allows for the following assumptions to be incorporated into the model: workers hold multiple jobs, choosing between different types of work rather than "working" or "not working" behaviour; they are aware of the tax rate they should pay; finally, there is a probability of not being paid on time, which is different for regular and extra work.

In general, the survey sample represents all the major characteristics of Ukrainian households as of August-October 2000, except for the discrepancy of official unemployment rate, which can be ignored owing to method of generating this number for the sample.[3]

Eventually, for the purposes of the analysis we are interested in labour composition differences between the official and unregistered economic sectors. The following Table 1 presents the detailed description of the unregistered and official labour force with the average hours and average earnings presented for the unregistered sector. Table B1, Appendix B reproduces Table 1 in percent terms.

[3] The number was computed from the personal answers on a question about the current working status rather than a question about the official registered status.

Table 1

Composition of official and unregistered labour force

Characteristics	Percent of total sample	Unregistered activity per month		
		Percent	Hours (average)	Earnings (average)
TOTAL	100.0	71.36	125	191
Sex:				
Male	45.23	45.70	131	178
Female	54.77	54.30	119	200
Age:				
16-24	16.15	14.65	117	169
25-39	28.21	28.40	129	245
40-59	32.18	32.69	126	139
60 and more	23.46	24.27	122	213
Education:				
Incomplete secondary	13.74	13.98	114	169
Secondary	20.65	20.64	123	213
Secondary vocational	35.00	35.65	135	182
Higher	26.26	26.21	116	213
Employment status:				
Employed	45.74	46.97	120	210
Pensioner	27.17	27.91	123	192
Housewife	3.07	3.09	123	278
Student /pupil	5.66	4.84	90	88
On maternity leave	1.86	1.51	99	180
On paid leave	0.30	0.30	86	72
On unpaid leave	0.80	0.67	126	54
Unemployed	9.68	10.11	158	184
Regular-work income per month[4] (in brackets – percent from answered):				
0-200	31.49 (71.49)	31.66 (71.16)	107	150
200-400	7.26 (16.48)	7.38 (16.59)	93	238
400-600	2.94 (6.67)	3.03 (6.81)	120	267
600-800	0.80 (1.82)	0.85 (1.91)	92	504
800-1000	0.39 (0.89)	0.42 (0.94)	55	1620
1000-2000	0.91 (2.07)	0.97 (2.18)	96	851
2000 and more	0.26 (0.59)	0.18 (0.40)	28	1548
Regular-work hours per month (in brackets – percent from answered):				
0-36	0.73 (1.70)	0.73 (1.69)	103	151
37-72	1.43 (3.34)	1.45 (3.36)	99	171
73-108	1.90 (4.43)	1.69 (3.92)	77	99

[4] In categories "Regular-work income", "Regular-work hours", "Ownership of regular place of work", "Industry of regular job" percents do not add up to 100% but reflect the proportion of the people answered.

109-144	2.98 (6.95)	3.03 (7.02)	124	346
145-180	17.06 (39.81)	17.31 (40.11)	104	169
181 and more	18.75 (43.76)	18.95 (43.91)	111	188
Wage arrears at regular work	14.47	15.13	116	153
Own land plot	61.90	60.05	121	187
Proceeds from crop	18.45	17.07	137	185
Ownership of regular place of work (in brackets – percent from answered):				
Private (Ukrainian)	11.79 (24.61)	12.17 (25.10)	148	207
Private (with foreign capital)	1.04 (2.17)	0.85 (1.75)	103	186
Collective	3.46 (7.22)	3.03 (6.25)	122	230
Joint Stock	5.01 (10.46)	4.90 (10.11)	117	486
State	26.61 (55.54)	27.54 (56.80)	105	155
Industry in the regular job (in brackets – percent from answered):				
Light and food	2.50 (6.38)	2.48 (6.26)	170	195
Woodworking and pulp and paper	0.30 (0.77)	0.30 (0.76)	52	88
Agriculture and forestry	2.72 (6.95)	2.42 (6.10)	112	117
Civil engineering	2.72 (6.95)	3.09 (7.80)	116	372
Transport and communication	4.80 (12.26)	4.96 (12.51)	117	131
Trade and public catering	6.78 (17.31)	7.26 (18.31)	132	180
Housing and communal services	2.89 (7.38)	2.78 (7.01)	118	170
Health, social security and sports	4.88 (12.46)	4.54 (11.45)	106	156
Education, culture, science, arts	6.39 (16.32)	6.42 (16.20)	91	214
Finance, social insurance, real estate	1.94 (4.95)	1.94 (4.89)	123	231
State administration, public organizations	3.24 (8.27)	3.45 (8.70)	110	122
Oblast:				
Dnipropetrovska	23.33	24.52	145	188
Crimea	4.02	4.30	123	384
Odesska	10.06	10.41	99	348
Lvivska	21.12	22.52	116	188
Kyivska	16.63	16.00	130	138
Donetska	24.84	22.28	123	125

Detailed analysis of labor composition in the two sectors reveals that approximately three quarters of the sample are involved in the unofficial

economic activities with distinct difference in gender groups. On average, men work more for the unofficial sector and women receives larger earnings from the unofficial sector. The largest percent of the unofficial workers are young aged at 25-39. There is no obvious interdependence between the unregistered employment and education level of a worker. Employed, housewives and unemployed constitute the greatest share of unofficial labor.

Surprisingly, there is a positive relationship between unregistered earnings and regular income and labor-hours up to some point, which suggests that the unofficial sector is a complement to the official employment rather than the substitute.

As it is mentioned previously, possession of a land plot has two different effects on the unregistered employment. On one hand, the land plot is the additional source of household's consumption, which implicitly increases the total registered income and lessens the necessity to search for other additional sources. On the other hand, proceeds from selling the crop from the land plot are the source of household unregistered earnings, which directly depends on the crops' quality. These two opposite effects are clearly obvious from the different amounts of average hours for the groups that possess land plots and sell the crops from own land plots.

Composition of ownership structure of regular job placements suggests that the workers of state and private enterprises with foreign capital are more likely to comply with tax system, whereas the shareowners of joint stock enterprises generate the largest unregistered earnings although not diverting the largest share of labor-hours to the unofficial sector. This result may imply that the joint stock shareholders accumulate their unregistered earnings right at their regular place of work.

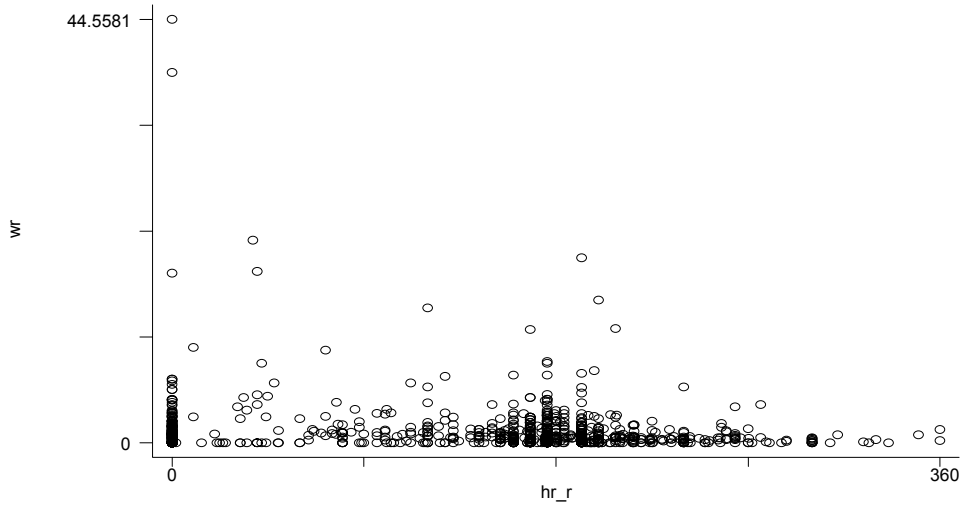
There is also some evidence that workers with primary job in civil engineering and finance and social insurance sectors are more likely to obtain unregistered earnings.

Finally, the analysis of the sample shows that Crimea and Odesska oblast inhabitants are on average more involved into the unregistered

economic activities. However, this result may be biased due to the coincidence of the questioning time with the seasonal peak of economic activities in the named regions.

Figure 2

Scatter of observed hourly wage rate (wr) and hours worked (hr_r) in the official sector



In order to determine the interdependencies of hours and net[5] hourly wage rates, we have looked at their combined scatters. Figure 2 depicts the scatter of hourly wage rate and official hours of work. Although we cannot infer that these scatters represent the labor supply curves, they are the testimony for the observed equilibrium settings in the two sectors. Consequently, we can infer from Figure 2 that the equilibrium hours of work in the official sector does not show an apparent connection to the observed hourly wage rate.

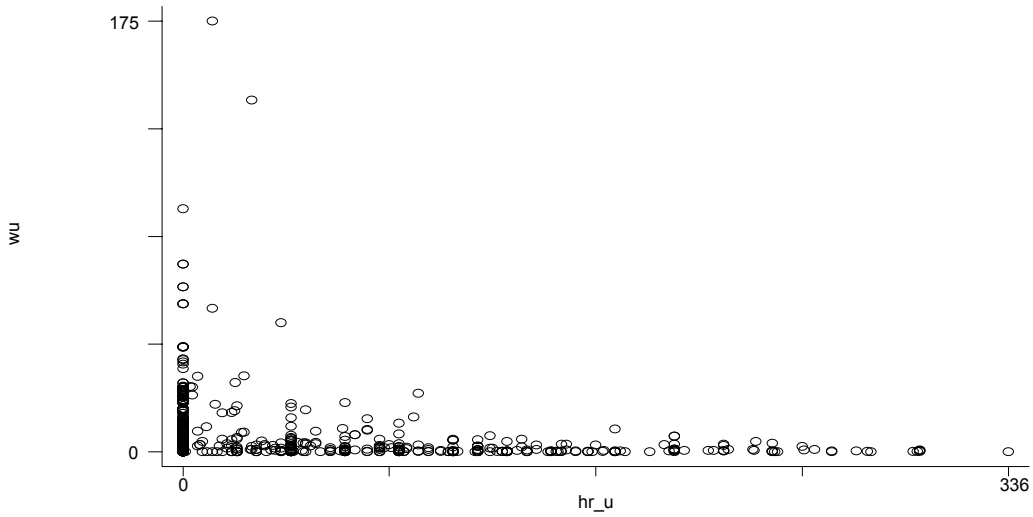
On the contrary, even a quick glance at Figure 3, which depicts the scatter of observed hourly wage rate and hours worked in the unofficial sector, is enough to infer negative non-linear relationship between the observed hourly wage rate and the hours worked in the irregular sector. There is a good possible explanation for the observed "surprising" relationship in

[5] Net hourly wage rate in the official sector is the hourly earnings after the tax deductions; net hourly wage rate in the unofficial sector is the expected hourly earnings in the unofficial sector (see Chapter 2).

the unofficial sector. An increase in the labor activities (here, an average hours of work per person rather than increasing total number of people working) in the irregular economy leads to the expansion of the unofficial production. At the same time, the scale of unofficial production can become such that it is easier for tax authorities to detect the unlawful businesses. In other words, the negative and non-linear interdependence in the unofficial sector is a consequence of an increasing probability of detection and the penalties for tax evasion. Thus, the reverse relationship between the hours and hourly wage rate in the unofficial sector is induced by the existence of the upper limit in the unofficial production capabilities.

Figure 3

Scatter of observed hourly wage rate (wu) and hours worked (hr_u) in the unofficial sector



Summarizing, the graphical analysis of the data suggests a vague relationship between the observed hourly wage rate and hours of work in the official sector and transparent negative non-linear interdependencies in the unofficial sector. Moreover, unregistered earnings and hours of work are influenced by gender, age, employment status, regular sector earnings, ownership structure and industry of regular job, and maybe biased in geographical aspect due to seasonal character of industry concentration in particular regions.

C h a p t e r 2

THEORY

Theoretical framework

Modelling of household behaviour in the thesis is based on a partial equilibrium analysis. The model of household behaviour includes the utility function of a worker and the stochastic budget constraint, described as a constraint on individual's consumption level. Therefore, the eventual decision of a worker is to maximize his/her utility function subject to imposed constraint on available leisure hours and composite consumption good. According to the initial descriptive analysis of the data sample, the model also assumes that a household will work in both sectors in economy (official and unofficial), which is pertinent to Ukrainian pattern of labour market. Finally, only personal income tax rate is included into the model. Tax revenues generated from personal income tax rate constituted approximately 24% of tax revenues to the consolidated budget of Ukraine in 2001.[6] The basic framework is adopted from Lemieux, Fortin and Frechette [1994] and Lacroix and Fortin [1992].

1. Official sector[7]

Based on the graphical analysis of observable hours and hourly wage rate, we assume that the monthly disposable earnings in the official sector (Y_o) are a linear function in both hours worked and hourly wage rate, whereas the latter is established by an employer (i.e. we assume that the gross hourly wage rate in the official sector is exogenous). Thus, in the official sector a worker is paid W_0 - gross hourly wage and he decides on h_0 hours per month devoted to the official sector. He is also expected to pay τ – ad valorem income tax

[6] Source: The State Committee of Statistics of Ukraine.

[7] Official sector includes regular and extra registered jobs.

rate – out of his gross labour earnings. Thus, the household's net official earnings:

$$Y_0 = W_0(1-\tau) * h_0 \qquad (2\text{-}1)$$

2. Unofficial sector

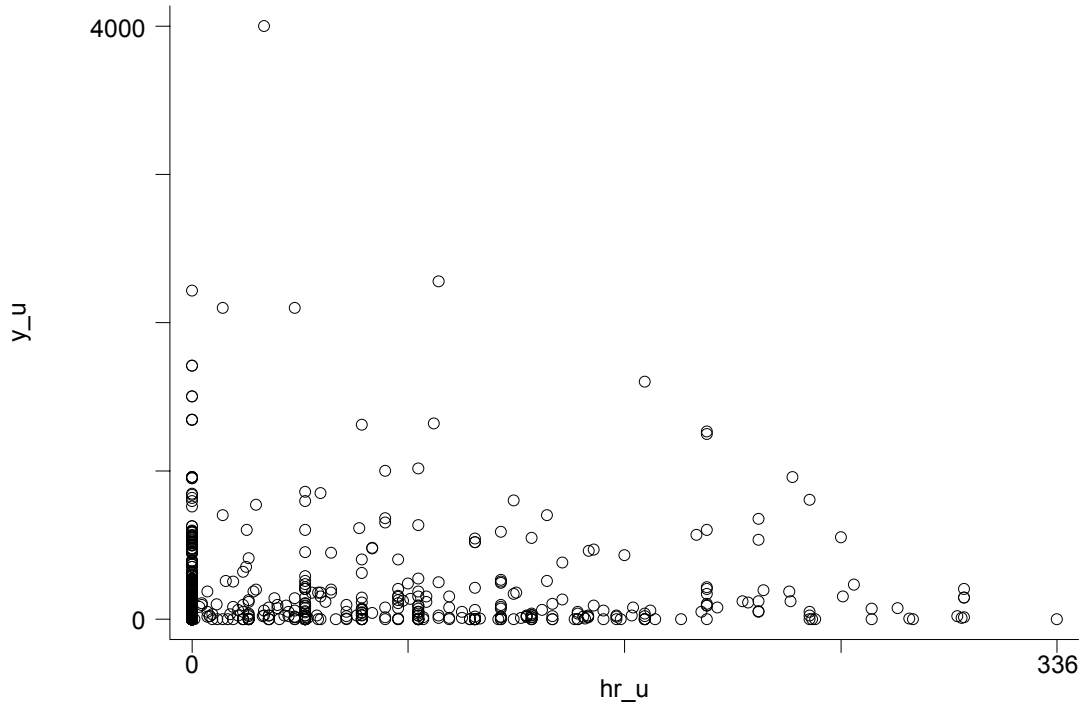
As our empirical findings suggest, the observed hourly wage rate in unofficial sector is a decreasing non-linear function in hours worked. Therefore, we assume that the hourly wage rate in unofficial sector is endogenous variable in our model:

$$w_u = h_u^{\gamma} \text{ , where } \gamma < 0 \text{ - "curvature" parameter .} \qquad (2\text{-}2)$$

Note that γ is the elasticity of hourly wage rate with respect to hours worked in the unofficial sector. Our assumption is not rejected by the observable interdependence between the unofficial earnings and labour-hours, presented in the following Figure 4.

Figure 4

Scatter of reported net earnings (y_u) and hours (hr_u) worked in the unofficial sector



Additionally, in order to incorporate individual heterogeneous attitudes toward tax evasion (hence, participation decision about the unofficial sector), we also include "revenue-shifter" parameter A, adopted in the model of Lemieux, Fortin and Frechette [1994].

Summarizing, the unregistered earnings become a function of the hours worked and a revenue-shifter parameter: $Y_u = Aw_u h_u = A * h_u^{1+\gamma}$, (2-3)

where $A = \alpha_1 + \sum_i x_i * \delta_i + \varepsilon$, (2-4)

x_i - factors that influence an individual's decision to evade taxes;

$\varepsilon \sim NID(0, \sigma^2)$.

From (2-3) it can be inferred that there is a further constraint on γ concerning the unregistered earnings elasticity with respect to the unofficial hours. For the unregistered earnings function to be decreasing in hours worked, we need: $1+\gamma < 0$, or $\gamma < -1$, where γ - is the hourly wage rate elasticity with respect to hours worked in unofficial sector.

Finally, there is an implicit tax rate for the individuals working unofficially. First of all, as we mentioned previously, the scope of unofficial output is limited due to the increasing probability of detection – p. Furthermore, in case of being caught and convicted in tax evasion, a worker is subject to the penalty rate - θ, which is assumed to be a factor of the official tax rate: $\theta = (1+\mu) * \tau$, where $\mu > 0$ (2-5)

Eventually, the expected earnings from the unofficial sector is a stochastic parameter function:

$$E(Y_u) = A * h_u^{1+\gamma} * (1 - p * \theta) \quad (2\text{-}6)$$

Hereafter, we would refer to the value (p*θ) as the marginal tax rate in unofficial sector.

3. Constraints

Working in both sectors, an individual generates his income and then distributes the earnings between consumption and savings. Of course, there are additional sources of earnings such as borrowings and savings from the

previous period and additional sources of spending such as lending to other individuals. For the modelling purposes we impose the budget constraint on the "net" individual consumption, which is the current period consumption (C) plus previous period savings (S_{-1}) and borrowings (B) and net of lending (L) and current period savings (S):

$$net\ C = C + S_{-1} + B - L - S \tag{2-7}$$

Hereafter, we would refer to the consumption (C) in sense of net consumption defined in (2-7). Proceeding, we assume that the individual expected consumption, which is also an individual static stochastic budget constraint, is a function of his labour income and non-labour income (Y_{NI}):

$$E(C) = W_0(1-\tau) * h_0 + A * h_u^{1+\gamma} * (1 - p * \theta) + Y_{NL} \tag{2-8}$$

In addition, there is a physical limit on total available hours per month. Therefore, the other constraint is imposed on monthly worker's time available:

$$l = T - h_u - h_o, \tag{2-9}$$

where l – leisure hours per month, T – total hours fund per month.

4. Utility function

In the models of household labour allocation, several utility functions were applied according to different assumptions on the household's preferences. For example, Lemieux, Fortin and Frechette [1994] used quasi-linear utility function separable in the arguments:

$U(l, C) = \gamma C + \upsilon(l)$, where $\upsilon(l)$ - is a strictly concave function.

Lacroix and Fortin [1992] analysed quadratic specification, which yields an attractive for estimation linear form in marginal utilities:

$$U(x) = \alpha' x + \frac{1}{2} x' \beta x, \text{ where } x = (h_o, h_u, C)'$$

Searching for a less-restrictive specification of a utility function, Burtless and Hausman [1978] derived the demand for leisure and composite consumption good through the indirect utility function and Roy's identity:

$$\upsilon(p, y) \equiv \max_x [U(x) : p * x \leq y]$$

The thesis approach allows us to overcome the restrictive specification of the household's preferences and similarly avoid complications incurred in indirect utility estimation. For the purposes of estimation, we only need to assume the necessary properties of a utility function, captured by the general form: *U=U (l, C)* (2-10)

whereas: C - composite consumption good, l – hours of leisure.

Hence, the regular conditions imply that consumption and leisure are "goods" rather than "bads", so that: $\frac{\partial U(\mathrm{C},\mathrm{l})}{\partial l} > 0$ [8], $\frac{\partial U(\mathrm{C},\mathrm{l})}{\partial C} > 0$ as first order conditions (2-11)

and $\frac{\partial^2 U(C,l)}{\partial l^2} < 0$, $\frac{\partial^2 U(C,l)}{\partial C^2} < 0$ as second order conditions. (2-12)

Utility function *U=U (l, C)* satisfying these conditions is a concave function in both arguments. Finally, it is also assumed that the incorporated components (composite consumption good and leisure) completely specify the household utility sources.

5. Constrained optimisation

An individual's labour-allocation problem is specified by the following equations:

$$\begin{cases} U(l, C) \to \max_{l,C} & (2\text{-}13) \\ E(C) = W_0(1-\tau) * h_0 + A * h_u^{1+\gamma} * (1 - p * \theta) + Y_{NL} & (2\text{-}14) \\ l = T - h_u - h_o & (2\text{-}15) \end{cases}$$

These three entries can be combined into a single Langrangian construction with λ_1 and λ_2 - Langrangian multipliers: (2-16)

$$L = U(l, C) + \lambda_1 * [E(C) - W_0(1-\tau)h_0 - Ah_u^{1+\gamma}(1 - p\theta) - Y_{NL}] + \lambda_2 * [l - T + h_0 + h_u]$$

Taking partial derivatives with respect to the variables of interest (composite

[8] These are total partial derivatives.

consumption good, hours of work devoted to both sectors and λs) yields the system of the following first order conditions:

$$\begin{cases} \frac{\partial L}{\partial C} = \frac{\partial U(C,l)}{\partial C} + \lambda_1 = 0, & (2\text{-}17) \\ \frac{\partial L}{\partial h_0} = \frac{\partial U(C,l)}{\partial l} * (-1) - \lambda_1 W_0 (1-\tau) + \lambda_2 = 0, & (2\text{-}18) \\ \frac{\partial L}{\partial h_u} = \frac{\partial U(C,l)}{\partial l} * (-1) - \lambda_1 (1+\gamma) A h_u^{\gamma} (1 - p\theta) + \lambda_2 = 0, & (2\text{-}19) \\ \frac{\partial L}{\partial \lambda_1} = E(C) - W_0 (1-\tau) h_0 - A h_u^{1+\gamma} (1 - p\theta) - Y_{NL} = 0, & (2\text{-}20) \\ \frac{\partial L}{\partial \lambda_2} = l - T + h_0 + h_u = 0, & (2\text{-}21) \end{cases}$$

After subtracting the third equation from the second and some further manipulations, we receive the formula for the optimal hours devoted to the unofficial sector as a function of the marginal tax rates in both sectors, gross regular wage rate and the revenue-shifter parameter:

$$h_u^* = \left[\frac{W_0 (1-\tau)}{(1+\gamma) A (1 - p\theta)} \right]^{\frac{1}{\gamma}} \quad (2\text{-}22)$$

Analysing the derived relationship, it can be concluded that:

1) The elasticity of unregistered hours of work with respect to regular gross hourly wage rate (W_0) and net regular hourly wage rate $\{W_0 * (1-\tau)\}$ is negative and bounded by: $-1 < \frac{1}{\gamma} < 0$
2) The elasticity of unregistered hours of work with respect to the marginal tax rate for the unofficial sector $(p * \theta)$ is negative since the elasticity of the unregistered hours with respect to the expected portion of the unregistered earnings is positive and bounded by the unity interval: $0 < -\frac{1}{\gamma} < 1$

While selecting the labour-hours devoted to the two sectors, an individual aims to maximize the aggregate earnings from both sectors. Therefore, rather than estimating a function of unofficial hours, we substitute the derived expression for them into the formula for the expected unregistered earnings and after some manipulations receive the final expression:

$$E(Y_u^*) = A^{-\frac{1}{\gamma}}(1-p\theta)^{-\frac{1}{\gamma}}(1+\gamma)^{-\frac{1+\gamma}{\gamma}} W_0^{\frac{1+\gamma}{\gamma}} (1-\tau)^{\frac{1+\gamma}{\gamma}} \quad (2\text{-}23)$$

In order to facilitate the exposition of unofficial earnings determinants and estimate the respective elasticities of the exogenous factors on the right-hand side, the optimal expected shadow earnings function is transformed into the logarithmic form: (2-24)

$$\ln E(Y_u^*) = -\frac{1}{\gamma}(\alpha_1 + \sum_i x_i\delta_i + \varepsilon) - \frac{1}{\gamma}\ln(1-p\theta) - \frac{1+\gamma}{\gamma}\ln(1+\gamma) + \frac{1+\gamma}{\gamma}\ln W_0 + \frac{1+\gamma}{\gamma}\ln(1-\tau)$$

Combining the constant terms under the single variable η, we can rewrite this function in the form applicable to the estimation: (2-25)

$$\ln E(Y_u^*) = -\frac{1}{\gamma}\eta - \frac{1}{\gamma}\sum_i x_i\delta_i - \frac{1}{\gamma}\ln(1-p\theta) + \frac{1+\gamma}{\gamma}\ln W_0 + \frac{1+\gamma}{\gamma}\ln(1-\tau) - \frac{1}{\gamma}\varepsilon,$$

where $\eta = \alpha_1 + (1+\gamma) * \ln(1+\gamma)$.

Moreover, such exposition allows for unique estimation of the curvature parameter γ, which is the elasticity of the hourly wage rate with respect to the worked hours in the unregistered sector; further, the estimated curvature parameter can be substituted to compute the elasticity of the unregistered earnings with respect the hours devoted to the unofficial sector: $(1+\gamma)$.

Estimation Strategy

1. Selection bias of hourly wage rate in the official sector

In cases of estimating a wage rate, a selection bias arises when the sample of the workers is a non-random sample of potential workers and we cannot observe wages for workers with higher reservation wages. The latter in

this case are more likely to be unemployed. Consequently, the coefficients of OLS based on the censored sample of workers would be biased due to ignoring the observation rule.

The full process of observing the wage rates for a sample of workers includes the following equations, as stated in Verbeek [2000]:

$w_i^* = x_i^{*\prime}\beta_1 + \varepsilon_{1i}$ - Linear wage equation, where the wage rate depends on some exogenous characteristics; (2-26)

$w_i^r = z_i'\gamma + \eta_i$ - Unobserved equation for the reservation wage of the individual; (2-27)

$\begin{cases} h_i = 1 \quad if\ w_i^* - w_i^r > 0 \\ h_i = 0 \quad if\ w_i^* - w_i^r \le 0 \end{cases}$ - Equation for labor supply decision: an individual would work only if the offered wage exceeds his reservation wage. (2-28)

We can rewrite equation (2-28) with unobserved error term and observed characteristics for the individual:

$$h_i^* \equiv w_i^* - w_i^r = x_i^{*\prime}\beta_1 - z_i'\gamma + (\varepsilon_{1i} - \eta_i) = x_{2i}'\beta_2 + \varepsilon_{2i} \tag{2-29}$$

If the error terms in (2-26) and (2-29) are correlated, we face a sample selection model, which is called Tobit II model:

$$\begin{cases} w_i^* = x_{1i}'\beta_1 + \varepsilon_{1i} & (2\text{-}30) \\ h_i^* = x_{2i}'\beta_2 + \varepsilon_{2i} & (2\text{-}31) \\ w_i = w_i^*, \quad h_i = 1 \text{ if } h_i^* > 0; & (2\text{-}32) \\ w_i \text{ is not observed}, \ h_i = 0 \text{ if } h_i^* \le 0; & \\ \text{where } \begin{pmatrix} \varepsilon_{1i} \\ \varepsilon_{2i} \end{pmatrix} \sim NID\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & 1 \end{pmatrix} \right), \text{and } \sigma_2^2 \text{ was normalized to 1 for} & \\ \text{the convenience of estimation.} & (2\text{-}33) \end{cases}$$

Therefore, at the first stage we check for the correlation coefficient between ε_{1i} and ε_{2i}: $\rho_{12} = \frac{\sigma_{12}}{\sigma_1}$. If the correlation coefficient is not statistically different from zero, we can consistently estimate (2-30) by OLS,

ignoring the selection rule (2-31). However, if the correlation coefficient does not equal zero, consistent estimates can be obtained from (2-30) by including into the regression the omitted variable - Heckman's lambda multiplied by σ_{12}:

$$E\{w_i \mid h_i = 1\} = x_{1i}'\beta_1 + E\{\varepsilon_{1i} \mid \varepsilon_{2i} > -x_{2i}'\beta_2\} = x_{1i}'\beta_1 + \frac{\sigma_{12}}{\sigma_2^2} * E\{\varepsilon_{1i} \mid \varepsilon_{2i} > -x_{2i}'\beta_2\},$$

$$E\{w_i \mid h_i = 1\} = x_{1i}'\beta_1 + \sigma_{12}\frac{\phi(x_{2i}'\beta_2)}{\Phi(x_{2i}'\beta_2)} = x_{1i}'\beta_1 + \sigma_{12} * \lambda(x_{2i}'\beta_2), \quad (2\text{-}34)$$

where $\phi(x_{2i}'\beta_2)$ and $\Phi(x_{2i}'\beta_2)$ are probability density and cumulative density of the normal distribution, $\lambda(x_{2i}'\beta_2)$ - Heckman's lambda, and σ_2^2 was normalized to 1. Also note, that $\sigma_{12} = \rho_{12} * \sigma_1$, thus, the sign of the omitted variable will depend on the sign of the correlation coefficient of the two error terms, since σ_1 and $\lambda(x_{2i}'\beta_2)$ are both positive.

Finally, it is obvious from (2-29) that x_{2i} should include at least all the exogenous variables contained in x_{1i}, and other exogenous variables, which presumably affect the reservation wage though not the offered wage for an individual. As noted in Verbeek [2000], such variables can only be included into x_{2i} and not into x_{1i}, if their expected coefficient in x_{1i} is zero. At the same time, as Heckman [1979] suggests, some variables in x_{2i}, which do not belong to the true structural equation, may appear statistically significant for w_i^* once they are included into regression (2-26), which is run on the selected sample.

2. Division bias of hourly wage rate in the official sector

The second possible problem with the hourly wage rate in both sectors is a division – biased estimates of the hourly wage rate, which is concerned with the way the hourly wage rate was generated. Due to the measurement error present in the reported hours of work, the bias is automatically

transmitted to the hourly wage rate. In this case the importance of the income effect is overemphasized, as was noted by Borjas [1996]. Therefore, we need to check and correct for the measurement error. We will do that in several steps.

First of all, we estimate the reduced log-linear form equation for the hours worked in the official sector by common OLS, including as regressors the computed hourly wage rate in the official sector and some exogenous variables (Lemieux, Fortin and Frechette [1994] suggest to include age, age squared as a proxy for the working experience, as well as education and gender dummies). Secondly, we find such instruments for the hourly wage rate, which are not correlated with the error term in the first OLS regression. Hence, age and education dummies should be excluded from the instruments list (see Literature Review for explanation). At the following stage, we run instrumental variable regression to compute the alternative coefficients for the initial log-hours equation. Finally, we check the validity of the instruments with over-identifying restrictions test and compare the OLS and IV estimators with Hausman test.

By OIR test we check a validity of the chosen instruments, i.e. whether the moment conditions for orthogonality of instruments and error terms are satisfied. Constructed OIR statistic asymptotically follows a chi-square distribution with degrees of freedom equal to the difference of the number of restrictions and the number of parameters to be estimated. If p-value of derived the OIR statistic is "high enough" (p-value > 0.1), then the condition for orthogonality is not rejected and instruments are accepted as valid.

Hausman test checks the null hypothesis of consistency of OLS estimators compared with the IV estimators, which are deemed to be consistent under both the null and the alternative hypotheses. The difference between the two estimators, weighted for appropriate variance-covariance matrix, asymptotically follows a chi-square distribution. Thus, we do not reject the consistency of OLS estimators if computed p-value exceeds the specific threshold level (we have looked for P-value>0.1).

3. Endogeneity of marginal tax rate in the official and unofficial sectors

Due to the same problem of the measurement error and unobservable wage rate for some workers, the marginal tax rate, computed from the estimated gross income and the reported disposable income (see Appendix C on proportional income tax rate and due tax amounts), is not an exogenous variable in the model. Therefore, we construct the corrected set for the post-tax earnings, taken into account the solutions to the previous two problems. Based on this corrected sample we estimate the gross income from the non-linear proportional procedure of imposing marginal tax rate, explained in Appendix C. Finally the marginal tax rate was computed according to the formula:

$$mtax^c = 1 - \frac{after - tax\ earnings^c}{gross\ earnings^c} \tag{2-35}$$

The advantages of this methodology are the following:

1) We corrected for the measurement error present in directly reported after-tax earnings;
2) We have computed the marginal income tax even for those workers, whose regular wage rate is not observable due to the self-selection.

Similar to the tax rate in the official sector, there is endogeneity involved in the marginal tax rate in the unofficial sector, which is equal in our model to the product of probability of being caught and the penalty rate for tax evasion.

Since the official declared penalties paid and the number of cases, when a worker was charged with tax evasion, underestimate the true parameters for these variables (because of underreporting), we have applied the commonly used technique in the literature. Penalties for tax evasion were generated as the proportion of the regular marginal tax rate, computed in (2-35):

$$penalty^c = \mu * mtax^c, \quad \mu > 1 \tag{2-36}$$

Probabilities of being caught were constructed on the basis of subjective qualitative responses to this question. Therefore, at the final stage, we generated four exogenous sets of penalties for tax evasion (with μ equal to 1.25, 1.5, 1.75 and 2 respectively) and three sets of subjective estimates of probabilities of being caught and charged with tax evasion.

4. Estimation of elasticities in log-linear equation for the expected unregistered earnings

Substituting $\alpha \equiv -\frac{1}{\gamma}$ into the final log-linear equation for the expected unregistered earnings (2-25), we obtain the logarithmic equation, which is linear in parameters: (2-37)

$$\ln E(Y_u^*) = \alpha * \eta + \alpha * \sum_i x_i \delta_i + \alpha * \ln(1 - p\theta) + (1-\alpha) * \ln W_0 + (1-\alpha) * \ln(1-\tau) + \alpha * \varepsilon,$$

which we can easily estimate with constrained OLS regression[9] since all the right-hand side variables are now exogenous.

We also compute the alternative estimates for standard errors of the coefficients with bootstrapping method. Bootstrapping estimates of the standard errors are based on the replications from the sample data, and, therefore, allow us to overcome the erroneous assumptions about the underlying distribution (Greene [2000], p.173, 843).

Finally, it is obvious that the estimated elasticities from equation (2-37) with respect to the marginal tax rates in both sectors do not equal parameter α, since:

$$\alpha \equiv \frac{\partial \ln E(Y_u^*)}{\partial \ln(1-\tau)} \neq \frac{\partial \ln E(Y_u^*)}{\partial \ln \tau} \quad (2\text{-}38)$$

Thus, the estimated elasticities with respect to the marginal tax rates in both sectors include the tax rates themselves:

[9] Constraints are imposed on the coefficients of the equation. See Chapter 4.

$$\frac{\partial \ln E(Y_u^*)}{\partial \ln \tau} = \frac{\partial \ln E(Y_u^*)}{\partial \ln(1-\tau)} * \frac{\partial \ln(1-\tau)}{\partial \ln \tau} = (1-\alpha) * \frac{\tau}{\tau - 1} \tag{2-39}$$

And similarly:

$$\frac{\partial \ln E(Y_u^*)}{\partial \ln(p\theta)} = \alpha * \frac{p\theta}{p\theta - 1} \tag{2-40}$$

$$\frac{\partial \ln h_u^*}{\partial \ln \tau} = -\alpha * \frac{\tau}{\tau - 1} \tag{2-41}$$

$$\frac{\partial \ln h_u^*}{\partial \ln(p\theta)} = \alpha * \frac{p\theta}{p\theta - 1} \tag{2-42}$$

The respective elasticities of the unofficial sector indicators with respect to the gross wage rate in the official sector are readily observable from the estimated regressions:

$$\frac{\partial \ln E(Y_u^*)}{\partial \ln W_0} = 1 - \alpha \tag{2-43}$$

$$\frac{\partial \ln h_u^*}{\partial \ln W_0} = -\alpha \tag{2-44}$$

Elasticities (2-39) - (2-42) will be estimated in two ways: at the sample means, and, as the mean of the sample elasticities in order to compare the predictions of the two.

Chapter 3

DETERMINANTS OF TAX EVASION

Estimating determinants of tax evasion

For estimation of tax evasion determinants a simple logit function is applied, which is designated by nature of the dependent variable. In order to obtain an accurate estimator for a probability of the tax evasion, logit functions are constructed for the two sets of dependent variables. In the first approach a dependent dummy variable indicates positive unregistered labor-hours, whereas in the second approach a dummy is based on an unregistered earnings. Application of the two estimators is instructive in the sense that it reveals different significant explanatory factors, which allow for more comprehensive conclusions on the incentives behind shadow labor activities. However, the first approach is suggested to be more reliable comparing to the second one for the several reasons:

(1) Dummy for unregistered hours is generated directly from the respondents' answers, whereas dummy for unregistered earnings is constructed in three stages, including division of unreported family earnings, and thus, the latter is more exposed to bias[10].

(2) Due to social and psychological reasons, people are more reluctant to expose their earnings than working hours, so that in general an estimator based on hour-dummy is closer to the true population parameter.

Both approaches consider the following logit function:

$$p(y=1) = \frac{1}{1+e^{-Z}}, \tag{3-1}$$

$$Z = \sum_j \alpha_j x_j + \varepsilon_j \tag{3-2}$$

[10] At the first stage, an individual defines whether the firm he is working for and his job is registered. At the second stage, an individual defines the shares of his income sources. At the third stage, the reported consumption expenditures, which exceed the reported total income of the household, are divided among the individuals, who report an unofficial occupation. If there are no such individuals, then the excessive expenditures are divided proportionally among the adults of the household.

whereas: $Z = \ln\left(\frac{p}{1-p}\right)$ – log of odds ratio in favour of tax evasion,

x_j - factors that influence individual's decision to work unofficially,

and p – probability of tax evasion.

A marginal impact of the particular factor in both approaches was evaluated at means of the explanatory variables and as a sample average of individual marginal effects. Since no significant difference was detected between the two estimates, only the sample average of the individual marginal effects is presented below. In the logit model a marginal effect of the particular factor depends on the values of the other explanatory variables:

$$\frac{\partial E[y \mid x\ ]}{\partial x} = \Lambda(\alpha' x) * [1 - \Lambda(\alpha' x)] * \alpha\ , \tag{3-3}$$

$$\text{where } \Lambda(\alpha' x) = \frac{1}{1 + e^{-Z}} = prob\ (y = 1) \tag{3-4}$$

Overall significance of the models are tested with likelihood ratio (LR test) and likelihood ratio index (LRI), as suggested in Greene [2000, p.831]. Under a null hypothesis of all zero coefficients, likelihood ratio statistics follows a chi-square distribution with the degrees of freedom equal to the number of the regressors in the model (except the constant term):

$$LR = -2(\ln L_r - \ln L_{ur}) \tag{3-5}$$

Both specifications rejected the null hypothesis of zero slopes in logit models with zero probability of LR statistics exceeding the critical values of chi-squared distribution (see Table 2, Table 3).

An analogue to the conventional R^2 in qualitative response models is McFadden's likelihood ratio index:

$$LRI = 1 - \frac{\ln L_{ur}}{\ln L_r} \tag{3-6}$$

However, it is suggested that high precision of the model can only be obtained with $\lim(\alpha' x) \to \pm\infty$. On the other hand, if it is the case, the model has a flaw. Therefore, for the purposes of the current analysis, LRI are

presented to compare the precision of the two specifications rather than the overall significance of the model. The same argument pertains to the predicted probabilities of tax evasion, evaluated at the estimated coefficients of logit function.

Specification for unregistered labor-hours

The maximum likelihood estimates for a specification of unregistered labour-hours are presented in Table D1, Appendix D. The Table 2 below summarizes the determinants of unregistered labour-hours and their marginal impact on a probability of the tax evasion.

Table 2

Determinants of unregistered labour-hours and their marginal impacts

Dependent variable: dummy for unregistered hours

Factors	**Coefficient**	**Odds ratio**	**p>|z|**	**Average marginal effect**
Hour_leisure	-0.0230738**	0.9771904	0.000	-0.00249
Hour_registered	-0.0379222**	0.9627879	0.000	-0.00409
Hour_r_registered	0.0148016**	1.014912	0.003	0.001597
Hour_housework	-0.0261387**	0.9742	0.000	-0.00282
Total earnings	0.0008528**	1.000853	0.000	9.2E-05
Registered income	-0.0013768**	0.9986241	0.001	-0.00015
Dnipropetrovska	-0.6258045**	0.534831	0.000	**-0.06753**
Kyivska	-0.845336**	0.4294131	0.000	**-0.09122**
W_trade&catering	0.7214868**	2.05749	0.002	**0.077852**
Excessive taxes	2.915855**	18.46459	0.000	**0.314636**
Urban	-0.3823666**	0.6822449	0.005	**-0.04126**
Not related extra job	2.097206**	8.143382	0.000	**0.226299**
Same position	1.584453**	4.876623	0.000	**0.170971**
Secondary vocational	0.4807258**	1.617248	0.003	**0.051873**
Higher education	0.6103662**	1.841105	0.001	**0.065862**
Number of adults	-0.2158014**	0.8058953	0.001	**-0.02329**
Age	-0.0075084*	0.9925197	0.067	-0.00081
Constant	10.7449**	na	0.000	na
Log likelihood unrestricted		-793.289		
Log likelihood restricted		-1318.59		
LR chi^2 (17)		1050.612		
Probability>chi^2		0.0000		
LR index		0.398383		

* - significance at 10% critical level, ** - significance at 1% critical level

Likelihood ratio test rejects the null hypothesis of joint insignificance of the included regressors. Likelihood ratio index indicates that approximately 39.83% of variation of the dependent variable is explained by the model. Finally, the last column of Table 2 describes the average marginal impact of the explanatory variables, holding all other factors fixed. Therefore, on average:

1) Excessive taxation is the most important factor for the tax evasion and it increases an individual's probability of working unofficially by 0.315 points.
2) If a job offer in shadow sector is not related to the regular work or the same position as at the regular work, then the probability of working unofficially rises by 0.226 and 0.171 points respectively.
3) Secondary vocational and higher education add 0.052 and 0.066 points to a probability of the unofficial occupation; however, these numbers do not suggest that education enhances incentives to work unofficially, but rather that education increases the chances to find a work at the secondary (shadow) job market ("moonlighting").
4) Self-employed individuals in a trading and public catering sector are more likely to perform the unofficial economic activities; and all else being equal, being employed in trade raises the chances of the tax evasion by 0.078 points.
5) An urban residence offers a larger job pool, which is reflected by the lower chances of being employed unofficially (-0.04126 points). The same argument is confirmed by a lower probability of the tax evasion in the largest oblasts of Ukraine – Kyivska and Dnipropetrovska (-0.0912 and –0.06753 points respectively). Therefore, large cities and an urban residence in general increase the chances of lawful behaviour through a larger pool of available jobs in the official sector.
6) Finally, an increase in the number of adults in a family raises the regular income per family member; therefore, a probability of the tax evasion falls by 0.02329 points.

The rest of the significant factors do not affect the likelihood of the shadow labour occupation considerably. However, it is interesting to note that the effect of hours devoted to regular registered sector positively impacts the probability of the tax evasion. We suggest that the sign can be explained by the multiple job holdings pertinent to Ukrainian labour force. Thus, the workers that perform the official economic activities are more likely to be employed at the unofficial sector.

Given the coefficients of the estimated logit, a probability of the tax evasion computed at the sample means of the regressors equals to 0.2028.

Specification for unregistered earnings

Table D2, Appendix D presents the logit regression results of unregistered earnings specification. The following Table 3 describes the coefficients of the significant factors and their marginal impact on a probability of generating unregistered earnings.

According to the specification of unregistered earnings, all other factors being equal, on average:

1) Wage arrears in the regular sector increases the probability of generating unregistered earnings by 0.056 points.
2) Possession of own land plot and occupation in an agriculture sector are the implicit sources of households' consumption. Consequently, these factors reduce the necessity to look for other unregistered sources of earnings by 0.044 and 0.136 points.
3) Having a share in the regular enterprise's income makes a shareowner interested in retaining his regular job. Therefore, a joint stock ownership of the regular job placement reduces the probability of unregistered earnings by 0.093 points.
4) High concentration of population (due to either seasonal character or historical industrial concentration) increases the available pool of jobs in primary official sector and the chances of earning the regular registered income. Thus, high regular seasonal earnings in Crimea and

heavy industry profile of Donetska oblast diminish a worker tax evasion.

Table 3

Determinants of unregistered earnings and their marginal impacts

Dependent variable: dummy for unregistered earnings

Factors	**Coefficient**	**Odds ratio**	**p>|z|**	**Average marginal effect**
Hour_leisure	0.0318887**	1.032403	0.000	0.006147127
Hour_unregistered	0.0317442**	1.032253	0.000	0.006119
Hour_registered	0.0322627**	1.032789	0.000	0.006219222
Hour_housework	0.030628**	1.031102	0.000	0.005904
Hour_land plot	0.0314823**	1.031983	0.000	0.006069
Total earnings	0.0082839**	1.008318	0.000	0.001597
Non-labour income	-0.0077505**	0.9922794	0.000	-0.00149
Registered_r income	-0.0082956**	0.9917387	0.000	-0.0016
Land plot	-0.2307416**	0.7939446	0.026	**-0.04448**
Proceeds_land plot	-0.0003455**	0.9996546	0.001	-6.7E-05
Bribe	-0.000046**	0.999954	0.045	-8.9E-06
Wage arrears	0.2922394**	1.339424	0.048	**0.056334**
Joint stock	-0.4819214**	0.6175956	0.036	**-0.0929**
Agriculture&forestry	-0.7077404**	0.4927564	0.025	**-0.13643**
W_housing&communal	-0.4486464*	0.6384918	0.055	**-0.08648**
Crimea	-0.725062**	0.4842945	0.012	**-0.13977**
Donetska	-0.4778805**	0.6200963	0.000	**-0.09212**
Constant	-16.95326**	na	0.000	na
Log likelihood unrestricted		-1256.79		
Log likelihood restricted		-1494.27		
LR chi^2 (17)		474.9569404		
Probability>chi^2		0.0000		
LR index		0.158926		

* - significance at 10% critical level, ** - significance at 5% critical level

5) A regular occupation in communal and housing services negatively contributes to the chances of generating unregistered earnings.

Likelihood ratio test does not support the null hypothesis of total insignificance of the included regressors. Likelihood ratio index implies that approximately 16% of variation in the dependent dummy is explained by the model factors. Finally, a probability of generating the unofficial earnings evaluated at the estimated coefficients and the sample means is 0.747.

The first approach to estimation of a probability of the tax evasion allows for the following conclusions. The most important factor that adds to probability of tax evasion is excessive taxation burden. Workers with secondary vocational and higher education levels are more likely to find an unofficial occupation at the secondary job market; thus, education increases the chances of getting a job. Existence of an extra-unregistered work at the same position at another enterprise considerably increases a probability of the tax evasion. However, if there is no relationship between the two jobs, a probability of accepting an offer for the unofficial work is also significantly high. Thus, the interdependence of the registered and unregistered work places cannot be established unambiguously. An occupation in trading and public catering increases the chances of working unofficially. Although it may seem unexpected than living in an urban area and particularly in Kyivska and Dnipropetrovska oblasts contributes to tax compliance, the evidence is obvious because of two facts. First of all, taxation enforcement in large cities is more strict and tight, which induces tax compliance. And secondly, inhabitants of an urban area have more opportunities to earn sufficient earnings from the legal sources, as well as a larger pool of jobs to choose from. Finally, every new member of the family contributes to tax compliance of the other members.

Summarizing the findings of the second approach, it can be concluded that the crops from own land plot are a substantial source for the private consumption, which is also confirmed with a lower probability of the unregistered earnings generated in agriculture and forestry sectors. Payment delay at registered work prompts searching for additional income sources. Similar to the first approach, the living area also determines the propensity to tax evasion, so that the inhabitants of Donetska oblast (pervasively involved in manufacturing) and Crimea (where most jobs are of seasonal character) have lower incentives to engage into unofficial economic activities due to higher incomes in the official sector. Granting a share in a firm's ownership

increases the chances of worker's loyalty to a regular enterprise, and therefore, reduces his prospective of generating unregistered earnings.

Although both specifications passed the likelihood ratio test, we consider the estimates of the average probability of tax evasion obtained by the first approach to be more reliable due to higher precision of the model specification (confirmed by higher likelihood ratio index) and presumably more reliable data on the unregistered hours.

C h a p t e r 4

ESTIMATION RESULTS

Selection bias of hourly wage rate in the official sector

An apparent exogenous regressor, which affects the reservation though not the offered wage rate, is the non-labor income of individual. Berndt [1990] suggests that that non-labor income for women should also include the earnings of their husbands, since the women decide to work conditional upon the income of their spouses. Similarly, this technique, called a recursive model of choice, was applied by Lacroix and Fortin [1992] to estimate the wage response of wives, considering their husbands' income as part of the non-labor income. Therefore, we run Heckman regressions (Tobit II model) for the sub-samples of men and women.

The following Table 4 presents the estimates of Heckman model for the sub-sample of women:

Table 4

Heckman selection model for the sub-sample of women

Dependent variable: log of hourly wage rate in the official sector

Wage equation:[11]			
	Coefficient	*Robust s.e.*	*P-value*
Light & food	-0.5050903	0.2958582	0.088
Agriculture & forestry	-0.9334143	0.3989027	0.019
Civil engineering	-0.8135923	0.3024302	0.007
Trade & public catering	-0.3609499	0.1927565	0.061
Health care, social security & sport	-0.4154025	0.1839104	0.024
Private ownership with Ukrainian capital	-1.085828	0.2575249	0.000
Private ownership with foreign capital	-1.818403	0.5263596	0.001
State ownership	-1.262092	0.2366741	0.000
Urban residence	0.6047598	0.131982	0.000
Constant	1.250856	0.3028775	0.000

[11] The excluded categories for the industries and the ownership structure of the regular place of work (the reference categories) can be found in Table 2, Section: Data Description.

Selection equation:			
Light & food	0.5761482	0.2717999	0.034
Agriculture & forestry	1.414065	0.5526455	0.011
Civil engineering	0.9287435	0.4070139	0.022
Trade & public catering	0.4672148	0.2016636	0.021
Health care, social security & sport	0.4086222	0.1816096	0.024
Private ownership with Ukrainian capital	1.122866	0.1989236	0.000
Private ownership with foreign capital	1.99086	0.4178021	0.000
State ownership	1.27768	0.1352441	0.000
Urban residence	-0.236006	0.1162425	0.042
Total number of people in household	-0.1690371	0.0455854	0.000
Age	0.0954072	0.0393201	0.015
Age squared	-0.0011866	0.0005357	0.027
Marital status	-0.425232	0.1230088	0.001
Non-labor income	0.0010841	0.0002511	0.000
Constant	-2.281952	0.6455021	0.000
ρ_{12}	-0.5606716	0.1374577	
σ_{12}	-0.603805	0.162844	
Number of observations	770		
Number of censored observations	536		
Log-likelihood	-658.0941		

Wald test of independent equations ($\rho_{12} = 0$):

Chi-squared statistic (1) = 10.00, probability>chi-squared = 0.0016

Thus, as the results of Heckman selection model for the sub-sample of women suggest there is a negative correlation between the error terms in structural and selection equations and OLS would provide inconsistent estimates for the wage equation. The negative correlation coefficient indicates that there is an omitted variable, which negatively impacts the decision to work and positively the offered wage rate, or vice versa. The negative sign in front of the omitted variable in this case implies that the coefficients of the initial wage equation underestimate the influence of the listed regressors.

In order to check the validity of exogenous variables added to the selection section, we run OLS regression with the regressors included in the first part of Table 4 and those in bold in the second part of Table 4 to verify that the coefficients are not statistically significant in the wage equation. If the

coefficients of the latter would not be statistically different from zero, then the exogenous variables can be potential regressors to the selection equation. The following Table 5 presents the selected results of F-test on the exogenous variables in bold in the second part of Table 4.

Table 5

Results of F-test for the selected exogenous variables in wage equation:

Dependent variable: log of hourly wage rate in the official sector

Regressors	F-statistic	Probability>F
Total number of people in household	F (1, 223) = 0.38	0.5375
Age	F (1, 223) = 0.25	0.6160
Age squared	F (1, 223) = 0.41	0.5209
Marital status	F (1, 223) = 0.45	0.5041
Non-labor income	F (1, 223) = 25.53	0.0000

Obviously, the null hypothesis of zero coefficients for the respective exogenous variable is not rejected in the first four cases, although it is statistically rejected for the non-labor income regressor. However, as mentioned in Heckman [1979], this result does not indicate the omitted variable in the wage equation, but rather a clear symptom of a sample selection bias since the non-labor income affects the decision to work rather than the offered wage. Therefore, we can conclude that there is a sample selection bias present in the wage equation for the sub-sample of women and we correct the observed hourly wage rate with the predicted values obtained from Heckman selection model.

Next, we repeat the same procedure for the sub-sample of men. The following Table 6 presents the Heckman model for the sub-sample of men.

For the sub-sample of men we could only find one exogenous variable distinctive to the selection equation – non-labor income. Nevertheless, even with one different regressor in the selection equation the Wald test rejects the independence of wage and selection equation, which is the evidence for the sample selection present.

Table 6

Heckman selection model for the sub-sample of men

Dependent variable: log of hourly wage rate in the official sector

Wage equation:			
	Coefficient	*Robust s.e.*	*P-value*
Age	0.0159673	0.0069661	0.022
Agriculture & forestry	1.081337	0.2722895	0.000
Civil engineering	0.5727477	0.188471	0.002
Transport & communication	0.4705549	0.2031462	0.021
Kyivska oblast	-0.5671017	0.1685937	0.001
Donetska oblast	-0.7767235	0.1569513	0.000
Private ownership with Ukrainian capital	0.6954403	0.2246985	0.002
State ownership	0.9626877	0.2457579	0.000
Total number of people in household	-0.2003657	0.0689089	0.004
Marital status	0.3768593	0.173683	0.030
Constant	-1.572112	0.4391952	0.000
Selection equation:			
Age	0.0285586	0.0044464	0.000
Agriculture & forestry	0.9658349	0.2982341	0.001
Civil engineering	0.4276048	0.2500669	0.087
Transport & communication	0.6561402	0.2255605	0.004
Kyivska oblast	-0.4049814	0.1444373	0.005
Donetska oblast	-0.3300413	0.1348044	0.014
Private ownership with Ukrainian capital	1.193089	0.1526772	0.000
State ownership	1.415201	0.1244436	0.000
Total number of people in household	-0.2546317	0.0485425	0.000
Marital status	0.3909561	0.134365	0.004
Non-labor income	-0.0113717	0.0043534	0.009
Constant	-1.285167	0.2453429	0.000
ρ_{12}	0.6801974	0.1445305	
σ_{12}	0.7780726	0.2077771	
Number of observations	720		
Number of censored observations	408		
Log-likelihood	-781.6619		

Wald test of independent equations ($\rho_{12} = 0$):

Chi-squared statistic (1) = 9.51, probability>chi-squared = 0.0020

A positive sign of the correlation coefficient suggests that there exists an omitted variable, which positively affects both the decision to accept the offered wage and the observable wage. Additionally, a positive coefficient in front of the omitted variable indicates that the coefficients in the initial wage equation overestimate the impact of the included regressors.

Again, as in the previous example for women, we test the zero coefficient for the non-labor income in the OLS regression for the wage equation:

Table 7

Results of F-test for the selected exogenous variable in wage equation:

Dependent variable: log of hourly wage rate in the official sector

Regressor	F-statistic	Probability>F
Non-labor income	F (1, 300) = 2.21	0.1381

F-test does not reject the zero coefficient for the non-labor income in wage equation, and therefore this exogenous variable is a valid regressor for the selection equation. Summarizing, Heckman model indicates the presence of a positive correlation between the error terms of the wage and the selection equation, which renders OLS coefficients obtained from the wage equation inconsistent. To treat this bias, we adjust our observations of the wage rate for men with their predicted values from the Heckman model.

Division bias of hourly wage rate in the official sector

Another problem, which arises in estimation of the hourly wage rate equation, is a division bias, i.e. hourly wage rate is subject to implicitly transmitted measurement error in the reported hours of work. In order to correct for this kind of error, we performed an instrumental variable regression.

At the first stage we used simple log-linear OLS regression of hourly wage rates on the reported hours in the official sector, including several other exogenous variables (see Table 8). If there is a division bias present, then the

log of hourly wage rate in the right-hand side of the equation is correlated with the error term and the reported coefficients are inconsistent.

Table 8

OLS log-linear regression of hourly wage rate on hours worked in the official sector

Dependent variable: log of hours worked in official sector

Regressors	*Coefficient*	*Robust s.e.*	*P-value*
Log (wage rate)	-0.2818405	0.0457019	0.000
Age	0.3154314	0.0337323	0.000
Age squared	-0.0038746	0.0004694	0.000
Incomplete secondary ed.	1.503297	0.3040007	0.000
Secondary education	2.083645	0.1748398	0.000
Secondary vocational ed.	2.64806	0.1530762	0.000
Higher education	3.125131	0.1554658	0.000
Marital status	0.4877053	0.1395676	0.000
Total number of people in household	-0.1335051	0.0503874	0.008
Constant	-5.543034	0.5850744	0.000
R-squared	0.1938	Number of observations	1490
F (9, 1480)	84.66	Probability>F	0.0000

The gender dummy appeared to be insignificant for OLS regression, therefore, we include the marital status dummy instead. At the next stage, we look for the valid instruments for the log of hourly wage rate in the official sector, requesting that they are uncorrelated with the error term in the first OLS regression. If we can find such instruments that could explain a sufficient part of variation in the hourly wage rate, they can be substituted into the initial OLS regression to correct the possible measurement error of the endogenous variable. Table 9 describes the chosen instruments and their respective F-test for zero coefficients in the initial OLS regression.

Apparently, these factors account for 66% of variation in the hourly wage rate and can be potential candidates for instrumental variable regression. F-test does not reject the validity of the chosen instruments in any case.

Table 9

OLS log-linear regression of hourly wage rate on its possible instruments

Dependent variable: log of hourly wage rate in official sector

Regressors	*Coefficient*	*Robust s.e.*	*P-value*
Gender dummy	-2.146663	.0402394	0.000
Non-labor income for women	.0002246	.0000664	0.001
Total number of adults in the household	-.0819246	.020671	0.000
Constant	1.160912	.0662856	0.000
R-squared	0.6565	Number of observations	1490
F (3, 1486)	953.11	Probability>F	0.0000
Result of F-test in the initial OLS regression			
Regressors		F-statistic	Probability>F
Gender dummy		F (1, 1479) = 0.05	0.8201
Non-labour income for women		F (1, 1479) = 0.65	0.4187
Total number of adults in the household		F (1, 1479) = 0.03	0.8628

Therefore, we can check the instrumental variable regression on hours worked, where the regressors include all the exogenous variables of initial OLS regression and the tested instruments presented at the Table 9. The following Table 10 summarizes the output of IV regression.

The results reveal only slight changes in the coefficients values. At the next stage we test the validity of the instruments for the hourly wage rate in the official sector with the over-identifying restrictions (OIR) test. Under the null hypothesis of zero correlation between the chosen instruments and the error term in the initial OLS regression (see Table 8), OIR-statistics follows a chi-squared distribution with 3 degress of freedom. The computed OIR-statistic is 1.8630329 with probability of observing such value p= 0.60131525. Therefore, OIR test suggests that we cannot reject the null hypothesis and the validity of the chosen instruments.

Finally, we check the presence of the systematic inconsistency in the OLS estimates comparing to those produced by IV regression with Hausman test.

Table 10

IV log-linear regression of hourly wage rate on hours worked in the official sector

Dependent variable: log of hours worked in official sector

Regressors	*Coefficient*	*Robust s.e.*	*P-value*
Log (wage rate)	-0.2692481	0.0560046	0.000
Age	0.3152693	0.0351213	0.000
Age squared	-0.0038726	0.0004795	0.000
Incomplete secondary ed.	1.505051	0.4079141	0.000
Secondary education	2.077356	0.2891415	0.000
Secondary vocational ed.	2.643463	0.2747628	0.000
Higher education	3.119554	0.2791224	0.000
Marital status	0.4862649	0.1385978	0.000
Total number of people in household	-0.1327893	0.0494869	0.007
Constant	-5.535831	0.6590826	0.000
R-squared	0.1937	Number of observations	1490
F (9, 1480)	37.84	Probability>F	0.0000

The computed Hausman statistic is 0.15 with the probability of observing such value p=0.6986. This result implies that there is no statistical evidence to reject the null hypothesis of consistency of OLS estimates.

We checked the existence of the division bias in the hourly wage rate with IV estimation and OIR and Hausman tests, which justify the applicability of the certain regression type. However, the tests' results reveal that we cannot reject the consistency of OLS estimates presented in Table 8. Thus, we may infer that the division bias does not affect the consistency and robustness of the estimates, based on the computed hourly wage rate in the official sector. Furthermore, the confirmed consistency of OLS estimates in the initial hours equation allows us to predict the estimated elasticity of hours worked to the hourly wage rate in the official sector. Thus, the doubling of hourly wage rate induces a 28% reduction of hours worked in the official sector, which is a slope coefficient in front of hourly wage rate presented in Table 8.

Estimation of elasticities in log-linear equation

for the expected unregistered earnings

Estimation of the parameters in the unregistered earnings equation (2-37) were done with the constrained OLS with the imposition of the following constraints on the regression coefficients:

$$\ln E(Y_u^*) = \alpha * \eta + \alpha * \sum_i x_i \delta_i + \alpha * \ln(1 - p\theta) + (1-\alpha) * \ln W_0 + (1-\alpha) * \ln(1-\tau) + \alpha * \varepsilon,$$
$$\ln E(Y_u^*) = \alpha_1 * \eta + \alpha_2 * \sum_i x_i \delta_i + \alpha_3 * \ln(1 - p\theta) + \alpha_4 * \ln W_0 + \alpha_5 * \ln(1-\tau) + \varepsilon_1, \qquad (4-1)$$

Constraints: $\alpha_3 + \alpha_4 = 1; \quad \alpha_4 = \alpha_5.$ (4-2)

Constrained OLS regressions were run for twelve different regimes of the marginal tax rate in the unofficial sector (see: Section *Estimation Strategy*), which included four different penalties' sets and three probabilities' sets. Appendix E lists all 12 regressions, including the alternative estimates for the standard errors. Coefficient α_3 was significant in all regression, whereas α_4 and α_5 were significant in 4 out of 12 regressions. The following Table 11 presents the combined results of these regressions for the parameter α with OLS standard errors and standard errors, obtained by bootstrapping technique. The last column of the Table indicates the estimated coefficient $\gamma = -\frac{1}{\alpha}$, which is a curvature parameter from the wage equation in the unofficial sector.

We may conclude from Table 11 that the estimated elasticities are consistent and robust to the assumption of the underlying distribution, which is inferred from similar values of standard errors provided by constrained OLS and bootstrapping technique. The last column of Table 11 indicates the hourly wage rate elasticity with respect to the hours worked in the unofficial sector.

Table 11

Constrained log-linear regression on expected unregistered earnings

Dependent variable: log of expected earnings in the unofficial sector

Regime[12]	α	Constrained OLS s.e.	P-value	Bootstrapping s.e.	γ
Regime 11	0.9520042	.0417975	0.000	.0381199	-1.050415
Regime 12	0.9505392	.041757	0.000	.0488653	-1.052034
Regime 13	0.9532467	.0418328	0.000	.0393315	-1.049046
Regime 21	0.9441311	.0414916	0.000	.0407263	-1.059175
Regime 22	0.9421473	.0414376	0.000	.0401857	-1.061405
Regime 23	0.9457429	.0415369	0.000	.0389864	-1.057369
Regime 31	0.934064	.0411155	0.000	.0361488	-1.070590
Regime 32	0.9314203	.0410447	0.000	.0382662	-1.073629
Regime 33	0.9361042	.0411725	0.000	.0428361	-1.068257
Regime 41	0.9201254	.0406177	0.000	.0416811	-1.086808
Regime 42	0.9166187	.0405252	0.000	.038367	-1.090966
Regime 43	0.9226578	.040688	0.000	.0368036	-1.083825

Thus, as we claimed in the theoretical part, the hourly wage rate in the unofficial sector is indeed convex and negatively related to hours worked. Apparently, a 1% in hours worked in unofficial sector would entail approximately 1.05% reduction in hourly wage rate.

Finally, Table 12 presents the estimated elasticities of the unofficial hours with respect to the gross wage rate in the official sector, marginal tax rates in both sectors, computed at the sample means and at the means of the sample estimates.

Table 12 indicates that the most influential factor for the decision of a worker to devote labor-efforts to the unofficial sector is the gross wage rate in the official sector. Thus, if the gross wage rate in the unofficial sector rises by 1%, the hours devoted to the unofficial sector would decline by 0.94% on average. The estimated mean elasticities with respect to the marginal tax rates in both sectors revealed higher values than the respective indicator estimated at the sample means of the marginal tax rates.

[12] Regimes are named such that the first number represents the penalties' set and the second number represents the probabilities' set. For further explanation, see Appendix E.

Table 12

Estimated elasticities of the unofficial hours

Regime	$\varepsilon(h_u, W_0)$	$\varepsilon(h_u, \bar{\tau})$	$\bar{\varepsilon}(h_u, \tau)$	$\varepsilon(h_u, \bar{p}\bar{\theta})$	$\bar{\varepsilon}(h_u, p\theta)$
Regime 11	-0.9520042	0.1440914	0.1651011	-0.0494285	-0.0613669
Regime 12	-0.9505392	0.1438697	0.164847	-0.0515167	-0.0641569
Regime 13	**-0.9532467**	**0.1442795**	**0.1653166**	-0.0473706	-0.0588777
Regime 21	-0.9441311	0.1428998	0.1637357	-0.0594409	-0.0792454
Regime 22	-0.9421473	0.1425995	0.1633917	-0.0619457	-0.0829288
Regime 23	-0.9457429	0.1431437	0.1640152	-0.0569634	-0.0760676
Regime 31	-0.934064	0.141376	0.1619898	-0.0693358	-0.1016842
Regime 32	-0.9314203	0.1409759	0.1615313	-0.0722388	-0.1065071
Regime 33	-0.9361042	0.1416849	0.1623436	-0.066447	-0.0977094
Regime 41	-0.9201254	0.1392664	0.1595725	-0.078895	-0.1343814
Regime 42	-0.9166187	0.1387356	0.1589644	**-0.0821569**	**-0.1407575**
Regime 43	-0.9226578	0.1396497	0.1600117	-0.0756154	-0.1294726
Average	**-0.9374002**	**0.1335477**	**0.1625684**	**-0.0642796**	**-0.0826998**

However, the conclusion is apparent: a worker decision is more sensitive to the marginal tax rate in the official sector than in unofficial sector. Consequently, a 1% increase in the marginal income tax would induce an increase in the hours worked in unofficial sector by 0.13 – 0.16%. A marginal increase in the punishment and enforcement efforts would only induce 0.06 – 0.08% reduction in the labor-efforts in the unofficial sector.

Policy implications and discussions

Summarizing the analysis of expected unregistered earnings, maximized at the optimal hours of work devoted to the unregistered sector, we can conclude:

1. Improvement in education allows for larger hours worked in the official sector. Combining this result with the hourly exogenous wage rate in the official sector, we may infer that given the same other social characteristics, education results in higher earnings in the official sector. At the same time, the results of logit estimation suggest that higher levels of education raise the chances to find an employment in the unofficial economy. Therefore, we may conclude

that increasing one's education status enhances the future prospects of finding a work, whether it would be at the regular or irregular job markets.

2. The hourly wage rate in the unofficial sector is negatively related to the unofficial hours of work, suggesting the existence of the upper limit on the possible output expansion in the unofficial sector.
3. Hours devoted to unofficial sector are highly sensitive (almost in one-to-one proportion) to the gross wage rate in the official sector, which implies that the enhancing of productivity and payments in the official sector can realize a transformation of the underground production to the official economy.
4. While choosing the instruments to lessen the worker's incentive to join the unofficial economy, hence to reduce the scope of shadow activities, the government should concentrate on the direct taxation in the regular sector, rather than enforcing the tax compliance. The results of the analysis suggest that Ukraine is on the backward-bending side of the Laffer curve, where an increase in the tax rate would entail a reduction of the generated tax revenues due to the transforming of the labor efforts to the unofficial economy.
5. Finally, the postulated results appears to be robust to the different penalties regimes and probabilities of being caught in tax evasion; moreover, the values of the estimated effects are consistent and robust to the underlying distribution of the error terms.

Although the inferences are based on the particular modal applied and there exist alternative methods to estimate the size and dynamics of unofficial economy, the results obtained are striking and straightforward to suggest the policy improvements. Taxation and low wage payments in the official sector are the major determinants of a worker's decision to search for employment in the unofficial economy, even if there exists a negatively sloped return function to the devoted efforts. Thus, not only from the businesses' point of view but also from the standpoint of employees themselves, imperfections in

the tax system undermine the official economy. These effects outweigh the effects of the poor institutions, presented in the model by enforcement efforts of tax compliance (probabilities and penalties for tax evasion).

Finally, the model can be improved in several aspects. First of all, the availability of the dynamic data would allow for more precise estimates and substantial results concerning the responses to the changes in income (income and substitution effects). Dynamic model can also incorporate other benefits of the regular sector, which have not been accounted for in our model, such as pension payments. Secondly, it is also desirable to find some alternative estimates for the subjective valuation of probabilities. Thirdly, there could be fixed costs of a regular work, which raise the reservation wage of a worker and make an irregular employment a more attractive alternative. Finally, a completely unexplored issue (although it may not be pertinent to the citizens of Ukraine) is the utility (or disutility) derived from honest behavior, such as working in the official sector.

The current model can also be considered in a general equilibrium setting, where the businesses and government introduced explicitly. A general equilibrium approach would include the aggregate labor demand (business' side) and aggregate labor supply (household's side) functions, as well as a cost-benefit analysis for the government of increasing the tax rates or enforcement efforts.

We attempt to explain the existence of the shadow economic sector in Ukraine from the point of view of its participants, in particular, workers. The statistics suggests that the shadow economy occupies 35.5% of the country's total GDP, and our data set reveals that approximately 71% of the respondents are involved into the shadow economy. There are negligible empirical investigations of the shadow economy in transition economies, which do not proceed further than defining the determinants of tax evasion or measuring the size of the shadow sector. Therefore, a lack of empirical application and enormous scope of shadow economy intensity in Ukraine motivate the topic of our research.

Preliminary analysis of the data set, generated upon the household survey in Ukraine in 2000, implies a vague relationship between the hourly earnings and hours worked in the official sector and a transparent negative non-linear interdependence between the same parameters of the unofficial economy. This peculiarity stipulates the assumptions of the theoretical framework applied in the thesis.

A household's decision to devote the labour-efforts to the regular and irregular economic sectors is modelled as a constrained utility optimisation of the composite consumption good and the hours of leisure left. The solution to the model yields an optimum for a function of the unregistered earnings, which depends on the set of exogenous and endogenous model's parameters.

The applied estimation strategies allow for a unique and consistent estimation of the parameters introduced in the unregistered earnings function. We find the evidence for a selection bias present in the generated data set, and correct for the hourly earnings with a Heckman selection procedure. In addition, we deal with the endogeneity of the marginal tax rates in the official and unofficial sectors by an instrumental variable technique. Although common for the reported hours worked, a division bias (measurement error

in hours worked, which is directly transmitted to the hourly earnings) is not detected in our analysis. The ultimate elasticities of the hours worked with respect to the variables of our interest are computed by the constrained OLS procedure, after the adjustment for the mentioned biases. Moreover, these elasticities appear to be robust to the different regimes of the marginal tax rate in the unofficial sector and the assumptions of the underlying distribution of the error terms.

We also analyse the marginal impacts of the determinants of tax evasion by constructing two logit functions, based on hours worked and computed unregistered earnings in the unofficial sector. The results of the logit estimation suggest that excessive tax burden, related extra job, regular employment in trade and public catering sectors and wage arrears at the regular work positively impact the probability of tax evasion. At the same time, urban residence, an increase in the number of adults in a household, possession of land plot or primary occupation in agriculture as well as a joint stock ownership of a regular work improve the chances of tax compliance.

The estimated parameters of the unregistered earnings function suggest that the labour-efforts devoted to the unofficial sector are highly sensitive to the gross hourly wage rate in the official sector (the elasticity equals to –0.94), marginal income tax rate in the official economy (the elasticity equals to –0.13) and the enforcement efforts directed toward tax compliance (the elasticity equals to –0.06). Thus, we claim that a transmission of the shadow sector into the officially producing economy is more likely to be affected by a reduction in the direct taxes rather than by enhancement of the enforcement efforts. The study also finds an evidence for the complementary relationship between the hours worked in the official and unofficial sectors. Additionally, the empirical results imply that an increase in education level improves the chances of a worker to find an employment in both economic sectors.

Finally, we suggest that the model can be improved by introducing a dynamic aspect, the better measures of the probabilities and punishment rates

for the tax evasion, an upper limit on hours worked in the official sector and a further construction of the general equilibrium model with the direct involvement of businesses and government.

BIBLIOGRAPHY

A World Bank country study. *Ukraine. Restoring growth with equity: A participatory country economic memorandum.* World Bank, 2000.

Analytical note: Tendencies of shadow economy in Ukraine. Ministry of Economy and European Integration, Economic Strategy Department, Issue 2, January 2002.

Basu Kaushik, Genicot Gerence, Stiglitz Joseph E. *Unemployment and wage rigidity when labour supply is a household decision.* Brookings Institution and Stanford University, 2000.

Becker S. Gary. *A theory of the allocation of time.* The Economic Journal, Vol.75, Issue 299 (September): 493-517, 1965.

Berndt, Ernst R. *The practice of econometrics: classic and contemporary.* Addison-Wesley Publishing Company, Inc., 1991.

Borjas George J. *Labour economics.* New York: McGraw-Hill, 1996.

Burtless Gary, Hausman Jerry A. *The effects of taxation on labour supply: Evaluating the Gary negative income tax experiment.* The Journal of Political Economy Vol.86, Issue 6 (December): 1103-1130, 1978.

Cagan Phillip. *The demand for currency relative to the total money supply.* The Journal of Political Economy Vol.66, Issue 4 (August): 303-328, 1958.

Clotfelter Charles T. *tax evasion and tax rates.* The review of economics and statistics, Vol.65, Issue 3 (August): 363-373, 1983.

Commander Simon, Tolstopyatenko Andrei. *A model of the informal economy in the transition setting.* EBRD and the World Bank, 1999.

Debraj Ray. *Development Economics.* Princeton University Press, 1998.

Decree of President of Ukraine #13-92. *On increasing non-taxable minimum and proportional income tax rates of citizens (with changes and additions).* December 26, 1992.

Earle John S. and Sakova Zuzana. *Entrepreneurship from scratch: lessons on the entry decision into self-employment from transition economies.* SITE Working paper #145, 1999.

Ehrenberg Ronald G., Smith Robert S. *Modern labour economics: Theory and public policy.* –7th ed. Addison-Wesley, 2000.

Eilat Yair and Zinnes Clifford. *The shadow economy in transition countries: Consequences for economic growth and donor assistance.* CAER II Discussion Paper No.83, 2000.

Frederiksen Anders, Graversen Ebbe Krogh, Smith Nina. *Overtime work, dual job holding and taxation.* IZA Discussion paper #323, July 2001.

Friedman Eric, Johnson Simon, Kaufmann Daniel, Zoido-Lobaton Pablo. *Dodging the grabbing hand: The determinants of unofficial activity in 69 countries.* World Bank, 1999.

Giles David E.A. *The underground economy: minimizing the size of government.* University of Victoria, March 1998.

Greene William H. *Econometric analysis*, 4th ed. Prentice Hall, New Jersey, 2000.

Gronau Reuben. *Leisure, home production, and work – the theory of the allocation of time revisited.* The Journal of Political Economy Vol.85, Issue 6 (December): 1099-1124, 1977.

Guariglia Alessandra, Kim Byung-Yeon. *The dynamics of moonlighting: What is happening in the Russian informal economy?* University of Essex, 2001.

Guasch Luis J., Hahn Robert W. *The costs and benefits of regulation: Some implications for developing countries.* Background Paper for World Bank Development Report, March 1997.

Hausman Jerry A., Poterba James M. *Household behaviour and the tax reform act of 1986.* The journal of economic perspectives, Vol.1, Issue 1 (Summer): 101-119, 1987.

Hausman Jerry A., Wise David A. *Social experimentation, truncated distributions, and efficient estimation.* Econometrics, Vol.45, Issue 4 (May): 919-938, 1977.

Hausman Jerry and Ruud Paul. *Family labour supply with taxes.* The American Economic Review, Vol.74, Issue 2 (May): 242-248, 1984.

Heckman James J. *A partial survey of recent research on the labour supply of women.* The American Economic Review, Vol.68, Issue 2 (May): 200-207, 1978.

Heckman James J. *Sample selection bias as a specification error.* Econometrica Vol.47, Issue 1 (January): 153-162, 1979.

Heckman James J. *Varieties of selection bias.* The American economic Review, Vol.80, Issue 2 (May): 313-318, 1990.

Heckman James J. *What has been learned about labour supply in the past twenty years?* The American Economic Review, Vol.83, Issue 2 (May): 116-121, 1993.

Johnson Simon, Kaufmann Daniel, McMillan John, Woodruff Christopher. *Why do firms hide? Bribes and unofficial activity after communism.* World bank, 1999.

Johnson Simon, Kaufmann Daniel, Zoido-Lobaton Pablo. *Corruption, public finances and the unofficial economy*. World Bank, 1998.

Johnston Jack and DiNardo John. *Econometric methods*. New York: McGraw-Hill, 1997.

Kaufmann Daniel and Kaliberda Aleksander. *Integrating the unofficial economy into the dynamics of post-socialist economies: A framework of analysis and evidence*. World Bank, 1996.

Lacroix Guy and Fortin Bernard. *Utility-based estimation of labour supply function in the regular and irregular sectors*. The Econometric Journal Vol.102, Issue 415 (November): 1407-1422, 1992.

Lacroix Guy and Fortin Bernard. *Utility-based estimation of labour supply functions in the regular and irregular sectors*. Working Paper #9020, Department d'economique, Universite Laval, 1990.

Lemieux Thomas, Fortin Bernard, Frechette Pierre. *The effect of taxes on labour supply in the underground economy*. The American economic review, Vol.84 (1), 1994.

Loyaza A. Norman. *The economics of the informal sector: A simple model and some empirical evidence from Latin America*. Policy Research Working Paper #1727, World Bank, 1997.

Monitoring of macroeconomic and industrial indicators. Ministry of Economy and European Integration, Issue 9, September 2001.

Novoseletska Oksana and Najman Boris. *On the origins and the nature of unofficial economic activities in Ukraine*. Tacis UEPLAC, 2001.

Perspective research: Economic statistics in Ukraine. The State Committee of Statistics of Ukraine, Issue 13, November 2000.

Quarterly predictions: Research of Economy in Ukraine. The State Committee of Statistics of Ukraine, June 2001.

Richter Wolfram F., Boadway Robin W. *Trading off tax distortion and tax evasion*. University of Dortmund, April 2001.

Schneider Friedrich and Enste Dominik. *Increasing shadow economies all over the world – fiction or reality?* Arbeitspapier 9819, 1998.

Schneider Friedrich. *Regulation and the underground economy*. The Liberty Briefing, Centre for the New Europe, Issue 1, April 2001.

Slinko Irina. *Multiple jobs, wage arrears, tax evasion and labour supply in Russia*. NES Working paper #BSP/99/018, Moscow, 1999.

Stata Reference Manual, Release 6. State Press, College Station, Texas, 1999.

Stiglitz, Joseph E. *Economics of public sector.* -3rd ed., W.W. Norton and Company, 1999.

Stuart Charles E. *Swedish tax rates, labour supply, and tax revenues.* The Journal of Political Economy, Vol.89, Issue 5 (October): 1020-1038, 1981.

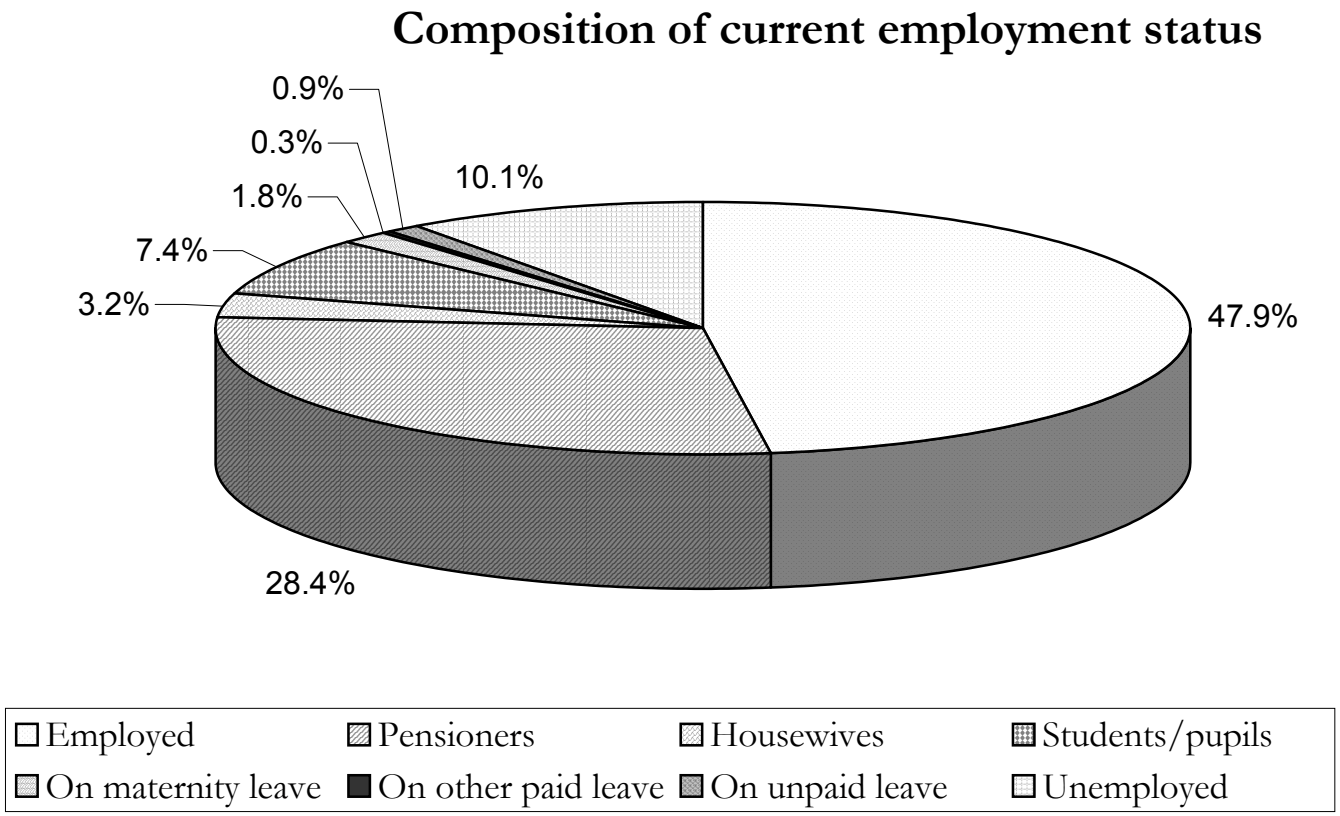
Tedds Lindsey M., Giles David E.A. *Modelling the underground economies in Canada and New Zealand: A comparative analysis.* Econometrics Working Paper EWP0003, University of Victoria, April 2000.

Triest Robert K. *The effect of income taxation on labour supply when deductions are endogenous.* University of California, Davis and National Bureau of Economic Research, 1990.

Verbeek, Marno. *A guide to modern econometrics.* Chichester, England: John Wiley and Sons, Ltd., 2000.

Appendix A

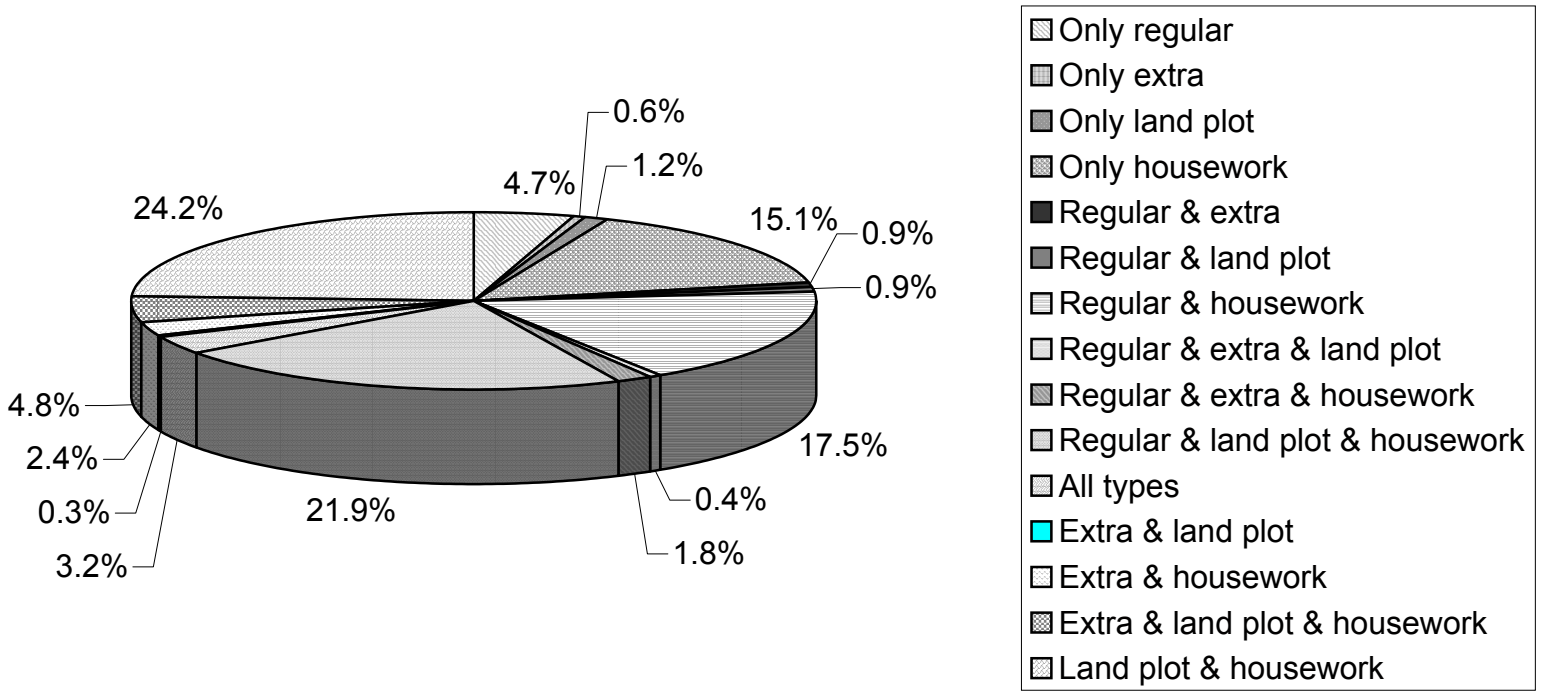
Figure A1



Appendix A

Figure A2

Composition of labor activities



Appendix B

Table B1

Composition of official and unregistered labour force in percents

Characteristics	Percent of total sample	Unregistered activity per month		
		Percent	Hours	Earnings
TOTAL	100.0	71.36	27.65	62.94
Sex:				
Male	45.23	45.70	48.91	45.50
Female	54.77	54.30	51.09	54.50
Age:				
16-24	16.15	14.65	11.56	15.24
25-39	28.21	28.40	28.51	27.45
40-59	32.18	32.69	34.22	32.94
60 and more	23.46	24.27	21.25	24.36
Education:				
Incomplete secondary	13.74	13.98	12.66	13.86
Secondary	20.65	20.64	20.31	20.80
Secondary vocational	35.00	35.65	41.41	34.39
Higher	26.26	26.21	24.84	27.18
Employment status:				
Employed	45.74	46.97	48.28	46.67
Pensioner	27.17	27.91	25.94	28.21
Housewife	3.07	3.09	2.81	3.29
Student /pupil	5.66	4.84	3.12	5.08
On maternity leave	1.86	1.51	1.09	1.65
On paid leave	0.30	0.30	0.47	0.34
On unpaid leave	0.80	0.67	1.41	0.55
Unemployed	9.68	10.11	15.00	9.40
Regular-work income per month:				
0-200	31.49	31.66	29.84	32.33
200-400	7.26	7.38	6.72	7.34
400-600	2.94	3.03	2.66	3.23
600-800	0.80	0.85	0.47	0.96
800-1000	0.39	0.42	0.31	0.41
1000-2000	0.91	0.97	0.94	0.96
2000 and more	0.26	0.18	0.16	0.21
Regular-work hours per month:				
0-36	0.73	0.73	0.78	0.69
37-72	1.43	1.45	0.94	1.51
73-108	1.90	1.69	1.41	1.85
109-144	2.98	3.03	2.81	2.95
145-180	17.06	17.31	17.50	16.82

181 and more	18.75	18.95	16.56	19.35
Wage arrears at regular work	14.47	15.13	14.84	15.24
Own land plot	61.90	60.05	60.94	59.78
Proceeds from crop	18.45	17.07	19.06	17.23
Ownership of regular place of work:				
Private (Ukrainian)	11.79	12.17	14.22	12.01
Private (with foreign capital)	1.04	0.85	0.78	0.82
Collective	3.46	3.03	2.97	2.95
Joint Stock	5.01	4.90	6.56	4.74
State	26.61	27.54	25.00	27.73
Industry in the regular job:				
Light and food	2.50	2.48	2.03	2.47
Woodworking and pulp and paper	0.30	0.30	0.47	0.27
Agriculture and forestry	2.72	2.42	2.81	2.26
Civil engineering	2.72	3.09	2.81	3.09
Transport and communication	4.80	4.96	5.94	4.80
Trade and public catering	6.78	7.26	8.75	7.00
Housing and communal services	2.89	2.78	3.91	2.54
Health, social security and sports	4.88	4.54	4.06	4.53
Education, culture, science, arts	6.39	6.42	4.53	7.07
Finance, social insurance, real estate	1.94	1.94	2.03	1.78
State administration, public organizations	3.24	3.45	2.50	3.84
Oblast:				
Dnipropetrovska	23.33	24.52	24.53	24.71
Crimea	4.02	4.30	5.31	3.91
Odesska	10.06	10.41	10.94	10.57
Lvivska	21.12	22.52	20.63	23.20
Kyivska	16.63	16.00	12.34	16.54
Donetska	24.84	22.28	26.25	21.07

Appendix C

Table C1

Proportional income tax rate for the regular sector job

Source: Decree of President of Ukraine "On increasing non-taxable minimum and proportional income tax rates of citizens" (with changes and additions) N 13-92 from December 26, 1992[13]

Aggregate monthly income (AMI)	Income tax rate and due tax amounts
1 minimum (equal to 17 hryvna)	0
10 minimum + 1hryvna – 20 minimum	10%* (AMI – 1 minimum)
20 minimum + 1 hryvna – 30 minimum	10% *9 minimum + + 20%*(AMI-10 minimum)
30 minimum + 1 hryvna – 50 minimum	10%*9 minimum + 20%*10 minimum + + 35%*(AMI – 20 minimum)
50 minimum + 1 hryvna – and more	10%*9 minimum + 20%*10 minimum + + 35%*10 minimum + + 50%* (AMI – 30 minimum)

[13] The scale changed in 2001, which is not relevant to the period of the current research.

Appendix D

Table D1

Result of logit estimation with the dummy dependent variable based on unofficial hours

Dependent variable: dummy for unregistered hours (unreg_hr =1,0)

Factors	Coefficient	Odds ratio	s.e.	z	p>z
Hour_leisure	-0.0230738**	0.9771904	0.0011374	-20.29	0.000
Hour_registered	-0.0379222**	0.9627879	0.0051126	-7.42	0.000
Hour_r_registered	0.0148016**	1.014912	0.004943	2.99	0.003
Hour_housework	-0.0261387**	0.9742	0.0016753	-15.60	0.000
Total earnings	0.0008528**	1.000853	0.0002103	4.05	0.000
Registered income	-0.0013768**	0.9986241	0.0004278	-3.22	0.001
Dnipropetrovska	-0.6258045**	0.534831	0.1656429	-3.78	0.000
Kyivska	-0.845336**	0.4294131	0.1946441	-4.34	0.000
W_trade&catering	0.7214868**	2.05749	0.2279637	3.16	0.002
Excessive taxes	2.915855**	18.46459	0.5388838	5.41	0.000
Urban	-0.3823666**	0.6822449	0.1374384	-2.78	0.005
Not related	2.097206**	8.143382	0.2870462	7.31	0.000
Same position	1.584453**	4.876623	0.2567608	6.17	0.000
Secondary vocational	0.4807258**	1.617248	0.162379	2.96	0.003
Higher education	0.6103662**	1.841105	0.1802713	3.39	0.001
Number of adults	-0.2158014**	0.8058953	0.0654495	-3.30	0.001
Age	-0.0075084*	0.9925197	0.0041062	-1.83	0.067
Constant	10.7449**	na	0.6575749	16.34	0.000
LR chi^2 (17)		1050.612			
Probability>chi^2		0.0000			
LR index		0.398383			

* - significance at 10% critical level, ** - significance at 1% critical level

Explanatory variables:

Hour_leisure – hours devoted to leisure per month;

Hour_registered – hours devoted to registered economic activities per month;

Hour_r_registered – hours devoted to regular registered economic activities per month;

Hour_house work – hours devoted to housework per month;

Total earnings – total personal income per month;

Registered income – total personal registered income per month;

Dnipropetrovska – dummy for Dnipropetrovska oblast residence;

Kyivska – dummy for Kyivska oblast residence;

W_trade&catering – official average hourly wage rate in trade and public catering sector;

Excessive taxes – dummy for self-accession of excessive taxes;

Urban – dummy for urban residence;

Not related – dummy for no relation between registered and unofficial economic activities;

Same position – dummy for same position at unofficial and registered occupation;

Secondary vocational – dummy for secondary vocational education;

Higher education – dummy for higher and incomplete higher education;

Number of adults – total number of adults in a family;

Age – age of the respondent.

Table D2

Result of logit estimation with the dummy dependent variable based on unregistered earnings

Dependent variable: dummy for unregistered earnings (unreg_y =1,0)

Factors	Coefficient	Odds ratio	s.e.	z	p>z
Hour_leisure	0.0318887**	1.032403	0.0077381	4.12	0.000
Hour_unregistered	0.0317442**	1.032253	0.0076063	4.17	0.000
Hour_registered	0.0322627**	1.032789	0.007738	4.17	0.000
Hour_housework	0.030628**	1.031102	0.0077531	3.95	0.000
Hour_land plot	0.0314823**	1.031983	0.0077386	4.07	0.000
Total earnings	0.0082839**	1.008318	0.0006306	13.14	0.000
Non-labour income	-0.0077505**	0.9922794	0.0009061	-8.55	0.000
Registered_r income	-0.0082956**	0.9917387	0.0006881	-12.06	0.000
Land plot	-0.2307416**	0.7939446	0.1033895	-2.23	0.026
Proceeds_land plot	-0.0003455**	0.9996546	0.0001024	-3.37	0.001
Bribe	-0.000046**	0.999954	0.000023	-2.00	0.045
Wage arrears	0.2922394**	1.339424	0.1474699	1.98	0.048
Joint stock	-0.4819214**	0.6175956	0.2304361	-2.09	0.036
Agriculture&forestry	-0.7077404**	0.4927564	0.3157465	-2.24	0.025
W_housing&communal	-0.4486464*	0.6384918	0.2340111	-1.92	0.055
Crimea	-0.725062**	0.4842945	0.2884583	-2.51	0.012
Donetska	-0.4778805**	0.6200963	0.1105142	-4.32	0.000
Constant	-16.95326**	na	4.174805	-4.06	0.000
LR chi^2 (17)		474.9569404			
Probability>chi^2		0.0000			
LR index		0.158926			

* - significance at 10% critical level, ** - significance at 5% critical level

Explanatory variables:

Hour_leisure – hours devoted to leisure per month;

Hour_unregistered – hours devoted to unregistered economic activities per month;

Hour_registered – hours devoted to registered economic activities per month;

Hour_house work – hours devoted to housework per month;

Hour_land plot – hours devoted to land plot activities per month;

Total earnings – total personal income per month;

Non-labour income – total individual non-labour income per month;

Registered_r income – total personal registered income at regular job per month;

Land plot – dummy for possession of personal land plot;

Proceeds_land plot – dummy for generating proceeds from selling own crops;

Bribe – dummy for paying bribes to different officials;

Wage arrears – dummy for occurrence of wage arrears at regular registered work;

Joint stock – joint stock ownership of regular place of work;

Agriculture&forestry – dummy for regular job at agriculture and forestry sector;

W_housing&communal – official average wage rate at housing and communal services sector;

Crimea – dummy for Crimea residence;

Donetska – dummy for Donetska oblast residence.

Appendix E

Table E1

Constrained OLS regression on expected unregistered earnings, Regime 11

Dependent variable: log of expected unregistered earnings

Regressors	Coefficients	OLS s.e.*	P-value
$\ln W_0$	0.0479958	0.0417975 (0.037996)	0.251
$\ln(1-\tau)$	0.0479958	0.0417975 (0.0467115)	0.251
$\ln(1-p\theta)$	0.9520042	0.0417975 (0.0381199)	0.000
Marital status	-0.6379125	0.1315167	0.000
Total number of adults in the household	-0.2348955	0.0560108	0.000
Non-labor income for wives	0.0016123	0.0001366	0.000
Age	0.0150328	0.0047266	0.002
Constant	2.87718	0.2472172	0.000

* Bootstrapping standard errors for the parameters of interest are indicated in the parenthesis.

Regime 11:

- Penalty rate: $\theta = 1.25 * \tau$, where τ - marginal proportional income tax rate in the regular sector;
- Probability of being caught in tax evasion: p=probability set (1).

Probability sets were generated upon the qualitative answers of the respondents to the question of the proportion of income declared to the tax authorities:

Generated probability sets

Answers:	Probability set (1)	Probability set (2)	Probability set (3)
All income	*1*	*1*	*1*
Larger part	*0.75*	*0.88*	*0.63*
Smaller part	*0.25*	*0.37*	*0.12*
Nothing	*0*	*0*	*0*
Undecided	*0.5*	*0.5*	*0.5*
Mean of probability set:	*0.3194631*	*0.3330268*	*0.3061812*

Table E2

Constrained OLS regression on expected unregistered earnings, Regime 12

Dependent variable: log of expected unregistered earnings

Regressors	Coefficients	OLS s.e.*	P-value
$\ln W_0$	0.0494608	0.041757 (0.0421737)	0.236
$\ln(1-\tau)$	0.0494608	0.041757 (0.0422318)	0.236
$\ln(1-p\theta)$	0.9505392	0.041757 (0.0488653)	0.000
Marital status	-0.6375496	0.1315371	0.000
Total number of adults in the household	-0.2346351	0.056018	0.000
Non-labor income for wives	0.0016122	0.0001366	0.000
Age	0.0150899	0.0047273	0.001
Constant	2.876517	0.2472527	0.000

* Bootstrapping standard errors for the parameters of interest are indicated in the parenthesis.

Regime 12:

- Penalty rate: $\theta = 1.25 * \tau$, where τ - marginal proportional income tax rate in the regular sector;
- Probability of being caught in tax evasion: p=probability set (2).

Table E3

Constrained OLS regression on expected unregistered earnings, Regime 13

Dependent variable: log of expected unregistered earnings

Regressors	Coefficients	OLS s.e.*	P-value
$\ln W_0$	0.0467533	0.0418328 (0.0376248)	0.264
$\ln(1-\tau)$	0.0467533	0.0418328 (0.0434794)	0.264
$\ln(1-p\theta)$	0.9532467	0.0418328 (0.0393315)	0.000
Marital status	-0.6382791	0.1315	0.000
Total number of adults in the household	-0.2350952	0.0560048	0.000
Non-labor income for wives	0.0016124	0.0001365	0.000
Age	0.014981	0.0047259	0.002
Constant	2.877627	0.2471883	0.000

* Bootstrapping standard errors for the parameters of interest are indicated in the parenthesis.

Regime 13:

- Penalty rate: $\theta = 1.25 * \tau$, where τ - marginal proportional income tax rate in the regular sector;
- Probability of being caught in tax evasion: p=probability set (3).

Table E4

Constrained OLS regression on expected unregistered earnings, Regime 21

Dependent variable: log of expected unregistered earnings

Regressors	Coefficients	OLS s.e.*	P-value
$\ln W_0$	0.0558689	0.0414916 (0.0407312)	0.178
$\ln(1-\tau)$	0.0558689	0.0414916 (0.0401369)	0.178
$\ln(1-p\theta)$	0.9441311	0.0414916 (0.0407263)	0.000
Marital status	-0.637671	0.1315501	0.000
Total number of adults in the household	-0.234838	0.0560249	0.000
Non-labor income for wives	0.0016141	0.0001366	0.000
Age	0.0151718	0.0047279	0.001
Constant	2.884335	0.2472915	0.000

* Bootstrapping standard errors for the parameters of interest are indicated in the parenthesis.

Regime 21:

- Penalty rate: $\theta = 1.5 * \tau$, where τ - marginal proportional income tax rate in the regular sector;
- Probability of being caught in tax evasion: p=probability set (1).

Table E5

Constrained OLS regression on expected unregistered earnings, Regime 22

Dependent variable: log of expected unregistered earnings

Regressors	Coefficients	OLS s.e.*	P-value
$\ln W_0$	0.0578527	0.0414376 (0.0354477)	0.163
$\ln(1-\tau)$	0.0578527	0.0414376 (0.0454996)	0.163
$\ln(1-p\theta)$	0.9421473	0.0414376 (0.0401857)	0.000
Marital status	-0.6372638	0.1315785	0.000
Total number of adults in the household	-0.2344553	0.056035	0.000
Non-labor income for wives	0.001614	0.0001366	0.000
Age	0.015243	0.004729	0.001
Constant	2.883355	0.2473412	0.000

* Bootstrapping standard errors for the parameters of interest are indicated in the parenthesis.

Regime 22:

- Penalty rate: $\theta = 1.5 * \tau$, where τ - marginal proportional income tax rate in the regular sector;
- Probability of being caught in tax evasion: p=probability set (2).

Table E6

Constrained OLS regression on expected unregistered earnings, Regime 23

Dependent variable: log of expected unregistered earnings

Regressors	Coefficients	OLS s.e.*	P-value
$\ln W_0$	0.0542571	0.0415369 (0.0388207)	0.192
$\ln(1-\tau)$	0.0542571	0.0415369 (0.0377972)	0.192
$\ln(1-p\theta)$	0.9457429	0.0415369 (0.0389864)	0.000
Marital status	-0.6380983	0.1315281	0.000
Total number of adults in the household	-0.2351193	0.0560171	0.000
Non-labor income for wives	0.0016141	0.0001366	0.000
Age	0.0151079	0.0047271	0.001
Constant	2.884992	0.2472532	0.000

* Bootstrapping standard errors for the parameters of interest are indicated in the parenthesis.

Regime 23:

- Penalty rate: $\theta = 1.5 * \tau$, where τ - marginal proportional income tax rate in the regular sector;
- Probability of being caught in tax evasion: p=probability set (3).

Table E7

Constrained OLS regression on expected unregistered earnings, Regime 31

Dependent variable: log of expected unregistered earnings

Regressors	Coefficients	OLS s.e.*	P-value
$\ln W_0$	0.065936	0.0411155 (0.048178)	0.109
$\ln(1-\tau)$	0.065936	0.0411155 (0.0388566)	0.109
$\ln(1-p\theta)$	0.934064	0.0411155 (0.0361488)	0.000
Marital status	-0.6382078	0.1316042	0.000
Total number of adults in the household	-0.2345075	0.0560475	0.000
Non-labor income for wives	0.001616	0.0001366	0.000
Age	0.015311	0.0047301	0.001
Constant	2.892192	0.2474097	0.000

* Bootstrapping standard errors for the parameters of interest are indicated in the parenthesis.

Regime 31:

- Penalty rate: $\theta = 1.75 * \tau$, where τ - marginal proportional income tax rate in the regular sector;
- Probability of being caught in tax evasion: p=probability set (1).

Table E8

Constrained OLS regression on expected unregistered earnings, Regime 32

Dependent variable: log of expected unregistered earnings

Regressors	Coefficients	OLS s.e.*	P-value
$\ln W_0$	0.0685797	0.0410447 (0.0334886)	0.095
$\ln(1-\tau)$	0.0685797	0.0410447 (0.0369908)	0.095
$\ln(1-p\theta)$	0.9314203	0.0410447 (0.0382662)	0.000
Marital status	-0.637789	0.1316432	0.000
Total number of adults in the household	-0.2339545	0.0560614	0.000
Non-labor income for wives	0.001616	0.0001367	0.000
Age	0.0153976	0.0047315	0.001
Constant	2.890774	0.2474783	0.000

* Bootstrapping standard errors for the parameters of interest are indicated in the parenthesis.

Regime 32:

- Penalty rate: $\theta = 1.75 * \tau$, where τ - marginal proportional income tax rate in the regular sector;
- Probability of being caught in tax evasion: p=probability set (2).

Table E9

Constrained OLS regression on expected unregistered earnings, Regime 33

Dependent variable: log of expected unregistered earnings

Regressors	Coefficients	OLS s.e.*	P-value
$\ln W_0$	0.0638958	0.0411725 (0.0418647)	0.121
$\ln(1-\tau)$	0.0638958	0.0411725 (0.0475847)	0.121
$\ln(1-p\theta)$	0.9361042	0.0411725 (0.0428361)	0.000
Marital status	-0.6386836	0.1315758	0.000
Total number of adults in the household	-0.2348936	0.0560375	0.000
Non-labor income for wives	0.0016161	0.0001366	0.000
Age	0.0152343	0.004729	0.001
Constant	2.893125	0.2473602	0.000

* Bootstrapping standard errors for the parameters of interest are indicated in the parenthesis.

Regime 33:

- Penalty rate: $\theta = 1.75 * \tau$, where τ - marginal proportional income tax rate in the regular sector;
- Probability of being caught in tax evasion: p=probability set (3).

Table E10

Constrained OLS regression on expected unregistered earnings, Regime 41

Dependent variable: log of expected unregistered earnings

Regressors	Coefficients	OLS s.e.*	P-value
$\ln W_0$	0.0798746	0.0406177 (0.0421329)	0.049
$\ln(1-\tau)$	0.0798746	0.0406177 (0.0448259)	0.049
$\ln(1-p\theta)$	0.9201254	0.0406177 (0.0416811)	0.000
Marital status	-0.6402985	0.1316969	0.000
Total number of adults in the household	-0.233643	0.0560858	0.000
Non-labor income for wives	0.0016182	0.0001367	0.000
Age	0.0154515	0.0047337	0.001
Constant	2.900985	0.2476092	0.000

* Bootstrapping standard errors for the parameters of interest are indicated in the parenthesis.

Regime 41:

- Penalty rate: $\theta = 2 * \tau$, where τ - marginal proportional income tax rate in the regular sector;
- Probability of being caught in tax evasion: p=probability set (1).

Table E11

Constrained OLS regression on expected unregistered earnings, Regime 42

Dependent variable: log of expected unregistered earnings

Regressors	Coefficients	OLS s.e.*	P-value
$\ln W_0$	0.0833813	0.0405252 (0.0443571)	0.040
$\ln(1-\tau)$	0.0833813	0.0405252 (0.04359)	0.040
$\ln(1-p\theta)$	0.9166187	0.0405252 (0.038367)	0.000
Marital status	-0.6399327	0.1317501	0.000
Total number of adults in the household	-0.2328471	0.0561048	0.000
Non-labor income for wives	0.0016181	0.0001368	0.000
Age	0.015555	0.0047358	0.001
Constant	2.89896	0.2477037	0.000

* Bootstrapping standard errors for the parameters of interest are indicated in the parenthesis.

Regime 42:

- Penalty rate: $\theta = 2 * \tau$, where τ - marginal proportional income tax rate in the regular sector;
- Probability of being caught in tax evasion: p=probability set (2).

Table E12

Constrained OLS regression on expected unregistered earnings, Regime 43

Dependent variable: log of expected unregistered earnings

Regressors	Coefficients	OLS s.e.*	P-value
$\ln W_0$	0.0773422	0.040688 (0.0395423)	0.058
$\ln(1-\tau)$	0.0773422	0.040688 (0.0412242)	0.058
$\ln(1-p\theta)$	0.9226578	0.040688 (0.0368036)	0.000
Marital status	-0.6408039	0.131661	0.000
Total number of adults in the household	-0.2341632	0.0560733	0.000
Non-labor income for wives	0.0016183	0.0001367	0.000
Age	0.0153612	0.0047324	0.001
Constant	2.902279	0.2475466	0.000

* Bootstrapping standard errors for the parameters of interest are indicated in the parenthesis.

Regime 43:

- Penalty rate: $\theta = 2 * \tau$, where τ - marginal proportional income tax rate in the regular sector;
- Probability of being caught in tax evasion: p=probability set (3).