



Lina Salanauskaitė
Maastricht Graduate School of Governance, the Netherlands

Viginta Ivaškaitė-Tamošiūnė
Vilnius University/ Institute for Social Research, Lithuania

Microsimulation of Lithuanian tax - benefit policies: what can EU-SILC and HBS data tell us?

A Work in progress – do not quote

Comments welcome at:
Lina.Salanauskaite@governance.unimaas.nl
Viginta@ktl.mii.lt

2008

CONTENTS

ABSTRACT	3
1. INTRODUCTION	4
2. REVIEW OF MICROSIMULATION AND MICRODATA RESOURCES IN LITHUANIA	4
3. HBS AND EU-SILC: REVIEW OF THE MAIN ATTRIBUTES	6
3.1. INTRODUCTION TO HBS FOR 2004 AND 2005	6
3.2. INTRODUCTION TO EU-SILC	7
3.3. COMPARATIVE ANALYSIS OF HBS AND EU-SILC DATABASES	8
3.3.1. <i>Concepts and design issues</i>	8
3.3.2. <i>Main demographic variables</i>	9
3.3.3. <i>Main income variables</i>	10
4. EU-SILC AND HBS CAPACITY TO SIMULATE LITHUANIAN TAX-BENEFIT SYSTEM.....	14
4.1. METHODOLOGY	14
4.2. SIMULATION CAPACITY OF DIRECT TAXES AND CONTRIBUTIONS	15
4.2.1. <i>Personal Income Tax</i>	15
4.2.2. <i>Social insurance contributions</i>	17
4.2.3. <i>Compulsory health insurance contributions</i>	18
4.3. SIMULATION CAPACITY OF NON-CONTRIBUTORY BENEFITS TO FAMILIES AND CHILDREN	19
4.3.1. <i>Family benefit</i>	19
4.3.2. <i>Benefit for families raising 3 and more children</i>	21
4.3.3. <i>Birth grant</i>	23
CONCLUSIONS.....	25
REFERENCES	26

Abstract

Microsimulation models for analysing Lithuanian tax-benefits system are not yet employed despite the fact that it is a valuable tool for the social policy analysis. Outcomes of ongoing social reforms are usually assessed for an „average“ or „statistical“ individual or family. Nevertheless, policy impacts on the population groups, coming from different economic or socio-cultural backgrounds are uncertain. If microsimulation models were employed to evaluate the consequences of proposed or actual changes in social policies, a more detailed picture of the potential winners and losers could be acquired

The purpose of this research is to compare the potential of national microdatabases for the construction of the static Lithuanian tax-benefit microsimulation model. We, therefore, review Lithuanian tax-benefit policies on a selected date (i.e. June 30, 2004), which is already used for the purpose of building EU-wide static tax benefit microsimulation model EUROMOD. Afterwards, we focus on comparison of two datasets containing rich information on the socio-economic and demographical situation in the country: the Household Budget Survey (HBS) for the years 2004 and 2005; and the European Community Statistics on Income and Living Conditions (EU-SILC) in Lithuania for the year 2005. This study particularly aims to assess both surveys in terms of their potential to simulate selected tax, social contribution and family policies.

The comparative analysis of both HBS and EU-SILC datasets shows that, despite embedded differences and shortages of the analysed data, these surveys contain valuable information for the purpose of simulation of the Lithuanian tax-benefit system. HBS offers higher possibilities of simulating Lithuanian benefit system, as it contains more detailed income categories (i.e. recipients of maternity/paternity insurance, which is not available in the EU-SILC) and provides more information on the non-cash household income. Thus, it allows evaluating a wider range of existing or proposed public policy measures. On the other hand, the use of EU-SILC data for Lithuania provides a higher comparative value in terms of tax and benefit simulation across the EU. In addition to this, our research suggests, that currently both datasets tend to have a higher potential in simulating Lithuanian benefits (especially, benefits for families and children) rather than direct taxes or social contributions (especially, personal income taxation). This situation could change in the future, when the fully implemented EU-SILC survey containing variables on employer's social contributions or income components in gross and net forms for the year of 2007 is going to be released.

1. Introduction

In 2006, a feasibility study on developing a Lithuanian tax-benefit microsimulation model has identified Household budget survey (HBS) as the most comprehensive micro-database for the purpose of building a microsimulation model (Stirling & Lazutka, 2006). Since August, 2007, the EU-SILC (Community Statistics on Income and Living Conditions) dataset for Lithuania has been released. This dataset collects information on a number of HBS covered statistics: population demographics, employment, income or housing conditions.

A study, prepared by Figari, Levy and Sutherland, has already suggested that the EU-SILC has a number of significant benefits in order to be used as the underlying database for the cross-country comparative studies using EUROMOD (Figari et al., 2007a). On the other hand, it has been also pointed out that national modules of EUROMOD tend to have a lower capacity to address the complexities of the tax benefit-systems if compared to the national microsimulation models (Figari et al., 2007b). A preliminary analysis of the EU-SILC and HBS questionnaires for Lithuania points to HBS, as a more comprehensive dataset for a construction of a national tax-benefit microsimulation mode. However, currently there is no research on the comparison of these two datasets as the potential sources for Lithuanian tax-benefit simulation.

The main purpose of this study is, thus, set to make a comparative study of these 2 datasets with respect to their data capacity in simulating Lithuanian tax-benefit policies. Particularly, this paper aims to identify which policies could be addressed by both, at least by one or by neither of the analysed datasets. In addition to this, this study aims to simulate selected tax-benefit policies using both EU-SILC and HBS microdata and to compare obtained simulation results with the administrative records.

Firstly, this paper presents a general description of microsimulation modelling and microdata resources in Lithuania. Secondly, it compares HBS and EU-SILC datasets on a number of issues: embedded concepts, data collection modes, demographic or income characteristics. Thirdly, the paper reviews Lithuania tax-benefit policies and examines if their rules could be simulated by the use of HBS and EU-SILC data. Finally, the conclusions on HBS and EU-SILC potential to simulate Lithuanian tax-benefit structure are made and recommendations for the further research with respect to identified data or simulation differences are suggested.

2. Review of microsimulation and microdata resources in Lithuania

Microsimulation modelling, as a tool for public policy analysis, is not yet actively employed in Lithuania. Davies indicates that in the last decade the unpredictable nature and rapid pace of expansion of the transition economies, data shortage and a lack of modelling resources were among the main reasons hindering the development of the microsimulation modelling in Lithuanian and other transition countries (2004). However, as the country's socio-economic environment, as well as micro databases, become much more predictable and reliable, microsimulation modelling receives an increasing attention among the policy analysts. Tax-benefit microsimulation models, as a kind of microsimulation modelling, play an especially significant role for the purpose of policy effect evaluation. These types of models are actively used in most of the EU countries.

Lithuania, however, has not yet developed a national tax-benefit microsimulation model (MSM) yet. Nevertheless, Lelkes reports (2007) that the country is currently in the process of constructing one. In addition to this, a feasibility study on the development of a Lithuanian tax-benefit MSM has been prepared in 2006 (Stirling & Lazutka, 2006). It was initiated by the I-CUE project - the project aimed at improving the capacity and usability of EUROMOD – the EU-wide tax-benefit static microsimulation model. This feasibility study reviews Lithuanian tax-benefit policies on June 30, 2005, and recommends the construction of a tax-benefit micro-simulation model based on a Household Budget Survey (HBS). In fact, HBS has also been identified as one of most feasible micro-databases for the purpose of tax-benefit microsimulation in the other new member states of the EU¹. The proto-types of EUROMOD for Estonia and Slovenia, for example, are (being) built using HBS.

This paper takes a further step in analysing the EU-SILC (Community Statistics on Income and Living Conditions) dataset as an alternative underlying micro database for the construction of a Lithuanian tax-benefit microsimulation model. The EU-SILC collects comparable cross sectional and longitudinal multidimensional micro data on income poverty and social exclusion in European countries (EUROSTAT, 2007a). Since 2005, the dataset covers 25 EU member states, plus Norway and Iceland, and is the largest comparative survey of European income and living conditions. The EU-SILC 2005 database, containing information on Lithuania, has been released in August, 2007.

The EU-SILC has already been assessed with respect to its potential on carrying out public policy simulations across the EU. The study, conducted by Figari, Levy and Sutherland, tests the usability of this database by linking it to the Spanish EUROMOD module (2007). The researchers suggest that despite limitations, associated with the variables contained in the EU-SILC, the database has a number of significant benefits: it could further increase model's comparability and reduce deficiencies, related to maintenance issues, or improve access to the EUROMOD utilized national data.

In addition to EU-SILC and HBS databases, other population representative microdatabases, such as *the Population and housing census*, *Population migration survey* or *Population employment survey*, are collected in Lithuania. However, these microdatabases aim at collecting in depth information on specific issues rather than providing a broad picture of the population income and living conditions. Besides, current Lithuanian population surveys could not be merged in order to form a more inclusive database. Such a practice of pooling relevant microdatabases is quite widespread for the construction of tax-benefit MSM. For example, the U.S. population based *Dynamic Simulation of Income Model* (DYNASIM), was developed on the host database, which matched Current Population Survey (CPS) with the social security earnings records of the individuals and federal income tax information (Anderson, 2003).

There is currently a limited potential in Lithuania of using pooled survey databases or administrative records. Therefore, this paper focuses on HBS and EU-SILC, as the most comprehensive surveys for the analysis of Lithuanian population incomes and living conditions.

¹ After the EU enlargement in 2004

3. HBS and EU-SILC: review of the main attributes

3.1. Introduction to HBS for 2004 and 2005

The construction of microsimulation model requires a wide range of characteristics of the underlying micro-units. Since 1996 Department of Statistics to the Government of the Republic of Lithuania (*Statistics Lithuania*) carries out the annual Household Budget Survey (HBS). “The objective of the survey is to obtain reliable information on the standards of living of Lithuanian households, i.e. incomes, expenditures, housing conditions as well as the acquisition of durable goods, etc.” (Statistics Lithuania, 2005:5).

Responding to its main objective, the survey collects socio-demographic information on individuals’ age and marital status, household size and its composition, information on employment status, individual and household income and its components (i.e. income in the form of wages from main or supplementary job, various types of pensions, social benefits, income from rent of property, non-cash income from agricultural production or charity), taxes and fees paid, expenditures and other types of population information. Generally, it can be stated that HBS collects data on:

- Demographic characteristics of households;
- Income;
- Expenditures;
- Employment;
- Consumption;
- Housing conditions.

The target population of the HBS is private households in Lithuania. Individuals living in the institutional establishments (nursing homes for elderly people, imprisonment institutions, compulsory military service installations, etc.) are excluded from the survey (Statistics Lithuania, 2005). The HBS households are being surveyed for the period of one month. As it is a cross-sectional survey, different households replace the previous target population after one month.

In order to assure that all social strata have equal possibilities to be randomly selected for the survey, the Population Register is used to draw the HBS target population. For this purpose, the territory of Lithuania is divided into 31 distinct strata, which cover 5 largest cities, small and middle size towns and countryside areas. HBS uses two data collection methods: the interview conducted by an interviewer and self-registration filled in by household members. For the purpose of data collection, two types of documents are used:

1. Main Household Questionnaire (filled in by an interviewer). The aim of this questionnaire is to collect social and economic information on household members, their living conditions and income.
2. Diary of Household Expenditure. This diary aims at the collection of household expenditure on food and non-food products and services. During the period of 15 days the household records these types of expenditures. The remaining days of the month are used for the registry on expenditure of non-food products and services. Based on the survey design, half of the surveyed households record expenditures on food in the first half of the month, while the other half – in the second half of the same month.

In 2004, HBS selected 10866 target households. Out of them, 7961 households (or 22611 individuals) participated in the survey. Non response ration in 2004 was 26.7%: it reached 41.0% in 5 largest cities and was at 13.4% level in the rural areas (Statistics Lithuania, 20005). In 2005, the selected sample size remained the same, but fewer households

participated: 7586 households, corresponding to 20633 individuals. This indicates that the non-response ratio increased in 2005 and reached 30.2% (45.5% in 5 largest cities and 16.2% in the rural areas) (Statistics Lithuania, 2006).

Despite the fact that HBS collects mainly the data for the year of the survey, but its data on expenditure and income from agricultural production of the household refers to the previous 12 months (Statistics Lithuania, 2005).

3.2. Introduction to EU-SILC

The EU-SILC (Community Statistics on Income and Living Conditions) collects information on private households and their members residing in the territory of the member states at the time of data collection. Persons living in collective households and in institutions, such as care or imprisonment establishments, are excluded from the target population. The main aim of the survey is to gather comparable information on the incomes, poverty and social exclusion in the different European countries. For this purpose, the survey collects information on household income, housing conditions, living conditions, employment, health status, access to health care, financial problems, housing problems, possibility to meet certain needs, as well as demographic and socio-economic characteristics of the population.

The EU-SILC was launched as a pilot survey in 2004 (EUROSTAT, 2007a). Since 2005, it incorporates 25 EU members (all members after the enlargement in 2004), plus Norway and Iceland. The full implementation of the survey is foreseen in 2007, when each participating country has to fully adapt their standards on the provision of imputed rent, employer social contributions or income component at gross level. EUROSTAT foresees Romania, Bulgaria, Turkey and Croatia being incorporated in the survey in the coming years.

Lithuanian data for the EU-SILC was first collected in 2005, during the II and the III quarters of the year. It was gathered via face-to-face interviewing of all respondents aged 16 and over. In addition to this, data from interviews was also combined with data from administrative sources. Lithuanian data on income is collected for the last calendar year (i.e. 2004), when demographic and other survey contained information refers to the year of the survey (i.e. 2005).

Surveyed Lithuanian households were selected from the Population Register using a random sampling method (Statistics Lithuania, 2007). According to *Statistics Lithuania*, the stratified sample design with a simple random sample was used in strata. This implies that the entire Lithuanian territory was divided into 7 non-overlapping strata, out of which a one-stage simple random sample is selected. The 5 largest cities of Lithuania, towns and rural areas are divided into separate strata. The survey sample size of households in every stratum is proportional to its share of population aged 16 and over.

In 2005, the EU-SILC targeted 6450 households, out of which 4441 participated in the survey. The participating households come from large cities (1755 households), towns (1232 households) and countryside regions (1454 households). *Statistics Lithuania* informs that the non-response ratios in large cities and countryside areas were equal to 43% and 17% respectively (2007). The issue of the non-response ratio was dealt by calibration in order to match the initial sample design and to allow some exact demographic estimates: number of persons in the stratum, number of persons by gender and some age groups.

The EU-SILC gathers both cross-sectional and longitudinal information. The latter one pertains to individual-level changes over time and is typically observed for a four years period. The EU-SILC 2005 data on Lithuania contains cross-sectional information only.

3.3. Comparative analysis of HBS and EU-SILC databases

3.3.1. Concepts and design issues

The main research object of both databases is a household, which is defined as:

- (HBS) One person or a group of persons living together in one apartment (house) and have a common budget or common expenses (Statistics Lithuania, 2005).
- (EU-SILC) A person living alone or a group of people who live together in the same dwelling and share expenditure including the joint provision of the essentials of living (Statistics Lithuania, 2007:13).

Household definitions, even though phrased differently, indicate analogous interpretation of the household concepts. In general, description of HBS and EU-SILC designs reveals a ray of similarities between the databases on the grounds of their aims, interviewing focus on individuals above age 16, and types of information collected. Nevertheless, considerable differences are also embedded in the surveys. For example:

1. The age of individual is recorded on the interview day in the HBS, while in the EU-SILC it is recorded by the quarter of the birth month. In addition to this, the EU-SILC provides information on person's age at the end of the year.
2. HBS data for 2004 and 2005 is collected monthly and represents the situation of income and demographical characteristics of the survey year (except of income from agricultural production, which refers to the previous 12 months). The EU-SILC data for 2005 was collected during the II and the III quarters (May-July) of the year. This database provides information on household composition, housing conditions, health, and debts reflecting the year of 2005, but its data on income and taxes represents the situation of previous calendar year.
3. HBS collects monthly income information, and only income related to agricultural production of the household or some state transfers, such as child benefit, are recorded as an aggregate for the previous 12 months. The EU-SILC presents only yearly income. In addition to information collected via interview, data from Social Insurance Fund and State Tax Inspectorate is used in the EU-SILC database.
4. HBS provides information on net income (after taxes), while the EU-SILC records both net and gross (before taxes) income data. The EU-SILC database also indicates in which form (net or gross) each income type was collected. In addition to this, the EU-SILC collects information on employers' social insurance contributions and monthly gross wages of employees. Due to the pilot nature of the survey, those 2 variables were not recorded in EU-SILC 2005 files for Lithuania yet.
5. HBS records information on expenditures of the households (for food or more expensive purchases), while the EU-SILC does not contain such information. However, the latter survey provides valuable information on the individuals' health and access to health care services.
6. HBS is a cross sectional survey, while the EU-SILC collects both cross-sectional and longitudinal information. Nevertheless, the EU-SILC 2005 for Lithuania still contains cross-sectional information only.

Statistics Lithuania also confirms this study observed and additional differences. It points specifically to the differences over income categories. For example, perquisite,

irregular transfers from other households, compensations for medicine, cost-free board for the children in the schools, concession for child care in the pre-school institutions for the single parent families are not recorded in the EU-SILC database, while all these income components are included in the disposable income of the HBS households (*Statistikos departamentas*, 2007). This means that the EU-SILC is not fully suitable to capture certain governmental measures on poverty reduction, especially if they are targeted at certain population groups. Due to differences in collected income components, *Statistics Lithuania* warns over potential differences in poverty indices or other social exclusion measures if using EU-SILC and HBS data (2007).

The following sections on comparison of the main demographic and income variables present the magnitude of the observed differences and also report on potential ways of aligning the main household income categories.

3.3.2. Main demographic variables

For the purposes of comparison of demographic attributes, a number of households and individuals participating in the survey, share of women and children, average age, average household size, and other population estimates are calculated using the EU-SILC and HBS databases (see Table 1). We use HBS data for 2005 and the EU-SILC for 2005 in order to make adequate demographic comparisons (these data refer to the same calendar year of 2005).

Table 1. Comparison of (weighted) demographical indices using EU-SILC and HBS data

	<i>EU-SILC (LT)</i> ²	<i>HBS 2005</i> ³
Individuals, million	3,41	3,43
Households, million	1,35	1,46
Women	53,5%	53,3%
Average age	38,37 (21,93)	38,18 (22,12)
Children (under the age of 18)	21,5%	21,8%
Adults (18 to 65)	63,1%	63,1%
Elderly (65+)	15,4%	15,2%
Average household size	2,53 (1,36)	2,34 (1,33)
Average equivalised household size	1,68 (0,59)	1,60 (0,58)
Households in the densely populated areas	40,8%	n.a. ⁴

Source: own calculations (standard deviations reported in parentheses)

The Table 1 shows that the EU-SILC indicates a slightly smaller Lithuanian population both in terms of individuals and households. These differences are especially reflected when reviewing household sizes. The EU-SILC also reports a lower number of households for relatively similar total population compared to HBS data. Therefore, its average household size is larger than the household size embedded in HBS database. According to the EU-SILC database an average household consists of 2,53 members, while according to HBS it contains 2,34 members. When using the equivalence scale recommended by EUROSTAT⁵, the gap

² Here and further on, STATA 9 statistical software is used for all calculations with EU-SILC data

³ Here and further on, SPSS No.10.0 statistical software is used for all calculations with HBS data

⁴ Densely populated areas are not specifically identified in the HBS terminology. Nevertheless, we could observe that households, living in the 5 largest cities, represent 40,2% of the population. This corresponds quite closely to the observed share of households residing in the densely populated areas if using EU-SILC data

⁵ This scale assigns a value of 1 to the household head, of 0.5 to each additional adult member, aged 14 and above, and of 0.3 to each child under the age of 13 (EUROSTAT, 2007a).

between household sizes widens even further: it decreases to around 1.7 using the EU-SILC data and to 1.6 members using the HBS data.

The Table 1 also suggests that both surveys contain rather similar representative weights on a share of women and adult population. However, the EU-SILC indicates slightly higher average age. This corresponds to the observation of a somewhat larger proportion of the elderly population in EU-SILC data. On the contrary, children in the EU-SILC database constitute a smaller share of the population compared to the findings in the HBS data.

3.3.3. Main income variables

The main purpose of income comparison is to identify if there are significant differences in the composite structures of the seemingly analogous income categories. We, therefore, firstly analyse the codebook provided descriptions on incomes. If differences are identified, we aim at constructing a variable that could narrow the definitional gap as close as possible. We tend to use EU-SILC definitions as thresholds and afterwards to construct aligned HBS variables. Such a decision is made due to the fact, that most of the EU-SILC contained variables tend to be aggregated to a higher degree compared to HBS data. This means, that HBS provided income list could be separated to a number of a lower degree variables (for example, income on social benefits could be separated to incomes from child benefit, social assistance, scholarships, etc), while EU-SILC data does not permit such disaggregation.

We use the HBS for the year 2004 in order to have the same reference point on incomes, as it is used in EU-SILC 2005 dataset. Income data, analysed in the abovementioned way, is reported in the Table 2. The last column of the table presents remarks and observed differences, which are crucial for the comparative purposes. It should be also noted, that only those income categories, which are more relevant for the analysis of the Lithuanian tax-benefit system, are considered within the scope of this research.

As Table 2 suggests, the following income categories have the closest definitional equivalents, if compared across both datasets⁶:

- cash income from rent, wealth, business and/or individual activities;
- employee non-cash income;
- social benefits for families and children;
- old age social benefits,
- social benefits for survivors and orphans,
- disability benefits;
- education related benefits.

The largest definitional differences are captured by comparing descriptions of:

- disposable income;
- cash social benefits;
- employee cash income,
- unemployment benefits;
- sickness benefits.

The main reasons behind the observed differences are associated with the choice of inclusion or exclusion of non-cash household agricultural production (e.g. within disposable income) and conceptually different classification of social benefits. For example, sickness

⁶ Based on the comparison of variable descriptions in the data support files

benefit is treated as a composite part of wages in EU-SILC, while HBS reports it as one of the social benefit types.

Comparative analysis of average household income values reveals the largest gap in the incomes from rent and wealth. The values of these incomes is about 4 times higher in HBS than in the EU-SILC database, while the number of recipients is about 5 times higher in the latter. These findings indicate, that despite relatively close definitional appearance of income categories, the EU-SILC and the HBS might report significantly different income values. In general, differences in average household income values might occur due to:

- Different explanation of income categories when collecting the data;
- Different data collection time, which may cause an inaccurate recall of income values (i.e. HBS collects monthly income data, while the EU-SILC collects annual data);
- Different data collection time, which might lead to deviations due to inflation;
- Low income incidence (i.e. less than 2%), associated with representation in terms of survey calibrated weights;
- Observed differences in average household sizes in the EU-SILC and HBS data.

In addition to this, the presentation of average values does not indicate the spread of data about the mean (i.e. variance or standard deviation). Therefore, interpretation of observed differences in means serves as an initial identification of the trends for the purpose of the further simulation analysis rather than an ultimate goal of the research.

Consequently, Table 4 suggests that the EU-SILC households tend to receive social transfers to families, such as social benefits for families and children, unemployment benefits, social benefits for orphans and survivors, and education related benefits, more often. This interpretation is reinforced by comparison of the number of households, receiving disposable income and disposable income, excluding social benefit and pensions. Based on EU-SILC data, only 74% of all households report positive disposable income, which excludes social benefits and pensions, while 87% of all households report this type of income in the HBS. Both databases indicate that around 99% of all households have some kind of positive total disposable income. Therefore, a conclusion could be drawn that social transfers and pensions play a more important role in the households, represented by EU-SILC data than by HBS data. Nevertheless, the values of these benefits seem to be lower in EU-SILC compared to HBS data. Due to abovementioned reasons, this difference in values might reflect, for example, imprecision in recall of the incomes or differences in household sizes (via calculation of per capita household income), embedded in EU-SILC and HBS data.

Comparison of social transfers to the elderly people, such as old age and disability benefits, indicates a somewhat opposite trend to the above reported development with respect to social transfers to families. Old age and disability benefits are reported less often in EU-SILC households, but their average values tend to be higher compared to HBS data.

Table 2. Comparison of monthly household incomes in EU-SILC and HBS (weighted)

Income categories	EU-SILC			HBS			Remarks/differences due to description of variables
	Variable	Households %	Average value, Lt ⁷	Variable	Households %	Average value, Lt	
Gross income	HY010 ⁸	99,1	1493 (1414)	n.a.	n.a.	n.a.	EU-SILC variable is aggregated using various income categories from individual files.
Disposable cash income (net)	HY020	99,1	1228 (1063)	disp_cas + inc_22k	99,5	1186 (921)	HBS excludes non cash income from agricultural activities while EU-SILC includes it. HBS doesn't separate income received by individuals under age 16 m.
Disposable income (excluding social benefits and pensions)	HY023	74,0	1257 (1130)	disp_cas + inc_5c ⁹ + inc_22k - inc_61c + inc_52c	87,2	1028 (982)	EU-SILC includes non-cash income from agricultural activities, but exclude scholarships (attributes it to category of social benefits). Sickness benefit in EU-SILC is attributed to category of employee cash income (while in HBS to social benefits).
Cash income from rent (net)	HY040N	2,5	111 (362)	inc_3c	0,6	431 (767)	No substantial differences
Cash income from property (net)	HY090N	1,8	135 (270)	inc_4c	0,4	638 (686)	No substantial differences
Social benefits (net)	Composite variable: py090g+ py100g+ py110g+ py130g+	56,2	318 (209)	inc_5c - inc_52c + inc_61c	59,4	450 (308)	Sickness benefit is included in category of cash social benefits in HBS, while EU-SILC does not. When comparing, this sickness benefit was excluded. Education related benefits are included in category of social benefits in EU-SILC.

⁷ Average size is calculated by accumulating monthly household income value and dividing it by households which receive corresponding income (positive income). 3.45 Litass (Lt) =3.45 euro (EUR)

⁸ All income variables of H file (HY...) are aggregated from individual files containing various income categories. Nevertheless some of individual categories such as sickness benefit are not recorded in individual files.

⁹ Letter c at the end of the variable name implies that variable represents cash income (HBS contains information about non-cash income components also)

Income categories	EU-SILC			HBS			Remarks/differences due to description of variables
	Variable	Households %	Average value, Lt ⁷	Variable	Households %	Average value, Lt	
	py140g						
Social benefits for families and children (net)	HY50N	15,3	148 (177)	inc_55c	11,64	219 (327)	No substantial differences
Employee cash income (net – including sickness benefits but excluding severance pay)	PY010N ¹⁰	62,4	1296 (1042)	inc_1c+ inc_52c – P05_4 ¹¹	62,3	1118 (858)	EU-SILC sickness benefit is included in employee cash income. HBS includes severance pays in this category; in the EU-SILC it is attributed to category of unemployment benefits.
Non cash employee's income ¹² (net)	PY020G	2,3	137 (129)	inc_1k	5,3	114 (131)	No substantial differences
Income from business, individual activity (net)	PY050N	16,1	629 (1003)	inc_2c	16,4	526 (988)	No substantial differences
Unemployment benefits (net)	PY090N	3,3	106 (127)	inc_56c + P05_4	2,1	184 (269)	Severance pay is attributed to employee cash income in HBS; in EU-SILC it is included in unemployment benefits.
Old age social benefits (net)	PY100G	38,6	541 (278)	inc_51c	39,1	513 (250)	No substantial differences
Social benefits for survivors and orphans (net)	PY110G	3,1	175 (104)	inc_54c	2,5	257 (303)	No substantial differences
Sickness benefit (net)	n.a.	n.a.	n.a.	inc_52c	2,6	85 (244)	No separate category of sickness benefit in EU-SILC, it is attributed to category of employee cash income.
Disability benefit (net)	PY130G	10,8	337 (194)	inc_53c	11,5	298 (181)	No substantial differences
Education related benefits (net)	PY140G	6,2	105 (117)	inc_61c	4,1	113 (97)	No substantial differences

Source: own calculations (standard deviations reported in parentheses)

¹⁰ This and other variables in the Table, starting with the letter p are of individual file level. They are recalculated to household level.

¹¹ Variables starting with letter P are of individual file level, only recalculated to household level.

¹² Non-cash employee income, old age benefits, social benefits for survivors and orphans, disability benefits and education related benefits are recorded only in gross in EU-SILC. Since these benefits are not taxable (under the Lithuanian law), gross income correspond net income.

4. EU-SILC and HBS capacity to simulate Lithuanian tax-benefit system

4.1. Methodology

In this chapter, a cross-sectional review of selected Lithuanian taxes and benefits in 2004 is presented. A cross-sectional evaluation of policy rules implies that policies are assessed on a specific point in time. Such an approach is used while constructing static microsimulation tax-benefit models. For example, most of the feasibility studies on EUROMOD extension to the new EU member states evaluate national rules of taxes and benefits on June 30, 2005 (Stirling and Lazutka, 2006). Following EUROMOD construction experience, we choose the same date to assess the HBS and the EU-SILC potential in simulation of Lithuanian tax-benefit policies in 2004.

For the purpose of further analysis, we use the HBS for the year of 2004 and the EU-SILC for 2005, as both datasets refer to the year of 2004 with respect to their income categories. This paper uses STATA 9 and SPSS No.10.0 versions of statistical software to simulate selected taxes and benefits using EU-SILC and HBS data respectively.

For the purpose of evaluation of the EU-SILC and the HBS capacity to simulate the Lithuanian tax-benefit system, we firstly review selected taxes and benefits with respect to their rules on June 30, 2004. Due to the limited scope of this analysis, we focus on evaluation of direct taxes and social contributions, as well as non-contributory social benefits.

Within the system of direct taxation, we examine the rules of personal income tax. This tax is one of the main sources of revenues to the state budget. According to the data of State Tax Inspectorate, revenues from personal income tax reached 21,3% of the total state budget in 2004 (*Valstybinė mokesčių inspekcija*, 2004). More revenues were only collected from indirect taxes, such as VAT (24,8%) and excise duties (22,7%).

Social insurance and health insurance contributions are analysed as the most important part of the Lithuanian system of social contributions. Additively, social insurance and health insurance constituted 10% of all revenues of Social Insurance Fund in 2004 (*Valstybinio socialinio draudimo fondo valdyba*, 2005). The decision of analysing both contributions is also made due to the fact that people working under employment contracts do not pay health contributions directly: it is designed as integral part of employer's social contributions, paid on behalf of the employee.

In addition to taxes and contributions, this paper reviews three types of non-contributory family policies: a family benefit, a benefit to families raising three or more children and a birth grant. The total value of these benefits amounted to 114 million Lt in 2004. This constitutes around 40% of all state transfers to families (*Socialinės apsaugos ir darbo ministerija*, 2007). The additional motive of focusing on the evaluation of these specific policies is an observation, made from the previous chapter. Income comparison has indicated that the EU-SILC reports a higher incidence of social transfers to families compared to HBS data. Furthermore, we have also found that on average households are larger in EU-SILC data than in HBS data. These developments are of particular importance when addressing and simulating state transfers to families and children.

Once the rules of the abovementioned policies are reviewed, we match information requirements, identified from the policy description, with the data availability in the EU-SILC and HBS datasets. Such an exercise reveals which dataset contains more suitable information

for the purpose of policy simulation. Due to the limited scope of this analysis, we only aim at actually simulating the reviewed benefits to families.

Finally, obtained simulation results are analysed mainly on the basis of their cumulative values in terms of aggregate expenses and a total number of potential beneficiaries. We use a 100% take-up rate assumption, which implies that simulation results indicate entitlement rights to the benefits rather than the actual utilization of these rights¹³. However, for the purpose of evaluating each dataset's potential in simulating benefits, we compare simulation results to the actual administrative data. The administrative data is used as a threshold indicator, revealing the deviation sign and magnitude of a simulated benefit is using the EU-SILC and HBS data. For the purpose of this analysis, we tend to use the rule of a thumb that simulated benefits are rather precise if they deviate from the administrative records within the 10% interval.

4.2. Simulation capacity of direct taxes and contributions

4.2.1. Personal Income Tax

Under *the Law of Personal Income Tax of the Republic of Lithuania (LR gyventojų pajamų mokesčio įstatymas, 2002)*, there were two income tax rates for residents on June 30, 2004. The 15% rate applied to income from profit, income of sportsmen and performers, honorarium, income under authorship contract or individual activity, income from rent, sale of property, and pensions from voluntary second pillar insurance, retrieval from contributions to voluntary pension or life insurance. All other income was taxed with 33% rate. When calculating taxable income, non-taxable income, income from activity under trade certificate, deductions referring to income from individual activity, the price of property acquisition and expenses in addition to main non-taxable exemptions and non-taxable values are deducted.

In Lithuania, the main non-taxable exemption (NTE) is 290 Lt per month. Other non-taxable exemption levels could be applied for certain population groups:

- For people of disability group I - 430 Lt;
- For people of disability group II - 380 Lt;
- For parents raising three or more children under the age of 18 or older, if in formal education - 430 Lt.; a non-taxable exemption increases by 46 Lt for the fourth and every subsequent child;
- For single parents having children under age 18 and older if in formal education - 335 Lt; NTE increases by 53 Lt for every subsequent child;
- For employees of agricultural activity subjects, when the subject's income from actual agricultural production per year is not less than 50% of total revenue - 330 Lt;
- For employees of farmers who have registered their farm - 330 Lt;
- For parents having one or two children under age 18 or older if in formal education – NTE equals to 29 Lt per eligible child.

All state social assistance benefits, paid from state and municipal budgets or *Social Insurance Fund*, except of sickness, pregnancy and maternity/paternity benefits, are not subjects to personal income taxation. Sickness, pregnancy and maternity/paternity benefits are taxed with 33% rate. A number of other non-taxable income types exist: these include interest from loans, stock, charity, scholarships and some other types of incomes.

The Table 3 presents variables needed for determination of taxable and non-taxable income, as well as groups of people entitled to different NTE levels. It depicts if this

¹³ Some studies indicate that actually the take-up rate of social benefits in the Baltic States is close to almost 100% (Aidukaite, 2006).

information could be identified using the EU-SILC and HBS data, and if potentially used variables provide equivalent information in both datasets.

Table 3. Information needed for simulation of personal income tax

Information needed	EU-SILC variable	HBS variable	Main differences in variables
15%			
Income from business, individual activity	PY050G (gross)	inc_2c (net)	PY050G includes income from interest, which is not taxed according to the personal income taxation rules. However, this income could not be separated out from the aggregate variable. This type of adjustment could be made using HBS data.
Cash income from rent	HY040G (gross)	inc_3c (net)	Variable HY040G is reported on the household level only. HBS contains individual and household level data.
Cash income from property	HY090G (gross)	inc_4c (net)	Variable Y090G is reported on the household level only. HBS contains individual and household level data.
33%			
Employee cash income	PY010G (gross)	inc_1c (net)	Variable PY010G does not include severance pay, but it includes sickness benefit. Variable Inc_1c includes severance pay (this category is recorded on the individual level, and thus could be separated out, if needed).
Sickness benefit	-	inc_52c (net)	This benefit is an inseparable part of PY010G variable in EU-SILC. HBS records sickness benefit on the individual level files.
Pregnancy, maternity/paternity benefit	-	P05_17 (net)	This benefit is integrated into variable HY050G in EU-SILC and cannot be separated. Variable P05_17 is reported at the individual level.
Severance and termination payment	-	P05_4 (net)	This benefit is included into unemployment benefit in EU-SILC variable PY90G and cannot be separated. Variable P05_4 in HBS is reported at the individual level.
Non-taxable exemptions (NTE)			
430 Lt: disability level I; 380 Lt: disability level II;	-	-	Disability groups could be partially identified by the receipt of disability benefits (variable PY130G in EU-SILC, variable inc_53c in HBS database).
430 Lt: parents having 3 or more children under age 18 or older if in formal education	-	-	Variable could be constructed based on household composition and employment data in both databases.
335 Lt: for single parents having children under age 18 or older if in	-	-	Variable may be constructed based on household composition and employment data in both databases.

Information needed	EU-SILC variable	HBS variable	Main differences in variables
formal education			
330 Lt: for employees at agricultural activity subjects, when the subject's income from realized production per year is not less than 50% of total revenue.	-	-	Both datasets have no data on the structure of company revenues. However, employment in the agricultural sector may be indicated based on NACE code in EU-SILC database (variable PL110) and based on info of work in agricultural sector in HBS data (variable P02_1 variable) ¹⁴ .
330 Lt: for employees of farmers who have registered their farm.	-	-	No data on employment in registered farms. Analogy to abovementioned assumption could be made on work in agricultural sector.
290 Lt: for parents having 1 or 2 children under age 18 and older if in formal education.	-	-	Variable may be constructed based on household composition and employment data in both databases.
290 Lt: for every other person	-	-	For all residents who receive income under the employment contract.

Source: data collected by authors

As Table 3 indicates, the information obtainable from both the EU-SILC and the HBS is not yet sufficient for simulation purposes of the personal income tax. Particularly, further information is needed to have a more precise estimation of:

- Disability groups;
- Full time education schools, universities;
- Employees of agricultural activity subjects, when the subject's income from realized production per year is not less than 50% of all income;
- Farmers who registered their farm and have employees;
- Accurate groups of social benefits: pregnancy and /or parental benefits (information recorded in HBS, but not in EU-SILC);
- Income from activity under trade certificates;
- Deductions related to income from personal activity;
- The value of property acquisition and expenditures.

4.2.2. Social insurance contributions

Social insurance contributions to *Social Insurance Fund* (SoDra) are paid by employers (for employees), employees (for themselves) and people working under self-employment contracts. The contribution ratio of social insurance is equal to 31% of gross income for employers, while it equals to 3% of gross income for employees. Annual social insurance contributions could not be lower than State defined Minimum monthly salary (MMS), which on June 30, 2004, was equal to:

- Minimum hourly wage - 2,95 Lt or
- Minimum monthly wage – 500 Lt¹⁵.

Self-employed persons, with the exception of the persons engaged in individual activities under business certificates, are covered on a compulsory basis only by pension social insurance to receive the basic and supplementary parts of a pension. These persons are insured

¹⁴ The variable indicating persons working in agricultural sector (personal, public or private) may be used as well Variable p01_18 (categories 1,2,3) is used.

¹⁵ Lietuvos Vyriausybės 2004-03-24 nutarimas Nr. 316 „Dėl minimaliojo darbo užmokesčio didinimo“ (Žin., 2004, Nr. 46-1511)

on a compulsory basis to receive the supplementary part of the pension if the annual aggregate amount of their income (after taxes) is equal to or exceeds the amount of 12 MMS. In case annual income is less than 12 MMS, the contributed amount should be equal to 50% of the state defined pension base, which amounted to 172 Lt per month¹⁶ on June 30, 2004. A 15% contribution rate is applied if the annual income is equal to or more than 12 MMS¹⁷.

Table 4. Information needed for simulation of social insurance contributions

	EU-SILC	NŪBT	Main differences of variables
3% from gross income			
Cash employee income	PY010G	inc_1c (excluding severance pay), plus sickness benefit	Variable PY010G does not include severance pay, but it includes sickness benefit. Variable Inc_1c includes severance pay (this category is recorded on the individual level, and thus could be separated out, if needed).
86 LT (50% from pension base)			
Persons engaged in individual activities	PL040	P02_3	PL040 describes employments status (coded 1 and 2 if a person is self employed). P02_3 indicates self-employed people.
15% from net annual income			
Persons engaged in individual activity, if annual income are higher than 6000 LT	PL040	P02_3	PL040 describes employments status (coded 1 and 2 if a person is self employed). P02_3 indicates self-employed people. Annual income from individual activities could be estimated by variables PY050G and inc_2c in EU-SILC and HBS respectively.

Source: data collected by authors

As Table 4 indicates, the information obtainable from both the EU-SILC and the HBS is sufficient enough for simulation purposes of social insurance contribution paid on the compulsory basis. Due to the limited scope of this analysis, this contribution is not going to be explored in terms of comparing simulation results with the actual administrative data.

4.2.3. Compulsory health insurance contributions

On 30th of June, 2004, compulsory health insurance contributions were paid by enterprise owners, members of partnership organisations, farmers, land users, users of personal farms, residents engaged in individual activities and individuals who insure themselves independently (*LR Sveikatos draudimo įstatymas*, 2003). People under employment contracts did not have to pay contributions as they were already included in employer paid social insurance contributions on behalf of the employees.

Persons engaged in individual activities, including those receiving honorariums and/or income from authorship contracts, were insured on a compulsory basis. Their contribution rates were equal to 30% of calculated sum of personal income tax. Contribution ratio could not be smaller than 1/12 of annual contributions to state budget for the individuals insured by state resources (in 2004 annual contribution was amounted to 221,4 LT, monthly contribution - 18,45 LT).

¹⁶ Lietuvos Vyriausybės 2004-03-24 nutarimas Nr. 325 „Dėl valstybinės socialinio draudimo bazinės pensijos didinimo“ (Žin., 2004, Nr. 46-1520)

¹⁷ Lietuvos Respublikos Valstybino socialinio draudimo įstatymas (2004), Nr.I-1336.

Every month farmers and other users of personal farms pay health insurance contributions of different values:

- 1) Farmers and land users (having 3 or more hectare of landed property) pay 3.5% of minimal monthly salary for themselves and other adult family members working in farm;
- 2) Personal farm users (having up to 3 ha of landed property) pay 1.5% of minimal monthly salary for themselves and other adult family members working in a farm.

The Table 5 describes the variables in EU-SILC and HBS datasets that are needed for the purpose of simulation of health insurance contribution.

Table 5. Information needed for simulation of compulsory health insurance

	EU-SILC	NÜBT	Main differences of variables
30% of simulated income tax (15% ratio) ¹⁸			
Persons engaged in individual activity	PL040	P02_3	PL040 describes employment status (coded 1 and 2 if a person is self-employed). P02_3 indicates self-employed people.
3.5% and 1.5% of minimal monthly salary			
Farmers and land users	-	-	In EU-SILC, farmers and land users could not be identified. In HBS, farmers and land users may be partially identified by: <ul style="list-style-type: none"> ▪ Owned (or used) arable land area: variable p04_2.2 (plot_agr). ▪ Working in the agriculture sector: variable P02_1.

Source: data collected by authors

As Table 5 indicates, the information obtainable from both the EU-SILC and the HBS is not sufficient enough for simulation purposes of health insurance contribution paid on the compulsory basis.

4.3. Simulation capacity of non-contributory benefits to families and children

4.3.1. Family benefit

On June 30, 2004, a family benefit was defined as a monthly benefit for families raising a child up to age 3. The benefit level was equal to 0.75 of minimum subsistence level (MSL), which was set at 125 Lt level per person per month during the course of 2004 (*Lietuvos Respublikos*, 1994). The payment of a family benefit was terminated if a family started receiving maternity (paternity) social insurance. This benefit was transformed to a child benefit since July 1, 2004, due to the introduction of a *new Law on the Allowances to Children* (Seimas of the Republic of Lithuania, 2004). Therefore, in 2004, the benefit was actually paid only for the first half of the year.

Table 6: Information needed for simulation of family benefit

	EU-SILC	HBS	Main differences of variables
Age of child	RX020	P01_age	RX020 reports age at the end of income reference period, while P01_age reports age at the time of the interview.
Mother and (or) father receives maternity (paternity)	-	P05_17	In EU-SILC, transfers to families and children are provided as an aggregate

¹⁸ But not less than 1/12 of annual contributions to state budget for the individuals insured by state resources for referring year (June 30 d., 2004)

	EU-SILC	HBS	Main differences of variables
social insurance			variable. Hence, identification of maternity (paternity) social insurance receipt is not feasible.

Source: data collected by authors

As the Table 6 indicates, the EU-SILC dataset has a lower potential to fully simulate the benefit rules. Due to the limited information on the receipt of maternity/paternity social insurance, a decision to perform simulation, based only on demographic criteria, is made. This implies that the entitlement right to the benefit is given to all children under the age of 3 (using age variables as indicated in the Table 6). If children qualify for this benefit based on the imposed age criteria, an assumption is made that the benefit is assigned for the whole year. The fact that the benefit was transformed to a child benefit since July 1, 2004, is taken into account by adjusting the values of administrative data. Such an approach indicates the hypothetical situation, describing what could have happened if benefit policies did not change during the course of the year.

According to the *Ministry of Social Security and Labour* (2005), the state spending on this benefit was equal to 47.3 million until July 1, 2004, and was distributed to 87100 beneficiaries. We make an assumption that the state would have spent the same amount on the second half of the year, if the reform on allowances to children would have not taken place. In such a case, the state annual spending on the benefit would have been equal to around 95 million Lt.

The simulation obtained results on the average benefit per household, the number of eligible households, simulated state expenses and actually observed state expenses on the benefit are reported in the Table 7.

Table 7. Simulation results versus administrative data: family benefit

	EU-SILC	HBS	Administrative data ¹⁹
Average benefit per household	101 Lt	98 Lt	-
% of all households entitled to benefit for 1 child	5.7%	6.2%	-
Annual state expenses	101 million Lt	103 million Lt	95 million Lt
Number of beneficiaries ²⁰	83565	87901	87100

Source: own calculation

Comparison of administrative data and simulation obtained results, using the EU-SILC and HBS data, indicates a relatively small deviation both in terms of state spending and a number of beneficiaries. The EU-SILC reports a somewhat higher benefit level per household, which points to the influence of the larger household size. However it forecasts a lower number of beneficiaries compared to administrative data and HBS based findings.

Simulation of the benefit is also feasible if taking into account the receipt of the maternity/paternity social insurance based on HBS data. The results of this simulation are presented in the Table 8.

¹⁹ This column indicates annual state spending, based on the assumptions in the abovementioned text.

²⁰ Number of beneficiaries equals to number of households that are entitled to receive the benefit

Table 8. Simulation of a family benefit assuming full rules

	EU-SILC	HBS
Annual state expenses	-	79 million Lt
Number of beneficiaries	-	68194

Source: own calculation

As it is reported in the Table 8, the use of maternity/paternity variable, while simulating family benefit, leads to a significant drop in a number of beneficiaries. Firstly, this indicates the relatively high incidence of maternity/paternity recipients among the households having children under the age of three. Secondly, the simulated decrease in a number of potential beneficiaries occurs due to the data incapacity to identify if maternity/paternity insurance is paid for a child under three or for a child, who is older than three, and thus not eligible for the family benefit.

In general, a gap between simulation results and administrative data could also occur due to:

- The assumption on the 100% take-up ratio;
- Specific simulation assumptions, such as benefit distribution for the entire year;
- Different references to age based on EU-SILC and HBS data.
- Different demographic structures embedded in EU-SILC 2005 and HBS 2004 data.

4.3.2. Benefit for families raising 3 and more children

Based on the rules, effective on June 30, 2004, this benefit is a monthly benefit to families, raising three or more children up to the age of 16. The benefit to children, who are older than 16, was paid, if they studied full-time in the secondary and higher educational system. The payment of this benefit was terminated if a child was placed under guardianship or if a child was disabled and turned 18.

Following the benefit rules, it was equal to 1.0 MSL for the 3rd child in the family. The benefit for the 4th and each additional child was increased by 0.3 MSL²¹. If entitlement to the benefit was granted, the benefit was paid monthly until the end of the calendar year.

This benefit was a means tested benefit with a conditional eligibility based on per capita family income being lower than 3 state supported income (SSI) levels. Family income was calculated based on the records of the preceding 3 months. The benefit for the 4th and each additional child was paid without income testing. This benefit was also eliminated by the introduction of a *new Law on the Allowances to Children*. Therefore, in 2004, the benefit was actually paid only for the first half of the year.

According to the *Ministry of Social Security and Labour* (2007), the state spending on this benefit amounted to 44.2 million Lt until July 1, 2004. It was paid to 48000 beneficiaries, out of whom 70% (or 33700 beneficiaries) were families raising three children. If making an assumption that the state would have spent the same amount during the second half of the year, annual state expenses would have constituted around 88 million Lt. The same rule applies for the partition of benefit expenses to spending for families with three children and spending for families with more than three children (see Table 10).

²¹ Lietuvos Respublikos Valtyniu pasalpu seimoms, auginancioms vaikus, istatymo 1,2,3,4⁽¹⁾ straipsniu pakeitimo ir papildymo istatymas, 2001 m. sausio 25 d. NR. IX-168, Vilnius

The Table 9 indicates that both the EU-SILC and the HBS cannot determine precisely if a person is studying full-time, while this is one of the criteria for the establishment of eligibility rights. In addition to this, both datasets do not fully provide with the income list that is established by the law for the purpose of means testing. Nevertheless, the available information allows making a number of plausible assumptions on the benefit dispersion using both EU-SILC and HBS data.

Table 9. Information needed for simulation of benefit for families raising 3 or more children

	EU-SILC	HBS	Main differences of variables
Age of child	RX020	P01_age	RX020 reports age at the end of income reference period, while P01_age reports age at the time of the interview.
Persons studying full-time in the secondary and higher educational system	PE010	P01_study (categories 4, 5, 6, 7, 8, 9, 10, 11)	Both variables do not indicate whether a person studies full-time or not.
Marital status	PB190	P01_marital (category 1)	Both variables could provide with marital status (single), which is particularly needed to establish eligibility right of children above the age of 18.
Household income	-	-	Income list, as defined by the law, is not fully computable by both datasets. An assumption is made that for means testing could be performed using per capita household income.
Household size	HX040	hh_size	Both datasets provide info on household structure and size.

Source: data collected by authors

As Table 10 depicts, simulation results obtained with EU-SILC data forecast a substantially higher number of potential beneficiaries and, thus, higher annual state expenses compared to the results obtained with HBS data. The division of benefit distribution to families raising three children and families raising more than three children suggests that the large part of the difference between simulation obtained results is due to the embedded influences of household sizes in EU-SILC and HBS datasets. Based on the EU-SILC data, around half of predicted state expenses would be used for families having more than three children. According to HBS data, less than 25% of predicted state expenses would be used for this population group. This suggests that large families in the EU-SILC database carry a much higher representation weight than the same type of families in HBS data.

Table 10. Benefit simulation results

	EU-SILC	HBS	Administrative data ²²
Average benefit per household	186 Lt	139 Lt	-
Number of households raising 3 and more children under age 24 in full-time education	58373	50788	-
Number of beneficiaries	51935	44986	48000
Annual state expenses	116 million Lt	75,3 million Lt	88 million Lt
Annual state expenses for households raising 3 children	59,3 million Lt	50,6 million Lt	57 million Lt
Annual state expenses for households raising more than 3 children	56.7 million Lt	24.7 million Lt	31 million Lt

Source: own calculation

²² This column indicates annual state spending, based on the assumptions in the abovementioned text.

In addition to examination of the HBS and the EU-SILC generated simulation results, comparison to administrative data confirms the abovementioned finding of EU-SILC inflation of the weights on the families raising more than three children. This dataset predicts quite precisely the potential spending for the families with three children (4% higher expenses), but departs extremely with the prediction of the benefits for families with more than three children (more than 80% higher expenses). On the other hand, HBS generates the annual amount with closer to the expected state annual spending for the year of 2004. However it predicts deflated amount of potential spending.

Consequently, simulation of the family benefit to large families based on EU-SICL data draws attention on the potential distortion of predicted results, while simulating other types of Lithuanian family benefits, in which household size plays a major role for granting eligibility rights. In addition to this observation, differences between simulation results and administrative data could also occur due to approximate income testing or a failure to distinguish between full-time and part-time studying. As it is already indicated in the previous sub-chapter, a gap between simulation results and administrative data could also occur due to:

- The assumption on the 100% take-up ratio;
- Specific simulation assumptions, such as benefit distribution for the entire year;
- Different references to age based on EU-SILC and HBS data.
- Different demographic structures embedded in EU-SILC 2005 and HBS 2004 data.

4.3.3. Birth grant

On June 30, 2004, the birth grant was defined as a universal lump-sum benefit paid to one of the parents raising a child or a guardian upon the birth of a child. It equalled to 6 MSL and was only paid if a child was not maintained in an institution financed by the state or municipality, or if a child did not decease before the time of application (*Lietuvos Respublikos, 1994*).

This benefit was sustained by the *new Law on the Allowances to Children* (Republic of Lithuania, 2004). Furthermore, the size of the benefit was only changed since January 1, 2005. Hence, the benefit of a value of 6 MSL was paid during the whole course of 2004.

According to the *Ministry of Social Security and Labour*, the state spent 22.2 million Lt for the birth grant in 2004, and it was distributed to 29600 beneficiaries (2007). The reform of July 1, 2004, did not affect the payment or eligibility rules of this benefit. Therefore, the reported administrative data presents the actual state spending on the benefit.

The Table 11 shows that simulation of this benefit requires only demographic data on a child's age. This data is available in both datasets, even though, it does not align uniformly in terms of their time reference points.

Table 11. Information needed for the simulation of a birth grand

	EU-SILC	HBS	Main differences of variables
Age of child	RX020	P01_age	RX020 reports age at the end of income reference period, while P01_age reports age at the time of the interview.

Source: data collected by authors

Simulation results that are presented in the Table 12 indicate significant differences between HBS and EU-SILC predicted results.

Table 12. Birth grant simulation results

	EU-SILC	HBS	Administrative data ²³
Average benefit per household	781 Lt	771 Lt	-
Number of beneficiaries	44518	29599	29600
Annual state expenses	34.8 million Lt	22.9 million Lt	22.2 million Lt

Source: own calculation

Analogously to the simulation of the benefit to families raising three or more children, the EU-SILC based simulation generates highly inflated state spending, which exceeds the actual administrative data by more than 50%. Oppositely, the use of the HBS data enables simulation, which forecasts very accurate representation of the state spending if compared to the administrative data. This precision applies both in terms of the number of beneficiaries and the annual state spending.

As this simulation is conducted on the basis of the age data only, the conclusions could be drawn that the observed differences are due to the different demographic structure embedded in EU-SILC and HBS data:

- Different yearly references points: the used EU-SILC data refers to the year of 2005, while HBS refers to the year of 2004.
- Different recording of the birth data: the used EU-SILC variable refers to the age at the income reference period (end of 2004), while HBS refers to the age on the day of the interview.
- Differences in the survey used calibration weights with respect to representation of the newborn population.

²³ This column indicates the actual state spending, as a birth grant was effective during the whole course of 2004.

Conclusions

Both Household Budget Survey (HBS) and Survey on Income and Living Conditions (EU-SILC) contain sufficient range of household and individual level variables for the purpose of simulation of selected Lithuanian tax and benefit policies. However, as this research suggests, these datasets tend to have a higher potential of benefit (especially, benefits for families and children) rather than tax (especially, personal income taxation) simulation.

Analysis of benefit simulation results points to the need of a further research on the representative weights of EU-SILC and HBS data on demographical, social and economic characteristics, which are not calibrated for vulnerable population groups during the initial design of the surveys. For example, when variable indicating person's age at the end of the year (in the EU-SILC database) is used in simulations, it causes significant differences in the outcomes of benefits, targeted at newborn population or families raising more than three children. Moreover, it is worth noticing that the EU-SILC database records demographical information reflecting the year of 2005, while its data on income represents the situation of previous calendar year. This issue reduces the strength of comparative research on the HBS and the EU-SILC for the same year of the data collection.

Analysis of the Lithuanian personal income taxation system indicates a need of additional information, which could be used to make more plausible assumptions on the taxpayers and taxable income types. Due to the limited scope of the analysis, this research does not simulate payments of the personal income tax. Nonetheless, such a simulation is feasible, as the practice of microsimulation of direct taxes by using Household Budget Surveys or the EU-SILC in the EU countries, as Austria, Poland or Slovenia, indicates.

In general a conclusion could be drawn, that HBS offers higher possibilities of simulating Lithuanian benefit system, as in addition to demographic and socio-economic description of the individuals, it offers variables on the beneficiaries of the state transfers or non-cash transfers. This complementary information could enable simulations on the changes in the social policy parameters, based on the observed distribution of the benefits. As some of the Lithuanian social policies are interlinked, HBS information on different types of beneficiaries also enables more precise estimation on the compliance with eligibility rules. On the other hand, the comparison of HBS and EU-SILC datasets highlights the difficulties and dangers of simulation results' interpretation, if the underlying data come from methodologically different datasets. Therefore, the use of EU-SILC data for Lithuania offers a higher comparative value in terms of tax and benefit simulation across the EU.

References

1. Aidukaite, J. (2006). *The formation of social insurance institutions of the Baltic States in the post-socialist era*. Journal of European Social Policy, 16: 259.
2. Anderson, J. M. (Capital Research Associates) (2003). *Models for Retirement Policy Analysis: Report to the Society of Actuaries*. [Accessed on 15 01 2008]. Access via website: <http://www.soa.org/research/pension/research-models-for-retirement-policy-analysis.aspx>
3. Davies, J.B. (2004). *Microsimulation, CGE and Macro Modelling for Transition and Developing Economies*. WIDER Discussion Paper No. 2004/08.
4. EUROSTAT. (2007a). *Description of SILC User database variables: Cross-sectional and Longitudinal (Version 2005.1 from 01-06-07)*.
5. EUROSTAT. (2007b). *EU-SILC User database description (Version 2005-2 from 27-06-07)*.
6. Figari, F., Levy, H., and Sutherland, H. (2007a). *Using EU-SILC for policy simulation: prospects, some limitations and suggestions*. EUROMOD Working Paper No. EM1/07.
7. Figari, F., Levy, H., Lietz, Ch., Mantovani, D., Paulus, A., and Sutherland, H. (2007b). *EUROMOD: the Tax-Benefit Microsimulation model for the EU*. Paper presented at the 1st International Microsimulation Association conference "Celebrating 50 years of Microsimulation", August 19-22, 2007.
8. Lelkes, O. (2007). *Tax-Benefit Microsimulation Models In Eastern Europe*. International Journal of Microsimulation, 1(1), 54-56.
9. *Lietuvos Respublikos Gyventojų pajamų mokesčio įstatymas* (2002). Vilnius: Lietuvos Respublikos Seimas, 2002m. liepos 2 d. Nr.IX-1007.
10. *Lietuvos Respublikos Išmokų vaikams Įstatymo 2, 5, 6, 8 ir 12 straipsnių pakeitimo Įstatymas* (2004). Vilnius: Lietuvos Respublikos Seimas, 2004 m. rugsėjo 29 d. Nr. IX-2470.
11. *Lietuvos Respublikos Sveikatos draudimo įstatymas* (2003). Vilnius: Lietuvos Respublikos Seimas, 2004 m. balandžio 22 d., Nr.I-1343.
12. Lietuvos Respublikos Valstybinių pašalpu šeimoms, auginančioms vaikus, įstatymo 1,2,3,4⁽¹⁾ straipsnių pakeitimo ir papildymo įstatymas, 2001 m. sausio 25 d. NR. IX-168, Vilnius
13. *Lietuvos Respublikos Valstybinių pašalpu šeimoms, auginančioms vaikus, įstatymas* (1994). Vilnius: Lietuvos Respublikos Seimas, 1994 m. lapkričio 3 d., Nr. I-621.
14. *Lietuvos Respublikos Valstybinio socialinio draudimo įstatymas* (2004). Vilnius: Lietuvos Respublikos Seimas, 2003 m. lapkričio 18 d., Nr.I-1336.
15. Lietuvos Vyriausybės 2004-03-24 nutarimas Nr. 316 „Dėl minimaliojo darbo užmokesčio didinimo“ (Žin., 2004, Nr. 46-1511).
16. Lietuvos Vyriausybės 2004-03-24 nutarimas Nr. 325 „Dėl valstybinės socialinio draudimo bazinės pensijos didinimo“ (Žin., 2004, Nr. 46-1520).
17. Ministry of Social Security and Labour (2005) *Social Report 2004*. Vilnius.
18. Seimas of the Republic of Lithuania (2004). *Law of the Republic of Lithuania on Benefits to Children*. May 18, 2004, No. IX-2237, Vilnius.

19. Statistics Lithuania (2005) *Household income and expenditure 2004*, Vilnius.
20. Statistics Lithuania (2006) *Household income and expenditure 2005*, Vilnius
21. Statistics Lithuania (2007) *Income and Living Conditions 2005*, Vilnius.
22. Statistikos departamentas prie Lietuvos Respublikos Vyriausybės (2007e). *Pajamų ir gyvenimo sąlygų tyrimas*. [Accessed on 30 01 2008]. Prieiga per internetą: http://www.stat.gov.lt/lt/methodology/view/?id=7&pq_id=1580
23. Stirling, T., Lazutka, R. (2006). *Lithuania, 2005 tax-benefit system*. Euromod, I-CUE Feasibility Study.
24. Valstybinė mokesčių inspekcija (2004). *VMI administruojamų nacionalinio biudžeto pajamų surinkimo apžvalga, 2004 m.*
25. Valstybinio socialinio draudimo fondo valdyba (2005). *Lietuvos respublikos valstybinio socialinio draudimo fondo biudžeto vykdymo ataskaita: 2004 m. Gruodžio 31 d.*