https://doi.org/10.15388/vu.thesis.217
https://orcid.org/0000-0003-4001-36779

VILNIUS UNIVERSITY

Nerijus
Černiauskas

Impact of selected taxes and transfers on income inequality in Lithuania

DOCTORAL DISSERTATION

Social Sciences,
Economics (S 004)

VILNIUS 2021

This dissertation was written between 2017 and 2021 at the Vilnius University. The research was supported by InGRID grant REFERENCE N. C06_02.

Academic supervisor:
Prof. Dr. Romas Lazutka (Vilnius University, Social Science, Economics – S 004).

This doctoral dissertation will be defended in a public meeting of the Dissertation Defence Panel:
Chairman – Prof. Dr. Teodroas Medaiskis (Vilnius University, Social Sciences, Economics – S 004).
Members:
Prof. Dr. Vincentas Giedraitis (Vilnius University, Social Sciences, Economics – S 004).
Dr. Jekaterina Navickė (Vilnius University, Social Sciences, Sociology – S 005).
Dr. Alari Paulus (Eesti Pank, Estonia, Social Sciences, Economics – S 004).
Prof. Dr. Karsten Staehr (Tallinn University of Technology, Social Sciences, Economics – S 004).

The dissertation shall be defended at a public meeting of the Dissertation Defence Panel at 13:00 on 17th of September 2021 in Room 403 of the Faculty of economics and business administration. Address: Saulėtekio al. 9, Vilnius, Lithuania Tel. +370 5 236 6126 ; e-mail: evaf@evaf.vu.lt.
The text of this dissertation can be accessed at the libraries of Vilnius university, as well as on the website of Vilnius University: www.vu.lt/lt/naujienos/ivykiu-kalendorius.

https://doi.org/10.15388/vu.thesis.217
https://orcid.org/0000-0003-4001-3677

VILNIAUS UNIVERSITETAS

Nerijus
Černiauskas

Pasirinktų mokesčių ir pervedimų poveikis pajamų nelygybei Lietuvoje

DAKTARO DISERTACIJA

Socialiniai mokslai,
ekonomika (S 004)

VILNIUS 2021

Disertacija rengta 2017 - 2021 metais Vilniaus universitete. Mokslinius tyrimus rėmė InGRID grant REFERENCE N. C06_02.

Mokslinis vadovas:
prof. dr. Romas Lazutka (Vilniaus universitetas, socialiniai mokslai, ekonomika – S 004).

Gynimo taryba:
Pirmininkas – **prof. dr. Teodroas Medaiskis** (Vilniaus universitetas, socialiniai mokslai, ekonomika – S 004).
Nariai:
prof. dr. Vincentas Giedraitis (Vilniaus universitetas, socialiniai mokslai, ekonomika – S 004).
dr. Jekaterina Navickė (Vilniaus universitetas, socialiniai mokslai, sociologija – S 005).
dr. Alari Paulus (Estijos centrinis bankas, Estija, socialiniai mokslai, ekonomika – S 004).
prof. dr. Karsten Staehr (Talino technikos universitetas, Estija, socialiniai mokslai, ekonomika – S 004).

Disertacija ginama viešame Gynimo tarybos posėdyje 2021 m. rugsėjo mėn. 17 d. 13 val. Vilniaus universiteto Ekonomikos ir verslo administravimo fakulteto 403 posėdžių salėje. Address: Saulėtekio al. 9, Vilnius, Lietuva, Tel. +370 5 236 6126 ; e-mail: evaf@evaf.vu.lt.
Disertaciją galima peržiūrėti Vilniaus universiteto bibliotekoje ir VU interneto svetainėje adresu: https://www.vu.lt/naujienos/ivykiu-kalendorius.

CONTENTS

INTRODUCTION

Relevance. Persistently high and increasing levels of income inequality within countries continue figuring in political and academic debates. In 2013 Barack Obama, the President of the United States, declared income inequality as the "defining issue of our time", since it has been undermining economic growth and social and political cohesion in the country. In 2017, on the other side of the Atlantic, the European Central Bank's President Mario Draghi labelled the inequality in Europe as highly destabilizing. Following the recessions sparked by COVID, governments in EU and US increased spending by 1.8 trillion euro and 1.9 trillion dollars respectively, and part of this spending is being justified by lowering inequalities. Contemporary academics such as Alvaredo et al. (2018), Piketty (2015), and Atkinson and Piketty (2010) all note rising income inequality within countries and carry out research to identify its causes, consequences, and efficient ways to address it.

Lithuania is also participating in this debate. Researchers and international organizations have found income inequality to be elevated and rising in Lithuania (Navickė 2020; Rakauskienė et al. 2020; Lazutka 2017; Navickė and Lazutka 2017, 2016; IMF 2016; Skučienė 2008) and above the ideal level of income inequality (Volodzkienė 2020; Grigoli and Robles 2017). In particular, authors note that taxes and transfers do relatively little to reduce income inequality in Lithuania (Navickė 2020; IMF 2016; Lazutka 2017). Reducing income inequality was even included in the 2016-2020 and 2020-2024 governmental agendas with concrete policies aiming to reduce it (LR Vyriausybė 2017; LR Seimas 2020). Nevertheless, a number of outstanding issues prevent applying the optimal policies to reduce income inequality.

While the above-mentioned authors and policy makers are in favour of reducing income inequality as one of the means for achieving a more equal society, it is important to mention that numerous other researchers regard income inequality as a tool for producing economic incentives. According to this approach, policy makers may overestimate the risks and underestimate the rewards associated with higher income inequality. However, even this approach cannot deny the need to better understand, monitor and manage the risks related to income inequality.

Assuming that both perfect (complete) income inequality and perfect income equality (perfectly equal income distribution) are undesirable (and unsustainable), defining and constructing a true "optimal" level of income inequality can be regarded as the main challenge for both scientists and policy makers. This is especially relevant when income is conceptualized as a broad concept that contains monetary (e.g., paper money) and nonmonetary (in-kind) income. There are strong arguments that at least a certain level of income inequality is desirable, since it rewards people for different

degrees of effort, risk, and forgone investment into physical or human capital. However, there are also strong arguments that "high" income inequality is undesirable, as it is difficult to reconcile with contemporary values such as democracy, equality of opportunity and "capabilities" – a term coined by Amartya Sen – a set of feasible outcomes that individuals can choose to achieve. Without income, the capability set is greatly reduced: one is less likely to receive a good education, receive quality health services, have time to devote to civil and cultural activities or even take the time to find a higher-paid job.

Current research tries to contribute to the process of addressing this challenge through application of sets of different techniques to analyse the impact of selected economic and demographic factors on growth and fluctuation of income inequality. Moreover, this dissertation also provides analysis and compares income inequality both within selected groups and between the groups of Lithuanian citizens, thus deepening our understanding of the variety and complexity of different factors relevant to different sectors of income inequality. Widening of the analytical scope, as well as elaboration and specification of factors that influence different income inequality indicators in Lithuania from different angles by using different techniques, contributes to a deeper understanding of interdependencies among economic variables relevant to general and specific forms of income inequality. The findings of this research might be useful for further analysis, including that of the impact of income inequality on the labour market and the development of the country's economy, as well as for the policy makers who are looking for effective means to reduce income inequality in Lithuania.

The constructive approach within this dissertation predetermines a focus on monetary income inequality rather than inequality in a broader sense (e. g. inequality of capabilities). First, income inequality is heavily associated with other capabilities, as already explained above. Therefore, a more equal society in terms of income is also likely to be a more equal society overall. Additionally, monetary income is continuous and easier to track and measure than other aspects, such as education, health, etc., so, it is more effective to focus first on the research of monetary income inequality before expanding the scope.

Aim and novelty. The aim of this dissertation, which was started in 2017, is to contribute to the debate on income inequality in Lithuania. It does so by assessing the impact of selected taxes and transfers on income inequality in Lithuania in the hope that policies will become more able to reduce the inequality. Nevertheless, the dissertation provides a positive understanding of the impact of selected taxes and transfers on income inequality in Lithuania rather than proposing a normative approach on

how the tax and benefit system should operate. This is because much more follow-up research is required to propose unambiguous policy recommendations. Therefore, the dissertation provides answers to the four relevant questions and suggests possible future research topics that could eventually lead to wanted policy recommendations.

Before outlining the four questions that are answered in the dissertation, let us clarify the definitions and meaning of income and selected taxes and transfers. In this dissertation, income refers to household disposable income, which is defined as household market income (labour income and capital income of all household members) plus public transfers (pensions, unemployment benefits, etc.) and private transfers (transfers between households such as gifts or inheritance) minus income taxes and social insurance contributions. The household income is equivalised (i.e., adjusted for household size) and therefore referred to as equivalised disposable income. This definition of (equivalised disposable) income is used by EUROSTAT and therefore allows comparing the results of this dissertation with official statistics, which is where income inequality in Lithuania was found to be elevated, and allows contributing to the debate. However, this means a focus only on selected taxes (income taxes and social contributions) and not on value added taxes and excise taxes which are relevant for a different definition of income (sometimes refereed to augmented disposable income). Similarly, I largely focus on transfers in cash, and not on transfers in kind, even though the provision of public health and public education clearly have important distribution effects and matter for a yet another definition of income (or well-being in general). In one case, however, I do consider transfers in kind specifically, government-funded pre-school education, when analysing whether this transfer in kind helps reduce market income differences between males and females around the birth of the child. Clearly, broadening the definition of income would open interesting future research avenues.

With these definitions in mind, the dissertation answers two broader questions and two specific questions to help evaluate the impact of selected taxes and transfers on income inequality in Lithuania. The first broad question is why income inequality in Lithuania was the highest among all EU-28 countries in 2015, the most recent data at the time, and to what extent can taxes and transfers explain this. The second: why did income inequality grow and fluctuate in Lithuania in 2007–2015 and, again, what role did taxes and transfers play in this period. While both of these have been tackled before (see, e.g., Navickė 2020; IMF 2016), the current research addresses them more extensively than previously, thereby providing a better understanding of the impact of taxes and transfers on income inequality. The information obtained from the two broader questions raised several specific questions, two of which are also answered in

the dissertation. The first specific question is why households pay a relatively small share of their income in taxes and social contributions and how does this differ from optimal amounts. The second specific question focuses on how the wage gap between males and females is affected when the child is born and can childcare provision, a type of transfer in kind, help reduce the wage gap. The reason for choosing these topics is that they all contribute to a better understanding of the impact of taxes and transfers on income inequality while maintaining the same definition of income (this is particularly because the same data sets were used to answer these questions).

Research chapters and methods. The four questions are tackled in four research chapters of the dissertation. These chapters present four separate studies that have been submitted in the form of articles to academic journals. Each chapter has its own introduction, methodology, conclusions, limitations, and suggestions for further studies. Three of the four articles have been written with co-authors and have already been accepted for publication, two of them have been published. The exact questions answered, and the methods used in the chapters are outlined below.

Several steps are taken to answer why was income inequality in Lithuania one of the highest among all EU-28 countries. First, the Gini Index is computed from the EU-SILC survey data for Lithuania, several other EU-28 countries, and the EU-28 as a whole[1]. Then, univariate decomposition techniques are used to decompose the Gini Indexes into their income components (e.g., labour income, social transfers, and capital income) and subgroup components (e.g., employed, self-employed, and retired). This allows breaking down overall income inequality between households into inequality contributions of specific incomes or subgroups. Finally, the redistribution impact of taxes and benefits (as well as social insurance contributions) in Lithuania is compared to that of the EU. This is done by calculating a redistribution index, which is also decomposed into the Kakwani progressivity index and level components for taxes, social insurance contributions and benefits.

The reasons for growing and fluctuating income inequality in Lithuania in 2007–2015 are investigated in chapter 2. Using a mixture of microsimulation and reweighting techniques, the impact of the change of a single factor at a time on the change of total income inequality found in EU-SILC data is quantified. The impact of changes of four broad factors are quantified in this way: changes in the economic policy related to tax and benefit rules, the labour market structure, the economic returns in labour and capital markets, and the demographics. Additionally, the demographic factor was further decomposed into granular components of education, marriage, etc.

1. In the chapter, other decomposition measures are also calculated, such as the Atkinson Index, but the conclusions do not change. Therefore, the Gini Index is used throughout the dissertation

The question of why do households pay a relatively small share of their income in taxes and social contributions and how does this differ from optimal taxes is addressed in chapter 3. First, the actual direct taxes and social contributions paid by households as a share of their income are compared with the statutory rates that should be paid. The actual rates are found in EU-SILC data. These are then compared to statutory rates, i.e., the rates that these individuals would have to pay according to the law. The statutory taxes and social contributions are obtained by taking income and eligibility relevant information from EU-SILC data and the tax and benefits rules from the tax and benefit simulator EUROMOD. Finally, optimal taxes were calculated using a model that considers the income distribution of the country, labour elasticities, government's budget constraint and society's tolerance of (in)equality.

The effect of the child penalty and childcare on the wage gap is investigated in chapter 4. Using the EU-SILC data set and an event study methodology, the earnings and other labour market variables of men and women are compared up to 3 years before and up to 3 years after the birth of the first child for the Baltic countries. Then, enrolment rates into early childhood education and care (ECEC) are regressed on earnings on men and women. This allows comparing whether male and female earnings react differently to their child's enrolment into ECEC.

Acknowledgement of contributions. As the majority of work in this dissertation is based on research papers compiled together with co-authors, I would like to acknowledge each of our contributions. In preparing the paper which formed the basis of chapter 1, I was responsible for obtaining the data, writing the literature review, drafting the text, and carrying out most of the calculations. Andrius Čiginas compiled the appendix, computed bootstrapped weights, edited the text, and advised on math-related issues. We both tackled reviewers' comments and agreed on the final version of the paper. In preparing the paper for chapter 2, Denisa Sologon and I were responsible for adapting Denisa Sologon's and her co-authors code (previously used to analyse changes of income inequality in Luxembourg) for Lithuania. Denisa also contributed to the introduction. In addition, I was responsible for data description and analysis of the data. The methodology was written by Cathal O'Donoghue. Linas Tarasonis described Lithuania's economic context. All four authors have contributed to drafting, editing, and reviewing the text following the responses from the referees. Regarding chapter 3, Alain Jousten supervised the project, and I carried out the computations. We both analysed the data and have co-written the draft and the final version of the paper. Chapter 4, as well as other non-body chapters, were compiled by me.

Statements for defence. Income inequality is higher in Lithuania than in the EU

irrespective of which popular inequality measure is used to estimate it and which of the popular equivalence scales is used.

Income inequality within subgroups contributes more to total income inequality than one between subgroups. This means that if we divide society into certain groups (e.g., males and females, the young and the elderly) differences within the groups will be much more significant than between the groups. Income inequality among the unemployed is particularly large.

Inequality between those who are in employment and those who are not is much larger in Lithuania when compared to other EU countries. This can be partly explained by unequal distribution of labour income and self-employment income as well as by the low level of public benefits and low progressivity of income taxes and social insurance contributions. Income taxes and social insurance contributions are effectively regressive for the self- employed in Lithuania unlike in the EU.

The income tax progressivity is much smaller for the self-employed than the employed. A large portion of self-employment income is not reported to the tax authorities in Lithuania and, hence, is never taxed. The optimal tax rates are roughly in line with the statutory rates for the employed but are much higher for the self-employed. This suggests that, following more detailed investigation, the tax rates on the self-employed can be raised.

In the 2007–2015 period, rising returns (wages and capital income) were most responsible for growing income inequality. The tax and benefit system offset the increase, but only during the period of 2007–2011. Income inequality also partly increased due to the reducing share of married households and partly due to rising share of people with higher education.

The male-female wage gap at least temporarily widens about twofold following the birth of the child in the Baltics. Women who make use of formal childcare tend to have higher earnings than those who do not even when earnings before the birth of the child are controlled for.

Dissertation outline. The dissertation consists of this introductory chapter and is followed by four research chapters. While each chapter has its own conclusions, more general conclusions are presented in separate conclusions chapter .

1. MEASUREMENT AND DECOMPOSITION OF LITHUANIA'S INCOME INEQUALITY

1.1. Introduction

Income inequality in Lithuania has been one of the largest in the EU and is still growing. Specifically, the Gini coefficient of equivalised disposable income, a common measure of inequality, stood at 36.9% in 2018 for Lithuania (Eurostat 2018d). This was the second-largest Gini coefficient among the surveyed EU countries, second to Bulgaria, and exceeded the EU average income inequality by over 6 Gini points. Additionally, income inequality in Lithuania has increased by 5 Gini points since 2012. All this happened in the context of more general concern over rising income inequality within major countries (Atkinson and Piketty 2010; OECD 2011a, 2015a, 2015b) and increasing empirical evidence that income inequality may hinder economic growth (Aghion et al. 1999; Berg and Osrty 2011; Ostry et al. 2014; Cingano 2014; Grigoli and Robles 2017). The size and dynamics of income inequality in Lithuania along with warnings about its possible negative consequences encouraged political and economic debate in Lithuania. There was an interest to re-examine whether income inequality in Lithuania is indeed one of the largest within the EU, what contributes to income inequality, and what policy could be efficient at reducing it. This study focuses on these questions: how confident are we in claiming that Lithuania's income inequality is high, what factors lay behind such inequality and how much can redistribution of direct taxes and public transfers reduce income inequality.

We first analysed the extent to which income inequality is high. Even though the Gini of equivilised income does suggest this, a high Gini is not sufficient for such a claim. Besides the issue of estimating standard errors and testing for different equivalent scales, which can also change the ranking of countries according to income inequality (Buhmann et al. 1988), the Gini index itself is subject to criticism. This is because the Gini index, just like any summary inequality measure, entails social judgements on the undesirability of inequality (Anthony B Atkinson et al. 1970). Specifically, the Gini is more sensitive to inequalities in the middle of the distribution rather than the tails. This is not necessarily a desirable property, especially for Lithuania, where the highest level of inequality was found in the tails (IMF 2016).

For this reason, we employed several statistical tests to examine whether we can claim that equivalised income inequality in Lithuania is one of the highest across the EU. First, we have evaluated the sampling errors to verify that conclusions from the sample data do not contradict the actual situation. Rao et al. (1992) bootstrapped stan-

dard errors based on survey design information reconstructed according to Goedemé (2013) and Zardo Trindade and Goedemé (2016) allow to estimate the likely biases. Second, we have adjusted household income by alternative equivalence scales. We use the OECD-modified equivalence scale and the square root equivalence scale. Third, we have calculated inequality with other summary measures, thereby explicitly focusing on different segments of the distribution rather than the middle. We have estimated inequality using alternative measures to the Gini index: the Atkinson index and the Generalized entropy index as in Jenkins (2017) with standard inequality preference parameter values. We found that income inequality is statistically larger than the income inequality in other countries regardless of the equivalence scale or the summary measure used. This also strengthens the following analysis which is based on the Gini index.

Next, we have investigated why equivalised income inequality is higher compared to other countries using univariate factor and subgroup decompositions that decompose inequality into parts. These decompositions are purely statistical: they do not incorporate agent responses to any covariate. Nevertheless, these decompositions help identify the households amongst which inequality is acute and suggest which aspects should be looked into deeper.

Factor component decomposition decomposes inequality measure by disaggregating it into mutually exclusive and exhaustive income components, for example, labour and capital income. Two versions of this method are well known: the natural decomposition as in Anthony F Shorrocks (1982) that focuses on the decomposition of the variance and the Lerman and Yitzhaki (1985a) decomposition that is used to decompose the Gini coefficient. We use the latter method, as the Gini is a more conventional index of inequality. This method was used by, for example, Garner and Terrell (1998) to examine income inequality in Slovakia and Czechia in the early transition period.

Subgroup decomposition decomposes inequality measures within and between mutually exclusive and exhaustive subgroups, for example, inequality between males and females and inequality amongst males and amongst females. There are many ways to decompose subgroups as illustrated in Cowell (2011) and Yitzhaki and Lerman (1991). We apply the Yitzhaki and Lerman (1991) method to decompose the Gini in a way that is closer to the chosen factor decomposition technique.

From the decompositions, we see that labour income inequality is much higher in Lithuania than elsewhere in Europe. Additionally, in line with previous findings (e.g. IMF 2016), the tax and public transfers system plays less of a redistributive role in Lithuania than in other countries. To understand why, we looked into marginal ef-

fects: how does a 1% change in tax and transfers affect income inequality. We also looked into redistributive effects: how much do taxes and transfers reduce inequality according to Kakwani (1977). Finally, we decompose the redistributive effect into the progressivity index and the average rate of tax and public transfers and compare this with that of the EU. This lets us calculate how much can inequality be reduced due to a change in progressivity and average tax and public transfer rates.

Overall, our results suggest that equivalised income inequality in Lithuania is one of the highest in the EU and this finding is robust to various statistical tests. The decompositions reveal large inequalities between and within many groups of households in Lithuania. The largest inequalities lie between the employed and the rest of the population, and this kind of inequality has been rising over time. Inequalities within the unemployed and those working in the agricultural sector are particularly distinct. The factor decomposition shows that labour income, especially self-employment income, is more unequally distributed in Lithuania than elsewhere. Public transfers and taxes seem to reduce income inequality in Lithuania less than in other countries. This is because taxes and public transfers in Lithuania are less progressive and the tax and public transfer rates are lower than in the EU. Income taxes and social contributions are effectively regressive for the self-employed in Lithuania unlike in the EU. It is found that to reduce income inequality in Lithuania via redistribution, the focus should be placed on increasing the progressivity of taxes and average public transfer rates.

The chapter is structured as follows: in Section 2, we give definitions of income and describe the data set used throughout the empirical investigation. The other three sections answer three research questions, each using its methodology and provide comments on the results. The final section concludes.

1.2. Definitions and data on income

We focus on equivalised disposable income inequality. Let us explain each term in more detail. *Income* is defined as a yearly disposable income. To get the disposable income we subtract taxes and social contributions from gross income. We include the social contributions of the employee and employer, as we see both of them affecting the demand for labour. In addition, a new law in 2019 requested employees to pay the majority of employees' social contributions (see SODRA 2020). Gross income is the sum of market income (labour income with social insurance contributions and capital income) and transfers (both private and public). In cases when we refer to public transfers to analyse redistribution, we add private transfers to the definition of market income. The unit of observation is a *household*. This assumes that household

members share their income and make joint decisions. To adjust for household size, an equivalence scale is used.

Focusing on equivalised income rather than individual income affects the results and this should be briefly justified. Research literature suggests that individuals make economic decisions taking themselves as well as their household members into consideration (see, among others, Vogler and Pahl 1994). For example, the income of all household members comprises a common budget constraint (Chiappori and Meghir 2015) thereby influencing each household member's behaviour. Additionally, some transfers are only granted at a household level (e.g. social assistance transfer) making the allocation of this transfer to any specific member artificial. Nevertheless, each household member has their preferences and a typically unequal control of the household's budget with evidence suggesting that decisions taken within a household are rarely joint and more often dominated by a specific household member (Pahl 1995). Therefore, while it is useful to look at equivalised income inequality to get a first idea of how unequally income is distributed within society, specific questions require looking into inequality within a household (for example, when determining how child transfers should be allocated if mothers are more likely to spend on children rather than fathers).

The data on income and covariates come from the yearly European Union Statistics on Income and Living Conditions (EU-SILC) instrument running since 2004. The data are compiled from a mixture of the survey and administrative sources. Each year around 5 thousand Lithuanian households with around 10 thousand persons over 16 years old who agree to share information on their income are included. The exact number of households and persons recorded in Lithuania and other countries in 2015 is shown in Table 1.1. Most of these persons provided all information on income, as can be seen from column 5 titled "Observations". As all EU member states collect data using the same methodology, we can compare the inequality in Lithuania with that of other EU countries.

While the data is explained by Eurostat (2018c), several features are mentioned here. The survey captures household income and, therefore, certain income components are available for the household rather than the individual level. Therefore, the income of all household members is summed up and allocated to each household member. While most covariates are recorded at the time of the interview, income is recorded for a previous year (the reference year). In this chapter, all years represent reference years. While the EU-SILC has a large survey component, some countries make use of register (administrative) data and are referred to as register countries. In 2015, the register countries included Cyprus, Denmark, Finland, Latvia, Lithua-

Table 1.1: EU-SILC summary statistics for 2015 income reference year

Country code	Country name	Households (thousands)	Household members (thousands)	Observations (thousands)	Average income (thousand euro)	Median income (thousand euro)	Gini (percent)
1	2	3	4	5	6	7	8
AT	Austria	6	10.8	10.8	26.1	23.7	27.2
BE	Belgium	5.9	11.1	11.1	24.3	22.3	26.3
BG	Bulgaria	7.3	15.6	15.5	3.9	3.2	37.7
CH	Switzerland	7.8	14.9	14.9	50.9	44.3	29.4
CY	Cyprus	4.2	9.4	9.4	16.9	14	32.1
CZ	Czech Republic	8.5	16.2	16.2	8.8	7.8	25.1
DE	Germany	13.3	23.3	23.1	23.9	21.2	29.8
DK	Denmark	6.3	11.8	11.8	32.1	28.7	27.7
EE	Estonia	6	12.5	12.5	10.1	8.6	32.7
EL	Greece	18.3	38	37.9	8.7	7.5	34.3
ES	Spain	14.2	30.7	30.7	15.8	13.7	34.5
FI	Finland	10.6	20.7	20.7	26.4	23.6	25.4
FR	France	11.5	21.3	21.3	25.3	21.7	29.2
HR	Croatia	7.6	17	17	6.3	5.7	29.8
HU	Hungary	8	15.9	15.8	5.4	4.8	28.2
IE	Ireland	5.2	10.2	10.2	25.5	22.4	29.5
IT	Italy	21.3	41.5	41.5	18.3	16.2	33.1
LT	Lithuania	4.8	9.6	9.6	7	5.6	37.0
LU	Luxembourg	3.8	8.2	8.2	39.8	34.4	31.5
LV	Latvia	6	11.6	11.6	7.5	6.4	34.5
NL	Netherlands	12.7	24.1	24.1	25.4	22.7	26.9
NO	Norway	6.9	13.6	13.6	43	39.6	24.9
PL	Poland	12	27.1	27.1	6.7	5.9	29.8
PT	Portugal	10.6	22.7	22.7	10.6	8.8	33.9
RO	Romania	7.4	15.8	15.7	2.7	2.4	34.8
RS	Serbia	5.6	15.1	15.1	3	2.6	38.6
SE	Sweden	5.8	11.2	11.2	27.3	25.2	27.6
SI	Slovenia	8.6	21.9	21.9	13.2	12.3	24.3
SK	Slovakia	5.7	14.1	14.1	7.4	7	24.3
UK	United Kingdom	9.7	17.8	17.6	24.6	21.1	31.5

The variables "Households" and "Household members" are the unique number of households and household members in the data set. The variable "Observations" refers to those household members for whom all income data is available. Columns 6 to 8 refer to the average, median and the Gini coefficient of the population estimate of equivalized household disposable income.

nia, the Netherlands, Northern Ireland, Norway, Slovenia, Sweden, and Switzerland. Finally, survey weights are used to form conclusions on the population from the sample data. The weights are further adjusted according to Eurostat (2018b): weights of household members over 16 years old are scaled up by distributing weights of those under 16.

1.3. Is income inequality in Lithuania high?

First, we have examined inequality from the full data sample and then analysed subgroup inequality (inequality between- and within- subgroups) in Lithuania.

1.3.1. Inequality

The most popular measure of the level of inequality is the Gini coefficient. The higher the Gini, the greater the level of inequality and it stood at $G = 0.37$ for Lithuania in 2015 (Eurostat 2018d). The Gini is represented, as in Lerman and Yitzhaki (1985a), by two times the covariance between income y and the rank of income $F(y)$ divided by average income μ,

$$G = \frac{2Cov(y, F(y))}{\mu}, \tag{1.1}$$

which describes inequality within the entire population. Since we have sample data only, we modify (1.1) to include sample weights, as shown in (5) in the Appendix.

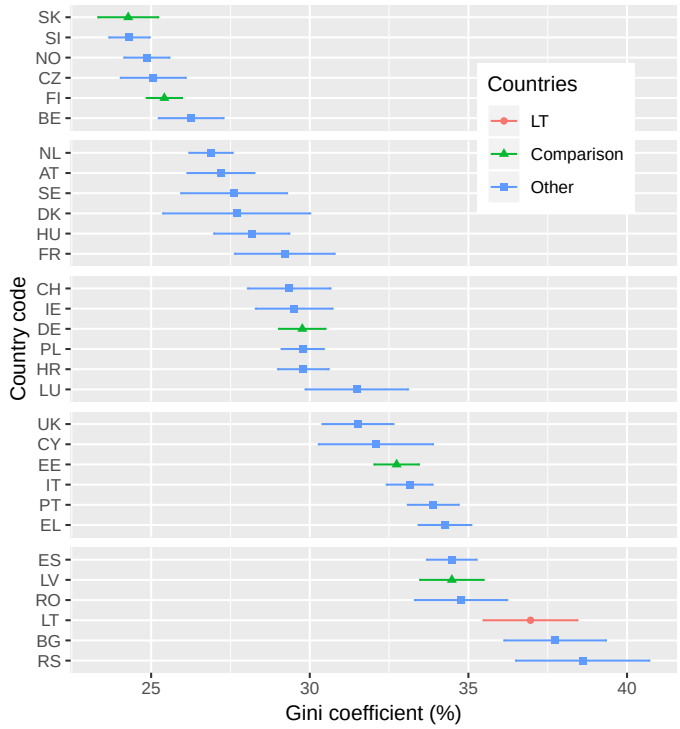
Lithuania's Gini coefficient has been compared with the Gini coefficients of all countries that are included in the EU-SILC data set for 2015 in Figure 1.1 and with the Gini coefficients for a subset of all countries in Table 1.2. The subset of countries includes the Baltic States, Finland as one of the Scandinavian countries, Germany– which represents the average inequality in the EU and Slovakia, where inequality is the lowest. As in previous studies (IMF 2016; Lazutka 2017), income inequality in Lithuania is one is of the highest according to the EU-SILC. The estimated confidence intervals (Figure 1.1) and standard errors (Table 1.2) indicate that this is statistically significant. For example, the Gini in Lithuania is about 7 Gini points higher than in Germany. The latter also happens to be the median in terms of inequality within the whole EU-SILC sample of countries.

Although Table 1.2 focuses on fewer countries, it provides more statistics on inequality than Figure 1.1. In Figure 1.1, household disposable income is equivalised by the OECD-modified equivalence scale. In Table 1.2, two different scales are used: the OECD-modified scale and the square root equivalence scale. The square root scale increases the Gini for Lithuania by 0.3 points, yet remains with the highest level of income inequality among all countries and 7 Gini points higher than the median country.

Furthermore, in Table 1.2, the generalized Gini coefficient, $G(v)$ (Yitzhaki 1983), where parameter v represents inequality aversion. This inequality parameter represents the dissatisfaction expressed towards inequality. With this parameter we can model different societal preferences. The value $v = 2$ gives the standard Gini, v between 1 and 2 represent lower inequality dissatisfaction and $v > 2$ indicates higher dissatisfaction. The measurement $G(1.5)$ results in lower Gini values in all countries for both equivalence scales (i.e. inequality is not as "bad"). Additionally, the difference between the Gini in Lithuania and the median country shrinks to 5 Gini points for both scales. Nevertheless, inequality in Lithuania remains significantly the highest out of the sample of 6 countries. Setting $v = 4$ increases the Gini index, but for Lithuania it remains the highest among the selected countries.

Finally, the Gini is compared with other measures of inequality. Other prominent measures include the Atkinson index (Atk) and General entropy index (GEI), see Das and Parikh (1982), Cowell (2000), and Plat (2012). Both of these measures show that

Figure 1.1: The Gini coefficients of equivalised disposable income in all EU-SILC countries



Household disposable income is equivalised by the OECD-modified scale (OECD 2012). Confidence intervals are estimated by using Rao et al. (1992) bootstrap methodology. Information on survey design is provided by Goedemé (2013) and Zardo Trindade and Goedemé (2016).

the higher the value, the greater the inequality. Both indexes also feature inequality aversion parameters. In the Atkinson index, a parameter value close to zero means indifference about inequality, while higher values show that people dislike it. In contrast, high GEI parameter values mean that people are indifferent about inequality. In all cases, inequality in Lithuania remained significantly the highest.

1.3.2. Subgroup inequality

The previous subsection has shown that inequality in Lithuania is large when compared to EU countries. Next, we will consider inequality between and within population subgroups, for example, between males and females and amongst males and females. Then we will estimate stratification–the extent to which income of one group

Table 1.2: Income inequality measures under different equivalence scales

Country	Equivalence scale	G(2)	G(1.5)	G(4)	GEI(0)	GEI(1)	GEI(2)	Atk(1)	Atk(0.1)
DE	OECD	29.764	19.602	46.279	0.157	0.158	0.220	0.146	0.016
		(0.373)	(0.318)	(0.388)	(0.004)	(0.006)	(0.016)	(0.004)	(0.001)
EE	OECD	32.738	21.096	51.419	0.192	0.171	0.188	0.175	0.017
		(0.358)	(0.256)	(0.463)	(0.006)	(0.004)	(0.005)	(0.005)	(0.000)
FI	OECD	25.416	16.897	40.216	0.112	0.116	0.150	0.106	0.011
		(0.283)	(0.236)	(0.340)	(0.003)	(0.004)	(0.009)	(0.003)	(0.000)
LT	OECD	36.957	24.644	55.797	0.254	0.233	0.306	0.224	0.023
		(0.755)	(0.609)	(0.801)	(0.012)	(0.011)	(0.022)	(0.009)	(0.001)
LV	OECD	34.479	22.756	53.403	0.217	0.202	0.255	0.195	0.020
		(0.511)	(0.432)	(0.563)	(0.007)	(0.008)	(0.019)	(0.006)	(0.001)
Median	OECD	29.764	19.719	46.279	0.158	0.163	0.228	0.146	0.016
		(0.373)	(0.538)	(0.388)	(0.008)	(0.011)	(0.039)	(0.006)	(0.001)
SK	OECD	24.277	15.624	40.310	0.115	0.106	0.136	0.109	0.011
		(0.482)	(0.383)	(0.682)	(0.006)	(0.007)	(0.022)	(0.005)	(0.001)
DE	Sqr. rt	30.224	19.873	47.169	0.163	0.162	0.223	0.150	0.016
		(0.379)	(0.324)	(0.389)	(0.004)	(0.006)	(0.016)	(0.004)	(0.001)
EE	Sqr. rt	33.158	21.305	52.399	0.199	0.175	0.190	0.180	0.017
		(0.354)	(0.253)	(0.451)	(0.006)	(0.004)	(0.005)	(0.005)	(0.000)
FI	Sqr. rt	25.918	17.202	41.213	0.117	0.120	0.155	0.110	0.012
		(0.288)	(0.240)	(0.347)	(0.003)	(0.004)	(0.009)	(0.003)	(0.000)
LT	Sqr. rt	37.383	24.854	56.684	0.261	0.237	0.307	0.230	0.023
		(0.763)	(0.625)	(0.790)	(0.012)	(0.011)	(0.023)	(0.009)	(0.001)
LV	Sqr. rt	35.039	23.063	54.513	0.226	0.207	0.259	0.202	0.021
		(0.521)	(0.447)	(0.553)	(0.007)	(0.008)	(0.021)	(0.006)	(0.001)
Median	Sqr. rt	29.699	19.873	47.169	0.161	0.162	0.223	0.149	0.016
		(0.662)	(0.324)	(0.389)	(0.008)	(0.006)	(0.016)	(0.007)	(0.001)
SK	Sqr. rt	25.000	16.043	41.302	0.120	0.109	0.132	0.113	0.011
		(0.447)	(0.350)	(0.622)	(0.006)	(0.006)	(0.016)	(0.005)	(0.001)

Table contains inequality measures for the median country (from all EU countries) and selected countries for each equivalence scale. $G(v)$ represents the Gini index with values $v = 2, 1.5, 4$ of parameter of inequality aversion, $GEI(a)$ stands for the General entrhopy index, and $Atk(b)$ is the Atkinson index, where $b = 1, 0.1$ and $a = 0, 1, 2$ represents the degree of inequality aversion. Bootstrapped standard errors are presented in the parenthesis.

overlaps income of other groups.

Continuing the discussion started in Section 1.2, the interpretation of a subgroup may not be straightforward, as we are dealing with equivalised income instead of individual income, but can be explained with the help of an example. Imagine a household composed of 1 male and 1 female. Then, comparing household income (i.e. adding up household members' income and allocating the summed household income to each member) implies no income inequality between the male and the female in that household. However, this is only true if all households have the same number of males and females. Some households are consisting of more males, while others have a higher number of female members. If males tend to earn more than females, households with more males will earn higher equivalised household income than equivalised households with more females. In aggregate, this will lead to inequalities between the subgroups. Inequality between this group should be interpreted as "inequality between male and female-dominated households". This way, we can combine information on

household income and the composition with individual characteristics. Of course, there could be other variables that are also correlated. For example, females tend to live longer and are therefore more likely to be retired and hence receive lower income. However, this approach abstracts from other variables.

The methodology used to estimate inequality between subgroups is similar to the one used by IMF (2016) and is based on Eurostat (2018a). The methodology for estimating inequality within subgroups and stratification are adapted from Yitzhaki and Lerman (1991). Additionally, the technique proposed by Yitzhaki and Lerman (1991) is used to decompose total inequality into *between*, *within* and *stratification* terms to see which of them contributes most to inequality.

Inequality between subgroups

Inequality between subgroups refers to measured inequality between households grouped under certain criteria. For example, households can be grouped by "Sex" into two subgroups $l = 1$ and $l = 2$: "Males" and "Females". To estimate between subgroup inequality, we first estimate the weighted average income of a subgroup $\widehat{m}^{(l)}$ and then divide by the average weighted income of all subgroups $\widehat{m}$, see (7) in Appendix .1, to get an income ratio $\widehat{m}^{(l)}/\widehat{m}$. We then compare the ratio with that of the EU, namely of its member states that joined the EU before 2004 (old EU states), and with those Member States that joined it after 2004 (new EU states). Our method is similar to that used in the IMF (2016), but has several differences: the IMF (2016) analyse weighted income decile ratios while we compare weighted average income ratios. The IMF (2016) compares Lithuania to the EU, while we additionally compare it to new and old EU states to control for the development of countries. Finally, we have more grouping criteria (a total of 9) and estimate standard errors.

Our findings are in line with those of IMF (2016), which also reviews between-subgroup inequality in Lithuania. The IMF (2016) reveals large inequalities between the top and the bottom income deciles, between the employed and the unemployed and non-labour market participants, between the elderly and other age subgroups, as well as between educated and less educated households subgroups, i.e. these ratios are much higher in Lithuania than in the entire EU.

In addition to these findings, the results presented in Table 1.3 allow adding the following points:

- Differences of ratios are significant between many subgroups in Lithuania. The subgroups include those grouped according to the IMF (2016) criteria (activity status, age bracket, number of dependants, education) as well as ratios in other subgroups. For example, we split households based on the main income source. Those who receive largely self-employment income tend, on average,

Table 1.3: Ratios of average subgroup incomes in 2015

Grouping	Subgroup	EU	EU new	EU old	LT
activity status	employed	113.2 (0.5) [25.5]	115.4 (0.9) [9.8]	112.6 (0.6) [15.5]	123.2 (1.3) [5.2]
activity status	unemployed	61.0 (1.3) [3.3]	59.5 (1.9) [1.2]	61.3 (1.6) [2.0]	54.0 (3.6) [0.6]
activity status	retired	97.3 (1.0) [9.5]	95.2 (1.3) [4.2]	97.9 (1.2) [5.6]	71.1 (1.7) [2.0]
activity status	study	85.5 (1.6) [3.5]	88.4 (2.4) [1.3]	84.8 (1.8) [2.2]	86.8 (3.2) [0.7]
activity status	other	77.4 (1.3) [5.7]	70.5 (1.9) [2.2]	79.2 (1.6) [3.5]	68.9 (3.1) [1.0]
nr working	0	78.3 (0.9) [12.8]	75.1 (1.2) [4.6]	79.0 (1.0) [8.0]	53.3 (1.6) [2.4]
nr working	1	93.4 (0.9) [15.1]	93.6 (1.6) [5.8]	93.3 (1.0) [9.2]	94.1 (2.6) [3.1]
nr working	2	119.3 (0.9) [16.7]	117.8 (1.4) [6.7]	119.7 (1.0) [10.1]	131.4 (2.8) [3.3]
nr working	3	116.9 (2.3) [2.5]	115.6 (4.1) [1.5]	117.5 (2.7) [1.3]	130.5 (4.8) [0.6]
nr working	4	124.3 (4.2) [0.6]	124.0 (5.3) [0.4]	124.5 (6.0) [0.3]	124.9 (8.6) [0.2]
main income	employment	108.7 (0.6) [24.1]	112.2 (0.9) [11.7]	107.4 (0.8) [13.6]	111.5 (1.4) [6.1]
main income	self-employment	106.3 (3.2) [3.4]	84.3 (3.2) [1.7]	114.8 (4.2) [1.9]	174.9 (10.7) [0.7]
main income	other	88.7 (0.6) [20.4]	79.3 (1.6) [5.6]	90.4 (0.7) [13.4]	57.6 (1.6) [2.8]
occupation	basic level	74.6 (1.3) [4.5]	78.1 (4.1) [1.9]	73.7 (1.2) [2.7]	69.6 (2.5) [1.4]
occupation	mid level	90.0 (0.6) [21.0]	91.9 (0.8) [9.6]	89.3 (0.7) [12.2]	88.7 (1.7) [4.1]
occupation	technicians , associates	112.9 (1.3) [6.0]	118.7 (2.1) [1.6]	111.9 (1.4) [3.9]	117.8 (4.6) [0.6]
occupation	professionals	139.1 (1.8) [7.3]	145.8 (2.6) [2.4]	137.7 (2.1) [4.6]	129.1 (3.1) [1.9]
occupation	managers	137.9 (3.3) [2.3]	153.0 (5.5) [0.8]	134.7 (3.8) [1.5]	162.7 (11.6) [0.6]
sector	agriculture	74.2 (2.6) [1.0]	66.4 (2.5) [1.0]	83.4 (4.7) [0.3]	99.5 (6.7) [0.3]
sector	industry	115.6 (1.9) [3.4]	116.2 (2.3) [2.0]	115.3 (2.5) [1.8]	113.9 (3.8) [0.8]
sector	IT, finance, RE , admin	134.4 (2.7) [3.4]	158.2 (7.9) [0.9]	130.6 (2.8) [2.2]	157.2 (10.9) [0.6]
sector	public admin, education , health	118.0 (1.3) [5.8]	127.1 (2.3) [1.7]	116.4 (1.5) [3.7]	124.6 (3.2) [1.1]
sector	other services	100.2 (1.1) [7.6]	109.5 (1.8) [3.0]	97.8 (1.3) [4.6]	119.3 (2.8) [1.8]
age bracket	under 19	83.8 (2.0) [1.8]	82.4 (3.5) [0.7]	84.2 (2.3) [1.1]	83.6 (4.1) [0.4]
age bracket	19 - 29	91.3 (1.3) [6.9]	95.5 (1.9) [3.0]	90.0 (1.5) [4.1]	104.2 (3.8) [1.7]
age bracket	30 - 64	103.9 (0.4) [30.1]	104.0 (0.6) [12.0]	103.9 (0.5) [18.1]	107.7 (1.2) [5.7]
age bracket	65+	96.9 (1.1) [9.1]	93.5 (1.4) [3.3]	97.7 (1.3) [5.6]	76.2 (1.9) [1.9]
dependents	0	105.2 (0.6) [24.5]	105.4 (1.1) [8.9]	105.2 (0.7) [15.1]	101.0 (1.8) [4.8]
dependents	1	102.6 (1.3) [9.2]	107.0 (1.9) [4.3]	101.1 (1.6) [5.3]	106.7 (3.2) [2.1]
dependents	2	94.2 (1.1) [9.7]	93.2 (1.8) [3.9]	94.5 (1.3) [5.8]	102.4 (4.8) [2.0]
dependents	3	82.7 (2.1) [3.2]	80.7 (3.7) [1.2]	83.2 (2.5) [2.0]	73.9 (6.9) [0.5]
dependents	4	73.0 (4.5) [0.8]	60.9 (6.2) [0.3]	75.9 (5.3) [0.5]	55.6 (5.5) [0.1]
dependents	5	50.9 (4.5) [0.4]	40.0 (3.5) [0.3]	58.7 (6.1) [0.2]	47.1 (11.8) [0.1]
education	up to secondary	87.8 (0.4) [32.4]	87.9 (0.6) [14.1]	87.8 (0.5) [19.0]	78.2 (1.5) [4.6]
education	post-secondary	103.0 (2.5) [1.5]	107.3 (3.9) [0.6]	101.7 (3.1) [0.9]	92.7 (2.3) [1.9]
education	tertiary education	129.6 (1.0) [13.4]	142.1 (2.0) [3.9]	127.3 (1.2) [8.6]	138.1 (2.8) [3.0]
sex	male	102.2 (0.5) [22.7]	102.8 (0.9) [9.0]	102.1 (0.7) [13.7]	104.5 (0.8) [4.3]
sex	female	98.0 (0.5) [25.2]	97.4 (0.8) [10.0]	98.1 (0.6) [15.2]	96.4 (0.6) [5.3]

Ratios are defined as weighted average income of a subgroup divided by weighted average income of all subgroups within that grouping. Bootstrapped standard errors are in the parenthesis. Number of observations in thousands in brackets.

to have more disposable income than those who work as employees or others–a trend not observed in the EU as a whole. Significant inequality also exists between subgroups grouped by the number of people working in the household (nr working) and the sector where one works (sector).

- Ratios between the majority of the 9 subgroups are also significantly different from the ratios between their EU counterparts. Besides the subgroups in the IMF (2016) (those grouped by activity status, age bracket, education), the self-employed in Lithuania on average earn proportionally more than their EU counterparts. Additionally, those who work in the information technologies, finance, real estate, and administration sector (IT, finance, RE, admin) earn, on average, relatively more income in Lithuania than one would in the EU.

- There are some groups between which inequality in Lithuania is smaller as compared to the EU. For example, those working within the agricultural sector are relatively better off in Lithuania compared to the EU. Additionally, income ratios in Lithuania are more similar to those in the new EU states. In particular, those who are under 19 years old have very similar relative incomes both in Lithuania and in the new EU states.

In general, ratios between subgroups are largely persistent and slightly widening since 2010. This can be seen in Table 1.4 which shows the ratio dynamics in Lithuania. For example, there was a slowly widening gap between the employed and the retired. This could be explained by rising market incomes due to a recovering economy that benefited the employed while statutory pensions, the main source of income for the retired, did not increase in the period due to budget consolidation (Černiauskas et al. 2020). Once the recovery began, wages in the private sector started rising, especially IT, finance, RE, admin sector, while the government started raising public sector wages (Public admin, education, health) much later. This could also explain the rising ratio difference between the two sectors.

Inequality within subgroups

Inequality exists within subgroups in Lithuania. A common way to measure it is to calculate inequality measures for subgroup income as is done for total income (see $\widehat{G}_l$ in Formula (8) in Appendix .1). We have calculated the Gini coefficients for Lithuania's subgroups and compared them with the Gini coefficients of the EU, new and old EU states in Table 1.5.

- Most of the within-subgroup Gini coefficients examined in Table 1.5 are higher in Lithuania than in the EU. Especially large subgroup inequality exists among

Table 1.4: Ratios of average subgroup incomes in Lithunia

Grouping	Subgroup	2006	2010	2015
activity status	employed	120.2 (0.9) [6.2]	122.1 (1.1) [5.4]	123.2 (1.3) [5.2]
activity status	unemployed	51.1 (2.7) [0.4]	57.7 (2.3) [1.1]	54.0 (3.6) [0.6]
activity status	retired	70.4 (1.2) [2.0]	86.4 (1.3) [2.3]	71.1 (1.7) [2.0]
activity status	study	89.1 (2.2) [1.1]	86.2 (2.3) [1.1]	86.8 (3.2) [0.7]
activity status	other	70.7 (2.2) [1.1]	74.3 (1.9) [1.0]	68.9 (3.1) [1.0]
nr working	0	51.4 (1.2) [2.1]	66.2 (1.3) [3.1]	53.3 (1.6) [2.4]
nr working	1	86.5 (1.8) [3.0]	89.7 (1.8) [3.4]	94.1 (2.6) [3.1]
nr working	2	122.0 (1.6) [4.7]	128.9 (2.1) [3.6]	131.4 (2.8) [3.3]
nr working	3	137.5 (5.6) [0.8]	144.7 (6.8) [0.7]	130.5 (4.8) [0.6]
nr working	4	138.7 (11.2) [0.2]	122.2 (11.5) [0.2]	124.9 (8.6) [0.2]
main income	employment	114.9 (0.9) [7.5]	115.3 (1.1) [6.7]	111.5 (1.4) [6.1]
main income	self-employment	106.1 (5.6) [0.6]	127.1 (10.8) [0.4]	174.9 (10.7) [0.7]
main income	other	59.0 (1.4) [2.8]	71.0 (1.2) [3.9]	57.6 (1.6) [2.8]
occupation	basic level	72.1 (1.8) [1.6]	72.6 (1.7) [1.5]	69.6 (2.5) [1.4]
occupation	mid level	90.0 (1.1) [5.1]	87.6 (1.1) [4.8]	88.7 (1.7) [4.1]
occupation	technicians , associates	120.0 (4.3) [0.7]	119.3 (3.8) [0.7]	117.8 (4.6) [0.6]
occupation	professionals	146.4 (3.4) [1.5]	138.9 (2.8) [1.9]	129.1 (3.1) [1.9]
occupation	managers	149.9 (4.6) [0.7]	146.0 (4.6) [0.7]	162.7 (11.6) [0.6]
sector	agriculture	85.1 (5.2) [0.5]	90.9 (5.9) [0.4]	99.5 (6.7) [0.3]
sector	industry	109.9 (2.6) [1.1]	113.8 (3.9) [0.9]	113.9 (3.8) [0.8]
sector	IT, finance, RE , admin	168.2 (8.0) [0.2]	137.2 (4.9) [0.5]	157.2 (10.9) [0.6]
sector	public admin, education , health	131.2 (2.7) [1.4]	138.1 (3.1) [1.3]	124.6 (3.2) [1.1]
sector	other services	117.0 (1.9) [2.2]	111.6 (2.2) [1.8]	119.3 (2.8) [1.8]
age bracket	under 19	82.0 (2.6) [0.6]	80.3 (2.8) [0.6]	83.6 (4.1) [0.4]
age bracket	19 - 29	111.2 (2.9) [2.2]	101.7 (2.6) [2.0]	104.2 (3.8) [1.7]
age bracket	30 - 64	105.1 (0.9) [6.4]	104.3 (0.8) [6.4]	107.7 (1.2) [5.7]
age bracket	65+	73.6 (1.4) [1.8]	90.4 (1.6) [2.0]	76.2 (1.9) [1.9]
dependents	0	104.0 (1.4) [5.3]	102.7 (1.4) [4.9]	101.0 (1.8) [4.8]
dependents	1	106.0 (2.3) [2.9]	102.9 (2.2) [2.8]	106.7 (3.2) [2.1]
dependents	2	95.0 (2.8) [2.0]	99.0 (3.3) [2.3]	102.4 (4.8) [2.0]
dependents	3	72.5 (4.6) [0.5]	81.8 (6.5) [0.7]	73.9 (6.9) [0.5]
dependents	4	45.7 (10.5) [0.1]	66.8 (10.1) [0.2]	55.6 (5.5) [0.1]
dependents	5	35.9 (6.4) [0.1]	95.5 (16.1) [0.1]	47.1 (11.8) [0.1]
education	up to secondary	81.5 (0.9) [5.8]	81.7 (1.0) [6.0]	78.2 (1.5) [4.6]
education	post-secondary	95.9 (1.6) [2.4]	96.1 (2.0) [2.1]	92.7 (2.3) [1.9]
education	tertiary education	148.5 (2.4) [2.5]	140.6 (2.2) [2.9]	138.1 (2.8) [3.0]
sex	male	104.9 (0.6) [4.9]	102.9 (0.7) [4.9]	104.5 (0.8) [4.3]
sex	female	96.0 (0.5) [6.0]	97.6 (0.5) [6.1]	96.4 (0.6) [5.3]

Ratios are defined as weighted average income of a subgroup divided by weighted average income of all subgroups within that grouping. Bootstrapped standard errors are in the parenthesis. Number of observations in thousands in brackets.

Table 1.5: The Gini coefficient of income of subgroups in 2015

Grouping	Subgroup	EU	new EU states	old EU states	LT
activity status	employed	28.6 (0.4) [25.5]	29.7 (0.7) [9.8]	28.3 (0.4) [15.5]	33.0 (0.9) [5.2]
activity status	unemployed	35.1 (0.9) [3.3]	36.9 (1.3) [1.2]	34.6 (1.1) [2.0]	47.8 (2.2) [0.6]
activity status	retired	27.4 (0.5) [9.5]	25.9 (0.5) [4.2]	27.8 (0.7) [5.6]	29.6 (1.0) [2.0]
activity status	student	32.2 (0.8) [3.5]	30.1 (1.0) [1.3]	32.7 (0.9) [2.2]	35.1 (1.6) [0.7]
activity status	other	32.5 (0.8) [5.7]	31.3 (1.1) [2.2]	32.7 (1.0) [3.5]	38.0 (1.5) [1.0]
nr working	0	32.2 (0.6) [12.8]	31.1 (0.9) [4.6]	32.4 (0.6) [8.0]	31.5 (1.5) [2.4]
nr working	1	31.6 (0.6) [15.1]	32.2 (1.2) [5.8]	31.4 (0.6) [9.2]	36.4 (1.1) [3.1]
nr working	2	26.5 (0.4) [16.7]	27.2 (0.6) [6.7]	26.3 (0.5) [10.1]	31.6 (1.3) [3.3]
nr working	3	25.1 (0.9) [2.5]	27.6 (1.7) [1.5]	23.8 (1.0) [1.3]	21.0 (1.4) [0.6]
nr working	4	24.6 (1.3) [0.6]	25.2 (1.9) [0.4]	24.2 (1.8) [0.3]	17.7 (3.3) [0.2]
main income	employment	27.2 (0.3) [24.1]	27.0 (0.4) [11.7]	27.3 (0.4) [13.6]	31.3 (0.9) [6.1]
main income	self-employment	42.6 (1.3) [3.4]	39.1 (1.6) [1.7]	43.0 (1.5) [1.9]	39.7 (2.2) [0.7]
main income	other	31.9 (0.4) [20.4]	33.9 (1.3) [5.6]	31.5 (0.5) [13.4]	33.0 (1.3) [2.8]
occupation	basic level	28.7 (1.0) [4.5]	33.1 (3.5) [1.9]	27.4 (0.7) [2.7]	35.6 (1.4) [1.4]
occupation	mid-level	27.5 (0.3) [21.0]	27.5 (0.5) [9.6]	27.5 (0.4) [12.2]	34.8 (0.9) [4.1]
occupation	technicians, associates	25.2 (0.6) [6.0]	24.4 (0.7) [1.6]	25.3 (0.7) [3.9]	31.0 (1.5) [0.6]
occupation	professionals	29.2 (0.8) [7.3]	25.9 (0.8) [2.4]	29.8 (0.9) [4.6]	31.0 (0.9) [1.9]
occupation	managers	31.9 (1.1) [2.3]	30.9 (1.7) [0.8]	31.9 (1.3) [1.5]	39.6 (2.5) [0.6]
sector	agriculture	35.2 (1.7) [1.0]	35.2 (1.5) [1.0]	34.0 (2.9) [0.3]	44.5 (2.4) [0.3]
sector	industry	27.1 (0.9) [3.4]	26.4 (0.8) [2.0]	27.4 (1.2) [1.8]	30.3 (1.4) [0.8]
sector	IT, finance, RE, admin	30.8 (1.0) [3.4]	31.5 (3.2) [0.9]	30.5 (1.0) [2.2]	35.1 (2.9) [0.6]
sector	public admin, education, health	25.7 (0.5) [5.8]	25.0 (0.8) [1.7]	25.7 (0.6) [3.7]	28.7 (1.0) [1.1]
sector	other services	27.9 (0.6) [7.6]	27.0 (0.8) [3.0]	27.9 (0.7) [4.6]	32.7 (1.2) [1.8]
age bracket	under 19	31.1 (1.1) [1.8]	32.8 (1.5) [0.7]	30.6 (1.3) [1.1]	34.8 (1.9) [0.4]
age bracket	19-29	30.8 (0.7) [6.9]	31.0 (1.0) [3.0]	30.6 (0.9) [4.1]	35.8 (1.4) [1.7]
age bracket	30-64	31.1 (0.3) [30.1]	31.7 (0.6) [12.0]	31.0 (0.4) [18.1]	37.4 (0.9) [5.7]
age bracket	65+	28.4 (0.7) [9.1]	26.7 (0.5) [3.3]	28.8 (0.8) [5.6]	30.8 (1.1) [1.9]
dependents	0	31.1 (0.4) [24.5]	30.8 (0.7) [8.9]	31.1 (0.4) [15.1]	38.4 (0.8) [4.8]
dependents	1	29.9 (0.6) [9.2]	29.2 (0.8) [4.3]	30.1 (0.8) [5.3]	30.8 (1.3) [2.1]
dependents	2	28.4 (0.6) [9.7]	29.2 (1.0) [3.9]	28.1 (0.7) [5.8]	37.9 (2.3) [2.0]
dependents	3	32.2 (1.1) [3.2]	32.2 (2.1) [1.2]	32.2 (1.3) [2.0]	36.2 (3.0) [0.5]
dependents	4	30.5 (2.6) [0.8]	34.6 (3.0) [0.3]	29.1 (3.0) [0.5]	20.7 (3.9) [0.1]
dependents	5	25.8 (4.0) [0.4]	21.9 (3.9) [0.3]	23.7 (5.7) [0.2]	24.7 (9.9) [0.1]
education	up to secondary	29.0 (0.3) [32.4]	29.7 (0.6) [14.1]	28.8 (0.4) [19.0]	35.0 (0.8) [4.6]
education	post-secondary	27.8 (1.0) [1.5]	28.2 (1.3) [0.6]	27.6 (1.2) [0.9]	34.7 (1.2) [1.9]
education	tertiary education	29.6 (0.5) [13.4]	27.1 (0.6) [3.9]	30.0 (0.6) [8.6]	33.1 (1.2) [3.0]
sex	male	31.0 (0.4) [22.7]	31.1 (0.6) [9.0]	30.9 (0.4) [13.7]	37.0 (0.9) [4.3]
sex	female	30.6 (0.4) [25.2]	30.9 (0.7) [10.0]	30.5 (0.5) [15.2]	36.8 (0.8) [5.3]

Bootstrapped standard errors are provided in the parenthesis. Number of observations are shown in thousands in brackets.

those working in the agricultural sector and the unemployed.

- The above-mentioned within-group inequalities are much higher in Lithuania than in the EU. Additionally, households, where the main source of income is self-employment income, are also unequal among themselves, even though similar inequality within subgroups exists in new EU states. The Gini of households with many children is relatively small and we know from the between analysis that these households earn a much lower income.

Over time, inequality within subgroups increased in many subgroups. Table 1.6 shows that the rise has been especially strong since 2010. In particular, the Gini coefficient of the unemployed rose from 39.8 in 2004 to 47.8 in 2015. This may be in part due to unequal economic recovery, where some of the unemployed were able to

Table 1.6: Gini of subgroup incomes in Lithuania

Grouping	Subgroup	2006	2010	2015
activity status	employed	30.9 (0.6) [6.2]	30.6 (0.7) [5.4]	33.0 (0.9) [5.2]
activity status	unemployed	42.1 (2.0) [0.4]	39.8 (1.8) [1.1]	47.8 (2.2) [0.6]
activity status	retired	25.1 (0.6) [2.0]	24.2 (0.7) [2.3]	29.6 (1.0) [2.0]
activity status	study	32.9 (1.3) [1.1]	34.0 (1.0) [1.1]	35.1 (1.6) [0.7]
activity status	other	34.8 (1.2) [1.1]	29.7 (1.2) [1.0]	38.0 (1.5) [1.0]
nr working	0	26.7 (1.2) [2.1]	29.8 (1.0) [3.1]	31.5 (1.5) [2.4]
nr working	1	31.9 (0.8) [3.0]	32.5 (1.0) [3.4]	36.4 (1.1) [3.1]
nr working	2	29.5 (0.8) [4.7]	27.8 (0.8) [3.6]	31.6 (1.3) [3.3]
nr working	3	24.7 (1.6) [0.8]	25.4 (1.9) [0.7]	21.0 (1.4) [0.6]
nr working	4	24.5 (3.8) [0.2]	19.4 (3.3) [0.2]	17.7 (3.3) [0.2]
main income	employment	30.4 (0.6) [7.5]	30.0 (0.6) [6.7]	31.3 (0.9) [6.1]
main income	self-employment	37.2 (2.1) [0.6]	44.3 (2.3) [0.4]	39.7 (2.2) [0.7]
main income	other	29.1 (1.2) [2.8]	29.3 (1.0) [3.9]	33.0 (1.3) [2.8]
occupation	basic level	31.0 (1.2) [1.6]	29.8 (0.9) [1.5]	35.6 (1.4) [1.4]
occupation	mid level	29.9 (0.6) [5.1]	30.2 (0.7) [4.8]	34.8 (0.9) [4.1]
occupation	technicians , associates	30.5 (1.8) [0.7]	28.3 (1.4) [0.7]	31.0 (1.5) [0.6]
occupation	professionals	31.7 (0.9) [1.5]	28.8 (0.9) [1.9]	31.0 (0.9) [1.9]
occupation	managers	31.9 (1.3) [0.7]	32.1 (1.5) [0.7]	39.6 (2.5) [0.6]
sector	agriculture	41.3 (2.0) [0.5]	37.2 (2.4) [0.4]	44.5 (2.4) [0.3]
sector	industry	28.3 (1.1) [1.1]	31.1 (1.6) [0.9]	30.3 (1.4) [0.8]
sector	IT, finance, RE , admin	29.8 (1.8) [0.2]	28.3 (1.8) [0.5]	35.1 (2.9) [0.6]
sector	public admin, education , health	29.4 (1.0) [1.4]	27.5 (0.9) [1.3]	28.7 (1.0) [1.1]
sector	other services	30.4 (1.0) [2.2]	31.5 (0.9) [1.8]	32.7 (1.2) [1.8]
age bracket	under 19	33.5 (1.5) [0.6]	33.1 (1.5) [0.6]	34.8 (1.9) [0.4]
age bracket	19 - 29	34.3 (1.3) [2.2]	32.4 (1.0) [2.0]	35.8 (1.4) [1.7]
age bracket	30 - 64	33.2 (0.6) [6.4]	34.7 (0.7) [6.4]	37.4 (0.9) [5.7]
age bracket	65+	27.5 (0.9) [1.8]	24.7 (0.9) [2.0]	30.8 (1.1) [1.9]
dependents	0	35.6 (0.7) [5.3]	33.2 (0.7) [4.9]	38.4 (0.8) [4.8]
dependents	1	30.7 (0.9) [2.9]	31.2 (1.0) [2.8]	30.8 (1.3) [2.1]
dependents	2	29.1 (1.4) [2.0]	33.5 (1.4) [2.3]	37.9 (2.3) [2.0]
dependents	3	31.0 (3.1) [0.5]	36.2 (2.5) [0.7]	36.2 (3.0) [0.5]
dependents	4	35.2 (9.1) [0.1]	29.2 (5.7) [0.2]	20.7 (3.9) [0.1]
dependents	5	32.1 (7.3) [0.1]	8.5 (8.3) [0.1]	24.7 (9.9) [0.1]
education	up to secondary	31.1 (0.7) [5.8]	30.6 (0.6) [6.0]	35.0 (0.8) [4.6]
education	post-secondary	29.6 (0.7) [2.4]	31.5 (1.0) [2.1]	34.7 (1.2) [1.9]
education	tertiary education	30.7 (0.8) [2.5]	29.5 (0.9) [2.9]	33.1 (1.2) [3.0]
sex	male	33.3 (0.6) [4.9]	33.3 (0.7) [4.9]	37.0 (0.9) [4.3]
sex	female	34.0 (0.6) [6.0]	32.7 (0.6) [6.1]	36.8 (0.8) [5.3]

Bootstrapped standard errors are in the brackets. Number of observations in thousands in brackets.

find some income sources, while others did not. Unemployment has risen substantially since the crisis and there have been many unemployment transfers handed out. However, these transfers were stopped to those who were unemployed for a longer time. Additionally, as the economy recovered, it became easier for the unemployed to be in employment for at least several months during the year. Similarly, there was a rise in inequality among those who are neither employed, unemployed, retired, or students (largely disabled). Additionally, there has been a rise in inequality among those who are over 65 and, to a lesser extent, those aged 30-64. Inequality increased within all the different education levels and within all occupations (managers in particular). Inequality increased in the agricultural sector as well as in the IT, finance, real estate and administration sectors (IT, finance, RE, admin).

Stratification between subgroups

Inequality is linked to stratification. Stratification measures whether the income of each member of a subgroup differs compared to the income of every member of all other subgroups. We use the methodology proposed by Yitzhaki and Lerman (1991), which measures stratification on a scale from -100 to 100. Value 100 indicates high stratification: all members of a subgroup have income that is different from members of other subgroups. Value 0 indicates no stratification–there is a perfect income overlap between the subgroups. Negative numbers indicate that the subgroup should actually be multiple subgroups, i.e. income of some subgroup members is much higher than that of members of other subgroups, however, some members also have much lower income than members of other subgroups. The estimates of measures of stratification in Table 1.7 allow us to make two more insights:

- Several subgroups in Lithuania are stratified. Families with more dependants are detached in terms of income from other subgroups and the difference is stark when compared to the EU. Households who are employed or have more employed members are stratified from the unemployed and those who do not participate in the labour market. Income stratification of these subgroups is greater in Lithuania than in the EU. Additionally, several subgroups are stratified in Lithuania to a similar extent as they are stratified in new EU states: subgroups characterised by occupation, education, and age bracket. This could signal that Lithuania, like in new EU states, is facing more labour market imbalances, where the demand for highly educated professionals is especially high, while redistribution channels are too weak to compensate for the income of those out of labour force (e.g. elderly).

- There are several subgroups that should form several smaller subgroups in

Table 1.7: Stratfication of subgroup income in 2015

Grouping	Subgroup	EU	new EU states	old EU states	LT
activity status	employed	17.7 (0.8) [25.5]	14.2 (1.5) [9.8]	18.6 (1.0) [15.5]	32.6 (2.0) [5.2]
activity status	unemployed	6.3 (2.3) [3.3]	7.3 (3.1) [1.2]	5.9 (2.8) [2.0]	-10.0 (5.8) [0.6]
activity status	retired	6.6 (0.8) [9.5]	10.4 (1.2) [4.2]	5.7 (1.0) [5.6]	11.7 (3.0) [2.0]
activity status	student	-3.5 (1.2) [3.5]	0.8 (2.0) [1.3]	-4.6 (1.5) [2.2]	1.3 (2.3) [0.7]
activity status	other	-1.6 (1.7) [5.7]	7.6 (2.2) [2.2]	-3.4 (1.9) [3.5]	-3.8 (3.3) [1.0]
nr working	0	-1.0 (1.3) [12.8]	6.1 (1.7) [4.6]	-2.7 (1.5) [8.0]	25.3 (6.6) [2.4]
nr working	1	-2.1 (0.8) [15.1]	-1.5 (1.5) [5.8]	-2.3 (0.9) [9.2]	-0.4 (2.2) [3.1]
nr working	2	23.2 (0.9) [16.7]	17.4 (1.7) [6.7]	24.9 (1.1) [10.1]	34.8 (2.6) [3.3]
nr working	3	15.8 (2.7) [2.5]	8.4 (5.6) [1.5]	19.8 (2.5) [1.3]	42.8 (3.7) [0.6]
nr working	4	18.0 (3.4) [0.6]	15.6 (5.2) [0.4]	19.5 (4.6) [0.3]	46.2 (9.1) [0.2]
main income	employment	16.3 (0.8) [24.1]	25.5 (1.6) [11.7]	14.4 (0.9) [13.6]	33.7 (2.4) [6.1]
main income	self-employment	-13.9 (1.3) [3.4]	-12.4 (2.2) [1.7]	-11.9 (1.8) [1.9]	14.8 (8.2) [0.7]
main income	other	-4.5 (0.8) [20.4]	-1.4 (2.2) [5.6]	-4.8 (0.9) [13.4]	18.4 (4.7) [2.8]
occupation	basic level	6.9 (2.3) [4.5]	1.1 (6.0) [1.9]	8.9 (1.9) [2.7]	2.0 (3.4) [1.4]
occupation	mid-level	7.0 (0.9) [21.0]	10.1 (1.4) [9.6]	6.4 (1.0) [12.2]	2.5 (1.8) [4.1]
occupation	technicians, associates	13.5 (1.1) [6.0]	14.9 (1.7) [1.6]	13.4 (1.2) [3.9]	13.9 (2.8) [0.6]
occupation	professionals	20.0 (1.6) [7.3]	27.0 (2.1) [2.4]	18.9 (1.9) [4.6]	22.3 (2.2) [1.9]
occupation	managers	10.4 (2.6) [2.3]	15.9 (4.2) [0.8]	9.4 (3.0) [1.5]	17.2 (4.3) [0.6]
sector	agriculture	-0.6 (3.8) [1.0]	9.9 (3.9) [1.0]	-5.1 (5.3) [0.3]	-19.7 (3.7) [0.3]
sector	industry	4.6 (1.2) [3.4]	7.6 (1.7) [2.0]	3.4 (1.6) [1.8]	3.7 (2.5) [0.8]
sector	IT, finance, RE, admin	5.5 (1.7) [3.4]	16.4 (3.7) [0.9]	4.1 (1.8) [2.2]	6.8 (3.9) [0.6]
sector	public admin, education, health	8.8 (1.2) [5.8]	12.6 (2.0) [1.7]	8.2 (1.4) [3.7]	9.8 (2.4) [1.1]
sector	other services	0.6 (1.1) [7.6]	5.8 (1.7) [3.0]	-0.2 (1.4) [4.6]	0.2 (2.1) [1.8]
age bracket	under 19	-1.0 (1.8) [1.8]	-3.6 (3.1) [0.7]	-0.3 (2.2) [1.1]	1.0 (3.0) [0.4]
age bracket	19-29	-1.4 (0.9) [6.9]	-1.1 (1.5) [3.0]	-1.4 (1.2) [4.1]	-0.6 (2.6) [1.7]
age bracket	30-64	0.0 (0.7) [30.1]	-1.2 (1.2) [12.0]	0.4 (0.8) [18.1]	2.7 (1.6) [5.7]
age bracket	65+	4.8 (0.9) [9.1]	7.9 (1.3) [3.3]	4.1 (1.0) [5.6]	7.9 (2.6) [1.9]
dependents	0	2.3 (0.7) [24.5]	4.3 (1.2) [8.9]	1.7 (0.9) [15.1]	-4.5 (2.0) [4.8]
dependents	1	1.7 (1.0) [9.2]	3.5 (1.5) [4.3]	1.3 (1.2) [5.3]	12.5 (2.9) [2.1]
dependents	2	3.8 (1.0) [9.7]	3.1 (1.8) [3.9]	4.0 (1.1) [5.8]	1.2 (2.8) [2.0]
dependents	3	-3.1 (2.0) [3.2]	-1.1 (4.2) [1.2]	-3.6 (2.2) [2.0]	-2.5 (7.0) [0.5]
dependents	4	1.3 (6.5) [0.8]	3.3 (7.3) [0.3]	0.7 (7.4) [0.5]	43.4 (7.1) [0.1]
dependents	5	26.1 (19.4) [0.4]	48.6 (9.1) [0.3]	16.9 (25.3) [0.2]	12.0 (37.0) [0.1]
education	up to secondary	7.4 (1.0) [32.4]	12.3 (2.0) [14.1]	6.5 (1.1) [19.0]	2.7 (2.1) [4.6]
education	post-secondary	2.9 (2.0) [1.5]	3.0 (2.6) [0.6]	3.1 (2.4) [0.9]	3.6 (1.9) [1.9]
education	tertiary education	17.2 (1.3) [13.4]	26.3 (1.9) [3.9]	15.8 (1.4) [8.6]	27.7 (2.4) [3.0]
sex	male	1.0 (0.6) [22.7]	0.6 (1.0) [9.0]	1.1 (0.7) [13.7]	2.4 (1.1) [4.3]
sex	female	-0.8 (0.6) [25.2]	-0.2 (1.0) [10.0]	-0.9 (0.7) [15.2]	-1.8 (1.0) [5.3]

Bootstrapped standard errors are provided in the parenthesis. Number of observations are shown in thousands in brackets.

Lithuania. The unemployed, for example, have a stratification value of -9.9, meaning that some unemployed are relatively well off, while others are not. This could reflect that some of the unemployed are still getting unemployment transfers, can take on part-time work, or are simply living in a high-income household, while others do not. Similar tendencies also exist in the agricultural sector, with some being much better off than others.

Stratification between groups has been increasing, especially since 2010 (see Table 1.8). This is particularly apparent when considering activity status: the stratification coefficient of those employed rose from 17.8% in 2010 to 32.6% in 2015. However, this could be largely attributed to a market correction, as the stratification coefficient was around 28.7% before the crisis.

Table 1.8: Stratfication of subgroup incomes in Lithuania

Grouping	Subgroup	2006	2010	2015
activity status	employed	28.7 (1.6) [6.2]	17.8 (1.9) [5.4]	32.6 (2.0) [5.2]
activity status	unemployed	6.5 (4.7) [0.4]	4.0 (4.3) [1.1]	-10.0 (5.8) [0.6]
activity status	retired	15.6 (2.2) [2.0]	21.0 (1.9) [2.3]	11.7 (3.0) [2.0]
activity status	study	0.2 (1.8) [1.1]	-4.5 (1.9) [1.1]	1.3 (2.3) [0.7]
activity status	other	0.5 (2.4) [1.1]	9.1 (2.2) [1.0]	-3.8 (3.3) [1.0]
nr working	0	44.3 (3.8) [2.1]	22.7 (2.5) [3.1]	25.3 (6.6) [2.4]
nr working	1	0.9 (1.8) [3.0]	-0.4 (2.1) [3.4]	-0.4 (2.2) [3.1]
nr working	2	31.3 (2.1) [4.7]	26.3 (2.6) [3.6]	34.8 (2.6) [3.3]
nr working	3	34.0 (5.7) [0.8]	33.1 (5.5) [0.7]	42.8 (3.7) [0.6]
nr working	4	27.4 (13.8) [0.2]	36.6 (10.4) [0.2]	46.2 (9.1) [0.2]
main income	employment	35.0 (2.4) [7.5]	17.1 (2.4) [6.7]	33.7 (2.4) [6.1]
main income	self-employment	-3.3 (3.0) [0.6]	-16.9 (5.2) [0.4]	14.8 (8.2) [0.7]
main income	other	23.3 (4.3) [2.8]	20.9 (2.5) [3.9]	18.4 (4.7) [2.8]
occupation	basic level	5.2 (3.1) [1.6]	11.4 (2.1) [1.5]	2.0 (3.4) [1.4]
occupation	mid level	9.0 (1.4) [5.1]	8.3 (1.7) [4.8]	2.5 (1.8) [4.1]
occupation	technicians , associates	13.9 (2.8) [0.7]	11.1 (3.7) [0.7]	13.9 (2.8) [0.6]
occupation	professionals	22.1 (2.7) [1.5]	22.0 (2.6) [1.9]	22.3 (2.2) [1.9]
occupation	managers	19.8 (3.4) [0.7]	14.7 (3.4) [0.7]	17.2 (4.3) [0.6]
sector	agriculture	-15.7 (4.3) [0.5]	-10.7 (5.1) [0.4]	-19.7 (3.7) [0.3]
sector	industry	5.3 (2.1) [1.1]	0.9 (2.7) [0.9]	3.7 (2.5) [0.8]
sector	IT, finance, RE , admin	17.0 (6.5) [0.2]	8.6 (4.4) [0.5]	6.8 (3.9) [0.6]
sector	public admin, education , health	10.2 (2.2) [1.4]	13.0 (2.5) [1.3]	9.8 (2.4) [1.1]
sector	other services	3.1 (1.9) [2.2]	-0.8 (1.9) [1.8]	0.2 (2.1) [1.8]
age bracket	under 19	-0.3 (2.1) [0.6]	0.3 (2.8) [0.6]	1.0 (3.0) [0.4]
age bracket	19 - 29	1.9 (2.2) [2.2]	-2.5 (2.1) [2.0]	-0.6 (2.6) [1.7]
age bracket	30 - 64	5.3 (1.3) [6.4]	-6.8 (1.3) [6.4]	2.7 (1.6) [5.7]
age bracket	65+	7.9 (2.5) [1.8]	19.3 (1.9) [2.0]	7.9 (2.6) [1.9]
dependents	0	-4.1 (1.7) [5.3]	3.5 (1.7) [4.9]	-4.5 (2.0) [4.8]
dependents	1	7.2 (2.1) [2.9]	2.1 (2.3) [2.8]	12.5 (2.9) [2.1]
dependents	2	8.4 (2.7) [2.0]	-3.1 (2.6) [2.3]	1.2 (2.8) [2.0]
dependents	3	3.1 (7.0) [0.5]	-9.6 (5.4) [0.7]	-2.5 (7.0) [0.5]
dependents	4	28.9 (10.5) [0.1]	5.6 (17.7) [0.2]	43.4 (7.1) [0.1]
dependents	5	48.1 (18.4) [0.1]	38.2 (25.9) [0.1]	12.0 (37.0) [0.1]
education	up to secondary	6.2 (1.6) [5.8]	10.6 (1.5) [6.0]	2.7 (2.1) [4.6]
education	post-secondary	8.0 (1.5) [2.4]	3.8 (1.7) [2.1]	3.6 (1.9) [1.9]
education	tertiary education	29.6 (2.5) [2.5]	22.8 (2.7) [2.9]	27.7 (2.4) [3.0]
sex	male	4.1 (1.0) [4.9]	0.0 (0.8) [4.9]	2.4 (1.1) [4.3]
sex	female	-3.0 (0.8) [6.0]	0.3 (0.7) [6.1]	-1.8 (1.0) [5.3]

Bootstrapped standard errors are in the parenthesis. Number of observations in thousands in brackets.

Subgroup decomposition

We have analysed between- and within-subgroup inequality and stratification separately. Now, we will identify how much each of the terms contributes to the Gini of disposable income in Lithuania and compare this to the EU, new and old EU states. To do this, we will use the methodology provided by Yitzhaki and Lerman (1991), outlined in Appendix .1.

The subgroup decomposition results are presented in Table 1.9. The Gini coefficient is decomposed into within, between, and stratification component for each of the 9 groupings considered before. The following conclusions can be drawn:

Table 1.9: Decomposition of the Gini coefficient in 2015

Grouping (1)	Decomposition (2)	EU (3)	new EU states (4)	old EU states (5)	LT (6)
activity status	between	3.3 (0.2) [47.5]	3.8 (0.3) [18.7]	3.1 (0.2) [28.8]	7.0 (0.5) [9.5]
activity status	stratification	-1.7 (0.1) [47.5]	-1.9 (0.2) [18.7]	-1.6 (0.1) [28.8]	-3.5 (0.3) [9.5]
activity status	within	29.2 (0.3) [47.5]	29.3 (0.5) [18.7]	29.2 (0.3) [28.8]	33.4 (0.7) [9.5]
nr working	between	3.5 (0.2) [47.9]	3.5 (0.3) [19.0]	3.5 (0.2) [28.9]	10.1 (0.7) [9.6]
nr working	stratification	-1.8 (0.1) [47.9]	-1.7 (0.2) [19.0]	-1.8 (0.1) [28.9]	-5.0 (0.4) [9.6]
nr working	within	29.1 (0.3) [47.9]	29.3 (0.5) [19.0]	29.0 (0.3) [28.9]	31.9 (0.8) [9.6]
main income	between	1.1 (0.1) [47.9]	2.9 (0.3) [19.0]	0.8 (0.1) [28.9]	8.4 (0.7) [9.6]
main income	stratification	-0.5 (0.1) [47.9]	-1.4 (0.2) [19.0]	-0.4 (0.1) [28.9]	-4.1 (0.4) [9.6]
main income	within	30.2 (0.3) [47.9]	29.6 (0.5) [19.0]	30.2 (0.3) [28.9]	32.6 (0.7) [9.6]
occupation	between	4.6 (0.2) [41.1]	5.3 (0.4) [16.2]	4.5 (0.3) [24.9]	5.5 (0.7) [8.6]
occupation	stratification	-2.4 (0.1) [41.1]	-2.6 (0.2) [16.2]	-2.4 (0.2) [24.9]	-2.8 (0.4) [8.6]
occupation	within	28.0 (0.3) [41.1]	27.5 (0.5) [16.2]	28.0 (0.3) [24.9]	34.1 (0.7) [8.6]
sector	between	1.9 (0.2) [21.2]	4.2 (0.5) [8.5]	1.5 (0.2) [12.7]	1.2 (0.5) [4.5]
sector	stratification	-1.0 (0.1) [21.2]	-2.1 (0.3) [8.5]	-0.7 (0.1) [12.7]	-0.5 (0.3) [4.5]
sector	within	27.9 (0.4) [21.2]	27.7 (0.7) [8.5]	27.8 (0.4) [12.7]	32.4 (0.9) [4.5]
age bracket	between	0.4 (0.1) [47.9]	0.3 (0.1) [19.0]	0.4 (0.1) [28.9]	1.4 (0.2) [9.6]
age bracket	stratification	-0.1 (0.1) [47.9]	-0.1 (0.1) [19.0]	-0.2 (0.1) [28.9]	-0.5 (0.1) [9.6]
age bracket	within	30.6 (0.3) [47.9]	30.8 (0.5) [19.0]	30.5 (0.3) [28.9]	36.1 (0.8) [9.6]
dependents	between	0.8 (0.1) [47.9]	1.6 (0.3) [19.0]	0.7 (0.1) [28.9]	1.1 (0.4) [9.6]
dependents	stratification	-0.4 (0.1) [47.9]	-0.7 (0.1) [19.0]	-0.3 (0.1) [28.9]	-0.4 (0.2) [9.6]
dependents	within	30.4 (0.3) [47.9]	30.2 (0.5) [19.0]	30.4 (0.3) [28.9]	36.3 (0.8) [9.6]
education	between	3.4 (0.2) [47.2]	4.6 (0.3) [18.7]	3.2 (0.2) [28.5]	5.9 (0.6) [9.5]
education	stratification	-1.8 (0.1) [47.2]	-2.3 (0.2) [18.7]	-1.7 (0.1) [28.5]	-3.1 (0.4) [9.5]
education	within	29.2 (0.3) [47.2]	28.9 (0.5) [18.7]	29.2 (0.3) [28.5]	34.2 (0.7) [9.5]
sex	between	0.0 (0.1) [47.9]	0.1 (0.1) [19.0]	0.0 (0.1) [28.9]	0.2 (0.1) [9.6]
sex	stratification	0.0 (0.0) [47.9]	0.0 (0.0) [19.0]	0.0 (0.0) [28.9]	-0.1 (0.0) [9.6]
sex	within	30.8 (0.3) [47.9]	31.0 (0.5) [19.0]	30.7 (0.3) [28.9]	36.9 (0.8) [9.6]

The first figure in columns (3-6) represents the contribution to Gini coefficient of equivalised household disposable income. Bootstrapped standard errors are provided in the parenthesis. Number of observations are shown in thousands in brackets.

- The majority of inequality decomposes into within-groups rather than between-groups in Lithuania. The largest between-contribution is observed between different households which have a different number of people working (nr working, 10 Gini points), but even here the within-contribution is 3 times higher. This finding is not surprising, as inequality within subgroups is often found to matter more (see Elbers et al. 2008), suggesting that the majority of variation

in income is between households of similar observable characteristics. Income inequality within groups is also more important for the EU. Additionally, several household characteristics seem to not contribute to inequality significantly in Lithuania, for example, sex.

- Except for education, labour market characteristics of the household are more important in explaining inequality than demographics. For example, the different number of people working, the main source of income of the household, and the occupation individually explain 5-10 Gini points. The between-contribution, when grouping people according to activity status is 7 Gini points. This means that if all household members were employed and would earn employment income, the Gini coefficient would fall by 7 points and become similar to the EU Gini coefficient. This between-contribution in Lithuania is about 2 times higher than the EU between-contribution, indicating that employment is much more important in terms of income in Lithuania than in the EU. Low redistribution (low taxes and transfers) in Lithuania could explain why it is very costly to not participate in the labour market (IMF 2016; Lazutka 2017). Furthermore, the number of those employed within a household matter in Lithuania. Demographic characteristics (age, number of dependents, sex) determine a relatively lower share (0.2-1.4 of Gini).

The within, between and stratification decomposition is decomposed further to reveal the importance of the employed to income inequality each year from 2005 to 2015. Specifically, the within-contribution of activity status is decomposed to the within contribution of the employed, unemployed, and non-participants. This decomposition, along with the between and stratification contributions, is shown in Table 1.10 for Lithuania. The rise in disposable income household inequality in Lithuania since 2011 can be primarily explained by a rise in income inequality among those who are employed. This is partly determined by the fact that a larger share of the population has become employed since the crisis (51% in 2011 and 55% in 2015), the employed are taking a larger share of income (from 62% to 68%) and are themselves more unequally distributed (the within-Gini rose from 29 to 33). To a lesser extent, inequality is also rising due to greater between-subgroup inequality and stratification, especially stratification of the employed vis-a-vis other groups. This is because average wages rose faster than non-labour income during this period.

Table 1.10: Decomposition of the first differences of the Gini coefficient of equivalised disposable income in Lithuania in 2015

year	employed	unemployed	other	between	stratification	sum
2005	-0.80	-0.30	-0.20	-0.30	0.00	-1.60
2006	-0.20	-0.50	-0.40	0.20	-0.10	-1.00
2007	0.70	0.00	0.40	-0.50	0.20	0.80
2008	0.40	0.30	0.70	-0.20	0.20	1.40
2009	0.20	0.70	0.00	0.00	0.00	0.90
2010	-4.10	0.60	-0.40	0.00	0.20	-3.70
2011	-0.50	-0.70	-0.20	0.60	-0.40	-1.20
2012	2.30	-0.10	0.60	-0.10	-0.10	2.60
2013	0.40	-0.20	0.00	0.20	0.00	0.40
2014	2.80	0.10	-0.20	0.50	-0.30	2.90
2015	-1.00	0.10	-0.30	0.40	-0.20	-1.00

The sum indicates the first differences of the Gini coefficient, while other columns show the contribution. In 2014, the Gini rose by 2.9 Gini points and 2.8 Gini points are explained by the change in the contribution of the employed.

1.4. Structure of income inequality by income factors

We estimate the structure of income inequality by decomposing household disposable income inequality by factors. Knowing which factors contribute to income inequality help explain why income inequality in Lithuania is high. The four components of disposable income are labour income, capital income, transfers, and taxes (including social transfers). These are further broken down by more granular income factors.

We use the Lerman and Yitzhaki (1985a) method to decompose the Gini coefficient. It allows decomposing $\widehat{G}$ into income factors $\sum_{k=1}^{K} \widehat{T}_k$, where k represents labour, capital, transfers and taxes. We further decompose $\widehat{T}_k$ into $(\widehat{R}_k/100)(\widehat{G}_k/100)\widehat{S}_k$. Here $\widehat{R}_k$ is the estimate of Gini correlation between household disposable income and factor k. The quantity $\widehat{R}_k$ ranges between -100 and 100. The value $\widehat{R}_k = 100$ refers to high positive correlation. This means that households with a lot of factor k also have a lot of total disposable income, while households with little factor k have small disposable income. If $\widehat{R}_k$ is close to -100, it means that households with little disposable income tend to have larger factor k income. Next, $\widehat{G}_k$ represents the Gini index of factor k and is approaching 100 if inequality of k is high. Finally, component $\widehat{S}_k$ is the share of factor k of the household disposable income, meaning that factors which constitute a larger share of income matter more for inequality. More details on this method are provided in Appendix .2. We provide the estimates for Lithuania and the EU. Unfortunately, 4 countries, including Germany,

did not provide all the necessary income factors, meaning that the data sample for the EU differs from the previous analysis.

Table 1.11 reveals the results for the decomposition of disposable income into $\widehat{T}_k$ for Lithuania and the EU by factors and the further decomposition into $\widehat{R}_k\widehat{G}_k\widehat{S}_k$ is available in Table 1.12.

Table 1.11: Factor decomposition of the Gini coefficient in 2015 by labour, capital, transfers, taxes and their subfactors

Variable	EU	new EU states	old EU states	LT
Gini	30.86 (0.30)	30.55 (0.44)	30.88 (0.38)	36.96 (0.76)
Labour	43.91 (0.63)	42.51 (0.78)	44.25 (0.8)	53.63 (1.28)
employment	30.95 (0.52)	33.90 (0.78)	30.06 (0.65)	34.48 (1.18)
employer's social insurance contribution	7.52 (0.17)	5.47 (0.16)	8.03 (0.20)	9.67 (0.38)
self-employment	5.27 (0.36)	3.06 (0.33)	5.97 (0.45)	9.29 (0.86)
company car	0.16 (0.02)	0.09 (0.01)	0.18 (0.03)	0.19 (0.05)
income received by people aged under 16	0.00 (0.00)	-0.01 (0.01)	0.01 (0.01)	0.00 (0.00)
Capital	2.76 (0.26)	0.96 (0.13)	3.29 (0.34)	1.32 (0.29)
interests, dividends, etc.	1.63 (0.23)	0.48 (0.09)	1.96 (0.29)	1.11 (0.28)
rental income	1.13 (0.09)	0.49 (0.08)	1.32 (0.11)	0.20 (0.05)
Transfers	4.66 (0.27)	2.56 (0.30)	5.24 (0.34)	-0.25 (0.30)
old-age benefits	5.07 (0.26)	3.28 (0.28)	5.62 (0.33)	-0.44 (0.25)
unemployment	0.19 (0.07)	0.02 (0.03)	0.20 (0.09)	0.04 (0.06)
survivor benefits	0.28 (0.06)	-0.04 (0.07)	0.37 (0.07)	-0.12 (0.03)
sickness benefits	0.06 (0.01)	0.10 (0.02)	0.05 (0.02)	0.47 (0.05)
education-related allowances	-0.04 (0.01)	-0.01 (0.02)	-0.05 (0.01)	0.00 (0.01)
family/children related allowances	-0.37 (0.04)	-0.18 (0.07)	-0.43 (0.05)	0.41 (0.15)
disability benefits	-0.11 (0.03)	-0.26 (0.07)	-0.06 (0.04)	0.03 (0.11)
social exclusion	-0.28 (0.02)	-0.18 (0.02)	-0.32 (0.03)	-0.41 (0.05)
housing allowances	-0.23 (0.02)	-0.05 (0.01)	-0.3 (0.02)	-0.03 (0.00)
regular inter-household cash transfers received	0.00 (0.06)	0.02 (0.04)	-0.01 (0.08)	-0.12 (0.05)
regular inter-household cash transfers paid	-0.14 (0.03)	-0.20 (0.08)	-0.12 (0.02)	-0.09 (0.04)
individual private pension	0.24 (0.04)	0.06 (0.01)	0.29 (0.05)	0.02 (0.01)
Taxes	-20.46 (0.31)	-15.49 (0.32)	-21.89 (0.39)	-17.74 (0.67)
tax on income and social insurance contributions	-12.79 (0.20)	-9.98 (0.25)	-13.69 (0.24)	-8.03 (0.30)
employer's social insurance contribution	-7.52 (0.17)	-5.47 (0.16)	-8.03 (0.20)	-9.67 (0.38)
regular taxes on wealth	-0.15 (0.01)	-0.04 (0.01)	-0.17 (0.02)	-0.04 (0.01)

Bootstraped standard errors are provided in the parentheses.

- Labour income contributes most to income inequality in Lithuania. It contributes 53.63 Gini points to total inequality. Labour income contributes most to income inequality on the EU level as well, yet about 9.72 Gini points less than in Lithuania. The labour component is especially large as it includes an employer's social insurance contributions. Capital contributes only 1.32 and transfers and taxes reduce income inequality by 0.25 and 17.74 points respectively.

- All labour sub-factors contributions are larger in Lithuania than in new and old EU states. The largest sub-factor contribution is employee income in Lithuania

(34.48 Gini points). The contribution is about 0.58 Gini points higher than in the new EU states and 4.42 higher than in the old EU states. Self-employed contribute less to inequality in Lithuania (9.29 Gini points). However, this is by 6.23 Gini points more than in new EU states and by 3.32 Gini points more than in the old EU states.

- Labour income has a greater contribution in Lithuania than in the EU largely because this income is more correlated with disposable income in Lithuania. In other words, those who get a lot of labour income tend to be the richest households in terms of disposable income also. This is seen from $\widehat{R}$, the value of which is equal to 90.61 in Lithuania, while it is under 79.58 in new EU states and 74.38 in old EU states. This is especially true for the self-employed: $\widehat{R}$ is equal to 70.11 and this is 25.46 points more compared to the EU. High $\widehat{R}$ means that self-employment income is especially important for self-employed households. This may give rise to concern, as such income is generally less stable than employment income. In contrast, for the labour income, $\widehat{G}$ in Lithuania is similar to $\widehat{G}$ observed in other EU countries while $\widehat{S}$ is only slightly larger.

- Taxes (and social contributions) negatively contribute to income inequality in Lithuania. Specifically, taxes reduce income inequality by 17.74 Gini points. This reduction is a couple of percentage points less than the EU and the old EU states in particular. The biggest difference is a lower $\widehat{S}$, which means that taxes constitute a smaller share of disposable income in Lithuanian than in the EU.

- Transfers seem to not contribute to income inequality in Lithuania. Specifically, transfers contribute -0.25 Gini points. At first this may seem surprising, as transfers are known to be of much greater effect in reducing income inequality (see, eg., Joumard et al. (2013)). However, it would be more correct to say that transfers do not contribute to inequality - i.e. they are not a part of the structure of inequality, instead of saying that transfers do not affect inequality. On the contrary, transfers can have a large effect. Upon closer inspection, we see the low contribution is due to a low $\widehat{R}$ which equals -1.84 for Lithuania. Upon multiplying $\widehat{R}$ by $\widehat{G}$ and $\widehat{S}$, the inequality contribution is close to zero. Therefore, the larger the $\widehat{S}$ going to transfers, the lower the inequality. Since transfers do not contribute to inequality and taxes reduce inequality, their relative effect on inequality is not comparable using this method. This leads us to Section 1.5 which discusses their relative effects.

Table 1.12: Factor decomposition of the of Gini of disposable income in 2015

Variable	Contribution	EU	new EU states	old EU states	LT
Labour	T	43.91 (0.63)	42.51 (0.78)	44.25 (0.8)	53.63 (1.28)
Labour	R	75.55 (0.59)	79.58 (0.63)	74.38 (0.75)	90.61 (0.48)
Labour	G	52.7 (0.35)	49.26 (0.53)	53.66 (0.43)	52.22 (0.91)
Labour	S	110.28 (0.67)	108.47 (0.74)	110.87 (0.86)	113.35 (1.13)
employee	T	30.95 (0.52)	33.9 (0.78)	30.06 (0.65)	34.48 (1.18)
employee	R	70.43 (0.66)	75.68 (0.8)	68.87 (0.84)	81.44 (1.13)
employee	G	56.49 (0.37)	53.9 (0.59)	57.37 (0.46)	55.16 (0.93)
employee	S	77.8 (0.6)	83.12 (0.82)	76.07 (0.73)	76.77 (1.24)
self-employment	T	5.27 (0.36)	3.06 (0.33)	5.97 (0.45)	9.29 (0.86)
self-employment	R	44.65 (1.84)	30.06 (2.51)	48.5 (2.14)	70.11 (2.44)
self-employment	G	92.06 (0.41)	89.43 (0.39)	92.73 (0.51)	91.13 (0.58)
self-employment	S	12.82 (0.43)	11.39 (0.46)	13.28 (0.54)	14.53 (1.01)
Capital	T	2.76 (0.26)	0.96 (0.13)	3.29 (0.34)	1.32 (0.29)
Capital	R	67.94 (2.14)	68.24 (3.15)	67.87 (2.33)	76.41 (5.08)
Capital	G	92.89 (0.33)	98.06 (0.14)	90.92 (0.44)	98.13 (0.29)
Capital	S	4.37 (0.28)	1.44 (0.13)	5.32 (0.36)	1.75 (0.31)
Transfer	T	4.66 (0.27)	2.56 (0.3)	5.24 (0.34)	-0.25 (0.3)
Transfer	R	21.22 (1.07)	13.35 (1.45)	22.98 (1.28)	-1.84 (2.23)
Transfer	G	66.88 (0.35)	64.69 (0.74)	67.27 (0.41)	57.26 (0.86)
Transfer	S	32.84 (0.4)	29.65 (0.52)	33.87 (0.5)	23.47 (0.66)
Tax	T	-20.46 (0.31)	-15.49 (0.32)	-21.89 (0.39)	-17.74 (0.67)
Tax	R	80.78 (0.47)	78.28 (0.65)	81.43 (0.56)	81.03 (1.14)
Tax	G	53.35 (0.37)	50.02 (0.56)	53.7 (0.45)	56.75 (0.94)
Tax	S	-47.49 (0.35)	-39.56 (0.3)	-50.07 (0.42)	-38.57 (0.71)

G is decomposed into income factors $\sum_{k=1}^{4} T_k$, where k represents labour, capital, transfers and taxes. We further decompose T_k into $(R_k/100)(G_k/100)S_k$. Here R_k is the Gini correlation between household disposable income and factor k that ranges between -100 and 100. The component G_k represents the Gini index of factor k and S_k is the share of factor k of the household disposable income. Bootstrapped standard errors are provided in the parentheses.

1.5. Marginal and redistribute effect of taxes and transfers on income inequality in Lithuania

In this section, we answer how much do transfers and taxes affect income inequality. We do so first by calculating the marginal effects: how does inequality respond to a percent change in an increase in taxes or transfers. Second, we estimate the redistributive effect of taxes and public transfers. Specifically, we analyse two ways in which taxes and public transfers can affect income inequality: by increasing their progressivity and their rate.

We use the Lerman and Yitzhaki (1985a) decomposition to shed light on the marginal contribution of each income factor to the Gini coefficient. We calculate the amount by which the Gini changes if we raise the factor contribution by a small value e_k and hold other income factors constant. This is approximately equal to evaluating

how many Gini points will the Gini coefficient change if we increase an income factor by 1%. The formula (10) in Appendix .2 quantifies the effects. If all income factors are raised by the same $e_k = e$, the Gini would not change, as summarised in the first row of Table 1.13.

Table 1.13: Marginal decomposition of the Gini coefficient in 2015 by labour, capital, transfers, taxes and their subfactors

Variable	EU	new EU states	old EU states	LT
Gini	0.0000 (0.0000)	0.0000 (0.0000)	0.0000 (0.0000)	0.0000 (0.0000)
Labour	0.0987 (0.0038)	0.0938 (0.0038)	0.1001 (0.0049)	0.1174 (0.0062)
employment	0.0694 (0.0035)	0.0851 (0.0042)	0.0657 (0.0043)	0.0611 (0.0072)
employer's social insurance contribution	0.0155 (0.0011)	0.0126 (0.0011)	0.0149 (0.0013)	0.0161 (0.0023)
self-employment	0.0131 (0.0025)	-0.0042 (0.0026)	0.0187 (0.0032)	0.0391 (0.005)
company car	0.0008 (0.0001)	0.0005 (0.0001)	0.0008 (0.0002)	0.001 (0.0003)
income received by people aged under 16	-0.0001 (0.0000)	-0.0002 (0.0001)	-0.0001 (0.0001)	0.0000 (0.0000)
Capital	0.0141 (0.0017)	0.0052 (0.0009)	0.0164 (0.0022)	0.0067 (0.0018)
interests, dividends, etc.	0.0087 (0.0015)	0.0028 (0.0006)	0.0102 (0.0019)	0.0063 (0.0018)
rental income	0.0054 (0.0006)	0.0024 (0.0006)	0.0062 (0.0007)	0.0004 (0.0003)
Transfers	-0.0547 (0.0022)	-0.0650 (0.0025)	-0.0522 (0.0028)	-0.0892 (0.0029)
old-age benefits	-0.0164 (0.002)	-0.0311 (0.0022)	-0.0118 (0.0026)	-0.0544 (0.0023)
unemployment	-0.0053 (0.0006)	-0.0020 (0.0003)	-0.0069 (0.0008)	-0.0031 (0.0005)
survivor benefits	-0.0031 (0.0005)	-0.0053 (0.0006)	-0.0025 (0.0006)	-0.0041 (0.0004)
sickness benefits	-0.0003 (0.0001)	-0.0001 (0.0001)	-0.0004 (0.0001)	-0.0007 (0.0003)
education-related allowances	-0.0012 (0.0001)	-0.0006 (0.0002)	-0.0014 (0.0002)	-0.0003 (0.0001)
family/children related allowances	-0.0114 (0.0005)	-0.0104 (0.0008)	-0.0116 (0.0006)	-0.0054 (0.0011)
disability benefits	-0.0072 (0.0004)	-0.0101 (0.0009)	-0.0062 (0.0004)	-0.0112 (0.0014)
social exclusion	-0.0047 (0.0003)	-0.0029 (0.0003)	-0.0053 (0.0004)	-0.0066 (0.0008)
housing allowances	-0.0042 (0.0002)	-0.0008 (0.0001)	-0.0054 (0.0003)	-0.0004 (0.0001)
regular inter-household cash transfers received	-0.0030 (0.0004)	-0.0029 (0.0004)	-0.0030 (0.0006)	-0.0037 (0.0008)
regular inter-household cash transfers paid	0.0010 (0.0003)	0.0007 (0.0008)	0.0011 (0.0003)	0.0007 (0.0004)
individual private pension	0.001 (0.0003)	0.0003 (0.0001)	0.0012 (0.0003)	0.0001 (0.0001)
Taxes	-0.0581 (0.0017)	-0.0340 (0.0018)	-0.0643 (0.0021)	-0.0348 (0.0039)
tax on income and social insurance contributions	-0.0440 (0.0011)	-0.0224 (0.0015)	-0.0511 (0.0014)	-0.0189 (0.0017)
employer's social insurance contribution	-0.0155 (0.0011)	-0.0126 (0.0011)	-0.0149 (0.0013)	-0.0161 (0.0023)
regular taxes on wealth	0.0014 (0.0001)	0.0009 (0.0001)	0.0017 (0.0002)	0.0002 (0.0001)

Bootstraped standard errors are provided in the parentheses.

Table 1.13 shows the marginal contributions to the Gini for Lithuania and the EU. Several conclusions can be drawn on taxes and transfers as well as labour and capital income.

- Transfers and taxes reduce income inequality. Raising transfers by 1% reduces inequality by 0.0892 Gini points while raising taxes (including social contributions) reduces income inequality by 0.0348 Gini points. Additionally, raising transfers has a larger effect in Lithuania than in the EU. Increasing old-age transfers alone would reduce inequality by 0.0544 Gini points–three times more than in the EU. Other transfers have a much smaller impact individually. Taxes, however, have less effect in Lithuania than in the EU, especially the old EU states. Specifically, a 1% rise in income taxes and social contributions paid by the household reduces inequality by 0.0348 Gini points–about half of

the impact in the old EU states, which is 0.0643. However, the tax situation in Lithuania is very similar to that of new EU states.

- Raising labour income would result in higher inequality in Lithuania and the effect is stronger for Lithuania than for the EU. A 1% increase in labour income means a 0.1147 rise in income inequality in Lithuania. This is almost 0.02 Gini points more than in the EU. The reason why inequality would rise more in Lithuania than in the EU is self-employment income. A 1% rise in self-employed income raises income inequality by 0.0391 Gini points in Lithuania as compared to 0.0131 Gini points in the EU. Raising employment income would raise income inequality by similar amounts in both economies.

The reasons why raising old-age benefits reduces inequality in Lithuania more than in the EU are most likely related to the design of the pension systems in Lithuania and the EU. First, the social expenditure on pensions in Lithuania is lower than in the EU (Lis 2018). Because of this, the retired have lower incomes as compared to the rest of the population and this difference is larger than for the EU (see Table 1.3). This means that any transfers to this group will on average reduce inequality more in Lithuania. Second, the old-age benefits that are handed out in Lithuania depend on previous contributions but are not very elastic to it. This means that the old-age benefits are relatively equally distributed amongst the retired and perhaps more so than in other countries. As a consequence, the retired are relatively more equal amongst themselves (see Table 1.5) as compared to inequality within other activity status groups. Therefore, increasing the income share of the pensioners, the most equal subgroup in society, will reduce overall income inequality also. However, whether the pensions in other EU countries are more or less elastic to previous contributions than Lithuania remains to be tested.

Similarly, the reasons why raising tax income would reduce income inequality in Lithuania less than in the EU is likely related to the design of the respective tax and social contribution systems. Lithuania's social contribution constitutes over $3/4$ labour taxes. But they are not progressive. The social contribution rates are flat without a ceiling and are therefore not redistributive among those who pay the contributions. Income tax constitutes just a quarter of labour taxes and, apart from a non-taxable minimum, has been non-progressive in 2005-2015 either. This means that while raising taxes will bring those with labour income closer to those without labour income, it will not reduce income inequality amongst those who have labour income.

The reason why raising labour income results in more inequality in Lithuania than in the EU may also be related to the tax system and tax evasion. In Lithuania, the

self-employed benefited from a lower taxable base. Additionally, the self-employed seem to evade taxes more often than employed in Lithuania (Černiauskas and Jousten 2020). As a result, there is very little redistribution for the self-employed taking place in Lithuania. Given that self-employment income is effectively not taxed, it correlates so well with disposable income and the Gini correlation coefficient $\widehat{R}$ was so high in Table 1.12.

Next, we estimate the redistributive effect of taxes and public transfers for the total population and self-employed separately. We follow Joumard et al. (2013), which is based on Kakwani (1977). This method also lets us decompose the redistribution effect into the progressivity and average rate of taxes or public transfers in Lithuania and compare these figures with the ones in the EU.

For i denoting taxes or transfers, the redistributive effect is decomposed as follows (Joumard et al. 2013):

$$redistribution_i = \frac{r_i}{100} progressivity_i, \tag{1.2}$$

where r_i represents the percent of taxes or public transfers in income and $progressivity_i$ takes the values from -100 to 100, where -100 indicates regressive i and 100 indicates progressive i.

Specifically, we apply the following calculations to get the average rate r_i and the progressivity index. To compute r_{tax}, we divide the total taxes paid by the disposable income of the population and multiply by 100. To compute $r_{transfers}$, we divide the public transfers received by the market income after transfers of the population and multiply by 100. To compute $progressivity_{tax}$, we subtract the concentration coefficient of market income after public transfers from the concentration coefficient of taxes. To compute the $progressivity_{transfers}$, we subtract the concentration coefficient of public transfers from the concentration coefficient of market income. The concentration coefficient is familiar to the Gini index. Like the Gini index, it is computed using (5), where y represents the variables tax or $transfers$. However, tax, $transfers$, and survey weights are sorted according to market income. It is also possible to sort by disposable income. In that case, the progressivity measures would be much smaller. However, we prefer sorting by market income, because we see the Lithuanian and EU system as transferring to and taxing from households primarily based on their market incomes.

The redistributive effects of taxes with social security contributions are similar to the redistributive effects of public transfers for Lithuania. The effects on the Gini of market income, as well as the components of the effects, are available in Table 1.14 for Lithuania and the EU in 2015. Both taxes and public transfers have a very simi-

lar effect on redistributing incomes. Interestingly, taxes excluding employer's social insurance contributions contribute much less to income redistribution in Lithuania and the EU. Since other studies typically disregard employer's social contributions, it could explain why they find taxes to be playing a small role in redistribution (see, e.g., OECD 2011a; Causa and Hermansen 2017).

Taxes have a high redistributive effect because of the average tax rate, while public transfers have a high effect because of their progressiveness in Lithuania. The average tax rate constitutes 38.6% of disposable income which is more than double the public transfer rates (16.7% of market income after transfers). However, taxes are much less progressive (31.4%) as compared to public transfers (78.7%). This means that raising tax progressivity will have a higher impact on reducing income inequality than raising public transfer progressivity, while raising the average public transfer rate will have a higher effect on income inequality than raising the average tax rate in Lithuania and, similarly, in the EU.

The redistributive effects of public transfers and taxes are much lower in Lithuania than in the EU. The redistributive impact of taxes in Lithuania is almost two times smaller than in the EU, while public transfers are about 50% smaller. All the subcomponents are smaller. Tax progressivity and the average rate of public transfers in particular are lower in Lithuania as compared to the EU.

Table 1.14: Progressivity index for market incomes in 2015

	EU	new EU states	old EU states	LT
Redistributive effect of public transfers	19.8	18.5	20.2	13.2
Redistributive effect of taxes	23.9	16.3	26.8	12.1
Redistributive effect of taxes without ESC[1]	8.9	7.1	9.6	3.8
Average tax rate	47.8	41.2	50.1	38.6
Average public transfer rate	21.9	21.3	22.1	16.7
Tax progressivity index	50.0	39.6	53.5	31.4
Public transfers progressivity index	90.6	86.9	91.7	78.7

The redistributive effects of public transfers and taxes are calculated by multiplying their progressivity index with the average rates as in (1.2). To compute the average tax rate, we divide the taxes paid by the disposable income of the population. To compute the average public transfer rate we divide the public transfers received by the market income after public transfers of the population. To compute the progressivity of taxes, we subtract the concentration coefficient of market income after public transfers from the concentration coefficient of taxes. To compute the progressivity of public transfers, we subtract the concentration coefficient of public transfers from the concentration coefficient of market income. Tax progressivity is measured using the Kakwani index, where 100 is a very progressive Tax system and -100 is a very regressive tax system. The same is applied to transfers.

[1]ESC - employer's social insurance contributions.

The tax system is much less distributive amongst the self-employed in Lithuania. The redistributive effect of taxes is negative in Lithuania as shown in Table 1.15. This means that the poorer households pay a larger share of their disposable income in taxes than the richer households. This is in line with previous findings (Černiauskas and Jousten 2020). We additionally see that this is very different when compared to the EU, wherein taxes do have a positive redistributive effect. Additionally, the average tax rate of the self-employed for Lithuania is less than a third of the EU and almost a quarter of the tax rates of the old EU states. Therefore, negative tax progressivity can explain why the self-employed contribute more to inequality in Lithuania than in other EU states.

Table 1.15: Progressivity index for market incomes in 2015 for self-employed

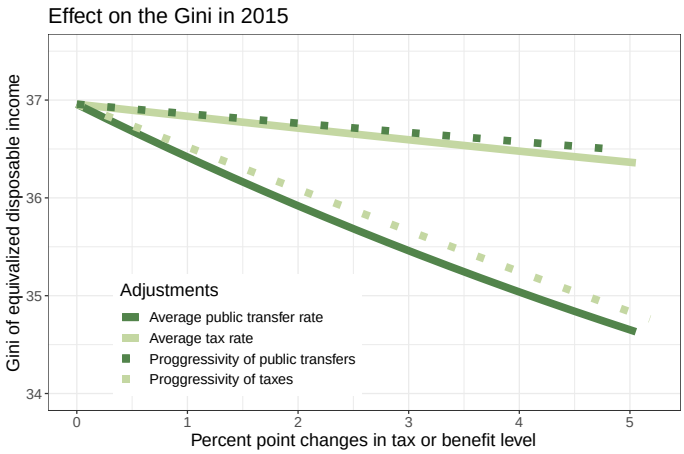
	EU	new EU states	old EU states	LT
Redistributive effect of transfers	1.3	1.9	1.1	2.0
Redistributive effect of taxes	4.9	1.7	5.6	-1.3
Redistributive effect of taxes without ESC	7.5	4.0	8.1	-0.5
Tax progressivity index	12.4	6.4	12.8	-11.3
Transfers progressivity index	19.1	20.5	17.1	25.8
Average tax rate	39.0	26.1	43.9	11.1
Average transfer rate	7.0	9.2	6.2	7.8

The redistributive effects of public transfers and taxes are calculated by multiplying their progressivity index with the average rates as in (1.2). To compute the average tax rate, we divide the taxes paid by the disposable income of the population. To compute the average public transfer rate we divide the public transfers received by the market income after public transfers of the population. To compute the progressivity of taxes, we subtract the concentration coefficient of market income after public transfers from the concentration coefficient of taxes. To compute the progressivity of public transfers, we subtract the concentration coefficient of public transfers from the concentration coefficient of market income. Tax progressivity is measured using the Kakwani index, where 100 is a very progressive Tax system and -100 is a very regressive tax system. The same is applied to transfers.
[1]ESC - employer's social insurance contributions.

The results suggest that raising tax progressivity and the average rate of public transfers should reduce income inequality most. We run a simulation (for the full population) to observe this. We simulate the effect of increasing the average rate and changing the progressivity of taxes and public transfers on Lithuania using EU-SILC data. The effect of changing the progressivity or average rate of tax and public transfers on the Gini of Lithuania is illustrated in Figure 1.2. We simulate the average rate of taxes by increasing the taxes for all those who are currently paying taxes. We do a similar simulation for public transfers. We increase taxes and transfers by up to 5 percentage points of market income after public transfers. We increase the

progressivity of taxes by increasing taxes by up to 5 percentage points for the top quintile of households that are sorted by market incomes and redistributing this gain to all other quintiles. The redistribution is also progressive. For example, if we were to increase taxes on the top quintile by 10%, then the 4th quintile will get to pay about 10% fewer taxes, the third will pay 20% less, the second 30% less and the first will pay 40% less. A scalar is added so that the reduction in taxes for the four bottom quintiles equals the increase in taxes for the top quantile. We increase the progressivity of public transfers by increasing transfers received by up to 5 percentage points for the bottom quintile of households that are sorted by market incomes and redistributing the cost to all other quintiles in a similar manner as for taxes. The simulations confirm that increasing the average rate of public transfers has a much higher effect on the Gini than raising taxes by the same amount. Increasing tax progressivity has a larger effect than increasing public transfer progressivity.

Figure 1.2: Simulating the effect of changes in progressivity and average rate of tax and public transfers on the Gini coefficient of equivalised disposable income in Lithuania



1.6. Conclusions

We have tackled three questions and each of them is elaborated in this study. We have also suggested possible improvements for future studies.

First, we have run three statistical tests and found that equivalised income inequality in Lithuania is in all cases one of the highest in the EU. Specifically, we

have tested for accuracy of estimates by estimating their standard errors, the inequality measure used as well as different equivalence scales. In all cases, equivalised income inequality in Lithuania is found to be one of the highest across the EU.

Second, we have investigated why equivalised income inequality in Lithuania is higher compared to the EU by using univariate decomposition techniques. We have found large inequalities between and within many groups of households in the country. In all cases, the within-group inequality contributes more to equivalised income inequality in Lithuania and the EU. It means that this inequality is higher within households of similar observable characteristics rather than between households of different characteristics. Inequalities within the unemployed and those working in the agricultural sector are especially prominent. Nevertheless, between-contributions are also significant for Lithuania, suggesting where policy can look into deeper. The largest between-group inequalities lie between the employed and the rest of the population. Moreover, this type of inequality has been rising over time. As the factor decomposition shows, the large between-group inequality contribution can be explained by unequal distribution of labour income, especially–self-employment income.

Third, we analysed the extent to which equivalised income inequalities stemming from the market income are offset by taxes and transfers. Specifically, we analysed the marginal and redistributive effects of Lithuania's taxes and transfers and compared this to the EU. The marginal decomposition of the Gini coefficient of equivalised disposable income by factors confirms that an increase in tax and transfer income reduces equivalised income inequality while an increase in labour income increases it. The way that the tax and transfer system is currently designed, the average marginal contribution is more than twice higher for transfers compared to taxes, and that among the transfers the role of the old-age pensions is the highest. Similarly, the analysis of the redistributive effect of the taxes and public transfer income also showed that these two income sources reduce income inequality. However, the redistributive impact of taxes in Lithuania is almost two times smaller than in the EU, while public transfers are about 50% smaller. The redistributive effect of taxes for the self-employed is negative in Lithuania and therefore reinforces income inequality, while taxes reduce inequality amongst the self-employed in the EU. This means that the current tax system and tax evasion/avoidance of higher-income households are likely to be responsible for a larger self-employment income contribution to inequality in Lithuania as opposed to EU.

We also decomposed the redistributive effect into the progressivity and the average rate of tax and public transfers effect. We find that the tax progressivity and the average rate of public transfers in particular are lower in Lithuania as compared to the

EU. The results suggest that raising tax progressivity and the average rate of public transfers would reduce equivalised income inequality most.

The estimates of equivalised income inequality may have several drawbacks. First, there is a large shadow economy in Lithuania, with some estimates exceeding 25% of GDP in 2013 and 2015 (see Schneider 2013; Žukauskas 2016). Even though survey respondents are informed that their data will not be used for tax purposes, some of them may still be unwilling to disclose information on their true income received. It remains unclear how this affects equivalised income inequality because it depends on the income distribution within the shadow economy together with the income distribution of the observed economy. Additionally, this estimate may cause problems when comparing households across countries, since the size of the shadow economy is particularly large in Lithuania. Second, as has been already pointed out various times, EU-SILC undersamples the income of rich individuals in all countries (especially capital income (Navickė and Lazutka 2017))–something that the survey weights do not correct for. Including the rich will result in higher measures of equivalised income inequality in Lithuania. However, equivalised income inequality will rise in other EU countries as well. Therefore, the relative position of Lithuania vis-a-vis other countries may not change so much. Nevertheless, the alternative Household Finance and Consumption Survey (HFCN 2019) could partly correct for both of these shortcomings, as it has data on consumption, which can be used to estimate the shadow economy and oversample the wealthy households for Lithuania along with many other EU countries. Furthermore, greater access to administrative data would be yet another path to take.

Future studies can also consider using an alternative methodology, for example, by using multivariate techniques to decompose equivalised income inequality. This was not the focus of the current study because the results of a multivariate decomposition depend on all variables by which the Gini is decomposed, and there is no consensus on which should be included. Furthermore, variables available to some countries are less available in others in the EU-SILC. Nevertheless, our additional check using a multivariate decomposition technique as in Social Situation Monitor (2017) does not contradict the results. Additionally, one may look into income inequality between individuals instead of households.

2. INCOME INEQUALITY AND REDISTRIBUTION IN LITHUANIA: THE ROLE OF POLICY, LABOUR MARKET, INCOME AND DEMOGRAPHICS

2.1. Introduction

Economic inequality has been rising since the 1980s in most advanced economies, as well as in post-Soviet countries and other emerging markets (Anthony B. Atkinson et al. 2011; OECD 2011b; Nolan et al. 2014; Alvaredo et al. 2018). Concerns about inequality have surged in the aftermath of the Great Recession, fuelled also by the rise in economic distress caused by the unequal distribution of gains stemming from globalization and economic growth. Rising inequalities in market incomes, changes in taxes and benefits, changes in the structure of the labour market (e.g. increasing female labour market participation or occupational structure dynamics), and changes in demographics (e.g. expansion of post-secondary education, spread of non-traditional family structures) are highlighted among the main determinants of the increase in income inequality in most OECD countries since the 1980s (e.g. Daly and Valletta 2006; OECD 2011b; Smeeding et al. 2011).

The role of tax-benefits systems in tackling inequality increases has been extensively studied, as disposable income is a product of both market incomes and tax-benefit rules. Much less research has examined why redistribution did not manage to tame the increase in inequality. The evidence is at odds with conclusions reached by the majority of studies that tax-benefit systems have become more redistributive since the 1980s (e.g. Immervoll and Richardson 2011). This is due to a methodological limitation that did not control for interactions between market incomes and tax-benefit rules. Failing to control for changes in market income distributions may lead one to wrongly conclude that redistribution has increased, when in fact the effect has been driven by increasing market income inequality; any progressive system will show an increase in redistribution with increasing inequalities in market incomes. The literature shows that inequality in market income grew twice as much as redistribution. This implies that the redistributive effect has weakened over time in most countries, which is consistent with redistributive policies' failure to tackle inequality increases (Immervoll and Richardson 2011; Alvaredo et al. 2018).

This question becomes even more important for countries where the increase in inequality was striking, especially in the recovery period following the Great Recession. The post-Soviet countries stand out in the European context with respect to their dramatic changes in the distribution of disposable income over time. Despite

this, they have received little attention in the inequality literature. We contribute to this literature with a systematic analysis that seeks to understand the trends in income inequality and the redistributive effects of the tax-benefit system in Lithuania by disentangling the role played by changes in policy design from changes in market income distributions (and their driving forces: labour market structure, returns, and demographics).

Since regaining independence from the Soviet Union in 1990, Lithuania has implemented numerous liberal reforms, which allowed the country to move rapidly from a centrally planned to a market economy. After joining the European Union (EU) along with the other Baltic states in 2004, Lithuania enjoyed high growth rates and economic convergence towards EU-15. The period of rapid economic expansion came to a halt in 2008, when the country was hit by a deep recession due to the Global Financial Crisis and the real GDP plummeted by almost 15% in 2009 as compared to 2008. A rapid recovery followed, with a GDP growth of 6% in 2011. Since then, the growth has stabilized but income inequality has shot up as well, despite numerous changes in the tax and benefit system. According to Eurostat, the Gini index of household equivalized disposable income in Lithuania grew by 5 points over the period 2011-2015, the highest growth rate of income inequality observed in the European Union (EU) (which saw an average increase in the Gini index of only 0.2 points over the same period).[1] As a result, as measured by the Gini index, Lithuanian income distribution was the second most unequal in the EU in 2015. While unequal economic growth could explain this rising inequality, there are also other possible explanations. The Lithuanian economy was affected by important secular demographic changes, namely, negative net migration, ageing, and declining marriage rates. The goal of this chapter is to quantify what factors drove large changes in Lithuanian income distributions over the period 2007-2015, which is a central issue for economic research and policy analysis.

In order to answer this question, we apply the latest methodological advancements in inequality decomposition techniques, which rely on counterfactual scenarios to isolate the impact of relevant factors. We build on the approach developed in Bourguignon et al. (2008) and Sologon et al. (2020), and adapt it to study changes in income distributions over time instead of differences in income distributions across countries at one given moment.[2] Traditional approaches compute one particular in-

1. Eurostat reports the Gini index based on the year the survey was conducted. By contrast, survey respondents are asked to provide their previous calendar year's income. Throughout the text, we report statistics of the income year, not the survey year.

2. Sologon et al. (2019) use the same approach to study changes in the income distribution in Portugal between 2007 and 2013, accounting for the distributional effects of the 2007-2008 crisis and the

equality summary index over time, and then decompose it into the contribution of specific characteristics, such as age, gender, labour market status or the source of income (see Reynolds and Smolensky (1977), A. Shorrocks (1980), A. Shorrocks (1982), Anthony F Shorrocks (1984) and Lerman and Yitzhaki (1985b)). Rather than looking at summary measures, the main object of our analysis are changes in the whole income distribution. Our method integrates micro-econometric and microsimulation approaches into a flexible parametric household income-generation process based on a system of equations for multiple income sources for the household and the European tax-benefit micro-simulation engine EUROMOD (Sutherland and Figari 2013). Such an infrastructure permits an accurate representation of the relationship between household characteristics, market incomes (from labour and capital), and tax-benefit rules. This is used to generate counterfactual distributions of household disposable incomes obtained via transformations of the income generation process, by "swapping" the characteristics between different periods along four dimensions: (i) labour market structure (e.g. employment, occupation, industry, sector), (ii) returns structure (e.g. labour income, capital incomes), (iii) demographic composition of the population, and (iv) tax-benefit rules. The comparison of these counterfactual distributions allows us to quantify the contribution of each factor to the changes in the income distribution observed over time.

By applying this approach, we provide a more detailed decomposition than existing studies that seek to unpack the drivers of inequality changes. Most research on the topic follows the approach proposed by Bargain and Callan (2010) and Bargain (2012), which uses two "swaps": market incomes and tax-benefit rules. For Lithuania, Navickė (2020), besides the policy and income effect, also added the demographic effect via re-weighting following DiNardo et al. (1996) to decompose the changes in the Gini index. The findings suggest that while the income effect dominated the increase in the Gini index, the rising income inequality was partly offset by the policy effects. Across the EU, Paulus and Tasseva (2020) identified the direct effect of policy changes, as well as the effect of automatic stabilizers and of changes in market incomes and demographics. For Australia, Li et al. (2020) identifies the policy, demographic, and market income effect, with the extension that the income effect captures both a semi-parametric re-weighting of the industrial and occupational distribution, besides the income adjustment, similar to the semi-parametric approach in Biewen and Juhasz (2012b). Tasseva (2020) decomposes disposable income changes in the United Kingdom, focusing on the role of education on income inequality. Specifically, the study used policy swaps to identify the tax and benefit

aftermath policies.

effect, re-weighting techniques to identify the compostion effect of education, and parametric techniques to identify the effect of returns to education, while other market income components were allocated to the residual. We, however, engage in a higher level of disaggregation by breaking up market income into institutional structures in terms of employment rates, the number of people with income sources, the distribution of income sources, the distribution of the returns, and the demographics using both parametric and semi-parametric techniques. [3] We clearly need to trade off parsimony and complexity. Given the novelty of the work, the computational time, and the limit of how much we can disaggregate, we tried to 'optimize' the balance between model complexity and degree of disaggregation. Future work will assess the sensitivity to different degrees of disaggregation. We have more disaggregation than Bargain and Callan (2010) and Bargain (2012), as we wanted to decompose the drivers of market incomes. The model is constructed on the basis of the European Union Statistics on Income and Living Conditions (EU-SILC) survey, a household survey that is available in a harmonised form for all European Union (EU) countries.

The next section presents the evolution of income inequality and of the economic climate in Lithuania. This is followed by section 2.3 which discusses the income generation model used to characterize and simulate the distribution of household disposable income, the decomposition methodology, and the data. Section 2.4 discribes the changes in the income distribution and redistribution between 2007 and 2015 in Lithuania. In section 2.5 we present the results of the decomposition analysis in Lithuania between 2007 and 2015, followed by a concluding section that discusses several policy implications.

2.2. Evolution of income inequality in Lithuania

Lithuania displayed one of the highest levels of income inequality across the European Union (EU) in 2015. According to the European Union Statistics on Income and Living Conditions (EU-SILC), the most reliable data on income inequality currently available, the Gini index of household equivalized disposable income was 37 Gini points in Lithuania in 2015 (see Figure 2.1). This made Lithuania the second most unequal country in the EU, ranking 6.2 Gini points higher than the EU average and a staggering 12.7 Gini points higher than Slovakia, a country with the most equal income distribution in the European Union and another country formerly behind the Iron Curtain.

3. We could potentially break it up even further, namely, in terms of individual markets; for example, we could swap different industries, swap different parts of the tax-benefit system, swap taxes and benefits separately.

Figure 2.1: Gini index, European Union, 2015

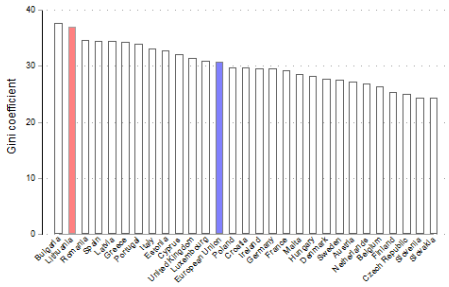
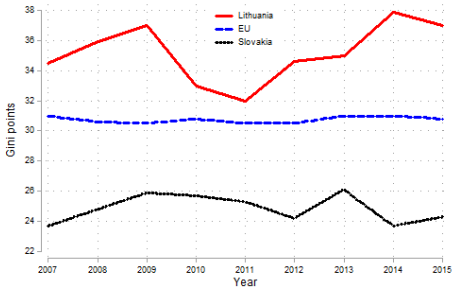


Figure 2.2: Gini index, Lithuania, 2007-2015



Source: Eurostat. Note: Gini index refers to equivalized disposable income.

Income inequality in Lithuania has been on the rise over the past two decades. Figure 2.2 portrays the dynamics of the Gini coefficient for Lithuania, Slovakia, and the European Union as a reference from 2007 to 2015. In Lithuania, the rise in income inequality (as measured by the Gini index) has not been monotonic, displayinga strong procyclicality. It fell during the crisis and then grew rapidly during the post-crisis expansion. Moreover, it appears to be significantly more volatile than the Gini coefficient in Slovakia. Overall, income inequality in Lithuania has consistently exceeded income inequality in Slovakia and the EU in general. In what follows, we discuss potential drivers of changes in the Lithuanian income distribution: demographics, structural and cyclical changes in the economy, and changes in the tax and benefits system.

Demographics

The demographic situation of Lithuania has been affected by three important trends over this period: negative net migration, ageing, and changing household composition. Outmigration, which accelerated significantly after Lithuania's accession to the EU, had a sizeable negative effect on the total size of the population. Specifically, the population of Lithuania decreased by 18% from 2004 to 2016, most of which was due to the negative net migration over the period. This trend has also affected the composition of the population: according to Statistics Lithuania, young workers (those between 15 and 34) are significantly more likely to migrate, causing an increase in the share of elderly in Lithuania. In addition, and similarly to most of Europe, life expectancy has been on the rise. As a result of these two trends, Lithuania's population has become older. In 2004, there were 22 people over 65 for every 100

working-age persons. This number rose to 28 by 2016. This shift might have had important consequences for income distribution, since a greater fraction of the population became dependent on pension income. Finally, the household composition in Lithuania changed. In 2007, almost 60% of households had dependent children, but this has fallen to 51% in 2015. Likewise, there were fewer (legally) married households: 48% of the households indicated that they were married in 2007, but only 39% in 2015. Since the income of married households tends to be more equallly distributed this could also contribute to income inequality.

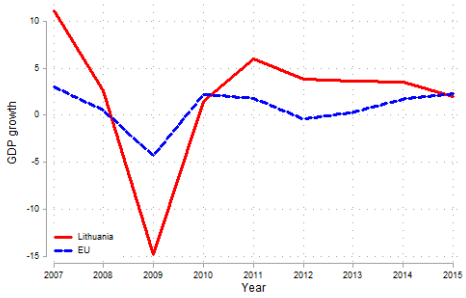
Cyclical and structural changes in the economy

Looking at Figure 2.2, the Gini coefficient in Lithuania appears to be strongly procyclical, much more so than in Slovakia or the average in the EU, which appears highly stable in the period under discussion. The Gini in Lithuania grew somewhat during 2005-2008, peaked at 37% in 2009 and then fell to a low of 32% in 2011, before starting to rise again, reaching 37% in 2015. This pattern coincides with the business cycle of Lithuania, with a bit of a lag.

The financial and economic turmoil that emerged in the global economy following the eruption of the 2007-2008 crisis in the US hit Lithuania particularly hard. Figure 2.3 portrays GDP growth of the Lithuanian economy versus the average in the EU. During the peak of the crisis in 2009, the Lithuanian economy contracted by almost 15% in real terms. Although similar contractions where observed in other Baltic states, this is about three times as severe as in the EU overall. The contraction in Lithuania was due to both internal and external reasons. The economic expansion preceding the crisis was characterized by significant imbalances: double-digit inflation, a housing boom, appreciating real exchange rates, and accelerating wage growth — which exceeded productivity growth. The domestic bubbles burst in early 2008, when the credit supply decelerated and banks started tightening credit conditions. The downturn was further exacerbated by negative developments in the external economic environment after the Lehman Brothers' bankruptcy. The sharp decline was followed by a rapid recovery in Lithuania, with growth rates above the EU average in the early 2010s.

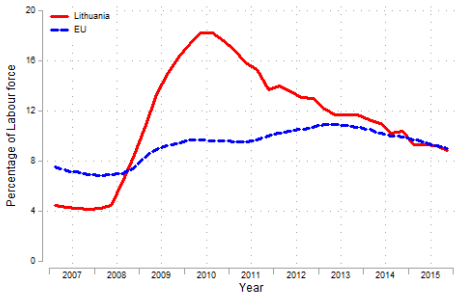
Labour market conditions following the the financial crisis of 2008 worsened dramatically. As shown in Figure 2.4, the unemployment rate rose steadily between 2008 and 2011, from 4% to almost 18%. For the sake of comparison, the fluctuations in the average unemployment rate in the EU were significantly less pronounced. Again, the labour market bounced back rather quickly during the expansion period: the unemployment rate fell below the EU average in 2015. In the face of economic turbulence,

Figure 2.3: GDP growth



Source: Eurostat.

Figure 2.4: Unemployment rate



Source: Eurostat.

the government of Lithuania had to choose between internal and actual devaluation. Internal devaluation was chosen to tackle the external and domestic macroeconomic instability. This generated sharp declines in public and private earnings in the labour market: top public salaries were cut by more than 20 percent and the gross average wages declined by 12.4 percent from the pre-crisis peak to the bottom.

The labour market has also experienced several important structural changes common to most developed countries. One of the most significant changes was a gradual move away from employment in agriculture towards employment in the service sector. The share of employed in agriculture almost halved, from 14% in 2004 to 8% in 2016. As agriculture is the least productive sector, these structural changes in the economy might have affected the income inequality. Additionally, around 8% of Lithuania's population is self-employed and subject to different tax regimes. The share of self-employed has been rising steadily since 2011.

Reforms in the tax and benefit system

The government implemented a large number of reforms in the tax and benefit system during this period.

In 2007-2009, many existing benefit levels were increased. The largest increase in benefit expenditure was due to old age pensions, which constituted 62% of all social protection benefits in 2007. This was partly due to the 35% increase in the state-approved social insurance basic monthly pension. Since pensioners are bunched at the bottom of the income distribution, this had an important redistributive impact. The second highest change in benefit expenditure was due to family/child benefits. The

length of parental leave benefit payout duration increased from one to two years. The effect was particularly strong because parental leave benefits are calculated based on the basis of average monthly reimbursable income (AMRI), which largely consisted of earnings. Since 1 July 2009, AMRI was averaged over 9 months, and since 1 October 2009 – over 12 months, one month before the right to parental leave benefits. This implies that payouts in 2009-2010 were paid based on the all-time-high pre-crisis earnings of 2007-2008. In addition, several child benefits were also increased in this period. The combined result was that expenditure on family/child benefits increased by 2.6 times in 2009 as compared to 2007, and constituted close to 16% of social protection benefits paid in 2009, up from 9% in 2007. State Supported Income, which affects social benefit payouts and unemployment insurance payouts, also increased by 70%.

The legislation which took effect in the 2007-2009 period was largely accepted prior to the crisis and proved unsustainablein a crumbling economy. Therefore, the government cut the spending on benefits substantially in an effort to stabilize the budget deficit by passing the Provisional Law on Recalculation and Payment of Social Benefits. The plan was to reduce the benefits, but only provisionally – between 1 January 2010 and 1 December 2011. The new law capped or reduced a number of benefits in Lithuania. For example, unemployment benefits were capped at 188 euro and old-age pensions either were frozen or decreased. Additionally, a lower ceiling was applied to parental leave benefits. While most of these temporary provisions expired at the end of 2011, several of them, such as reduced state pensions for officers, soldiers, and academic workers, remained in effect until the end of 2013.

During 2011-2015, the benefit system gradually recovered and extended payouts. The Provisional Law on Recalculation and Payment of Social Benefits ended, resulting in higher payout ceilings. Additionally, in 2015, the sickness benefit, which is paid from the State Social Insurance Fund, was increased. Moreover, the economy started to recover, leading to higher earnings and payouts linked to them.

Overall, benefit payouts increased in nominal terms much more in the 2007-2011 period as compared to 2011-2015. The average benefit payouts for the two periods are found in Table 2.1. As can be seen, social assistance increased by 95%, maternity and paternity benefits by 83%, and old age pensions by 26% in the first period in nominal terms. This means that the increase in benefits in 2007-2009 greatly outweighed the provisional cuts in 2009-2011. In contrast, we see much milder increases or even declines in average payouts in the 2011-2015 period (with sickness benefits being the exception). Changes in real terms are harder to interpret in this case, as they crucially depend on the deflator. The relative decline of real growth rates would be just as

apparent if we deflate the benefit payouts by wages (e.g. old age pensions grew in the first period by 11%, but fell in the second by 6%), but less apparent if we deflate by the harmonized index of consumer prices (e.g. old-age pensions grew by 3% in the first and by 9% in the second). This is because wages have grown much faster than inflation in Lithuania since 2011. This table does not allow us identify the extent to which changes in the tax structure (such as changing social insurance basic monthly pension or prolonging parental leave benefits) and market forces (such as dynamics of earnings) affected these payouts. However, it is expected that both factors should play a strong role and that the decomposition procedure should help disentangle the two.

Table 2.1: Nominal growth of average benefit levels

	2007-2011	2011-2015
Old-age pension	26%	13%
Work incapacity and invalidity pensions	27%	4%
Maternity and paternity benefits	83%	-29%
Sickness benefit	1%	31%
Social assistance	95%	-13%
Benefits for bringing up children	49%	1%

Notes: the figures represent percent changes over the period 2007-2011 and 2011-2015 for average social protection expenditures in current prices by selected benefit types. Old-age pension refers to average old-age state social insurance pension payout per person per month. Sickness benefits refer to average expenditure on state social insurance sickness benefit per sick day. Other calculations are available on request. The statistics were calculated by the authors using administrative data on social protection from Statistics Lithuania.

There were important changes in retirement policies over the period. First, from 2006 to 2011, the old-age pension age in Lithuania was 62.5 for men and 60 women. From 1 January 2012, the state pension age gradually increased by 4 months annually, from 60 to 65 years, for women, and by 2 months annually, from 62.5 to 65 years, for men. In 2015, it was 63 years and 2 months for men and 61 years and 4 months for women. Second, in 2004, the pension system was reformed to allow for an opportunity to accumulate and invest a part of the funds in the private sector. Every person insured for full pension insurance was allowed to voluntarily choose either to stay only in the public social insurance system or to switch to the 2nd pension pillar by directing a part of social insurance contributions to a personal account in a chosen privately managed pension fund.

In addition, there have been a number of reforms in the tax system. The personal income tax rate was decreased from 33 to 24% during the course of 2005-2008. Since 2011, all income, except income from distributed profit and income which is subject to a tax rate of 5%, is subject to a uniform tax rate of 15%. During the period of 2011-2013, income from distributed profit was taxed at a 20% rate. Since 1 January 2014, this tax rate was lowered to 15%.

There were also changes in one of the largest components of labour costs, namely, social insurance contributions. These contributions are flat-rate without ceilings, but they differ for employees and self-employed. Employees contribute 3% of gross wages and salaries as contributions to pension social insurance and, since 2009, an additional 6% to health social insurance. Employers, for their part, pay on behalf of their employees 31% of gross wages and salaries to pension social insurance, sickness and maternity social insurance, unemployment social insurance, health insurance, employment injuries, and occupational diseases social insurance. Until 2009, self-employed persons paid contributions only to pension social insurance, depending on their income. Since 2009, self-employed persons additionally contribute to sickness and maternity social insurance. Starting in 2009, social insurance contributions had to be paid on income from sports, performing or authorship/copyright agreements (until 2009, those were only taxed by the personal income tax).

In what follows, we focus on the period between 2007 and 2015, which was a very dynamic period for the Lithuanian economy. We further divide this period into two sub-periods. We are particularly interested in the 2011-2015 sub-period for two reasons. First, the business cycle was in the economic expansion phase throughout the period, making the results easier to interpret. Second, income inequality in Lithuania has increased dramatically during this period. This naturally leaves us with the 2007-2011 period as the second sub-period, which is dominated by the financial crisis of 2008.[4]

2.3. Methodology and data

The objective of this chapter is to decompose changes in the income distribution over time in Lithuania. Given the complex drivers of the income distribution, including demographics, factor markets, market income, and public policy, we require a multi-dimensional framework to undertake the decomposition. Decomposing by population characteristics, income sources, and policy drivers, we utilise the simulation-based approach developed in Sologon et al. (2020)(for disposable income) and Bourguignon et al. (2008) (for market income) for the purpose of cross-national decompositions and

4. We also avoid analysing the period before 2007 as fewer variables were available in EU-SILC.

extended in O'Donoghue et al. (2020) to "nowcast" the distributional impact of the COVID-19 crisis. We used the generic household income-generation model (IGM) developed by Sologon et al. (2020) to simulate the labour market situation and household market income distribution as a function of personal and household attributes and to generate counterfactual distributions under alternative scenarios. The IGM relies on a system of hierarchically structured, multiple equation models for detailed income sources, combining: a set of personal characteristics, parameters describing how the receipt and level of income sources vary with personal characteristics, and residuals linking model predictions to observed income sources. Taxes and benefits are partly calculated using the EUROMOD microsimulation model (Sutherland and Figari 2013) and partly with the help of equation models, as done for IGM. The framework is flexible in comparing disposable income distributions across countries, across regions, or over time to disentangle the role of labour market structure, returns, demographics, and tax-benefit rules. The same factors identified as driving cross-national differences are valid when studying changes in inequality over time.

This framework allows us to decompose overall changes in inequality to changes into 4 factors. The first factor is called "demographics", which captures changes in income distribution due to changes in the distribution of demographic characteristics such as age, sex, and family composition. The second factor is called "labour market structure", which assesses the impact of a changing distribution of the employed, unemployed, the industry at which people work, and their occupations on the income distribution. The third factor is called "prices and returns". This factor quantifies the returns due to demographic factors and labour market factors. Therefore, it includes wages per hour worked, returns for a given occupation, industry, and capital returns (the price of rent and dividends). The fourth factor is the "tax-benefit" system. It quantifies changes in the tax-benefit policy rules on the distribution of disposable income.

All 4 factors and their components vary over time, and crucial to consider when trying to disentangle the factors influencing the distribution of income over time. The methodology simulates counterfactual incomes associated with market, policy, and demographic characteristics of the alternative year, and assesses the impact of changes in these individual components on the total household disposable income distribution. Specifically, we take the underlying demographic structure in time period (s) and simulate the presence of counter-factual market incomes and their level as well as incomes from public policies that exist in the alternative year (t). Doing this in sequence allows us to assess the impact of replacing the market structure, the distribution of market incomes, or the structure of public policies of time (s) with time

(t), holding all other components constant. This enables us to work out how much of the change in the distribution of disposable income was due to individual components (see Sologon et al. 2020) for a detailed description of the micro-simulation micro-econometric approach using the household income distribution model).[5]

In this section, we describe the individual simulation components of the IGM and the mechanism for decomposing disposable income inequality.

2.3.1. Components of the income distribution

We consider 5 broad components of disposable income:

- gross labour incomes, y_h^L (including employee, self-employed incomes),
- household capital incomes, y_h^K (including capital, rental incomes),
- and other household non-benefit pre-tax incomes, y_h^O(including private pension, private transfers, and other incomes),
- public benefits, y_h^B , and
- household direct taxes, y_h^T, which include social security contributions.

We define household disposable income as:

$$y_h = \underbrace{y_h^L + y_h^K + y_h^O}_{Market} + \underbrace{y_h^B - y_h^T}_{Non-market} \tag{2.1}$$

Some market income components are aggregates of smaller components, which are modelled separately to achieve a fine level of disaggregation. *Gross labour income* is aggregated from employment and self-employment income, while *capital income* - from investment and property income. Each component of market income is estimated at the individual level. For each household, the incomes of all individual members are added to obtain the household's income. For each income source, we follow two steps. First, we estimate a binary participation indicator $I_{hi}()$ equal to one if the individual i of household h receives that type of income, and zero otherwise.Second, for the individuals receiving it, we estimate the level $y_{hi}()$. For labour income, we first estimate a binary indicator equal to one if the individual is working, and zero otherwise. Then, for those individuals working, we assign the estimated

5. It is important to note that model parameters do not capture causal relationships between the various endogenous and exogenous variables considered. Rather, parametric relationships are reduced-form projections which describe statistical relationships between basic conditioning variables and various components of income.

income, either from employment or self-employment. *Other non-benefit pre-tax income* are not modelled at such a granular level because too few households had such income. Formally, this is represented by:

$$y_h^L = \sum_{i=1}^{n_h} I_{hi}^{lab} \left(I_{hi}^{emp} y_{hi}^{emp} + I_{hi}^{semp} y_{hi}^{semp} \right) \tag{2.2}$$

$$y_h^K = \sum_{i=1}^{n_h} \left(I_{hi}^{inv} y_{hi}^{inv} + I_{hi}^{prop} y_{hi}^{prop} \right) \tag{2.3}$$

$$y_h^O = \sum_{i=1}^{n_h} I_{hi}^{O} y_{hi}^{O} \tag{2.4}$$

where: n_h is the total number of individuals in household h; I_{hi}^{lab} is an indicator equal to one if individual i belonging to household h (individual hi from now on) is working; and for $S \in$ {emp, semp, inv, prop, other}, I_{hi}^{S} is an indicator equal to one if individual hi receives any income from source S, and y_{hi}^{S} refers to the level of income received from source S.

To simulate counterfactual distributional characteristics, we first statistically estimate individual equations for the presence and level of each of the income sources. For the presence of a market incomes source, we first estimate a binary participation using a logistic model. We model occupation (8 categories, based on the ISCO-08 classification) and industry (primary, secondary, or tertiary) using a multinomial logistic regression model.

For the distribution of wages, we utilise individual characteristics conditional on the whole wage distribution and not only on the conditional mean, as in the regressions used for other income sources; assuming a Singh-Maddala distribution, F_X:

$$F_{X=z}(w) = \text{SM}(w; a(z), b(z), q(\text{z})) = 1 - \left[1 + \left(\frac{w}{b(z)} \right)^{a(z)} \right]^{-q(z)} \tag{2.5}$$

where X indicates that the distribution is conditional on a vector of characteristics z; $q(z)$ is a shape parameter for the 'upper tail'; $a(z)$ is a shape parameter ('spread') affecting both tails of the distribution, and $b(z)$ is a scale parameter. $a, b$and q parameters are allowed to vary log-linearly with individual characteristics, following Biewen and Jenkins (2005) and Van Kerm (2013). The approach utilises a flexible unimodal three-parameter distribution which provides a good fit to wage distributions (Van Kerm et al. 2016). The wage, estimated separately for males and females, is then

given by:

$$w_{hi} = F^{-1}_{X=z}(\upsilon^{emp}_{hi}) = b(z)[(1-\upsilon^{emp}_{hi})^{-\frac{1}{q(z)}} - 1]^{\frac{1}{a(z)}} \tag{2.6}$$

where υ^{emp}_{hi} is a random term uniformly distributed and z contains both demographic variables, x_{hi} occupation, occ_{hi} and industry sector, ind_{hi}. The female wage model is participation-corrected (Van Kerm 2013). The level of non-wage income sources are estimated using a log-linear model for those individuals receiving the income source.

Non-market incomes resulting from public policy such as income taxes, social insurance contributions, social assistance benefits (including housing support), social insurance benefits, and universal benefits are simulated using the EUROMOD tax-benefit microsimulation model (see Sutherland and Figari (2013)). EUROMOD incorporates the tax-benefit schemes of EU member countries, with harmonised input datasets. It simulates social benefits, taxes, and social insurance contribution entitlements, utilising the actual legal rules of the individual policies. Encompassing present and historic tax-benefit policies, EUROMOD allows a user to swap policies between different periods (see e.g. Levy et al. (2007),Bargain and Callan (2010) and Bargain (2012)). We sum income derived from household benefits (y^B_h) and household direct taxes (y^T_h) individually. Household benefits are defined as the sum of household pension income, means-tested benefits and non-means tested benefits:

$$y^B_h = y^{pens}_h + y^{mtb}_h + y^{nmtb}_h \tag{2.7}$$

Direct taxes are defined as a combination of income taxes and social security contributions (ssc):

$$y^T_h = y^{tax}_h + \sum_{i=1}^{n_h} y^{ssc}_{hi} \tag{2.8}$$

All direct taxes and some of the benefits are modelled by EUROMOD. We use regression techniques to model the partially simulated and non-simulated variables. A summary of the variables modelled by EUROMOD and by regression models is available in Appendix Tables A5 and A6.

2.3.2. Simulating counterfactual distributions

As outlined at the start of the section, we utilise these market and non-market models to simulate counterfactual distributions and to undertake a decomposition of changes in the income distribution over time, between period t and period s. The income generation model (IGM) can be defined as:

$$Y = m(X, \Upsilon; \xi) \tag{2.9}$$

where:

- Y is household disposable income,
- X is a vector of exogenous characteristics,
- ξ is the vector of parameter values and
- Υ is a vector of unobserved heterogeneity terms.

The income generating process is not a 'structural' model, but rather a statistical representation of the structure of the presence and the level of market incomes, as well as the the tax-benefit rules.

The objective of this approach is to understand how the distribution F of a random variable Y (such as disposable income) as well as any functional of interest $\theta(F)$ (such as inequality indices, quantiles) varies over time, to answer the question: 'What would the income distribution of time t be if its IGM was the one of time s along one or more of the dimensions considered?'. In particular, we are interested in the degree to which changes in the individual components affect changes in the distribution of disposable income.

The change depends on the (joint) distribution of X and Υ in the population through m and ξ resulting from differences in the distributions of observable characteristics as well as unobservable residual heterogeneity and differences in the model's parametric structure and parameter values. We assume that all years can be represented by a common parametric model of the form m but that years differ in the values taken by the parameters ξ. We undertake the decomposition in the income distribution over time by swapping individual income components between periods, one at a time. To do this, we estimate the IGM for each year separately and calibrate transformations so as to replace components of the IGM of year t with components of the IGM of year s.This is analogous to the standard Oaxaca-Blinder decomposition but implemented in a multiple equations model and over time.

In swapping components between periods, there are many combinations that are possible, given the range of different incomes and income components. In this study, we focus on four 'transformations':

- a labour market structure transformation;
- a returns transformation;
- a demographic transformation; and
- a tax-benefit system transformation.

Below we outline the transformation in a general form and leave the exact variables on which the transformations are applied to the Appendix tables A5 and A6 (see columns "variables" and "factors" in particular). We also included the main model tables (Tables A8 to A23 in the Appendix) while the rest of the model tables are available on request.

The labour market structure transformation changes important characteristics of the labour market structure such as employment, occupation, and industry sector, and involves swapping between periods the elements of the parameter vector ξ characterising the labour market to simulate an alternative parameter vector, $\tilde{l}(\xi)$,which will result in an alternative outcome Y^l:

$$Y^l = m(X, \Upsilon; \tilde{l}(\xi)). \tag{2.10}$$

Y^l is the counterfactual distribution that would prevail in period t if we "import" the labour market structure of period s, while keeping everything else the same.

The returns transformation acts through the parameter vector ξ, changing the parameters of the equations for each market income source (employment income, self-employment income, capital income, modelled benefit income, other income) to produce an alternative parameter vector, $\tilde{r}(\xi)$, which would result in an alternative outcome Y^r :

$$Y^r = m(X, \Upsilon; \tilde{r}(\xi)). \tag{2.11}$$

Y^r is the counterfactual distribution that would prevail in period t if we "import" the structure of returns of period s, while keeping everything else constant. This follows the logic of the manipulation of the vector of coefficients in Mincerian earnings regressions aimed to capture 'price' effects (as distinct from 'composition' effects) in traditional Oaxaca-Blinder decomposition exercises. It resembles the decomposition of Juhn et al. (1993) in the way residual variances are accounted for: it swaps the variance terms by rescaling the residuals of time t for each of the five income components, but preserves the rank correlation of the residuals.

The demographic transformation changes the values of variables relating to socio-demographic characteristics of the population (education, age, sex, number of children by age, legal marital status, citizenship, and whether the individual is over 65 without any children to account for single elderly households) and involves a modification of the distribution of the random variables in X as in Sologon et al. (2020). We reweigh the population at time t to resemble the population structure at time s by a factor obtained semi-pametrically following DiNardo et al. (1996) and Barsky

et al. (2002):

$$\omega(X) = \frac{\Pr(X|s)}{\Pr(X|t)} = \frac{\Pr(s|X)}{\Pr(t|X)} \frac{\Pr(t)}{\Pr(s)} \tag{2.12}$$

The alternative distribution of $\tilde{X}(X)$ results in obtaining a counterfactual outcome for income,Y^d:

$$Y^d = m(\tilde{X}(X), \Upsilon; \xi). \tag{2.13}$$

Y^d is the counterfactual distribution that would prevail in period t if we "import" the demographic structure of period s, while keeping everything else constant.

The tax-benefit system transformation modifies the level and eligibility of benefits and tax liabilities, simulated by EUROMOD, to produce an alternative parameter vector $\tilde{tb}(\xi)$. This involves swapping model parameters as above for the equations describing the benefits not fully simulated by EUROMOD, and using EUROMOD to apply the tax-benefit rules and parameters of period s onto the market incomes and household characteristics of period t. For these simulations, pre-fiscal monetary variables are inflated (deflated) to the year of the tax-benefit system being considered by using the EUROMOD uprating indices. Most non-benefit monetary variables, including employment and self-employment incomes, are uprated by the average gross monthly earnings index. Several income components, such as investment income, private pensions, private transfers, and some benefit monetary variables are uprated by the harmonized index of consumer prices. Most benefit monetary variables are uprated by benefit specific indices (for example, social assistance benefits are uprated by an index that captures the change in the average amount of monetary social assistance benefit received between years). Similar swapping of tax-benefit policy rules and parameters were implemented for analysing trends in income distributions (see Herault and Azpitarte 2016; Bargain and Callan 2010; Bargain 2012; Paulus and Tasseva 2020) and cross-country differences (see Sologon et al. 2020; Levy et al. 2007; Dardanoni and Lambert 2002).

The resulting counterfactual is formalized as:

$$Y^{tb} = m(X, \Upsilon; \tilde{tb}(\xi)) \tag{2.14}$$

Y^{tb} is the counterfactual distribution that would prevail in period t if we import the tax-benefit rules of period s, while keeping everything else constant.

For each of the four transformations, the impact is assessed by comparing the original distribution in period t with each counterfactual. We can compute the impact on any distribution functional of interest, θ, such as the Gini index or the quantiles.

This type of measure is called a partial distributional policy effect in Rothe (2012) or simply a policy effect in Firpo et al. (2009). For transformation k with $k \in$ {l, r, d, tb}, this impact is given by:

$$\Delta_{\theta}^{k}(F) = \theta(F) - \theta(F^{k}). \tag{2.15}$$

In our approach, the incomes obtained in the simulations are aligned to the year of the tax-benefit system being applied. For example, when we apply the period t tax-benefit system, the resulting incomes are in period t values. This implies that counterfactuals obtained by importing in period t the demographics, labour market structure, and returns from period s are aligned with period t values.

When we import the tax-benefit rules from period s, however, the resulting simulated incomes are aligned with period s, in terms of both productivity level and prices. As we need to compare this counterfactual with the original t distribution using a scale-variant distributional functional, such as the quantiles, we need to index disposable incomes by the average market income index to ensure all incomes are expressed in period t values (in terms of productivity and prices), in line with the other counterfactual differences. As the aim of the tax-benefit transformation is to evaluate actual policy changes, we use a distributional neutral benchmark given by the actual change in average market income levels between period t and s (Bargain and Callan, 2010). Specifically, we adjust the simulated incomes expressed in 2011 values by the ratio between the mean market income in 2015 and 2011. We perform a similar adjustment for 2007. This way we account for the price changes and for the productivity growth between the years. This ensures that the tax and benefit effect measures the change in relative position of those who do get market incomes and those who do not (e.g. welfare payments), thereby capturing the change in generosity of the system. That is, we measure the marginal effect of the tax and benefit system on disposable income when we control for the level of productivity growth and prices (as well as demography and labour market structure).When we compare distributions using scale-invariant distribution functionals, such as the Gini index, inflating (deflating) disposable incomes has no impact on the comparison.

2.3.3. Decomposition of changes in the income distribution over time

Next, we decompose the observed differences between income distributions and their corresponding functionals in years t and s. We compute a certain functional $\theta(F)$ for each of the two years, $\theta(F^t)$ and $\theta(F^s)$. Our procedure aims to decompose the total observed difference, $\theta(F^t) - \theta(F^s)$, into the contributions of each of the individual

determinants k of a set K:

$$\Delta_\theta(F^t, F^s) = \theta(F^t) - \theta(F^s) = \sum_{k=1}^{K} \Delta_\theta^k(F^t, F^s) \tag{2.16}$$

One approach is to apply each transformation sequentially, one after the other, from the original distribution, F^t, to the target distribution, F^s, and take the difference between two consecutive steps of the sequence. The drawback of such a sequential decomposition is path-dependence, i.e. the estimated contribution of each factor depends on the chosen sequence. To reduce issues of path-dependence[6], we focus on 'direct effects' following Biewen and Juhasz (2012a) and Biewen (2014). The direct effect assesses the impact of each factor from the same initial benchmark distribution:

$$D_\theta^k(F^t, F^s) = \theta(F^t) - \theta(F_t^k) \tag{2.17}$$

where F_t^k is the counterfactual distribution obtained by applying one transformation k to the initial distribution F^t. Comparing direct effects is a natural way to assess the effects of alternative transformations (Biewen and Juhasz 2012a). The sum of all direct effects and unexplained factors does not add up to the overall observed difference. The discrepancy reflects interactions between components.

In the context of our decomposition, we have four direct effects of each transformation, the unexplained component, and an interaction term:

$$D_\theta^l(F^t, F^s) = \theta(F^t) - \theta(F_t^l) \tag{2.18}$$

$$D_\theta^r(F^t, F^s) = \theta(F^t) - \theta(F_t^r) \tag{2.19}$$

$$D_\theta^d(F^t, F^s) = \theta(F^t) - \theta(F_t^d) \tag{2.20}$$

$$D_\theta^{tb}(F^t, F^s) = \theta(F^t) - \theta(F_t^{tb}) \tag{2.21}$$

$$\Delta\Upsilon_\theta(F^t, F^s) = \theta(F^t) - \theta(F_s^{l,r,d,tb}) \tag{2.22}$$

$$I_\theta(F^t, F^s) = \big(\theta(F^t) - \theta(F^s)\big) - \left[\left(\sum_{k\in\{l,r,d,tb\}} D_\theta^k(F^t, F^s)\right) + \Delta\Upsilon_\theta(F^t, F^s)\right]. \tag{2.23}$$

The term $\Delta\Upsilon_\theta(F^t, F^s)$ captures the contribution of differences in the distribution of scaled residual or unobserved heterogeneity terms Υ between period t and s. Fol-

6. We do not eliminate path-dependence completely. For example, our results are conditional on the choice of the reference year.

lowing the original approach in Sologon et al. (2020), we did not perform specific transformations involving the residual terms since they do not have clear-cut economic interpretations: they mostly reflect the correlation of scaled residuals across all income sources and differences over time in residual distributions that may be due to unmodelled heteroscedasticity or model misspecification.[7] $I_\theta(F^t, F^s)$ is an interaction term. Following Biewen (2014) and Sologon et al. (2020), it is calculated as the total difference in θ (net of the unexplained effect) minus the sum of direct effects, accounting for all two-way and three-way interactions between the four components in the model.

The total observed change over time is decomposed into:

$$\Delta_\theta(F^t, F^s) = D^l_\theta(F^t, F^s) + D^r_\theta(F^t, F^s) + D^d_\theta(F^t, F^s) + D^{tb}_\theta(F^t, F^s) + \Delta\Upsilon_\theta(F^t, F^s) + I_\theta(F^t, F^s) \quad (2.24)$$

As a robustness check, we also use the Shapley value approach, as in Anthony F Shorrocks (2013) and Sastre and Trannoy (2002) ((see, e.g., Deutsch et al. 2018, for a recent application)). The procedure calculates marginal contributions of each component in all possible decompositions, and then averages them out. We report the Shapley value decomposition results for the full sample period in the Appendix, while we use the direct effects throughout the text. Our conclusions are robust across the two approaches.

2.3.4. Data

We use the nationally representative household survey for Lithuania: the European Union Statistics of Income and Living Conditions (EU-SILC) for period 2007 to 2015. This yearly survey contains detailed information about income in the preceding year as well as the socio-economic characteristics of households and their members, largely during the survey year. Therefore, we focus on 2008, 2012 and 2016 EU-SILC survey waves for Lithuania.

Given that a central component of our income generation process is the tax-benefit microsimulation engine EUROMOD, we use the 'EUROMOD input data' versions of the EU-SILC dataset for Lithuania, which have been standardized for common definitions of income variables and household characteristics (Sutherland and Figari 2013). The disposable household income in EUROMOD is composed of the sum across all household members of market incomes and public pensions plus cash benefits, minus

7. $\Delta\Upsilon_\theta(F^t, F^s)$ is obtained by swapping residuals across periods. Starting from time s we jointly apply all four transformations calibrated to period t parameters. Subtracting this construct from time t's original distribution we capture the difference between the residuals of period t and period s.

taxes and social insurance contributions. The 'EUROMOD input data' that we feed to EUROMOD are already modified by the IGM. That is, the labour market transformation, the returns transformation, and part of tax and benefit transformation (for values not modelled by EUROMOD) have been applied to the data to derive hypothetical income distributions.

Additionally, the values have been uprated (i.e. indexed to nominal averages of respective system years) before being fed to EUROMOD. The uprating values differ depending on the monetary value (for example, pensions are uprated to average statutory pension each year, while labour income is uprated to average gross wages of that year). Then, direct taxes, social insurance contributions and a part of cash benefits are calculated by EUROMOD. EUROMOD assumes full take-up of benefits (no tax evasion). All incomes are expressed in single adult equivalent by dividing total household income by the square root of household size. Sample sizes exceed 10 thousand individuals, corresponding to just under 5 thousand households in each year.

Table 2.2: Population socio-economic characteristics (shares of total population)

	2007	2011	2015	2007-2011	2011-2015
Demographic					
Tertiary Education	0.287	0.332	0.358	0.045 (0.014)	0.026 (0.015)
People 16-65	0.684	0.670	0.665	-0.014 (0.008)	-0.005 (0.008)
People >65	0.148	0.173	0.179	0.024 (0.007)	0.006 (0.008)
Child 0-3	0.038	0.037	0.039	-0.001 (0.005)	0.002 (0.005)
Child 4-11	0.080	0.073	0.081	-0.007 (0.006)	0.008 (0.006)
Child 12-15	0.049	0.047	0.036	-0.002 (0.004)	-0.011 (0.004)
Married	0.578	0.530	0.469	-0.048 (0.011)	-0.061 (0.012)
Citizen	0.995	0.995	0.992	0.000 (0.002)	-0.002 (0.002)
Male	0.444	0.450	0.451	0.006 (0.007)	0.000 (0.007)
Household size	3.316	3.091	2.991	-0.225 (0.074)	-0.101 (0.068)
Labour market structure					
Months worked	6.629	5.903	6.479	-0.726 (0.121)	0.576 (0.124)
Employee/Self-Employed	0.897	0.942	0.910	0.045 (0.007)	-0.032 (0.007)
Occupation					
Managers	0.139	0.115	0.115	-0.024 (0.009)	0.000 (0.009)
Professionals	0.168	0.233	0.229	0.064 (0.012)	-0.003 (0.013)
Associate Prof.	0.104	0.084	0.071	-0.021 (0.008)	-0.013 (0.007)
Clerks	0.041	0.038	0.043	-0.003 (0.005)	0.005 (0.005)
Service	0.118	0.125	0.122	0.007 (0.010)	-0.003 (0.009)
Craft	0.204	0.193	0.189	-0.011 (0.011)	-0.003 (0.011)
Plant	0.112	0.103	0.103	-0.009 (0.008)	-0.001 (0.008)
Unskilled	0.113	0.110	0.129	-0.003 (0.008)	0.018 (0.009)
Industry					
Agriculture	0.078	0.058	0.052	-0.020 (0.007)	-0.006 (0.006)
Industry	0.246	0.155	0.151	-0.091 (0.012)	-0.003 (0.010)
Services	0.676	0.788	0.797	0.111 (0.013)	0.009 (0.012)
Business certificate	0.262	0.191	0.215	-0.071 (0.040)	0.024 (0.038)
Price and returns					
With wage income	0.615	0.606	0.653	-0.009 (0.011)	0.047 (0.011)
Wages	4.263	3.750	4.624	-0.513 (0.097)	0.874 (0.105)
With capital income	0.085	0.075	0.164	-0.010 (0.007)	0.089 (0.008)
Capital income	9.004	4.883	9.174	-4.122 (2.620)	4.291 (2.035)
Nr. of observations	12130	12659	10895		

Notes: The estimates are weighted. The shares for education refer to age-group 25-64; for married, sex to age >= 16; for months worked to ages 16 to 80; for employees, occupation, industry and sector to those in work aged 16 to 80; for citizen to the entire sample; for buisness sertificates to self-employed. The shares for capital refer to age>= 16. Wages and capital income deflated by the harmonized index of consuer prices. Standard errors in parenthesis.

The demographic, cyclical, and structural changes discussed previously are visible in the EU-SILC data. Table 2.2 shows several population socio-economic characteristics for each of the three years, based on the samples in our database.

In terms of demographics characteristics, we see a relative increase in education attainment and life expectancy and a decline in the presence of children, especially those aged 12 to 15 and a relative decline in (legal) marriage rates (from 58 percent in 2007 to 47 percent in 2015).

Changes in the labour market structure are more nuanced. In 2007, an average respondent worked for 6.6 months during the year; this fell to 5.9 in 2011. This constitutes a greater than 10% reduction in employment time during the crisis years. The economy recovered in 2015, when an average person worked for 6.5 months. The crisis has also changed the composition of employees and self-employed among those who were employed. In 2011, self-employment plummeted by about half, reflecting the vulnerability of this type of work during turbulent times. The distribution of workers across types of occupation also experienced some changes: the economy experienced an increase in the share of professionals and a decrease in the share of associate professionals.

This change in composition of occupations relates to an increase in the share of people with tertiary education: a larger share of high-skilled workers were able to take more qualified jobs. There was also a large shift towards the service sector at the expense of the agricultural and industry sectors, as expected.

Finally, looking at the participation and returns in the labour and capital markets, we can see that the share of people with capital income doubled since 2007. Average capital income increased by about 87 percent after accounting for inflation, while it decreased by 46 percent during the first sub-period. We observe similar dynamics in the labour market: wages have fallen by 12 percent and increased by 23 percent during the first and second sub-periods, respectively.

We take this as evidence of significant changes in the returns of investments in both the labour and the capital markets.

2.4. Changes in the income distribution and redistribution between 2007 and 2015 in Lithuania

2.4.1. Changes in disposable income inequality

We start by characterizing the changes in the distribution of equivalised household disposable incomes in Lithuania between 2007 and 2015, considering both the period 2007-2015 as a whole and two sub-periods: 2007-2011 and 2011-2015.

Table 2.3 shows the mean and median monthly disposable incomes and the Gini

index associated with each of these periods. We present both nominal and uprated values in order to assess the evolution of incomes relative to price developments (the harmonized index of consumer prices, HICP). Nominal values do not differ a lot between 2007 and 2011, but there is a rapid increase in 2015. The HICP uprated mean and median income values, however, were significantly lower in 2011 as compared to 2007. Therefore, we observe a decline in purchasing power during the economic crisis period.

In contrast, the mean and median income rose roughly by 34% increase since 2011. The Gini moved in tandem with real incomes. It slightly fell between 2007 to 2011, but then increased by 2.9 Gini points in 2015.

Table 2.3: Summary statistics of household disposable income

	Nominal		HICP adjusted		
	Mean	Median	Mean	Median	Gini
2007	433	369	549	467	0.339
	(4.34)	(3.84)	(5.50)	(4.87)	(0.0041)
2011	438	364	455	378	0.331
	(3.59)	(3.89)	(3.73)	(5.63)	(0.003)
2015	611	508	611	508	0.360
	(6.66)	(5.82)	(6.66)	(5.82)	(0.0039)

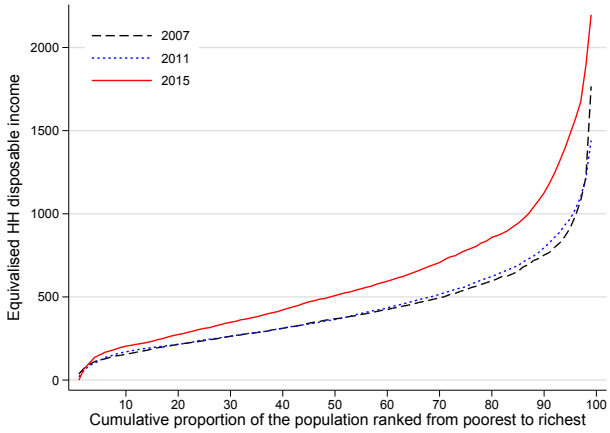
Note: Standard errors in parenthesis.

The rise of the Gini alongside rising mean and median incomes suggests that incomes rose unevenly for the population, particularly from 2011 to 2015. We see this in Figure 2.5 in the form of Pen's parades. When comparing the distributions of 2007 and 2015, it can be seen that almost all quantiles experienced an income increase, including the quantiles at the bottom of the income distribution.

Furthermore, income increased the most since 2011 and barely changed in the previous period. What we also see is that the income of different quantiles increased by different absolute amounts - those at the top gained significantly more than those at the bottom.

The relative increase in income is presented in two panels of Figure 2.6. Panel 2.6a shows the pairwise differences between the three distributions shown in Figure 2.5, as a percentage of the 2015 distribution. For each percentile, the change between 2007 and 2015 is equal to the sum of the change between 2007 and 2011 and the change between 2011 and 2015. Therefore, for each percentile, the change over the whole period can be decomposed into the contributions of each of the two sub-periods. The 2007-2015 period comprised two very distinct sub-periods in what

Figure 2.5: Distribution of equivalised household disposable income (nominal values)

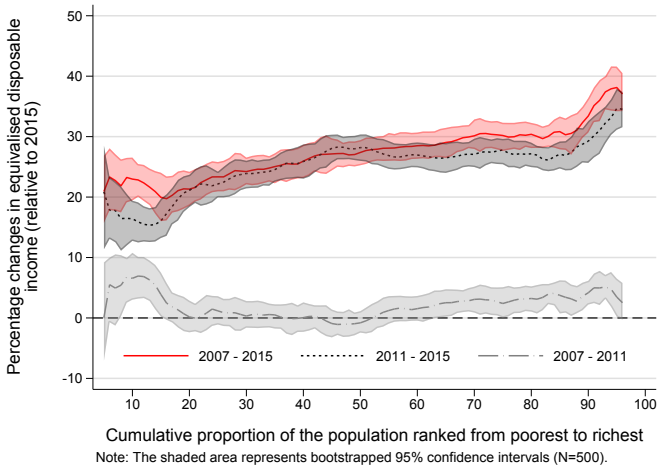


concerns the evolution of incomes across the income distribution. The years between 2007 and 2011 brought mild increases in the income of some of the poorer and the richer, while the bottom 5% and the 40-50% actually lost income. This contrasts with the 2011-2015 period, where income of the entire distribution rose. However, the rise in income in 2011-2015 period differs along the distribution: it rose by around 20% for the bottom 20% of the population and around 30% for the top of the population. The top 10% of the distribution gained even more than 30% in their disposable income. Therefore, the economic upturn increased inequality between the tails of the distribution.

Panel 2.6b contains the HICP deflated quantile differences. Therefore, it captures the drop in purchasing power from 2007 to 2011 across the income distribution. Even though this was followed by a rapid recovery, incomes at the bottom of the income distribution increased by far less that those at the top. As a result, the purchasing power of those at the bottom of 25 percent of income distribution was the same in 2015 as in 2007.

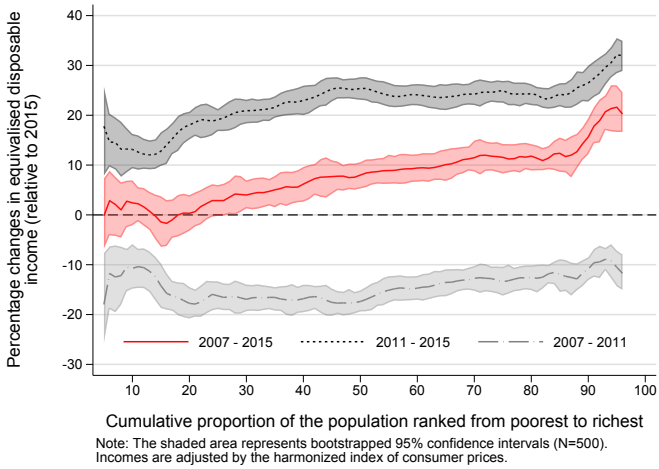
Figure 2.6: Relative changes in the distribution of equivalised household disposable income

(a) Nominal values



Note: The shaded area represents bootstrapped 95% confidence intervals (N=500).

(b) HICP adjusted to 2015 prices



Note: The shaded area represents bootstrapped 95% confidence intervals (N=500).
Incomes are adjusted by the harmonized index of consumer prices.

2.4.2. The redistributive effect of the tax and transfer system

An important determinant of the disposable income distribution is the redistributive action of the tax and transfer system, which typically cushions developments in the market income distribution. In Table 2.4 we provide summary indicators of the effect of the system as a whole, as well as the partial effects of taxes and transfers. The

effectiveness of the system as a whole is measured by net redistribution, which is defined as the difference between the Gini of market income and the Gini of disposable income. Next, the effectiveness of each component of redistribution, i.e. transfers and taxes, is evaluated separately. Specifically, we present measures of (i) redistribution, given by the Reynolds-Smolensky index; (ii) average tax (transfer) rates, defined as the ratio between the total amount of taxes (transfers) paid (received) and the total pre-tax (transfer) income; and (iii) progressivity/regressivity effect, measured by the Kakwani index[8].

Table 2.4: The redistributive effect of the tax and transfer system

	2007	2011	2015	2007-2011	2011-2015	2007-2015
Gini Market Income (1)	0.473 [0.463 - 0.483]	0.513 [0.505 - 0.521]	0.515 [0.505 - 0.525]	0.040 [0.028 - 0.053]	0.002 [-0.011 - 0.015]	0.042 [0.028 - 0.056]
Gini Disposable Income (2)	0.339 [0.33 - 0.349]	0.331 [0.325 - 0.338]	0.360 [0.352 - 0.368]	-0.008 [-0.020 - 0.002]]	0.029 [0.017 - 0.038]	0.021 [0.008 - 0.032]
Net Redistribution (1)-(2)	0.134 [0.128 - 0.139]	0.182 [0.175 - 0.188]	0.155 [0.149 - 0.161]	0.048 [0.039 - 0.057]	-0.026 [-0.036 - -0.018]	0.021 [0.012 - 0.029]
Gini Market Income (+ Transfers) (3)	0.369 [0.36 - 0.379]	0.364 [0.358 - 0.371]	0.391 [0.383 - 0.399]	-0.005 [-0.017 - 0.005]	0.026 [0.015 - 0.037]	0.021 [0.009 - 0.034]
Average Transfer Rate	0.186 [0.178 - 0.195]	0.252 [0.241 - 0.263]	0.223 [0.213 - 0.233]	0.066 [0.053 - 0.081]	-0.029 [0.016 - 0.046]	0.037 [-0.052 - -0.024]
Transfer Regressivity	0.768 [0.745 - 0.791]	0.845 [0.832 - 0.860]	0.801 [0.782 - 0.820]	0.078 [0.05 - 0.104]	-0.044 [-0.066 - 0.021]	0.034 [0.003 - 0.062]
Transfer Redistribution (RS) (1)-(3)	0.104 [0.099 - 0.108]	0.148 [0.142 - 0.154]	0.124 [0.119 - 0.129]	0.045 [0.037 - 0.053]	-0.024 [-0.032 - -0.016]	0.021 [0.013 - 0.028]
Gini Market Income (+ Transfers - Taxes) (4)	0.341 [0.332 - 0.350]	0.343 [0.337 - 0.349]	0.372 [0.364 - 0.381]	0.002 [-0.01 - 0.012]	0.030 [0.018 - 0.040]	0.032 [0.02 - 0.043]
Average Tax Rate	0.177 [0.175 - 0.179]	0.100 [0.099 - 0.101]	0.107 [0.105 - 0.108]	-0.077 [-0.08 - -0.075]	0.007 [0.005 - 0.008]	-0.070 [-0.073 - -0.068]
Tax Progressivity (K)	0.144 [0.139 - 0.149]	0.199 [0.193 - 0.205]	0.161 [0.154 - 0.165]	0.055 [0.047 - 0.063]	-0.038 [-0.047 - -0.032]	0.017 [0.009 - 0.024]
Tax Redistribution (RS) (3)-(4)	0.029 [0.028 - 0.03]	0.022 [0.021 - 0.022]	0.019 [0.018 - 0.019]	-0.007 [-0.009 - -0.006]	-0.003 [-0.004 - -0.002]	-0.010 [-0.012 - -0.009]

Notes: K = Kakwani; RS = Reynolds-Smolensky. Bootstrapped 95% confidence intervals (N=500) are reported in squared brackets.

The analysis of these indicators suggests several findings. First, in terms of overall redistribution, the tax and transfer system as a whole was a crucial determinant of the level of disposable income inequality in Lithuania. In each of the three years considered, the net redistributive effect was around 15 Gini points, or about 30% of the Gini of market income. However, the system was not equally redistributive throughout the whole period. The tax and benefit system became more redistributive in 2011 as compared to 2007, as the net redistributive effect increased by 35%, from 0.134 to 0.182. The effect was large enough to dominate the increase in market income inequality by more than 13%: the resulting disposable income inequality was smaller than in 2007. The system, however, became less redistributive in 2015 as compared to 2011: disposable income inequality increased, even though market income inequality

8. Note that in the case of transfers, higher regressivity means more transfers being *received* by lower income households, while in the case of taxes, higher regressivity means more taxes being *paid* by lower-income households. Therefore, an increase in transfer regressivity increases redistribution while an increase in tax progressivity (and therefore a decrease in tax regressivity) increases redistribution.

did not change during this period.

Second, considering the redistributive effects of each part of the system, one can see that the bulk of redistribution was due to transfers. In 2007, transfers accounted for 78% of the total redistribution effect, whereas the tax system was responsible only for 21%. In addition, transfers became even more important in 2011: the average transfer rate and the benefit regressivity increased as compared to 2007 while average taxes fell. However, the increase in the importance of benefits in 2011 was partly undone by 2015, when average transfer rates and regressivity decreased.

2.5. Drivers of changes in the income distribution between 2007 and 2015

This section decomposes the changes of total income inequality presented in Subsection 2.4 into the contributions of the main factors considered in our model, as described in Subsection 2.3.3. This helps us understand why income inequality changed.

Decomposing changes in incomes

Figure 2.7 shows the contribution of each factor to the total changes in income distributions (i.e., the decomposition of the total changes in income distribution that were depicted in Figure 2.6b). Analogously to the results presented in Figure 2.6, for each percentile in each graph, the change in the period 2007-2015 is equal to the sum of the changes in the periods 2007-2011 and 2011-2015. Furthermore, for each percentile, and each period, the total change in the income distribution given in Figure 2.6 is equal to the sum of the four factor contributions as portrayed in Figure 2.7 as well as the the interaction effects and the residuals. The joint effect of the latter two can be found in Figure A1 in the Appendix.

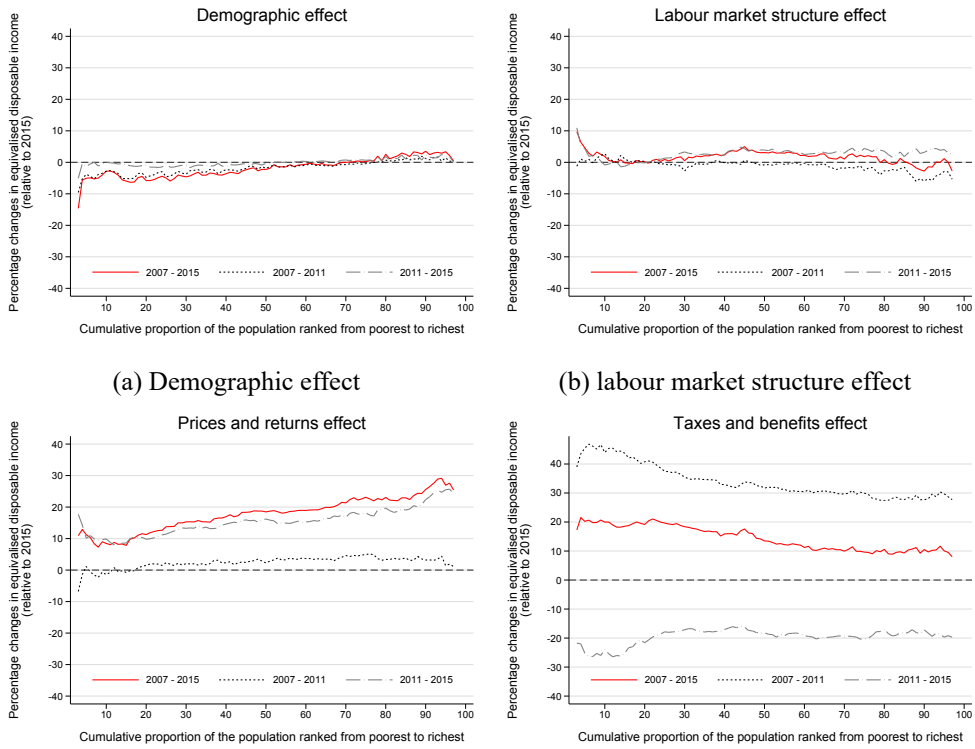
All four factors contributed to changing household disposable income distribution in Lithuania. The biggest effect was due to the price and returns effect, as well as changes in the generosity of the tax and benefit system. Changes in price and returns increased disposable income of the median household by about20% during the whole period, whereas changes in the tax and benefit system generosity contributed another 12%. Changes in labour market structure increased income by 5% and the demographic effect generated a negative change in the disposable income of the median household.

Changes in the transfer system, the prices and returns as well as the demographics, appear to have affected the income inequality: the size (and the sign in some cases) of the effects vary, depending on the position on the income distribution. As expected, changes in the tax and benefits increased the income of the bottom deciles more than

the top of the income distribution. The effect generated a decrease in income inequality. Nonetheless, the top of the income distribution has benefited significantly more from the changes to the price and returns of the markets, which has contributed to the rise of the income inequality. Although the demographic effect had a smaller impact on the level of disposable income, its effect on inequality appears to be very significant over the analysed period. This is because changes in the demographics of the population affect the bottom and the top of the income distribution unequally: due to the demographic effect, the income of the bottom 30% of the population was 5% lower in 2015, whereas the income of the top has increased by 5%. The size of the contribution of the demographic effect to increasing income inequality is comparable to the size of the tax and benefit effect acting in the opposite direction.

Looking at the two sub-periods, neither changes in the tax and benefit system nor the prices and returns had the same effect throughout the whole period. The largest gains for the bottom of the income distribution was due to the changes in tax and benefits over the crisis period. This was partly because benefits were substantially raised in this period, as well as because market incomes have dropped. In contrast, the tax and benefit became much less generous during the upturn, because benefits increased less or not at all, while market incomes rose rapidly. Furthermore, benefits that target the bottom of the income distribution, such as social assistance, actually decreased during the 2011-2015 period and as a result the bottom 20% benefited less than the rest of the distribution. In contrast, the price and returns played a modest role in 2007-2011; most of the effect came during the years of economic expansion. This speaks to the nature of the prices and returns effect and is consistent with a procyclical nature of that effect. Overall, the emerging picture implies that the measures adapted by the tax and benefit system could not deliver sufficient redistribution at a time when incomes were rising rapidly, i.e. during the upturn of the business cycle. In contrast, the demographic effect appears to be less sensitive to the business cycle conditions. It slowly but gradually increased inequality in both sub-periods, likely due to the secular nature of the demographic shifts.

Finally, the effect of changes in the labour market structure appear to be mostly concentrated at the bottom of the income distribution. There is a positive effect on the bottom 5% of households: their income increased by almost 10% during the whole period, with most change happening in second period. The income of households in the middle of the income distribution also increased slightly. Interestingly, the top of the income distribution either did not gain or lost income because of changes in the labour market structure.



(a) Demographic effect

(b) labour market structure effect

(c) Prices and returns effect

(d) Taxes and benefits effect

Figure 2.7: Decomposition of changes in the distribution of equivalised household disposable income

Further decomposing of the demographic effect

To further decompose the demographic effect, we calculated the contribution of each observable that we use to calculate the demographic effect. The results are presented in Table A2 in the Appendix. For the sake of brevity, we only report the contributions of the most important demographic factors: age, education, and marital status. Table A2 discloses that declining (legal) marriage rates contributed the most while increasing rates of education, defined as tertiary level education versus all lower education levels, rates had an important role as well. The marriage effect generated a very unequal and negative effect across almost the whole of the income distribution. The rising education rates, by contrast, resulted in a positive and significantly more equal effect across the distribution.

The combination of these two effects, displayed in Figure A2d, explains the totality of the demographic effect. Interestingly, the ageing of society does not appear to

have played a significant role in explaining the recent increase in income inequality.

The reason why marriage had a large effect on income inequality seems to stem from the fact that inequality among married households is smaller than among households with a single adult. This means that as a smaller share of population becomes married, income inequality increases. This finding is consistent with Burtless (1999), who found a similar result for the United States in the late 20th century. There are several factors that might generate this effect.

The low earnings of one partner can be offset by the higher earnings of the second partner — an insurance mechanism that non-married households (which are largely also single-person households) do not have. Alternatively, marriage can be a "luxury" into which higher income earners self-select. Additionally, our results show that fewer married households are linked to lower household disposable income.

While our decompositions are not causal, other studies do tend to suggest that this link may be causal: as summarized by Lundberg et al. (2016), married men tend to work longer hours and get higher earnings. This is related to changing behaviour (reducing risky activities such as drug use or drinking and increasing preferences for work). As such, falling marriage rates among the poorest households maybe especially problematic, as this pushes them into even deeper poverty.

Education has an overall positive effect for incomes, although the effect is slightly stronger in the upper tail of the income distribution. This means that the rise in the number of people with tertiary education is associated with higher income overall, even though the richest individuals benefit more. This result is in line with Magda et al. (2020), who show that education contributed to wage inequality in Lithuania in the similar period due to the composition effect (more educated people).

Importantly, the demographic effect only captures a part of the overall education effect. This is because the demographic effect only considers the share of people with tertiary education, but does not consider the returns to education. Magda et al. (2020) finds that returns to education in terms of wages indeed declined in Lithuania in a similar period, and that this decline was strong enough to offset the composition effect (more people getting tertiary education). As a result, more education also means more equal wages and subsequently more equal incomes. Additionally, those with higher education tend to receive more equal incomes than those who do not have it (Černiauskas and Čiginas 2020). Because of this, even if between-group inequality increases (that is, the income gap between those with a tertiary degree and those without rises), the higher share of educated results in lower the within-group inequality in Lithuania. This finding contrasts Lemieux (2006) results obtained for the US, where the more educated (albeit defined at a more disaggregated level) tended to be more

unequal than those who were educated, in which case more education means less equality. One possible explanation for different levels of within inequality could relate to more homogeneous universities in Lithuania than in the US, resulting in more equal outcomes. For example, all but one university in Lithuania are public, and the government (now and in soviet times) provides heavy subsidies to enter. However, this should be explored further. Similarly, the stronger effect for the top of income distribution could be examined further. This would be problematic if higher income families have certain privileges of obtaining education. Contemporary reports do not suggest unequal access to higher education per se, but there are signs that supply of early education is unequal, which could allow wealthier families to access it, and then subsequently find it easier to enroll into higher education (OECD 2017a).

Decomposing changes in inequality and redistribution

Here we quantify the contributions of the four factors as well as their interactions to the changes in income inequality and net redistribution. That is, we decompose Table 2.4, found in section 2.4.2. Table 2.5 shows the contributions to the changes in Gini of disposable income and the Gini of market income. The contributions to the changes of the net redistribution, which is the difference between Gini of market income and the Gini of disposable income, is found in Table 2.6. All decompositions are based on direct effects, as shown in Section 2.3.3. As a robustness check, the decompositions based on the Shapley value can be found in Table A7 in the Appendix.

Starting with the contributions to the changes in disposable income inequality, we can see that the effects of the four factors were heterogeneous. In terms of the size of the effect, the contribution of the prices and returns factor was the most important, and the totality of the effect is concentrated in the second period. Over the period of economic recovery, the Gini of disposable income rose by 3.2 pp due to higher prices and returns. This number is consistent with Figure 2.7c, which shows that the upper tail of the disposable income distribution benefited significantly more than the lower tail. Demographic changes were another important contributor to the growing income inequality in Lithuania. Unlike the effect of prices and returns, trends in the Lithuanian demographic situation appear to be secular and independent from the business cycle conditions: the impact in both periods is similar quantitatively, amounting to a total contribution of 1.3 pp to the Gini index.

The remaining two factors acted in the opposite direction and were responsible for taming the growing income inequality due to the returns and the demographic effects. Specifically, changes in the tax and benefit system managed to counter half of the increase in market income inequality. Its contribution to *reducing* the Gini of

disposable income amounted to 2.0 pp, and the effect is concentrated in the period of financial crisis. As discussed in Section 2.2, no additional measures were implemented during the years of economic expansion, as most of the transfers, such as pensions, were frozen.

This means that the amount of redistribution remained the same, and the tax and benefit system was not able to accommodate rising disposable income inequality during the economic upturn. Finally, the labour market structure is shown to make a smaller but also significant contribution to lowering income inequality, which occurred during the first sub-period.

Moving to market income inequality, one can observe that it has grown significantly over the whole period, but most of it occurred during the financial crisis of 2008: the Gini grew by 4.2 pp, with 95% of the growth concentrated in the first sub-period.[9] Interestingly, demographics was the most important factor, contributing to about half of this increase. Going back to Table 2.2, this was a period when the share of married households decreased while the number of those with tertiary education increased, suggesting that household and education composition was behind this rise in inequality.

Not surprisingly, the effect of prices and returns in the labour and capital markets on income inequality portrays procyclicality. The effect of prices and returns was negative during the crisis years (-0.7 pp) but positive and significant in size during the years of economic expansion (1.3 pp). Looking at the whole period, the two phases cancel each other out, and the total effect is only 0.6 pp.

Changes in labour market structure appear to be the only factor that has reduced market income inequality substantially, and the effect is mainly concentrated in the first sub-period. It is important to note that the component unexplained by our methodology amounts to a significant share of the total change, especially so during the first sub-period. This implies that factors not modelled by our methodology (e.g., regional composition of workers and jobs) also played a role.

9. The small effect of the tax-benefit transformation on market income inequality is due to adjustments to minimum wages, which are included in the taxes and benefit transformation. Regarding the compliance rules implemented in EUROMOD, we are not bringing in second-order non-compliance and we are not assuming differential compliance. We assume, thus, no tax-evasion, compliance in benefit take-up, and compliance with minimum wage regulations.

Table 2.5: Decomposition of changes in equivalised income inequality

	Gini Disposable Income			Gini Market Income		
	2007-2015	2007-2011	2011-2015	2007-2015	2007-2011	2011-2015
Total change	**0.021**	**-0.008**	**0.029**	**0.042**	**0.04**	**0.002**
Demographics	0.013	0.008	0.006	0.020	0.017	0.003
Labour Market Structure	-0.012	-0.017	0.005	-0.015	-0.015	-0.001
Prices and Returns	0.030	-0.002	0.032	0.006	-0.007	0.013
Taxes and Benefits	-0.020	-0.021	0.000	0.002	0.004	-0.001
Interactions	0.017	0.002	0.016	0.003	0.009	-0.006
Unexplained	-0.008	0.022	-0.030	0.026	0.033	-0.007

Notes: Columns indicate the time period over which statistics where calculated (e.g. 2007-2011 refers to the change from 2007 to 2011).

Next, we examine net redistribution to assess whether the changes in the income distribution were due to changes in policy design or changes in the distribution of market incomes. Here, market incomes refer to all factors (except the tax and benefit factor) plus interactions and the unexplained residual. We decompose the changes in the redistributive indices marking the transition from market to disposable income. Specifically, we decompose redistributive indexes into total a) net redistribution, b) benefit redistribution (benefit regressivity, average benefit rate), and c) tax redistribution (tax progressivity, average tax rate). Our infrastructure allows us to assess to what extent the observed changes in these indices are due to changes in policy design over time, as captured by the tax-benefit effect in Table 2.6. Controlling for changes in market income distributions between 2007 and 2015, we find that net redistribution increased. The increase was driven by an increase in benefit redistribution, as seen in Table 2.7a, where all the increase took place in the period 2007-2011. In contrast, changes to the tax system reduced income redistribution. Again, the policy changes predominantly took place in 2007-2011.

We then split the benefit and tax redistribution further into average tax/benefit rate effect and a progressivity effect with the help of the STATA package compiled by Peichl and Kerm (2007). The results are found in Table 2.8a. From this, we see that the benefit redistribution increased due to higher benefit generosity. Higher benefits were paid out, particularly in the period of 2007-2011. Had market incomes not risen in that period also, inequality would have been even lower. Although benefits became less regressive, benefit redistribution increased. As shown in Černiauskas and Čiginas (2020), this is because it is more effective to change benefit level than benefit regressivity, as benefits are already regressive. Although tax rates did become more progressive (partly due to rising tax-exempt amount of income), the level

of taxes decreased substantially due to lower tax rates. As a result, taxes became less redistributive.

The tax-benefit system during the period from 2011 to 2015 did not generate sufficient redistribution for prevention of income inequality, which resulted from rapid increases in market incomes. Comparatively low levels of benefits and reluctance to introduce an increasingly progressive income tax were the main factors of rising income inequality.

2.6. Concluding remarks

This chapter studies the drivers of changes in the income distribution in Lithuania from 2007 to 2015 by adapting a methodology developed by Sologon et al. (2020). We assess the role played by changes in the labour market structure, the economic returns in labour and capital markets, the demographics, and the economic policy related to tax and benefit rules. The case of Lithuania is especially interesting, given the country's recent transition from a planned economy to a market one, its ongoing convergence to the EU-15, and large fluctuations in disposable income over the business cycle. During the period under discussion, the Lithuanian economy experienced a global financial crisis which significantly affected household disposable income, a series of tax and benefit reforms, and a changing demographic structure.

Income inequality reached unprecedented levels as a result. To address this challenge, one must first understand the factors that contribute to income inequality and determine whether the tax and benefit system in place is able to reduce it.

Our results suggest that the growing returns in the labour and capital markets, as well as large structural changes in the demographics of the population, played the main role in explaining the observed increase in income inequality. Changes in the tax and benefit system reduced income inequality overall, but only during the period 2007-2011. In particular, the benefit system became more redistributive because of larger benefit pay-outs that were increased in this period. By the year 2011, those who lost work had access to relatively high unemployment benefits, parental benefits, sick leaves, old age pension, and other benefits, as compared to 2007. However, benefits only slightly increased thereafter, while in some cases (e.g., due to increasing pension age) fewer benefits have been handed out altogether. Tax rates have been lowered in 2007-2011 and were not raised in the later period. As a consequence, disposable income inequality increased sharply over the next period. Although the returns effect was the main contributor to increasing income inequality, especially during 2011-2015 period, other important factors played a significant role as well. Our results show that the demographic effect persistently increased income inequality over the

analysed period. Specifically, we found that declining marriage rates were mostly responsible for the increase.

Several lessons can be drawn from the analysis of the Lithuanian economy during 2007-2015. First, implementing fiscal consolidation by reducing the generosity of the benefits system can have important negative distributional consequences. Falling regressivity of benefits during the economic expansion in the aftermath of the financial crisis was one of the main contributors to increasing disposable income inequality in Lithuania. Second, the Lithuanian tax system is designed in such a way that its progressivity declines in response to unequal growth in income distribution. As the economy continues to converge towards EU-15, we can expect this mechanism to continue unless the tax system is reconsidered. Third, changing demographic composition of the population can have important consequences on the income inequality as well. As marriage rates continue to decline (most likely due to a change in the preferences with respect to the size of the household), we can expect to see rising income inequality in the future.

Table 2.6: Decomposition of the changes in redistribution

(a) Net redistribution, benefit redistribution and tax redistribution

	Net redistribution			Benefit redistribution			Tax redistribution		
	2007-2015	2007-2011	2011-2015	2007-2015	2007-2011	2011-2015	2007-2015	2007-2011	2011-2015
Total	**0.021**	**0.048**	**-0.026**	**0.021**	**0.045**	**-0.024**	**-0.010**	**-0.007**	**-0.003**
Taxes and Benefits	0.023	0.024	-0.001	0.027	0.028	-0.002	-0.008	-0.008	0.000
Market incomes	-0.001	0.024	-0.025	-0.006	0.017	-0.022	-0.002	0.001	-0.003

Notes: Columns indicate the time period over which statistics where calculated (e.g. 2007-2011 refers to the change from 2007 to 2011).

(a) Detailed tax and benefit redistribution

	Benefit regressivity			Average benefit rate			Tax progressivity			Average tax rate		
	2007-2015	2007-2011	2011-2015	2007-2015	2007-2011	2011-2015	2007-2015	2007-2011	2011-2015	2007-2015	2007-2011	2011-2015
Total	**0.034**	**0.078**	**-0.044**	**0.037**	**0.066**	**-0.029**	**0.017**	**0.055**	**-0.038**	**-0.070**	**-0.077**	**0.007**
Taxes and Benefits	-0.017	0.015	-0.032	0.070	0.060	0.010	0.029	0.033	-0.004	-0.071	-0.071	0.000
Market incomes	0.051	0.062	-0.012	-0.033	0.006	-0.039	-0.012	0.022	-0.035	0.000	-0.006	0.006

Notes: Columns indicate the time period over which statistics where calculated (e.g. 2007-2011 refers to the change from 2007 to 2011).

3. STATUTORY, EFFECTIVE AND OPTIMAL NET TAX SCHEDULES IN LITHUANIA

3.1. Introduction

Although it is widely accepted that taxes are necessary to finance government expenditures and social transfer programs, there is a great deal of disagreement concerning who should be paying these taxes. Regarding labor income taxation, the optimal tax literature considers three factors (income distribution, labour elasticities and society's preferences) when determining who should pay taxes and how much should they pay (see, e.g., Saez 2001). Oftentimes, however, statutory tax rates - the rates that are inscribed in the law - are smaller and less progressive than optimal ones (Saez 2002). Additionally, tax avoidance and fraud lead to further divergence between taxes that are actually paid (i.e. effective tax rates) and optimal ones. These three concepts (optimal, effective and statutory rates) are interrelated in a complex way: optimal taxes inform us about the desirable rate structure, whereas effective rates show how the tax system effectively taxes people based on rules set out by statutory rates, as prescribed by law. The interplay between these concepts is key to addressing urgent public policy questions: how do statutory rates effectively impact on individuals?; how does the tax system fare as compared to optimality principles?; etc.

We perform this analysis for the case of Lithuania. Our objective is twofold: first, we establish the extent to which the real world labour tax structure of the country is aligned with lessons from the optimal tax literature. Second, we compare the three schedules for employees and the self-employed. Governments utilise the tax system to encourage various types of behaviours - including the choice of self-employment. While this may have favourable effects on the labour supply or taxable income of those concerned, it may cause additional difficulties. For example, the self-employed usually face lower statutory income tax rates and are more likely to evade taxes as compared to employees, which leads to smaller government coffers and questions of social justice (Milanez and Bratta 2019). Lithuania is a particularly interesting case study in this regard. First, it applies rather distinct rules for employees and the self-employed. Second, it enjoys good survey and administrative data availability.

This chapter relates to two bodies of tax literature. The first is the optimal tax literature, particularly the sub-branch which compares optimal tax schedules with statutory ones. The literature of optimal taxation started with partial equilibrium models based on individuals, most notably Mirlees (1971). He demonstrated that higher marginal tax rates generate labour responses that cause employees to spend less time

in employment. The Mirlees model was modified by Saez (2001) by replacing theoretical labour responses with observable income-dependent labour supply elasticities. This methodology was first used to argue that optimal gross income (which excludes social contributions) tax rates of top incomes in the USA could exceed 50%. More recent studies have replaced the labour elasticity with elasticities of taxable income. These are considered broader than labour elasticity, as they include other behaviour responses, such as tax evasion and avoidance, and not only labour supply. Klemm et al. (2018), also using Saez (2001) methodology and estimates of taxable elasticities, suggest that optimal income tax rates for top incomes exceed 60% for 27 global countries. A slightly modified version by Saez (2002) considers optimal tax rates at the bottom of the income distribution, by incorporating labour market responses at the intensive and extensive margin throughout the income distribution.

Subsequent authors have shown that optimal taxes rates differ, depending on the optimal tax schedule model. For example, Immervoll et al. (2011) extends Saez (2002) model (which includes only individuals) to couples, and suggests lower taxes on secondary earners versus primary earners for a sample of 15 EU countries. Additionally, the income tax schedule also depends on the existence of non-income tax schedules. For example, Huang and Rios (2016) shows that countries with a non-linear income tax and a linear non-income tax (such as the value-added tax in Russia) should have lower marginal income tax rates. However, if a country also exhibits high income under-reporting, then marginal income taxes should be lifted again. Using general equilibrium models, other authors such as Heathcote et al. (2017) find that incorporating skill investment and public good provision suggest lower progressivity (although high poverty rates that prevent skill investment undermine such claims). There are also models that look at employment and self-employment simultaneously, for example, Zawisza (2019). This model incorporates own-elasticities to declare employment or self-employment income and evaluates the cross-elasticities of switching between employment and self-employment. He found the elasticities of the self-employed to be three times higher than the elasticities of the employed in Poland. The lack of consensus leaves the researcher puzzled as to which model to use, but the lack of elasticity and other parameter estimates constrains the model choice to that of Saez (2002). This means that we work with the same elasticity for the self-employed and employed, which may lead to an over-estimation of the optimal tax schedule for the self-employed.

Furthermore, the optimal tax literature has attempted to analyse different tax and income concepts. Mirlees (1971), Saez (2001), and Immervoll et al. (2011) focused on income tax and employment income. Saez (2002) considered net taxes (income

taxes minus public benefits), which means that individuals take into consideration their income taxes and (instantaneous) benefits when making employment decisions. This is useful when analysing optimal taxes at the bottom of the income distribution, since high public benefits (such as unemployment benefits) may discourage work as much as high taxes. However, for most OECD countries (OECD 2019), income tax constitutes a small part of the ”tax” burden. For them, social contributions are both higher and not necessarily actuarially fair, meaning that this, too, can be seen as a tax.

We also relate to the tax literature which examines statutory and effective tax rate differences between employees and the self-employed. Studies focusing on labour taxation show that statutory tax schedules for employees (OECD 2019) and for the self-employed (Milanez and Bratta 2019) vary across OECD countries and across different households types within countries. Estimates of effective tax rates largely come from the tax evasion literature, which implicitly compares statutory and effective tax rates, although the focus is often on the individual. The closest work to our chapter is by Matsaganis et al. (2013), who estimate income misreporting in Greece of wages and of self-employment income in the period 2005-2009. They find that about 43% of self-employment income was under-reported in 2009 and that the tails of income distribution under-reported income more often. They do this by comparing EU-SILC data on income coupled with administrative data on income, and use EUROMOD to streamline the definitions. They face the challenge of having different samples of people in the EU-SILC and the administrative records. In a different study, Johns and Slemrod (2010) finds that top income-earners tend to avoid taxes, leading to lower effective tax rates in the USA, and Alstadsæter et al. (2017) find that the most wealthy Scandinavians also exhibit a similar trend. Even though the evidence suggests that employees do evade income, up to 20% of the top incomes in Estonia do so (Paulus 2015), the self-employed tend to engage in tax evasion and avoidance substantially more (see, e.g. Baldini et al. 2009; Slemrod 2016) with some estimates showing that more than half of income may be concealed from the authorities (Artavanis et al. 2016).

We find that the three net tax schedules diverge much more for the self-employed than for employees. In fact, the optimal, statutory and effective tax rates for employees largely coincide for all but the tails of the income distribution. In contrast, for the self-employed, the effective tax rates are well below the statutory tax rates, while statutory rates are also below the optimal rates for most of the income distribution.

The chapter is structured as follows. In section 3.2, we present the data sources and the definitions used throughout the chapter. The following three sections cover the statutory, effective and optimal net tax schedules. The results are presented and

discussed in section 3.6, while the conclusions, recommendations and limitations are presented in section 3.7.

3.2. Data and definitions

We use the European Union Statistics on Income and Living Conditions (EU-SILC) dataset to estimate statutory, effective, and optimal net tax schedules for Lithuania. This is the only publicly available source of data with sufficient information for our analysis in one dataset for Lithuania, as it contains key information on employment income, taxes, benefits, household composition and information that can help to classify individuals as employees or self-employed. The yearly EU-SILC has been running since 2004 and is the reference for comparable data on personal income in Europe. Each year, around five thousand households encompassing around 10 thousand household members over 16 years of age who agree to share information on their incomes are included[1]. We pool data from surveys carried out in 2015-2016, which contain income data (reference years) of 2014-2015. While the data is well explained on the Eurostat website[2], some features are mentioned here.

Firstly, certain income components are only available for the household level in the survey. Notably, income tax and social contributions are calculated at the household level. This restricts the analysis to the concept of household (equivalised) income rather than individual income, which can be considered a blessing or a curse. On one hand, the literature suggests that individuals make economic decisions taking themselves as well as their household members into consideration (see, among others, Vogler and Pahl 1994). For example, the incomes of all household members comprise a common budget constraint (Chiappori and Meghir 2015), thereby influencing each household member's behaviour. Additionally, some benefits are only granted at the household level (e.g. social assistance benefit), making the allocation of this benefit to any specific household member artificial. Nevertheless, each household member has his/her own preferences and a typically unequal control of the household's budget, with evidence suggesting that decisions within households are rarely joint and more often taken by specific household members (Pahl 1995)[3]. To partly account for the limitations of working with household data we carry out an analysis of single's households as a robustness check, but our results still hold.

1. For 2015 reference year, 5142 households out of 6161 households participated in the survey-interview. This means that at least one respondent was willing to fill in the survey on behalf of the household. For those 5142 households, information on all household members was collected.

2. https://ec.europa.eu/eurostat/web/microdata/european-union-statistics-on-income-and-living-conditions

3. Future studies should also compare them with net tax schedules for individual incomes or the interactions between individuals within a household.

Second, EU-SILC has a large survey component, but, since 2012, Lithuania has made heavy use of register (administrative) data. The State Social Insurance Fund Board data and the State Tax Inspectorate under the Ministry of Finance of the Republic of Lithuania data have been linked to sample data and used for checking cash or near-cash employee income, social insurance contributions and taxes on income, as well as old age benefits. Maternity and maternity/paternity allowances, care allowance, social assistance, old-age, and survivor's pensions have been taken from the administrative data. See country report[4] for more information. Register data is directly imputed from the registers for households which agree to participate in the survey. If register data is not available, then survey data is used. In the case of income, particularly employment income and income from self-employment, data is taken from both administrative and survey sources, and the greater value of the two is used. This "true" income is later used to estimate statutory taxes. In this way, we can observe actual incomes and not just income that has been reported to the tax authorities. In the case of taxes and benefits, we mainly rely on administrative data.

Third, survey weights are used to partly adjust for probability of selection, non-response and, as appropriate, to adjust the sample to external data. Currently, the sample is adjusted for demographic and geographic external data only. The weights are further adjusted according to Eurostat 2018b: weights of household members who are over age 16 are scaled up by distributing weights of those under age 16. For most of the calculations, we only considered households that had at least one non-student household member aged 18-62. This means we kept one observation per household whose weight was the sum of the individual weights in that household.

Fourth, there is evidence that income inequality is underestimated in EU-SILC (Hlasny and Verme 2018; Törmälehto 2017). Callan et al. (2020) find that in Ireland only the top 1% of income is missing from household surveys as compared to register data, after accounting for concept differences. In line with this, Navickė and Lazutka (2016) show that capital income is under-reported for Lithuania in EU-SILC, which is usually concentrated at the top of income distribution, while other income components are much less under-reported. A study of Estonian Household Finance and Consumption Survey by Meriküll and Room (2019) showed that the rich as well as the poor usually do respond to surveys and so unit non-response is a smaller problem, but income is under-reported due to item-non-response. That is, the richer individuals do participate in household questioners but tend to avoid questions related to specific income/wealth questions. Since employee income, taxes and social contributions for those who agree to participate in the survey are taken from regis-

4. https://circabc.europa.eu/faces/jsp/extension/wai/navigation/container.jsp

ters in Lithuania, item non-response should be a smaller problem here. One major exception is self-employment income, which is not imputed from registers and has been often found omitted in the mentioned study. Unfortunately, Meriküll and Room (2019) do not succeed in replicating register data with survey data using data imputation techniques, citing lack of common support as the key issue. Furthermore, as self-employment income is often under-reported to tax authorities, such imputation are unhelpful in the first place. As we focus on labour rather than capital income and we see self-employment income as problematic to weight, we refrain from reweighing our data.

Finally, EU-SILC is compatible with Euromod. Euromod is a European tax-benefit simulator which takes in EU-SILC data and calculates how much tax each individual should pay or how many benefits he should receive based on his market income and other characteristics (e.g. age, whether he has any dependants, employment status, etc.). This allows us to estimate statutory tax schedules. It should be noted that while EU-SILC is used for EUROMOD, there are adjustments made in the process[5].

This data and Euromod allows us to estimate the three net tax schedules. Specifically, we estimate household equivalised net taxes as a share of household equivalised gross employment income. Let us explain each term in more detail. *Gross employment income* is defined as yearly gross employee and self-employed income (including social contributions of the employee, the self-employed and the employer). *Net tax* is the difference between taxes paid and public transfers received. Gross employment income minus net taxes is net labour income. The unit of observation is a household to which we allocate an equivalised income[6]. To obtain equivalised income, we first sum the incomes of all household members for a given household. Then, we adjust the sum by an OECD-modified equivalence scale, where 1 is attributed to the first household member, 0.5 to the second and each subsequent person aged 14 and over and 0.3 to each child aged under age 14. Henceforth, any reference to income or taxes in this text relates to equivalised household income and taxes. Finally, we construct a working sample which includes only households with at least one member who is not a student and is between 18 and 62 years of age. This allows us to focus on the

5. For example, 20 household members who were born after the income period were removed in the EUROMOD 2015 and 2016 input files. This meant that survey weights add up to different totals and equivalence scales also differ for those households. Unfortunately, the household id's differ in the two data sets and we were not able to identify which are the household members that should be removed from EU-SILC to generate the same weights.

6. The alternative would be to have different tax rates for different types of households (e.g. single, married, married with children) as done in Guner et al. (2014), but using equivalised income allows us to have a single summary statistic and worry less about sample size.

Table 3.1: Average yearly equivalised income and net taxes in Lithuania, euro

Variable	Full sample (population)	Working sample (18-62, non student)
gross employment income	7663	8952
net taxes (minus)	1045	1944
net labour income	6618	7008
number of households	9657	6459

Data for 2014-2015 income reference years comes from the EU-SILC dataset. Gross employment income and net taxes include employer's and employee's social contributions.

working-age population and excludes pensioners - implicitly also reducing the role of these benefits in household income. We do not remove them completely, because many households have at least one pensioner or student, and they contribute to the household income. Income and net tax statistics from EU-SILC for 2014 and 2015 reference years for the full sample, which represents Lithuania's population, and the working sample is summarised in Table 3.1.

We focus mainly on gross employment income and net labour income. These variables relate most closely with one's work incentives[7]. EU-SILC has more income variables that also relate to work incentives, but we refrain from discussing those[8] Gross employee income is defined as the total remuneration in cash payable by an employer to an employee in return for work done by the latter during the income reference period, plus the employer's social insurance contribution. Gross self-employment income is defined as the income received during the income reference period by individuals, for themselves or in respect of their family members, as a result of their current or former involvement in self-employed work. Self-employed work covers those jobs where the remuneration is directly dependent upon the profits (or the potential for profits) derived from the goods and services produced (where own consumption is considered to be part of profits).

7. Other possible strategies could include looking at taxes only or net taxes, taking into account inter-temporal benefit accrual such as for pensions. We leave this aside for future research.

8. For example, disposable income includes all the variables that fall under net labour income as well as other incomes, such as private transfers, and other taxes, notably capital tax. These variables play a minor role in this survey and do not impact the results. EU-SILC also includes several non-cash items that may have a larger impact on income and decision-making, but it is not clear to what extent this can be taxed. For example, non-cash items, especially imputed rent, which is the approximate income one would receive if one was to rent his/her residence, constitutes about 18% percent of gross employment income of the working sample.

We include social contributions and all benefits in our definition of net tax in order to better reflect the incentives Lithuania's households face when participating in the labour market. Social contributions constitute a relatively large share of labour costs as well as the biggest source of revenue for the government (11.9% of GDP in 2015 according to Eurostat, while income tax makes up only 5.4%, even lower than VAT - 7.7%). Although contributions are used to finance social benefits, and could be seen as tax-neutral, there are also reasons to think of them, at least in part, as a tax. In their book, Frölich et al. (2014) argue that some people may either not want the benefits associated with social contributions or want less of them, in which case only the difference between the desired benefits and the paid contributions should be considered as tax. For example, using USA data Chetty et al. (2016) finds that poor people tend to live shorter lives, meaning they have less chance of getting any benefits despite their contributions. Knowing that the largest share of social contributions is to insure against old-age, not paying social insurance contributions may be a very rational response for these people. In such cases, people may either work less if the contributions are perceived as too high or turn to informal work to avoid paying them (Frölich et al. 2014). Since we cannot identify the part of social contributions that are paid willingly, or how much of other taxes people willingly pay in exchange for public goods and services, we include social contributions into our definition of tax. We include all benefits (old-age, sickness/health, disability, family, unemployment and other benefits) into the definition of equivalised income.

Detailed statistics of income and net taxes as a percent of gross employment income are shown in Table 3.2. 88% of gross employment income is derived from gross employee income, with the residual derived from self-employment income. Public transfers increase income, resulting in 21% higher gross labour income than gross employment income for the full sample, but only 13 % in the working sample. Public transfers increase income by less in the working sample because we exclude a large share of pensioners together with their old-age public transfers. Other public transfers[9] still constitute a sizable share of income in the working sample. Tax on income and social insurance contributions reduce gross employment income by just over a third. As a result, net labour income is 86% of gross employment income on average (resp. 78% of working sample). Therefore, the net taxes as a percent of income gross employment income is 34% in the working sample.

As Lithuania's tax system treats employees and the self-employed differently, we also examine different types of households. In total, there are three non-overlapping

9. Disability benefits and family/children related allowances each constitute about a third of the other public transfers. While unemployment benefits only make up 10% of other public transfers.

Table 3.2: Detailed equivalized income and net tax in Lithuania, % of gross employment income

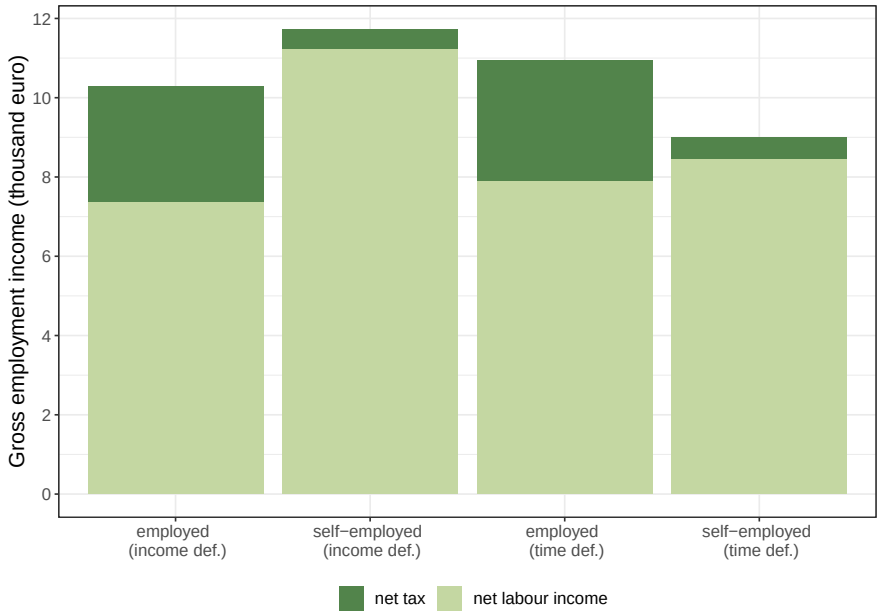
Variable	Full sample (population)	Working sample (18-62, non student)
gross employee income	88	88
gross self-employed income	12	12
gross employment income	100	100
old-age public transfer	12	4
other public transfers	9	8
gross labour income	121	113
tax on income and social insurance contributions	-34	-34
net labour income	86	78
number of households	9657	6459

All variables are in percent of gross employment income. Data for the income reference years of 2014-2015 comes from EU-SILC. Gross employment income and it's components include employer's and employee's social contributions.

groups of households: employees, self-employed and other. We use two definitions to define a household. The preferred is the *Income* definition, where we sum household members' gross labour income components (employee, self-employed and public transfer income) in a household and see which of the three components is dominant. Additionally, employee/self-employed households must have received or made a loss of at least 10 euros of gross employee/self-employed income in the reference year; otherwise, they are classified as "other". The alternative is the *Time* definition, where the total household member's months spent in an activity is considered. Specifically, each household member had to identify his/her main activity in each month of the income reference year, be it an employee, self-employed or other. We then sum all the months of all household members, note which is the largest, and label that household accordingly.

Using the income definition results in a higher net labour income of the self-employed households, as summarised in Figure 3.1. Under the income definition, self-employed households receive around 14% more gross employment income than employee households, but pay only 17% of the net taxes that employee households pay. This results in 52% higher net labour income of the self-employed as compared to employees. Under the time definition, the self-employed pay less net taxes than employees, but they also earn much less gross employment income. More generally, while self-employment is not the activity that households report spending most of their time on collectively, it is the one that generates the largest net labour income. In-

Figure 3.1: Average equivalised yearly income in Lithuania for 2014-2015 reference years



Bars represent average equivalised income for employee and self-employed households under two grouping definitions: income definition and time definitions. The sum of equivalised net labour income and equivalised net tax is equivalised gross employment income. Calculations are based on the working sample. There are 264 households that fall under the time definition for the self-employed and 545 under the income definition (4566 and 4889 for the employees respectively).

deed, only 3.3% of households report spending most of their time in self-employment, while 7.4% report gaining most of their gross labour income from self-employment. This is largely because over half of household members who earn their own self-employment income also earn employee income, and 60% cohabit with someone who earns employee income. Those who earn their own employee income are much less likely to earn self-employment income (10%) or cohabit with someone who does (14%).

Finally, we compute average and marginal tax rates throughout the chapter. The formula for the average tax rate for the gross employment income decile $i = 1, 2, ..., 10$ is

$$atr_i = \frac{\sum_{k=1_i}^{n_i} taxes_k * w_k}{\sum_{k=1_i}^{n_i} income_k * w_k}$$

defined by the sum of taxes paid by households $k = 1_i, 2_i, ..., n_i$ and n_i would mean the n^{th} household member of decile i. We adjust the distribution of taxes using survey weights w_k. Then, we divide the weighted taxes paid by the income of all households multiplied by their weights in decile i.

Similarly, marginal taxes for gross employment income decile $i = 2, 3, ..., 10$

$$mtr_i = \frac{\sum_{k=1_i}^{n_i} taxes_k * w_k - \sum_{k=1_{i-1}}^{n_{i-1}} taxes_k * w_k}{\sum_{k=1_i}^{n_i} income_k * w_k - \sum_{k=1_{i-1}}^{n_{i-1}} income_k * w_k}$$

3.3. Statutory net tax schedule

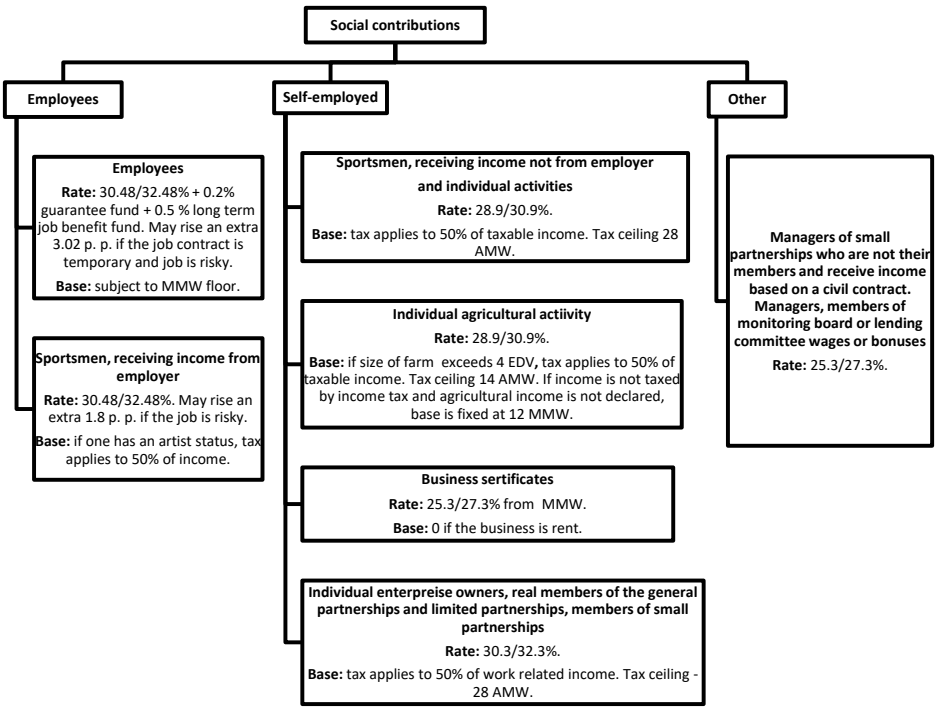
We proxy the characteristics of the statutory net tax schedule in Lithuania by applying the tax and benefit rules applicable in the country to the observations from EU-SILC. Specifically, we utilise Euromod - a tax and benefit simulator - to estimate the amounts of taxes and benefits that would be due if we simply apply the statutory rules to the data at hand for all households, and for the separate groups of employees and self-employed. We use the income definition to allocate households into employee and self-employed throughout this section. Finally, we present statutory average tax schedules for Lithuania for the two groups.

Lithuania's tax and benefit system is complex. First, it incorporates various taxes, social contributions and benefits. We consider income tax, all social contributions, and a wide range of benefits. Most benefits, including pensions, are related to household members' previous income, although various coefficients, ceilings and floors ensure some income redistribution in the system. Second, there are various household-member and household-level characteristics that determine how much net taxes a household member should pay. This results in a wide range of net taxes to consider.

Figure 3.2 presents the statutory social contribution rates and bases that we derive for the household member in our sample[10]. Different contribution rates and bases are applied to employees and the self-employed; gross employee taxable income is subject to a monthly minimum wage (MMW) floor, while most forms of self-employed income benefit from a 50 percent tax base reduction. Therefore, the effective taxes paid by the self-employed can be much smaller than those paid by employees. A likely possible weakness of our data is that some tax-relevant information for properly applying the statutory rules may not be factored in, hence inducing a potential bias of an a priori unknown sign. For example, the self-employed may benefit from carried-forward losses, a factor that would effectively further widen the difference in statutory rates between employees and the self-employed[11].

10. These social contributions were effective before a large tax reform that took place in 2019.
11. Other examples are tax exemptions for specific disabilities, economic activity or information that

Figure 3.2: Statutory social insurance contributions excluding the statutory health insurance contributions prior to 2019 reform



MMW - monthly minimum wage, AMW - average monthly wage. Sources: based on state tax inspectorate of Lithuania.

Euromod and EU-SILC dataset for Lithuania is able to estimate the majority of taxes and a portion of benefits[12]. For example, family benefits that depend on the number of children and their ages are simulated. Furthermore, simulations are made for a number of contributory (social insurance-based) benefits, such as maternity leave or benefits assigned to low-income household members. A number of benefits with entitlement rights dependent on contribution history (i.e. pensions, sickness benefit,

is not collected in EU-SILC survey. If these were fully accounted for, the statutory rates would be lower and closer to the effective tax rates. It is also likely that these specificities will be more important for the self-employed.

12. Euromod input files are slightly modified versions of EU-SILC data. In the case of Lithuania, 10 household members that were not yet born in the reference period were dropped in the 2016 and 2015 survey. Euromod also reads country-specific files which describe the statutory taxes and benefits of those countries that are then applied on the input files. More information on Euromod can be found at http://www.euromod.ac.uk and in Navickė and Čižauskaitė (2018) in particular.

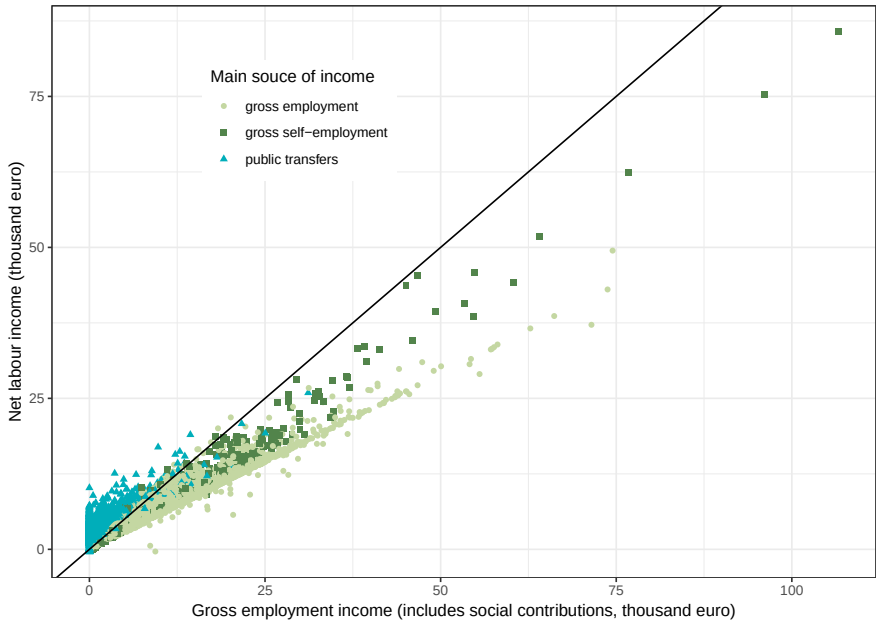
disability benefits, etc.) are not simulated due to the lack of data on previous employment history and salaries received, some event occurrence (i.e. disability or accident at work), or lack of information on previous partner entitlements (i.e. survival pensions). In those cases where potential benefits are not simulated, they are replaced with effective benefits from the input file. We run the simulations at a household-member level, after which we aggregate to household-level and adjust incomes by an equivalence scale. Finally, we construct a working sample by keeping households with at least one household member who is 18-62 years of age and is not a student. We use EUROMOD version's H1.0+ 2014 and 2015 Lithuania's system files on LT_2015_a1 and LT_2016_a2 input data respectively.

The relation between two simulated variables is plotted in Figure 3.3. On the x-axis is the gross employment income, and on the y-axis is net labour income (gross employment income minus net taxes). The figure thus links the mechanisms which transform gross employment income into net labour income. The diagonal line represents no transformation: what a household earns from employment income becomes its net labour income. Anything below the diagonal line refers to income that is taxed away. Anything above the diagonal line means that the household received public transfers that exceed paid taxes. The colours and shapes of the points represent the groups according to their main source of income: gross employee income, gross self-employment income, and public transfer income.

Many households that receive hardly any gross employment income are legally entitled to substantial public transfers which raise their net labour income above the diagonal line. This is largely because some or all households are able to apply for old-age benefits or disability benefits. Once households start earning some gross employment income, their net labour income becomes dispersed and their main source of income is increasingly likely to be employment income. As gross employment income rises, the majority of households tend to be below the diagonal line, as they have to pay taxes and receive fewer benefits.

The self-employed households receive higher net labour income as compared to the employee households, especially at higher gross employment income levels. This is because employees are legally subject to higher statutory average tax rates than the self-employed for the same level of gross employment income. In part, this is due to the lower taxable base of the self-employed. Furthermore, the self-employed have access to more types of tax treatment. For example, the self-employed may purchase business certificates. This requires their holders to pay a one-off fee determined by the municipality if they receive under 4500 euro from the activity. For a couple with two

Figure 3.3: Statutory equivalised incomes of households grouped using the income definition in Lithuania



Equivalised gross employment income and net labour income is in thousands of euro per year. Households are allocated to groups according to the income definition for 2014-2015 income reference years and are represented by dots in the graph (see Section 3.2). The diagonal line illustrates that household employment income is equal to net labour income. Any dot above the diagonal line illustrates that the household receives additional benefits, while dots under the horizontal line means that the household pays additional taxes or social contributions..

business certificates, this could lead to $9000/1.5 = 6000$ euro equivalised income that is barely taxed, while other types of incomes could be declared under different activity forms or taxed at a different rate thereafter.

Not only do the self-employed earn more net labour income on average due to lower taxes, but self-employed households are concentrated at the top of the income distribution. For example, in the bottom 20% of the net labour income distribution, only 5% of households can be considered self-employed under the income definition. The share of households that are self-employed almost triples in the top 20% of the income distribution, and reaches 30% for the top 5% in Lithuania. Such a distribution of self-employed households also encourages us to make stronger claims on the richer self-employed rather than the poorer ones. Nevertheless, the data suggests that the self-employed are faring worse at the bottom of the income distribution. As seen in Table 3.3, employee households grouped using the income definition in the second

(pseudo) decile receive 1.77 thousand euro net labour income, and do not pay any net taxes. The self-employed receive less net labour income (0.83 thousand euro) and pay more net taxes. This is because the self-employed receive fewer benefits as compared to employees at the bottom of the income distribution, but they pay similar taxes.

Table 3.3: Statutory equivalised gross employment income, net taxes and net labour income in thousand euro per year.

percentile	gross employment income	net taxes			net labour income		
		all	employees	self-employed	all	employees	self-employed
0-7	0.00	-2.25			2.25		
7-20	1.41	-1.12	0.25	0.42	2.53	1.72	1.19
20-30	3.59	0.07	0.76	0.70	3.52	2.86	2.95
30-40	5.10	0.78	1.29	1.01	4.32	3.81	4.14
40-50	6.61	1.54	1.80	1.32	5.07	4.79	5.27
50-60	8.22	2.29	2.48	1.86	5.93	5.73	6.34
60-70	10.08	3.03	3.16	2.83	7.05	6.96	7.03
70-80	12.36	4.08	4.28	2.91	8.28	8.08	9.39
80-90	15.69	5.67	5.86	4.09	10.02	9.84	11.50
90-100	27.65	10.29	10.91	7.02	17.36	16.65	21.21

Data is sorted according to equivalised gross employment income (includes social contributions). Net taxes include tax minus public benefits (public transfers). Net labour income is gross employment income plus benefits minus taxes. We report averages of percentile ranges. Gross employment income is taken from EU-SILC, while net taxes are estimated by Euromod, which takes into account various individual and household characteristics (e.g. age, health status). All figures are taken from Euromod and are weighted to include only those households with at least one member who is aged 18-62 and is not a student. The number of observations per decile is available in Table A1 in the Appendix.

Table 3.4 contains data on the composition of average statutory net tax rates. As gross employment income rises, average net tax rates rise as well. In particular, average net taxes are negative for the bottom percentiles (as people receive more benefits than they pay in taxes), and they rise to 36.7% of gross employment income.

At the bottom of the income distribution, both groups pay similar taxes as a share of gross employment income, even though reasons differ[13]. As gross employment income rises, employees receive less benefits and start paying more taxes as a share of gross employment income (due to the diminishing effect of non-taxable minimum for employees). The self-employed also receive less benefits but are not required to pay higher taxes. As a result, the richest employee households pay 39.3% for the their income in tax, while the self-employed households pay 23.8%.

Similar observations can be made when considering marginal net tax rates. Statutory marginal net tax rates increase from 39% to 43% for employee households, while they fluctuate around 25% for most self-employed households. Two observations, in

13. The employees pay less tax because of a non-taxable minimum, which gradually diminishes as income rises. The self-employed tend to pay less social contributions because of a lower tax base and exemptions.

Table 3.4: Household statutory average net tax rates in Lithuania, net taxes as a share of gross employment income

percentile	net taxes			taxes		public transfers	
	all	employees	self-employed	employees	self-employed	employees	self-employed
0-7							
7-20	-1.121	-0.106	0.171	0.336	0.340	0.442	0.169
20-30	-0.018	0.076	0.104	0.369	0.303	0.293	0.199
30-40	0.130	0.168	0.104	0.379	0.314	0.211	0.210
40-50	0.222	0.237	0.138	0.395	0.287	0.158	0.149
50-60	0.270	0.281	0.200	0.400	0.286	0.119	0.086
60-70	0.293	0.301	0.277	0.400	0.318	0.099	0.041
70-80	0.326	0.335	0.226	0.411	0.293	0.076	0.068
80-90	0.355	0.366	0.257	0.412	0.300	0.046	0.043
90-100	0.367	0.393	0.238	0.419	0.276	0.026	0.037

Percentiles are sorted by gross employment income (includes social contributions). Taxes include income tax and social contributions. public transfers include old-age, disability, unemployment and other benefits. Net taxes are taxes minus public benefits. Gross employment income is taken from EU-SILC, while all other figures are estimated by Euromod, which takes into account various individual and household characteristics (e.g. age, health status). Number of observations per decile is available in Table A1 in the Appendix.

particular, are worth mentioning. The first is that the self-employed in the sixth decile face marginal taxes as high as 46%. This is partly related to public transfers which are capped at these levels. The second observation is that business certificates are no longer allowed at such high levels, and income composition changes. If we remove all households which have both self-employed and employee incomes and remove households with business certificates, the marginal statutory tax rates fluctuate between 24 and 36% for the self-employed.

Our results may be influenced by income concepts and definitions used. In particular, taxes are applied to individuals and not to households in Lithuania, so it is important to consider individuals in the analysis instead of equivalized households. As mentioned, EU-SILC data bundles several income components at the household level, most notability income taxes and social contributions, which are difficult to disentangle. Thus, while EUROMOD can model individual level taxes, the same does not apply to effective taxes[14].

14. As a validity check, we restrict the sample of households to those where there is only one person aged 18-62, although other individuals can also live there. We label this as single households as this removes the issue of married households. The results presented in Table A3 in the appendix are similar to our previous analysis even though the small sample size requires smaller bins. In particular tax rates are very similar in magnitude. Public transfer rates, however, became larger when considering single households as these households tend to contain pensioners or dependants. If we further restrict households to pure one person households, where the one person must be aged 18-62, tax rates remain similar but public transfer rates become closer to those observed in Table 3.4. Moving to single households removes both the issue of focusing on households instead of individuals and remove equivalization effects. Unfortunately, there would be too few observations to allow reporting. Nevertheless, it seems that the results are not significantly affected by the choice of the concepts and we proceed further with our initial ones.

3.4. Effective net tax schedule

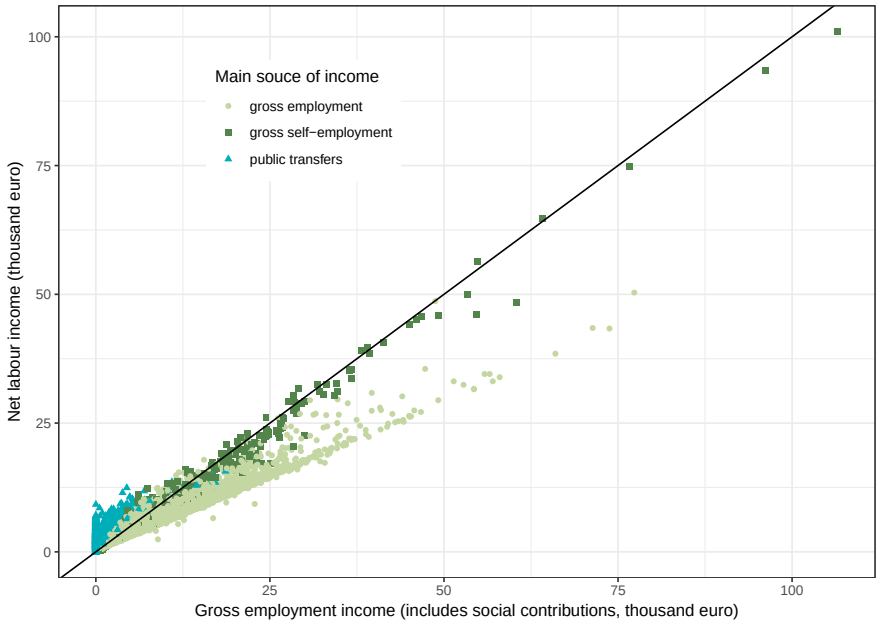
We estimate effective equivalised net tax schedule for Lithuania in a similar fashion as was done for statutory tax section. We use EU-SILC data for the period 2014-2015 and simply compare the net taxes that each household paid with the gross employment income that each household received. The vast majority of net taxes paid by households in EU-SILC come from administrative sources and therefore represent effective taxes paid. Gross employment income in the EU-SILC represents actual income, rather than the income that the tax authorities observe[15]. We find that there is little difference between the statutory and effective net tax schedules for employee households, but the self-employed households pay even less net tax than statutory rates predict. For example, self-employed households effectively pay 29 percent points less net tax on average than employee households at the top of the income distribution.

As in Section 3.3, we plot gross employment income against net labour income for different employment groups in Figure 3.4. In many respects, the effective graph depicting effective tax schedule is similar to Figure 3.3 depicting the statutory tax schedule. The main difference is that self-employed households receive even greater net labour income than employee households. Furthermore, in Figure 3.4, for a large number of households, gross employment income is equal to or even above the diagonal line, irrespective of the amount of gross employment income they earn.

The effective average net tax rates for the self-employed are much below the statutory rates. The top decile of the self-employed pay 7% of their gross employment income as net taxes, as shown in Table 3.5, even though statutory rates suggest that they should be paying 24% (see Table 3.4). While statutory rates might be somewhat overstating taxes because of carried-forward losses, or other tax-relevant features imperfectly captured by the EU-SILC data, the difference is sufficiently large to be noteworthy. In fact, the main drivers are lower effective taxes and social contributions paid by the self-employed (whereas effective and statutory benefits received by the self-employed are similar). In contrast, the statutory and effective net tax rates for employees are similar. This results in a large effective net tax rate difference between the two groups: effective average net tax rates are up to five times lower for the self-employed as compared to employees. Additionally, self-employed average net tax

15. This is because in producing the EU-SILC data for Lithuania, households are asked to report their gross employment income in the questionnaire. Gross employment income is also taken from administrative records for the same household. The two sources (administrative and survey) are compared for each household by the EU-SILC team, and only the larger value of gross employment income is kept in the EU-SILC data that is available to us. Therefore, if respondents revealed more gross employment income in the questionnaire than to authorities, a gap arises between the effective and statutory net tax schedules.

Figure 3.4: Effective equivalised incomes of households grouped using the income definition in Lithuania



Equivalised gross employment income and net labour income is in thousands of euro per year. Households are allocated to employment groups according to the income definition for the 2014-2015 income reference years and are represented by dots in the graph (see Section 3.2). The diagonal line illustrates that what a household earns from employment is what it receives as its net labour income. Any dot above the diagonal line illustrates that the household receives additional benefits, while dots under the horizontal line means that the households pay additional taxes or social contributions.

rates are less progressive: effective average tax rates are flat, with some progressivity coming from public transfers. The lack of progresivity of effective tax rates for the self-employed can be seen in Figure 3.5.

As in the section of statutory net taxes, our concepts and definitions may influence the results, but moving to an individual analysis is problematic given our EU-SILC data. Nevertheless, the results for this restricted subsample of "single" households as in section 3.3 are in line with the fuller sample - if anything the difference in net tax rates between employees and self-employed is larger meaning that our results can be seen as conservative[16].

16. We test a subsample of households with one household member aged 18-62, but allow older and younger members to coexist in Table A4 in the appendix. We label this as single's households. Again, tax rates are similar as to the previous analysis, although the difference in tax rates between the employed

Table 3.5: Household average effective net tax rates in Lithuania, net taxes as a share of gross employment income

percentile	net taxes			taxes		public transfers	
	all	employees	self-employed	employees	self-employed	employees	self-employed
0-7							
7-20	-0.812	-0.095	-0.113	0.320	0.122	0.416	0.235
20-30	-0.041	0.041	-0.101	0.331	0.130	0.290	0.231
30-40	0.074	0.119	-0.144	0.327	0.167	0.207	0.310
40-50	0.191	0.205	0.006	0.352	0.141	0.147	0.135
50-60	0.219	0.241	-0.005	0.365	0.103	0.123	0.108
60-70	0.242	0.267	0.083	0.364	0.124	0.097	0.041
70-80	0.278	0.297	0.048	0.373	0.106	0.076	0.058
80-90	0.313	0.336	0.057	0.387	0.098	0.050	0.041
90-100	0.313	0.359	0.070	0.385	0.101	0.026	0.032

Percentiles are sorted by gross employment income (which includes social contributions). Taxes include income tax and social contributions. Public transfers include old-age, disability, unemployment and other public benefits. Net taxes are taxes minus public benefits. All figures are taken from EU-SILC and are weighted to include only those households with at least one member aged 18-62 and is not a student. The number of observations per decile is available in Table A2 in the Appendix.

3.5. Optimal net tax schedule

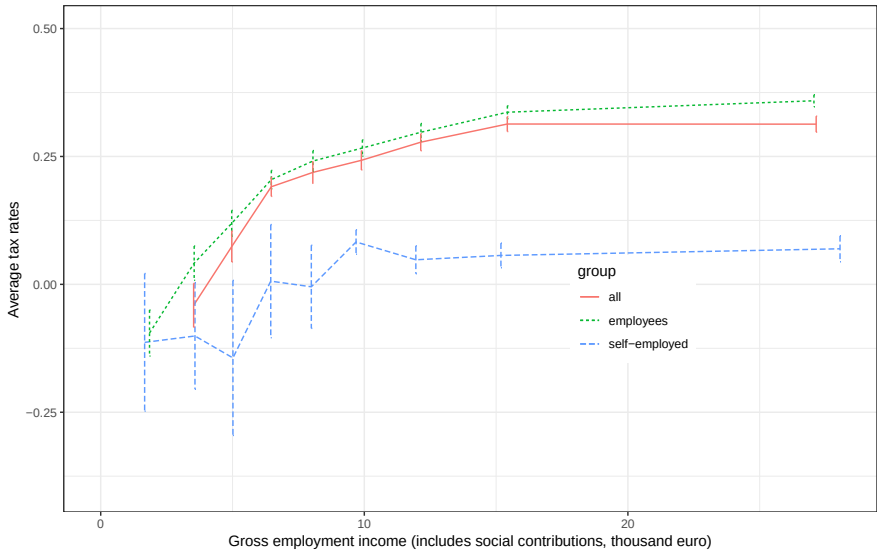
In this section, we estimate an optimal net tax schedule for Lithuania. This allows us to evaluate whether the statutory net tax schedule described in Section 3.3 is in line with the economic fundamentals of the country. Additionally, it could shed light on whether the difference between the effective and statutory rates is likely due to overly high (economically unsustainable) statutory rates, or to low tax compliance. We use a model developed by Saez (2002): it provides the whole optimal net tax schedule given a number of elasticities, government preferences for redistribution and its budget, and a pre-existing income distribution.

3.5.1. The model

The model is taken from Saez (2002), with the exception that individuals are replaced with households (see Section 3.2). The model starts by indexing households by $m \in M$. The measure of households on M is denoted by $dv(m)$. The household's utility depends positively on net labour income c, and the chosen occupation $i \in 0, 1, .., 10$, thus $u(c, i)$. $i = 0$ denotes unemployed or inactive households. The higher the i, the higher the gross employment income w associated with that occupation and the higher the net labour income. In our study, the i represents the same (pseudo) deciles used in Sections 3.3 and 3.4.

and the self-employed becomes even more apparent. Public transfer rates (and subsequently net tax rates) become higher in the single's households because of many co-inhabiting pensioners or dependants. Removing them and keeping only households with a single household member aged 18-62 brings public transfers closer to initial estimates found in Table 3.5, while tax rates are closer to those found in Table A4 in the appendix. These results are not presented due to too few observations.

Figure 3.5: Net taxes are higher and more progressive for employees than for the self-employed



Standard errors were compiled taking into account survey design with the help of codes from Goedemé 2013; Zardo Trindade and Goedemé 2016 and computed using Lumley 2018 R package as described in Lumley 2004. We use 95% confidence intervals.

The fraction of households choosing i is denoted by $h(c_0, c_1, ..., c_I)$, meaning that households weight the net disposable income associated with each job before choosing the best one for them.

The government chooses the net taxes, T_i, that each household should pay or the benefits it should receive and maximises welfare:

$$W = \int_M \mu^m u_m(w_i - T_i, i) dv(m)$$

where μ^m are positive weights and subject to a budget constraint (3.1) described below.

The rest of the derivations are found in Saez (2002), but they eventually lead to a system of three equations

$$\sum_{i=0}^{I} h_i T_i = H \tag{3.1}$$

$$\sum_{i=0}^{I} h_i g_i = 1 \tag{3.2}$$

$$\frac{T_i - T_{i-1}}{c_i - c_{i-1}} = \frac{1}{\zeta_i h_i} \sum_{j=i}^{10} h_j [1 - g_j - \eta_j \frac{T_j - T_0}{c_j - c_0}] \tag{3.3}$$

that show how the government chooses T to maximise W. Let us go through each equation separately.

Equation (3.1) is the government's budget constraint mentioned previously. H is the per capital government's budget net of redistribution. In the simulation, $h_i(c_i - c_0)$, meaning that each household considers the relative gain in net labour income of becoming employed $c_i - c_0$.

Equation (3.2) is a normalisation of the welfare function expressed in terms of social welfare weights. Specifically, g_i denotes the value (in terms of public funds) of giving an additional dollar to a household in occupation i. That is, the government is indifferent regarding giving one more dollar to a household in occupation i and getting g_i of public funds. The higher the g_i, the happier the government is to give money to this occupation and, assuming the government values redistribution, g_i decreases as i increases. Additionally, g depends on net labour income c, the marginal value of public funds p and the distributional tastes of the government v as shown in (3.4). If c is already equally distributed, then there is less reason to further redistribute and so g should be equal across i's. The higher the p, the more the government values its public funds and the less keen it is to redistribute income. The higher the v, the keener the government is to give money to the poorest members of society instead of to the wealthiest.

$$g_i = \frac{1}{p c_i^v} \tag{3.4}$$

Equation (3.3) defines the optimal net tax schedule of a change in net tax rate for occupation i by a small amount dT. Three effects are at work here, which have to be balanced to reach optimal net tax rates. First, there is the mechanical effect of a change in net tax rate. The rise in T_i causes the government to collect more revenue from all those in occupation i and all richer occupations $i + 1, i + 2, ..., 10$.

This is represented by $\sum_{j=i}^{10} h_j$. Second, we include the effect of social weights, g_i attached to each occupation. This is done by stating that the government values each dollar collected by occupation i at $1 - g_i$, since the government may prefer not taking money from some groups in the first place (e.g. the very poor). Third, it includes two behavioural responses: the extensive response and the intensive response.

The extensive response is captured by the extensive labour supply elasticity (technically, the extensive mobility elasticity),

$$\eta_i = \frac{c_i - c_0}{h_i} \frac{\partial h_i}{\partial (c_i - c_0)} \tag{3.5}$$

which refers to T_i becoming so large that some people working in i may choose to become unemployed or inactive (i_0). It measures the percentage change in number of employed in occupation i when the difference between net labour incomes of employed in occupation i and unemployed/inactive changes by 1%. For example $\eta_i = 0.5$ means that if $c_i - c_0$ increases by 1%, employment in i will rise by 0.5%.

The intensive response is captured by the intensive mobility elasticity (akin to the intensive labour supply elasticity)

$$\zeta_i = \frac{c_i - c_{i-i}}{h_i} \frac{\partial h_i}{\partial (c_i - c_{i-i})} \tag{3.6}$$

which refers to people moving from one occupation to another in search of lower net taxes. It measures the percentage increase in supply of job i when $c_i - c_{i-1}$ is increased by 1%. This specification ignores income effects, or the effect of rising incomes for all occupations simultaneously. In the literature, however, income effects are in any case found to have a small impact (Saez 2002).

Finally, h_i represents the optimal i distribution given the empirically observed h_i^0 distribution

$$h_i = h_i^0 (\frac{c_i - c_0}{c_i^0 - c_0^0})^{\eta_i} \tag{3.7}$$

where the h_i^0 are reconfigured to account for the extensive response to change in net taxes. Here, c^0, represent the actual net income and c_i represent the optimal net income which is estimated simultaneously with (3.1, 3.2, 3.3). Whenever net taxes are lowered for households of occupation i, so that $c_i - c_0$ becomes bigger, more households should be working in i, given extensive elasticity η_i and actual net incomes

$c_i^0 - c_0^0$.

3.5.2. The parameters

There are several parameters that need to be chosen for Lithuania: the labour supply elasticities (or, actually, long-run taxable income elasticities), societies' preferences and other. We use taxable income elasticities, e_z, defined as

$$e_z = \frac{1-\tau}{z}\frac{\delta z}{\delta(1-\tau)}$$

the percent in reported income when the net-of-tax rate increases by 1 percent. The benefit of this "sufficient" elasticity is to capture directly all behavioural effects or raising taxes, including real responses (e.g. labour supply adjustments), tax avoidance (e.g. claiming deductions or (legal) income shifting between tax bases) and illegal tax evasion behaviour (see Saez et al. 2012, for example). Nevertheless, we also rely on the available labour supply elasticity estimates for Lithuania.

Elasticities

We start with choosing (uncompensated) intensive and extensive labour mobility elasticities for (3.5) and (3.6) respectively. Income effects are usually found to be small on aggregate (Saez 2002; Bargain et al. 2014), which justifies considering uncompensated labour supply elasticity instead of compensated labour supply elasticity. Additionally, we require different extensive and intensive mobility elasticities for high and low income households. If these differ, this should produce a kink in the optimal tax schedule: higher extensive elasticities for low incomes calls for subsidies to the poor.

First, it should be noted that ζ is not observed empirically, but can be calculated

$$\zeta_i = \frac{\epsilon_i w_i}{w_i - w_{i-1}}$$

by first estimating

$$\epsilon_i = \frac{1-\tau}{w}\frac{\delta w}{\delta(1-\tau)}$$

where ϵ show how much wage responds to the net-of-tax rate change.

Second, as the magnitude of elasticities is uncertain, Saez (2002) proposed a wider range of ϵ's and η's for the upper and lower tail of distribution based on the summary of literature (see Table 3.6). Unfortunately, the ranges are large, are based mainly on US data, are ambiguous about being short- or long-run elasticities and refer to labour supply responses only (i.e. are not elasticities of taxable income). This has been partly remedied by newer studies.

Table 3.6: intensive and extensive elasticities as proposed by Saez (2002)

	High income ($w \geq 20,000$\$)	Low income ($w < 20,000$\$)
η	0	[0 - 1]
ϵ	[0.25 - 0.5]	[0.25 - 0.5]

The table indicates a range of possible elasticities for the United States.

Barrios et al. (2019) estimated Lithuania's short-run labour supply elasticity,

$$e_h = \frac{w}{h}\frac{\delta h}{\delta w} \tag{3.8}$$

denoting a percent change in net-wage on the number of hours worked, to be between 0.15 for high-skill individuals and 0.3 for low-skill individuals. This elasticity captures the main behaviour effect: the real response of labour employment and work duration (the sum of ϵ and η). While there are no estimates for Lithuania's intensive, e_{hi}, and extensive, e_{he}, margin, Bargain et al. (2014) study these distributions across income quantiles countries largely comparable to Lithuania, such as Estonia, Hungary, Finland and Poland. For the four countries, the extensive labour elasticities for the lower quantiles, e_{hel} range between 0.08 to 0.26 (an exception is Finland, with 0.8). For the higher end, e_{heh} range between 0.05 to 0.23. For the same four countries, intensive labour elasticities range between 0 to 0.03 for the lower, e_{hil}, and -0.04 to 0.03 for the higher e_{hih} deciles. The extensive elasticity was found to vary between 0.3 to 0.65 in Staehr (2008) for Estonia, while intensive elasticity was negligible. This suggests that for Lithuania, also, most of the labour supply would come from the extensive margin for both the lower and higher income households, even though there may not be large differences between the upper and bottom income distributions.[17]

17. The unresponsiveness of elasticities to income deciles was explained in a more recent study for Slovakia by Senaj et al. (2014). There, e_{heh} falls to 0.06 from 0.16 e_{hel} when only prime age workers

Lithuania's long-run labour supply elasticity could be much higher, and long-run taxable income elasticities are larger still. We opt for long-run elasticities to capture long-run effects on the economy. Jäntti et al. (2015), who has access to long-term data for largely Scandinavian countries, finds e_{he} to range between 0 and 0.4, while e_{hi} ranges between 0 and 0.28. This suggests that a fair long-run range for Lithuania's e_h is 0.1 to 0.7. It is expected that $e_z >= e_h$. Empirical studies such as Jongen and Stoel (2019) for the Netherlands show that e_h is only 0.05, while e_z is 0.21 in the long run. Lithuania's long-run elasticity of taxable income should also have a similar range, but is more likely to be from 0.2 to 0.8, with the most likely elasticities at 0.5 at the top and the bottom of the income distribution (the intensive margin more relevant for the top and the extensive margin for the bottom). This falls within the range of e_z estimates, although it exceeds the average of 0.3 (Neisser 2017).

One reason for the larger e_z in Lithuania could be the tax system. The narrower the tax base, hence many tax avoidance possibilities, the higher is the elasticity (Saez et al. 2012). The statutory net tax of Lithuania shows that avoidance possibilities exist, especially for the self-employed. Another reason could be the low level of law enforcement (Saez et al. 2012). The large shadow economy in Lithuania suggests that tax rules there are not enforced sufficiently. The final list of e_z is presented in Table 3.7. We assumed that the high income corresponds to 12000.

While elasticities in Table 3.7 apply to the general population, which is dominated by employed households, it does not necessarily apply to average self-employed households. For instance, tax evasion can be higher amongst the self-employed, since they are not subject to third-party reporting. Indeed, the elasticities for the self-employed are found to be up to three times larger in Spain byAlmunia and Lopez-Rodriguez (2019) and in Poland by Zawisza (2019). Other studies also show that elasticities of self-employed income are roughly two times higher than for other types of income (Neisser 2017). However, since we have no available elasticities for Lithuania, we leave this for future work.

Society's preferences and other parameters

Another parameter is the society's preference parameter v. Saez (2002) in most cases used $v = 1$, which already has a high preference for redistribution, while $v = 0.25$ would be a lower point estimate. According to surveys, 92% of Lithuanians believe income inequality is too high, one of the leading countries in the EU. Additionally, Lithuania's government explicitly tries to reduce poverty and income

are considered, but not when a larger share of older workers are included. For Lithuania, then, where pensions are relatively low compared to the average wage, potential pensioners are also more likely to respond strongly to wages.

Table 3.7: Ranges of elasticities of taxable income for Lithuania

	High income ($w \geq 12000euro$)	Low income ($w < 12000euro$)
η	[0.2, **0.3**, 0.5]	[0.2, **0.4**, 0.6]
ϵ	[0.1, **0.2**, 0.3]	[0.02, **0.1**, 0.2]

The preferred taxable income elasticities for Lithuania are **bolded** while the range of possible elasticities are in brackets. w is equivalised employment income, which includes employer's and employee's social contributions.

inequality (LR Vyriausybė 2017). Therefore, v should be clearly positive and relatively high. We set $v = 1$ in the baseline and $v = 0.7$ as an alternative scenario.

The other parameters are derived from EU-SILC data itself. $H = 2199$ as this was the sum of net transfers from the EU-SILC survey, c_i^0 and h_i^0 was taken from the EU-SILC survey as well. $i = 1, 2, ..., 10$ so that each occupation constitutes about 10% of population, although the first bin is smaller, so that $w_0 = 0$.

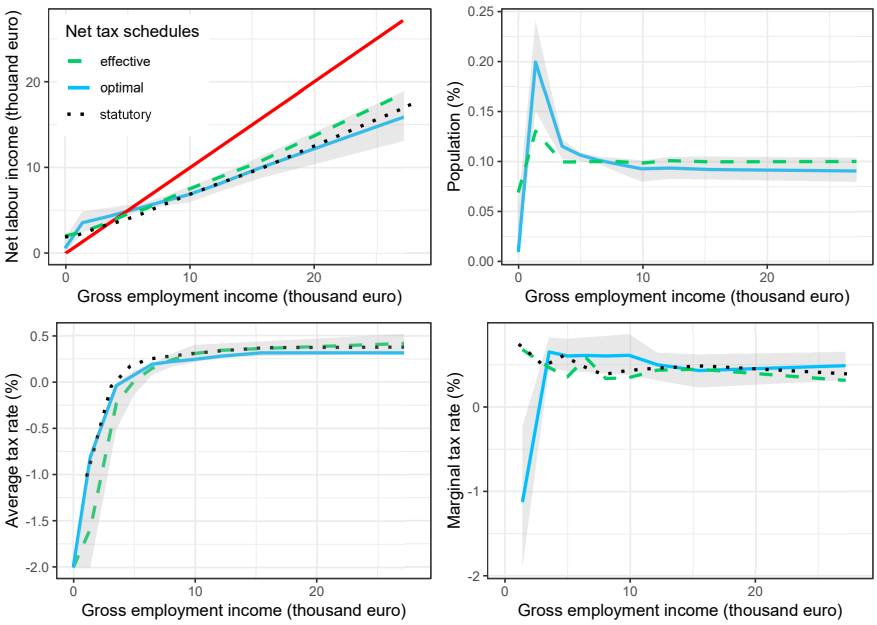
3.5.3. The simulations

Given the model and the parameters, we utilise an R-package by Hasselman (2018) to run the simulations for Lithuania. We obtain four key variables: net labour income, population distribution by income, and average and marginal net tax rates. Information about each variable is presented in four graphs in Figure 3.6 and in Table 3.8. In each graph, the preferred parameter specification is depicted by a blue line, and alternative parameter choices are presented as a shaded area around the blue line. The green dashed line represents the effective net tax schedule, and the black dotted line is the statutory tax schedule. Let us go through what messages each graphs suggest in turn.

The effective and statutory net tax schedule coincides with the optimal net schedule for the middle of the income distribution, but less for the tails. The figure on the top- left holds the transformation from gross employment income to net labour income. Effective net labour income and statutory schedule coincides with the optimal net labour income for middle (gross employment) incomes, and, in most cases, falls within the range of optimal schedules. At higher incomes, the optimal net labour income is slightly below the net labour income of the statutory and well below the effective net tax schedules. For those earning little gross employment income, the optimal tax rates suggest that more can be done to increase labour market participation and reduce unemployment: less income should be directed to the very poorest, and in-work credits should be provided. Unemployment and non-participation would

then drop (from 6.9% to close to 0.9%) while the share of households employed at lower income levels rises (from 13.0% to 20.0%) because of more in-work credit, as illustrated by the top-right figure. The unemployment and non-participation drop should be taken with caution. The optimal net tax model does not distinguish between work capacity and household preferences. For example, some households may suffer from severe disability or wish to attend to their own children. In these cases, it may not make sense to fully remove benefits or expect that in-work incentives would encourage these people to work.

Figure 3.6: Optimal, effective and statutory net tax schedules



In each graph, the optimal net tax schedule with the preferred parameter specification (see Table 3.7) is depicted by a blue line while alternative parameter choices are presented as a shaded area around the blue line. The green dashed line and the black dotted line represents the variables distributions in line with the effective and statutory tax schedules respectively. The diagonal red line on the top-left figure is a 45 degree line depicting zero net taxes.

Effective/statutory average and marginal tax rates are close to their optimal levels in the middle of the income distribution, but not the tails. Optimal marginal tax rates for the bottom deciles are strongly negative: 112.4% of their gross employment income. This contrasts markedly with the effective positive 67.7% marginal tax rate for the bottom deciles. Additionally, the optimal marginal tax rate for the top of the

gross employment income distribution is 48.7% while the effective marginal tax rate is 31.3% or about 11% below statutory. Empirical studies suggest that optimal tax rates tend to be much higher than statutory rates at top incomes. Saez (2002) shows that the majority of estimates of optimal tax rates for top incomes for the USA lie above 50%. Klemm et al. (2018) also find that the top optimal marginal tax rates exceed 50% and tend to be 10 - but sometimes even 30 - percent points above the statutory marginal tax rates in 27 countries. Therefore, the 11% difference is on the lower side of the estimates. Part of the reason for the gap is the large extensive labour elasticity in Lithuania for top incomes, which prevents taxing high incomes too high. Another reason is a large presence of self-employed.

Table 3.8: Effective and optimal variables for Lithuania

percentile	gross employment income	net labour income[a]	net labour income[b]	percent of households[a]	percent of households[b]	average tax rate[a]	average tax rate[b]	marginal tax rate[a]	marginal tax rate[b]
0-7	0.0	2.1	0.6	6.9	0.9				
7-20	1.4	2.5	3.6	13.0	20.0	-81.2	-158.5	67.7	-112.4
20-30	3.5	3.7	4.4	10.0	11.5	-4.1	-23.5	46.4	64.8
30-40	5.0	4.6	4.9	10.0	10.7	7.4	0.8	35.4	60.1
40-50	6.5	5.2	5.5	10.0	10.1	19.1	14.7	57.9	60.7
50-60	8.0	6.3	6.2	10.0	9.8	21.9	23.6	33.4	60.1
60-70	9.9	7.5	6.9	9.9	9.3	24.2	30.5	34.6	60.9
70-80	12.1	8.8	8.0	10.1	9.4	27.8	34.1	43.3	49.7
80-90	15.4	10.6	9.9	10.0	9.2	31.3	36.0	44.5	42.9
90-100	27.1	18.6	15.9	10.0	9.1	31.3	41.4	31.3	48.7

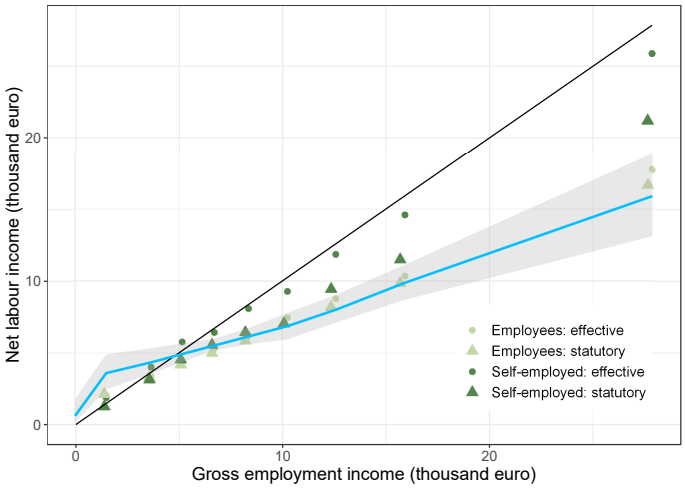
[a] Effective variable. [b] Optimal variable. Gross employment income and net labour income are in thousand euro per equivalized household in Lithuania in the period 2014-2015. Share of households, average tax rates and marginal tax rates are in percentages. Number of observations per decile is available in Table A2 in the Appendix.

3.6. Statutory, effective and optimal net tax schedules for employees and the self-employed

Here, we compare statutory, effective and optimal (equivalised) net tax schedules for employee and self-employed households. The three net tax schedules coincide more for employees than the self-employed. This can be seen in Figure 3.7 where the two groups are distinguished. The effective and statutory net tax schedules for employees lie close to the optimal tax schedule, while the self-employed are further away - in most cases, outside of the optimal net tax schedules range. The self-employed are subject to lower statutory net tax rates which exceed the range of optimal net tax schedules for higher gross employment income deciles. Additionally, the self-employed effectively pay even lower effective tax rates than they are required. This holds true for the whole gross employment income distribution. There is also a smaller difference between the two groups at the bottom. The self-employed face relatively higher net tax rates than employees due to lower public transfers (compare Tables 3.4 and 3.5).

There are several possible ways to explain the large gap between the effective

Figure 3.7: Statutory, effective and optimal equivalised tax schedules for households grouped according to income



The graph illustrates how household's equivalised gross employment income translates into equivalised net labour income for three tax schedules: effective, statutory and optimal. The effective and statutory tax schedules are presented as points (for each decile) for employees and the self-employed. The optimal tax schedule is calculated for the the total population aged 18-62, and excludes students. The shaded area around the dashed blue line illustrates a range of optimal tax schedules using a range of parameters as shown in Table 3.7. The data comes from EU-SILC, the simulation was carried out with the help of Euromod and the optimal tax schedule was computed along the lines of Saez (2002).

and statutory net tax schedules for the self-employed. The most likely explanation is tax evasion. In Lithuania, there is a tendency to under-report self-employment income or to not declare being self-employed at all, as previously noted by Navickė and Čižauskaitė (2018). Assuming that survey respondents are more willing to reveal their true self-employed incomes in questionnaires, we can compare the effective and statutory tax rates to obtain an estimate for evaded taxes in Lithuania, as done in Table 3.9. Employee households may not pay up to about 5.6% to 14.4% of their taxes, while the self-employed may evade as much as 69.9%, depending on the gross employment income distribution. Assuming that missing taxes arise from under-reported income, we see that these numbers are high, but plausible, given the empirical literature. An estimate for Lithuania is found in Kukk et al. (2019), who estimated income under-reporting of the self-employed in surveys to be around 25% to 30%, depending on the definition of "self-employed". The study, however, uses the consumption approach to estimate tax evasion, which should give a lower bound of under-reporting estimates. Also, income under-reporting in surveys does not necessarily mean that people equally under-report income to authorities. For example, the same study estimated that, in Estonia, the self-employed under-report 22% of their income, while Paulus (2015) estimated that as much as 71% of self-employment income is unreported to authorities, which is what matters for tax collection. Estimates from other countries are generally in line with what we expect given our results. Paulus (2015) finds that, in Estonia, up to 20% of employees under-report income. Paulus (2015) also finds that under-reporting is greatest at the tails of the income distribution, something also found by Johns and Slemrod (2010) for the USA. While there is greater under-reporting at the lower percentiles for Lithuania, the message is less clear for the top. However, this may be due to failure to capture top incomes in the survey for Lithuania. Many more studies find that the self-employed evade much more taxes than employees by under-reporting income. Baldini et al. (2009) finds that, in Italy, the self-employed tend to evade more income tax than employees. Pissarides and Weber (1989) find that the self-employed in UK actually have 1.55 times the reported income, meaning that they under-report income by 35% in the UK, while Slemrod (2016) cites IRS studies in the USA, where 56% of income may be unreported for the self-employed. A study by Artavanis et al. (2016) in Greece shows that the self-employed in certain professions, such as doctors, lawyers, engineers and scientists, as well as accountants and financial service agents, under-report more than half of their income.

Even though tax evasion is a likely explanation for the difference between effective and statutory tax rates for the self-employed, it is also reasonable to assume that

Table 3.9: Estimated difference between statutory and effective tax schedules per equivalised household per year

percentile	% of statutory tax		th. euro	
	employees	self-employed	employees	self-employed
0-7				
7-20	4.76	64.12	0.02	0.28
20-30	10.30	57.10	0.14	0.62
30-40	13.72	46.82	0.27	0.75
40-50	10.89	50.87	0.28	0.96
50-60	8.75	63.99	0.29	1.50
60-70	9.00	61.01	0.36	1.96
70-80	9.25	63.95	0.47	2.33
80-90	6.07	67.33	0.39	3.17
90-100	8.11	63.41	0.94	4.83

The figures are derived from the difference between statutory and effective average tax rates from Tables 3.4 and 3.5 respectively. Percentiles are sorted according to the equivalised household gross employment income of all non-students aged 16-62.

some of this difference is due to measurement error. However, it is not clear if in aggregate the error under- or overestimates the difference. First, Euromod does not model all taxes and contributions, which would result in lower tax evasion. Second, there might still be some income, particularly self-employment income, that is not reported to the authorities and not revealed in the questionnaire, which would mean greater tax evasion.

The difference in statutory rates between employees and the self-employed could be accounted for in several ways. For example, the government may perceive the self-employed more favourably than employees. There could be at least two reasons for this. One is that the self-employed would not be able to become employees, and this scenario is better than being unemployed. A second reason is that the government believes that the self-employed tend to contribute more to society, either by themselves producing significantly more earnings due to lower taxes, by supporting the rest of the economy by being entrepreneurs and eventually hiring more labour, or by producing other positive externalities (see Scheuer and Slemrod (2019)). However, the first theory does not stand up to the data and the literature, while the second lacks credible evidence. Regarding the first reason, the self-employed are bunched at the top of the income distribution. If these households tend to earn high incomes, it is not clear why they could not become employees or pay higher taxes as self-employed. Regarding

the second reason, a minority of the self-employed, according to EU-SILC, could be considered entrepreneurs and less than 10% of self-employed at the top of income distribution have employees of their own.

This leaves the possibility that the self-employed are especially responsive to tax rate changes or bring about large positive externalities - something that has not yet been tested for Lithuania. At the same time, a review of the literature suggests that a major reason for becoming self-employed is not entrepreneurship, but greater tax evasion/avoidance opportunities (Baliamoune-Lutz and Garello 2014). Additionally, the empirical literature is mixed concerning whether the self-employed respond to tax changes, thereby placing lower statutory rates into question (Baliamoune-Lutz and Garello 2014). For example, Bruce (2002) show that higher statutory tax rates on self-employed income in the USA did not lead to the closing of small businesses. On the contrary: higher proportional taxes on the self-employed, together with the possibility of offsetting losses, actually encourage entrepreneurship via a risk-sharing channel, as first explained by Domar and Musgrave (1944) and later found in empirical work (e.g. Baliamoune-Lutz and Garello 2014). What seems to deter self-employment is progressive self-employment taxes, as shown by Gentry and Hubbard (2000) for the USA and by Baliamoune-Lutz and Garello (2014) in Europe.

3.7. Conclusions, limitations and recommendations

We compared the statutory, effective and optimal net tax schedules for Lithuania for the period 2014-2015. We did this for all Lithuanian households and then looked at employee and self-employed households separately to investigate different forms of employment.

We found that the three schedules largely coincide for the middle of the income distribution for all households. The three diverge, however, at the tails of the income distribution. At the bottom of the income distribution, the optimal net tax schedule suggests that more in-work benefits should be provided for the least paid, to encourage employment. At the top of the income distribution, more effort could be made to extract tax revenue in order to improve tax compliance. The results for employee households were similar to that of all households.

We found that the three net tax schedules coincide more for employee households than for self-employed households. Except for those at the very bottom of the income distribution, the self-employed are subject to lower statutory net tax rates and very low progressivity, as compared to employees. Unfortunately, using the same elasticities for the employed and the self-employed does not allow us to draw strong conclusions about optimal taxes for the self-employed. Nevertheless, the self-employed do ef-

fectively pay much lower taxes than the statutory tax schedule would suggest. This holds throughout the income distribution and could mean that as much as 70% of self-employed taxes are not paid.

Our conclusion can be viewed as a conservative one. If we were to exclude pension contributions or consider all social contributions as generating actuarially fair benefits, the inadequacy in taxation levels would likely be even larger. The divergence would be greater still if we were to consider income taxes only, and not social contributions or benefits. Additionally, we considered a budget-neutral tax schedule. Finally, the fact that statutory rates differ substantially can explain why optimal taxes are also relatively low. Were there fewer opportunities to avoid taxes by having a broader tax base, measured elasticities would be smaller and optimal taxes would be higher.

As this is an initial step in comparing the three schedules, there are ways to improve the estimates. First, the EU-SILC is known to poorly capture top incomes; greater access to administrative data could help solve this problem. Second, the fact that the statutory tax schedule differs from the effective tax schedule for the self-employed means that the household misreport their employment status and incomes to the authorities, to EU-SILC or both. Third, we were not able to find Lithuania-specific long-run estimated elasticities, meaning that the current ones had to be taken from other studies. Nonetheless, such elasticities can be eventually estimated, particularly as a large income tax reform took effect in 2019. Obtaining taxable income elasticities for the self-employed and the employed separately would be especially beneficial. Fourth, one may consider a different set of elasticities or/and preferences for the optimal net tax schedules of employees and the self-employed. For example, society could value the self-employed more, or they themselves could be more responsive to wages.

The findings presented in this chapter point to several recommendations.

First, the effective net tax schedule indicates that less taxes and social contributions are collected than households are statutorily required provide. Therefore, more effort can be placed on the auditing of households, especially at the upper tail of the income distribution, to extract more government revenue. Before doing so, the marginal cost of the audit and the marginal value of public funds should be estimated.

Second, the optimal net tax schedule recommends providing tax credits to those who receive low wages. Upon obtaining better estimates of the bottom of the distribution, this policy could be considered further. This is especially relevant with the resurgence of discussions on universal incomes, which counters in-work credit suggestions.

Third, the optimal tax schedule recommends less benefits to unemployed and non-active households. With the combination of lower out of work benefits and higher tax credits, households would be more inclined to seek employment. However, one would first have to consider at least the health and preferences of households, as many benefits relate to health, disability and children.

Fourth, the benefits of the current lower statutory taxes for the self-employed should be closely weighted alongside the associated costs of lower tax revenue. As the majority of the self-employed are found at the upper tail of the income distribution, a great deal of tax revenue is not collected. Furthermore, international evidence shows that some companies start hiring and individuals start choosing self-employment purely for the purpose of paying less taxes. In such cases, it may be in the general interest to raise statutory tax rates for the self-employed closer to, or even above, the tax rates of employees.

4. CHILD PENALTY AND CHILDCARE IN THE BALTICS

4.1. Introduction

Inequality in earnings between genders remains large and relatively unexplained in the Baltics, which can cause a headache for policy makers. In 2018, the unadjusted gender pay gap[1] was 21.8% in Estonia (the highest in EU), 19.6% in Latvia and 14% in Lithuania which is close to EU27 average of 14.1% (Eurostat 2021). A measure of the ”adjusted” gender pay gap, when observables such as industry, occupation, hours worked, age and education are controlled for, reduced the wage gap for Estonia and Latvia but increased that of Lithuania (Boll and Lagemann 2018)[2]. Without understanding where the gender gap is coming from, it is difficult for policy makers to choose adequate policies.

In this chapter, I measure the impact of having children on the wage gap in the Baltics. That is, I quantify the child penalty – the income forgone due to bringing up a child – and its contribution to the wage gap in the first years after childbirth. Following Kleven, Landais, and Søgaard (2019), I conduct an event study by observing the dynamics of men's and woman's earnings up to 3 years before and after childbirth. The EU-SILC data allows me to observe a 29-53% decline of earnings for females when the first born child is up to 3 years old, while male earnings fall by half of this amount (0 - 27%).

Noting the widening wage gap after childbirth I further quantify whether access to early childhood education and care (ECEC) [3] could offset it. For this, I merge the longitudinal EU-SILC data with the cross-section version of EU-SILC data, encompassing respondent's answers to questions on the use of formal ECEC, which includes pre-school education care provided by professionals in other locations, as well as informal childcare provided by non-parents (e.g. grandparents, friends, etc.). Regression analysis suggests that formal and informal ECEC is associated with lower wage child penalty for women, but not for men. I then test whether expansion of pre-school ECEC results in a lower wage gap. Specifically, I test whether higher overall enrolments, which are driven by higher access to pre-school ECEC, affect earnings

1. The difference between male and female average gross hourly earnings as a percent of average gross hourly earnings of men

2. The adjusted gender pay gap may be higher than the unadjusted one if women have better endowments than men (e.g. are more educated), but receive less returns from the endowments.

3. In the EU context, ECEC is defined as ”provision for children from birth through to compulsory and primary education that falls within a national regulatory framework, i.e. it has to comply with a set of rules, minimum standards and/or undergo accreditation procedures” (Vandenbroeck et al. 2018). However, in this chapter, ECEC is defined more broadly to also include informal ECEC.

of men and women differently. Unfortunately, the results are model-dependent, although the more flexible regressions suggest that women benefited more than men.

In doing so, I contribute to the empirical literature on the wage gap in the Baltics. Limited work has been done on estimating the child penalty or the wage gap just before and after childbirth in the region directly, despite it being recognized as the major cause of the wage gap in developed countries, such as Denmark (Kleven, Landais, and Søgaard 2019). Anspal et al. (2011) found that women with children earn on average 1.2% less than women without children in Estonia, while no comparable loss of earnings was observed for men. However, unlike the methodology used in this chapter, the methodology used by Anspal et al. (2011) cannot address two issues related to the correlation between earnings and childbirth. The first is the issue of reverse causality. That it, that earnings could influence childbirth and not vice versa. The second is the issue of unobserved heterogeneity. This refers to the issue that women with children are systematically different from women without children in terms of innate ability or willingness to work. Instead, the empirical literature looked at other factors behind the wage gap in the Baltics.

Other factors contributing to the wage gap that were studied in the Baltics include labour market segmentation and the reasons for it. Researchers conclude that horizontal and vertical labour market segregation – i.e., males and females concentrating in certain industries and occupations – contributes to about a third of the gap (see, e.g., Rakauskienė et al. 2020; Boll and Lagemann 2018; Anspal et al. 2011; Anspal et al. 2015)[4]. Several reasons for this segregation have been evaluated. The human capital theory-based explanation, that men invest more into their human capital than women and are able to pick the better paid industry, is at least partly rejected for the Baltics. This is because women tend to be more educated than men and have similar work experience (Anspal 2015; Boll and Lagemann 2018). Instead, more evidence favours a gender identity theory, whereby women's preferences and perceptions differ from those of men. One way in which these preferences differ is that women seem to associate themselves with certain societal expectations and derive utility if they meet these expectations. In line with this theory, Redmond and McGuinness (2019) found that job preferences explain around 10% of the wage gap for the Baltics. Moreover, the higher the income, the more the job preferences contribute to the wage gap. Another study that supports the identity story is by Meriküll and Mõtsmees (2017). They found that women in Estonia ask for lower wages before entering employment as compared to men and the magnitude of the requested wage gap matches the ad-

4. Although Boll and Lagemann (2018) find that this segmentation can explain much less of the wage gap for Lithuania than in the other two Baltic countries.

justed wage gap. Therefore, this may suggest that women expect lower wages and undervalue themselves or that they value being employed more than men do, even if it means forgoing higher wages. Vahter and Masso (2019) found that the wage gap is larger within foreign firms compared to domestic firms and provide tentative evidence that higher willingness to work among men compared to women allows men to be compensated for working longer or inconvenient hours. While gender identity theory is the strongest candidate to explain the wage gap, it is not conclusive. For example, the results of Meriküll and Mõtsmees (2017) can also be interpreted to support that discrimination exists in the labour market. It could be that women know that they will be discriminated against and not bother asking for higher wages. Even though discrimination is difficult if not impossible to observe directly, Rakauskienė et al. (2020) attempted to estimate it within medium sized firms in Lithuania but no systematic evidence was found. Finally, there is the institutional explanation that certain arrangements can favour women rather than men. While legal matters have been attended to by Rakauskienė et al. (2020), there exist differences in maternity/paternity benefits which are yet to be examined.

Additionally, I contribute to the empirical literature on ECEC. Studies find that formal ECEC affects outcomes of the child, the parents and the society in various dimensions (see, e.g., Vandenbroeck et al. 2018, for an overview of the literature). I only focus on the effects of ECEC on the parents' labour market outcomes.

Studies, which examine the effect of formal ECEC on parents' employment, must resolve the endogeneity issue and the issue of prices of formal ECEC and informal care (see Blau and Currie 2006). That is, researchers have to know whether those parents who choose to take their children to formal childcare arrangements are not special compared to those that do not (e.g., have higher potential earnings, like to work longer) and that parents do not substitute formal childcare with informal (such as that provided by grandparents). Without this, the often-observed correlation between formal ECEC and positive labour outcomes cannot be considered causal and, therefore, expansion of formal ECEC would not necessarily improve labour market outcomes. To sidestep this problem, researchers sought out quasi-experiments, where large reforms were implemented. Some studies look at reforms that affected only part of the country. In this case, outcomes of parents in the affected part of the country could be compared with outcomes in the unaffected parts of the country. Examples of such studies include Baker et al. (2008) and Lefebvre and Merrigan (2008) for Canada, Gathmann and Sass (2018) for Germany and Carta and Rizzica (2018) for Italy. Other studies look at country wide expansion, but exploit either different expansion times in municipalities, as the case for Norway by Havnes and Mogstad (2011),

or school eligibility conditions, as the case for USA by Gelbach (2002). All of the mentioned studies, apart from Havnes and Mogstad (2011), find large positive effects on the labour market.

The remainder of the chapter is structured as follows. The child penalty in the Baltics is estimated in section 4.2, where the data used, the methodology applied and the results obtained are presented in separate subsections. Section 4.2 also includes a subsection on parental leave policies in the Baltics, so that these results could be easier compared to results in other countries. The effects on parents' earnings of the expansion of formal ECEC in the Baltics is examined in section 4.3. Finally, the conclusions are formulated in section 4.4.

4.2. Estimating child penalty in the Baltics

Child penalty in the Baltics is estimated in this section. First, I describe the EU-SILC dataset which is used in this section and section 4.3.1. Then, I present the relevant descriptive statistics which I use throughout this section only. Furthermore, I present the methodology used to estimate the child penalty and the estimation results. I also include a short subsection on maternity, paternity, and parental leave policies at the end.

4.2.1. EU-SILC data

The data on earnings of men and women as well as the year of birth of a child and other relevant demographic and labour market variables come from the yearly European Union Statistics on Income and Living Conditions (EU-SILC) instrument. In this chapter, I use data from 2005 to 2017. The data are compiled from a mixture of the survey and administrative sources. Since all EU member states collect data using the same methodology, I obtain comparable data on the three Baltic states.

Including three Baltic countries allows utilizing a larger comparable sample. Each year, around 5 thousand households per a Baltic country with around 10 thousand persons over 16 years old, who agree to share information on their income, are included. My sample is much smaller, as I only focus on those households which experience a first childbirth during one of the years. Specifically, I need to observe households at least 1 year before childbirth ($t = -1$) and one year after childbirth ($t = 1$) to see how their incomes change. Table 4.1 shows that the sample size for an individual country would be small but remain reasonable for the region as a whole. The sample size is particularly small for Lithuania and that can be explained by an overall small sample size of the survey and no explicit targeting of such observations. Because the countries are often seen as very similar but not identical, this also allows exploiting additional variation in the study.

Table 4.1: Number of observations by event time and sex

	Female					Male				
Event time t	-2	-1	0	1	2	-2	-1	0	1	2
EE	154	287	288	180	88	106	211	211	139	71
LT	61	117	117	76	43	35	77	77	50	30
LV	118	242	258	167	84	96	199	211	144	68
Baltics	333	646	663	423	215	237	487	499	333	169

Note: Event time denotes the number of years before $(t < 0)$, during $(t = 0)$ and after $(t > 0)$ birth of first child.

As the EU-SILC is a yearly survey, it is not possible to state the actual age of the child in $t = 1$ or $t = 2$. If the child is born at the start of the previous year, he or she will be 1 year and 1 day old at the beginning of $t = 1$ and 2 years old at the end of $t = 1$. If born at the end of the year, then the child may be just 1 day old at the start of $t = 1$ and 1 year and 1 day by the end of $t = 1$. On average, the child is expected to be 1 year old, but the parent will have periods when the child is much younger and older. Due to similar reasoning, the child is expected to be 2 years old in period $t = 2$.

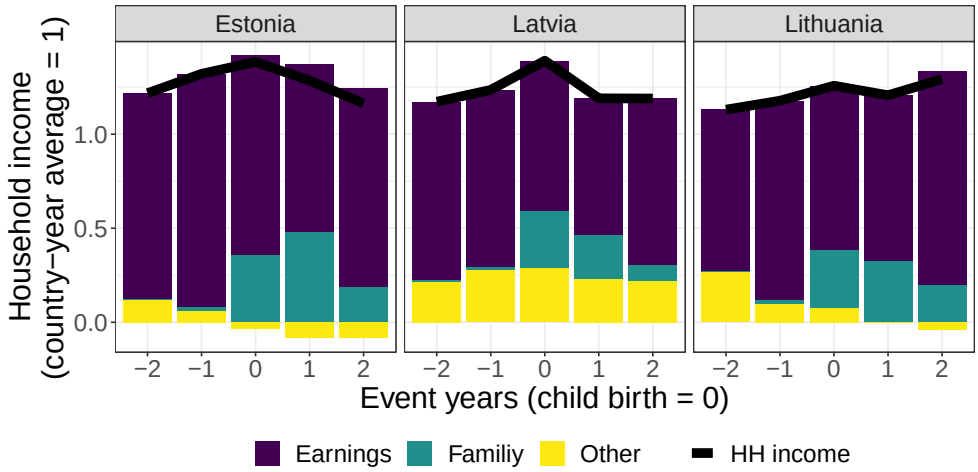
EU-SILC data is available in two versions which I merge for the chapter. One is the longitudinal rotational panel. In it, the same household appears up to 4 times for up to 4 different years. Each year a quarter of the households (those which have already been tracked for 4 years) are replaced with new ones. The second is cross-sectional, which contains all the respondents in a single year. Both versions reference the same individuals, but there are variables available in the cross-section variant (e.g., industry, variables related to childcare) that are not present in the longitudinal version. I merge them with several common variables.

While the data is explained by Eurostat (2018c), several features are mentioned here. Most covariates are recorded at the time of the interview, but income and months worked are recorded for the previous year (the reference year). In this chapter, all years represent reference years. Age is recorded at the end of the reference year. Unfortunately, the number of hours worked is recorded during the survey year, which means that it is not possible to calculate actual wage rates for the reference year. Finally, survey weights are used to form conclusions on the population from the sample data.

4.2.2. Descriptive statistics

Before presenting statistics of earnings of males and females separately, it is worthwhile to observe the general dynamics of earnings and income in general at the household level. Therefore, household earnings, along with other types of income for the Baltics are presented in Figure 4.1. Event time is plotted on the x-axis denoting the number of years before $(t < 0)$, during $(t = 0)$ and after $(t > 0)$ birth of the first child. Income, which consists of earnings, family benefits, and other income, is plotted on the y-axis. Income is demeaned to the average income of that country for that year. The general observation is that overall household income is similar before and after birth. Even though earnings fall in all three countries at $t = 0$, this effect is offset by rising family benefits. Therefore, households as a whole do not seem to lose income. The next step is to see how earnings evolve for males and females.

Figure 4.1: Household income and event time



The descriptive statistics from the EU-SILC survey are presented in Table 4.2. The columns indicate event time. The rows refer to values of demographic and labour market variables related to females and to the difference of values of variables related to females and males. Three observations can be made from the descriptive statistics.

The descriptive statistics point to a large impact of childbirth on labour market activities of females. For example, 77-78% of females earn positive earnings 2 years and 1 year prior to the birth of the child – a sign of high participation in the labour

market. At year 0, the year of the birth of the first child, fewer females earn any income (62%) and the share falls further in $t = 1$ to 35%. The decline is followed by a partial recovery in $t = 2$ as 59% of females participate in the labour market. Other labour market activity variables (normalized mean earnings, months worked and weekly hours) also follow a similar pattern and can be found in Annex Tables.

Descriptive statistics also indicate a widening labour market activity gap between females and males around the time the child is born. Two years prior to the birth of a child, 13 percentage points fewer females earn positive earnings as compared to men. This gap widens to 48 percentage points by $t = 1$. Widening of the labour market outcome gaps can be observed by comparing other labour market variables as well, such as mean monthly earnings, mean yearly earnings, mean months worked and mean weekly hours worked. This suggests that females' labour market participation is much more affected by childbirth as compared to that of men.

Finally, descriptive statistics suggest that several female variables that do not measure labour market outcomes also differ from those of men before childbirth. It is important to control for these differences to correctly estimate the effect of childbirth on labour market outcomes. On average, females are younger, more of them are still in education, have less work experience and tend to be more educated than men. Females and males also work in different sectors and in different occupations (see Table A24 in appendix). Males tend to work in construction, transport, and industry sectors while females tend to work in public administration, education, and the health sector. Similarly, men tend to occupy more management positions, while women occupy more professional positions before childbirth.

4.2.3. Methodology

The child penalty is estimated by treating childbirth as an event time study following Kleven, Landais, and Søgaard (2019). That is, I observe the individual's labour market variables before $t < 0$, during $t = 0$ and after $t > 0$ birth of first child[5]. I do this separately for men and women using a baseline equation specification in (4.1),

$$Y^g_{icst} = \sum_{j \neq -1} \alpha^g_j \cdot \mathbb{I}[j = t] + \sum_k \beta^g_k \cdot \mathbb{I}[k = age_{ics}] + \sum_y \gamma^g_y \mathbb{I} \cdot [y = s] + \sum_{cnt} \delta^g_{cnt} \mathbb{I} \cdot [cnt = c] + \nu^g_{icst} \quad (4.1)$$

5. This implies that I can only use data for which the event did occur - i.e. individuals with a first child birth. Including individuals without children would not allow me looking at the changes to earnings before and after childbirth.

Table 4.2: Descriptive statistics by sex and event time

Event time t	-2	-1	0	1	2
Positive earnings during the reference year					
Females	0.77	0.78	0.62	0.35	0.59
Females-Males	-0.13***	-0.12***	-0.28***	-0.48***	-0.31***
Mean earnings during the reference year (1 represents mean earnings for each country)					
Females	0.45	0.46	0.20	0.11	0.26
Females-Males	-0.31***	-0.36***	-0.60***	-0.62***	-0.66***
Mean monthly earnings during the reference year					
Females	1.03	1.02	0.75	0.48	0.68
Females-Males	-0.52***	-0.48***	-0.62***	-0.87***	-0.68***
Mean months worked during the reference year					
Females	8.25	8.62	5.83	4.84	6.86
Females-Males	-1.91***	-1.51***	-4.52***	-5.07***	-4.32***
Weekly hours worked during the survey year					
Females	27.26	22.64	12.34	18.00	24.56
Females-Males	-6.83***	-12.45***	-23.57***	-15.95***	-11.84***
Highest eduction attained					
Females	0.48	0.50	0.53	0.52	0.51
Females-Males	0.12**	0.09**	0.12***	0.09	0.01
Age at the end of the reference year					
Females	25.22	26.02	26.93	27.84	28.59
Females-Males	-3.33***	-3.17***	-3.14***	-2.62***	-3.45***
Percent living in an urban region					
Females	0.53	0.57	0.57	0.56	0.56
Females-Males	-0.04	-0.02	-0.02	0.01	0.02
Number of months spent in paid work during lifetime					
Females	54.40	58.18	62.37	68.67	74.88
Females-Males	-43.48***	-40.65***	-46.57***	-40.32***	-52.42***
Number of months spent studying in the reference year					
Females	2.29	1.80	0.88	0.25	0.19
Females-Males	1.87***	1.39***	0.75***	0.18**	0.13

Note: Table shows females statistics and the differences between female and male statistics. Event time denotes the number of years before $(t < 0)$, during $(t = 0)$ and after $(t > 0)$ birth of first child. Stars (***, **, *) indicate signicance at the 1%, 5% and 10% respectively and are calculated using r Survey package and Z critical values.

where Y represents earnings or another labour market outcome variable (e.g., hours or months worked), g represents gender and thus the equation is estimated separately for males and females. At one moment in time, we observe individual i in country c at year s and event time t. The key variable of interest is α in front of the event time dummies, $\mathbb{I}[j = t]$. If α_j takes on different values for $j = t \geq 0$ as compared to $j = t < 0$ this would mean that the birth of a child had an impact on labour market outcomes. Without any other controls, the results of (4.1) would be the same as those in Table 4.2. However, adding controls allows to more accurately estimate the child penalty. The remainder of the terms are age, year and country dummies and the coefficients in front of them are denoted β, γ and δ respectively. The intercept is also included although not shown in the equation. Dummy trap is prevented by omitting the event time $t = -1$ from the regression. This technique effectively sets α_{-1} to zero, allows comparing the effect of childbirth on labour market activities before and after childbirth.

While (4.1) is the **baseline** specification and similar to the equation used by Kleven, Landais, and Søgaard (2019), with the addition of country dimension, I construct an **extended** specification to account for additional controls. I fix the controls at their $t = -1$ levels. This allows controlling for occupation, education level and sector before the child was born. Admittedly, it is not clear whether these should be included. On the one hand, they allow to control for more factors, but on the other hand, they may also increase the risk of under evaluating the child penalty. This would be the case if men and women prepare for childbirth differently and well in advance.

I then follow Kleven, Landais, and Søgaard (2019) to use (4.1) to visualize the child penalty. I use the regression coefficients from the equation to estimate two versions of $\hat{Y}^g_{icst}$ for males and females. In the fist version, I model the counter factual $\hat{Y}^g$ where there is no child penalty (i.e. set $\alpha_j = 0, \forall j$). In the second, I estimate the child penalty by setting all other estimated coefficients to zero.

While equation (4.1) allows looking at how labour market activity changed for men and women separately, it does not allow to test whether the changes are significantly different for men and females. For this, I modify (4.1) by pooling men and women together. I add a dummy variable $male$ which takes the value of 1 if the observation is a man and zero otherwise. I also interact the $male$ dummy with event time dummies as summarized in (4.2)

$$Y_{icst} = male_{icst} + \sum_{j \neq -1} g_j \cdot \mathbb{I}[j = t \cap male = 1] + \sum_{j \neq -1} \alpha_j \cdot \mathbb{I}[j = t] + ... \quad (4.2)$$

While occupation and sector controls are also considered in alternative regressions, one has to be aware that including then can underestimate the effect of child penalty. This is the case if the choice of profession and sector is chosen expecting childbirth. For example, one can decide on a less rewarding but stable career if one was expecting to have children.

4.2.4. Results

The results suggest that there is a large child penalty in the Baltics for women. As shown in Table 4.3, there is no difference between normalised earnings for females leading up to childbirth(the coefficient in front of $t = -2$ is not statistically different from zero, hence, from $t = -1$). However, normalised earnings start declining rapidly thereafter. Female earnings fall by around a third in the year the first child is born (29.7 to 34.9 percentage points depending on the number of controls). The decline is even greater in one year after childbirth with about half of earnings declining. The decrease is still felt two years after childbirth, even though the effect is milder. There, again, women forgo around a third of their earnings following childbirth.

There is also a smaller child penalty for males. In Table 4.3 we again see that male earnings do not change much leading up to childbirth, but then fall by 11 to 15 percent points (depending on the number of controls) in $t = 0$. This can be explained by males often taking at least a paternity leave and sometimes paternity benefits (e.g., in Lithuania in 2011, around 60% of fathers took a paternity leave and 7% took parental leave based on national statistics and author's calculations). Male earnings further decline in $t = 1$ and become around a quarter smaller than pre-childbirth male earnings. The results are ambiguous about what happens with men earnings in $t = 2$. Using the baseline model specification, the earnings return to their pre-birth levels. However, the extended model specification suggests that male earnings remain 27 percent points lower than they were before the first child was born.

As a result, the labour market activity gap doubles for the first two years after child birth. In the final two columns of Table 4.3 we see that males tend to earn 25.6 percent points more than females in the Baltics under the baseline specification and 37 percent points more under the extended specification. These gaps widen in all post birth periods. By $t = 2$ the gaps increase by 29.6 percent points and 26.5 percent points respectively. The gap is also shown visually in Figure 4.2.

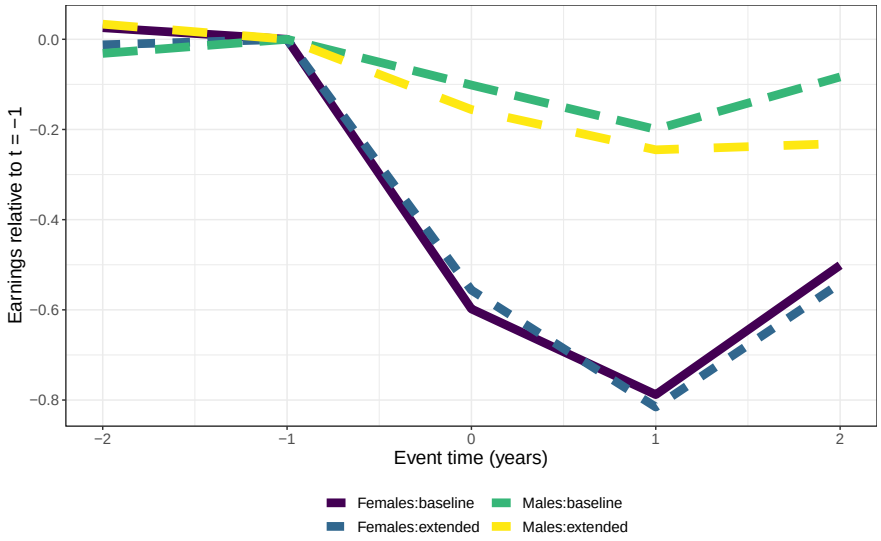
Table 4.3: Estimating child penalty on earnings

	Dependent variable:					
	Normalised earnings (1 represents average earnings for the country in that year)					
	Female		Male		Pooled	
	(1)	(2)	(3)	(4)	(5)	(6)
t=-2	0.027	−0.021	0.010	0.030	0.040	0.024
	(0.024)	(0.057)	(0.045)	(0.088)	(0.025)	(0.059)
t=0	−0.297***	−0.349***	−0.110**	−0.150**	−0.324***	−0.437***
	(0.019)	(0.037)	(0.046)	(0.059)	(0.022)	(0.051)
t=1	−0.410***	−0.531***	−0.191***	−0.241**	−0.461***	−0.631***
	(0.026)	(0.045)	(0.068)	(0.104)	(0.033)	(0.071)
t=2	−0.294***	−0.359***	−0.124	−0.270**	−0.348***	−0.429***
	(0.043)	(0.059)	(0.096)	(0.124)	(0.050)	(0.097)
male					0.256***	0.370***
					(0.062)	(0.123)
t=-2:male					−0.061	−0.018
					(0.053)	(0.095)
t=0:male					0.246***	0.304***
					(0.045)	(0.084)
t=1:male					0.306***	0.408***
					(0.069)	(0.125)
t=2:male					0.296***	0.265*
					(0.111)	(0.160)
Age dummies	T	T	T	T	T	T
Year dummies	T	T	T	T	T	T
Country dummies	T	T	T	T	T	T
Sector dummies		T		T		T
Occupation dummies		T		T		T
Education dummies		T		T		T
Constant	0.012	−0.220	−0.408***	1.167*	−0.153	0.424
	(0.135)	(0.335)	(0.155)	(0.592)	(0.113)	(0.669)
Observations	2,280	485	1,725	603	4,005	1,088
Log Likelihood	−1,139.439	−175.680	−2,133.676	−626.591	−3,952.701	−1,013.283
Akaike Inf. Crit.	2,388.877	483.360	4,397.353	1,413.182	8,051.401	2,198.566

*Note: Standard errors (found in parenthesis) incorporate information on the strata (country) and primary sample unit (individual) and therefore are clustered at the individual level. *p<0.1; **p<0.05; ***p<0.01*

While the above examples are centred around earnings, similar conclusions can be drawn for other labour market variables. Appendix Tables A25, A26, A27, A28 contain results for regressing on positive earnings, number of months worked, monthly wages and hours respectively. Females suffer a child penalty in all cases and the respective gaps widen. Men usually also experience a penalty, although they are more likely to get lower monthly wages or work fewer hours than not work at all. Women, however, are very likely to work fewer months or even not work at all during the year.

Figure 4.2: Child penalty on earnings



Note: Baseline specification refer to regression specification (1) and (3) in Table 4.3 for females and males respectively, while extended specifications refers to columns (2) and (4).

4.2.5. Maternity, paternity and parental leave policies

As the child penalty estimated here only focuses on the first three years of childbirth the results may be dependent on the leave policies. While I do not explicitly control for them (except via an intercept), the reader may be interested to know how the leave policies in the Baltics compare to his or her country. In particular, one should consider the duration, eligibility and compensation of three leave policies: maternity leave (when only the mother can temporarily leave work to take care of the child), paternity leave (when only the father can temporarily leave work to take care of the child), and parental leave (when either parent can temporarily leave work to take care

of the child). Since paternal policies last longer, they are likely to have the biggest impact on labour outcomes. I comment on policies that existed between the years of 2005 to 2017 (the year of the data that is investigated). In general, the leave policies were quite similar throughout the 2000s but became less generous due to cuts and ceilings introduced as a consequence of the 2008-2009 economic crisis.

The duration of parental leave lasted up to 3 years in the Baltics and the benefits were generous for those with previous earnings, at least for the first year (see, e.g., Karu and Bražienė 2014). In Estonia, the parent was compensated for 100% of previous earnings for the first 1.5 years and received a small monthly payment until the child was 3 years old. In Latvia, the parent had a choice whether to receive 60% of previous earnings for the first year and a small amount for the next half of the year or receive 43% of previous earnings for 1.5 years. In Lithuania, parents could choose whether to receive 100% of earnings for the first year or receive 70% for the first and 40% for the second. All countries had benefit ceilings or reduced payouts after a threshold and all have benefit floors. In Estonia, all parents are eligible, in Latvia – only the employed/self-employed and in Lithuania – only those with an insurance record. In most cases, parental leave benefits are reduced if parents choose to work while on parental leave. In Estonia, earnings are reduced if they exceed a threshold. In Latvia, only 30% of benefits are paid. In Lithuania, 1st year benefits are reduced, while second year benefits are not.

Mothers took parental leave much more often than fathers, although fathers were increasingly willing to do so. According to national statistics, in 2011, for example, mothers were 16 times more likely to go on parental leave than fathers in Estonia. However, this ratio fell steadily to 10 by 2017. In Lithuania, the situation is more nuanced. As long as the child did not reach 1 year of age, the age at which paternity benefits are reduced upon receiving earnings, mothers in Lithuania were 12 times more likely to be the ones on parental leave throughout the 2011–2017 period. However, mothers are increasingly less likely to be on parental leave in the second year after childbirth (ratio fell from 15 in 2011 to 1,9 in 2017), presumably because the father can continue working and still receive benefits.

In the Baltics, maternity benefits provide more incentives for mothers to take care of children than do paternity benefits for fathers in the first few months around the birth of the child (see, e.g.,). Maternity leave policies typically last 3 to 4 months around the birth of the child while paternity benefits last 28 days in Lithuania and 10 days in other two Baltic states. In all countries maternity and paternity benefits are all linked to previous earnings (typically 80% to 100% of previous earnings, except for 2009–2011 in Estonia, when no paternal benefits were paid). There are ceilings

(except for maternity benefits in Estonia) and no floors (except for maternity benefits in Estonia). Eligibility rules are the same for both paternity and maternity benefits: all employed and self-employed are eligible, while parents in Lithuania must have an insurance history.

4.3. Offsetting the child penalty with formal early childhood education and care

In this section, I test whether the child penalty can be reduced by providing formal early childhood education and care (ECEC) to families. Specifically, I compare the labour market outcomes for parents with children of ages 1 and 2 who did and who did not receive formal ECEC services. I control for both contemporaneous variables and labour market outcomes before birth. Including the latter allows accounting for the issue of reverse causality (that higher earnings lead to higher use of formal ECEC services and not vice versa). I also test whether expansion of formal ECEC services in the Baltics resulted in lower wage gap. For this, I apply the instrumental variable (IV) approach.

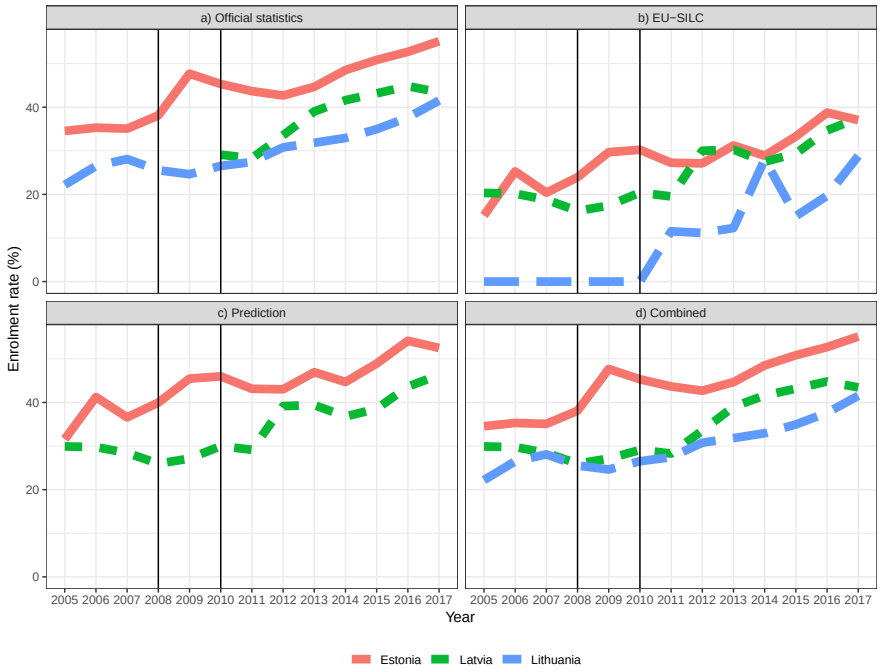
4.3.1. Early childhood education and care in the Baltics

ECEC services play an important role in the Baltic countries, even for children of 1 and 2 years old. Panel a in Figure 4.3 shows register data from official statistics, according to which the enrolment rate into official ECEC in Lithuania has increased from about 20% of all children aged 1-2 in 2005 to 40% in 2017. Similarly, the rates increased from around 37% to 50% in Estonia. Enrolment rates rose in Latvia also, although less data is available.

Even though register data on enrolment rates is not available for Latvia before 2010, I predict it using EU-SILC enrolment rates. Panel b of Figure 4.3 shows EU-SILC survey data. Although enrolment rates are lower (especially for Lithuania for the period before 2010) in EU-SILC data as compared to the register data, the trends are similar. Therefore, I predict register values for Estonia and Latvia in panel c based on the correlation between enrolment rates in the register and EU-SILC data. Upon combining register data from panel a with data predictions from panel c, I obtain estimates of register data prior to 2010 for Latvia. The final results are in panel d.

In all three states the vast majority of ECEC services are public and are provided by the municipalities (OECD 2016, 2017b, 2020). Additionally, the public ECEC services are offered at a very low cost. ECEC is free in Latvia (from 18 months) and in Lithuania (from 0 months, but only 20 hours per week, although some municipalities offer 40-hour free services) and the fees are very low in Estonia, where parent contributions are capped at 20% of the minimum wage and low income families are

Figure 4.3: Enrolment rates of children aged from 1 to 3 (excluding 3) into formal ECEC in pre-schools in the Baltics



Note: all four panels present information on enrolment rates of children of ages of 1 to 3 (excluding 3) into formal ECEC in pre-schools. Panel a contains information from official statistics (actual number enrolled into pre-school institutions, establishments or into pre-school education divided by the population in the respective age group). In panel b the enrolments rates are calculated from EU-SILC survey. In panel c enrolment rates of official statistics are predicted using EU-SILC enrolment rates for Estonia and Latvia with country fixed effects. Panel d contains the same information as panel a with the missing values for Latvia taken from panel c.

exempted (Motiejunaite-Schulmeister et al. 2019). Parents in all countries have to pay additional expenses (such as transport and food), but these are relatively low and should not be a barrier to entry (Motiejunaite-Schulmeister et al. 2019). Provision of ECEC services has become guaranteed for children 18 months old and older in Estonia since 1999 and in Latvia since 2011 (OECD 2016), which means that the municipalities must provide enough places for the children if parents request it, and the provision of services has expanded since 2009 (Motiejunaite-Schulmeister et al. 2019)

Under such conditions, demand for ECEC services has been exceeding supply.

Table 4.4: Places at pre-school institutions and the number of pre-school institutions

	Years				Growth rate (%)	
country	2006	2010	2013	2016	2006-2010	2010-2016
Places, ths.						
LT	87	92	113	123	6	34
Institutions, nr.						
EE	620	638	652	635	3	0
LT	652	626	675	737	-4	18
LV	557	591	617	647	6	9

Source: National statistics and author calculations.

Parents who wanted to take their kids to formal ECEC had to wait in long queues, especially in the cities (Aidukienė 2014; OECD 2016, 2017b, 2020). To meet the demand, governments have been increasing the supply of ECEC services. In Estonia since 2009, extra funds have been allocated to create new places for children under 3 in urban areas. Even though formal ECEC is not guaranteed in Lithuania for children of such a young age, the municipalities in Lithuania, along with the municipalities in other Baltic countries, have expanded provision of services either themselves or via co-financing arrangements with third parties (OECD 2016, 2017b, 2020). This is most visible in the case of Lithuania, where the number of places at preschool education rose by 34% in 2010–2016 alone (see Table 4.4) and 80% of this expansion was due to public places[6]. Unfortunately, such statistics are not available for other Baltic countries, which requires looking at the number of institutions as a proxy for the dynamics of capacity[7]. We see in Table 4.4 that the number of institutions increased by 18%, 9% and 0% in Lithuania, Latvia and Estonia respectively since 2010.

While enrolment into ECEC services of ages 1-2 have been growing during in the 2005–2017 period in all three Baltic states, the growth was observed in different times. For example, there is a clear jump in enrolment rates in Estonia from 2008 to 2009, which coincides with the expansion of ECEC in urban areas. Judging from panel d of Figure 4.3, ECEC enrolment rates decreased in this period in the two other Baltic countries. There was a gradual increase in enrolments since 2011 in Lithuania and

6. Private pre-schools started playing a bigger role only very recently. There were very few private preschools before 2010.

7. Pre-school institutions provide ECEC services for children of all ages, not just 1-2. Additionally, there have been simultaneous structural changes. Many rural pre-school education institutions closed down as people moved to the cities. This is why there is a net decrease in the number of pre-school education institutions in the 2006-2010 period in Lithuania and no increase in the number of institutions for Estonia in the 2010-2016 period.

Latvia, which also coincides with expansion of ECEC services, as seen in Table 4.4. This enforces the assumption that rising enrolment rates were driven by supply factors (a mixture of legislature changes and increasing capacity).

Besides formal childcare, informal childcare should be also accounted for parent market outcomes. The only data on non-preschool childcare that exists comes from EU-SILC, although this data is not very reliable (Vandenbroeck et al. 2018). Nevertheless, EU- SILC data suggests that unofficial childcare was also important in this period, especially in Lithuania and Estonia. Unofficial childcare is childcare by grand-parents, other household members outside of parents, other relatives, friends, or neighbours. Additionally, unofficial childcare may have increased in this period. In 2016, 42% of Lithuanian children aged 1 and 2 and a third of Estonian children received such childcare. This compares to a mere 4% and 14% in 2006 respectively. Unofficial childcare grew in Estonia too, although the growth was much less pronounced. Even though other forms of official care outside of ECEC exist, its role was negligible by 2010 in all three Baltic countries.

4.3.2. Estimating the effect of early childhood education and care

In this section, I estimate the effect of formal ECEC in pre-school on earnings of parents. First, I run several regressions and then I implement an instrumental variable (IV) approach. In both cases, I focus on the outcome of parent's earnings at the time when their first child was, on average, 1 or 2 years old. The data comes from the EU-SILC data set for the Baltic states described in subsection 4.2.1. In the estimations, I include the original survey weights and stratify the data at a country level.

II first regress whether or not a parent enrolled his or her child into formal ECEC on that parent's earnings. Specifically, the dependent variable is the level of earnings of individual i of sex g in country c in year s. The independent variable is an enrolment dummy, signifying whether or not his/her first 1–2-year-old child was enrolled in ECEC. I also include country, year, and age dummies. While this regression is simple, its results should be treated with caution as it does not yet address the issues of reverse causality.

The results of the first regression suggest that pre-school has a greater impact on female earnings than on male earnings. The correlation is positive and significant for females with a point estimate of 0.1 (see column (1) of Table 4.6) but not significant for males (see column (3) of Table 4.6).

Next, I supplement the regression with additional controls. Specifically, I add education dummies and earnings 1 year before the birth of a child. The second control is particularly important, as it is a quick test for reverse causality – it tests whether peo-

Table 4.5: Early education and childcare of ages 1-2 in EU-SILC, percent

	Years				Difference	
country	2006	2010	2013	2016	2006-2010	2010-2016
Education at pre-school						
EE	13	25	21	36	12	10
LT	0	10	22	54	10	44
LV	14	16	29	48	3	32
Other official childcare						
EE	3	2	0	8	0	6
LT	22	1	7	0	-20	-1
LV	1	3	0	4	2	1
Unofficial childcare						
EE	14	25	23	32	11	6
LT	4	16	20	42	13	25
LV	3	10	11	12	6	2
Total						
EE	26	40	38	56	14	16
LT	25	26	38	80	1	53
LV	18	27	33	52	9	26

Note: other official childcare refers to childcare at day-care centre, by a professional child-minder at child's home or at child-minder's home. Unofficial childcare refers to childcare by grand-parents, other household members (outside parents), other relatives, friends or neighbours.

Table 4.6: Estimation of the effect of pre-school enrolment on parent's earnings

	Dependent variable:					
	Normalised earnings (1 represents average earnings for the country in that year)					
	Female		Male		Pooled	
	(1)	(2)	(3)	(4)	(5)	(6)
enrolment	0.101***	0.149**	0.053	−0.012	0.096***	0.122*
	(0.022)	(0.061)	(0.076)	(0.056)	(0.024)	(0.067)
enrolment:male					−0.031	−0.156*
					(0.078)	(0.087)
male					0.547***	0.266***
					(0.035)	(0.063)
earnings (before child)		0.167**		0.499***		0.095
		(0.070)		(0.081)		(0.074)
other official childcare		0.219*		0.355***		0.295***
		(0.117)		(0.134)		(0.094)
unofficial childcare		0.068*		0.056		0.073*
		(0.041)		(0.078)		(0.042)
male:earnings (before child)						0.432***
						(0.101)
age dummies	T	T	T	T	T	T
year dummies	T	T	T	T	T	T
country dummies	T	T	T	T	T	T
education dummies		T		T		T
Constant	0.214***	−0.291***	1.116***	0.346	0.380***	−0.146
	(0.069)	(0.109)	(0.159)	(0.400)	(0.088)	(0.232)
Observations	2,767	568	2,223	460	4,990	1,028
Log Likelihood	−1,426.255	−161.722	−2,735.288	−359.343	−4,933.933	−656.962
Akaike Inf. Crit.	2,950.509	423.443	5,592.575	832.686	9,995.867	1,435.924

*Note: Standard errors (found in parenthesis) incorporate information on the strata (country) and primary sample unit (individual) and therefore are clustered at the individual level. *p<0.1; **p<0.05; ***p<0.01*

ple with higher earnings are the ones who used ECEC services, which would explain why their earnings are higher. Additionally, I add dummies for use of other official and unofficial means of childcare.

The results of the second regression indicate that child's attendance of pre-school education increases earnings of females even more than suggested by the previous regression. Not only does the coefficient in front of the enrolment dummy remain significant for females and insignificant for males, but the point estimate for females becomes bigger – it rises to 0.15 (see columns (2) and (4) of Table 4.6).

In the third regression, I run a pooled regression model with males and females in a single regression. The upside of this approach is that I test whether earnings of females increase more than males while simultaneously benefiting from a higher number of degrees of freedom. The downside is that this model is more restrictive, because it assumes that all other coefficients are equal for men and women (e.g., the effect of education is the same for males and females). In this model, earnings of females increase significantly more than that of males, but only if additional controls are included (see columns (6) and (5) of Table 4.6). It seems especially important to include previous male earnings to avoid suffering from severe forms of reverse causality.

Although the above regressions support the hypothesis that women benefit more from ECEC than men, it does not suggest that increasing formal ECEC in pre-schools has benefited women earnings. For example, it could be that more pre-school education simply crowded out other forms of childcare. Indeed, the regression estimates suggest that other official and unofficial childcare is also positively correlated with earnings. On one hand, the EU-SILC data suggests that there was a simultaneous increase in official ECEC and unofficial ECEC meaning that this is unlikely. On the other, EU-SILC data on unofficial ECEC is less reliable. Therefore, an additional test could help investigate this.

To address this issue, I take an instrumental variable approach. In theory, this allows me to see what happened to earnings of males and females at a time when enrolments into formal pre-school ECEC increased and all other factors remained the same. In other words, we expect to see earnings increase only if parents would have otherwise stayed home with their kids. This is done in two stages.

In the first stage, I regress whether individual i of sex g in country c in year s has enrolled his 1-2 year old child into formal pre-school ECEC on the overall formal pre-school enrolment rate in the country in year ENR^g_{ics}. Since most of the changes in pre-school education can be explained by supply side effects, this assumption seems reasonable. This implies that greater access to ECEC services in the

country should have increased individual's probability of enrolling his or her child to preschool. Therefore, equation (4.3) allows me to obtain an estimate of whether an individual did enrol his or her 1–2-year-old child into pre-school, $\hat{enrolment}^g_{ics}$, at times when access increased. I also add controls that are the same as those in the second stage.

$$enrolment^g_{ics} = \alpha_1 + \beta^g_1 ENR^g_{cs} + \gamma_1 controls + \nu^g_{1ics} \tag{4.3}$$

In the second stage, I estimate whether $\hat{enrolment}^g_{ics}$ had an impact on earnings. Specifically, I regress individual's i of sex g in country c in year s predicted child's enrolment $\hat{enr}^g_{ics}$ on the individual's earnings Y^g_{ics}

$$Y^g_{ics} = \alpha_2 + \beta^g_2 \hat{enrolment}^g_{ics} + \gamma_2 controls + \nu^g_{2ics} \tag{4.4}$$

and examine the value of β_2. A positive β_2 suggests that earnings increased due to enrolment. I control for age, country, education. I try adding both year dummies and more precise variables such as unemployment rate and GDP per capita.

The estimates for the first stage suggest positive correlation between country enrolment rates and individual enrolment rates. Furthermore, the coefficient of enrolment is close to 1 when only the instrument and the constant are included (see columns (1) and (3) in Table 4.7). This is expected, as the average enrolment rate in the country should match the average probability of an individual enrolling his/her child in EU-SILC data. Additionally, the instrument is strong, as indicated by the F-test. For an instrument to be strong, the F test of removing the instrument should be above 10, while in this case, it is 88 and 74 respectively.

Nevertheless, the instrument becomes weak if more controls are added. In columns (2) and (4), the F-statistic of excluding the instrument fall to 5.7 and 1.0 for females and males respectively. Therefore, the results are only suggestive. Interestingly, adding more control variables reduces the magnitude of the enrolment coefficients, but they remain significant.

Moving to the second stage, the results are very dependent on the model specification. If we do not include earnings before the birth of the child, then results suggest that enrolment rates do not significantly affect earnings of women but increase the earnings of men (see line 1 on column (1) and (3) in Table 4.8). If we do include earnings before the birth of the child, the results reverse: higher enrolments lead to higher earnings for women, but not for men (see column (2) and (4)). The message

Table 4.7: First stage: effect of pre-school institutions on enrolments

	Dependent variable:					
	enrolment into pre-school					
	Female		Male			
	(1)	(2)	(3)	(4)	(5)	(6)
enrolment (combined)	1.039***	0.991**	0.975***	0.506	1.188***	0.837**
	(0.110)	(0.414)	(0.113)	(0.494)	(0.180)	(0.330)
age		0.088**		0.047**		0.053***
		(0.040)		(0.019)		(0.019)
age^2		−0.001**		−0.001**		−0.001***
		(0.001)		(0.0003)		(0.0003)
male					0.012	0.063
					(0.051)	(0.111)
real GDP		0.0001*		0.00005		0.0001**
		(0.00003)		(0.00003)		(0.00002)
unemployment rate		0.017**		0.014**		0.015***
		(0.008)		(0.007)		(0.006)
earnings (before child)		0.014		−0.040		−0.001
		(0.070)		(0.028)		(0.067)
combined:male					−0.068	−0.212
					(0.157)	(0.281)
male:earnings (before child)						−0.022
						(0.069)
country dummies		T		T		T
education dummies		T		T		T
Constant	−0.178***	−2.534***	−0.167***	−1.693***	−0.273***	−1.940***
	(0.037)	(0.864)	(0.037)	(0.530)	(0.077)	(0.462)
F-test of removing instruments	88.49	5.72	74.13	1.05	43.73	6.42
Observations	2,767	568	2,223	460	4,990	1,028
Log Likelihood	−1,691.368	−350.893	−1,307.790	−279.049	−2,975.133	−634.570
Akaike Inf. Crit.	3,386.736	727.786	2,619.581	584.098	5,962.266	1,301.141

*Note: Standard errors (found in parenthesis) incorporate information on the strata (country) and primary sample unit (individual) and therefore are clustered at the individual level. *p<0.1; **p<0.05; ***p<0.01*

is even less clear in the pooled model. If earnings before the birth of a child are not included, then earnings of both men and women increase upon enrolment, but men's earnings increase more (see the line 2 on column (5) in Table 4.8). If earnings before the child are included, then women do not seem to benefit from earnings at all, while men do (column (6)). Therefore, it is not clear whether this is an issue of the pooled model being too restrictive or that women indeed do not benefit as much as suggested. Even though the results are not shown, adding year dummies makes the results also insignificant.

Table 4.8: IV regression: effect of pre-school enrolment on parent's Earnings

	Dependent variable:					
	earnings					
	Female		Male		Pooled	
	(1)	(2)	(3)	(4)	(5)	(6)
enrolment	0.139	0.164*	1.560*	−0.183	0.459**	−0.125
	(0.109)	(0.100)	(0.811)	(0.772)	(0.225)	(0.226)
age	0.062***	−0.010	0.099***	0.034	0.070***	0.028*
	(0.007)	(0.013)	(0.030)	(0.043)	(0.013)	(0.017)
age2	−0.001***	0.0004*	−0.001**	−0.0004	−0.001***	−0.0003
	(0.0001)	(0.0002)	(0.0005)	(0.001)	(0.0002)	(0.0002)
unemployment rate	−0.006***	−0.009***	−0.017	0.001	−0.011**	−0.004
	(0.002)	(0.003)	(0.012)	(0.014)	(0.004)	(0.006)
GDP per capita (real)	−0.00001	−0.00001	−0.0001*	0.0001	−0.00005***	0.00002
	(0.00001)	(0.00001)	(0.00005)	(0.0001)	(0.00002)	(0.00003)
earnings (pre-birth)		0.154***		0.591***		0.194***
		(0.023)		(0.062)		(0.032)
male					0.547***	0.134**
					(0.093)	(0.054)
enrolment:male					0.682*	0.494*
					(0.369)	(0.299)
male:earnings (pre-birth)						0.392***
						(0.054)
Constant	−0.860***	0.244	0.035	−1.219	−0.580*	−0.688
	(0.157)	(0.304)	(0.901)	(1.602)	(0.317)	(0.615)
Degrees of freedom	1814	365	1461	292	3281	665
Observations	2,767	575	2,223	464	4,990	1,039
R^2	0.109	0.215	−0.468	0.528	0.088	0.558
Adjusted R^2	0.106	0.204	−0.473	0.520	0.086	0.553
Residual Std. Error	10.236 (df = 2759)	9.642 (df = 566)	27.691 (df = 2215)	17.193 (df = 455)	18.670 (df = 4978)	14.221 (df = 1027)

*Note: Standard errors (found in parenthesis) incorporate information on the strata (country) and primary sample unit (individual) and therefore are clustered at the individual level. *p<0.1; **p<0.05; ***p<0.01*

4.4. Conclusion

This chapter examines the size of the child penalty at times when the first child is up to three years old in the Baltics. I find that women's labour market outcomes respond to childbirth much more strongly than those of males. Earnings of women decrease by 30–43% while male earnings decrease by 5–27% on average two years

after childbirth, depending on the model specification. Since there exists a 26–37% wage gap before childbirth, this effectively doubles the gap for several years. The widening wage gap is observed on other variables (months and hours worked, wages). Men usually also experience a penalty, although they are more likely to get lower monthly wages or work fewer hours than not work at all. Women, however, are very likely to work fewer months or even not work at all during the year.

I then examine how much does enrolment into pre-school education (a type of formal early childhood education and care) offset the child penalty when the child is up to 3 years old. I find that women tend to earn up to 15% more if they enrol children into preschool while men do not, even when controlling for earnings before childbirth. Men do not seem to benefit in this case. This suggests that ECEC can reduce the widening wage gap by up to a half. However, other forms of childcare seem also to be prevalent in the Baltics and widening access to formal ECEC can simply means that informal childcare becomes less used with no actual effect on earnings. On the one hand, this does not seem to be the case as there is a rise in both formal and informal childcare. Additionally, an instrumental variable regression, when country-wide enrolment rates are included as a proxy for access to child care, are sometimes in line with simple regressions. On the other hand, the instrumental variable regressions are very sensitive to the model and the controls that are included, suggesting a need for more granular data to cement the conclusions.

Therefore, this study faces several data limitations. I do not have access to longer term data, which prevents examining when and to what extent does this additional wage gap fade. Additionally, the instrumental variable test could benefit from more granular data (e.g., on ECEC capacity and children's enrolment into ECEC by region). Finally, I only consider the effect of ECEC on parent's earnings. Other factors are also important to consider when evaluating whether more formal ECEC is indeed a good tool to avoid the widening wage gap several years after childbirth. But this requires more data on the children who enrol into ECEC. Hopefully, increasing access to the register in the Baltic countries will address these issues.

CONCLUSION

The overall conclusion is that income inequality in Lithuania has been relatively high as compared to other countries in 2007 to 2015 and this is partly due to a low role played by the tax and benefit system.

As shown in chapter 1, equivalised disposable income inequality in Lithuania has been higher than in the EU and even the new EU states (those Member States that joined it after 2004). Income inequality has been high irrespective of the inequality

measure used or the equivalence scale used. This means that households differed in terms of their incomes in Lithuania more than in other EU states.

Additionally, income inequality was found to be higher due to within-group inequality rather than between-group inequality. For example, there was high income inequality amongst the unemployed, i.e., some households with unemployed household members had very little incomes while others had quite high incomes, and the inequality within this subgroup was larger than in other EU countries. A similar observation was made for those working within the agricultural sector which seems to be particularly unequal in Lithuania. However, inequalities within most subgroups were larger in Lithuania than in the EU, suggesting that reducing income inequality within a particular subgroup would have a limited effect on overall income inequality.

Among the groupings that were constructed in this research, income inequality between the employed and the rest of the population was especially large. This was partly due to the relatively unequally distributed labour income in Lithuania, which became even more unequal. In addition, there has been little progressivity in the taxation of labour income (the tax system redistributes income two times less in Lithuania than in the EU). Furthermore, there were few alternative sources of income available for those not employed. Many households with little or no labour income were dependent on public transfers (e.g., pensions or unemployment benefits). However, these were about two times lower in Lithuania than in the EU.

There are potentially two ways to reduce income inequality between those in and out of employment. One is to increase transfers to those who are out of employment. The other, is to encourage those out of work to start working. Based on the model proposed in this dissertation (see chapter 3), it is optimal to encourage people to start working. This means that out of work benefits (unemployment benefits, poverty relief benefits, family benefits and other benefits) should be reduced for working-age people while in-work benefits, e.g. via income tax credit, should be provided. If more people become employed, income inequality should diminish. However, it is important to note that the proposed model suffers from several limitations. It does not consider the supply of jobs or the ability of those currently out of work to start working. The importance of supply of jobs is especially relevant during crisis periods, when the number of vacancies falls, and unemployment rises. While this special case can be avoided by increasing unemployment benefits during crisis periods, this model should be improved prior to enacting it in practice.

While labour income has been more unequally distributed and less redistributed in Lithuania as compared to the EU, the situation has been even more acute for self-employment income. There were three main issues identified with self-employment

income and the taxation of it. First, up to 2/3 of self-employment income may have not been revealed to tax authorities, which meant a large portion of this income was untaxed. Second, the tax rates for the self-employment income are much lower than for the employed. Third, the tax rates are not progressive for the self-employed and are regressive in some cases (e.g., health insurance contributions have floors while business certificates are lump sum taxes). Finally, a large share of those who have self-employment income tend to be at the top of the income distribution. Therefore, closer auditing of the self-employed and reconsideration of the term structure would have been, and still is, appropriate.

Higher tax rates and a more progressive tax regime for the self-employed should reduce income inequality in Lithuania. The self-employed are a very heterogeneous group, with many self-employed earning little and relatively few earning very high incomes. Therefore, income inequality amongst the self-employed is large. Tax and benefit policies should recognize this. While, upon obtaining reliable tax rate elasticities of taxable self-employment income, it should make sense to raise the tax rates of high-earning self-employed, this is not true for the self-employed earning little incomes. The case for treating low-earning self-employed differently would be even stronger if in-work benefits would be provided for the employed.

Future studies could also investigate how tax and benefit policies impact within- and between-subgroup inequality of other groupings. For example, it would be interesting to learn why income inequality amongst the unemployed is large in Lithuania. One possible reason is that unemployment benefits are of relatively short duration and are dependent on work history. Therefore, only about a third of registered unemployed were eligible for unemployment benefits during COVID-19 pandemic in 2020 (Lazutka 2020). Additionally, unemployment benefits depend on previous earnings. This means that the short-term unemployed with high previous earnings would enjoy much higher benefits than others. Another reason is that there is a lack of information on unemployment benefits as well as unwillingness to apply for unemployment benefits. A fourth issue is the issue of long term unemployed who are not eligible for unemployment benefits and often suffer from complex conditions preventing them from re-entering the labour market (Ministry of Social Secuirty and Labour 2020). With the help of a specialized data set, one can quantify the relative contributions of these issues. Similar in-depth studies would be needed for each separate grouping.

Not only has income inequality in Lithuania been high as compared to other countries, but it has also tended to increase over the period. The main reason for the increase has been growing returns – wages and capital income. As average wages rose during the period, particularly after 2011, there was an increasingly widening

gap between those who had access to labour income and those who did not. During this period social transfers were not increased sufficiently (they increased less than the nominal wage growth). This was both because public transfers were not indexed to nominal wage growth and because discretionary spending was not adjusted accordingly. The situation was even more difficult for those who were not eligible for transfers, such as the long term unemployed.

While rising wages and the slowly-responding tax-and-transfer system were found to be the main reasons behind rising income inequality in 2007–2016, unfavourable demographic dynamics also increased income inequality. As decomposition techniques have revealed, income inequality was also associated with declining marriage rates in Lithuania and rising higher education attainment rates.

There are several possible ways to explain the link between declining marriage rates and income inequality. It is possible that income inequality affects marriage rates: if richer individuals would be less inclined to marry poorer individuals, then higher income inequality would limit partner choice and lead to fewer marriages. However, it is also likely that declining marriages rates influence income inequality. Because income inequality is larger amongst non-married households than amongst married households, a rising share of non-married households drives up income inequality.

Unfortunately, the methodology was not able to reveal why income inequality amongst married households is smaller than between non-married households. One possible explanation is purely mathematical. Non-married households typically have a single working aged individual who may face an extreme outcome (be very rich or very poor). Married households have two or more individuals and their outcomes (positive or negative) average out. Therefore, married households have a lower probability to face extreme outcomes, because this would require both household members to be very rich or very poor. Another explanation is a behavioural one: married males tend to work longer and more intensely than not married households, which means that fewer married households should deal with unemployment or low incomes. A third explanation could be that each partner can better compensate for each other's weaknesses and lead to fewer mistakes that lead to very poor outcomes.

Similarly, the reasons why rising education increased income inequality are also not clear. One likely explanation is the transition effect: until a certain fraction of individuals receives higher education, income inequality will increase. But once that fraction is reached, income inequality will start decreasing. However, other possibilities also exist, and this could be an interesting avenue for future work.

Not only does income inequality change between and within subgroups, but it

also changes across the lifetime of a person. For example, earnings were reduced upon retirement and low pension replacement rates resulted in a wide gap between those with and those without labour income. Incomes were also affected by the birth of the child. Women earnings declined by almost 50 percent 2 years after the birth of the child (largely because they left the labour market), double the 25 percent decline observed in Scandinavian and English-speaking countries, but slightly below that of German speaking countries (Kleven, Landais, Posch, et al. 2019).

Even though most research within this dissertation was focused on the tax-and-benefit policy to lower income inequality in Lithuania in 2007–2016, it is not the only way forward. In many cases specific targeted measures can be better. For example, the rising inequality in earnings between males and females following childbirth discussed in chapter 4 can be offset by measures other than providing more benefits. In fact, higher parental benefits alone may further increase income inequality between genders by keeping women away from the labour market longer. Instead, more focus can be placed on redesigning the parental benefits so that fathers participate in childcare more, make it easier for parents to simultaneously work and take care of small children or provide early childhood education and care (ECEC). Even though conclusive evidence was not obtained during the investigation, the collected evidence suggests that women did tend to suffer smaller earning losses if they used ECEC services. Furthermore, the limited available evidence suggests that societal norms, rather than benefits, have a far larger impact on child penalty (Kleven, Landais, Posch, et al. 2019).

A lot is still left to do on the subject on income inequality in Lithuania. With the advent of opening of administrative data, the methods applied in this dissertation can be carried out with much greater accuracy and detail. One will be able to obtain a more accurate estimate of tax rate elasticity of taxable income for Lithuania, more granular decompositions, better estimates of income inequality at the top of the income distribution and an estimate of long-term child penalty. Moreover, other factors behind income inequality can also be examined, such as the effect of the change of minimal wage or the decision to freeze public wages. One will be also able to go beyond individuals and investigate income inequality at the firm level. This will allow broadening the concept of inequality even further by including wealth, public goods, and consumption taxes. Therefore, this seems like the beginning and not the end of this empirical examination of income inequality in Lithuania.

Finally, since income inequality seems to be an important indicator for Lithuania, regular analysis of its dynamics could be useful. For example, income inequality has been declining since 2015 in Lithuania. Gini coefficient of equivalised disposable

income before social transfers declined from 52.6 to 50.3 in 2018. Similarly, Gini coefficient of equivalised disposable income fell from 37.0 to 35.4. Such a decline is very welcome by those advocating for less income inequality. However, future studies are needed to understand why this has happened.

BIBLIOGRAPHY

Aghion, Philippe, Eve Caroli, and Cecilia Garcia-Penalosa. 1999. Inequality and Economic Growth: the Perspective of the New Growth Theories. *Journal of Economic Literature* 37 (4): 1615–1660.

Aidukienė, Teresa. 2014. Lithuanian experience: towards accessible and inclusive early childhood education. Conference: Bucharest, 7–8 April 2014. https://www.worldbank.org/content/dam/Worldbank/Event/ECA/romania/Teresa%20Aidukiene-Lithuania-2014-04-02-.pdf.

Almunia, Miguel, and David Lopez-Rodriguez. 2019. The elasticity of taxable income in spain: 1999–2014. *SERIEs,* 1–40.

Alstadsæter, Annette, Niels Johannesen, and Gabriel Zucman. 2017. *Tax Evasion and Inequality.* NBER Working Paper 23772. Cambridge: National Bureau of Economic Research.

Alvaredo, Facundo, Lucas Chancel, Thomas Piketty, Emmanuel Saez, and Gabriel Zucman. 2018. *World inequality report 2018.* Harvard University Press.

Anspal, Sten. 2015. Essays on gender wage inequality in the Estonian labour market. PhD diss.

Anspal, Sten, et al. 2015. Gender wage gap in Estonia: a non-parametric decomposition. *Baltic Journal of Economics* 15 (1): 1–16.

Anspal, Sten, Tairi Rõõm, S Anspal, L Kraut, and T Rõõm. 2011. Gender pay gap in Estonia: empirical analysis. *Report for the Estonian ministry of social affairs. Tallinn: Ministry of Social Affairs.*

Artavanis, Nikolaos, Adair Morse, and Margarita Tsoutsoura. 2016. Measuring income tax evasion using bank credit: evidence from greece. *The Quarterly Journal of Economics* 131 (2): 739–798.

Atkinson, Anthony B, et al. 1970. On the measurement of inequality. *Journal of economic theory* 2 (3): 244–263.

Atkinson, Anthony B., and Thomas Piketty. 2010. *Top Incomes: A Global Perspective.* Oxford University Press.

Atkinson, Anthony B., Thomas Piketty, and Emmanuel Saez. 2011. Top incomes in the long run of history. *Journal of Economic Literature* 49, no. 1 (March): 3–71.

Baker, Michael, Jonathan Gruber, and Kevin Milligan. 2008. Universal child care, maternal labor supply, and family well-being. *Journal of political Economy* 116 (4): 709–745.

Baldini, Massimo, Paolo Bosi, and Michele Lalla. 2009. Tax Evasion and Misreporting in Income Tax Returns and Household Income Surveys. *Politica economica* 25 (3): 333–348.

Baliamoune-Lutz, Mina, and Pierre Garello. 2014. Tax Structure and Entrepreneurship. *Small Business Economics* 42 (1): 165–190.

Bargain, O. 2012. Back to the future: decomposition analysis of distributive policies using behavioural simulations. *International Tax and Public Finance* 19 (5): 708–731.

Bargain, O., and T. Callan. 2010. Analysing the effects of tax-benefit reforms on income distribution: a decomposition approach. *Journal of Economic Inequality* 8 (1): 1–21.

Bargain, Olivier, Kristian Orsini, and Andreas Peichl. 2014. Comparing labor supply elasticities in europe and the united states new results. *Journal of Human Resources* 49 (3): 723–838.

Barrios, Salvador, Viginta Ivaškaitė-Tamošiūnė, Anamaria Maftei, Edlira Narazani, and Janos Varga. 2019. *Progressive Tax Reforms in Flat Tax Countries.* Euromod Working Paper Series 2. Essex: University of Essex.

Barsky, Robert, John Bound, Charles Kerwin, and Joseph Lupton. 2002. Accounting for the black-white wealth gap: a nonparametric approach. *Journal of the American Statistical Association* 97 (459).

Berg, Andrew G, and Jonathan D. Osrty. 2011. *Inequality and Unsustainable Growth: Two Sides of the Same Coin?* International Monetary Fund.

Berger, Yves G. 2008. *A Note on the Asymptotic Equivalence of Jackknife and Linearization Variance Estimation for the Gini coefficient.*

Biewen, Martin. 2014. A general decomposition formula with interaction effects. *Applied Economics Letters* 21 (9): 636–642.

Biewen, Martin, and Stephen P Jenkins. 2005. A framework for the decomposition of poverty differences with an application to poverty differences between countries. *Empirical Economics* 30 (2): 331–358.

Biewen, Martin, and Andos Juhasz. 2012a. Understanding rising income inequality in germany, 1999/2000–2005/2006. *Review of Income and Wealth* 58 (4): 622–647.

———. 2012b. Understanding rising income inequality in Germany, 1999/2000-2005/2006. *Review of Income and Wealth* 58 (4): 622–647. ISSN: 1475-4991. https://doi.org/10.1111/j.1475-4991.2012.00514.x. http://dx.doi.org/10.1111/j.1475-4991.2012.00514.x.

Blau, David, and Janet Currie. 2006. Pre-school, day care, and after-school care: who's minding the kids? *Handbook of the Economics of Education* 2:1163–1278.

Boll, Christina, and Andreas Lagemann. 2018. Gender pay gap in EU countries based on SES (2014). *Luxembourg, Publication Office of the European Union. doi* 10:978935.

Bourguignon, François, Francisco HG Ferreira, and Phillippe G Leite. 2008. Beyond oaxaca–blinder: accounting for differences in household income distributions. *The Journal of Economic Inequality* 6 (2): 117–148.

Bruce, Donald. 2002. Taxes and Entrepreneurial Endurance: Evidence from the Self-Employed. *National Tax Journal,* 5–24.

Buhmann, Brigitte, Lee Rainwater, Guenther Schmaus, and Timothy M Smeeding. 1988. Equivalence Scales, Well-being, Inequality, and Poverty: Sensitivity Estimates Across Ten Countries Using the Luxembourg Income Study (LIS) Database. *Review of Income and Wealth* 34 (2): 115–142.

Burtless, Gary. 1999. Effects of growing wage disparities and changing family composition on the us income distribution. *European Economic Review* 43 (4-6): 853–865.

Callan, Tim, Karina Doorley, and Alyvia McTague. 2020. Top incomes in ireland: reconciling evidence from tax records and household survey data.

Carta, Francesca, and Lucia Rizzica. 2018. Early kindergarten, maternal labor supply and children's outcomes: evidence from italy. *Journal of Public Economics* 158:79–102.

Causa, Orsetta, and Mikkel Hermansen. 2017. Income redistribution through taxes and transfers across oecd countries, no. 1453, https://doi.org/https://doi.org/https://doi.org/10.1787/bc7569c6-en. https://www.oecd-ilibrary.org/content/paper/bc7569c6-en.

Černiauskas, Nerijus, and Andrius Čiginas. 2020. Measurement and decomposition of lithuania's income inequality. *Baltic Journal of Economics* 20 (2): 139–169.

Černiauskas, Nerijus, and Alain Jousten. 2020. *Statutory, Effective and Optimal Net Tax Schedules in Lithuania.* Bank of Lithuania Working Paper Series 72. Bank of Lithuania, February. https://ideas.repec.org/p/lie/wpaper/72.html.

Černiauskas, Nerijus, Denisa M. Sologon, Cathal O'Donoghue, and Linas Tarasonis. 2020. *Changes in Income Inequality in Lithuania: the Role of Policy, Labour Market Structure, Returns and Demographics.* Bank of Lithuania Working Paper Series 71. Bank of Lithuania, January. https://ideas.repec.org/p/lie/wpaper/71.html.

Chetty, Raj, Michael Stepner, Sarah Abraham, Shelby Lin, Benjamin Scuderi, Nicholas Turner, Augustin Bergeron, and David Cutler. 2016. The association between income and life expectancy in the united states, 2001-2014. *Jama* 315 (16): 1750–1766.

Chiappori, Pierre-André, and Costas Meghir. 2015. Intrahousehold Inequality. In *Handbook of income distribution,* 2:1369–1418. Elsevier.

Cingano, Federico. 2014. Trends in Income Inequality and its Impact on Economic Growth, no. 163, https://EconPapers.repec.org/RePEc:oec:elsaab:163-en.

Cowell, Frank A. 2000. Measurement of Inequality. *Handbook of Income Distribution* 1:87–166.

———. 2011. *Measuring Inequality.* Oxford University Press.

Daly, Mary C., and Robert G. Valletta. 2006. Inequality and poverty in United States: the effects of rising dispersion of men's earnings and changing family behaviour. *Economica* 73 (289): 75–98.

Dardanoni, Valentino, and Peter J. Lambert. 2002. Progressivity comparisons. *Journal of Public Economics* 86 (1): 99–122.

Das, Tarun, and Ashok Parikh. 1982. Decomposition of Inequality Measures and a Comparative Analysis. *Empirical Economics* 7 (1): 23–48.

Deutsch, Joseph, María Noel Pi Alperin, and Jacques Silber. 2018. Using the shapley decomposition to disentangle the impact of circumstances and efforts on health inequality. *Social Indicators Research* 138 (2): 523–543.

DiNardo, John, Nicole M Fortin, and Thomas Lemieux. 1996. Labor market institutions and the distribution of wages, 1973-1992: a semiparametric approach. *Econometrica* 64 (5): 1001–1044.

Domar, Evsey D, and Richard A Musgrave. 1944. Proportional Income Taxation and Risk-taking. *The Quarterly Journal of Economics* 58 (3): 388–422.

Elbers, Chris, Peter Lanjouw, Johan A Mistiaen, and Berk Özler. 2008. Reinterpreting Between-group Inequality. *The Journal of Economic Inequality* 6 (3): 231–245.

Eurostat. 2018a. *EU statistics on income and living conditions (EU-SILC) methodology - distribution of income.* https://ec.europa.eu/eurostat/statistics-explained/index.php?title=EU_statistics_on_income_and_living_conditions_(EU-SILC)_methodology_-_distribution_of_income.

———. 2018b. *EU Statistics on Income and Living Conditions (EU-SILC) Methodology – Concepts and Contents.* http://ec.europa.eu/eurostat/statistics-explained/index.php?title=EU_statistics_on_income_and_living_conditions_(EU-SILC)_methodology_%E2%80%93_concepts_and_contents#Adjusted_cross_sectional_weight_.28RB050a.29.

———. 2018c. *European Union Statistics on Income and Living Conditions (EU-SILC).* Online. https://ec.europa.eu/eurostat/web/microdata/european-union-statistics-on-income-and-living-conditions.

———. 2018d. Gini Coefficient of Equivalised Disposable Income - EU-SILC Survey, https://ec.europa.eu/eurostat/data/database.

———. 2021. Gender pay gap in unadjusted form. Accessed January 6, 2021. https://ec.europa.eu/eurostat/databrowser/view/sdg_05_20/default/table?lang=en.

Firpo, Sergio, Nicole M Fortin, and Thomas Lemieux. 2009. Unconditional Quantile Regressions. *Econometrica* 77 (3): 953–973.

Frölich, Markus, David Kaplan, Carmen Pagés, Jamele Rigolini, and David Robalino. 2014. *Social insurance, informality, and labour markets: how to protect workers while creating good jobs.* 7879. Oxford University Press.

Garner, Thesia I, and Katherine Terrell. 1998. A Gini Decomposition Analysis of Inequality in the Czech and Slovak Republics During the Transition 1. *Economics of Transition* 6 (1): 23–46.

Gathmann, Christina, and Björn Sass. 2018. Taxing childcare: effects on childcare choices, family labor supply, and children. *Journal of labor Economics* 36 (3): 665–709.

Gelbach, Jonah B. 2002. Public schooling for young children and maternal labor supply. *American Economic Review* 92 (1): 307–322.

Gentry, William M, and R Glenn Hubbard. 2000. Tax Policy and Entry into Entrepreneurship. *American Economic Review* 90 (2): 283–287.

Goedemé, Tim. 2013. How Much Confidence Can We Have in EU-SILC? Complex Sample Designs and the Standard Error of the Europe 2020 Poverty Indicators. *Social Indicators Research* 110 (1): 89–110. https://doi.org/10.1007/s11205-011-9918-2.

Grigoli, Francesco, and Adrian Robles. 2017. *Inequality Overhang.* International Monetary Fund.

Guner, Nezih, Remzi Kaygusuz, and Gustavo Ventura. 2014. Income taxation of us households: facts and parametric estimates. *Review of Economic Dynamics* 17 (4): 559–581.

Hasselman, Berend. 2018. Nleqslv: Solve Systems of Nonlinear Equations. R package version 3.3.2, https://CRAN.R-project.org/package=nleqslv.

Havnes, Tarjei, and Magne Mogstad. 2011. Money for nothing? Universal child care and maternal employment. *Journal of Public Economics* 95 (11-12): 1455–1465.

Heathcote, Jonathan, Kjetil Storesletten, and Giovanni L Violante. 2017. Optimal Tax Progressivity: an Analytical Framework. *The Quarterly Journal of Economics* 132 (4): 1693–1754.

Herault, Nicolas, and Francisco Azpitarte. 2016. Understanding changes in the distribution and redistribution of income: a unifying decomposition framework. *Review of Income and Wealth* 62 (2).

HFCN. 2019. Household Finance and Consumption Network (HFCN). ECB. https://www.ecb.europa.eu/pub/economic-research/research-networks/html/researcher_hfcn.en.html.

Hlasny, Vladimir, and Paolo Verme. 2018. Top incomes and inequality measurement: a comparative analysis of correction methods using the eu silc data. *Econometrics* 6 (2): 30.

Huang, Jason, and Juan Rios. 2016. Optimal Tax Mix with Income Tax Noncompliance. *Journal of Public Economics* 144:52–63.

IMF. 2016. *Republic of Lithuania, Selected Issues.*

Immervoll, Herwig, Henrik Jacobsen Kleven, Claus Thustrup Kreiner, and Nicolaj Verdelin. 2011. Optimal Tax and Transfer Programs for Couples with Extensive Labor Supply Responses. *Journal of Public Economics* 95 (11-12): 1485–1500.

Immervoll, Herwig, and Linda Richardson. 2011. Redistribution policy and inequality reduction in oecd countries: what has changed in two decades? *OECD Social, Employment, and Migration Working Papers,* no. 122.

Jäntti, Markus, Jukka Pirttilä, and Håkan Selin. 2015. Estimating labour supply elasticities based on cross-country micro data: a bridge between micro and macro estimates? *Journal of Public Economics* 127:87–99.

Jenkins, Stephen. 2017. *The Measurement of Income Inequality,* 17–52. Routledge.

Johns, Andrew, and Joel Slemrod. 2010. The distribution of income tax noncompliance. *National Tax Journal* 63 (3): 397.

Jongen, Egbert L. W., and Maaike Stoel. 2019. The elasticity of taxable labour income in the netherlands. *De Economist* 167, no. 4 (December): 359–386. https://doi.org/10.1007/s10645-019-09349-7. https://doi.org/10.1007/s10645-019-09349-7.

Joumard, Isabelle, Mauro Pisu, and Debbie Bloch. 2013. Tackling income inequality. *OECD Journal: Economic Studies* 2012 (1): 37–70.

Juhn, Chinhui, Kevin M. Murphy, and Brooks Pierce. 1993. Wage inequality and the rise in returns to skill. *Journal of Political Economy* 101 (3): 410–442. ISSN: 00223808. http://www.jstor.org/stable/2138770.

Kakwani, Nanak C. 1977. Measurement of Tax Progressivity: an International Comparison. *The Economic Journal* 87 (345): 71–80.

Karu, Marre, and Rūta Bražienė. 2014. Comparing leave policies across the baltic states. 11th International LP & R seminar in Tallinn. https://www.leavenetwork.org/annual-seminars/.

Klemm, Alexander D, Li Liu, Victor Mylonas, and Philippe Wingender. 2018. *Are Elasticities of Taxable Income Rising?* IMF Working Paper 18/132. International Monetary Fund.

Kleven, Henrik, Camille Landais, Johanna Posch, Andreas Steinhauer, and Josef Zweimüller. 2019. Child penalties across countries: evidence and explanations. In *Aea papers and proceedings,* 109:122–26.

Kleven, Henrik, Camille Landais, and Jakob Egholt Søgaard. 2019. Children and gender inequality: Evidence from Denmark. *American Economic Journal: Applied Economics* 11 (4): 181–209.

Kukk, Merike, Alari Paulus, and Karsten Staehr. 2019. Cheating in europe: underreporting of self-employment income in comparative perspective. *International Tax and Public Finance* (August).

Lazutka, Romas. 2017. Pervedimų, Valstybės Teikiamų Prekių ir Paslaugų Vaidmuo. In *Pajamų Nelygybė Lietuvoje, October 5th.* Lietuvos Nacionalinė Martyno Mažvydo Biblioteka: Bank of Lithuania. https://www.lb.lt/uploads/documents/files/Lazutka.pdf.

———. 2020. *Lietuvos socialinės politikos reakcija į nedarbo augimą dėl covid-19.* Technical report.

Lefebvre, Pierre, and Philip Merrigan. 2008. Child-care policy and the labor supply of mothers with young children: A natural experiment from Canada. *Journal of Labor Economics* 26 (3): 519–548.

Lemieux, Thomas. 2006. Increasing residual wage inequality: composition effects, noisy data, or rising demand for skill? *American Economic Review* 96 (3): 461–498.

Lerman, Robert I, and Shlomo Yitzhaki. 1985a. Income Inequality Effects by Income Source: a New Approach and Applications to the United States. *The Review of Economics and Statistics,* 151–156.

Lerman, Robert I, and Shlomo Yitzhaki. 1985b. Income inequality effects by income source: a new approach and applications to the united states. *The Review of Economics and Statistics* 67 (1): 151–156.

Levy, Horacio, Christine Lietz, and Holly Sutherland. 2007. Swapping policies: alternative tax-benefit strategies to support children in Austria, Spain and the UK. *Journal of Social Policy* 36 (4): 625–647.

Li, Jinjing, Hai Anh La, and Denisa M Sologon. 2020. Policy, demography, and market income volatility: what shaped income distribution and inequality in australia between 2002 and 2016? *Review of Income and Wealth.*

Lis, Maciej. 2018. *Lithuania'S Pension System in the International Persective.* Technical report. September.

LR Seimas. 2020. Nutarimas dėl aštuonioliktosios Lietuvos Respublikos vyriausybės programos [Resolution on the Approval of the Program Implementation Plan of the 18th Government of the Republic of Lithuania], https://e-seimas.lrs.lt/portal/legalAct/lt/TAD/973c87403bc311eb8c97e01ffe050e1c?jfwid=i3h7ws9m4.

LR Vyriausybė. 2017. Nutarimas dėl Lietuvos Respublikos Vyriausybės Programos Įgyvendinimo Plano Patvirtinimo [Resolution on the Approval of the Program Implementation Plan of the Government of the Republic of Lithuania], https://www.e-tar.lt/portal/lt/legalAct/2389544007bf11e79ba1ee3112ade9bc.

Lumley, Thomas. 2004. Analysis of Complex Survey Samples. R package verson 2.2, *Journal of Statistical Software* 9 (1): 1–19.

———. 2018. Survey: Analysis of Complex Survey Samples. R package version 3.34.

Lundberg, Shelly, Robert A Pollak, and Jenna Stearns. 2016. Family inequality: diverging patterns in marriage, cohabitation, and childbearing. *Journal of Economic Perspectives* 30 (2): 79–102.

Magda, Iga, Jan Gromadzki, and Simone Moriconi. 2020. Firms and wage inequality in central and eastern europe. *Journal of Comparative Economics.*

Matsaganis, M, C Leventi, and M Flevotomou. 2013. *Distributional implications of tax evasion and the crisis in greece.* EUROMOD Working Paper EM17/13. Euromod.

Meriküll, Jaanika, and Pille Mõtsmees. 2017. Do you get what you ask? The gender gap in desired and realised wages. *International Journal of Manpower.*

Meriküll, Jaanika, and Tairi Room. 2019. *Are survey data underestimating the inequality of wealth?* Technical report. Bank of Estonia.

Milanez, Anna, and Barbara Bratta. 2019. *Taxation and the future of work.* 41. https://doi.org/https://doi.org/https://doi.org/10.1787/20f7164a-en. https://www.oecd-ilibrary.org/content/paper/20f7164a-en.

Ministry of Social Secuirty and Labour. 2020. *2021–2030 m. plėtros programos valdytojo lietuvos respublikos socialinės apsaugos ir darbo ministerijos pajamų nelygybės mažinimo plėtros programos pagrindimas [justification for the development programme for reducing income inequality by the ministry of social security and labour as the manager of the 2021–2030 development programme 2020].* Technical report.

Mirlees, James. 1971. An Exploration into the Theory of Optimal Taxation. *Review of Economic Studies* 38.

Motiejunaite-Schulmeister, Akvile, Marie-Pascale Balcon, and Isabelle de Coster. 2019. Key Data on Early Childhood Education and Care in Europe, 2019 Edition. Eurydice Report. *Education, Audiovisual and Culture Executive Agency, European Commission.*

Navickė, Jekaterina. 2020. Factors behind the changes in income distribution in the baltics: income, policy, demography. *Journal of Baltic Studies* 51 (2): 137–157.

Navickė, Jekaterina, and Aušra Čižauskaitė. 2018. *EUROMOD Country Report – LITHUANIA (LT) 2015-2018.* https://www.euromod.ac.uk/sites/default/files/country-reports/year9/Y9_CR_LT_Final.pdf.

Navickė, Jekaterina, and Romas Lazutka. 2016. Functional and Personal Income Distribution in the Baltics: Comparison of National and Households Accounts. In *Business and Management 2016: the 9th International Scientific Conference, May 12-13, 2016, Vilnius, Lithuania.* Technika.

———. 2017. Distributional Implications of the Economic Development in the Baltics: Reconciling Micro and Macro Perspectives (May): 1–20.

Neisser, Carina. 2017. The elasticity of taxable income: a meta-regression analysis.

Nolan, Brian, Wiemer Salverda, Daniele Checchi, Ive Marx, Abigail McKnight, István György Tóth, and Herman G. van de Werfhorst, eds. 2014. *Changing inequalities and societal impacts in rich countries: Thirty countries' experiences.* Oxford University Press.

O'Donoghue, Cathal, Denisa M. Sologon, Iryna Kyzyma, and John McHale. 2020. Modelling the distributional impact of the covid-19 crisis*. *Fiscal Studies* 41 (2): 321–336. https://doi.org/https://doi.org/10.1111/1475-5890.12231. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/1475-5890.12231. https://onlinelibrary.wiley.com/doi/abs/10.1111/1475-5890.12231.

OECD. 2011a. *Divided We Stand.* 388. https://doi.org/https://doi.org/https://doi.org/10.1787/9789264119536-en. https://www.oecd-ilibrary.org/content/publication/9789264119536-en.

———. 2011b. *Divided we stand: why inequality keeps rising.* Paris: OECD.

———. 2012. *What are equivalence scales?* http://www.oecd.org/eco/growth/OECD-Note-EquivalenceScales.pdf.

———. 2015a. *All on board.* 200. https://doi.org/https://doi.org/https://doi.org/10.1787/9789264218512-en. https://www.oecd-ilibrary.org/content/publication/9789264218512-en.

———. 2015b. *In It Together: Why Less Inequality Benefits All.* 336. https://doi.org/https://doi.org/https://doi.org/10.1787/9789264235120-en.

———. 2016. *Early childhood education and care in Latvia.* https://doi.org//https://doi.org/10.1787/9789264250628-5-en. https://www.oecd-ilibrary.org/content/component/9789264250628-5-en.

———. 2017a. *Education in lithuania.* 196. https://doi.org/https://doi.org/https://doi.org/10.1787/9789264281486-en. https://www.oecd-ilibrary.org/content/publication/9789264281486-en.

———. 2017b. *Šalių švietimo politikos apžvalgos: Švietimas Lietuvoje.* https://www.upc.smm.lt/naujienos/europa/apzvalga/Svietimas-Lietuvoje-EBPO-.pdf.

———. 2019. *Taxing Wages 2019.* OECD Publishing. Paris. https://doi.org/https://doi.org/10.1787/tax_wages-2019-en.

OECD. 2020. *Early Learning and Child Well-being in Estonia.* https://doi.org/https://doi.org/10.1787/15009dbe-en. https://www.oecd-ilibrary.org/content/publication/15009dbe-en.

Ostry, Jonathan D., Andrew Berg, and Charalambos G Tsangarides. 2014. *Redistribution, Inequality, and Growth.* International Monetary Fund.

Pahl, Jan. 1995. His Money, Her Money: Recent Research on Financial Organisation in Marriage. *Journal of economic psychology* 16 (3): 361–376.

Paulus, Alari. 2015. *Tax Evasion and Measurement Error: An Econometric Analysis of Survey Data Linked with Tax Records.* ISER Working Paper Series 2015-10. Essex: University of Essex.

Paulus, Alari, and Iva Valentinova Tasseva. 2020. Europe through the crisis: discretionary policy changes and automatic stabilizers. *Oxford Bulletin of Economics and Statistics* 82 (4): 864–888.

Peichl, Andreas, and Philippe van Kerm. 2007. *PROGRES: Stata module to measure distributive effects of an income tax.* Statistical Software Components, Boston College Department of Economics, September. https://ideas.repec.org/c/boc/bocode/s456867.html.

Piketty, Thomas. 2015. About Capital in the Twenty-First Century. *American Economic Review* 105 (5): 48–53.

Pissarides, Christopher A, and Guglielmo Weber. 1989. An expenditure-based estimate of britain's black economy. *Journal of public economics* 39 (1): 17–32.

Plat, Didier. 2012. IC2: Inequality and Concentration Indices and Curves. R package version 1.0-1, https://CRAN.R-project.org/package=IC2.

Rakauskienė, Ona Gražina, Eglė Krinickienė, and Vaida Servetkienė. 2020. *Moterų ir vyrų pajamų atotrūkis Lietuvoje (Income gap by gender in Lithuania).*

Rao, JNK, CFJ Wu, and K Yue. 1992. Some Recent Work on Resampling Methods for Complex Surveys. *Survey Methodology* 18 (2): 209–217.

Redmond, Paul, and Seamus McGuinness. 2019. The gender wage gap in Europe: Job preferences, gender convergence and distributional effects. *Oxford Bulletin of Economics and Statistics* 81 (3): 564–587.

Reynolds, Morgan, and Eugene Smolensky. 1977. Public expenditures, taxes and the distribution of income: the United States, 1950, 1961, 1970. *Academic Press, New York.*

Rothe, Christoph. 2012. Partial Distributional Policy Effects. *Econometrica* 80 (5): 2269–2301.

Saez, Emmanuel. 2001. Using Elasticities to Derive Optimal Income Tax Rates. *The review of economic studies* 68 (1): 205–229.

———. 2002. Optimal Income Transfer Programs: Intensive Versus Extensive Labor Supply Responses. *The Quarterly Journal of Economics* 117 (3): 1039–1073.

Saez, Emmanuel, Joel Slemrod, and Seth H Giertz. 2012. The elasticity of taxable income with respect to marginal tax rates: a critical review. *Journal of economic literature* 50 (1): 3–50.

Sastre, Mercedes, and Alain Trannoy. 2002. Shapley Inequality Decomposition by Factor Components: Some Methodological Issues. *Journal of Economics* 77 (1): 51–89.

Scheuer, Florian, and Joel Slemrod. 2019. *Taxation and the superrich.* Technical report. National Bureau of Economic Research.

Schneider, Friedrich. 2013. *Size and Development of the Shadow Economy of 31 European and 5 Other OECD Countries from 2003 to 2013: a Further Decline.* http://politeia.org.ro/wp-content/uploads/2013/05/ShadEcEurope31_Jan2013.pdf.

Senaj, Matúš, Zuzana Siebertová, Norbert Švarda, and Jana Valachyová. 2014. Estimates of labour supply elasticities in slovakia1.

Shorrocks, AF. 1980. The class of additively decomposable inequality measures. *Econometrica* 48 (3): 613–625.

———. 1982. Inequality decomposition by factor components. *Econometrica* 50 (1): 193–211.

Shorrocks, Anthony F. 1982. Inequality Decomposition by Factor Components. *Econometrica: Journal of the Econometric Society,* 193–211.

———. 1984. Inequality decomposition by population subgroups. *Econometrica* 52 (6): 1369–85.

Shorrocks, Anthony F. 2013. Decomposition Procedures for Distributional Analysis: a Unified Framework Based on the Shapley Value. *The Journal of Economic Inequality* 11 (1): 99–126.

Skučienė, Daiva. 2008. Pajamų nelygybė lietuvoje. *FILOSOFIJA. SOCIOLOGIJA,* no. 4, 22–33.

Slemrod, Joel. 2016. *Tax compliance and enforcement: new research and its policy implications.* Ross School of Business Working Paper 1302. Michigan: University of Michigan.

Smeeding, Timothy, Robert Erikson, and Markus Jäntti. 2011. *Persistence, privilege, and parenting: the comparative study of intergenerational mobility.* Russell Sage Foundation.

Social Situation Monitor. 2017. Research findings - Social Situation Monitor - Measurement and methodology, https://ec.europa.eu/social/main.jsp?catId=1050&intPageId=1888&langId=en.

SODRA. 2020. Contribution rates for employees. SODRA. https://www.sodra.lt/en/benefits/contribution-rates/contribution-rates-for-employees.

Sologon, Denisa M., Vanda Almeida, and Philipe Van Kerm. 2019. Accounting for the distributional effects of the 2007-2008 crisis and the economic adjustment program in Portugal. *LISER Working Papers no. 2019-05, 52 p.*

Sologon, Denisa M., Philippe Van Kerm, Jinjing Li, and Cathal O'Donoghue. 2020. Accounting for differences in income inequality across countries: tax-benefit policy, labour market structure, returns and demographics. *Journal of Economic Inequality,* https://doi.org/10.1007/s10888-020-09454-7.

Staehr, Karsten. 2008. Microfoundations of Economic Success: Lessons from Estonia. Chap. Estimates of Employment and Welfare Effects of Personal Labour Income Taxation in a Flat-Tax Country: The Case of Estonia, edited by David Mayes, 242–290. Cheltenham: Edward Elgar Publishing.

Sutherland, H., and F. Figari. 2013. EUROMOD: the European Union tax-benefit microsimulation model. *International Journal of Microsimulation* 6 (1).

Tasseva, Iva Valentinova. 2020. The changing education distribution and income inequality in great britain. *Review of Income and Wealth,* https://doi.org/https://doi.org/10.1111/roiw.12486.

Törmälehto, VM. 2017. *High income and affluence: evidence from the european union statistics on income and living conditions (eu-silc).* Statistical Working Papers, Eurostat, Luxembourg.

Vahter, Priit, and Jaan Masso. 2019. The contribution of multinationals to wage inequality: foreign ownership and the gender pay gap. *Review of World Economics* 155 (1): 105–148.

Van Kerm, Philippe. 2013. Generalized measures of wage differentials. *Empirical Economics* 45 (1): 465–482.

Van Kerm, Philippe, Seunghee Yu, and Chung Choe. 2016. Decomposing quantile wage gaps: a conditional likelihood approach. *Journal of the Royal Statistical Society: Series C (Applied Statistics)* 65 (4): 507–527.

Vandenbroeck, Michel, Karolien Lenaerts, and M Beblavý. 2018. Benefits of early childhood education and care and the conditions for obtaining them. *European Expert Network on Economics of Education* 32:1–86.

Vogler, Carolyn, and Jan Pahl. 1994. Money, Power and Inequality within Marriage. *The Sociological Review* 42 (2): 263–288.

Volodzkienė, Lina. 2020. Ekonominės nelygybės poveikio socialinei ekonominei pažangai vertinimas es šalyse. PhD diss., Mykolo Romerio universitetas.

Yitzhaki, Shlomo. 1983. On an Extension of the Gini Inequality Index. *International Economic Review,* 617–628.

Yitzhaki, Shlomo, and Robert I Lerman. 1991. Income Stratification and Income Inequality. *Review of Income and Wealth* 37 (3): 313–329.

Zardo Trindade, Lorena, and Tim Goedemé. 2016. *Notes on Updating the EU-SILC UDB Sample Design variables, 2012-2014.* Working Papers 1602. Herman Deleeck Centre for Social Policy, University of Antwerp. https://EconPapers.repec.org/RePEc:hdl:wpaper:1602.

Zawisza, Tomasz. 2019. Optimal taxation of employment and self-employment: evidence from poland and implications.

Žukauskas, Vytautas. 2016. *Lietuvos Šešėlinė Ekonomika.* Lietuvos Laisvos Rinkos Institutas.

APPENDIX

$\mathcal{U} = \{1, \ldots, N\}$ is the set representing elements of the finite survey population, and $y_1, \ldots, y_N$ are values of the variable of interest (income) in $\mathcal{U}$. The subset $s = \{i_1, \ldots, i_n\}$ of $\mathcal{U}$ is the sample, while w_i and $i \in s$ are the corresponding survey weights. We use the estimator

$$\widehat{G} = \frac{2}{\hat{\mu}}\Big(\frac{1}{\widehat{N}}\sum_{i\in s} w_i y_i \widehat{F}(y_i) - \frac{1}{\widehat{N}^2}\sum_{i\in s} w_i y_i \sum_{i\in s} w_i \widehat{F}(y_i)\Big) \tag{5}$$

of the Gini coefficient (3.2), constructed in line with Berger (2008), where $\widehat{F}(y_i)$ are values of the estimated distribution function

$$\widehat{F}(y) = \frac{1}{\widehat{N}}\sum_{j\in s} w_j \mathbb{I}\{y_j \le y\} \quad \text{with} \quad \widehat{N} = \sum_{j\in s} w_j, \tag{6}$$

and

$$\hat{\mu} = \frac{1}{\widehat{N}}\sum_{j\in s} w_j y_j.$$

Here $\mathbb{I}\{\cdot\}$ stands for the indicator function. Estimators of the subgroup and factor decompositions are constructed using similar plug-in principles.

.1. Subgroup decompositions

We give the decomposition of (5) by groups as in Yitzhaki and Lerman (1991). Let $s = s_1 \cup \cdots \cup s_L$ be a division of the sample by non-overlapping groups. Denote

$$\widehat{N}_l = \sum_{j\in s_l} w_j, \quad \widehat{P}_l = \frac{\widehat{N}_l}{\widehat{N}}, \quad \hat{m} = \frac{1}{\widehat{N}}\sum_{j\in s} w_j y_j,$$
$$\hat{m}^{(l)} = \frac{1}{\widehat{N}_l}\sum_{j\in s_l} w_j y_j, \quad \widehat{F}^{(l)} = \frac{1}{\widehat{N}_l}\sum_{i\in s_l} w_i \widehat{F}(y_i), \tag{7}$$

where $\widehat{N}_l$ is the estimated population size in the subgroup l, the quantity $\widehat{P}_l$ is the estimated population share, $\hat{m}$ is the estimated mean of the survey variable in $\mathcal{U}$, $\hat{m}^{(l)}$ is the estimated mean in the subgroup, and $\widehat{F}^{(l)}$ is the estimate of the average of global ranks in the subgroup l. Consider the values $\widehat{F}_l(y_i)$ and $\widehat{F}_{L\setminus l}(y_i)$, $i \in s$, of the estimated distribution functions

$$\widehat{F}_l(y) = \frac{1}{\widehat{N}_l}\sum_{j\in s_l} w_j \mathbb{I}\{y_j \le y\} \quad \text{and} \quad \widehat{F}_{L\setminus l}(y) = \frac{1}{\widehat{N} - \widehat{N}_l}\sum_{j\in s\setminus s_l} w_j \mathbb{I}\{y_j \le y\}$$

in the subgroup l and outside this subgroup, respectively. Introduce the notations

$$\widehat{\mathrm{cov}}(m^{(l)}, F^{(l)}) = \sum_{l=1}^{L} \widehat{P}_l \Big(\hat{m}^{(l)} - \sum_{l=1}^{L} \widehat{P}_l \hat{m}^{(l)}\Big)\Big(\widehat{F}^{(l)} - \sum_{l=1}^{L} \widehat{P}_l \widehat{F}^{(l)}\Big)$$

and

$$\widehat{\mathrm{cov}}_l(y, F_l(y)) = \frac{1}{\widehat{N}_l} \sum_{i \in s_l} w_i (y_i - \hat{m}^{(l)})\Big(\widehat{F}_l(y_i) - \frac{1}{\widehat{N}_l} \sum_{i \in s_l} w_i \widehat{F}_l(y_i)\Big),$$

and

$$\widehat{\mathrm{cov}}_l(y, F_l(y) - F_{L\setminus l}(y)) = \widehat{\mathrm{cov}}_l(y, F_l(y)) - \frac{1}{\widehat{N}_l} \sum_{i \in s_l} w_i (y_i - \hat{m}^{(l)})\Big(\widehat{F}_{L\setminus l}(y_i) - \frac{1}{\widehat{N}_l} \sum_{i \in s_l} w_i \widehat{F}_{L\setminus l}(y_i)\Big).$$

Then the estimated decomposition by groups is written as

$$\widehat{G} = \underbrace{\sum_{l=1}^{L} \widehat{S}_l \widehat{G}_l}_{\text{Within}} + \underbrace{\sum_{l=1}^{L} \widehat{S}_l \widehat{G}_l \widehat{Q}_l (\widehat{P}_l - 1)}_{\text{Stratification}} + \underbrace{\frac{2\,\widehat{\mathrm{cov}}(m^{(l)}, F^{(l)})}{\hat{m}}}_{\text{Between}}, \tag{8}$$

where

$$\widehat{S}_l = \widehat{P}_l \frac{\hat{m}^{(l)}}{\hat{m}}, \quad \widehat{G}_l = \frac{2\,\widehat{\mathrm{cov}}_l(y, F_l(y))}{\hat{m}^{(l)}}, \quad \widehat{Q}_l = \frac{\widehat{\mathrm{cov}}_l(y, F_l(y) - F_{L\setminus l}(y))}{\widehat{\mathrm{cov}}_l(y, F_l(y))}.$$

Here the component $\widehat{S}_l$ represents the share of the survey variable, $\widehat{G}_l$ is the estimated within-group Gini coefficient, and the part $\widehat{Q}_l$ is the estimated stratification term.

.2. Factor decompositions

We write down an estimate of the factor decomposition by Lerman and Yitzhaki (1985a). Write $y_i = \sum_{k=1}^{K} y_i^{(k)}$, where k is a factor of the survey variable. Consider the values $\widehat{F}(y_i)$ and $\widehat{F}(y_i^{(k)})$, $i \in s$, of distribution function (6) and denote the expressions

$$\widehat{\mathrm{cov}}(y^{(k)}, F(y)) = \frac{1}{\widehat{N}} \sum_{i \in s} w_i y_i^{(k)} \widehat{F}(y_i) - \frac{1}{\widehat{N}^2} \sum_{i \in s} w_i y_i^{(k)} \sum_{i \in s} w_i \widehat{F}(y_i)$$

and

$$\widehat{\text{cov}}(y^{(k)}, F(y^{(k)})) = \frac{1}{\widehat{N}} \sum_{i \in s} w_i y_i^{(k)} \widehat{F}(y_i^{(k)}) - \frac{1}{\widehat{N}^2} \sum_{i \in s} w_i y_i^{(k)} \sum_{i \in s} w_i \widehat{F}(y_i^{(k)}).$$

Also, introduce the weighted means

$$\hat{\mu}^{(k)} = \frac{1}{\widehat{N}} \sum_{j \in s} w_j y_j^{(k)}.$$

Then the estimated decomposition by factors is

$$\widehat{G} = \sum_{k=1}^{K} \widehat{T}_k = \sum_{k=1}^{K} \widehat{R}_k \widehat{G}_k \widehat{S}_k, \tag{9}$$

where

$$\widehat{R}_k = \frac{\widehat{\text{cov}}(y^{(k)}, F(y))}{\widehat{\text{cov}}(y^{(k)}, F(y^{(k)}))}, \quad \widehat{G}_k = \frac{2\,\widehat{\text{cov}}(y^{(k)}, F(y^{(k)}))}{\hat{\mu}^{(k)}}, \quad \widehat{S}_k = \frac{\hat{\mu}^{(k)}}{\hat{\mu}}.$$

Here $\widehat{R}_k$ is the estimate of the so-called Gini correlation between the survey variable and its kth component, $\widehat{G}_k$ represents the Gini index of factor k, and $\widehat{S}_k$ is the share of factor. For a small change in the kth factor, the expression of marginal effects is

$$\frac{\partial \widehat{G}}{\partial e_k} = \widehat{S}_k(\widehat{R}_k \widehat{G}_k - \widehat{G}_k), \tag{10}$$

see Lerman and Yitzhaki (1985a).

Table A1: Number of observations per decile from EUROMOD output

percentile	gross employment income	total observations	employed	self-employed
0-7	0	690	0	0
7-20	1,413	784	174	20-49
20-30	3,588	649	385	59
30-40	5,102	649	462	20-49
40-50	6,609	641	543	20-49
50-60	8,219	635	559	20-49
60-70	10,080	626	557	20-49
70-80	12,357	615	545	52
80-90	15,690	635	578	20-49
90-100	27,651	572	461	107

Data is sorted according to equivalised gross employment income (includes social contributions which are evaluated by Euromod). All figures are taken from Euromod and are weighted to include only those households with atleast 1 member who is 18-62 year old and is not a student. Deciles are based on weighted observations, which results in different number of observations per quantile. 20-49 indicates that there are between 20 and 49 (inclusive) number of observations, although the number is not publishable due to confidentially reasons. The first 7 percentiles do not have any gross employment income.

Table A2: Number of observations per decile from EU-SILC

percentile	gross employment income	total observations	employed	self-employed
0-7	0	680	0	0
7-20	1,395	787	300	20-49
20-30	3,525	629	501	60
30-40	4,972	636	557	20-49
40-50	6,470	653	602	20-49
50-60	8,047	631	576	20-49
60-70	9,888	623	569	20-49
70-80	12,141	619	557	55
80-90	15,425	625	569	20-49
90-100	27,143	576	467	107

Data is sorted according to equivalised gross employment income (includes social contributions). All figures are taken from Euromod and are weighted to include only those households with atleast 1 member who is 18-62 year old and is not a student. Deciles are based on weighted observations, which results in different number of observations per quantile. 20-49 indicates that there are between 20 and 49 (inclusive) number of observations, although the number is not publishable due to confidentially reasons. The first 7 percentiles do not have any gross employment income.

Table A3: Household statutory average net tax rates in Lithuania, net taxes as a share of gross employment income. Sample restricted to households with 1 household member aged 18-62, but can include older and younger household members as well.

percentile	net taxes			taxes		public transfers	
	all	employees	self-employed	employees	self-employed	employees	self-employed
0-24							
24-50	-0.946 [996]	-0.006 [281]	-0.012 [60]	0.354 [281]	0.327 [60]	0.361 [281]	0.339 [60]
50-75	0.116 [720]	0.170 [599]	0.003 [53]	0.398 [599]	0.303 [53]	0.229 [599]	0.300 [53]
75-100	0.293 [653]	0.312 [577]	0.207 [68]	0.419 [577]	0.295 [68]	0.106 [577]	0.088 [68]

Percentiles are sorted by gross employment income (includes social contributions). Taxes include income tax and social contributions. public transfers include old-age, disability, unemployment and other benefits. Net taxes are taxes minus public benefits. Gross employment income is taken from EU-SILC, while all other figures are estimated by Euromod, which takes into account various individual and household characteristics (e.g. age, health status). All figures are taken from EU-SILC and are weighted to include only those households with one member aged 18-62 and is not a student, but older and younger household members may be present. The number of observations per quantile is in [].

Table A4: Household average effective net tax rates in Lithuania, net taxes as a share of gross employment income. Sample restricted to households with 1 household member aged 18-62, but can include older and younger household members as well.

percentile	net taxes			taxes		public transfers	
	all	employees	self-employed	employees	self-employed	employees	self-employed
0-24							
24-50	-0.522 [996]	-0.064 [287]	-0.190 [55]	0.316 [287]	0.117 [55]	0.380 [287]	0.307 [55]
50-75	0.142 [721]	0.173 [605]	-0.115 [51]	0.362 [605]	0.105 [51]	0.188 [605]	0.221 [51]
75-100	0.285 [653]	0.323 [576]	0.018 [68]	0.399 [576]	0.082 [68]	0.076 [576]	0.064 [68]

Percentiles are sorted by gross employment income (which includes social contributions). Taxes include income tax and social contributions. Public transfers include old-age, disability, unemployment and other public benefits. Net taxes are taxes minus public benefits. All figures are taken from EU-SILC and are weighted to include only those households with one member aged 18-62 and is not a student. The number of observations per quantile is in [].

.3. Additional tables and figures

Tables A5 and A6 list income generation process components. Table A5 contains the examined income sources and states whether the variable was aggregated or modelled. In case the variable is modelled, it contains the conditioning variables. The corresponding model transformation is also included. The same is done for demographic and labour market variables in Table A6.

Table A5: Definition of income components and summary modelling information

Variable	Definition	Level	Treatment	Factor	Model	Conditioning variables
y_h	total household disposable income	household	aggregate		–	–
y_h^L	gross labour income	household	aggregate		–	–

Continued on next page

Table A5 – continued from previous page

Variable	Definition	Level	Treatment	Factor	Model	Conditioning variables
I_{hi}^{emp}, y_{hi}^{emp}	employee income (wage*hours)	individual	aggregate	Returns (wage rates) and /LM structure (hours)	–	–
I_{hi}^{se}, y_{hi}^{se}	self-employment income (receipt, amount)	individual	modelled	Returns	logit,log-linear	x_{hi}, $firm\_size_{hi}$, occ_{hi}, ind_{hi}, $work\_history_{hi}$, lse_{hi}, $lsepf_{hi}$
y_h^K	capital income (investment, property)	household	aggregate	Returns	–	
I_h^{inv}, y_h^{inv}	investment income (receipt, amount)	individual	modelled	Returns	logit,log-linear	x_{hi}
I_h^{prop}, y_h^{prop}	property income (receipt, amount)	individual	modelled	Returns	logit,log-linear	x_{hi}
y_{hi}^O	other incomes (receipt, amount)	individual	aggregate, modelled	Returns	logit, log-linear	x_{hi}
y_h^B	public transfers	household	aggregate	TB	–	–
$I_{hi}^{sickness}$, $y_{hi}^{sickness}$	sickness (receipt, amount)	individual	modelled	TB	logit, log-linear	x_{hi}
$I_h^{housing}$, $y_h^{housing}$	housing benefits (receipt, amount)	household	modelled	TB	logit, log-linear	x_h

Continued on next page

Table A5 – continued from previous page

Variable	Definition	Level	Treatment	Factor	Model	Conditioning variables
I_h^{sa}, y_h^{sa}	social assistance (receipt, amount)	household	modelled	TB	logit, log-linear	x_h
I_{hi}^{ed}, y_{hi}^{ed}	education benefit (receipt, amount)	individual	modelled	TB	logit, log-linear	x_h
y_h^{mb}	maternity and paternity benefits	individual	modelled	TB	EUROMOD	x_{hi},y_{hi}^L,y_{hi}^K, $work-history_{hi}$
y_h^{pcb}	pregnancy and childcare benefit	individual	modelled	TB	EUROMOD	x_{hi},y_{hi}^L,y_{hi}^K, $work-history_{hi}$
I_{hi}^{unemp}, y_{hi}^{unemp}	unemployment benefits (receipt, amount)	individual	aggregate, modelled	TB	logit, log-linear, EUROMOD	x_{hi}, $un-employed_{hi}$ (for receipt)
I_{hi}^{pens}, y_{hi}^{pens}	state old age benefits (receipt, amount)	individual	aggregate, modelled	TB	logit, log-linear, EUROMOD	x_{hi}, $work-history_{hi}$, $retired_{hi}$ (for receipt)
$I_{hi}^{disability}$ $y_{hi}^{disability}$	disability (receipt and amount)	individual	aggregate, modelled	TB	logit, log-linear, EUROMOD	x_{hi}, $disabled_{hi}$
I_{hi}^{surv}, y_{hi}^{surv}	survivor benefits (receipt, amount)	individual	aggregate, modelled	TB	logit, log-linear, EUROMOD	x_{hi}
t_h	taxes and social security contributions	individual and household	aggregate, modelled	TB	EUROMOD	y_{hi}^L, y_{hi}^K, y_{hi}^O,y_{hi}^B,x_{hi}, $expen-diture_h$,lbl_{hi}
y_h^{ca}	child allowance	family	modelled	TB	EUROMOD	x_h,y_h^L,y_h^K, y_h^B,$work-history_{hi}$

Continued on next page

Table A5 – continued from previous page

Variable	Definition	Level	Treatment	Factor	Model	Conditioning variables
y_h^{bg}	birth grant	Individual	modelled	TB	EUROMOD	x_{hi},y_{hi}^L,y_{hi}^K, $work - history_{hi}$
y_{hi}^{sb}	social benefit	individual	modelled	TB	EUROMOD	x_{hi},y_{hi}^L,y_h^K, y_{hi}^B, $asset_h$,$work-history_{hi}$

Table A6: Demographic and labour market variable

Variable	Definition	Level	Treatment	Factor	Model	Conditioning variables
x_h	household-level demographic characteristics (number of children aged 0–3, 4–11 and 12–15) and individual characteristics of the household head (marital status, gender, age and age squared, university education), assets	household	observed	Demo	–	–

Continued on next page

Table A6 – continued from previous page

Variable	Definition	Level	Treatment	Factor	Model	Conditioning variables
x_{hi}	individual-level characteristics: gender, age and age squared, university education, marital status, number of children in the household (aged 0–3, 4–11 and 12–15), citizenship, age*university, age squared*university, sex, sex*university, age*sex, work-history	individual	observed	Demo	–	–
occ_{hi}	occupation (1-digit ISCO); for working individuals only	individual	modelled	LM Struct	multinomial logit	x_{hi}
ind_{hi}	industry sector (primary, secondary or tertiary); for working individuals only	individual	modelled	LM Struct	multinomial logit	x_{hi}
s_{hi}	number of hours worked	individual	modelled	LM Struct	linear	x_{hi}
w_{hi}	average wage rate; for employees only	individual	modelled	Returns	Singh-Maddala	x_{hi} occ_{hi} ind_{hi}
$retired_{hi}$	retired	individual	modelled	LM Struct	logit	x_{hi}
$ue - nemployed_{hi}$	unemployed	individual	modelled	LM Struct	logit	x_{hi}

Continued on next page

Table A6 – continued from previous page

Variable	Definition	Level	Treatment	Factor	Model	Conditioning variables
$occ - pension_{hi}$	pays voluntary pension	individual	modelled	LM Struct	logit	x_{hi}
lse_{hi}	owner of enterprise with employees (sub-group of self-employed)	individual	modelled	LM Struct	logit	x_{hi}
lbl_{hi}	has business certificate (sub-group of self-employed)	individual	modelled	LM Struct	logit	x_{hi}
$lsepf_{hi}$	engaged in individual activities (sub-group of self-employed)	individual	modelled	LM Struct	logit	x_{hi}

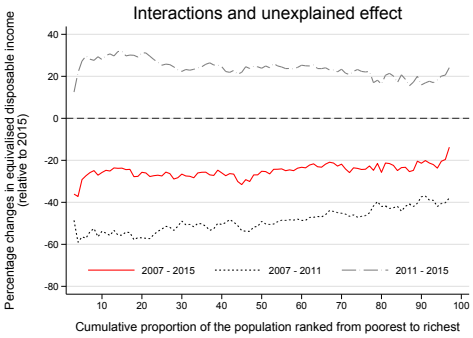
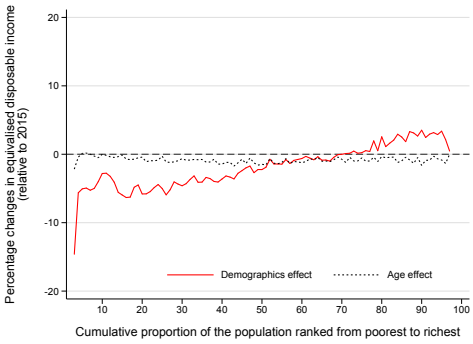
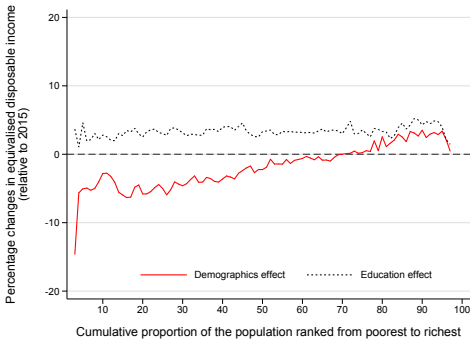


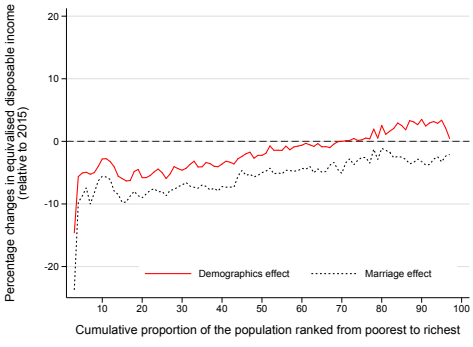
Figure A1: Interactions and unexplained effect



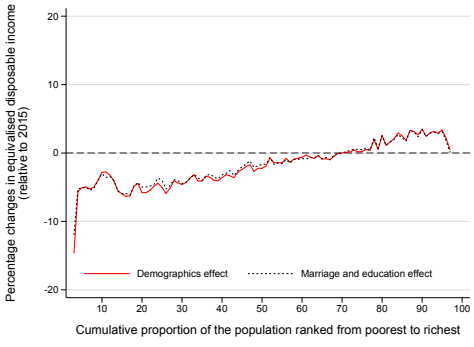
(a) Age effect



(b) Education effect



(c) Marriage effect



(d) Marriage and education effect

Figure A2: Decomposition of the demographic effect

Table A7: Comparison of direct effects and Shapley value effects

	Direct effect	Shapley
Demographics	0.013	0.014
Labour Market Structure	-0.012	-0.006
Prices and Returns	0.030	0.037
Taxes and Benefits	-0.020	-0.021
Unexplained and interactions	0.010	-0.004

Table A8: Employment

	Females						Males					
	2007		2011		2015		2007		2011		2015	
	coef	se	coef	se	coef	se	coef	se	coef	se	coef	se
univ	-0.183	3.133	1.168	2.550	2.454	2.197	5.899*	2.702	4.547	2.858	-1.032	2.136
nch03	0.435	0.427	-0.723	0.500	-0.377	0.302	0.344	0.357	-0.004	0.42	-0.263	0.418
nch411	-0.246	0.229	-0.092	0.330	-0.424	0.240	-0.282	0.168	-0.374	0.258	-0.199	0.173
nch1215	-0.308	0.224	-1.007***	0.256	0.393	0.350	-0.318	0.199	-0.302	0.444	0.142	0.265
marr	0.173	0.225	-0.018	0.233	-0.347	0.188	-0.564	0.36	0.420	0.315	0.022	0.255
age	0.038	0.058	0.068	0.044	0.132**	0.042	0.161**	0.058	0.003	0.053	0.049	0.046
age2	-0.001	0.001	-0.002***	0.000	-0.002***	0.000	-0.002***	0.001	-0.001	0.001	-0.001*	0
ageuniv	0.023	0.154	-0.017	0.109	-0.056	0.092	-0.279*	0.121	-0.157	0.123	0.053	0.092
age2univ	0.000	0.002	0.000	0.001	0.001	0.001	0.003*	0.001	0.002	0.001	-0.000	0.001
constant	2.618*	1.238	3.537***	1.036	1.117	1.030	0.298	1.23	4.104***	1.142	2.180*	0.953
N	2834		3166		2856		2844		2945		2630	
chi2	101.809		245.04		153.779		144.341		81.85		51.297	
p	0		0		0		0		0		0	

Table A9: Average wage rate- females

	2007		2011		2015	
	coef	se	coef	se	coef	se
a						
age: (.,.)	-0.016	0.031	0.057***	0.010	0.049	0.044
age: (30,.)	0.010	0.040	-0.098***	0.019	-0.065	0.067
age: (40,.)	0.009	0.020	0.085***	0.017	0.058	0.046
age: (50,.)	-0.021	0.011	-0.087***	0.013	-0.081**	0.031
univ	-0.145	0.078	-0.280***	0.074	-0.242***	0.067
marr	0.224***	0.056	0.079	0.064	-0.027	0.125
nch03	0.413	0.231	-0.115	0.116	-0.641***	0.162
nch411	0.056	0.092	0.048	0.088	0.132	0.086
nch1215	-0.062	0.069	0.145	0.107	0.142	0.113
firm_size1	-0.253***	0.075	-0.352***	0.068	-0.154	0.130
Constant	1.730	0.913	0.128	0.246	0.187	0.989
b						
age: (.,.)	0.056	0.030	0.015	0.016	-0.004	0.043
age: (30,.)	-0.056	0.041	0.000	0.023	-0.013	0.058
age: (40,.)	-0.018	0.027	-0.062***	0.019	-0.016	0.025
age: (50,.)	-0.006	0.024	0.056**	0.018	0.032	0.019
univ	0.615***	0.076	0.414***	0.085	0.445***	0.088
marr	-0.134*	0.068	-0.171**	0.056	-0.053	0.070
nch03	-0.431***	0.079	-0.236**	0.089	0.227	0.162
nch411	-0.067	0.081	-0.077	0.064	-0.191***	0.055
nch1215	0.026	0.100	-0.141	0.073	-0.053	0.084
firm_size1	0.070	0.120	-0.043	0.066	-0.070	0.102
occ_eur== 1.0000	0.129*	0.054	0.397***	0.084	0.461***	0.048
occ_eur== 2.0000	0.095*	0.047	0.289***	0.051	0.308***	0.063
occ_eur== 3.0000	0.033	0.045	0.157**	0.051	0.210***	0.049
occ_eur== 5.0000	-0.240***	0.047	-0.115*	0.054	-0.105	0.057
occ_eur== 6.0000	-0.157	0.111	-0.109	0.062	0.027	0.045
occ_eur== 7.0000	-0.177*	0.087	0.085	0.074	-0.082	0.052
occ_eur== 8.0000	-0.315***	0.055	-0.243***	0.049	-0.133***	0.040
ind_saps== 1.0000	-0.337***	0.085	-0.085	0.101	-0.077	0.059
ind_saps== 3.0000	-0.075	0.055	0.057	0.041	0.004	0.038
Work history (length of time in months)	0.001	0.000	0.001*	0.000	0.001***	0.000
Constant	-0.796	0.804	0.258	0.452	1.227	1.171
citizen			0.095	0.136	0.015	0.169
q						
age: (.,.)	0.053	0.064	-0.033	0.029	-0.060	0.097
age: (30,.)	-0.054	0.086	0.074	0.049	0.037	0.137
age: (40,.)	-0.040	0.052	-0.139**	0.045	-0.062	0.073
age: (50,.)	0.027	0.036	0.145***	0.036	0.120*	0.055
univ	0.373*	0.160	0.427*	0.182	0.518**	0.161
marr	-0.265*	0.124	-0.370**	0.141	-0.042	0.189
nch03	-0.894***	0.270	-0.248	0.242	0.506**	0.164
nch411	-0.113	0.185	-0.119	0.153	-0.363**	0.139
nch1215	0.029	0.223	-0.202	0.226	-0.227	0.206
firm_size1	0.360	0.251	0.290	0.151	0.078	0.284
Work history (length of time in months)	0.001	0.001	0.001	0.001	0.001**	0.001
Constant	-1.443	1.799	0.657	0.721	1.722	2.267
m						
age: (.,.)	-0.163***	0.023	-0.132***	0.024	-0.118***	0.011
age: (30,.)	0.160***	0.045	0.134**	0.049	0.152***	0.024
age: (40,.)	0.015	0.042	-0.018	0.031	-0.099***	0.025
age: (50,.)	0.102**	0.034	0.144***	0.013	0.194***	0.015
univ	-0.748***	0.083	-0.645***	0.078	-0.742***	0.059
nch03	0.269*	0.109	0.692***	0.109	0.557***	0.121
nch411	0.173**	0.066	0.039	0.078	0.195***	0.058
nch1215	0.043	0.074	0.204**	0.068	0.108	0.076
citizen	-0.329	0.443	-0.649**	0.226	-0.281	0.206
no_partner	-0.184	0.130	-0.262**	0.080	-0.159	0.089
spuniv	-0.141	0.125	-0.143	0.130	-0.053	0.102
spinwork	-0.309**	0.119	-0.288**	0.105	-0.271**	0.101
Constant	4.490***	0.757	3.903***	0.561	3.035***	0.312
theta						
Constant	0.536	0.321	-8.847***	0.515	-9.105***	0.571
N	5365		5864		5039	
p	.		.		0	

Table A10: Average wage rate - male

	2007		2011		2015	
	coef	se	coef	se	coef	se
a						
age: (..)	-0.012	0.020	0.022	0.016	0.061***	0.018
age: (30,.)	-0.001	0.025	-0.016	0.024	-0.113***	0.030
age: (40,.)	0.017	0.022	-0.026	0.021	0.096**	0.029
age: (50,.)	-0.001	0.020	-0.005	0.015	-0.075***	0.023
univ	0.210**	0.080	-0.138*	0.064	-0.061	0.091
marr	0.191	0.144	0.198**	0.063	0.015	0.097
No of children	-0.008	0.044	-0.097*	0.040	-0.006	0.049
firm_size1	-0.071	0.068	-0.280***	0.058	-0.142	0.086
Constant	1.645**	0.503	0.864*	0.406	0.122	0.440
b						
age: (..)	0.025	0.025	0.001	0.020	-0.008	0.014
age: (30,.)	-0.051	0.030	-0.020	0.029	0.032	0.021
age: (40,.)	-0.000	0.022	0.012	0.019	-0.066**	0.020
age: (50,.)	-0.000	0.019	0.006	0.016	0.030	0.016
univ	0.020	0.065	0.266***	0.071	0.143*	0.057
marr	-0.002	0.155	-0.113	0.069	0.113	0.059
No of children	0.006	0.043	0.022	0.040	-0.081**	0.030
citizen	-0.011	0.099	0.135	0.109	0.048	0.108
firm_size1	-0.184**	0.068	-0.185**	0.058	-0.176***	0.051
occ_eur== 1.0000	0.339***	0.063	0.213***	0.058	0.213**	0.068
occ_eur== 2.0000	0.155*	0.063	0.189***	0.056	0.226***	0.066
occ_eur== 3.0000	0.095	0.065	0.018	0.058	0.172*	0.068
occ_eur== 5.0000	-0.191**	0.066	-0.127*	0.061	-0.038	0.069
occ_eur== 6.0000	0.041	0.058	-0.149**	0.053	0.027	0.063
occ_eur== 7.0000	0.021	0.060	-0.161**	0.053	-0.027	0.063
occ_eur== 8.0000	-0.182**	0.063	-0.219***	0.057	-0.132*	0.066
ind_saps== 1.0000	-0.285***	0.042	-0.232***	0.045	-0.273***	0.044
ind_saps== 3.0000	0.013	0.020	-0.066**	0.023	-0.062**	0.024
Work history (length of time in months)	0.001***	0.000	0.001***	0.000	0.001**	0.000
Constant	0.553	0.641	1.048	0.538	1.284***	0.386
q						
age: (..)	0.024	0.058	-0.047	0.041	-0.119***	0.034
age: (30,.)	-0.076	0.072	0.010	0.062	0.203***	0.054
age: (40,.)	0.046	0.053	0.045	0.047	-0.171**	0.055
age: (50,.)	-0.013	0.046	0.013	0.035	0.111*	0.043
univ	-0.798***	0.169	0.180	0.154	-0.108	0.157
marr	-0.191	0.353	-0.343*	0.150	0.081	0.174
No of children	0.081	0.108	0.004	0.087	-0.180*	0.087
firm_size1	-0.101	0.168	0.067	0.141	0.108	0.152
Work history (length of time in months)	0.001	0.001	0.002**	0.001	0.001	0.001
Constant	-0.202	1.473	1.291	1.060	2.404**	0.873
N	2514		2784		2383	
p	0.039		0		0	

Table A11: Number of hours worked

	Females						Males					
	2007		2011		2015		2007		2011		2015	
	coef	se	coef	se	coef	se	coef	se	coef	se	coef	se
univ	17.884	14.378	6.467	7.498	18.910**	7.090	5.319	8.247	32.844***	7.1	16.718*	8.219
nch03	-4.652***	1.156	-4.896***	1.293	-5.998***	1.364	-0.492	1.347	1.430	1.083	-0.520	1.124
nch411	-0.898	0.728	-1.091	0.686	-1.219	0.692	-0.039	0.6	-0.576	0.707	-0.046	0.743
nch1215	0.083	0.959	-1.162	0.905	-2.189*	0.902	-1.272	0.664	0.223	0.813	-2.001	1.109
marr	1.231	0.786	2.273**	0.788	-0.389	0.658	3.775***	0.934	2.804*	1.207	4.106***	0.962
age	1.103**	0.344	1.750***	0.206	1.929***	0.185	0.639**	0.214	1.715***	0.206	1.265***	0.226
age2	-0.014***	0.004	-0.019***	0.002	-0.020***	0.002	-0.007**	0.002	-0.018***	0.002	-0.014***	0.003
ageuniv	-1.014	0.732	0.118	0.354	-0.435	0.325	-0.153	0.385	-1.175***	0.326	-0.529	0.375
age2univ	0.014	0.009	-0.003	0.004	0.003	0.004	0.001	0.004	0.011**	0.004	0.005	0.004
Constant	15.120	7.749	-8.673*	4.278	-12.369**	3.977	22.461***	4.155	-7.899	4.041	4.687	4.754
N	2834		3166		2856		2844		2945		2630	
r2	0.076		0.153		0.172		0.04		0.133		0.109	
p	0		0		0		0		0		0	

Table A12: Self-employment income (receipt, amount)

For receipt, see A21,A22 and A23.

(a) Amount

	Females						Males					
	2007		2011		2015		2007		2011		2015	
	coef	se	coef	se	coef	se	coef	se	coef	se	coef	se
univ	4.248*	2.019	-1.988	2.960	-6.305	3.663	0.881	1.241	-1.989	6.384	-1.474	1.712
nch03	-0.033	0.256	-0.178	0.317	0.003	0.271	0.109	0.194	0.088	0.334	0.326	0.284
nch411	-0.146	0.209	-0.215	0.185	0.002	0.210	-0.101	0.11	0.185	0.208	0.049	0.179
nch1215	0.063	0.153	0.104	0.167	0.309	0.473	-0.194	0.135	0.103	0.095	0.247	0.208
marr	0.108	0.166	0.389*	0.167	-0.115	0.204	0.807***	0.224	0.282	0.225	0.164	0.241
age	0.039	0.029	0.033	0.052	0.108*	0.045	-0.023	0.025	0.071	0.084	0.063	0.045
age2	-0.000	0.000	-0.001	0.000	-0.002***	0.000	-0.000	0	-0.001	0.001	-0.001*	0
citizen	2.181***	0.645	0.291	0.648	-0.132	0.369	-0.484	0.303	0.000	0	-1.035	0.603
firm_size1	0.633	0.541	-0.768	0.807	-0.585	0.437	-0.467	0.263	0.000	0	-0.017	0.359
occ_eur== 1.0000	0.136	0.593	0.210	0.339	-0.154	0.389	1.386	0.907	-0.364	0.956	0.951	0.515
occ_eur== 2.0000	0.345	0.652	0.269	0.580	0.166	0.364	1.474	0.926	1.301	0.988	0.590	0.403
occ_eur== 3.0000	-0.851	0.612	-0.133	0.780	-1.034	0.614	1.539	0.923	0.125	1.034	-0.410	0.55
occ_eur== 5.0000	0.247	0.576	0.797*	0.389	0.253	0.334	0.849	0.931	0.399	0.986	1.045*	0.448
occ_eur== 6.0000	0.431	0.639	-0.091	0.418	-0.121	0.405	1.500	0.906	-0.422	0.943	-0.360	0.425
occ_eur== 7.0000	-0.201	0.597	-5.543***	0.355	-0.823	0.574	1.611	0.951	-0.176	0.941	-0.774	0.485
occ_eur== 8.0000	-0.048	0.593	0.070	0.359	-0.095	0.460	0.479	0.932	-0.577	0.948	0.598	0.476
ind_saps== 1.0000	1.333***	0.359	-0.235	0.491	0.039	0.550	0.238	0.181	0.215	0.602	-0.623	0.495
ind_saps== 3.0000	1.186***	0.282	-0.759	0.469	-0.361	0.437	0.200	0.144	-0.268	0.36	-0.688*	0.305
Work history (length of time in months)	-0.000	0.001	0.002*	0.001	0.005***	0.001	-0.000	0.001	0.001	0.001	0.002	0.002
Owner of individual enterprise with employees: 1 yes, 0 no employees	0.763***	0.180					0.615***	0.176				
People engaged in individual activities: 1 yes, 0 no	0.250	0.263	-0.470	0.867	-0.064	0.276	0.293	0.183	0.015	0.384	-0.287	0.369
ageuniv	-0.195*	0.092	0.057	0.128	0.255	0.137	-0.036	0.052	0.010	0.226	0.114	0.073
age2univ	0.002*	0.001	-0.000	0.001	-0.002*	0.001	0.000	0.001	0.000	0.002	-0.002*	0.001
Constant	0.390	1.156	4.994**	1.669	4.149**	1.435	5.446***	1.18	4.050	2.266	5.970***	1.231
Owner of inidvidual enterprise with employees: 1 yes, 0 no employees			0.264	0.480	1.557***	0.322			1.347**	0.489	-0.032	0.282
N	290		245		307		330		161		247	
r2	0.619		0.341		0.414		0.413		0.329		0.441	
p	0		.		.		0		0		.	

Table A13: Investment income (receipt, amount)

(a) Receipt

	Females						Males					
	2007		2011		2015		2007		2011		2015	
	coef	se	coef	se	coef	se	coef	se	coef	se	coef	se
univ	0.899	2.498	1.891	3.161	0.259	1.768	3.395	2.66	5.302*	2.142	2.795	1.858
nch03	-0.371	0.521	0.466	0.357	0.037	0.246	-0.346	0.54	0.319	0.346	-0.182	0.269
nch411	-0.180	0.232	0.121	0.203	0.163	0.154	-0.155	0.257	0.182	0.201	0.067	0.169
nch1215	0.010	0.252	0.193	0.236	0.038	0.196	-0.004	0.262	0.165	0.242	0.050	0.223
marr	1.059***	0.238	0.638***	0.188	0.811***	0.137	1.270**	0.4	0.742*	0.336	1.106***	0.21
age	0.086	0.049	0.259**	0.082	0.216***	0.045	0.107	0.067	0.306***	0.066	0.244***	0.045
age2	-0.001	0.000	-0.002**	0.001	-0.002***	0.000	-0.001	0.001	-0.002***	0.001	-0.002***	0
ageuniv	-0.010	0.095	0.010	0.108	0.043	0.065	-0.105	0.098	-0.130	0.081	-0.074	0.07
age2univ	0.000	0.001	-0.000	0.001	-0.000	0.001	0.001	0.001	0.001	0.001	0.001	0.001
constant	-6.217***	1.258	-11.904***	2.424	-8.186***	1.255	-6.984***	1.668	-12.951***	1.665	-8.752***	1.209
N	5659		6109		5346		4814		5115		4268	
chi2	74.507		132.948		245.965		119.573		105.83		195.298	
p	0		0		0		0		0		0	

(b) Amount

	Females						Males					
	2007		2011		2015		2007		2011		2015	
	coef	se	coef	se	coef	se	coef	se	coef	se	coef	se
univ	1.022	2.884	-0.619	3.660	-2.083	4.129	-0.493	3.306	5.49	5.7	-4.793	4.664
nch03	-0.877	0.529	-0.476	0.449	-0.561	0.662	-0.684	0.469	-0.29	0.38	-0.178	0.602
nch411	-0.178	0.413	-0.165	0.373	-0.076	0.446	-0.110	0.444	0.00	0.417	0.093	0.435
nch1215	-0.650	0.572	-0.731	0.398	-0.291	0.436	-0.927	0.548	-0.17	0.386	0.058	0.563
marr	-0.494	0.302	0.347	0.264	0.418	0.333	-0.930	0.621	-0.78	0.585	-1.355**	0.488
age	0.094	0.076	-0.237	0.125	0.182	0.108	0.194**	0.073	0.10	0.176	0.321*	0.127
age2	-0.001	0.001	0.002*	0.001	-0.001	0.001	-0.002**	0.001	0.00	0.001	-0.002*	0.001
citizen	-1.452***	0.405	-2.421***	0.458	-4.101***	0.655	-1.744***	0.343	0.00	0	0.819	1.84
ageuniv	-0.027	0.117	0.033	0.135	0.140	0.149	0.084	0.129	-0.21	0.201	0.294	0.177
age2univ	0.000	0.001	-0.000	0.001	-0.001	0.001	-0.001	0.001	0.00	0.002	-0.003	0.002
Constant	1.194	1.764	10.142**	3.634	-3.552	3.311	-0.950	2.177	-0.04	5.125	-10.499*	4.119
N	359		328		739		314		275.00		597	
r2	0.097		0.082		0.133		0.219		0.08		0.189	
p	.		.		0		0		0.09		0	

Table A14: Property income (receipt, amount)

(a) Receipt

	Females						Males					
	2007		2011		2015		2007		2011		2015	
	coef	se	coef	se	coef	se	coef	se	coef	se	coef	se
univ	1.411	2.847	3.541	2.332	-4.055	2.536	-2.652	3.21	2.830	2.93	-0.827	2.491
nch03	-0.262	0.362	0.248	0.367	0.629	0.338	-0.016	0.386	0.265	0.427	0.384	0.403
nch411	0.427*	0.199	-0.150	0.280	0.307	0.241	0.410*	0.207	-0.217	0.312	0.387	0.272
nch1215	0.482*	0.244	0.402	0.262	0.266	0.345	0.419	0.269	0.565	0.292	0.104	0.365
marr	0.732**	0.240	0.386	0.215	0.524*	0.226	1.113**	0.378	0.253	0.291	0.564	0.327
age	0.132*	0.058	0.211***	0.054	0.140*	0.064	0.117*	0.052	0.239***	0.054	0.201***	0.054
age2	-0.001	0.001	-0.002**	0.000	-0.001*	0.001	-0.001	0	-0.002***	0	-0.002**	0.001
ageuniv	-0.051	0.106	-0.112	0.085	0.171	0.091	0.123	0.12	-0.044	0.108	0.059	0.097
age2univ	0.001	0.001	0.001	0.001	-0.002	0.001	-0.001	0.001	0.000	0.001	-0.001	0.001
constant	-8.452***	1.550	-10.282***	1.455	-7.511***	1.942	-8.685***	1.391	-11.374***	1.559	-9.819***	1.372
N	5659		6109		5346		4814		5115		4268	
chi2	64.267		66.376		40.714		87.378		58.955		68.519	
p	0		0		0		0		0		0	

(b) Amount

	Females						Males					
	2007		2011		2015		2007		2011		2015	
	coef	se	coef	se	coef	se	coef	se	coef	se	coef	se
univ	-2.481	2.736	-0.023	5.118	12.321***	3.624	-7.30	5.019	-10.844*	4.546	2.980	5.276
nch03	0.795*	0.387	0.744	0.663	-0.843*	0.332	0.85	0.465	2.104**	0.669	-1.230**	0.469
nch411	-0.194	0.182	0.107	0.310	-0.117	0.243	-0.16	0.211	0.090	0.365	-0.655*	0.266
nch1215	-0.302	0.240	0.318	0.291	0.018	0.220	-0.25	0.331	0.687**	0.26	-0.476	0.558
marr	-0.936***	0.241	-0.713***	0.184	-0.222	0.375	-0.21	0.333	-0.566	0.298	0.754	0.696
age	-0.010	0.062	0.053	0.072	-0.123	0.082	-0.06	0.088	0.066	0.074	-0.514**	0.193
age2	0.000	0.001	-0.001	0.001	0.001	0.001	0.00	0.001	-0.000	0.001	0.004**	0.002
citizen	0.000	0.000	0.000	0.000	0.000	0.000	0.00	0	0.000	0	-1.357*	0.606
ageuniv	0.106	0.108	0.002	0.182	-0.388**	0.141	0.32	0.184	0.407*	0.162	-0.005	0.195
age2univ	-0.001	0.001	0.000	0.002	0.003*	0.001	0.00	0.002	-0.003*	0.001	-0.000	0.002
Constant	3.811*	1.857	1.751	2.152	5.510**	1.811	3.91	2.538	0.529	2.049	17.523**	5.294
N	234		226		213		183.00		169		166	
r2	0.122		0.122		0.272		0.15		0.218		0.319	
p	0.006		0		0		0.05		0.001		.	

Table A15: Other incomes (receipt, amount)

(a) Receipt

	Females						Males					
	2007		2011		2015		2007		2011		2015	
	coef	se	coef	se	coef	se	coef	se	coef	se	coef	se
univ	2.272	2.064	1.273	1.992	-5.307*	2.295	2.087	4.146	5.874	5.172	0.784	4.337
nch03	-0.825	0.582	0.046	0.353	0.675	0.502	0.906	0.476	0.727*	0.332	-0.524	0.637
nch411	-0.080	0.235	0.557*	0.267	-0.357	0.325	-0.254	0.401	-0.111	0.348	0.416	0.288
nch1215	0.005	0.293	0.063	0.294	-0.650	0.469	0.127	0.342	-0.790	0.731	-0.184	0.423
marr	-1.389***	0.243	-2.680***	0.332	-2.355***	0.349	0.603	0.601	-0.427	0.482	-0.711	0.456
age	0.069*	0.031	0.173***	0.031	0.016	0.031	0.099	0.069	0.131	0.069	0.040	0.042
age2	-0.001*	0.000	-0.002***	0.000	-0.000	0.000	-0.002	0.001	-0.002	0.001	-0.000	0
ageuniv	-0.120	0.087	-0.079	0.081	0.210*	0.094	-0.115	0.172	-0.233	0.261	0.020	0.197
age2univ	0.001	0.001	0.001	0.001	-0.002*	0.001	0.001	0.002	0.002	0.003	-0.000	0.002
constant	-4.719***	0.820	-3.565*	1.570	-3.142*	1.380	-5.790***	1.261	-4.931**	1.807	-3.541*	1.8
citizen			-3.037**	1.100	-0.150	1.063			-1.054	1.107	-1.408	1.222
N	6044		6434		5606		5275		5451		4533	
chi2	44.681		105.278		66.101		35.459		57.426		22.011	
p	0		0		0		0		0		0.015	

(b) Amount

	Females						Males					
	2007		2011		2015		2007		2011		2015	
	coef	se	coef	se	coef	se	coef	se	coef	se	coef	se
univ	0.233	2.553	-2.211	2.556	-4.937**	1.664	3.311	2.43	1.622	3.216	-1.955	2.22
nch03	-0.101	0.377	0.267	0.576	0.533	0.345	-0.030	0.255	0.010	0.367	-0.702	0.689
nch411	-0.667**	0.224	0.043	0.299	0.371	0.256	-0.219	0.206	0.008	0.268	-0.667	0.401
nch1215	-1.620***	0.466	-0.504*	0.214	0.898**	0.310	-0.200	0.297	-0.283	0.692	-1.884**	0.533
marr	-0.432	0.302	0.051	0.200	-0.127	0.285	0.683*	0.323	0.472	0.303	0.852	0.448
age	0.005	0.043	-0.018	0.063	-0.052	0.038	0.077	0.045	-0.084	0.043	-0.072	0.073
age2	-0.000	0.000	0.000	0.001	0.000	0.000	-0.001*	0.001	0.001	0	0.000	0.001
ageuniv	-0.042	0.100	0.106	0.102	0.188**	0.068	-0.206	0.118	-0.099	0.152	0.128	0.106
age2univ	0.001	0.001	-0.001	0.001	-0.002*	0.001	0.003*	0.001	0.001	0.002	-0.002	0.001
Constant	4.907***	1.304	4.847**	1.760	6.136***	0.799	2.322*	0.992	5.751***	0.881	6.270***	1.523
N	151		162		123		68		73		46	
r2	0.434		0.143		0.41		0.247		0.148		0.508	
p	0		0.007		0		0.016		0.482		0	

Table A16: State old age benefits (receipt and amount)

(a) Receipt

	Females						Males					
	2007		2011		2015		2007		2011		2015	
	coef	se	coef	se	coef	se	coef	se	coef	se	coef	se
univ	63.687	171.574	3086.470*	1562.159	-97.252	164.432	1265.274*	523.081	-62.800	47.803	336.982	631.387
marr	-0.110	0.317	-0.231	0.269	-0.039	0.270	0.204	0.383	0.066	0.353	0.372	0.41
age	12.328***	2.421	5.512***	0.732	8.940***	1.567	2.753	6.968	4.458***	0.763	1.584	4.351
age2	-0.086***	0.017	-0.038***	0.005	-0.061***	0.011	-0.016	0.055	-0.030***	0.005	-0.007	0.035
Work history (length of time in months)	-0.001	0.001	0.001	0.001	0.004***	0.001	0.002	0.001	0.000	0.001	0.001	0.002
retired	0.329	0.315	0.894**	0.347	0.249	0.352	0.406	0.309	0.481	0.251	1.310**	0.4
ageuniv	-1.855	4.987	-105.619*	53.440	2.844	4.747	-42.209*	17.274	1.805	1.381	-11.394	20.692
age2univ	0.013	0.036	0.903*	0.457	-0.020	0.034	0.351*	0.142	-0.013	0.01	0.096	0.169
constant	-429.473***	83.277	-192.893***	24.765	-318.412***	54.046	-112.334	221.588	-160.557***	26.357	-73.696	136.532
N	2354		2681		2551		1668		1971		1682	
chi2	111.231		234.056		305.054		188.982		156.366		136.932	
p	0		0		0		0		0		0	

(b) Amount

	Females						Males					
	2007		2011		2015		2007		2011		2015	
	coef	se	coef	se	coef	se	coef	se	coef	se	coef	se
univ	-4.470	4.726	-14.793*	7.286	-1.947	6.108	11.415	9.545	19.228**	6.818	5.814	12.849
marr	-0.085***	0.023	-0.004	0.022	0.003	0.025	-0.064	0.04	0.031	0.036	0.161***	0.048
age	0.107*	0.054	0.326***	0.055	0.289***	0.080	0.366**	0.117	0.987***	0.144	0.486*	0.203
age2	-0.001	0.000	-0.002***	0.000	-0.002***	0.001	-0.002**	0.001	-0.007***	0.001	-0.003*	0.001
citizen	0.079	0.083	-0.287	0.210	0.244	0.238	0.213	0.129	-0.145	0.223	-0.397*	0.186
Work history (length of time in months)	0.001***	0.000	0.000***	0.000	0.001***	0.000	0.000*	0	0.000**	0	0.000*	0
ageuniv	0.131	0.133	0.416*	0.204	0.068	0.169	-0.315	0.266	-0.534**	0.19	-0.155	0.355
age2univ	-0.001	0.001	-0.003*	0.001	-0.001	0.001	0.002	0.002	0.004**	0.001	0.001	0.002
Constant	0.894	1.941	-6.704***	1.975	-6.096*	2.897	-8.375*	4.196	-31.076***	5.206	-12.546	7.353
N	1947		2339		2074		1168		1466		1166	
r2	0.128		0.22		0.196		0.138		0.389		0.236	
p	0		0		0		0		0		0	

Table A17: Occupation (1-digit ISCO); for working individuals only

(a) 2-3

	2						3					
	2007		2011		2015		2007		2011		2015	
	coef	se	coef	se	coef	se	coef	se	coef	se	coef	se
univ	1.557	1.789	2.706	1.785	6.263**	2.018	0.960	1.903	1.134	1.843	7.850***	2.187
nch03	-0.385	0.245	-0.227	0.243	-0.145	0.241	-0.093	0.252	-0.156	0.275	-0.059	0.295
nch411	-0.626***	0.16	-0.250	0.144	0.008	0.146	-0.155	0.159	-0.360	0.195	-0.108	0.182
nch1215	-0.362	0.186	-0.380*	0.175	-0.111	0.211	-0.453*	0.219	-0.638**	0.221	0.351	0.234
marr	-0.075	0.189	-0.053	0.164	0.208	0.171	-0.291	0.218	0.039	0.197	-0.232	0.212
age	-0.091	0.05	-0.085	0.055	-0.035	0.054	-0.072	0.051	-0.109*	0.053	-0.000	0.058
age2	0.000	0.001	0.000	0.001	-0.000	0.001	0.000	0	0.001	0.001	0.000	0.001
ageuniv	0.001	0.081	-0.046	0.076	-0.172*	0.084	-0.055	0.087	-0.044	0.084	-0.329***	0.094
age2univ	0.000	0.001	0.001	0.001	0.002*	0.001	0.000	0.001	0.000	0.001	0.003**	0.001
sex	1.197	0.632	0.789	0.628	1.060	0.741	0.428	0.691	1.447*	0.673	1.431	0.828
sexuniv	-0.455	0.325	-0.371	0.308	-0.747*	0.32	0.241	0.356	-0.102	0.337	-0.203	0.374
agesex	0.010	0.013	0.017	0.012	0.017	0.014	0.009	0.014	-0.016	0.013	-0.014	0.016
constant	0.200	1.259	0.883	1.479	-1.425	1.519	1.347	1.446	1.461	1.398	-1.768	1.622
N	5674		6091		5475		5674		6091		5475	
chi2	1405.208		1628.927		1470.587		1405.208		1628.927		1470.587	
p	0		0		0		0		0		0	

(b) 4-5

	4						5					
	2007		2011		2015		2007		2011		2015	
	coef	se	coef	se	coef	se	coef	se	coef	se	coef	se
univ	0.461	2.772	-1.916	2.727	6.014*	2.655	-2.189	2.896	-0.604	2.017	5.934**	2.169
nch03	-0.053	0.359	-0.156	0.327	0.043	0.329	0.568*	0.273	-0.298	0.281	0.216	0.271
nch411	-0.235	0.236	-0.010	0.211	0.116	0.196	-0.154	0.155	-0.063	0.162	-0.105	0.163
nch1215	-0.713*	0.295	-1.047**	0.321	-0.596	0.352	-0.086	0.197	-0.527**	0.188	-0.277	0.243
marr	-0.462	0.294	-0.086	0.244	0.146	0.254	-0.317	0.218	-0.054	0.185	0.167	0.198
age	-0.086	0.063	-0.054	0.074	-0.057	0.066	-0.086	0.047	-0.026	0.051	-0.022	0.049
age2	0.001	0.001	0.000	0.001	0	0.001	0.000	0	0.000	0.001	0.000	0
ageuniv	-0.009	0.139	0.054	0.127	-0.238	0.123	0.090	0.145	-0.041	0.088	-0.295**	0.095
age2univ	-0.000	0.002	-0.001	0.001	0.002	0.001	-0.001	0.002	0.001	0.001	0.003**	0.001
sex	2.211**	0.824	1.388	0.825	1.241	1.045	1.810**	0.67	3.711***	0.663	3.065***	0.754
sexuniv	-0.277	0.533	0.455	0.507	-0.742	0.576	-0.785	0.469	-0.524	0.366	-1.068**	0.401
agesex	-0.017	0.017	-0.002	0.017	0.022	0.02	-0.003	0.014	-0.046***	0.013	-0.026	0.015
constant	-0.853	1.637	-0.641	1.859	-1.586	2.191	0.520	1.343	-1.565	1.41	-1.667	1.417
N	5674		6091		5475		5674		6091		5475	
chi2	1405.208		1628.927		1470.587		1405.208		1628.927		1470.587	
p	0		0		0		0		0		0	

(c) 6-7

	6						7					
	2007		2011		2015		2007		2011		2015	
	coef	se	coef	se	coef	se	coef	se	coef	se	coef	se
univ	-1.156	1.982	-2.661	1.915	3.019	2.365	1.413	3.424	1.868	3.156	2.228	2.769
nch03	0.158	0.27	-0.072	0.281	-0.072	0.272	0.474	0.316	0.018	0.333	0.061	0.298
nch411	-0.200	0.141	-0.250	0.151	0.111	0.147	-0.040	0.166	-0.280	0.188	0.157	0.169
nch1215	-0.309	0.179	-0.552**	0.176	-0.091	0.209	-0.148	0.203	-0.450*	0.207	-0.262	0.254
marr	0.088	0.223	-0.193	0.182	-0.251	0.187	-0.189	0.235	0.220	0.252	0.022	0.21
age	-0.058	0.042	-0.151**	0.047	-0.092*	0.045	0.121*	0.05	0.033	0.057	0.091	0.055
age2	0.000	0	0.002***	0	0.001*	0	-0.001*	0	-0.000	0.001	-0.000	0.001
ageuniv	-0.061	0.088	0.024	0.083	-0.191*	0.097	-0.173	0.14	-0.106	0.133	-0.125	0.126
age2univ	0.001	0.001	-0.000	0.001	0.002*	0.001	0.002	0.002	0.001	0.001	0.001	0.001
sex	0.260	0.609	1.778**	0.586	1.768**	0.685	-0.567	0.861	0.516	0.93	1.626	0.938
sexuniv	0.183	0.425	-0.234	0.375	-0.990*	0.446	0.431	0.779	-1.071	0.826	-0.794	0.588
agesex	-0.024	0.013	-0.048***	0.012	-0.040**	0.013	-0.024	0.017	-0.046*	0.018	-0.060***	0.018
constant	3.466**	1.193	4.637***	1.255	3.023*	1.243	-0.661	1.466	0.987	1.582	-1.319	1.607
N	5674		6091		5475		5674		6091		5475	
chi2	1405.208		1628.927		1470.587		1405.208		1628.927		1470.587	
p	0		0		0		0		0		0	

(d) 8

	8					
	2007		2011		2015	
	2007		2011		2015	
	coef	se	coef	se	coef	se
univ	-1.385	2.388	-2.964	2.346	3.082	2.704
nch03	0.028	0.282	-0.182	0.299	0.331	0.279
nch411	-0.014	0.154	-0.115	0.172	0.374*	0.163
nch1215	-0.132	0.209	-0.543**	0.202	-0.201	0.225
marr	-0.470*	0.208	-0.454*	0.186	-0.630***	0.191
age	-0.092*	0.043	-0.121**	0.047	-0.036	0.045
age2	0.001	0.000	0.001*	0.000	-0.000	0.000
ageuniv	-0.071	0.106	-0.044	0.107	-0.231*	0.113
age2univ	0.001	0.001	0.001	0.001	0.003*	0.001
sex	-0.104	0.609	0.081	0.632	0.520	0.745
sexuniv	0.099	0.478	0.171	0.456	-0.693	0.461
agesex	0.014	0.012	0.005	0.012	0.009	0.014
constant	2.374*	1.205	3.749**	1.290	1.379	1.279
N	5674		6091		5475	
chi2	1405.208		1628.927		1470.587	
p	0		0		0	

Table A18: Industry sector (primary - control, secondary or tertiary); for working individuals only

	Secondary						Tertiary					
	2007		2011		2015		2007		2011		2015	
	coef	se	coef	se	coef	se	coef	se	coef	se	coef	se
univ	-2.685	3.066	-3.740	3.304	1.553	3.52	-2.589	2.901	-3.487	3.063	-0.682	3.243
nch03	0.144	0.291	-0.052	0.277	-0.627*	0.316	0.016	0.278	-0.042	0.230	-0.499	0.284
nch411	-0.204	0.148	-0.031	0.212	0.018	0.227	-0.311*	0.142	0.069	0.193	0.184	0.217
nch1215	-0.533**	0.185	-0.493*	0.205	-0.187	0.261	-0.406*	0.161	-0.357*	0.177	-0.121	0.229
marr	0.231	0.224	0.081	0.24	0.413	0.221	0.354	0.207	0.024	0.207	0.472*	0.204
age	-0.076	0.06	0.043	0.058	0.079	0.059	-0.127*	0.054	-0.195***	0.048	-0.134**	0.050
age2	0.001	0.001	-0.000	0.001	-0.001*	0.001	0.001*	0.001	0.002***	0.000	0.002***	0.000
ageuniv	0.041	0.13	0.174	0.144	-0.039	0.146	0.080	0.121	0.178	0.132	0.066	0.131
age2univ	0.000	0.001	-0.002	0.002	0.000	0.002	-0.000	0.001	-0.002	0.001	-0.001	0.001
sex	0.734	0.723	1.939**	0.721	1.203	0.78	0.850	0.670	1.564*	0.612	2.088**	0.668
sexuniv	1.099	0.662	0.688	0.48	0.230	0.505	1.025	0.638	1.039*	0.434	0.565	0.462
agesex	-0.019	0.015	-0.037*	0.015	-0.011	0.016	-0.008	0.014	-0.018	0.013	-0.022	0.013
constant	3.204*	1.545	-1.077	1.483	-1.325	1.605	4.307**	1.427	4.976***	1.205	2.975*	1.366
N	5008		6111		5486							
chi2	184.93		252.821		224.973							
p	0		0		0							

Table A19: Retired

	Females						Males					
	2007		2011		2015		2007		2011		2015	
	coef	se	coef	se	coef	se	coef	se	coef	se	coef	se
univ	211.092	210.749	138.328	81.220	347.443*	143.265	-303.337**	108.961	-39.94	76.364	-27.926	68.644
nch015	-0.223	0.225	-0.354	0.259	1.613	1.938	-0.172	0.33	0.79	0.478	0.009	1.081
marr	-0.196	0.462	-0.130	0.337	-0.943*	0.397	1.351**	0.497	-0.35	0.331	0.342	0.409
age	4.937*	2.027	4.216***	0.602	5.856***	0.998	-8.925*	3.63	0.03	2.255	-3.727*	1.501
age2	-0.032*	0.016	-0.029***	0.004	-0.040***	0.007	0.082**	0.032	0.00	0.018	0.036**	0.013
ageuniv	-7.094	7.418	-4.417	2.758	-11.404*	4.959	10.644**	3.759	1.63	2.449	1.100	2.346
age2univ	0.060	0.065	0.035	0.023	0.094*	0.043	-0.093**	0.032	-0.02	0.02	-0.010	0.02
constant	-178.829**	63.060	- 148.812***	20.693	- 208.951***	33.906	238.498*	103.493	-19.59	68.944	94.473*	43.407
N	2001		2212		1927		1241		1490.00		1183	
chi2	116.48		120.981		176.091		50.27		122.91		76.962	
p	0		0		0		0		0.00		0	

Table A20: Unemployed

	Females						Males					
	2007		2011		2015		2007		2011		2015	
	coef	se	coef	se	coef	se	coef	se	coef	se	coef	se
univ	4.546	4.870	11.144***	3.361	1.164	4.122	7.699	5.214	8.645*	4.178	3.084	6.556
nch03	-0.421	0.533	-0.179	0.339	-0.446	0.375	1.034*	0.5	-0.973	0.769	0.400	0.542
nch411	-0.599	0.327	-0.021	0.255	0.069	0.228	0.209	0.246	-0.241	0.417	0.188	0.35
nch1215	0.162	0.432	-0.262	0.238	-0.336	0.330	-0.190	0.624	-0.005	0.264	-0.022	0.389
marr	-0.151	0.395	-0.369	0.334	-0.315	0.288	-0.247	0.392	0.549	0.331	-0.369	0.369
age	0.376***	0.099	0.414***	0.071	0.400***	0.074	0.379***	0.068	0.607***	0.058	0.254***	0.055
age2	-0.004**	0.001	-0.005***	0.001	-0.005***	0.001	-0.004***	0.001	-0.008***	0.001	-0.003***	0.001
ageuniv	-0.227	0.265	-0.620***	0.183	-0.151	0.199	-0.368	0.286	-0.474*	0.226	-0.019	0.367
age2univ	0.003	0.003	0.008***	0.002	0.003	0.002	0.004	0.003	0.006*	0.003	-0.001	0.005
constant	-10.341***	1.797	-8.787***	1.156	-9.030***	1.240	-8.927***	1.239	-10.982***	0.925	-6.000***	0.941
N	1067		985		773		932		980		694	
chi2	43.774		72.056		49.107		51.5		139.529		44.03	
p	0		0		0		0		0		0	

Table A21: Owner of enterprise with employees (sub-group of self-employed)

	Females						Males					
	2007		2011		2015		2007		2011		2015	
	coef	se	coef	se	coef	se	coef	se	coef	se	coef	se
univ	15.539	10.995	-17.43	17.263	1.27	6.464	0.20	4.674	-38.14	39.628	-5.63	10.429
nch015	0.181	0.420	0.68	0.535	-1.49	0.826	0.22	0.226	0.25	0.576	0.64	0.43
marr	-0.596	0.429	0.28	0.957	-0.12	0.701	-0.33	0.384	-0.34	0.998	0.46	0.719
age	-0.251**	0.090	0.08	0.355	0.25	0.171	0.04	0.073	1.16	0.71	0.15	0.335
age2	0.003***	0.001	0.00	0.004	0.00	0.002	0.00	0.001	-0.01	0.006	0.00	0.003
ageuniv	-0.517	0.454	0.96	0.857	-0.21	0.300	0.08	0.184	1.70	1.424	0.28	0.386
age2univ	0.005	0.004	-0.01	0.010	0.00	0.003	0.00	0.002	-0.02	0.013	0.00	0.003
constant	3.307	2.133	-3.63	9.115	-4.34	2.858	-3.34	1.934	-37.74	21.163	-7.59	9.437
N	290		245.00		307.00		330.00		161.00		247.00	
chi2	50.482		36.75		5.64		31.35		21.41		20.40	
p	0		0.00		0.58		0.00		0.00		0.00	

Table A22: Has business certificate (sub-group of self-employed)

	Females						Males					
	2007		2011		2015		2007		2011		2015	
	coef	se	coef	se	coef	se	coef	se	coef	se	coef	se
univ	6.36	7.000	-1.31	7.416	1.573	6.919	-13.879	14.549	2.772	18.96	10.720	11.692
nch03	-0.38	0.898	-1.16	0.767	0.458	0.701	-0.703	0.709	0.656	0.971	0.087	0.65
nch411	0.28	0.459	-1.74	1.093	-0.518	0.504	-0.115	0.396	0.116	0.622	0.163	0.371
nch1215	-0.30	0.461	0.72	0.537	-2.086	1.123	-0.547	0.424	-0.546	0.475	1.092	0.595
marr	0.61	0.521	0.41	0.532	-0.867	0.480	1.386	0.865	0.894	0.716	0.786	0.615
age	0.02	0.100	0.11	0.213	0.315**	0.099	-0.228*	0.103	0.348	0.182	0.258*	0.122
age2	0.00	0.001	0.00	0.002	-0.004***	0.001	0.001	0.001	-0.004*	0.002	-0.003*	0.001
ageuniv	-0.31	0.292	0.03	0.309	-0.060	0.298	0.623	0.819	-0.088	0.763	-0.398	0.616
age2univ	0.00	0.003	0.00	0.003	0.000	0.003	-0.008	0.011	0.000	0.007	0.003	0.008
constant	0.07	2.245	-1.92	5.130	-5.688*	2.430	5.510*	2.434	-8.504*	3.847	-7.150*	2.868
N	290.00		245.00		307		330		161		247	
chi2	25.62		20.49		32.656		32.517		18.124		21.8	
p	0.00		0.02		0		0		0.034		0.01	

Table A23: Engaged in individual activities (sub-group of self-employed)

	Females						Males					
	2007		2011		2015		2007		2011		2015	
	coef	se	coef	se	coef	se	coef	se	coef	se	coef	se
univ	-2.406	7.419	10.951	8.554	9.164*	4.025	7.820	8.176	5.224	5.038	-0.671	5.077
nch411	-1.419	0.814	0.263	0.341	-0.103	0.317	0.248	0.276	0.248	0.478	-0.004	0.325
nch1215	1.194***	0.356	-1.142	0.768	0.162	0.544	0.362	0.334	-0.450	0.721	0.765*	0.37
marr	-0.718	0.571	2.254*	1.072	-0.076	0.456	-0.464	0.469	-0.083	1.003	0.755	0.829
age	0.221	0.240	0.592*	0.249	0.185	0.102	0.644*	0.27	0.057	0.109	0.022	0.175
age2	-0.002	0.002	-0.006*	0.003	-0.002	0.001	-0.007*	0.003	-0.000	0.001	-0.001	0.002
ageuniv	0.090	0.334	-0.504	0.399	-0.361*	0.160	-0.263	0.382	-0.096	0.229	0.054	0.227
age2univ	-0.000	0.004	0.006	0.004	0.004*	0.002	0.002	0.004	0.000	0.002	0.000	0.002
constant	-9.380	5.372	-20.512***	5.816	-8.548**	2.606	-18.596**	6.149	-7.474**	2.321	-4.614	3.533
nch03			0.184	0.484	-0.458	0.809			0.537	0.34	-0.381	0.713
N	2834		3166		2856		2844		2945		2630	
chi2	19.052		41.257		15.598		15.866		21.89		57.401	
p	0.015		0		0.076		0.044		0.009		0	

Table A24: Descriptive statistics by sex and event time

	Event time t	-2	-1	0	1	2
Sector (Nace 1.1)						
Agriculture	Females	0.02	0.01	0.05	0.08	0.02
Agriculture	Females-Males	-0.02	-0.02	0.02	0.07*	-0.03
Construction	Females	0.01	0.02	0.01	0.01	0.00
Construction	Females-Males	-0.12***	-0.09***	-0.10***	-0.17***	-0.20**
Financial	Females	0.06	0.05	0.04	0.04	0.01
Financial	Females-Males	0.04	0.04**	0.02	0.04*	0.01
Hotels and restaurants	Females	0.09	0.04	0.06	0.02	0.19
Hotels and restaurants	Females-Males	0.07	-0.01	0.04	-0.01	0.18
Industry	Females	0.06	0.10	0.10	0.17	0.21
Industry	Females-Males	-0.25***	-0.19***	-0.16**	-0.11	-0.03
Other services	Females	0.09	0.09	0.09	0.03	0.01
Other services	Females-Males	0.05	0.03	0.06	-0.01	-0.02
Public admin, education and health	Females	0.31	0.35	0.39	0.32	0.35
Public admin, education and health	Females-Males	0.18**	0.19**	0.24*	0.22**	0.19
Real estate and renting	Females	0.11	0.09	0.04	0.04	0.07
Real estate and renting	Females-Males	0.10***	0.03	-0.01	-0.03	-0.06
Transport, storage and communication	Females	0.04	0.06	0.06	0.10	0.09
Transport, storage and communication	Females-Males	-0.14***	-0.04	-0.05	-0.02	0.03
Wholesale and retail	Females	0.22	0.18	0.17	0.17	0.06
Wholesale and retail	Females-Males	0.10	0.04	-0.06	0.02	-0.09
Occupation						
Armed forces	Females	0.02	0.01	0.01	0.01	0.01
Armed forces	Females-Males	0.00	-0.01	-0.03***	-0.03**	-0.03
Managers	Females	0.08	0.05	0.05	0.08	0.05
Managers	Females-Males	-0.11**	-0.12***	-0.06**	-0.03	-0.08
Professionals	Females	0.28	0.30	0.34	0.25	0.35
Professionals	Females-Males	0.12**	0.12**	0.13**	0.01	0.13
Technical	Females	0.19	0.18	0.20	0.21	0.20
Technical	Females-Males	0.06	0.05	0.06	0.09*	0.05
Clerical	Females	0.11	0.09	0.06	0.07	0.04
Clerical	Females-Males	0.09***	0.08***	0.04**	0.06**	0.01
Service	Females	0.20	0.20	0.19	0.22	0.26
Service	Females-Males	0.15***	0.12***	0.12***	0.16***	0.17*
Skiled agricultural	Females	0.01	0.01	0.02	0.02	0.01
Skiled agricultural	Females-Males	-0.02	-0.01	-0.01	0.00	0.00
Craft	Females	0.02	0.04	0.03	0.03	0.02
Craft	Females-Males	-0.23***	-0.17***	-0.18***	-0.16***	-0.16***
Plant workers	Females	0.03	0.03	0.03	0.04	0.02
Plant workers	Females-Males	-0.10***	-0.09***	-0.09***	-0.11***	-0.10**
Elementary	Females	0.06	0.08	0.08	0.06	0.04
Elementary	Females-Males	0.03*	0.02	0.03	0.01	0.01

Note: Table shows females statistics and the differences between female and male statistics. Event time denotes the number of years before $(t < 0)$, during $(t=0)$ and after $(t >0)$ birth of first child. Stars (***, **, *) indicate signicance at the 1%, 5% and 10% respectively and are calculated using r Survey package and Z critical values.

Table A25: Estimating child penalty on labour market participation (positive earnings)

	Dependent variable:					
	positive earnings					
	Female		Male		Pooled	
	(1)	(2)	(3)	(4)	(5)	(6)
t=-2	0.034	−0.040	0.027	−0.015	0.033	−0.048
	(0.026)	(0.046)	(0.021)	(0.017)	(0.026)	(0.037)
t=0	−0.225***	−0.121***	−0.023	−0.005	−0.215***	−0.113***
	(0.022)	(0.039)	(0.018)	(0.008)	(0.023)	(0.037)
t=1	−0.539***	−0.613***	−0.086***	−0.016	−0.522***	−0.615***
	(0.037)	(0.066)	(0.030)	(0.014)	(0.038)	(0.065)
t=2	−0.349***	−0.273***	−0.054	−0.009	−0.327***	−0.326***
	(0.051)	(0.078)	(0.035)	(0.014)	(0.052)	(0.082)
male					0.033	0.007
					(0.027)	(0.025)
t=-2:male					−0.015	0.033
					(0.033)	(0.040)
t=0:male					0.184***	0.122***
					(0.029)	(0.038)
t=1:male					0.410***	0.591***
					(0.055)	(0.068)
t=2:male					0.233***	0.321***
					(0.062)	(0.085)
Age dummies	T	T	T	T	T	T
Year dummies	T	T	T	T	T	T
Country dummies	T	T	T	T	T	T
Sector dummies		T		T		T
Occupation dummies		T		T		T
Education dummies		T		T		T
Constant	0.091	−0.627*	−0.227*	−0.621***	0.054	−0.719***
	(0.159)	(0.349)	(0.130)	(0.197)	(0.135)	(0.259)
Observations	2,280	485	1,725	603	4,005	1,088
Log Likelihood	−1,662.537	−198.703	−658.897	334.236	−2,489.512	−208.934
Akaike Inf. Crit.	3,435.075	529.406	1,447.794	−508.471	5,125.024	589.868

*Note: Standard errors (found in parenthesis) incorporate information on the strata (country) and primary sample unit (individual) and therefore are clustered at the individual level. *p<0.1; **p<0.05; ***p<0.01*

Table A26: Estimating child penalty on months worked

	Dependent variable:					
	months_t					
	Female		Male		Pooled	
	(1)	(2)	(3)	(4)	(5)	(6)
t=-2	0.194	−0.995*	0.192	−0.578*	0.170	−0.749
	(0.286)	(0.514)	(0.281)	(0.332)	(0.272)	(0.457)
t=0	−3.332***	−2.948***	−0.047	0.022	−3.316***	−2.971***
	(0.231)	(0.413)	(0.203)	(0.148)	(0.229)	(0.440)
t=1	−4.896***	−5.021***	−0.305	−0.076	−4.769***	−4.557***
	(0.373)	(0.679)	(0.362)	(0.252)	(0.384)	(0.691)
t=2	−3.348***	−1.931**	0.594*	−0.083	−3.130***	−1.895***
	(0.540)	(0.775)	(0.332)	(0.352)	(0.540)	(0.640)
male					0.376	0.046
					(0.322)	(0.342)
t=-2:male					0.008	0.155
					(0.393)	(0.585)
t=0:male					3.192***	3.009***
					(0.298)	(0.481)
t=1:male					4.364***	4.358***
					(0.526)	(0.677)
t=2:male					3.390***	1.622**
					(0.595)	(0.658)
Age dummies	T	T	T	T	T	T
Year dummies	T	T	T	T	T	T
Country dummies	T	T	T	T	T	T
Sector dummies		T		T		T
Occupation dummies		T		T		T
Education dummies		T		T		T
Constant	1.374	5.186	−2.142	15.021***	0.498	3.023
	(1.348)	(3.670)	(1.581)	(1.791)	(1.190)	(1.920)
Observations	2,280	485	1,725	603	4,005	1,088
Log Likelihood	−7,049.392	−1,335.255	−4,964.239	−1,301.012	−12,130.670	−2,842.121
Akaike Inf. Crit.	14,208.780	2,802.509	10,058.480	2,762.024	24,407.340	5,856.242

*Note: Standard errors (found in parenthesis) incorporate information on the strata (country) and primary sample unit (individual) and therefore are clustered at the individual level. *p<0.1; **p<0.05; ***p<0.01*

Table A27: Estimating child penalty on monthly wage

	Dependent variable:					
	monthly wage					
	Female		Male		Pooled	
	(1)	(2)	(3)	(4)	(5)	(6)
t=-2	0.024	0.016	0.109	0.221	0.053	0.109
	(0.052)	(0.095)	(0.090)	(0.135)	(0.050)	(0.087)
t=0	−0.327***	−0.260***	−0.274***	−0.187**	−0.374***	−0.325***
	(0.049)	(0.081)	(0.081)	(0.083)	(0.052)	(0.093)
t=1	−0.560***	−0.549***	−0.396***	−0.348**	−0.680***	−0.841***
	(0.063)	(0.092)	(0.124)	(0.133)	(0.082)	(0.132)
t=2	−0.437***	−0.519***	−0.488***	−0.585***	−0.563***	−0.605***
	(0.103)	(0.120)	(0.163)	(0.217)	(0.127)	(0.171)
male					0.381***	0.528***
					(0.102)	(0.192)
t=-2:male					0.033	0.135
					(0.099)	(0.164)
t=0:male					0.138	0.180
					(0.087)	(0.139)
t=1:male					0.351**	0.522***
					(0.141)	(0.190)
t=2:male					0.155	0.176
					(0.196)	(0.281)
Age dummies	T	T	T	T	T	T
Year dummies	T	T	T	T	T	T
Country dummies	T	T	T	T	T	T
Sector dummies		T		T		T
Occupation dummies		T		T		T
Education dummies		T		T		T
Constant	−0.211	1.819**	−0.138	−1.680**	−0.482***	−0.940
	(0.217)	(0.779)	(0.237)	(0.680)	(0.165)	(0.961)
Observations	1,416	412	1,569	589	2,985	1,001
Log Likelihood	−1,836.418	−438.476	−2,680.794	−867.591	−4,689.092	−1,421.803
Akaike Inf. Crit.	3,768.836	1,004.952	5,489.589	1,895.183	9,518.184	3,013.606

*Note: Standard errors (found in parenthesis) incorporate information on the strata (country) and primary sample unit (individual) and therefore are clustered at the individual level. *p<0.1; **p<0.05; ***p<0.01*

Table A28: Estimating child penalty on hours

	Dependent variable:					
	hours					
	Female		Male		Pooled	
	(1)	(2)	(3)	(4)	(5)	(6)
t=-2	6.609***	−0.989	−0.330	−3.140***	6.314***	−0.945
	(1.200)	(1.773)	(1.311)	(1.183)	(1.151)	(1.544)
t=0	−12.072***	−23.743***	0.212	−1.395*	−11.917***	−22.507***
	(1.022)	(2.136)	(0.961)	(0.791)	(1.017)	(2.336)
t=1	−8.219***	−20.380***	−0.924	−2.884***	−7.703***	−17.320***
	(1.416)	(2.233)	(1.368)	(1.095)	(1.402)	(2.240)
t=2	−3.770*	−11.064***	−0.087	−2.298	−2.772	−10.032***
	(1.936)	(3.358)	(1.566)	(1.677)	(1.933)	(2.959)
male					9.384***	3.116**
					(1.366)	(1.432)
t=-2:male					−6.531***	−1.050
					(1.739)	(1.857)
t=0:male					11.749***	20.282***
					(1.342)	(2.499)
t=1:male					6.170***	14.114***
					(2.050)	(2.399)
t=2:male					1.019	6.204**
					(2.274)	(3.099)
Age dummies	T	T	T	T	T	T
Year dummies	T	T	T	T	T	T
Country dummies	T	T	T	T	T	T
Sector dummies		T		T		T
Occupation dummies		T		T		T
Education dummies		T		T		T
Constant	1.722	18.092	1.453	47.060***	0.216	13.963
	(5.290)	(14.899)	(4.709)	(6.616)	(3.556)	(11.006)
Observations	2,280	485	1,725	603	4,005	1,088
Log Likelihood	−10,021.260	−1,986.942	−7,448.975	−2,246.604	−17,570.000	−4,371.186
Akaike Inf. Crit.	20,152.520	4,105.883	15,027.950	4,653.207	35,286.000	8,914.372

*Note: Standard errors (found in parenthesis) incorporate information on the strata (country) and primary sample unit (individual) and therefore are clustered at the individual level. *p<0.1; **p<0.05; ***p<0.01*

SUMMARY

Vertinant pajamų nelygybę Lietuvoje pabrėžiama, kad čia, kaip ir daugumoje pasaulio valstybių, ji yra didelė ir dar didėja. 2017 m., kai buvo pradėti tyrimai šiai disertacijai, namų ūkių disponuojamų pajamų nelygybė Lietuvoje buvo viena didžiausių iš 28 Europos Sąjungos (ES) šalių. Didelė pajamų nelygybė gali sulėtinti ekonominį

augimą ir skaldyti visuomenę, todėl ši tema yra aktuali tiek tarp politikų, tiek tarp mokslininkų. Pavyzdžiui, Jungtinių Amerikos Valstijų prezidentas Barakas Obama paskelbė, kad pajamų nelygybė yra viena iš opiausių šių laikų problemų („defining issue of our time“), o Lietuvos Respublikos 17 ir 18 vyriausybių vienas iš pagrindinių tikslų buvo ir yra mažinti pajamų nelygybę Lietuvoje. Mokslininkai, besigilinantys į nelygybę Lietuvoje, dažnai akcentuoja, kad čia mokesčių ir socialinių išmokų sistema palyginti nedaug prisideda prie pajamų nelygybės mažinimo.

Šios disertacijos tikslas yra geriau suprasti, kokią įtaką namų ūkių disponuojamų pajamų nelygybei Lietuvoje daro tiesioginiai mokesčiai ir socialinės išmokos ir tam tikri pervedimai natūra. Disponuojamos pajamos – tai pajamos, skirtos vartojimui, kurios susideda iš darbo ir kapitalo pajamų, privačių pervedimų tarp namų ūkių, socialinių išmokų (pvz., pensijų išmokų, nedarbo pašalpų) atėmus tiesioginius mokesčius (pvz., gyventojų pajamų mokestį ir socialines įmokas). Nors nesunku nuspėti, kad didesni progresiniai mokesčiai arba didesnės socialinės išmokos turėtų mažinti pajamų nelygybę, iki šiol nebuvo žinoma, kiek santykinai mažesnės išmokos ir mažesni progresiniai mokesčiai Lietuvoje paaiškina, kodėl pajamų nelygybė čia tokia didelė. Taip pat nėra aišku, kiek dėl santykinai mažų socialinių išmokų ir mokesčių padidėjo pajamų nelygybė Lietuvoje. Ne mažiau svarbūs ir asignavimai natūra – pavyzdžiui, valstybės finansuojamos vietos darželiuose. Atsiradus daugiau vietų darželiuose, bus daugiau galimybių abiem tėvams dirbti. Kadangi vaikus dažniausiai prižiūri mamos, didesnis darželių prieinamumas turėtų mažinti darbo pajamų skirtumus tarp vyrų ir moterų.

Šiam tikslui pasiekti buvo atlikti keturi tyrimai ir, naudojant gautus rezultatus, parašyti keturi moksliniai straipsniai. Šie tyrimai sudaro teikiamos disertacijos pagrindą. Atliekant tuos tyrimus buvo analizuojami skirtingi pajamų nelygybės aspektai, susiję su tiesioginiais mokesčiais ir išmokomis, buvo taikomos skirtingos metodikos ir pasiekti atitinkami rezultatai.

Pirmasis tyrimas leidžia paaiškinti, kodėl pajamų nelygybė 2015[8] m. Lietuvoje buvo didesnė negu kitose ES šalyse. Ypatingas dėmesys buvo skiriamas socialinių pervedimų ir tiesioginių mokesčių poveikiui pajamų nelygybei. Šiam poveikiui įvertinti buvo naudojamos nelygybės indeksų, apskaičiuotų naudojant pajamų ir gyvenimo sąlygų statistinio tyrimo duomenis, dekompozicijos. Nustatyta, kad santykinai didelę pajamų nelygybę paaiškina santykinai mažesnės nei kitose ES šalyse socialinės išmokos ir mažesni progresiniai mokesčiai. Pavyzdžiui, socialinių išmokų poveikis pajamų nelygybei Lietuvoje yra apie 50 proc. mažesnis, palyginti su 28 ES valstybėmis, o tiesioginių mokesčių (kurie apima ir socialines įmokas) mažesnis

8. Pradėjus rašyti disertaciją, tai buvo patys naujausi duomenys.

beveik du kartus. Tiesioginių mokesčių poveikis savarankiškai dirbantiesiems Lietuvoje yra neigiamas, tai yra pajamų nelygybė padidėja atskaičius mokesčius, nors kitose ES šalyse pajamų nelygybė atskaičius mokesčius tarp savarankiškai dirbančiųjų sumažėja. Taip pat nustatyta, kad, nors ir yra nelygybė tarp skirtingų populiacijos grupių (pavyzdžiui, tarp dirbančiųjų ir nedirbančiųjų), svarbesnis pajamų nelygybės veiksnys yra nelygybė pačiose grupėse. Pavyzdžiui, Lietuvoje pajamų nelygybė tarp bedarbių yra santykinai didesnė nei kitose šalyse. Prie šios nelygybės iš dalies prisideda santykinai trumpalaikės bedarbio išmokos, kurios taip pat yra susietos su buvusiu atlyginimu. Todėl ilgalaikiai bedarbiai arba trumpalaikiai bedarbiai, kurių atlyginimas buvo mažas, negaus nedarbo išmokų arba gaus jas daug mažesnes nei trumpalaikiai bedarbiai, kurių pajamos buvo didelės.

Antro tyrimo metu paaiškinama, kodėl pajamų nelygybė Lietuvoje 2007–2015 m. kito ir kokią įtaką tam kitimui turėjo tiesioginiai mokesčiai ir socialinės išmokos. Šiam tikslui buvo naudotos mikrosimuliacijos (įskaitant ir EUROMOD simuliacijas) ir dekompozicijos. Buvo identifikuota, kad pajamų nelygybė, ypač tarp 2011 m. ir 2015 m., didėjo dėl to, kad didėjo atlyginimai . Dėl to didėjo pajamų atotrūkis tarp dirbančiųjų ir nedirbančiųjų (pensininkų, bedarbių, vaikus prižiūrinčiųjų ir pan.). O mokesčių ir socialinių išmokų sistema per lėtai didino pajamas tų, kurie darbo pajamų neturėjo. Todėl pajamų nelygybė didėjo.

Trečiojo tyrimo metu paaiškinta, kodėl Lietuvoje pajamų nelygybė tarp dirbančiųjų, ypač savarankiškai dirbančiųjų, yra didesnė negu daugumoje ES šalių. Buvo tikrinama, kiek ir kodėl prie šios nelygybės prisideda „grynieji tiesioginiai mokesčiai“ (tiesioginiai mokesčiai, atėmus socialines išmokas). Buvo skaičiuojamas efektyvus, statutinis ir optimalus grynųjų mokesčių progresyvumas tarp dirbančiųjų pagal darbo sutartis ir savarankiškai dirbančiųjų ir šie trys rodikliai buvo palyginti tarpusavyje. Paaiškėjo, kad dirbantiems pagal darbo šių trijų skaičiavimų rezultatai yra panašūs (t. y. dirbantieji pagal darbo sutartis moka tiek mokesčių, kiek turėtų pagal įstatymą ir pagal optimalių mokesčių modelį). Tačiau savarankiškai dirbantieji moka gerokai mažiau mokesčių nei dirbantieji pagal darbo sutartis. Tai lėmė mažesni tarifai ir, tikėtina, mokesčių vengimas.

Ketvirto tyrimo metu buvo nagrinėjama vyrų ir moterų atlyginimų nelygybė gimus vaikui Baltijos šalyse. Tam buvo naudojama tam tikro pobūdžio „prieš ir po“ analizė (angl. *event study*), t. y. buvo matuojami vyrų ir moterų atlyginimai prieš gimstant pirmajam vaikui ir jam gimus. Buvo aptikta, kad moterų atlyginimai pirmus dvejus metus po vaiko gimimo labai sumažėja, o vyrų lieka nedaug pakitę. Taip pat buvo tiriama, ar didesnis valstybinių lopšelių-darželių skaičius (tam tikri asignavimai natūra) Baltijos šalyse sumažino šį pajamų atotrūkį. Rezultatai buvo skirtingi.

Šių tyrimų pagrindu yra teikiamos rekomendacijos dėl tolesnių tyrimų, padėsiančių dar geriau įvertinti tiesioginių mokesčių ir socialinių išmokų poveikį pajamų nelygybei Lietuvoje. Pirma, rekomenduojama šiuos įverčius skaičiuoti ėmus ne vien apklausos duomenis, bet ir administracinius duomenis. Tai leis įvertinti, kiek prie pajamų nelygybės prisideda daugiausia pajamų turintieji namų ūkiai. Taip pat tai leistų tiksliau įvertinti lopšelių-darželių skaičiaus įtaką vyrų ir moterų pajamų atotrūkiui. Antra, reikėtų tiksliau įvertinti dirbančiųjų pagal darbo sutartis ir savarankiškai dirbančiųjų tiesioginių mokesčių tarifų elastingumą pajamų atžvilgiu. Šie elastingumai padėtų paaiškinti, ar tikrai vertėtų didinti tarifus, ypač savarankiškai dirbantiesiems, siekiant surinkti daugiau mokesčių. Šiam tikslui būtų galima nagrinėti 2018 m. reformą, kurios metu savarankiškai dirbančiųjų tarifai pasikeitė. Taip pat vertėtų atlikti naudos ir sąnaudų analizę, siekiant įvertinti, kiek verta dėti pastangų, kad būtų išieškoti mokesčiai iš savarankiškai dirbančiųjų. Trečia, vertėtų analizuoti mokesčių ir socialinių išmokų poveikį ir kitiems pajamų konceptams. Pavyzdžiui, reikėtų įvertinti, kokia yra pajamų nelygybė sumokėjus pridėtinės vertės mokestį ir akcizo mokestį (šie mokesčiai dažniausiai būna regresiniai). Taip pat vertėtų atsižvelgti į tai, kad daugybę paslaugų valstybė teikia nemokamai. Pavyzdžiui, Lietuvoje valstybė teikia daug nemokamų švietimo ir sveikatos paslaugų, kuriomis labiau naudojasi santykinai mažai pajamų turintys gyventojai (mokiniai, studentai, pensininkai).

Trumpas gyvenimo aprašymas Nerijus Černiauskas domisi pajamų nelygybės, socialinių išmokų ir mokesčių sistemų temomis. 2012 m. įgijo ekonomikos ir politikos mokslų bakalauro laipsnį Birmingamo universitete, Anglijoje, o 2014 m. viešojo sektoriaus ekonomikos magistro laipsnį Amsterdamo universitete, Nyderlanduose. 2018 m. įstojo į doktorantūros studijas Vilniaus universitete. Nuo 2009 iki 2017 m. dirbo analitiku įmonėse „Sveikatos ekonomikos centras“ ir „Euromonitor International“, o nuo 2015 m. iki 2020 m. ekonomistu Lietuvos banke. Nuo 2018 m. pradėjo dėstyti Vilniaus universitete. Nuo 2020 m. dirba vyr. specialistu Lietuvos Respublikos finansų ministerijoje.

ACKNOWLEDGMENT

I am grateful to the patience and support of my supervisor Prof. Romas Lazutka, the researchers and staff of Faculty of economics and business administration of Vilnius University, the Macroeconomics and Forecasting Division, the Applied Macroeconomics Research Division, the Center for Excellence in Finance and Economic Research at the Bank of Lithuania, the editors of Baltic Journal of economics, IZA journal of Labour policy, the editor of the special issue of Review of Income and Wealth and the anonymous referees of the respective journals. I am indebted to my coauthors

who have worked day and night of various parts of the articles.

Above all, I thank my mother, father, sister, and friends who supported me along the way.

This dissertation is based on data from Eurostat, EU Statistics on Income and Living Conditions [2004:2018]. The responsibility for all conclusions drawn from the data lies entirely with me.

PUBLICATION LIST AND THEIR COPIES

1st publication

Measurement and decomposition of Lithuania's income inequality

N. Černiauskas, A. Čiginas.

Baltic Journal of Economics **20**, 139-169 (2020)

DOI: 10.1080/1406099X.2020.1780693

Baltic Journal of Economics

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/rbec20

Measurement and decomposition of Lithuania's income inequality

Nerijus Černiauskas & Andrius Čiginas

To cite this article: Nerijus Černiauskas & Andrius Čiginas (2020) Measurement and decomposition of Lithuania's income inequality, Baltic Journal of Economics, 20:2, 139-169, DOI: 10.1080/1406099X.2020.1780693

To link to this article: https://doi.org/10.1080/1406099X.2020.1780693

BALTIC JOURNAL OF ECONOMICS
2020, VOL. 20, NO. 2, 139–169
https://doi.org/10.1080/1406099X.2020.1780693

Routledge
Taylor & Francis Group

Measurement and decomposition of Lithuania's income inequality*

Nerijus Černiauskas [a] and Andrius Čiginas [b]

[a]Economics Department, Bank of Lithuania & Faculty of Economics and Business Administration, Vilnius University, Vilnius, Lithuania; [b]Economics Department, Bank of Lithuania & Faculty of Mathematics and Informatics, Vilnius University, Vilnius, Lithuania

ABSTRACT

Even though Lithuania's household income inequality is among the highest in the European Union (EU), little empirical work has been carried out to explain such disparities. We investigate it using the EU Statistics on Income and Living Conditions sample microdata. We confirm that income inequality in Lithuania is high compared to the EU average. Our decompositions reveal that the number of employed household members in Lithuania's households affects income inequality more as compared to the EU. It is related to a larger labour income, and self-employment income, in particular, contribution to inequality in Lithuania. Moreover, taxes, social contributions, and transfers reduce income inequality in Lithuania less than in the EU. Specifically, income taxes and social contributions are less progressive while transfers constitute a smaller share of income in Lithuania than in the EU. Income taxes and social contributions are effectively regressive for the self-employed in Lithuania.

ARTICLE HISTORY
Received 19 June 2019
Accepted 8 June 2020

KEYWORDS
Income inequality; factor decomposition; subgroup decomposition; Lithuania; European Union; Kakwani

JEL CLASSIFICATION
D31

1. Introduction

Income inequality in Lithuania has been one of the largest in the EU and is still growing. Specifically, the Gini coefficient of equivalised disposable income, a common measure of inequality, stood at 36.9% in 2018 for Lithuania (Eurostat, 2020). This was the second-largest Gini coefficient among the surveyed EU countries, second to Bulgaria, and exceeded the EU average income inequality by over 6 Gini points. Additionally, income inequality in Lithuania has increased by 5 Gini points since 2012. All this happened in the context of more general concern over rising income inequality within major countries (Atkinson & Piketty, 2010; OECD, 2011, 2015a, 2015b) and increasing empirical evidence

*We express our gratitude to Ignas Goštautas, with whom this project was originally started and for the patience and support of the Macroeconomics and Forecasting Division, the Applied Macroeconomics Research Division, the Center for Excellence in Finance and Economic Research at the Bank of Lithuania, the Faculty of economics and business administration of Vilnius University, the editor of BJE and the two anonymous referees. This paper is based on data from Eurostat, EU Statistics on Income and Living Conditions [2004:2016]. The responsibility for all conclusions drawn from the data lies entirely with the authors. The views expressed in this paper are those of the authors and do not necessarily represent those of the Bank of Lithuania or the Eurosystem.

CONTACT Nerijus Černiauskas nerijus.cerniauskas@evaf.vu.lt, nercern@gmail.com

that income inequality may hinder economic growth (Aghion et al., 1999; Cingano, 2014; Grigoli & Robles, 2017; Ostry & Berg, 2011; Ostry et al., 2014). The size and dynamics of income inequality in Lithuania along with warnings about its possible negative consequences encouraged political and economic debate in Lithuania. There was an interest to re-examine whether income inequality in Lithuania is indeed one of the largest within the EU, what contributes to income inequality, and what policy could be efficient at reducing it. This study focuses on these questions: how confident are we in claiming that Lithuania's income inequality is high, what factors lay behind such inequality and how much can redistribution of direct taxes and public transfers reduce income inequality.

We first analysed the extent to which income inequality is high. Even though the Gini of equivilised income does suggest this, a high Gini is not sufficient for such a claim. Besides the issue of estimating standard errors and testing for different equivalent scales, which can also change the ranking of countries according to income inequality (Buhmann et al., 1988), the Gini index itself is subject to criticism. This is because the Gini index, just like any summary inequality measure, entails social judgements on the undesirability of inequality (Atkinson, 1970). Specifically, the Gini is more sensitive to inequalities in the middle of the distribution rather than the tails. This is not necessarily a desirable property, especially for Lithuania, where the highest level of inequality was found in the tails (IMF, 2016).

For this reason, we employed several statistical tests to examine whether we can claim that equivalised income inequality in Lithuania is one of the highest across the EU. First, we have evaluated the sampling errors to verify that conclusions from the sample data do not contradict the actual situation. Rao et al. (1992) bootstrapped standard errors based on survey design information reconstructed according to Goedemé (2013) and Zardo and Goedemé (2016) allow to estimate the likely biases. Second, we have adjusted household income by alternative equivalence scales. We use the OECD-modified equivalence scale and the square root equivalence scale. Third, we have calculated inequality with other summary measures, thereby explicitly focusing on different segments of the distribution rather than the middle. We have estimated inequality using alternative measures to the Gini index: the Atkinson index and the Generalized entropy index as in Jenkins (2017) with standard inequality preference parameter values. We found that income inequality is statistically larger than the income inequality in other countries regardless of the equivalence scale or the summary measure used. This also strengthens the following analysis which is based on the Gini index.

Next, we have investigated why equivalised income inequality is higher compared to other countries using univariate factor and subgroup decompositions that decompose inequality into parts. These decompositions are purely statistical: they do not incorporate agent responses to any covariate. Nevertheless, these decompositions help identify the households amongst which inequality is acute and suggest which aspects should be looked into deeper.

Factor component decomposition decomposes inequality measure by disaggregating it into mutually exclusive and exhaustive income components, for example, labour and capital income. Two versions of this method are well known: the natural decomposition as in Shorrocks (1982) that focuses on the decomposition of the variance and the Lerman and Yitzhaki (1985) decomposition that is used to decompose the Gini coefficient. We use the latter method, as the Gini is a more conventional index of inequality. This

method was used by, for example, Garner and Terrell (1998) to examine income inequality in Slovakia and Czechia in the early transition period.

Subgroup decomposition decomposes inequality measures within and between mutually exclusive and exhaustive subgroups, for example, inequality between males and females and inequality amongst males and amongst females. There are many ways to decompose subgroups as illustrated in Cowell (2011) and Yitzhaki and Lerman (1991). We apply the Yitzhaki and Lerman (1991) method to decompose the Gini in a way that is closer to the chosen factor decomposition technique.

From the decompositions, we see that labour income inequality is much higher in Lithuania than elsewhere in Europe. Additionally, in line with previous findings (e.g. IMF, 2016), the tax and public transfers system plays less of a redistributive role in Lithuania than in other countries. To understand why, we looked into marginal effects: how does a 1% change in tax and transfers affect income inequality. We also looked into redistributive effects: how much do taxes and transfers reduce inequality according to Kakwani (1977). Finally, we decompose the redistributive effect into the progressivity index and the average rate of tax and public transfers and compare this with that of the EU. This lets us calculate how much can inequality be reduced due to a change in progressivity and average tax and public transfer rates.

Overall, our results suggest that equivalised income inequality in Lithuania is one of the highest in the EU and this finding is robust to various statistical tests. The decompositions reveal large inequalities between and within many groups of households in Lithuania. The largest inequalities lie between the employed and the rest of the population, and this kind of inequality has been rising over time. Inequalities within the unemployed and those working in the agricultural sector are particularly distinct. The factor decomposition shows that labour income, especially self-employment income, is more unequally distributed in Lithuania than elsewhere. Public transfers and taxes seem to reduce income inequality in Lithuania less than in other countries. This is because taxes and public transfers in Lithuania are less progressive and the tax and public transfer rates are lower than in the EU. Income taxes and social contributions are effectively regressive for the self-employed in Lithuania unlike in the EU. It is found that to reduce income inequality in Lithuania via redistribution, the focus should be placed on increasing the progressivity of taxes and average public transfer rates.

The paper is structured as follows: in Section 2, we give definitions of income and describe the data set used throughout the empirical investigation. The other three sections answer three research questions, each using its methodology and provide comments on the results. The final section concludes.

2. Definitions and data on income

We focus on equivalised disposable income inequality. Let us explain each term in more detail. *Income* is defined as a yearly disposable income. To get the disposable income we subtract taxes and social contributions from gross income. We include the social contributions of the employee and employer, as we see both of them affecting the demand for labour. In addition, a new law in 2019 requested employees to pay the majority of employees' social contributions (see SODRA, 2020). Gross income is the sum of market income (labour income with social insurance contributions and capital income) and

transfers (both private and public). In cases when we refer to public transfers to analyse redistribution, we add private transfers to the definition of market income. The unit of observation is a *household*. This assumes that household members share their income and make joint decisions. To adjust for household size, an equivalence scale is used.

Focusing on equivalised income rather than individual income affects the results and this should be briefly justified. Research literature suggests that individuals make economic decisions taking themselves as well as their household members into consideration (see, among others, Vogler & Pahl, 1994). For example, the income of all household members comprises a common budget constraint (Chiappori & Meghir, 2015) thereby influencing each household member's behaviour. Additionally, some transfers are only granted at a household level (e.g. social assistance transfer) making the allocation of this transfer to any specific member artificial. Nevertheless, each household member has their preferences and a typically unequal control of the household's budget with evidence suggesting that decisions taken within a household are rarely joint and more often dominated by a specific household member (Pahl, 1995). Therefore, while it is useful to look at equivalised income inequality to get a first idea of how unequally income is distributed within society, specific questions require looking into inequality within a household (for example, when determining how child transfers should be allocated if mothers are more likely to spend on children rather than fathers).

The data on income and covariates come from the yearly European Union Statistics on Income and Living Conditions (EU-SILC) instrument running since 2004. The data are compiled from a mixture of the survey and administrative sources. Each year around 5 thousand Lithuanian households with around 10 thousand persons over 16 years old who agree to share information on their income are included. The exact number of households and persons recorded in Lithuania and other countries in 2015 is shown in Table 1. Most of these persons provided all information on income, as can be seen from column 5 titled 'Observations'. As all EU member states collect data using the same methodology, we can compare the inequality in Lithuania with that of other EU countries.

While the data is explained by Eurostat (2018c), several features are mentioned here. The survey captures household income and, therefore, certain income components are available for the household rather than the individual level. Therefore, the income of all household members is summed up and allocated to each household member. While most covariates are recorded at the time of the interview, income is recorded for a previous year (the reference year). In this paper, all years represent reference years. While the EU-SILC has a large survey component, some countries make use of register (administrative) data and are referred to as register countries. In 2015, the register countries included Cyprus, Denmark, Finland, Latvia, Lithuania, the Netherlands, Northern Ireland, Norway, Slovenia, Sweden, and Switzerland. Finally, survey weights are used to form conclusions on the population from the sample data. The weights are further adjusted according to Eurostat (2018b): weights of household members over 16 years old are scaled up by distributing weights of those under 16.

3. Is income inequality in Lithuania high?

First, we have examined inequality from the full data sample and then analysed subgroup inequality (inequality between- and within-subgroups) in Lithuania.

Table 1. EU-SILC summary statistics for 2015 income reference year.

Country code	Country name	Households (thousands)	Household members (thousands)	Observations (thousands)	Average income (thousand euro)	Median income (thousand euro)	Gini (percent)
1	2	3	4	5	6	7	8
AT	Austria	6	10.8	10.8	26.1	23.7	27.2
BE	Belgium	5.9	11.1	11.1	24.3	22.3	26.3
BG	Bulgaria	7.3	15.6	15.5	3.9	3.2	37.7
CH	Switzerland	7.8	14.9	14.9	50.9	44.3	29.4
CY	Cyprus	4.2	9.4	9.4	16.9	14	32.1
CZ	Czech Republic	8.5	16.2	16.2	8.8	7.8	25.1
DE	Germany	13.3	23.3	23.1	23.9	21.2	29.8
DK	Denmark	6.3	11.8	11.8	32.1	28.7	27.7
EE	Estonia	6	12.5	12.5	10.1	8.6	32.7
EL	Greece	18.3	38	37.9	8.7	7.5	34.3
ES	Spain	14.2	30.7	30.7	15.8	13.7	34.5
FI	Finland	10.6	20.7	20.7	26.4	23.6	25.4
FR	France	11.5	21.3	21.3	25.3	21.7	29.2
HR	Croatia	7.6	17	17	6.3	5.7	29.8
HU	Hungary	8	15.9	15.8	5.4	4.8	28.2
IE	Ireland	5.2	10.2	10.2	25.5	22.4	29.5
IT	Italy	21.3	41.5	41.5	18.3	16.2	33.1
LT	Lithuania	4.8	9.6	9.6	7	5.6	37.0
LU	Luxembourg	3.8	8.2	8.2	39.8	34.4	31.5
LV	Latvia	6	11.6	11.6	7.5	6.4	34.5
NL	Netherlands	12.7	24.1	24.1	25.4	22.7	26.9
NO	Norway	6.9	13.6	13.6	43	39.6	24.9
PL	Poland	12	27.1	27.1	6.7	5.9	29.8
PT	Portugal	10.6	22.7	22.7	10.6	8.8	33.9
RO	Romania	7.4	15.8	15.7	2.7	2.4	34.8
RS	Serbia	5.6	15.1	15.1	3	2.6	38.6
SE	Sweden	5.8	11.2	11.2	27.3	25.2	27.6
SI	Slovenia	8.6	21.9	21.9	13.2	12.3	24.3
SK	Slovakia	5.7	14.1	14.1	7.4	7	24.3
UK	United Kingdom	9.7	17.8	17.6	24.6	21.1	31.5

Notes: The variables 'Households' and 'Household members' are the unique number of households and household members in the data set. The variable 'Observations' refers to those household members for whom all income data is available. Columns 6–8 refer to the average, median and the Gini coefficient of the population estimate of equivalized household disposable income.

3.1. Inequality

The most popular measure of the level of inequality is the Gini coefficient. The higher the Gini, the greater the level of inequality and it stood at G=0.37 for Lithuania in 2015 (Eurostat, 2020). The Gini is represented, as in Lerman and Yitzhaki (1985), by two times the covariance between income y and the rank of income $F(y)$ divided by average income μ,

$$G = \frac{2Cov(y, F(y))}{\mu}, \tag{1}$$

which describes inequality within the entire population. Since we have sample data only, we modify (1) to include sample weights, as shown in (A1) in Appendix.

Lithuania's Gini coefficient has been compared with the Gini coefficients of all countries that are included in the EU-SILC data set for 2015 in Figure 1 and with the Gini coefficients for a subset of all countries in Table 2. The subset of countries includes the Baltic States, Finland as one of the Scandinavian countries, Germany – which represents the average inequality in the EU and Slovakia, where inequality is the lowest. As in previous studies

(IMF, 2016; Lazutka, 2017), income inequality in Lithuania is one is of the highest according to the EU-SILC. The estimated confidence intervals (Figure 1) and standard errors (Table 2) indicate that this is statistically significant. For example, the Gini in Lithuania is about 7 Gini points higher than in Germany. The latter also happens to be the median in terms of inequality within the whole EU-SILC sample of countries.

Although Table 2 focuses on fewer countries, it provides more statistics on inequality than Figure 1. In Figure 1, household disposable income is equivalised by the OECD-modified equivalence scale. In Table 2, two different scales are used: the OECD-modified scale and the square root equivalence scale. The square root scale increases the Gini for Lithuania by 0.3 points, yet remains with the highest level of income inequality among all countries and 7 Gini points higher than the median country.

Furthermore, in Table 2, the generalized Gini coefficient, $G(v)$ (Yitzhaki, 1983), where parameter v represents inequality aversion. This inequality parameter represents the dissatisfaction expressed towards inequality. With this parameter we can model different

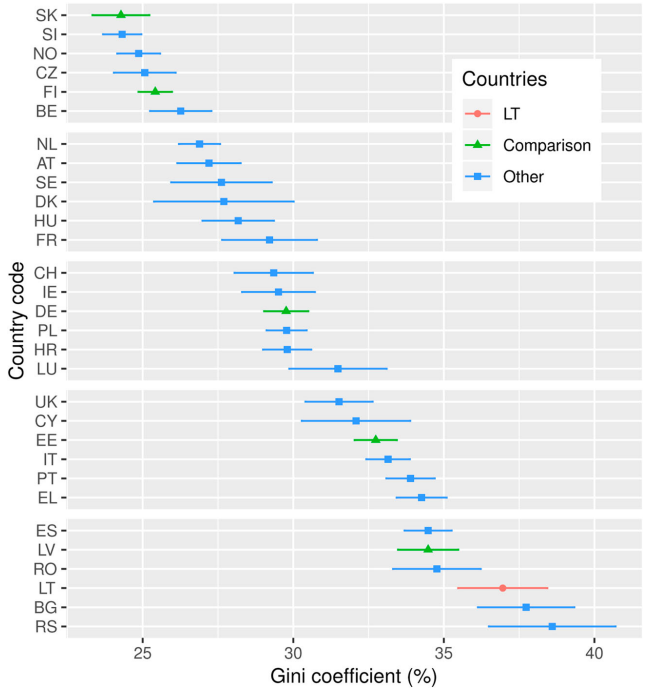


Figure 1. The Gini coefficients of equivalised disposable income in all EU-SILC countries. Household disposable income is equivalised by the OECD-modified scale. Confidence intervals are estimated by using Rao et al. (1992) bootstrap methodology. Information on survey design is provided by Goedemé (2013) and Zardo and Goedemé (2016).

Table 2. Income inequality measures under different equivalence scales.

Country	Equivalence scale	G(2)	G(1.5)	G(4)	GEI(0)	GEI(1)	GEI(2)	Atk(1)	Atk(0.1)
DE	OECD	29.764	19.602	46.279	0.157	0.158	0.220	0.146	0.016
		(0.373)	(0.318)	(0.388)	(0.004)	(0.006)	(0.016)	(0.004)	(0.001)
EE	OECD	32.738	21.096	51.419	0.192	0.171	0.188	0.175	0.017
		(0.358)	(0.256)	(0.463)	(0.006)	(0.004)	(0.005)	(0.005)	(0.000)
FI	OECD	25.416	16.897	40.216	0.112	0.116	0.150	0.106	0.011
		(0.283)	(0.236)	(0.340)	(0.003)	(0.004)	(0.009)	(0.003)	(0.000)
LT	OECD	36.957	24.644	55.797	0.254	0.233	0.306	0.224	0.023
		(0.755)	(0.609)	(0.801)	(0.012)	(0.011)	(0.022)	(0.009)	(0.001)
LV	OECD	34.479	22.756	53.403	0.217	0.202	0.255	0.195	0.020
		(0.511)	(0.432)	(0.563)	(0.007)	(0.008)	(0.019)	(0.006)	(0.001)
Median	OECD	29.764	19.719	46.279	0.158	0.163	0.228	0.146	0.016
		(0.373)	(0.538)	(0.388)	(0.008)	(0.011)	(0.039)	(0.006)	(0.001)
SK	OECD	24.277	15.624	40.310	0.115	0.106	0.136	0.109	0.011
		(0.482)	(0.383)	(0.682)	(0.006)	(0.007)	(0.022)	(0.005)	(0.001)
DE	Sqr. rt	30.224	19.873	47.169	0.163	0.162	0.223	0.150	0.016
		(0.379)	(0.324)	(0.389)	(0.004)	(0.006)	(0.016)	(0.004)	(0.001)
EE	Sqr. rt	33.158	21.305	52.399	0.199	0.175	0.190	0.180	0.017
		(0.354)	(0.253)	(0.451)	(0.006)	(0.004)	(0.005)	(0.005)	(0.000)
FI	Sqr. rt	25.918	17.202	41.213	0.117	0.120	0.155	0.110	0.012
		(0.288)	(0.240)	(0.347)	(0.003)	(0.004)	(0.009)	(0.003)	(0.000)
LT	Sqr. rt	37.383	24.854	56.684	0.261	0.237	0.307	0.230	0.023
		(0.763)	(0.625)	(0.790)	(0.012)	(0.011)	(0.023)	(0.009)	(0.001)
LV	Sqr. rt	35.039	23.063	54.513	0.226	0.207	0.259	0.202	0.021
		(0.521)	(0.447)	(0.553)	(0.007)	(0.008)	(0.021)	(0.006)	(0.001)
Median	Sqr. rt	29.699	19.873	47.169	0.161	0.162	0.223	0.149	0.016
		(0.662)	(0.324)	(0.389)	(0.008)	(0.006)	(0.016)	(0.007)	(0.001)
SK	Sqr. rt	25.000	16.043	41.302	0.120	0.109	0.132	0.113	0.011
		(0.447)	(0.350)	(0.622)	(0.006)	(0.006)	(0.016)	(0.005)	(0.001)

Notes: Table contains inequality measures for the median country (from all EU countries) and selected countries for each equivalence scale. $G(v)$ represents the Gini index with values v=2,1.5,4 of parameter of inequality aversion, $GEI(a)$ stands for the General entrhopy index, and $Atk(b)$ is the Atkinson index, where b=1,0.1 and a=0,1,2 represents the degree of inequality aversion. Bootstrapped standard errors are presented in the parenthesis.

societal preferences. The value v=2 gives the standard Gini, v between 1 and 2 represent lower inequality dissatisfaction and v>2 indicates higher dissatisfaction. The measurement G(1.5) results in lower Gini values in all countries for both equivalence scales (i.e. inequality is not as 'bad'). Additionally, the difference between the Gini in Lithuania and the median country shrinks to 5 Gini points for both scales. Nevertheless, inequality in Lithuania remains significantly the highest out of the sample of six countries. Setting v=4 increases the Gini index, but for Lithuania it remains the highest among the selected countries.

Finally, the Gini is compared with other measures of inequality. Other prominent measures include the Atkinson index (Atk) and General entropy index (GEI), see Das and Parikh (1982), Cowell (2000) and Plat (2012). Both of these measures show that the higher the value, the greater the inequality. Both indexes also feature inequality aversion parameters. In the Atkinson index, a parameter value close to zero means indifference about inequality, while higher values show that people dislike it. In contrast, high GEI parameter values mean that people are indifferent about inequality. In all cases, inequality in Lithuania remained significantly the highest.

3.2. Subgroup inequality

The previous subsection has shown that inequality in Lithuania is large when compared to EU countries. Next, we will consider inequality between and within population subgroups,

for example, between males and females and amongst males and females. Then we will estimate stratification – the extent to which income of one group overlaps income of other groups.

Continuing the discussion started in Section 2, the interpretation of a subgroup may not be straightforward, as we are dealing with equivalised income instead of individual income, but can be explained with the help of an example. Imagine a household composed of one male and one female. Then, comparing household income (i.e. adding up household members' income and allocating the summed household income to each member) implies no income inequality between the male and the female in that household. However, this is only true if all households have the same number of males and females. Some households are consisting of more males, while others have a higher number of female members. If males tend to earn more than females, households with more males will earn higher equivalised household income than equivalised households with more females. In aggregate, this will lead to inequalities between the subgroups. Inequality between this group should be interpreted as 'inequality between male and female-dominated households'. This way, we can combine information on household income and the composition with individual characteristics. Of course, there could be other variables that are also correlated. For example, females tend to live longer and are therefore more likely to be retired and hence receive lower income. However, this approach abstracts from other variables.

The methodology used to estimate inequality between subgroups is similar to the one used by IMF (2016) and is based on Eurostat (2018a). The methodology for estimating inequality within subgroups and stratification are adapted from Yitzhaki and Lerman (1991). Additionally, the technique proposed by Yitzhaki and Lerman (1991) is used to decompose total inequality into *between*, *within* and *stratification* terms to see which of them contributes most to inequality.

*Inequality between subgroups*Inequality between subgroups refers to measured inequality between households grouped under certain criteria. For example, households can be grouped by 'Sex' into two subgroups l=1 and l=2: 'Males' and 'Females'. To estimate between subgroup inequality, we first estimate the weighted average income of a subgroup $\widehat{m}^{(l)}$ and then divide by the average weighted income of all subgroups $\widehat{m}$, see (A3) in Appendix 1, to get an income ratio $\widehat{m}^{(l)}/\widehat{m}$. We then compare the ratio with that of the EU, namely of its member states that joined the EU before 2004 (old EU states), and with those Member States that joined it after 2004 (new EU states). Our method is similar to that used in the IMF (2016), but has several differences: the IMF (2016) analyse weighted income decile ratios while we compare weighted average income ratios. The IMF (2016) compares Lithuania to the EU, while we additionally compare it to new and old EU states to control for the development of countries. Finally, we have more grouping criteria (a total of nine) and estimate standard errors.

Our findings are in line with those of IMF (2016), which also reviews between-subgroup inequality in Lithuania. The IMF (2016) reveals large inequalities between the top and the bottom income deciles, between the employed and the unemployed and non-labour market participants, between the elderly and other age subgroups, as well as between educated and less educated households subgroups, i.e. these ratios are much higher in Lithuania than in the entire EU.

In addition to these findings, the results presented in Table 3 allow adding the following points:

- Differences of ratios are significant between many subgroups in Lithuania. The subgroups include those grouped according to the IMF (2016) criteria (activity status, age bracket, number of dependants, education) as well as ratios in other subgroups. For example,

Table 3. Ratios of average subgroup incomes in 2015.

Grouping	Subgroup	EU	EU new	EU old	LT
activity status	employed	113.2 (0.5) [25.5]	115.4 (0.9) [9.8]	112.6 (0.6) [15.5]	123.2 (1.3) [5.2]
activity status	unemployed	61.0 (1.3) [3.3]	59.5 (1.9) [1.2]	61.3 (1.6) [2.0]	54.0 (3.6) [0.6]
activity status	retired	97.3 (1.0) [9.5]	95.2 (1.3) [4.2]	97.9 (1.2) [5.6]	71.1 (1.7) [2.0]
activity status	study	85.5 (1.6) [3.5]	88.4 (2.4) [1.3]	84.8 (1.8) [2.2]	86.8 (3.2) [0.7]
activity status	other	77.4 (1.3) [5.7]	70.5 (1.9) [2.2]	79.2 (1.6) [3.5]	68.9 (3.1) [1.0]
nr working	0	78.3 (0.9) [12.8]	75.1 (1.2) [4.6]	79.0 (1.0) [8.0]	53.3 (1.6) [2.4]
nr working	1	93.4 (0.9) [15.1]	93.6 (1.6) [5.8]	93.3 (1.0) [9.2]	94.1 (2.6) [3.1]
nr working	2	119.3 (0.9) [16.7]	117.8 (1.4) [6.7]	119.7 (1.0) [10.1]	131.4 (2.8) [3.3]
nr working	3	116.9 (2.3) [2.5]	115.6 (4.1) [1.5]	117.5 (2.7) [1.3]	130.5 (4.8) [0.6]
nr working	4	124.3 (4.2) [0.6]	124.0 (5.3) [0.4]	124.5 (6.0) [0.3]	124.9 (8.6) [0.2]
main income	employment	108.7 (0.6) [24.1]	112.2 (0.9) [11.7]	107.4 (0.8) [13.6]	111.5 (1.4) [6.1]
main income	self-employment	106.3 (3.2) [3.4]	84.3 (3.2) [1.7]	114.8 (4.2) [1.9]	174.9 (10.7) [0.7]
main income	other	88.7 (0.6) [20.4]	79.3 (1.6) [5.6]	90.4 (0.7) [13.4]	57.6 (1.6) [2.8]
occupation	basic level	74.6 (1.3) [4.5]	78.1 (4.1) [1.9]	73.7 (1.2) [2.7]	69.6 (2.5) [1.4]
occupation	mid-level	90.0 (0.6) [21.0]	91.9 (0.8) [9.6]	89.3 (0.7) [12.2]	88.7 (1.7) [4.1]
occupation	technicians, associates	112.9 (1.3) [6.0]	118.7 (2.1) [1.6]	111.9 (1.4) [3.9]	117.8 (4.6) [0.6]
occupation	professionals	139.1 (1.8) [7.3]	145.8 (2.6) [2.4]	137.7 (2.1) [4.6]	129.1 (3.1) [1.9]
occupation	managers	137.9 (3.3) [2.3]	153.0 (5.5) [0.8]	134.7 (3.8) [1.5]	162.7 (11.6) [0.6]
sector	agriculture	74.2 (2.6) [1.0]	66.4 (2.5) [1.0]	83.4 (4.7) [0.3]	99.5 (6.7) [0.3]
sector	industry	115.6 (1.9) [3.4]	116.2 (2.3) [2.0]	115.3 (2.5) [1.8]	113.9 (3.8) [0.8]
sector	IT, finance, RE, admin	134.4 (2.7) [3.4]	158.2 (7.9) [0.9]	130.6 (2.8) [2.2]	157.2 (10.9) [0.6]
sector	public admin, education, health	118.0 (1.3) [5.8]	127.1 (2.3) [1.7]	116.4 (1.5) [3.7]	124.6 (3.2) [1.1]
sector	other services	100.2 (1.1) [7.6]	109.5 (1.8) [3.0]	97.8 (1.3) [4.6]	119.3 (2.8) [1.8]
age bracket	under 19	83.8 (2.0) [1.8]	82.4 (3.5) [0.7]	84.2 (2.3) [1.1]	83.6 (4.1) [0.4]
age bracket	19–29	91.3 (1.3) [6.9]	95.5 (1.9) [3.0]	90.0 (1.5) [4.1]	104.2 (3.8) [1.7]
age bracket	30–64	103.9 (0.4) [30.1]	104.0 (0.6) [12.0]	103.9 (0.5) [18.1]	107.7 (1.2) [5.7]
age bracket	65+	96.9 (1.1) [9.1]	93.5 (1.4) [3.3]	97.7 (1.3) [5.6]	76.2 (1.9) [1.9]
dependents	0	105.2 (0.6) [24.5]	105.4 (1.1) [8.9]	105.2 (0.7) [15.1]	101.0 (1.8) [4.8]
dependents	1	102.6 (1.3) [9.2]	107.0 (1.9) [4.3]	101.1 (1.6) [5.3]	106.7 (3.2) [2.1]
dependents	2	94.2 (1.1) [9.7]	93.2 (1.8) [3.9]	94.5 (1.3) [5.8]	102.4 (4.8) [2.0]
dependents	3	82.7 (2.1) [3.2]	80.7 (3.7) [1.2]	83.2 (2.5) [2.0]	73.9 (6.9) [0.5]
dependents	4	73.0 (4.5) [0.8]	60.9 (6.2) [0.3]	75.9 (5.3) [0.5]	55.6 (5.5) [0.1]
dependents	5	50.9 (4.5) [0.4]	40.0 (3.5) [0.3]	58.7 (6.1) [0.2]	47.1 (11.8) [0.1]
education	up to secondary	87.8 (0.4) [32.4]	87.9 (0.6) [14.1]	87.8 (0.5) [19.0]	78.2 (1.5) [4.6]
education	post-secondary	103.0 (2.5) [1.5]	107.3 (3.9) [0.6]	101.7 (3.1) [0.9]	92.7 (2.3) [1.9]
education	tertiary education	129.6 (1.0) [13.4]	142.1 (2.0) [3.9]	127.3 (1.2) [8.6]	138.1 (2.8) [3.0]
sex	male	102.2 (0.5) [22.7]	102.8 (0.9) [9.0]	102.1 (0.7) [13.7]	104.5 (0.8) [4.3]
sex	female	98.0 (0.5) [25.2]	97.4 (0.8) [10.0]	98.1 (0.6) [15.2]	96.4 (0.6) [5.3]

Notes: Ratios are defined as weighted average income of a subgroup divided by weighted average income of all subgroups within that grouping. Bootstrapped standard errors are in the parenthesis. Number of observations in thousands in brackets.

we split households based on the main income source. Those who receive largely self-employment income tend, on average, to have more disposable income than those who work as employees or others – a trend not observed in the EU as a whole. Significant inequality also exists between subgroups grouped by the number of people working in the household (nr working) and the sector where one works (sector).

- Ratios between the majority of the nine subgroups are also significantly different from the ratios between their EU counterparts. Besides the subgroups in the IMF (2016) (those grouped by activity status, age bracket, education), the self-employed in Lithuania on average earn proportionally more than their EU counterparts. Additionally, those who work in the information technologies, finance, real estate, and administration sector (IT, finance, RE, admin) earn, on average, relatively more income in Lithuania than one would in the EU.
- There are some groups between which inequality in Lithuania is smaller as compared to the EU. For example, those working within the agricultural sector are relatively better off in Lithuania compared to the EU. Additionally, income ratios in Lithuania are more similar to those in the new EU states. In particular, those who are under 19 years old have very similar relative incomes both in Lithuania and in the new EU states.

In general, ratios between subgroups are largely persistent and slightly widening since 2010. This can be seen in Table 4 which shows the ratio dynamics in Lithuania. For example, there was a slowly widening gap between the employed and the retired. This could be explained by rising market incomes due to a recovering economy that benefited the employed while statutory pensions, the main source of income for the retired, did not increase in the period due to budget consolidation (Černiauskas et al., 2020). Once the recovery began, wages in the private sector started rising, especially IT, finance, RE, admin sector, while the government started raising public sector wages (Public admin, education, health) much later. This could also explain the rising ratio difference between the two sectors.

*Inequality within subgroups*Inequality exists within subgroups in Lithuania. A common way to measure it is to calculate inequality measures for subgroup income as is done for total income (see $\widehat{G}_I$ in Formula (A4) in Appendix 1). We have calculated the Gini coefficients for Lithuania's subgroups and compared them with the Gini coefficients of the EU, new and old EU states in Table 5.

- Most of the within-subgroup Gini coefficients examined in Table 5 are higher in Lithuania than in the EU. Especially large subgroup inequality exists among those working in the agricultural sector and the unemployed.
- The above-mentioned within-group inequalities are much higher in Lithuania than in the EU. Additionally, households, where the main source of income is self-employment income, are also unequal among themselves, even though similar inequality within subgroups exists in new EU states. The Gini of households with many children is relatively small and we know from the between analysis that these households earn a much lower income.

Over time, inequality within subgroups increased in many subgroups. Table 6 shows that the rise has been especially strong since 2010. In particular, the Gini coefficient of the unemployed rose from 39.8 in 2004 to 47.8 in 2015. This may be in part due to

Table 4. Ratios of average subgroup incomes in Lithunia.

Grouping	Subgroup	2006	2010	2015
activity status	employed	120.2 (0.9) [6.2]	122.1 (1.1) [5.4]	123.2 (1.3) [5.2]
activity status	unemployed	51.1 (2.7) [0.4]	57.7 (2.3) [1.1]	54.0 (3.6) [0.6]
activity status	retired	70.4 (1.2) [2.0]	86.4 (1.3) [2.3]	71.1 (1.7) [2.0]
activity status	study	89.1 (2.2) [1.1]	86.2 (2.3) [1.1]	86.8 (3.2) [0.7]
activity status	other	70.7 (2.2) [1.1]	74.3 (1.9) [1.0]	68.9 (3.1) [1.0]
nr working	0	51.4 (1.2) [2.1]	66.2 (1.3) [3.1]	53.3 (1.6) [2.4]
nr working	1	86.5 (1.8) [3.0]	89.7 (1.8) [3.4]	94.1 (2.6) [3.1]
nr working	2	122.0 (1.6) [4.7]	128.9 (2.1) [3.6]	131.4 (2.8) [3.3]
nr working	3	137.5 (5.6) [0.8]	144.7 (6.8) [0.7]	130.5 (4.8) [0.6]
nr working	4	138.7 (11.2) [0.2]	122.2 (11.5) [0.2]	124.9 (8.6) [0.2]
main income	employment	114.9 (0.9) [7.5]	115.3 (1.1) [6.7]	111.5 (1.4) [6.1]
main income	self-employment	106.1 (5.6) [0.6]	127.1 (10.8) [0.4]	174.9 (10.7) [0.7]
main income	other	59.0 (1.4) [2.8]	71.0 (1.2) [3.9]	57.6 (1.6) [2.8]
occupation	basic level	72.1 (1.8) [1.6]	72.6 (1.7) [1.5]	69.6 (2.5) [1.4]
occupation	mid-level	90.0 (1.1) [5.1]	87.6 (1.1) [4.8]	88.7 (1.7) [4.1]
occupation	technicians, associates	120.0 (4.3) [0.7]	119.3 (3.8) [0.7]	117.8 (4.6) [0.6]
occupation	professionals	146.4 (3.4) [1.5]	138.9 (2.8) [1.9]	129.1 (3.1) [1.9]
occupation	managers	149.9 (4.6) [0.7]	146.0 (4.6) [0.7]	162.7 (11.6) [0.6]
sector	agriculture	85.1 (5.2) [0.5]	90.9 (5.9) [0.4]	99.5 (6.7) [0.3]
sector	industry	109.9 (2.6) [1.1]	113.8 (3.9) [0.9]	113.9 (3.8) [0.8]
sector	IT, finance, RE, admin	168.2 (8.0) [0.2]	137.2 (4.9) [0.5]	157.2 (10.9) [0.6]
sector	public admin, education, health	131.2 (2.7) [1.4]	138.1 (3.1) [1.3]	124.6 (3.2) [1.1]
sector	other services	117.0 (1.9) [2.2]	111.6 (2.2) [1.8]	119.3 (2.8) [1.8]
age bracket	under 19	82.0 (2.6) [0.6]	80.3 (2.8) [0.6]	83.6 (4.1) [0.4]
age bracket	19–29	111.2 (2.9) [2.2]	101.7 (2.6) [2.0]	104.2 (3.8) [1.7]
age bracket	30–64	105.1 (0.9) [6.4]	104.3 (0.8) [6.4]	107.7 (1.2) [5.7]
age bracket	65+	73.6 (1.4) [1.8]	90.4 (1.6) [2.0]	76.2 (1.9) [1.9]
dependents	0	104.0 (1.4) [5.3]	102.7 (1.4) [4.9]	101.0 (1.8) [4.8]
dependents	1	106.0 (2.3) [2.9]	102.9 (2.2) [2.8]	106.7 (3.2) [2.1]
dependents	2	95.0 (2.8) [2.0]	99.0 (3.3) [2.3]	102.4 (4.8) [2.0]
dependents	3	72.5 (4.6) [0.5]	81.8 (6.5) [0.7]	73.9 (6.9) [0.5]
dependents	4	45.7 (10.5) [0.1]	66.8 (10.1) [0.2]	55.6 (5.5) [0.1]
dependents	5	35.9 (6.4) [0.1]	95.5 (16.1) [0.1]	47.1 (11.8) [0.1]
education	up to secondary	81.5 (0.9) [5.8]	81.7 (1.0) [6.0]	78.2 (1.5) [4.6]
education	post-secondary	95.9 (1.6) [2.4]	96.1 (2.0) [2.1]	92.7 (2.3) [1.9]
education	tertiary education	148.5 (2.4) [2.5]	140.6 (2.2) [2.9]	138.1 (2.8) [3.0]
sex	male	104.9 (0.6) [4.9]	102.9 (0.7) [4.9]	104.5 (0.8) [4.3]
sex	female	96.0 (0.5) [6.0]	97.6 (0.5) [6.1]	96.4 (0.6) [5.3]

Notes: Ratios are defined as weighted average income of a subgroup divided by weighted average income of all subgroups within that grouping. Bootstrapped standard errors are in the parenthesis. Number of observations in thousands in brackets.

unequal economic recovery, where some of the unemployed were able to find some income sources, while others did not. Unemployment has risen substantially since the crisis and there have been many unemployment transfers handed out. However, these transfers were stopped to those who were unemployed for a longer time. Additionally, as the economy recovered, it became easier for the unemployed to be in employment for at least several months during the year. Similarly, there was a rise in inequality among those who are neither employed, unemployed, retired, or students (largely disabled). Additionally, there has been a rise in inequality among those who are over 65 and, to a lesser extent, those aged 30–64. Inequality increased within all the different education levels and within all occupations (managers in particular). Inequality increased in the agricultural sector as well as in the IT, finance, real estate and administration sectors (IT, finance, RE, admin).

*Stratification between subgroups*Inequality is linked to stratification. Stratification measures whether the income of each member of a subgroup differs compared to

Table 5. The Gini coefficient of income of subgroups in 2015.

Grouping	Subgroup	EU	new EU states	old EU states	LT
activity status	employed	28.6 (0.4) [25.5]	29.7 (0.7) [9.8]	28.3 (0.4) [15.5]	33.0 (0.9) [5.2]
activity status	unemployed	35.1 (0.9) [3.3]	36.9 (1.3) [1.2]	34.6 (1.1) [2.0]	47.8 (2.2) [0.6]
activity status	retired	27.4 (0.5) [9.5]	25.9 (0.5) [4.2]	27.8 (0.7) [5.6]	29.6 (1.0) [2.0]
activity status	student	32.2 (0.8) [3.5]	30.1 (1.0) [1.3]	32.7 (0.9) [2.2]	35.1 (1.6) [0.7]
activity status	other	32.5 (0.8) [5.7]	31.3 (1.1) [2.2]	32.7 (1.0) [3.5]	38.0 (1.5) [1.0]
nr working	0	32.2 (0.6) [12.8]	31.1 (0.9) [4.6]	32.4 (0.6) [8.0]	31.5 (1.5) [2.4]
nr working	1	31.6 (0.6) [15.1]	32.2 (1.2) [5.8]	31.4 (0.6) [9.2]	36.4 (1.1) [3.1]
nr working	2	26.5 (0.4) [16.7]	27.2 (0.6) [6.7]	26.3 (0.5) [10.1]	31.6 (1.3) [3.3]
nr working	3	25.1 (0.9) [2.5]	27.6 (1.7) [1.5]	23.8 (1.0) [1.3]	21.0 (1.4) [0.6]
nr working	4	24.6 (1.3) [0.6]	25.2 (1.9) [0.4]	24.2 (1.8) [0.3]	17.7 (3.3) [0.2]
main income	employment	27.2 (0.3) [24.1]	27.0 (0.4) [11.7]	27.3 (0.4) [13.6]	31.3 (0.9) [6.1]
main income	self-employment	42.6 (1.3) [3.4]	39.1 (1.6) [1.7]	43.0 (1.5) [1.9]	39.7 (2.2) [0.7]
main income	other	31.9 (0.4) [20.4]	33.9 (1.3) [5.6]	31.5 (0.5) [13.4]	33.0 (1.3) [2.8]
occupation	basic level	28.7 (1.0) [4.5]	33.1 (3.5) [1.9]	27.4 (0.7) [2.7]	35.6 (1.4) [1.4]
occupation	mid-level	27.5 (0.3) [21.0]	27.5 (0.5) [9.6]	27.5 (0.4) [12.2]	34.8 (0.9) [4.1]
occupation	technicians, associates	25.2 (0.6) [6.0]	24.4 (0.7) [1.6]	25.3 (0.7) [3.9]	31.0 (1.5) [0.6]
occupation	professionals	29.2 (0.8) [7.3]	25.9 (0.8) [2.4]	29.8 (0.9) [4.6]	31.0 (0.9) [1.9]
occupation	managers	31.9 (1.1) [2.3]	30.9 (1.7) [0.8]	31.9 (1.3) [1.5]	39.6 (2.5) [0.6]
sector	agriculture	35.2 (1.7) [1.0]	35.2 (1.5) [1.0]	34.0 (2.9) [0.3]	44.5 (2.4) [0.3]
sector	industry	27.1 (0.9) [3.4]	26.4 (0.8) [2.0]	27.4 (1.2) [1.8]	30.3 (1.4) [0.8]
sector	IT, finance, RE, admin	30.8 (1.0) [3.4]	31.5 (3.2) [0.9]	30.5 (1.0) [2.2]	35.1 (2.9) [0.6]
sector	public admin, education, health	25.7 (0.5) [5.8]	25.0 (0.8) [1.7]	25.7 (0.6) [3.7]	28.7 (1.0) [1.1]
sector	other services	27.9 (0.6) [7.6]	27.0 (0.8) [3.0]	27.9 (0.7) [4.6]	32.7 (1.2) [1.8]
age bracket	under 19	31.1 (1.1) [1.8]	32.8 (1.5) [0.7]	30.6 (1.3) [1.1]	34.8 (1.9) [0.4]
age bracket	19–29	30.8 (0.7) [6.9]	31.0 (1.0) [3.0]	30.6 (0.9) [4.1]	35.8 (1.4) [1.7]
age bracket	30–64	31.1 (0.3) [30.1]	31.7 (0.6) [12.0]	31.0 (0.4) [18.1]	37.4 (0.9) [5.7]
age bracket	65+	28.4 (0.7) [9.1]	26.7 (0.5) [3.3]	28.8 (0.8) [5.6]	30.8 (1.1) [1.9]
dependents	0	31.1 (0.4) [24.5]	30.8 (0.7) [8.9]	31.1 (0.4) [15.1]	38.4 (0.8) [4.8]
dependents	1	29.9 (0.6) [9.2]	29.2 (0.8) [4.3]	30.1 (0.8) [5.3]	30.8 (1.3) [2.1]
dependents	2	28.4 (0.6) [9.7]	29.2 (1.0) [3.9]	28.1 (0.7) [5.8]	37.9 (2.3) [2.0]
dependents	3	32.2 (1.1) [3.2]	32.2 (2.1) [1.2]	32.2 (1.3) [2.0]	36.2 (3.0) [0.5]
dependents	4	30.5 (2.6) [0.8]	34.6 (3.0) [0.3]	29.1 (3.0) [0.5]	20.7 (3.9) [0.1]
dependents	5	25.8 (4.0) [0.4]	21.9 (3.9) [0.3]	23.7 (5.7) [0.2]	24.7 (9.9) [0.1]
education	up to secondary	29.0 (0.3) [32.4]	29.7 (0.6) [14.1]	28.8 (0.4) [19.0]	35.0 (0.8) [4.6]
education	post-secondary	27.8 (1.0) [1.5]	28.2 (1.3) [0.6]	27.6 (1.2) [0.9]	34.7 (1.2) [1.9]
education	tertiary education	29.6 (0.5) [13.4]	27.1 (0.6) [3.9]	30.0 (0.6) [8.6]	33.1 (1.2) [3.0]
sex	male	31.0 (0.4) [22.7]	31.1 (0.6) [9.0]	30.9 (0.4) [13.7]	37.0 (0.9) [4.3]
sex	female	30.6 (0.4) [25.2]	30.9 (0.7) [10.0]	30.5 (0.5) [15.2]	36.8 (0.8) [5.3]

Notes: Bootstrapped standard errors are provided in the parenthesis. Number of observations are shown in thousands in brackets.

the income of every member of all other subgroups. We use the methodology proposed by Yitzhaki and Lerman (1991), which measures stratification on a scale from −100 to 100. Value 100 indicates high stratification: all members of a subgroup have income that is different from members of other subgroups. Value 0 indicates no stratification – there is a perfect income overlap between the subgroups. Negative numbers indicate that the subgroup should actually be multiple subgroups, i.e. income of some subgroup members is much higher than that of members of other subgroups, however, some members also have much lower income than members of other subgroups. The estimates of measures of stratification in Table 7 allow us to make two more insights:

- Several subgroups in Lithuania are stratified. Families with more dependants are detached in terms of income from other subgroups and the difference is stark when compared to the EU. Households who are employed or have more employed

Table 6. Gini of subgroup incomes in Lithuania.

Grouping	Subgroup	2006	2010	2015
activity status	employed	30.9 (0.6) [6.2]	30.6 (0.7) [5.4]	33.0 (0.9) [5.2]
activity status	unemployed	42.1 (2.0) [0.4]	39.8 (1.8) [1.1]	47.8 (2.2) [0.6]
activity status	retired	25.1 (0.6) [2.0]	24.2 (0.7) [2.3]	29.6 (1.0) [2.0]
activity status	study	32.9 (1.3) [1.1]	34.0 (1.0) [1.1]	35.1 (1.6) [0.7]
activity status	other	34.8 (1.2) [1.1]	29.7 (1.2) [1.0]	38.0 (1.5) [1.0]
nr working	0	26.7 (1.2) [2.1]	29.8 (1.0) [3.1]	31.5 (1.5) [2.4]
nr working	1	31.9 (0.8) [3.0]	32.5 (1.0) [3.4]	36.4 (1.1) [3.1]
nr working	2	29.5 (0.8) [4.7]	27.8 (0.8) [3.6]	31.6 (1.3) [3.3]
nr working	3	24.7 (1.6) [0.8]	25.4 (1.9) [0.7]	21.0 (1.4) [0.6]
nr working	4	24.5 (3.8) [0.2]	19.4 (3.3) [0.2]	17.7 (3.3) [0.2]
main income	employment	30.4 (0.6) [7.5]	30.0 (0.6) [6.7]	31.3 (0.9) [6.1]
main income	self-employment	37.2 (2.1) [0.6]	44.3 (2.3) [0.4]	39.7 (2.2) [0.7]
main income	other	29.1 (1.2) [2.8]	29.3 (1.0) [3.9]	33.0 (1.3) [2.8]
occupation	basic level	31.0 (1.2) [1.6]	29.8 (0.9) [1.5]	35.6 (1.4) [1.4]
occupation	mid-level	29.9 (0.6) [5.1]	30.2 (0.7) [4.8]	34.8 (0.9) [4.1]
occupation	technicians, associates	30.5 (1.8) [0.7]	28.3 (1.4) [0.7]	31.0 (1.5) [0.6]
occupation	professionals	31.7 (0.9) [1.5]	28.8 (0.9) [1.9]	31.0 (0.9) [1.9]
occupation	managers	31.9 (1.3) [0.7]	32.1 (1.5) [0.7]	39.6 (2.5) [0.6]
sector	agriculture	41.3 (2.0) [0.5]	37.2 (2.4) [0.4]	44.5 (2.4) [0.3]
sector	industry	28.3 (1.1) [1.1]	31.1 (1.6) [0.9]	30.3 (1.4) [0.8]
sector	IT, finance, RE, admin	29.8 (1.8) [0.2]	28.3 (1.8) [0.5]	35.1 (2.9) [0.6]
sector	public admin, education, health	29.4 (1.0) [1.4]	27.5 (0.9) [1.3]	28.7 (1.0) [1.1]
sector	other services	30.4 (1.0) [2.2]	31.5 (0.9) [1.8]	32.7 (1.2) [1.8]
age bracket	under 19	33.5 (1.5) [0.6]	33.1 (1.5) [0.6]	34.8 (1.9) [0.4]
age bracket	19–29	34.3 (1.3) [2.2]	32.4 (1.0) [2.0]	35.8 (1.4) [1.7]
age bracket	30–64	33.2 (0.6) [6.4]	34.7 (0.7) [6.4]	37.4 (0.9) [5.7]
age bracket	65+	27.5 (0.9) [1.8]	24.7 (0.9) [2.0]	30.8 (1.1) [1.9]
dependents	0	35.6 (0.7) [5.3]	33.2 (0.7) [4.9]	38.4 (0.8) [4.8]
dependents	1	30.7 (0.9) [2.9]	31.2 (1.0) [2.8]	30.8 (1.3) [2.1]
dependents	2	29.1 (1.4) [2.0]	33.5 (1.4) [2.3]	37.9 (2.3) [2.0]
dependents	3	31.0 (3.1) [0.5]	36.2 (2.5) [0.7]	36.2 (3.0) [0.5]
dependents	4	35.2 (9.1) [0.1]	29.2 (5.7) [0.2]	20.7 (3.9) [0.1]
dependents	5	32.1 (7.3) [0.1]	8.5 (8.3) [0.1]	24.7 (9.9) [0.1]
education	up to secondary	31.1 (0.7) [5.8]	30.6 (0.6) [6.0]	35.0 (0.8) [4.6]
education	post-secondary	29.6 (0.7) [2.4]	31.5 (1.0) [2.1]	34.7 (1.2) [1.9]
education	tertiary education	30.7 (0.8) [2.5]	29.5 (0.9) [2.9]	33.1 (1.2) [3.0]
sex	male	33.3 (0.6) [4.9]	33.3 (0.7) [4.9]	37.0 (0.9) [4.3]
sex	female	34.0 (0.6) [6.0]	32.7 (0.6) [6.1]	36.8 (0.8) [5.3]

Notes: Bootstrapped standard errors are in the brackets. Number of observations in thousands in brackets.

members are stratified from the unemployed and those who do not participate in the labour market. Income stratification of these subgroups is greater in Lithuania than in the EU. Additionally, several subgroups are stratified in Lithuania to a similar extent as they are stratified in new EU states: subgroups characterized by occupation, education, and age bracket. This could signal that Lithuania, like in new EU states, is facing more labour market imbalances, where the demand for highly educated professionals is especially high, while redistribution channels are too weak to compensate for the income of those out of labour force (e.g. elderly).

- There are several subgroups that should form several smaller subgroups in Lithuania. The unemployed, for example, have a stratification value of −9.9, meaning that some unemployed are relatively well off, while others are not. This could reflect that some of the unemployed are still getting unemployment transfers, can take on part-time work, or are simply living in a high-income household, while others do not. Similar tendencies also exist in the agricultural sector, with some being much better off than others.

Table 7. Stratification of subgroup income in 2015.

Grouping	Subgroup	EU	new EU states	old EU states	LT
activity status	employed	17.7 (0.8) [25.5]	14.2 (1.5) [9.8]	18.6 (1.0) [15.5]	32.6 (2.0) [5.2]
activity status	unemployed	6.3 (2.3) [3.3]	7.3 (3.1) [1.2]	5.9 (2.8) [2.0]	−10.0 (5.8) [0.6]
activity status	retired	6.6 (0.8) [9.5]	10.4 (1.2) [4.2]	5.7 (1.0) [5.6]	11.7 (3.0) [2.0]
activity status	student	−3.5 (1.2) [3.5]	0.8 (2.0) [1.3]	−4.6 (1.5) [2.2]	1.3 (2.3) [0.7]
activity status	other	−1.6 (1.7) [5.7]	7.6 (2.2) [2.2]	−3.4 (1.9) [3.5]	−3.8 (3.3) [1.0]
nr working	0	−1.0 (1.3) [12.8]	6.1 (1.7) [4.6]	−2.7 (1.5) [8.0]	25.3 (6.6) [2.4]
nr working	1	−2.1 (0.8) [15.1]	−1.5 (1.5) [5.8]	−2.3 (0.9) [9.2]	−0.4 (2.2) [3.1]
nr working	2	23.2 (0.9) [16.7]	17.4 (1.7) [6.7]	24.9 (1.1) [10.1]	34.8 (2.6) [3.3]
nr working	3	15.8 (2.7) [2.5]	8.4 (5.6) [1.5]	19.8 (2.5) [1.3]	42.8 (3.7) [0.6]
nr working	4	18.0 (3.4) [0.6]	15.6 (5.2) [0.4]	19.5 (4.6) [0.3]	46.2 (9.1) [0.2]
main income	employment	16.3 (0.8) [24.1]	25.5 (1.6) [11.7]	14.4 (0.9) [13.6]	33.7 (2.4) [6.1]
main income	self-employment	−13.9 (1.3) [3.4]	−12.4 (2.2) [1.7]	−11.9 (1.8) [1.9]	14.8 (8.2) [0.7]
main income	other	−4.5 (0.8) [20.4]	−1.4 (2.2) [5.6]	−4.8 (0.9) [13.4]	18.4 (4.7) [2.8]
occupation	basic level	6.9 (2.3) [4.5]	1.1 (6.0) [1.9]	8.9 (1.9) [2.7]	2.0 (3.4) [1.4]
occupation	mid-level	7.0 (0.9) [21.0]	10.1 (1.4) [9.6]	6.4 (1.0) [12.2]	2.5 (1.8) [4.1]
occupation	technicians, associates	13.5 (1.1) [6.0]	14.9 (1.7) [1.6]	13.4 (1.2) [3.9]	13.9 (2.8) [0.6]
occupation	professionals	20.0 (1.6) [7.3]	27.0 (2.1) [2.4]	18.9 (1.9) [4.6]	22.3 (2.2) [1.9]
occupation	managers	10.4 (2.6) [2.3]	15.9 (4.2) [0.8]	9.4 (3.0) [1.5]	17.2 (4.3) [0.6]
sector	agriculture	−0.6 (3.8) [1.0]	9.9 (3.9) [1.0]	−5.1 (5.3) [0.3]	−19.7 (3.7) [0.3]
sector	industry	4.6 (1.2) [3.4]	7.6 (1.7) [2.0]	3.4 (1.6) [1.8]	3.7 (2.5) [0.8]
sector	IT, finance, RE, admin	5.5 (1.7) [3.4]	16.4 (3.7) [0.9]	4.1 (1.8) [2.2]	6.8 (3.9) [0.6]
sector	public admin, education, health	8.8 (1.2) [5.8]	12.6 (2.0) [1.7]	8.2 (1.4) [3.7]	9.8 (2.4) [1.1]
sector	other services	0.6 (1.1) [7.6]	5.8 (1.7) [3.0]	−0.2 (1.4) [4.6]	0.2 (2.1) [1.8]
age bracket	under 19	−1.0 (1.8) [1.8]	−3.6 (3.1) [0.7]	−0.3 (2.2) [1.1]	1.0 (3.0) [0.4]
age bracket	19–29	−1.4 (0.9) [6.9]	−1.1 (1.5) [3.0]	−1.4 (1.2) [4.1]	−0.6 (2.6) [1.7]
age bracket	30–64	0.0 (0.7) [30.1]	−1.2 (1.2) [12.0]	0.4 (0.8) [18.1]	2.7 (1.6) [5.7]
age bracket	65+	4.8 (0.9) [9.1]	7.9 (1.3) [3.3]	4.1 (1.0) [5.6]	7.9 (2.6) [1.9]
dependents	0	2.3 (0.7) [24.5]	4.3 (1.2) [8.9]	1.7 (0.9) [15.1]	−4.5 (2.0) [4.8]
dependents	1	1.7 (1.0) [9.2]	3.5 (1.5) [4.3]	1.3 (1.2) [5.3]	12.5 (2.9) [2.1]
dependents	2	3.8 (1.0) [9.7]	3.1 (1.8) [3.9]	4.0 (1.1) [5.8]	1.2 (2.8) [2.0]
dependents	3	−3.1 (2.0) [3.2]	−1.1 (4.2) [1.2]	−3.6 (2.2) [2.0]	−2.5 (7.0) [0.5]
dependents	4	1.3 (6.5) [0.8]	3.3 (7.3) [0.3]	0.7 (7.4) [0.5]	43.4 (7.1) [0.1]
dependents	5	26.1 (19.4) [0.4]	48.6 (9.1) [0.3]	16.9 (25.3) [0.2]	12.0 (37.0) [0.1]
education	up to secondary	7.4 (1.0) [32.4]	12.3 (2.0) [14.1]	6.5 (1.1) [19.0]	2.7 (2.1) [4.6]
education	post-secondary	2.9 (2.0) [1.5]	3.0 (2.6) [0.6]	3.1 (2.4) [0.9]	3.6 (1.9) [1.9]
education	tertiary education	17.2 (1.3) [13.4]	26.3 (1.9) [3.9]	15.8 (1.4) [8.6]	27.7 (2.4) [3.0]
sex	male	1.0 (0.6) [22.7]	0.6 (1.0) [9.0]	1.1 (0.7) [13.7]	2.4 (1.1) [4.3]
sex	female	−0.8 (0.6) [25.2]	−0.2 (1.0) [10.0]	−0.9 (0.7) [15.2]	−1.8 (1.0) [5.3]

Notes: Bootstrapped standard errors are provided in the parenthesis. Number of observations are shown in thousands in brackets.

Stratification between groups has been increasing, especially since 2010 (see Table 8). This is particularly apparent when considering activity status: the stratification coefficient of those employed rose from 17.8% in 2010 to 32.6% in 2015. However, this could be largely attributed to a market correction, as the stratification coefficient was around 28.7% before the crisis.

*Subgroup decomposition*We have analysed between- and within-subgroup inequality and stratification separately. Now, we will identify how much each of the terms contributes to the Gini of disposable income in Lithuania and compare this to the EU, new and old EU

Table 8. Stratification of subgroup incomes in Lithuania.

Grouping	Subgroup	2006	2010	2015
activity status	employed	28.7 (1.6) [6.2]	17.8 (1.9) [5.4]	32.6 (2.0) [5.2]
activity status	unemployed	6.5 (4.7) [0.4]	4.0 (4.3) [1.1]	−10.0 (5.8) [0.6]
activity status	retired	15.6 (2.2) [2.0]	21.0 (1.9) [2.3]	11.7 (3.0) [2.0]
activity status	study	0.2 (1.8) [1.1]	−4.5 (1.9) [1.1]	1.3 (2.3) [0.7]
activity status	other	0.5 (2.4) [1.1]	9.1 (2.2) [1.0]	−3.8 (3.3) [1.0]
nr working	0	44.3 (3.8) [2.1]	22.7 (2.5) [3.1]	25.3 (6.6) [2.4]
nr working	1	0.9 (1.8) [3.0]	−0.4 (2.1) [3.4]	−0.4 (2.2) [3.1]
nr working	2	31.3 (2.1) [4.7]	26.3 (2.6) [3.6]	34.8 (2.6) [3.3]
nr working	3	34.0 (5.7) [0.8]	33.1 (5.5) [0.7]	42.8 (3.7) [0.6]
nr working	4	27.4 (13.8) [0.2]	36.6 (10.4) [0.2]	46.2 (9.1) [0.2]
main income	employment	35.0 (2.4) [7.5]	17.1 (2.4) [6.7]	33.7 (2.4) [6.1]
main income	self-employment	−3.3 (3.0) [0.6]	−16.9 (5.2) [0.4]	14.8 (8.2) [0.7]
main income	other	23.3 (4.3) [2.8]	20.9 (2.5) [3.9]	18.4 (4.7) [2.8]
occupation	basic level	5.2 (3.1) [1.6]	11.4 (2.1) [1.5]	2.0 (3.4) [1.4]
occupation	mid-level	9.0 (1.4) [5.1]	8.3 (1.7) [4.8]	2.5 (1.8) [4.1]
occupation	technicians, associates	13.9 (2.8) [0.7]	11.1 (3.7) [0.7]	13.9 (2.8) [0.6]
occupation	professionals	22.1 (2.7) [1.5]	22.0 (2.6) [1.9]	22.3 (2.2) [1.9]
occupation	managers	19.8 (3.4) [0.7]	14.7 (3.4) [0.7]	17.2 (4.3) [0.6]
sector	agriculture	−15.7 (4.3) [0.5]	−10.7 (5.1) [0.4]	−19.7 (3.7) [0.3]
sector	industry	5.3 (2.1) [1.1]	0.9 (2.7) [0.9]	3.7 (2.5) [0.8]
sector	IT, finance, RE, admin	17.0 (6.5) [0.2]	8.6 (4.4) [0.5]	6.8 (3.9) [0.6]
sector	public admin, education, health	10.2 (2.2) [1.4]	13.0 (2.5) [1.3]	9.8 (2.4) [1.1]
sector	other services	3.1 (1.9) [2.2]	−0.8 (1.9) [1.8]	0.2 (2.1) [1.8]
age bracket	under 19	−0.3 (2.1) [0.6]	0.3 (2.8) [0.6]	1.0 (3.0) [0.4]
age bracket	19–29	1.9 (2.2) [2.2]	−2.5 (2.1) [2.0]	−0.6 (2.6) [1.7]
age bracket	30–64	5.3 (1.3) [6.4]	−6.8 (1.3) [6.4]	2.7 (1.6) [5.7]
age bracket	65+	7.9 (2.5) [1.8]	19.3 (1.9) [2.0]	7.9 (2.6) [1.9]
dependents	0	−4.1 (1.7) [5.3]	3.5 (1.7) [4.9]	−4.5 (2.0) [4.8]
dependents	1	7.2 (2.1) [2.9]	2.1 (2.3) [2.8]	12.5 (2.9) [2.1]
dependents	2	8.4 (2.7) [2.0]	−3.1 (2.6) [2.3]	1.2 (2.8) [2.0]
dependents	3	3.1 (7.0) [0.5]	−9.6 (5.4) [0.7]	−2.5 (7.0) [0.5]
dependents	4	28.9 (10.5) [0.1]	5.6 (17.7) [0.2]	43.4 (7.1) [0.1]
dependents	5	48.1 (18.4) [0.1]	38.2 (25.9) [0.1]	12.0 (37.0) [0.1]
education	up to secondary	6.2 (1.6) [5.8]	10.6 (1.5) [6.0]	2.7 (2.1) [4.6]
education	post-secondary	8.0 (1.5) [2.4]	3.8 (1.7) [2.1]	3.6 (1.9) [1.9]
education	tertiary education	29.6 (2.5) [2.5]	22.8 (2.7) [2.9]	27.7 (2.4) [3.0]
sex	male	4.1 (1.0) [4.9]	0.0 (0.8) [4.9]	2.4 (1.1) [4.3]
sex	female	−3.0 (0.8) [6.0]	0.3 (0.7) [6.1]	−1.8 (1.0) [5.3]

Notes: Bootstrapped standard errors are in the parenthesis. Number of observations in thousands in brackets.

states. To do this, we will use the methodology provided by Yitzhaki and Lerman (1991), outlined in Appendix 1.

The subgroup decomposition results are presented in Table 9. The Gini coefficient is decomposed into within, between, and stratification component for each of the nine groupings considered before. The following conclusions can be drawn:

- The majority of inequality decomposes into within-groups rather than between-groups in Lithuania. The largest between-contribution is observed between different households which have a different number of people working (nr working, 10 Gini points), but even here the within-contribution is 3 times higher. This finding is not surprising, as inequality within subgroups is often found to matter more (see Elbers et al., 2008), suggesting that the majority of variation in income is between households of similar observable characteristics. Income inequality within groups is also more important for the EU. Additionally, several household characteristics seem to not contribute to inequality significantly in Lithuania, for example, sex.

Table 9. Decomposition of the Gini coefficient in 2015.

Grouping (1)	Decomposition (2)	EU (3)	new EU states (4)	old EU states (5)	LT (6)
activity status	between	3.3 (0.2) [47.5]	3.8 (0.3) [18.7]	3.1 (0.2) [28.8]	7.0 (0.5) [9.5]
activity status	stratification	−1.7 (0.1) [47.5]	−1.9 (0.2) [18.7]	−1.6 (0.1) [28.8]	−3.5 (0.3) [9.5]
activity status	within	29.2 (0.3) [47.5]	29.3 (0.5) [18.7]	29.2 (0.3) [28.8]	33.4 (0.7) [9.5]
nr working	between	3.5 (0.2) [47.9]	3.5 (0.3) [19.0]	3.5 (0.2) [28.9]	10.1 (0.7) [9.6]
nr working	stratification	−1.8 (0.1) [47.9]	−1.7 (0.2) [19.0]	−1.8 (0.1) [28.9]	−5.0 (0.4) [9.6]
nr working	within	29.1 (0.3) [47.9]	29.3 (0.5) [19.0]	29.0 (0.3) [28.9]	31.9 (0.8) [9.6]
main income	between	1.1 (0.1) [47.9]	2.9 (0.3) [19.0]	0.8 (0.1) [28.9]	8.4 (0.7) [9.6]
main income	stratification	−0.5 (0.1) [47.9]	−1.4 (0.2) [19.0]	−0.4 (0.1) [28.9]	−4.1 (0.4) [9.6]
main income	within	30.2 (0.3) [47.9]	29.6 (0.5) [19.0]	30.2 (0.3) [28.9]	32.6 (0.7) [9.6]
occupation	between	4.6 (0.2) [41.1]	5.3 (0.4) [16.2]	4.5 (0.3) [24.9]	5.5 (0.7) [8.6]
occupation	stratification	−2.4 (0.1) [41.1]	−2.6 (0.2) [16.2]	−2.4 (0.2) [24.9]	−2.8 (0.4) [8.6]
occupation	within	28.0 (0.3) [41.1]	27.5 (0.5) [16.2]	28.0 (0.3) [24.9]	34.1 (0.7) [8.6]
sector	between	1.9 (0.2) [21.2]	4.2 (0.5) [8.5]	1.5 (0.2) [12.7]	1.2 (0.5) [4.5]
sector	stratification	−1.0 (0.1) [21.2]	−2.1 (0.3) [8.5]	−0.7 (0.1) [12.7]	−0.5 (0.3) [4.5]
sector	within	27.9 (0.4) [21.2]	27.7 (0.7) [8.5]	27.8 (0.4) [12.7]	32.4 (0.9) [4.5]
age bracket	between	0.4 (0.1) [47.9]	0.3 (0.1) [19.0]	0.4 (0.1) [28.9]	1.4 (0.2) [9.6]
age bracket	stratification	−0.1 (0.1) [47.9]	−0.1 (0.1) [19.0]	−0.2 (0.1) [28.9]	−0.5 (0.1) [9.6]
age bracket	within	30.6 (0.3) [47.9]	30.8 (0.5) [19.0]	30.5 (0.3) [28.9]	36.1 (0.8) [9.6]
dependents	between	0.8 (0.1) [47.9]	1.6 (0.3) [19.0]	0.7 (0.1) [28.9]	1.1 (0.4) [9.6]
dependents	stratification	−0.4 (0.1) [47.9]	−0.7 (0.1) [19.0]	−0.3 (0.1) [28.9]	−0.4 (0.2) [9.6]
dependents	within	30.4 (0.3) [47.9]	30.2 (0.5) [19.0]	30.4 (0.3) [28.9]	36.3 (0.8) [9.6]
education	between	3.4 (0.2) [47.2]	4.6 (0.3) [18.7]	3.2 (0.2) [28.5]	5.9 (0.6) [9.5]
education	stratification	−1.8 (0.1) [47.2]	−2.3 (0.2) [18.7]	−1.7 (0.1) [28.5]	−3.1 (0.4) [9.5]
education	within	29.2 (0.3) [47.2]	28.9 (0.5) [18.7]	29.2 (0.3) [28.5]	34.2 (0.7) [9.5]
sex	between	0.0 (0.1) [47.9]	0.1 (0.1) [19.0]	0.0 (0.1) [28.9]	0.2 (0.1) [9.6]
sex	stratification	0.0 (0.0) [47.9]	0.0 (0.0) [19.0]	0.0 (0.0) [28.9]	−0.1 (0.0) [9.6]
sex	within	30.8 (0.3) [47.9]	31.0 (0.5) [19.0]	30.7 (0.3) [28.9]	36.9 (0.8) [9.6]

Notes: The first figure in columns (3–6) represents the contribution to Gini coefficient of equivalised household disposable income. Bootstrapped standard errors are provided in the parenthesis. Number of observations are shown in thousands in brackets.

- Except for education, labour market characteristics of the household are more important in explaining inequality than demographics. For example, the different number of people working, the main source of income of the household, and the occupation individually explain 5–10 Gini points. The between-contribution, when grouping people according to activity status is 7 Gini points. This means that if all household members were employed and would earn employment income, the Gini coefficient would fall by 7 points and become similar to the EU Gini coefficient. This between-contribution in Lithuania is about 2 times higher than the EU between-contribution, indicating that employment is much more important in terms of income in Lithuania than in the EU. Low redistribution (low taxes and transfers) in Lithuania could explain why it is very costly to not participate in the labour market (IMF, 2016; Lazutka, 2017). Furthermore, the number of those employed within a household matter in Lithuania. Demographic characteristics (age, number of dependents, sex) determine a relatively lower share (0.2–1.4 of Gini).

The within, between and stratification decomposition is decomposed further to reveal the importance of the employed to income inequality each year from 2005 to 2015. Specifically, the within-contribution of activity status is decomposed to the within contribution of the employed, unemployed, and non-participants. This decomposition, along with the between and stratification contributions, is shown in Table 10 for Lithuania.

Table 10. Decomposition of the first differences of the Gini coefficient of equivalised disposable income in Lithuania in 2015.

year	employed	unemployed	other	between	stratification	sum
2005	−0.80	−0.30	−0.20	−0.30	0.00	−1.60
2006	−0.20	−0.50	−0.40	0.20	−0.10	−1.00
2007	0.70	0.00	0.40	−0.50	0.20	0.80
2008	0.40	0.30	0.70	−0.20	0.20	1.40
2009	0.20	0.70	0.00	0.00	0.00	0.90
2010	−4.10	0.60	−0.40	0.00	0.20	−3.70
2011	−0.50	−0.70	−0.20	0.60	−0.40	−1.20
2012	2.30	−0.10	0.60	−0.10	−0.10	2.60
2013	0.40	−0.20	0.00	0.20	0.00	0.40
2014	2.80	0.10	−0.20	0.50	−0.30	2.90
2015	−1.00	0.10	−0.30	0.40	−0.20	−1.00

Notes: The sum indicates the first differences of the Gini coefficient, while other columns show the contribution. In 2014, the Gini rose by 2.9 Gini points and 2.8 Gini points are explained by the change in the contribution of the employed.

The rise in disposable income household inequality in Lithuania since 2011 can be primarily explained by a rise in income inequality among those who are employed. This is partly determined by the fact that a larger share of the population has become employed since the crisis (51% in 2011 and 55% in 2015), the employed are taking a larger share of income (from 62% to 68%) and are themselves more unequally distributed (the within-Gini rose from 29 to 33). To a lesser extent, inequality is also rising due to greater between-subgroup inequality and stratification, especially stratification of the employed vis-a-vis other groups. This is because average wages rose faster than non-labour income during this period.

4. Structure of income inequality by income factors

We estimate the structure of income inequality by decomposing household disposable income inequality by factors. Knowing which factors contribute to income inequality help explain why income inequality in Lithuania is high. The four components of disposable income are labour income, capital income, transfers, and taxes (including social transfers). These are further broken down by more granular income factors.

We use the Lerman and Yitzhaki (1985) method to decompose the Gini coefficient. It allows decomposing $\widehat{G}$ into income factors $\sum_{k=1}^{K} \widehat{T}_k$, where k represents labour, capital, transfers and taxes. We further decompose $\widehat{T}_k$ into $(\widehat{R}_k/100)(\widehat{G}_k/100)\widehat{S}_k$. Here $\widehat{R}_k$ is the estimate of Gini correlation between household disposable income and factor k. The quantity $\widehat{R}_k$ ranges between −100 and 100. The value $\widehat{R}_k = 100$ refers to high positive correlation. This means that households with a lot of factor k also have a lot of total disposable income, while households with little factor k have small disposable income. If $\widehat{R}_k$ is close to −100, it means that households with little disposable income tend to have larger factor k income. Next, $\widehat{G}_k$ represents the Gini index of factor k and is approaching 100 if inequality of k is high. Finally, component $\widehat{S}_k$ is the share of factor k of the household disposable income, meaning that factors which constitute a larger share of income matter more for inequality. More details on this method are provided in Appendix 2. We provide the estimates for Lithuania and the EU. Unfortunately, 4 countries, including Germany, did not provide all

Table 11. Factor decomposition of the Gini coefficient in 2015 by labour, capital, transfers, taxes and their sub-factors.

Variable	EU	new EU states	old EU states	LT
Gini	30.86 (0.30)	30.55 (0.44)	30.88 (0.38)	36.96 (0.76)
Labour	43.91 (0.63)	42.51 (0.78)	44.25 (0.8)	53.63 (1.28)
employment	30.95 (0.52)	33.90 (0.78)	30.06 (0.65)	34.48 (1.18)
employer's social insurance contribution	7.52 (0.17)	5.47 (0.16)	8.03 (0.20)	9.67 (0.38)
self-employment	5.27 (0.36)	3.06 (0.33)	5.97 (0.45)	9.29 (0.86)
company car	0.16 (0.02)	0.09 (0.01)	0.18 (0.03)	0.19 (0.05)
income received by people aged under 16	0.00 (0.00)	−0.01 (0.01)	0.01 (0.01)	0.00 (0.00)
Capital	2.76 (0.26)	0.96 (0.13)	3.29 (0.34)	1.32 (0.29)
interests, dividends, etc.	1.63 (0.23)	0.48 (0.09)	1.96 (0.29)	1.11 (0.28)
rental income	1.13 (0.09)	0.49 (0.08)	1.32 (0.11)	0.20 (0.05)
Transfers	4.66 (0.27)	2.56 (0.30)	5.24 (0.34)	−0.25 (0.30)
old-age benefits	5.07 (0.26)	3.28 (0.28)	5.62 (0.33)	−0.44 (0.25)
unemployment	0.19 (0.07)	0.02 (0.03)	0.20 (0.09)	0.04 (0.06)
survivor benefits	0.28 (0.06)	−0.04 (0.07)	0.37 (0.07)	−0.12 (0.03)
sickness benefits	0.06 (0.01)	0.10 (0.02)	0.05 (0.02)	0.47 (0.05)
education-related allowances	−0.04 (0.01)	−0.01 (0.02)	−0.05 (0.01)	0.00 (0.01)
family/children related allowances	−0.37 (0.04)	−0.18 (0.07)	−0.43 (0.05)	0.41 (0.15)
disability benefits	−0.11 (0.03)	−0.26 (0.07)	−0.06 (0.04)	0.03 (0.11)
social exclusion	−0.28 (0.02)	−0.18 (0.02)	−0.32 (0.03)	−0.41 (0.05)
housing allowances	−0.23 (0.02)	−0.05 (0.01)	−0.3 (0.02)	−0.03 (0.00)
regular inter-household cash transfers received	0.00 (0.06)	0.02 (0.04)	−0.01 (0.08)	−0.12 (0.05)
regular inter-household cash transfers paid	−0.14 (0.03)	−0.20 (0.08)	−0.12 (0.02)	−0.09 (0.04)
individual private pension	0.24 (0.04)	0.06 (0.01)	0.29 (0.05)	0.02 (0.01)
Taxes	−20.46 (0.31)	−15.49 (0.32)	−21.89 (0.39)	−17.74 (0.67)
tax on income and social insurance contributions	−12.79 (0.20)	−9.98 (0.25)	−13.69 (0.24)	−8.03 (0.30)
employer's social insurance contribution	−7.52 (0.17)	−5.47 (0.16)	−8.03 (0.20)	−9.67 (0.38)
regular taxes on wealth	−0.15 (0.01)	−0.04 (0.01)	−0.17 (0.02)	−0.04 (0.01)

Note: Bootstraped standard errors are provided in the parentheses.

the necessary income factors, meaning that the data sample for the EU differs from the previous analysis.

Table 11 reveals the results for the decomposition of disposable income into $\widehat{T}_k$ for Lithuania and the EU by factors and the further decomposition into $\widehat{R}_k\widehat{G}_k\widehat{S}_k$ is available in Table 12.

- Labour income contributes most to income inequality in Lithuania. It contributes 53.63 Gini points to total inequality. Labour income contributes most to income inequality on the EU level as well, yet about 9.72 Gini points less than in Lithuania. The labour component is especially large as it includes an employer's social insurance contributions. Capital contributes only 1.32 and transfers and taxes reduce income inequality by 0.25 and 17.74 points respectively.
- All labour sub-factors contributions are larger in Lithuania than in new and old EU states. The largest sub-factor contribution is employee income in Lithuania (34.48 Gini points). The contribution is about 0.58 Gini points higher than in the new EU states and 4.42 higher than in the old EU states. Self-employed contribute less to inequality in Lithuania (9.29 Gini points). However, this is by 6.23 Gini points more than in new EU states and by 3.32 Gini points more than in the old EU states.
- Labour income has a greater contribution in Lithuania than in the EU largely because this income is more correlated with disposable income in Lithuania. In other words,

Table 12. Factor decomposition of the of Gini of disposable income in 2015.

Variable	Contribution	EU	new EU states	old EU states	LT
Labour	*T*	43.91 (0.63)	42.51 (0.78)	44.25 (0.8)	53.63 (1.28)
Labour	*R*	75.55 (0.59)	79.58 (0.63)	74.38 (0.75)	90.61 (0.48)
Labour	*G*	52.7 (0.35)	49.26 (0.53)	53.66 (0.43)	52.22 (0.91)
Labour	*S*	110.28 (0.67)	108.47 (0.74)	110.87 (0.86)	113.35 (1.13)
employee	*T*	30.95 (0.52)	33.9 (0.78)	30.06 (0.65)	34.48 (1.18)
employee	*R*	70.43 (0.66)	75.68 (0.8)	68.87 (0.84)	81.44 (1.13)
employee	*G*	56.49 (0.37)	53.9 (0.59)	57.37 (0.46)	55.16 (0.93)
employee	*S*	77.8 (0.6)	83.12 (0.82)	76.07 (0.73)	76.77 (1.24)
self-employment	*T*	5.27 (0.36)	3.06 (0.33)	5.97 (0.45)	9.29 (0.86)
self-employment	*R*	44.65 (1.84)	30.06 (2.51)	48.5 (2.14)	70.11 (2.44)
self-employment	*G*	92.06 (0.41)	89.43 (0.39)	92.73 (0.51)	91.13 (0.58)
self-employment	*S*	12.82 (0.43)	11.39 (0.46)	13.28 (0.54)	14.53 (1.01)
Capital	*T*	2.76 (0.26)	0.96 (0.13)	3.29 (0.34)	1.32 (0.29)
Capital	*R*	67.94 (2.14)	68.24 (3.15)	67.87 (2.33)	76.41 (5.08)
Capital	*G*	92.89 (0.33)	98.06 (0.14)	90.92 (0.44)	98.13 (0.29)
Capital	*S*	4.37 (0.28)	1.44 (0.13)	5.32 (0.36)	1.75 (0.31)
Transfer	*T*	4.66 (0.27)	2.56 (0.3)	5.24 (0.34)	−0.25 (0.3)
Transfer	*R*	21.22 (1.07)	13.35 (1.45)	22.98 (1.28)	−1.84 (2.23)
Transfer	*G*	66.88 (0.35)	64.69 (0.74)	67.27 (0.41)	57.26 (0.86)
Transfer	*S*	32.84 (0.4)	29.65 (0.52)	33.87 (0.5)	23.47 (0.66)
Tax	*T*	−20.46 (0.31)	−15.49 (0.32)	−21.89 (0.39)	−17.74 (0.67)
Tax	*R*	80.78 (0.47)	78.28 (0.65)	81.43 (0.56)	81.03 (1.14)
Tax	*G*	53.35 (0.37)	50.02 (0.56)	53.7 (0.45)	56.75 (0.94)
Tax	*S*	−47.49 (0.35)	−39.56 (0.3)	−50.07 (0.42)	−38.57 (0.71)

Notes: G is decomposed into income factors $\sum_{k=1}^{4} T_k$, where k represents labour, capital, transfers and taxes. We further decompose T_k into $(R_k/100)(G_k/100)S_k$. Here R_k is the Gini correlation between household disposable income and factor k that ranges between −100 and 100. The component G_k represents the Gini index of factor k and S_k is the share of factor k of the household disposable income. Bootstrapped standard errors are provided in the parentheses.

those who get a lot of labour income tend to be the richest households in terms of disposable income also. This is seen from $\widehat{R}$, the value of which is equal to 90.61 in Lithuania, while it is under 79.58 in new EU states and 74.38 in old EU states. This is especially true for the self-employed: $\widehat{R}$ is equal to 70.11 and this is 25.46 points more compared to the EU. High $\widehat{R}$ means that self-employment income is especially important for self-employed households. This may give rise to concern, as such income is generally less stable than employment income. In contrast, for the labour income, $\widehat{G}$ in Lithuania is similar to $\widehat{G}$ observed in other EU countries while $\widehat{S}$ is only slightly larger.

- Taxes (and social contributions) negatively contribute to income inequality in Lithuania. Specifically, taxes reduce income inequality by 17.74 Gini points. This reduction is a couple of percentage points less than the EU and the old EU states in particular. The biggest difference is a lower $\widehat{S}$, which means that taxes constitute a smaller share of disposable income in Lithuanian than in the EU.
- Transfers seem to not contribute to income inequality in Lithuania. Specifically, transfers contribute −0.25 Gini points. At first this may seem surprising, as transfers are known to be of much greater effect in reducing income inequality (see, e.g. Joumard et al., 2013). However, it would be more correct to say that transfers do not contribute to inequality – i.e. they are not a part of the structure of inequality, instead of saying that transfers do not affect inequality. On the contrary, transfers can have a large effect. Upon closer inspection, we see the low contribution is due to a low $\widehat{R}$ which equals −1.84 for Lithuania. Upon multiplying $\widehat{R}$ by $\widehat{G}$ and $\widehat{S}$, the inequality contribution is close to zero. Therefore, the larger the $\widehat{S}$ going to transfers, the lower the inequality. Since transfers do not

contribute to inequality and taxes reduce inequality, their relative effect on inequality is not comparable using this method. This leads us to Section 5 which discusses their relative effects.

5. Marginal and redistribute effect of taxes and transfers on income inequality in Lithuania

In this section, we answer how much do transfers and taxes affect income inequality. We do so first by calculating the marginal effects: how does inequality respond to a percent change in an increase in taxes or transfers. Second, we estimate the redistributive effect of taxes and public transfers. Specifically, we analyse two ways in which taxes and public transfers can affect income inequality: by increasing their progressivity and their rate.

We use the Lerman and Yitzhaki (1985) decomposition to shed light on the marginal contribution of each income factor to the Gini coefficient. We calculate the amount by which the Gini changes if we raise the factor contribution by a small value e_k and hold other income factors constant. This is approximately equal to evaluating how many Gini points will the Gini coefficient change if we increase an income factor by 1%. The formula (A6) in Appendix 2 quantifies the effects. If all income factors are raised by the same $e_k = e$, the Gini would not change, as summarized in the first row of Table 13.

Table 13 shows the marginal contributions to the Gini for Lithuania and the EU. Several conclusions can be drawn on taxes and transfers as well as labour and capital income.

- Transfers and taxes reduce income inequality. Raising transfers by 1% reduces inequality by 0.0892 Gini points while raising taxes (including social contributions) reduces income inequality by 0.0348 Gini points. Additionally, raising transfers has a larger effect in Lithuania than in the EU. Increasing old-age transfers alone would reduce inequality by 0.0544 Gini points – three times more than in the EU. Other transfers have a much smaller impact individually. Taxes, however, have less effect in Lithuania than in the EU, especially the old EU states. Specifically, a 1% rise in income taxes and social contributions paid by the household reduces inequality by 0.0348 Gini points – about half of the impact in the old EU states, which is 0.0643. However, the tax situation in Lithuania is very similar to that of new EU states.
- Raising labour income would result in higher inequality in Lithuania and the effect is stronger for Lithuania than for the EU. A 1% increase in labour income means a 0.1147 rise in income inequality in Lithuania. This is almost 0.02 Gini points more than in the EU. The reason why inequality would rise more in Lithuania than in the EU is self-employment income. A 1% rise in self-employed income raises income inequality by 0.0391 Gini points in Lithuania as compared to 0.0131 Gini points in the EU. Raising employment income would raise income inequality by similar amounts in both economies.

The reasons why raising old-age benefits reduces inequality in Lithuania more than in the EU are most likely related to the design of the pension systems in Lithuania and the EU. First, the social expenditure on pensions in Lithuania is lower than in the EU (Lis, 2018). Because of this, the retired have lower incomes as compared to the rest of the population

Table 13. Marginal decomposition of the Gini coefficient in 2015 by labour, capital, transfers, taxes and their sub-factors.

Variable	EU	new EU states	old EU states	LT
Gini	0.0000 (0.0000)	0.0000 (0.0000)	0.0000 (0.0000)	0.0000 (0.0000)
Labour	0.0987 (0.0038)	0.0938 (0.0038)	0.1001 (0.0049)	0.1174 (0.0062)
employment	0.0694 (0.0035)	0.0851 (0.0042)	0.0657 (0.0043)	0.0611 (0.0072)
employer's social insurance contribution	0.0155 (0.0011)	0.0126 (0.0011)	0.0149 (0.0013)	0.0161 (0.0023)
self-employment	0.0131 (0.0025)	−0.0042 (0.0026)	0.0187 (0.0032)	0.0391 (0.005)
company car	0.0008 (0.0001)	0.0005 (0.0001)	0.0008 (0.0002)	0.001 (0.0003)
income received by people aged under 16	−0.0001 (0.0000)	−0.0002 (0.0001)	−0.0001 (0.0001)	0.0000 (0.0000)
Capital	0.0141 (0.0017)	0.0052 (0.0009)	0.0164 (0.0022)	0.0067 (0.0018)
interests, dividends, etc.	0.0087 (0.0015)	0.0028 (0.0006)	0.0102 (0.0019)	0.0063 (0.0018)
rental income	0.0054 (0.0006)	0.0024 (0.0006)	0.0062 (0.0007)	0.0004 (0.0003)
Transfers	−0.0547 (0.0022)	−0.0650 (0.0025)	−0.0522 (0.0028)	−0.0892 (0.0029)
old-age benefits	−0.0164 (0.002)	−0.0311 (0.0022)	−0.0118 (0.0026)	−0.0544 (0.0023)
unemployment	−0.0053 (0.0006)	−0.0020 (0.0003)	−0.0069 (0.0008)	−0.0031 (0.0005)
survivor benefits	−0.0031 (0.0005)	−0.0053 (0.0006)	−0.0025 (0.0006)	−0.0041 (0.0004)
sickness benefits	−0.0003 (0.0001)	−0.0001 (0.0001)	−0.0004 (0.0001)	−0.0007 (0.0003)
education-related allowances	−0.0012 (0.0001)	−0.0006 (0.0002)	−0.0014 (0.0002)	−0.0003 (0.0001)
family/children related allowances	−0.0114 (0.0005)	−0.0104 (0.0008)	−0.0116 (0.0006)	−0.0054 (0.0011)
disability benefits	−0.0072 (0.0004)	−0.0101 (0.0009)	−0.0062 (0.0004)	−0.0112 (0.0014)
social exclusion	−0.0047 (0.0003)	−0.0029 (0.0003)	−0.0053 (0.0004)	−0.0066 (0.0008)
housing allowances	−0.0042 (0.0002)	−0.0008 (0.0001)	−0.0054 (0.0003)	−0.0004 (0.0001)
regular inter-household cash transfers received	−0.0030 (0.0004)	−0.0029 (0.0004)	−0.0030 (0.0006)	−0.0037 (0.0008)
regular inter-household cash transfers paid	0.0010 (0.0003)	0.0007 (0.0008)	0.0011 (0.0003)	0.0007 (0.0004)
individual private pension	0.001 (0.0003)	0.0003 (0.0001)	0.0012 (0.0003)	0.0001 (0.0001)
Taxes	−0.0581 (0.0017)	−0.0340 (0.0018)	−0.0643 (0.0021)	−0.0348 (0.0039)
tax on income and social insurance contributions	−0.0440 (0.0011)	−0.0224 (0.0015)	−0.0511 (0.0014)	−0.0189 (0.0017)
employer's social insurance contribution	−0.0155 (0.0011)	−0.0126 (0.0011)	−0.0149 (0.0013)	−0.0161 (0.0023)
regular taxes on wealth	0.0014 (0.0001)	0.0009 (0.0001)	0.0017 (0.0002)	0.0002 (0.0001)

Note: Bootstraped standard errors are provided in the parentheses.

and this difference is larger than for the EU (see Table 3). This means that any transfers to this group will on average reduce inequality more in Lithuania. Second, the old-age benefits that are handed out in Lithuania depend on previous contributions but are not very elastic to it. This means that the old-age benefits are relatively equally distributed amongst the retired and perhaps more so than in other countries. As a consequence, the retired are relatively more equal amongst themselves (see Table 5) as compared to inequality within other activity status groups. Therefore, increasing the income share of the pensioners, the most equal subgroup in society, will reduce overall income inequality also. However, whether the pensions in other EU countries are more or less elastic to previous contributions than Lithuania remains to be tested.

Similarly, the reasons why raising tax income would reduce income inequality in Lithuania less than in the EU is likely related to the design of the respective tax and social contribution systems. Lithuania's social contribution constitutes over 3/4 labour taxes. But they are not progressive. The social contribution rates are flat without a ceiling and are therefore not redistributive among those who pay the contributions. Income tax constitutes just a quarter of labour taxes and, apart from a non-taxable minimum, has been non-progressive in 2005–2015 either. This means that while raising taxes will bring those with labour income closer to those without labour income, it will not reduce income inequality amongst those who have labour income.

The reason why raising labour income results in more inequality in Lithuania than in the EU may also be related to the tax system and tax evasion. In Lithuania, the self-employed benefited from a lower taxable base. Additionally, the self-employed seem to evade taxes more often than employed in Lithuania (Černiauskas & Jousten, 2020). As a result, there is very little redistribution for the self-employed taking place in Lithuania. Given that self-employment income is effectively not taxed, it correlates so well with disposable income and the Gini correlation coefficient $\widehat{R}$ was so high in Table 12.

Next, we estimate the redistributive effect of taxes and public transfers for the total population and self-employed separately. We follow Joumard et al. (2013), which is based on Kakwani (1977). This method also lets us decompose the redistribution effect into the progressivity and average rate of taxes or public transfers in Lithuania and compare these figures with the ones in the EU.

For i denoting taxes or transfers, the redistributive effect is decomposed as follows (Joumard et al., 2013):

$$redistribution_i = \frac{r_i}{100} progressivity_i, \tag{2}$$

takes the values from −100 to 100, where −100 indicates regressive i and 100 indicates progressive i.

Specifically, we apply the following calculations to get the average rate r_i and the progressivity index. To compute r_{tax}, we divide the total taxes paid by the disposable income of the population and multiply by 100. To compute $r_{transfers}$, we divide the public transfers received by the market income after transfers of the population and multiply by 100. To compute $progressivity_{tax}$, we subtract the concentration coefficient of market income after public transfers from the concentration coefficient of taxes. To compute the $progressivity_{transfers}$, we subtract the concentration coefficient of public transfers from the concentration coefficient of market income. The concentration coefficient is familiar to the Gini index. Like the Gini index, it is computed using (A1), where y represents the variables *tax* or *transfers*. However, *tax*, *transfers*, and survey weights are sorted according to market income. It is also possible to sort by disposable income. In that case, the progressivity measures would be much smaller. However, we prefer sorting by market income, because we see the Lithuanian and EU system as transferring to and taxing from households primarily based on their market incomes.

The redistributive effects of taxes with social security contributions are similar to the redistributive effects of public transfers for Lithuania. The effects on the Gini of market income, as well as the components of the effects, are available in Table 14 for Lithuania

and the EU in 2015. Both taxes and public transfers have a very similar effect on redistributing incomes. Interestingly, taxes excluding employer's social insurance contributions contribute much less to income redistribution in Lithuania and the EU. Since other studies typically disregard employer's social contributions, it could explain why they find taxes to be playing a small role in redistribution (see, e.g. Causa & Hermansen, 2017; OECD, 2011).

Taxes have a high redistributive effect because of the average tax rate, while public transfers have a high effect because of their progressiveness in Lithuania. The average tax rate constitutes 38.6% of disposable income which is more than double the public transfer rates (16.7% of market income after transfers). However, taxes are much less progressive (31.4%) as compared to public transfers (78.7%). This means that raising tax progressivity will have a higher impact on reducing income inequality than raising public transfer progressivity, while raising the average public transfer rate will have a higher effect on income inequality than raising the average tax rate in Lithuania and, similarly, in the EU.

The redistributive effects of public transfers and taxes are much lower in Lithuania than in the EU. The redistributive impact of taxes in Lithuania is almost two times smaller than in the EU, while public transfers are about 50% smaller. All the subcomponents are smaller. Tax progressivity and the average rate of public transfers in particular are lower in Lithuania as compared to the EU.

The tax system is much less distributive amongst the self-employed in Lithuania. The redistributive effect of taxes is negative in Lithuania as shown in Table 15. This means that the poorer households pay a larger share of their disposable income in taxes than the richer households. This is in line with previous findings (Černiauskas & Jousten, 2020). We additionally see that this is very different when compared to the EU, wherein taxes do have a positive redistributive effect. Additionally, the average tax rate of the self-employed for Lithuania is less than a third of the EU and almost a quarter of the tax rates of the old EU states. Therefore, negative tax progressivity can explain why the self-employed contribute more to inequality in Lithuania than in other EU states.

The results suggest that raising tax progressivity and the average rate of public transfers should reduce income inequality most. We run a simulation (for the full population) to

Table 14. Progressivity index for market incomes in 2015.

	EU	new EU states	old EU states	LT
Redistributive effect of public transfers	19.8	18.5	20.2	13.2
Redistributive effect of taxes	23.9	16.3	26.8	12.1
Redistributive effect of taxes without ESC[a]	8.9	7.1	9.6	3.8
Average tax rate	47.8	41.2	50.1	38.6
Average public transfer rate	21.9	21.3	22.1	16.7
Tax progressivity index	50.0	39.6	53.5	31.4
Public transfers progressivity index	90.6	86.9	91.7	78.7

Notes: The redistributive effects of public transfers and taxes are calculated by multiplying their progressivity index with the average rates as in (2). To compute the average tax rate, we divide the taxes paid by the disposable income of the population. To compute the average public transfer rate we divide the public transfers received by the market income after public transfers of the population. To compute the progressivity of taxes, we subtract the concentration coefficient of market income after public transfers from the concentration coefficient of taxes. To compute the progressivity of public transfers, we subtract the concentration coefficient of public transfers from the concentration coefficient of market income. Tax progressivity is measured using the Kakwani index, where 100 is a very progressive Tax system and −100 is a very regressive tax system. The same is applied to transfers.

[a]ESC – employer's social insurance contributions.

Table 15. Progressivity index for market incomes in 2015 for self-employed.

	EU	new EU states	old EU states	LT
Redistributive effect of transfers	1.3	1.9	1.1	2.0
Redistributive effect of taxes	4.9	1.7	5.6	−1.3
Redistributive effect of taxes without ESC[a]	7.5	4.0	8.1	−0.5
Tax progressivity index	12.4	6.4	12.8	−11.3
Transfers progressivity index	19.1	20.5	17.1	25.8
Average tax rate	39.0	26.1	43.9	11.1
Average transfer rate	7.0	9.2	6.2	7.8

Notes: The redistributive effects of public transfers and taxes are calculated by multiplying their progressivity index with the average rates as in (2). To compute the average tax rate, we divide the taxes paid by the disposable income of the population. To compute the average public transfer rate we divide the public transfers received by the market income after public transfers of the population. To compute the progressivity of taxes, we subtract the concentration coefficient of market income after public transfers from the concentration coefficient of taxes. To compute the progressivity of public transfers, we subtract the concentration coefficient of public transfers from the concentration coefficient of market income. Tax progressivity is measured using the Kakwani index, where 100 is a very progressive Tax system and −100 is a very regressive tax system. The same is applied to transfers.

[a]ESC – employer's social insurance contributions.

observe this. We simulate the effect of increasing the average rate and changing the progressivity of taxes and public transfers on Lithuania using EU-SILC data. The effect of changing the progressivity or average rate of tax and public transfers on the Gini of Lithuania is illustrated in Figure 2. We simulate the average rate of taxes by increasing the taxes for all those who are currently paying taxes. We do a similar simulation for public transfers. We increase taxes and transfers by up to 5 percentage points of market income after public transfers. We increase the progressivity of taxes by increasing taxes by up to 5 percentage points for the top quintile of households that are sorted by market incomes and redistributing this gain to all other quintiles. The redistribution is also progressive. For example, if we were to increase taxes on the top quintile by 10%, then the 4th quintile will get to pay about 10% fewer taxes, the third will pay 20% less, the second 30% less and the first will

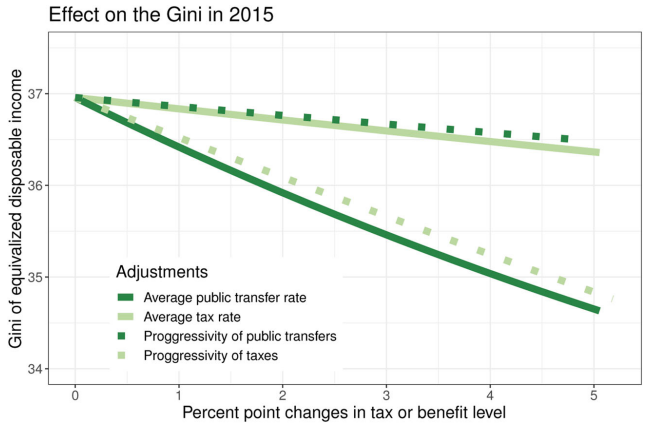


Figure 2. Simulating the effect of changes in progressivity and average rate of tax and public transfers on the Gini coefficient of equivalised disposable income in Lithuania.

pay 40% less. A scalar is added so that the reduction in taxes for the four bottom quintiles equals the increase in taxes for the top quantile. We increase the progressivity of public transfers by increasing transfers received by up to 5 percentage points for the bottom quintile of households that are sorted by market incomes and redistributing the cost to all other quintiles in a similar manner as for taxes. The simulations confirm that increasing the average rate of public transfers has a much higher effect on the Gini than raising taxes by the same amount. Increasing tax progressivity has a larger effect than increasing public transfer progressivity.

6. Conclusions

We have tackled three questions and each of them is elaborated in this study. We have also suggested possible improvements for future studies.

First, we have run three statistical tests and found that equivalised income inequality in Lithuania is in all cases one of the highest in the EU. Specifically, we have tested for accuracy of estimates by estimating their standard errors, the inequality measure used as well as different equivalence scales. In all cases, equivalised income inequality in Lithuania is found to be one of the highest across the EU.

Second, we have investigated why equivalised income inequality in Lithuania is higher compared to the EU by using univariate decomposition techniques. We have found large inequalities between and within many groups of households in the country. In all cases, the within-group inequality contributes more to equivalised income inequality in Lithuania and the EU. It means that this inequality is higher within households of similar observable characteristics rather than between households of different characteristics. Inequalities within the unemployed and those working in the agricultural sector are especially prominent. Nevertheless, between-contributions are also significant for Lithuania, suggesting where policy can look into deeper. The largest between-group inequalities lie between the employed and the rest of the population. Moreover, this type of inequality has been rising over time. As the factor decomposition shows, the large between-group inequality contribution can be explained by unequal distribution of labour income, especially – self-employment income.

Third, we analysed the extent to which equivalised income inequalities stemming from the market income are offset by taxes and transfers. Specifically, we analysed the marginal and redistributive effects of Lithuania's taxes and transfers and compared this to the EU. The marginal decomposition of the Gini coefficient of equivalised disposable income by factors confirms that an increase in tax and transfer income reduces equivalised income inequality while an increase in labour income increases it. The way that the tax and transfer system is currently designed, the average marginal contribution is more than twice higher for transfers compared to taxes, and that among the transfers the role of the old-age pensions is the highest. Similarly, the analysis of the redistributive effect of the taxes and public transfer income also showed that these two income sources reduce income inequality. However, the redistributive impact of taxes in Lithuania is almost two times smaller than in the EU, while public transfers are about 50% smaller. The redistributive effect of taxes for the self-employed is negative in Lithuania and therefore reinforces income inequality, while taxes reduce inequality amongst the self-employed in the EU. This means that the current tax

system and tax evasion/avoidance of higher-income households are likely to be responsible for a larger self-employment income contribution to inequality in Lithuania as opposed to EU.

We also decomposed the redistributive effect into the progressivity and the average rate of tax and public transfers effect. We find that the tax progressivity and the average rate of public transfers in particular are lower in Lithuania as compared to the EU. The results suggest that raising tax progressivity and the average rate of public transfers would reduce equivalised income inequality most.

The estimates of equivalised income inequality may have several drawbacks. First, there is a large shadow economy in Lithuania, with some estimates exceeding 25% of GDP in 2013 and 2015 (see Schneider, 2013; Žukauskas, 2016). Even though survey respondents are informed that their data will not be used for tax purposes, some of them may still be unwilling to disclose information on their true income received. It remains unclear how this affects equivalised income inequality because it depends on the income distribution within the shadow economy together with the income distribution of the observed economy. Additionally, this estimate may cause problems when comparing households across countries, since the size of the shadow economy is particularly large in Lithuania. Second, as has been already pointed out various times, EU-SILC undersamples the income of rich individuals in all countries (especially capital income (Navickė & Lazutka, 2018)) – something that the survey weights do not correct for. Including the rich will result in higher measures of equivalised income inequality in Lithuania. However, equivalised income inequality will rise in other EU countries as well. Therefore, the relative position of Lithuania vis-a-vis other countries may not change so much. Nevertheless, the alternative Household Finance and Consumption Survey (HFCN, 2019) could partly correct for both of these shortcomings, as it has data on consumption, which can be used to estimate the shadow economy and oversample the wealthy households for Lithuania along with many other EU countries. Furthermore, greater access to administrative data would be yet another path to take.

Future studies can also consider using an alternative methodology, for example, by using multivariate techniques to decompose equivalised income inequality. This was not the focus of the current study because the results of a multivariate decomposition depend on all variables by which the Gini is decomposed, and there is no consensus on which should be included. Furthermore, variables available to some countries are less available in others in the EU-SILC. Nevertheless, our additional check using a multivariate decomposition technique as in Social Situation Monitor (2017) does not contradict the results. Additionally, one may look into income inequality between individuals instead of households.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Notes on contributors

Nerijus Černiauskas, Phd student of Economics at Vilnius University working on income inequality in Lithuania.

Dr. Andrius Čiginas, Researcher on statistics and probability at Vilnius University.

ORCID

Nerijus Černiauskas http://orcid.org/0000-0003-4001-3677
Andrius Čiginas http://orcid.org/0000-0001-8509-5034

References

Aghion, P., Caroli, E., & Garcia-Penalosa, C. (1999). Inequality and economic growth: The perspective of the new growth theories. *Journal of Economic Literature*, *37*(4), 1615–1660. https://doi.org/10.1257/jel.37.4.1615

Atkinson, A. B. (1970). On the measurement of inequality. *Journal of Economic Theory*, *2*(3), 244–263. https://doi.org/10.1016/0022-0531(70)90039-6

Atkinson, A. B., & Piketty, T. (2010). *Top incomes: A global perspective*. Oxford University Press.

Berger, Y. G. (2008). A note on the asymptotic equivalence of Jackknife and linearization variance estimation for the Gini coefficient. Journal of official statistics. 24(4). https://www.researchgate.net/publication/277826713_A_Note_on_the_Asymptotic_Equivalence_of_Jackknife_and_Linearization_Variance_Estimation_for_the_Gini_Coefficient

Buhmann, B., Rainwater, L., Schmaus, G., & Smeeding, T. M. (1988). Equivalence scales, well-being, inequality, and poverty: Sensitivity estimates across ten countries using the Luxembourg income study (LIS) database. *Review of Income and Wealth*, *34*(2), 115–142. https://doi.org/10.1111/roiw.1988.34.issue-2

Causa, O., & Hermansen, M. (2017). *Income redistribution through taxes and transfers across OECD countries* (No. 1453). https://doi.org/10.1787/bc7569c6-en. https://www.oecd-ilibrary.org/content/paper/bc7569c6-en

Černiauskas, N., & Jousten, A. (2020, February). *Statutory, effective and optimal net tax schedules in Lithuania* (Bank of Lithuania Working Paper Series 72). Bank of Lithuania. https://ideas.repec.org/p/lie/wpaper/72.html

Černiauskas, N., Sologon, D. M., O'Donoghue, C., & Tarasonis, L. (2020, January). *Changes in income inequality in Lithuania: The role of policy, labour market structure, returns and demographics* (Bank of Lithuania Working Paper Series 71). Bank of Lithuania. https://ideas.repec.org/p/lie/wpaper/71.html

Chiappori, P.-A., & Meghir, C. (2015). Intrahousehold Inequality. In A.B Atkinson & F Bourguignon (Eds.), *Handbook of Income Distribution* (Vol. 2, pp. 1369–1418). Elsevier.

Cingano, F. (2014). *Trends in income inequality and its impact on economic growth* (No. 163). https://EconPapers.repec.org/RePEc:oec:elsaab:163-en

Cowell, F. A. (2000). Measurement of inequality. *Handbook of Income Distribution*, *1*, 87–166. https://doi.org/10.1016/S1574-0056(00)80005-6

Cowell, F. (2011). *Measuring inequality*. Oxford University Press.

Das, T., & Parikh, A. (1982). Decomposition of inequality measures and a comparative analysis. *Empirical Economics*, *7*(1), 23–48. https://doi.org/10.1007/BF02506823

Elbers, C., Lanjouw, P., Mistiaen, J. A., & Özler, B. (2008). Reinterpreting between-group inequality. *The Journal of Economic Inequality*, *6*(3), 231–245. https://doi.org/10.1007/s10888-007-9064-x

Eurostat. (2018a). *EU statistics on income and living conditions (EU-SILC) methodology -- Distribution of income*. https://ec.europa.eu/eurostat/statistics-explained/index.php?title=EU_statistics_on_income_and_living_conditions_(EU-SILC)_methodology_-_distribution_of_income

Eurostat. (2018b). *EU statistics on income and living conditions (EU-SILC) methodology – Concepts and contents*. https://ec.europa.eu/eurostat/statistics-explained/index.php?title=EU_statistics_on_income_and_living_conditions_(EU-SILC)_methodology_%E2%80%93_concepts_and_contents#Adjusted_cross_sectional_weight_.28RB050a.29

Eurostat. (2018c). *European union statistics on income and living conditions (EU-SILC)*. https://ec.europa.eu/eurostat/web/microdata/european-union-statistics-on-income-and-living-conditions

Eurostat. (2020). *Gini coefficient of equivalised disposable income – EU-SILC survey*. https://ec.europa.eu/eurostat/data/database

Garner, T. I., & Terrell, K. (1998). A gini decomposition analysis of inequality in the Czech and Slovak Republics during the transition. *The Economics of Transition*, *6*(1), 23–46. https://doi.org/10.1111/ecot.1998.6.issue-1

Goedemé, T. (2013). How much confidence can we have in EU-SILC? Complex sample designs and the standard error of the Europe 2020 poverty indicators. *Social Indicators Research*, *110*(1), 89–110. https://doi.org/10.1007/s11205-011-9918-2

Grigoli, F., & Robles, A. (2017). Inequality overhang. International Monetary Fund.

HFCN. (2019). *Household finance and consumption network (HFCN)*. ECB. https://www.ecb.europa.eu/pub/economic-research/research-networks/html/researcher_hfcn.en.html

IMF. (2016). *Republic of Lithuania, selected issues* (IMF Country Report No. 16/126).

Jenkins, S. (2017). *The measurement of income inequality* (pp. 17–52). Routledge.

Joumard, I., Pisu, M., & Bloch, D. (2013). Tackling income inequality: the role of taxes and transfers. *OECD Journal: Economic Studies*, *2012*(1), 37–70. https://doi.org/10.1787/eco_studies-2012-5k95xd6l65lt

Kakwani, N. C. (1977). Measurement of tax progressivity: An international comparison. *The Economic Journal*, *87*(345), 71–80. https://doi.org/10.2307/2231833

Lazutka, R. (2017). Pervedimų, Valstybės Teikiamų Prekiuų ir Paslaugų Vaidmuo. In *Pajamų Nelygybė Lietuvoje, October 5th*, Lietuvos Nacionalinė Martyno Mažvydo Biblioteka. Bank of Lithuania. https://www.lb.lt/uploads/documents/files/Lazutka.pdf

Lerman, R. I., & Yitzhaki, S. (1985). Income inequality effects by income source: A new approach and applications to the United States. *The Review of Economics and Statistics*, *67*(1), 151–156. https://doi.org/10.2307/1928447

Lis, M. (2018, September). *Lithuania's pension system in the international persective* (Technical report).

Navickė, J., & Lazutka, R. (2018). Distributional implications of the economic development in the baltics: Reconciling micro and macro perspectives. *Social Indicators Research*, *138*(1), 187–206. https://doi.org/10.1007/s11205-017-1645-x

OECD. (2011). *Divided we stand*. https://doi.org/10.1787/9789264119536-en. https://www.oecd-ilibrary.org/content/publication/9789264119536-en

OECD. (2015a). *All on board*. https://doi.org/10.1787/9789264218512-en. https://www.oecd-ilibrary.org/content/publication/9789264218512-en

OECD. (2015b). *In it together: Why less inequality benefits all*. https://doi.org/10.1787/9789264235120-en. https://www.oecd-ilibrary.org/content/publication/9789264235120-en

Ostry, J. D., & Berg, A. (2011). Inequality and unsustainable growth: Two sides of the same coin? https://www.imf.org/en/Publications/Staff-Discussion-Notes/Issues/2016/12/31/Inequalityand-Unsustainable-Growth-Two-Sides-of-the-Same-Coin-24686

Ostry, J. D., Berg, A., & Tsangarides, C. G. (2014). *Redistribution, inequality, and growth*. International Monetary Fund.

Pahl, J. (1995). His money, her money: Recent research on financial organisation in marriage. *Journal of Economic Psychology*, *16*(3), 361–376. https://doi.org/10.1016/0167-4870(95)00015-G

Plat, D. (2012). *IC2: Inequality and concentration indices and curves* (R package version 1.0-1). https://cran.r-project.org/web/packages/IC2/index.html

Rao, J. N. K., Wu, C. F. J., & Yue, K. (1992). Some recent work on resampling methods for complex surveys. *Survey Methodology*, *18*(2), 209–217. https://doi.org/10.1007/978-3-0348-8930-8_11

Schneider, F. (2013). *Size and development of the shadow economy of 31 european and 5 other OECD countries from 2003 to 2013: A further decline*. http://politeia.org.ro/wp-content/uploads/2013/05/ShadEcEurope31_Jan2013.pdf

Shorrocks, A. F. (1982). Inequality decomposition by factor components. *Econometrica: Journal of the Econometric Society*, *50*(1), 193–211. https://doi.org/10.2307/1912537

Social Situation Monitor. (2017). *Research findings – Social situation monitor – Measurement and methodology*. https://ec.europa.eu/social/home.jsp?langId=en

SODRA. (2020). *Contribution rates for employees*. SODRA. https://www.sodra.lt/en/benefits/contributionrates/contribution-rates-for-employees

Vogler, C., & Pahl, J. (1994). Money, power and inequality within marriage. *The Sociological Review*, *42* (2), 263–288. https://doi.org/10.1111/j.1467-954X.1994.tb00090.x
Yitzhaki, S. (1983). On an extension of the Gini inequality index. *International Economic Review*, *24*(3), 617–628. https://doi.org/10.2307/2648789
Yitzhaki, S., & Lerman, R. I. (1991). Income stratification and income inequality. *Review of Income and Wealth*, *37*(3), 313–329. https://doi.org/10.1111/roiw.1991.37.issue-3
Zardo, T. L., & Goedemé, T. (2016). *Notes on updating the EU-SILC UDB sample design variables, 2012–2014*.
Žukauskas, V. (2016). Lietuvos Šešėlinė Ekonomika, Lietuvos laisvosios rinkos institutas, Vilnius.

Appendices

$\mathcal{U} = \{1, \ldots, N\}$ is the set representing elements of the finite survey population, and $y_1, \ldots, y_N$ are values of the variable of interest (income) in $\mathcal{U}$. The subset $s = \{i_1, \ldots, i_n\}$ of $\mathcal{U}$ is the sample, while w_i and $i \in s$ are the corresponding survey weights. We use the estimator

$$\widehat{G} = \frac{2}{\hat{\mu}}\left(\frac{1}{\widehat{N}}\sum_{i\in s} w_i y_i \widehat{F}(y_i) - \frac{1}{\widehat{N}^2}\sum_{i\in s} w_i y_i \sum_{i\in s} w_i \widehat{F}(y_i)\right) \tag{A1}$$

of the Gini coefficient (1), constructed in line with Berger (2008), where $\widehat{F}(y_i)$ are values of the estimated distribution function

$$\widehat{F}(y) = \frac{1}{\widehat{N}}\sum_{j\in s} w_j \mathbb{I}\{y_j \le y\} \quad \text{with} \quad \widehat{N} = \sum_{j\in s} w_j, \tag{A2}$$

and

$$\hat{\mu} = \frac{1}{\widehat{N}}\sum_{j\in s} w_j y_j.$$

Here $\mathbb{I}\{\cdot\}$ stands for the indicator function. Estimators of the subgroup and factor decompositions are constructed using similar plug-in principles.

Appendix 1. Subgroup decompositions

We give the decomposition of (A1) by groups as in Yitzhaki and Lerman (1991). Let $s = s_1 \cup \cdots \cup s_L$ be a division of the sample by non-overlapping groups. Denote

$$\widehat{N}_l = \sum_{j\in s_l} w_j, \quad \widehat{P}_l = \frac{\widehat{N}_l}{\widehat{N}}, \quad \hat{m} = \frac{1}{\widehat{N}}\sum_{j\in s} w_j y_j, \quad \hat{m}^{(l)} = \frac{1}{\widehat{N}_l}\sum_{j\in s_l} w_j y_j, \quad \widehat{F}^{(l)} = \frac{1}{\widehat{N}_l}\sum_{i\in s_l} w_i \widehat{F}(y_i), \tag{A3}$$

where $\widehat{N}_l$ is the estimated population size in the subgroup l, the quantity $\widehat{P}_l$ is the estimated population share, $\hat{m}$ is the estimated mean of the survey variable in $\mathcal{U}$, $\hat{m}^{(l)}$ is the estimated mean in the subgroup, and $\widehat{F}^{(l)}$ is the estimate of the average of global ranks in the subgroup l. Consider the values $\widehat{F}_l(y_i)$ and $\widehat{F}_{L\setminus l}(y_i)$, $i \in s$, of the estimated distribution functions

$$\widehat{F}_l(y) = \frac{1}{\widehat{N}_l}\sum_{j\in s_l} w_j \mathbb{I}\{y_j \le y\} \quad \text{and} \quad \widehat{F}_{L\setminus l}(y) = \frac{1}{\widehat{N} - \widehat{N}_l}\sum_{j\in s\setminus s_l} w_j \mathbb{I}\{y_j \le y\}$$

in the subgroup l and outside this subgroup, respectively. Introduce the notations

$$\widehat{cov}(m^{(l)}, F^{(l)}) = \sum_{l=1}^{L} \widehat{P}_l\left(\hat{m}^{(l)} - \sum_{l=1}^{L} \widehat{P}_l \hat{m}^{(l)}\right)\left(\widehat{F}^{(l)} - \sum_{l=1}^{L} \widehat{P}_l \widehat{F}^{(l)}\right)$$

and

$$\widehat{cov}_l(y, F_l(y)) = \frac{1}{\widehat{N}_l}\sum_{i\in s_l} w_i(y_i - \hat{m}^{(l)})\left(\widehat{F}_l(y_i) - \frac{1}{\widehat{N}_l}\sum_{i\in s_l} w_i\widehat{F}_l(y_i)\right),$$

and

$$\widehat{cov}_l(y, F_l(y) - F_{L\setminus l}(y)) = \widehat{cov}_l(y, F_l(y)) - \frac{1}{\widehat{N}_l}\sum_{i\in s_l} w_i(y_i - \hat{m}^{(l)})\left(\widehat{F}_{L\setminus l}(y_i) - \frac{1}{\widehat{N}_l}\sum_{i\in s_l} w_i\widehat{F}_{L\setminus l}(y_i)\right).$$

Then the estimated decomposition by groups is written as

$$\widehat{G} = \underbrace{\sum_{l=1}^{L}\widehat{S}_l\widehat{G}_l}_{\text{Within}} + \underbrace{\sum_{l=1}^{L}\widehat{S}_l\widehat{G}_l\widehat{Q}_l\left(\widehat{P}_l - 1\right)}_{\text{Stratification}} + \underbrace{\frac{2\widehat{cov}(m^{(l)}, F^{(l)})}{\hat{m}}}_{\text{Between}}, \tag{A4}$$

where

$$\widehat{S}_l = \widehat{P}_l\frac{\hat{m}^{(l)}}{\hat{m}}, \quad \widehat{G}_l = \frac{2\widehat{cov}_l(y, F_l(y))}{\hat{m}^{(l)}}, \quad \widehat{Q}_l = \frac{\widehat{cov}_l(y, F_l(y) - F_{L\setminus l}(y))}{\widehat{cov}_l(y, F_l(y))}.$$

Here the component $\widehat{S}_l$ represents the share of the survey variable, $\widehat{G}_l$ is the estimated within-group Gini coefficient, and the part $\widehat{Q}_l$ is the estimated stratification term.

Appendix 2. Factor decompositions

We write down an estimate of the factor decomposition by Lerman and Yitzhaki (1985). Write $y_i = \sum_{k=1}^{K} y_i^{(k)}$, where k is a factor of the survey variable. Consider the values $\widehat{F}(y_i)$ and $\widehat{F}(y_i^{(k)})$, $i \in s$, of distribution function (A2) and denote the expressions

$$\widehat{cov}(y^{(k)}, F(y)) = \frac{1}{\widehat{N}}\sum_{i\in s} w_i y_i^{(k)}\widehat{F}(y_i) - \frac{1}{\widehat{N}^2}\sum_{i\in s} w_i y_i^{(k)}\sum_{i\in s} w_i\widehat{F}(y_i)$$

and

$$\widehat{cov}(y^{(k)}, F(y^{(k)})) = \frac{1}{\widehat{N}}\sum_{i\in s} w_i y_i^{(k)}\widehat{F}(y_i^{(k)}) - \frac{1}{\widehat{N}^2}\sum_{i\in s} w_i y_i^{(k)}\sum_{i\in s} w_i\widehat{F}(y_i^{(k)}).$$

Also, introduce the weighted means

$$\hat{\mu}^{(k)} = \frac{1}{\widehat{N}}\sum_{j\in s} w_j y_j^{(k)}.$$

Then the estimated decomposition by factors is

$$\widehat{G} = \sum_{k=1}^{K}\widehat{T}_k = \sum_{k=1}^{K}\widehat{R}_k\widehat{G}_k\widehat{S}_k, \tag{A5}$$

where

$$\widehat{R}_k = \frac{\widehat{cov}(y^{(k)}, F(y))}{\widehat{cov}(y^{(k)}, F(y^{(k)}))}, \quad \widehat{G}_k = \frac{2\widehat{cov}(y^{(k)}, F(y^{(k)}))}{\hat{\mu}^{(k)}}, \quad \widehat{S}_k = \frac{\hat{\mu}^{(k)}}{\hat{\mu}}.$$

Here $\widehat{R}_k$ is the estimate of the so-called Gini correlation between the survey variable and its kth component, $\widehat{G}_k$ represents the Gini index of factor k, and $\widehat{S}_k$ is the share of factor. For a small change in the

*k*th factor, the expression of marginal effects is

$$\frac{\partial \widehat{G}}{\partial e_k} = \widehat{S}_k(\widehat{R}_k\widehat{G}_k - \widehat{G}_k), \tag{A6}$$

see Lerman and Yitzhaki (1985).

2nd publication

Estimating statutory, effective and optimal net tax schedules in Lithuania

N. Černiauskas, A. Jousten.

IZA Journal of Labor Policy **11**, - (2021)

DOI: 10.2478/izajolp-2021-0005

Nerijus Cerniauskas[1,*] *and Alain Jousten*[2]

Statutory, effective, and optimal net tax schedules in Lithuania

Abstract

We estimate effective and optimal net income tax schedules and compare them to the estimated statutory rates for the case of Lithuania for the period 2014–2015. Values of effective net tax rates are estimated from the survey of EU Statistics on Income and Living Conditions; the statutory net tax rates are estimated with the European tax-benefit simulator EUROMOD, whereas optimal net taxes are calculated via Saez (2002) methodology. We find that the three net tax schedules are similar for employees in the middle of the income distribution. At the bottom of the income distribution, optimal net tax schedules suggest higher in-work benefits. The net tax schedules diverge substantially for the self-employed. At the top of the income distribution, where the majority of self-employed are concentrated, the self-employed are required to pay 15 cents less net taxes per Euro than employees—and they effectively pay 29 cents less.

Current version: February 16, 2021
Keywords: Optimal tax schedule, effective tax schedule, statutory tax schedule, taxes, transfers, employees, self-employed, Lithuania
JEL codes: H2, H21
Corresponding author: Nerijus Cerniauskas
nerijus.cerniauskas@evaf.vu.lt

1 Vilnius University (Faculty of Economics and Business Administration, Saulėtekio al. 9, Vilnius, Lithuania)

2 University of Liège (Faculty of Law, Tax Institute and HEC Liège, Rue Louvrex 14, 4000 Liège, Belgium), IZA (Schaumburg-Lippe-Straße 5-9, 53113 Bonn, Germany) and NETSPAR (Warandelaan 2, 5037 AB Tilburg, The Netherlands). Email: alain@jousten.net

Cite as: Cerniauskas and Jousten. *IZA Journal of Labor Policy* (2021) 11:5.
https://doi.org/10.2478/izajolp-2021-0005

1 Introduction

Although it is widely accepted that taxes are necessary to finance government expenditures and social transfer programs, there is a great deal of disagreement concerning who should be paying these taxes. Regarding labor income taxation, the optimal tax literature considers three factors (income distribution, labor elasticities, and society's preferences) when determining who should pay taxes and how much should they pay (see, e.g., Saez 2001). Oftentimes, however, statutory tax rates—the rates that are inscribed in the law—are smaller and less progressive than optimal ones (Saez 2002). Additionally, tax avoidance and fraud lead to further divergence between taxes that are actually paid (i.e., effective tax rates) and the optimal ones. These three concepts (optimal, effective, and statutory rates) are interrelated in a complex way: optimal taxes inform us about the desirable rate structure, whereas effective rates show how the tax system effectively taxes people based on rules set out by statutory rates as prescribed by law. The interplay between these concepts is key to addressing urgent public policy questions such as how statutory rates effectively impact on individuals, how the tax system fares as compared to optimality principles, etc.

We perform this analysis for the case of Lithuania. Our objective is twofold: first, we establish the extent to which the real world labor tax structure of the country is aligned with lessons from the optimal tax literature. Second, we compare the three schedules for employees and the self-employed. Governments utilize the tax system to encourage various types of behaviors—including the choice of self-employment. While this may have favorable effects on the labor supply or taxable income of those concerned, it may cause additional difficulties. For example, the self-employed usually face lower statutory income tax rates and are more likely to evade taxes as compared to employees, which leads to smaller government coffers and questions of social injustice (Milanez and Bratta 2019). Lithuania is a particularly interesting case study in this regard. First, it applies rather distinct rules for employees and the self-employed. Second, it enjoys good survey and administrative data availability.

This article relates to two bodies of tax literature. The first is the optimal tax literature, particularly the subbranch, which compares optimal tax schedules with statutory ones. The literature of optimal taxation started with partial equilibrium models based on individuals, most notably Mirrlees (1971). He demonstrated that higher marginal tax rates generate labor responses that cause employees to spend less time in employment. The Mirrlees model was modified by Saez (2001) by replacing theoretical labor responses with observable income-dependent labor supply elasticities. This methodology was first used to argue that optimal gross income (which excludes social contributions) tax rates of top incomes in the United States could exceed 50%. More recent studies have replaced the labor elasticity with elasticities of taxable income. These are considered broader than labor elasticity, as they include other behavior responses, such as tax evasion and avoidance, and not only labor supply. Klemm et al. (2018), also using Saez (2001) methodology and estimates of taxable elasticities, suggest that optimal income tax rates for top incomes exceed 60% for 27 global countries. A slightly modified version by Saez (2002) considers optimal tax rates at the bottom of the income distribution, by incorporating labor market responses at the intensive and extensive margin throughout the income distribution.

Subsequent authors have shown that optimal taxes rates differ, depending on the optimal tax schedule model. For example, Immervoll et al. (2011) extends Saez (2002) model

(which includes only individuals) to couples, and suggest lower taxes on secondary earners versus primary earners for a sample of 15 EU countries. Additionally, the income tax schedule also depends on the existence of non-income tax schedules. For example, Huang and Rios (2016) show that countries with a nonlinear income tax and a linear non-income tax (such as the value-added tax in Russia) should have lower marginal income tax rates. However, if a country also exhibits underreporting of high income, then marginal income taxes should be lifted again. Using general equilibrium models, other authors such as Heathcote et al. (2017) find that incorporating skill investment and public good provision suggests lower progressivity (although high poverty rates that prevent skill investment undermine such claims). There are also models that look at employment and self-employment simultaneously, for example, Zawisza (2019). This model incorporates own-elasticities to declare employment or self-employment income and evaluates the cross-elasticities of switching between employment and self-employment. He found the elasticities of the self-employed to be three times higher than the elasticities of the employed in Poland. The lack of consensus leaves the researcher puzzled as to which model to use, but the lack of elasticity and other parameter estimates constrain the model choice to that of Saez (2002). This means that we work with the same elasticity for the self-employed and employed, which may lead to an over-estimation of the optimal tax schedule for the self-employed.

Furthermore, the optimal tax literature has attempted to analyze different tax and income concepts. Mirrlees (1971), Saez (2001), and Immervoll et al. (2011) focused on income tax and employment income. Saez (2002) considered net taxes (income taxes minus public benefits), which means that individuals take into consideration their income taxes and (instantaneous) benefits when making employment decisions. This is useful when analyzing optimal taxes at the bottom of the income distribution, since high public benefits (such as unemployment benefits) may discourage work as much as high taxes. However, for most developed countries who belong to the Organisation for Economic Co-operation and Development (OECD), income tax constitutes a small part of the "tax" burden (OECD 2019). For them, social contributions are both higher and not necessarily actuarially fair, meaning that this, too, can be seen as a tax.

We also relate to the tax literature that examines statutory and effective tax rate differences between employees and the self-employed. Studies focusing on labor taxation show that statutory tax schedules for employees (OECD 2019) and for the self-employed (Milanez and Bratta 2019) vary across OECD countries and across different household types within countries. Estimates of effective tax rates largely come from the tax evasion literature, which implicitly compares statutory and effective tax rates, although the focus is often on the individual. The work that offers the closest parallel to our article is the one by Leventi et al. (2013), who estimate income misreporting in Greece of wages and of self-employment income in the period 2005–2009. They find that about 43% of self-employment income was underreported in 2009 and that the tails of income distribution underreported income more often. They do this by comparing European Union Statistics on Income and Living Conditions (EU-SILC) data on income coupled with administrative data on income, and use EUROMOD to streamline the definitions. They face the challenge of having different samples of people in the EU-SILC and the administrative records. In a different study, Johns and Slemrod (2010) find that top income-earners tend to avoid taxes, leading to lower effective tax rates in the United States, and Alstadsæter et al. (2017) find that the wealthiest Scandinavians also exhibit a similar trend. Even

though the evidence suggests that employees do evade income, up to 20% of the top incomes in Estonia do so (Paulus 2015), the self-employed tend to engage in tax evasion and avoidance substantially more (see, e.g., Baldini et al. 2009; Slemrod 2016) with some estimates showing that more than half of the income may be concealed from the authorities (Artavanis et al. 2016).

We find that the three net tax schedules diverge much more for the self-employed than for employees. In fact, the optimal, statutory, and effective tax rates for employees largely coincide for all but the tails of the income distribution. In contrast, for the self-employed, the effective tax rates are well below the statutory tax rates, while statutory rates are also below the optimal rates for most of the income distribution.

The article is structured as follows. In Section 2, we present the data sources and the definitions used throughout the article. The following three sections cover the statutory, effective, and optimal net tax schedules. The results are presented and discussed in Section 6, while the conclusions, recommendations, and limitations are presented in Section 7.

2 Data and Definitions

We use the EU-SILC dataset to estimate statutory, effective, and optimal net tax schedules for Lithuania. This is the only publicly available source of data with sufficient information for our analysis in one dataset for Lithuania, as it contains key information on employment income, taxes, benefits, household composition, and information that can help to classify individuals as employees or self-employed. The yearly EU-SILC has been running since 2004 and is the reference for comparable data on personal income in Europe. Each year, around 5,000 households encompassing around 10,000 household members over 16 years of age who agree to share information on their incomes are included.[1] We pool data from surveys carried out in 2015–2016, which contain income data (reference years) of 2014–2015. Though the data is well explained on the Eurostat website,[2] some features are mentioned here.

First, only certain income components are available for the household level in the survey. Notably, income tax and social contributions are calculated at the household level. This restricts the analysis to the concept of household (equivalized) income rather than individual income, which can be considered a blessing or a curse. On the one hand, the literature suggests that individuals make economic decisions taking themselves as well as their household members into consideration (see, among others, Vogler and Pahl 1994). For example, the incomes of all household members comprise a common budget constraint (Chiappori and Meghir 2015), thereby influencing each household member's behavior. Additionally, only some benefits are granted at the household level (e.g., social assistance benefit), making the allocation of this benefit to any specific household member artificial. Nevertheless, each household member has his/her own preferences and a typically unequal control of the household's budget, with evidence suggesting that decisions within households are rarely joint and more often taken by

1 For the reference year 2015, 5,142 households out of 6,161 households participated in the survey-interview. This means that at least one respondent was willing to fill in the survey on behalf of the household. For those 5,142 households, information on all household members was collected.

2 The website can be accessed at: https://ec.europa.eu/eurostat/web/microdata/european-union-statistics-on-income-and-living-conditions.

specific household members (Pahl 1995).[3] To partly account for the limitations of working with household data, we carry out an analysis of singles' households as a robustness check, but our results still hold.

Second, EU-SILC has a large survey component, but, since 2012, Lithuania has made heavy use of register (administrative) data. The State Social Insurance Fund Board data and the State Tax Inspectorate under the Ministry of Finance of the Republic of Lithuania data have been linked to sample data and used for checking cash or near-cash employee income, social insurance contributions and taxes on income, as well as old-age benefits. Maternity and maternity/paternity allowances, care allowance, social assistance, old-age, and survivor's pensions have been taken from the administrative data. (See country report[4] for more information.) Register data is directly imputed from the registers for households that agree to participate in the survey. If register data is not available, then survey data is used. In the case of income, particularly employment income and income from self-employment, data is taken from both administrative and survey sources, and the greater value of the two is used. This "true" income is later used to estimate statutory taxes. In this way, we can observe actual incomes and not just income that has been reported to the tax authorities. In the case of taxes and benefits, we mainly rely on administrative data.

Third, survey weights are used to partly adjust for probability of selection, nonresponse and, as appropriate, to adjust the sample to external data. Currently, the sample is adjusted for demographic and geographic external data only. The weights are further adjusted according to Eurostat (2018): weights of household members who are over age 16 are scaled up by distributing weights of those under age 16. For most of the calculations, we only considered households that had at least one nonstudent household member aged 18–62. This means we kept one observation per household whose weight was the sum of the individual weights in that household.

Fourth, there is evidence that income inequality is underestimated in EU-SILC (Hlasny and Verme 2018; Törmälehto 2017). Callan et al. (2020) find that in Ireland only the top 1% of income is missing from household surveys as compared to register data, after accounting for concept differences. In line with this, Navicke and Lazutka (2016) show that capital income is underreported for Lithuania in EU-SILC, which is usually concentrated at the top of income distribution, while other income components are much less underreported. A study of Estonian Household Finance and Consumption Survey by Meriküll and Room (2019) showed that the rich as well as the poor usually do respond to surveys and so unit nonresponse is a smaller problem, but income is underreported due to item nonresponse. In other words, the richer individuals do participate in household questioners but tend to avoid questions related to specific income/wealth questions. Since employee income, taxes, and social contributions for those who agree to participate in the survey are taken from registers in Lithuania, item nonresponse should be a smaller problem here. One major exception is self-employment income, which is not imputed from registers and has been often found omitted in the mentioned study. Unfortunately, Meriküll and Room (2019) do not succeed in replicating register data with survey data using data imputation techniques, citing lack of common support as the key issue. Furthermore, as self-employment income is often underreported to tax authorities, such imputations

3 Future studies should also compare them with net tax schedules for individual incomes or the interactions between individuals within a household.

4 The report can be accessed at: https://circabc.europa.eu/faces/jsp/extension/wai/navigation/container.jsp

are unhelpful in the first place. As we focus on labor rather than capital income and we see self-employment income as problematic to weight, we refrain from reweighing our data.

Finally, EU-SILC is compatible with EUROMOD. EUROMOD is a European tax-benefit simulator that takes in EU-SILC data and calculates how much tax each individual should pay or how many benefits he should receive based on his market income and other characteristics (e.g., age, whether there are any dependents, and employment status). This allows us to estimate statutory tax schedules. It should be noted that while EU-SILC is used for EUROMOD, there are adjustments made in the process.[5]

This data and EUROMOD allows us to estimate the three net tax schedules. Specifically, we estimate household equivalized net taxes as a share of household equivalized gross employment income. Let us explain each term in more detail. *Gross employment income* is defined as yearly gross employee and self-employed income (including social contributions of the employee, the self-employed, and the employer). *Net tax* is the difference between taxes paid and public transfers received. Gross employment income minus net taxes is net labor income. The unit of observation is a household to which we allocate an equivalized income.[6] To obtain equivalized income, we first sum the incomes of all household members for a given household. Then, we adjust the sum by an OECD-modified equivalence scale, where 1 is attributed to the first household member, 0.5 to the second and each subsequent person aged 14 and over, and 0.3 to each child aged under age 14. Henceforth, any reference to income or taxes in this text relates to equivalized household income and taxes. Finally, we construct a working sample that includes only households with at least one member who is not a student and is between 18 and 62 years of age. This allows us to focus on the working-age population and excludes pensioners—implicitly also reducing the role of these benefits in household income. We do not remove them completely, because many households have at least one pensioner or student, and they contribute to the household income. Income and net tax statistics from EU-SILC for 2014 and 2015 reference years for the full sample, which represents Lithuania's population, and the working sample is summarized in Table 1.

Table 1 Average yearly equivalized income and net taxes in Lithuania, Euro

Variable	Full sample (population)	Working sample (18–62, nonstudent)
Gross employment income	7,663	8,952
Net taxes (minus)	1,045	1,944
Net labor income	6,618	7,008
Number of households	9,657	6,459

Note: income variables are bolded while the number of households is in italics. Figures come from the EU-SILC dataset for Lithuania for income reference years of 2014-2015. Data for 2014–2015 income reference years comes from the EU-SILC dataset. Gross employment income and net taxes include employer's and employee's social contributions.

5 For example, 20 household members who were born after the income period were removed in the EUROMOD 2015 and 2016 input files. This meant that survey weights add up to different totals and equivalence scales also differ for those households. Unfortunately, the household IDs differ in the two data sets and we were not able to identify those household members that should be removed from EU-SILC to generate the same weights.

6 The alternative would be to have different tax rates for different types of households (e.g., single, married, married with children) as done in Guner et al. (2014), but using equivalized income allows us to have a single summary statistic and worry less about sample size.

We focus mainly on gross employment income and net labor income. These variables relate most closely with one's work incentives.[7] EU-SILC has more income variables that also relate to work incentives, but we refrain from discussing those.[8] Gross employee income is defined as the total remuneration in cash payable by an employer to an employee in return for work done by the latter during the income reference period, plus the employer's social insurance contribution. Gross self-employment income is defined as the income received during the income reference period by individuals, for themselves or in respect of their family members, as a result of their current or former involvement in self-employed work. Self-employed work covers those jobs where the remuneration is directly dependent upon the profits (or the potential for profits) derived from the goods and services produced (where own consumption is considered to be part of profits).

We include social contributions and all benefits in our definition of net tax to better reflect the incentives Lithuania's households face when participating in the labor market. Social contributions constitute a relatively large share of labor costs as well as the biggest source of revenue for the government (11.9% of GDP in 2015 according to Eurostat, whereas income tax makes up only 5.4%, even lower than VAT—7.7%). Although contributions are used to finance social benefits, and could be seen as tax-neutral, there are also reasons to think of them, at least in part, as a tax. In their book, Frölich et al. (2014) argue that some people may either not want the benefits associated with social contributions or want less of them, in which case only the difference between the desired benefits and the paid contributions should be considered as tax. For example, using US data, Chetty et al. (2016) find that poor people tend to live shorter lives, meaning they have less chance of getting any benefits despite their contributions. Knowing that the largest share of social contributions is to insure against old-age, not paying social insurance contributions may be a very rational response for these people. In such cases, people may either work less if the contributions are perceived as too high or turn to informal work to avoid paying them (Frölich et al. 2014). Since we cannot identify the part of social contributions that are paid willingly, or how much of other taxes people willingly pay in exchange for public goods and services, we include social contributions into our definition of tax. We include all benefits (old-age, sickness/health, disability, family, unemployment, and other benefits) into the definition of equivalized income.

Detailed statistics of income and net taxes as a percent of gross employment income are shown in Table 2. In all, 88% of gross employment income is derived from gross employee income, with the residual derived from self-employment income. Public transfers increase income, resulting in 21% higher gross labor income than gross employment income for the full sample, but only 13% in the working sample. Public transfers increase income by less in the working sample because we exclude a large share of pensioners together with their old-age

7 Other possible strategies could include looking at taxes only or net taxes, taking into account inter-temporal benefit accrual, such as for pensions. We reserve this for future research.

8 For example, disposable income includes all the variables that fall under net labor income as well as other incomes, such as private transfers, and other taxes, notably capital tax. These variables play a minor role in this survey and do not impact the results. EU-SILC also includes several noncash items that may have a larger impact on income and decision-making, but it is not clear to what extent this can be taxed. For example, noncash items, especially imputed rent, which is the approximate income one would receive if one was to rent his/her residence, constitutes about 18% of gross employment income of the working sample.

Table 2 Detailed equivalized income and net tax in Lithuania, percentage of gross employment income

Variable	Full sample (population)	Working sample (18–62, nonstudent)
Gross employee income	88	88
Gross self-employed income	12	12
Gross employment income	100	100
Old-age public transfer	12	4
Other public transfers	9	8
Gross labor income	121	113
Tax on income and social insurance contributions	−34	−34
Net labor income	86	78
Number of households	9,657	6,459

Note: income variables are bolded when they are aggregates of preceding variables, while the number of households is in italics. Figures come from the EU-SILC dataset for Lithuania for income reference years of 2014-2015.
All variables are in percent of gross employment income. Data for the income reference years of 2014–2015 comes from EU-SILC. Gross employment income and its components include employer's and employee's social contributions.

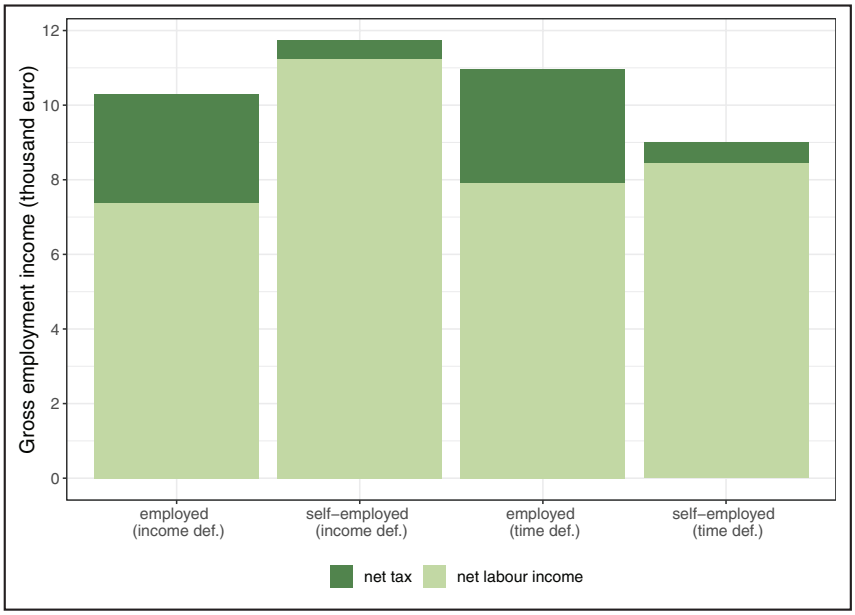
public transfers. Other public transfers[9] still constitute a sizable share of income in the working sample. Tax on income and social insurance contributions reduce gross employment income by just over a third. As a result, net labor income is 86% of gross employment income on average (78% of working sample). Therefore, the amount of net taxes as a percent of gross employment income is 34% in the working sample.

As Lithuania's tax system treats employees and the self-employed differently, we also examine different types of households. In total, there are three nonoverlapping groups of households: employees, self-employed, and other. We use two definitions to define a household. The preferred is the *Income* definition, where we sum household members' gross labor income components (employee, self-employed, and public transfer income) in a household and see which of the three components is dominant. Additionally, employee/self-employed households must have received or made a loss of at least 10 Euros of gross employee/self-employed income in the reference year; otherwise, they are classified as "other." The alternative is the *Time* definition, where the total household member's months spent in an activity is considered. Specifically, each household member had to identify his/her main activity in each month of the income reference year, be it an employee, self-employed, or other. We then sum all the months of all household members, note which is the largest, and label that household accordingly.

Using the income definition results in a higher net labor income of the self-employed households, as summarized in Figure 1. Under the income definition, self-employed households receive around 14% more gross employment income than employee households, but pay only 17% of the net taxes that employee households pay. This results in 52% higher net labor income of the self-employed as compared to employees. Under the time definition, the self-employed pay less net taxes than employees, but they also earn much less gross employment income. More generally, while self-employment is not the activity that households report spending most of

9 Disability benefits and family/children-related allowances each constitute about a third of the other public transfers. On the other hand, unemployment benefits only make up 10% of other public transfers.

Figure 1 Average equivalized yearly income in Lithuania for 2014–2015 reference years.



Bars represent average equivalized income for employee and self-employed households under two grouping definitions: income definition and time definitions. The sum of equivalized net labor income and equivalized net tax is equivalized gross employment income. Calculations are based on the working sample. There are 264 households that fall under the time definition for the self-employed and 545 under the income definition (4,566 and 4,889 for the employees, respectively).

their time on collectively, it is the one that generates the largest net labor income. Indeed, only 3.3% of households report spending most of their time in self-employment, whereas 7.4% report gaining most of their gross labor income from self-employment. This is largely because over half of household members who earn their own self-employment income also earn employee income, and 60% cohabit with someone who earns employee income. Those who earn their own employee income are much less likely to earn self-employment income (10%) or cohabit with someone who does (14%).

Finally, we compute average and marginal tax rates throughout the article. The formula for the average tax rate for the gross employment income decile $i = 1,2,...,10$ is the following equation:

$$atr_i = \frac{\sum_{k=1_i}^{n_i} taxes_k * w_k}{\sum_{k=1_i}^{n_i} income_k * w_k},$$

defined by the sum of taxes paid by households $k = 1_i, 2_i, ..., n_i$ and n_i would mean the nth household member of decile i. We adjust the distribution of taxes using survey weights w_k. Then, we divide the weighted taxes paid by the income of all households multiplied by their weights in decile i.

Similarly, marginal taxes for gross employment income decile $i = 2,3,...,10$ are given by the following equation:

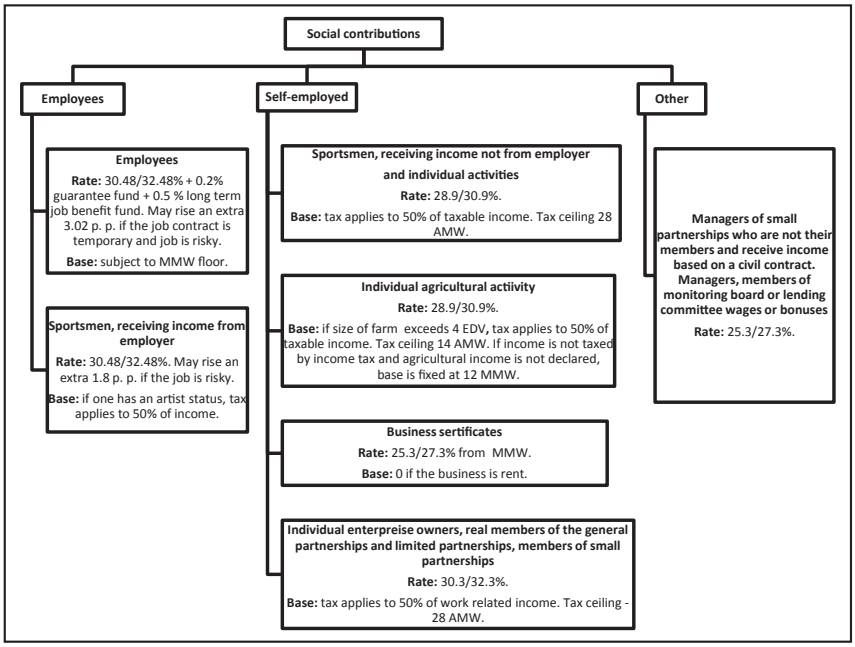
$$mtr_i = \frac{\sum_{k=1_i}^{n_i} taxes_k * w_k - \sum_{k=1_{i-1}}^{n_{i-1}} taxes_k * w_k}{\sum_{k=1_i}^{n_i} income_k * w_k - \sum_{k=1_{i-1}}^{n_{i-1}} income_k * w_k}$$

3 Statutory Net Tax Schedule

We proxy the characteristics of the statutory net tax schedule in Lithuania by applying the tax and benefit rules applicable in the country to the observations from EU-SILC. Specifically, we utilize EUROMOD—a tax and benefit simulator—to estimate the amounts of taxes and benefits that would be due if we simply apply the statutory rules to the data at hand for all households, and for the separate groups of employees and self-employed. We use the income definition to allocate households into employee and self-employed throughout this section. Finally, we present statutory average tax schedules for Lithuania for the two groups.

Lithuania's tax and benefit system is complex. First, it incorporates various taxes, social contributions, and benefits. We consider income tax, all social contributions, and a wide range of benefits. Most benefits, including pensions, are related to household members' previous income, although various coefficients, ceilings, and floors ensure some income redistribution

Figure 2 Statutory social insurance contributions excluding the statutory health insurance contributions prior to 2019 reform.



AMW, average monthly wage; MMW, monthly minimum wage. Sources: Based on state tax inspectorate of Lithuania.

in the system. Second, there are various household-member and household-level characteristics that determine how much net taxes a household member should pay. This results in a wide range of net taxes to consider.

Figure 2 presents the statutory social contribution rates and bases that we derive for the household member in our sample.[10] Different contribution rates and bases are applied to employees and the self-employed; gross employee taxable income is subject to a monthly minimum wage (MMW) floor, while most forms of self-employed income benefit from a 50% tax base reduction. Therefore, the effective taxes paid by the self-employed can be much smaller than those paid by employees. A likely possible weakness of our data is that some tax-relevant information for properly applying the statutory rules may not be factored in, hence inducing a potential bias of an a priori unknown sign. For example, the self-employed may benefit from carried-forward losses, a factor that would effectively further widen the difference in statutory rates between employees and the self-employed.[11]

EUROMOD and EU-SILC dataset for Lithuania is able to estimate the majority of taxes and a portion of benefits.[12] For example, family benefits that depend on the number of children and their ages are simulated. Furthermore, simulations are made for a number of contributory (social insurance-based) benefits, such as maternity leave or benefits assigned to low-income household members. A number of benefits with entitlement rights dependent on contribution history (i.e. pensions, sickness benefit, disability benefits, etc.) are not simulated due to the lack of data on previous employment history and salaries received, some event occurrence (i.e. disability or accident at work), or lack of information on previous partner entitlements (i.e. survival pensions). In those cases where potential benefits are not simulated, they are replaced with effective benefits from the input file. We run the simulations at a household-member level, after which we aggregate to household-level and adjust incomes by an equivalence scale. Finally, we construct a working sample by keeping households with at least one household member who is 18–62 years of age and is not a student. We use EUROMOD version's H1.0+ 2014 and 2015 Lithuania's system files on LT_2015_a1 and LT_2016_a2 input data, respectively.

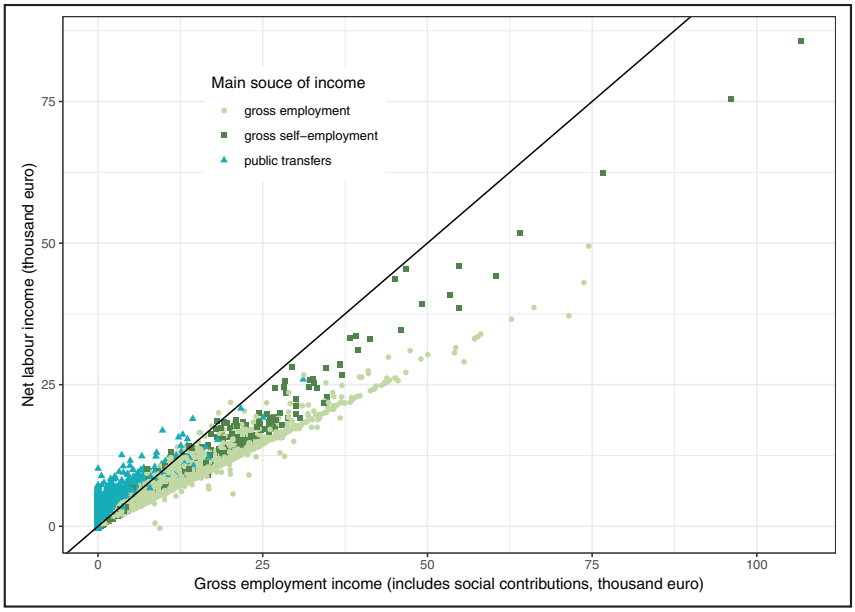
The relation between two simulated variables is plotted in Figure 3. On the *x*-axis is the gross employment income, and on the *y*-axis is net labor income (gross employment income minus net taxes). The figure thus links the mechanisms that transform gross employment income into net labor income. The diagonal line represents no transformation: what a household earns from employment income becomes its net labor income. Anything below the diagonal line refers to income that is taxed away. Anything above the diagonal line means that the household received public transfers that exceed paid taxes. The colors and shapes of the points represent the groups according to their main source of income: gross employee income, gross self-employment income, and public transfer income.

10 These social contributions were effective before a large tax reform that took place in 2019.

11 Other examples are tax exemptions for specific disabilities, economic activity, or information that is not collected in EU-SILC survey. If these were fully accounted for, the statutory rates would be lower and closer to the effective tax rates. It is also likely that these specificities will be more important for the self-employed.

12 EUROMOD input files are slightly modified versions of EU-SILC data. In the case of Lithuania, 10 household members who were not yet born in the reference period were dropped in the 2016 and 2015 surveys. EUROMOD also reads country-specific files, which describe the statutory taxes and benefits of those countries that are then applied on the input files. More information on EUROMOD can be found at: http://www.euromod.ac.uk and in Navicke and Cižauskaite (2018) in particular.

Figure 3 Statutory equivalized incomes of households grouped using the income definition in Lithuania.



Equivalized gross employment income and net labor income is in thousands of Euros per year. Households are allocated to groups according to the income definition for 2014–2015 income reference years and are represented by dots in the graph (see Section 2). The diagonal line illustrates that household employment income is equal to net labor income. Any dot above the diagonal line illustrates that the household receives additional benefits, whereas dots under the horizontal line mean that the household pays additional taxes or social contributions.

Many households that receive hardly any gross employment income are legally entitled to substantial public transfers, which raise their net labor income above the diagonal line. This is largely because some or all households are able to apply for old-age benefits or disability benefits. Once households start earning some gross employment income, their net labor income becomes dispersed and their main source of income is increasingly likely to be employment income. As gross employment income rises, the majority of households tend to be below the diagonal line, as they have to pay taxes and receive fewer benefits.

The self-employed households receive higher net labor income as compared to the employee households, especially at higher gross employment income levels. This is because employees are legally subject to higher statutory average tax rates than the self-employed for the same level of gross employment income. In part, this is due to the lower taxable base of the self-employed. Furthermore, the self-employed have access to more types of tax treatment. For example, the self-employed may purchase business certificates. This requires their holders to pay a one-off fee determined by the municipality if they receive less than 45,000 Euro from the activity. For a couple with two business certificates, this could lead to 90,000/1.5 = 60,000 Euro equivalized income that is barely taxed, whereas other types of incomes could be declared under different activity forms or taxed at a different rate thereafter.

Table 3 Statutory equivalized gross employment income, net taxes, and net labor income in thousand Euro per year.

Percentile	Gross employment	Net taxes			Net labor income		
	Income	All	Employees	Self-employed	All	Employees	Self-employed
0–7	0.00	−2.25			2.25		
−20	1.41	−1.12	0.25	0.42	2.53	1.72	1.19
−30	3.59	0.07	0.76	0.70	3.52	2.86	2.95
−40	5.10	0.78	1.29	1.01	4.32	3.81	4.14
−50	6.61	1.54	1.80	1.32	5.07	4.79	5.27
−60	8.22	2.29	2.48	1.86	5.93	5.73	6.34
−70	10.08	3.03	3.16	2.83	7.05	6.96	7.03
−80	12.36	4.08	4.28	2.91	8.28	8.08	9.39
−90	15.69	5.67	5.86	4.09	10.02	9.84	11.50
−100	27.65	10.29	10.91	7.02	17.36	16.65	21.21

Data is sorted according to equivalized gross employment income (includes social contributions). Net taxes include tax minus public benefits (public transfers). Net labor income is gross employment income plus benefits minus taxes. We report averages of percentile ranges. Gross employment income is taken from EU-SILC, whereas net taxes are estimated by EUROMOD, which takes into account various individual and household characteristics (e.g., age, health status). All figures are taken from EUROMOD and are weighted to include only those households with at least one member who is aged 18–62 and is not a student. The number of observations per decile is available in Table A1 in Appendix.

Not only do the self-employed earn more net labor income on average due to lower taxes, but self-employed households are concentrated at the top of the income distribution. For example, in the bottom 20% of the net labor income distribution, only 5% of households can be considered self-employed under the income definition. The share of households that are self-employed almost triples in the top 20% of the income distribution, and reaches 30% for the top 5% in Lithuania. Such a distribution of self-employed households also encourages us to make stronger claims on the richer self-employed rather than the poorer ones. Nevertheless, the data suggests that the self-employed are faring worse at the bottom of the income distribution. As seen in Table 3, employee households grouped using the income definition in the second (pseudo) decile receive 1,720 Euro net labor income, and do not pay any 250 net taxes. The self-employed receive less net labor income (1,190 Euro) and pay more net taxes. This is because the self-employed receive fewer benefits as compared to employees at the bottom of the income distribution, but they pay similar taxes.

Table 4 contains data on the composition of average statutory net tax rates. As gross employment income rises, average net tax rates rise as well. In particular, average net taxes are negative for the bottom percentiles (as people receive more benefits than they pay in taxes), and they rise to 36.7% of gross employment income.

At the bottom of the income distribution, both groups pay similar taxes as a share of gross employment income, even though reasons differ.[13] As gross employment income rises, employees receive less benefits and start paying more taxes as a share of gross employment income (due to the diminishing effect of nontaxable minimum for employees). The

13 The employees pay less tax because of a nontaxable minimum, which gradually diminishes as income rises. The self-employed tend to pay less social contributions because of a lower tax base and exemptions.

Table 4 Household statutory average net tax rates in Lithuania, net taxes as a share of gross employment income

2*percentile	Net taxes			Taxes		Public transfers	
	All	Employees	Self-employed	Employees	Self-employed	Employees	Self-employed
0–7							
−20	−1.121	−0.106	0.171	0.336	0.340	0.442	0.169
−30	−0.018	0.076	0.104	0.369	0.303	0.293	0.199
−40	0.130	0.168	0.104	0.379	0.314	0.211	0.210
−50	0.222	0.237	0.138	0.395	0.287	0.158	0.149
−60	0.270	0.281	0.200	0.400	0.286	0.119	0.086
−70	0.293	0.301	0.277	0.400	0.318	0.099	0.041
−80	0.326	0.335	0.226	0.411	0.293	0.076	0.068
−90	0.355	0.366	0.257	0.412	0.300	0.046	0.043
−100	0.367	0.393	0.238	0.419	0.276	0.026	0.037

Percentiles are sorted by gross employment income (includes social contributions). Taxes include income tax and social contributions. Public transfers include old-age, disability, unemployment, and other benefits. Net taxes are taxes minus public benefits. Gross employment income is taken from EU-SILC, while all other figures are estimated by EUROMOD, which takes into account various individual and household characteristics (e.g., age, health status). Number of observations per decile is available in Table A1 in Appendix.

self-employed also receive less benefits but are not required to pay higher taxes. As a result, the richest employee households pay 39.3% for their income in tax, while the self-employed households pay 23.8%.

Similar observations can be made when considering marginal net tax rates. Statutory marginal net tax rates increase from 39% to 43% for employee households, whereas they fluctuate around 25% for most self-employed households. Two observations, in particular, are worth mentioning. The first is that the self-employed in the sixth decile face marginal taxes as high as 46%. This is partly related to public transfers that are capped at these levels. The second observation is that business certificates are no longer allowed at such high levels, and income composition changes. If we remove all households that have both self-employed and employee incomes and remove households with business certificates, the marginal statutory tax rates fluctuate between 24% and 36% for the self-employed.

Our results may be influenced by income concepts and definitions used. In particular, taxes are applied to individuals and not to households in Lithuania, so it is important to consider individuals in the analysis instead of equivalized households. As mentioned, EU-SILC data bundles several income components at the household level, most notability income taxes and social contributions, which are difficult to disentangle. Thus, while EUROMOD can model individual level taxes, the same does not apply to effective taxes.[14]

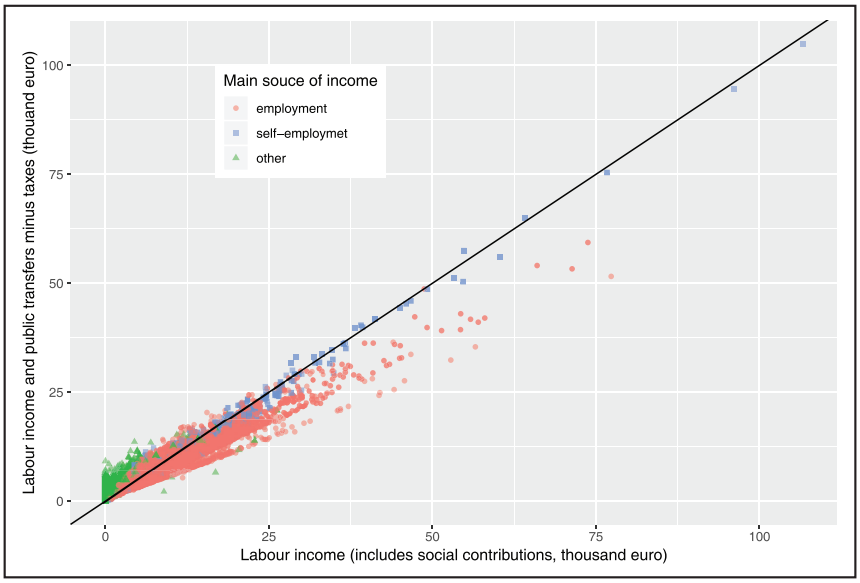
14 As a validity check, we restrict the sample of households to those where there is only one person aged 18–62, although other individuals can also live there. We label this type of households as single households, since this removes the issue of married households. The results presented in Table A3 in Appendix are similar to our previous analysis even though the small sample size requires smaller bins. In particular, tax rates are very similar in magnitude. Public transfer rates, however, became larger when considering single households, as these households tend to contain pensioners or dependents. If we further restrict households to strictly one person households, where that one person must be aged 18–62, tax rates remain similar but public transfer rates become closer to those observed in Table 4. Moving to single households both removes the issue of focusing on households instead of individuals and removes equivalization effects. Unfortunately, there would be too few observations to allow reporting. Nevertheless, it seems that the results are not significantly affected by the choice of the concepts and we proceed further with our initial ones.

4 Effective Net Tax Schedule

We estimate effective equivalized net tax schedule for Lithuania in a similar fashion as was done for statutory tax section. We use EU-SILC data for the period 2014–2015 and simply compare the net taxes that each household paid with the gross employment income that each household received. The vast majority of net taxes paid by households in EU-SILC comes from administrative sources and therefore represents effective taxes paid. Gross employment income in the EU-SILC represents actual income, rather than the income that the tax authorities observe.[15] We find that there is little difference between the statutory and effective net tax schedules for employee households, but the self-employed households pay even less net tax than statutory rates predict. For example, self-employed households effectively pay 29 percentage points less net tax on average than employee households at the top of the income distribution.

As in Section 3, we plot gross employment income against net labor income for different employment groups in Figure 4. In many respects, the effective graph depicting effective tax

Figure 4 Effective equivalized incomes of households grouped using the income definition in Lithuania.



Equivalized gross employment income and net labor income is in thousands of Euro per year. Households are allocated to employment groups according to the income definition for the 2014–2015 income reference years and are represented by dots in the graph (see Section 2). The diagonal line illustrates that what a household earns from employment is what it receives as its net labor income. Any dot above the diagonal line illustrates that the household receives additional benefits, while dots under the horizontal line means that the households pay additional taxes or social contributions.

15 This is because in producing the EU-SILC data for Lithuania, households are asked to report their gross employment income in the questionnaire. Gross employment income is also taken from administrative records for the same household. The two sources (administrative and survey) are compared for each household by the EU-SILC team, and only the larger value of gross employment income is kept in the EU-SILC data that is available to us. Therefore, if respondents revealed more gross employment income in the questionnaire than to authorities, a gap arises between the effective and statutory net tax schedules.

Table 5 Household average effective net tax rates in Lithuania, net taxes as a share of gross employment income

2*percentile	Net taxes			Taxes		Public transfers	
	All	Employees	Self-employed	Employees	Self-employed	Employees	Self-employed
0–7							
−20	−0.812	−0.095	−0.113	0.320	0.122	0.416	0.235
−30	−0.041	0.041	−0.101	0.331	0.130	0.290	0.231
−40	0.074	0.119	−0.144	0.327	0.167	0.207	0.310
−50	0.191	0.205	0.006	0.352	0.141	0.147	0.135
−60	0.219	0.241	−0.005	0.365	0.103	0.123	0.108
−70	0.242	0.267	0.083	0.364	0.124	0.097	0.041
−80	0.278	0.297	0.048	0.373	0.106	0.076	0.058
−90	0.313	0.336	0.057	0.387	0.098	0.050	0.041
−100	0.313	0.359	0.070	0.385	0.101	0.026	0.032

Percentiles are sorted by gross employment income (which includes social contributions). Taxes include income tax and social contributions. Public transfers include old-age, disability, unemployment, and other public benefits. Net taxes are taxes minus public benefits. All figures are taken from EU-SILC and are weighted to include only those households with at least one member aged 18–62 and who is not a student. The number of observations per decile is available in Table A2 in Appendix.

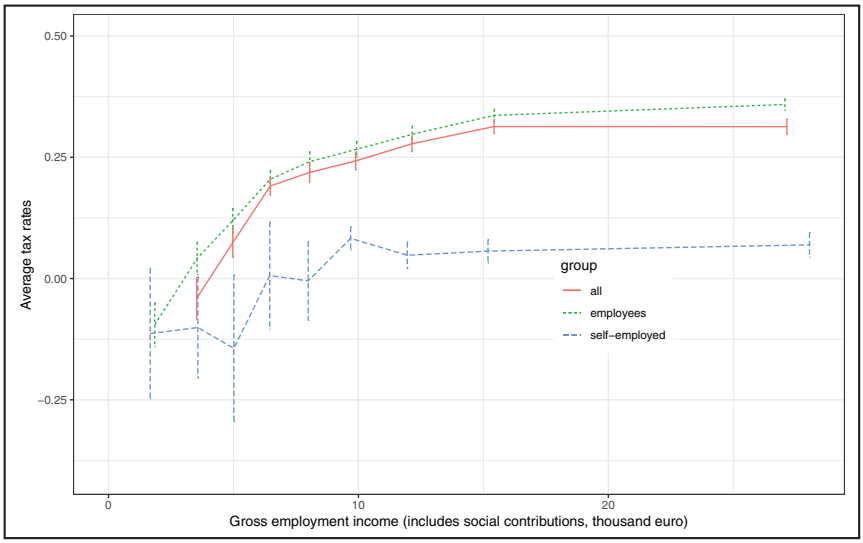
schedule is similar to Figure 3 depicting the statutory tax schedule. The main difference is that self-employed households receive even greater net labor income than employee households. Furthermore, in Figure 4, for a large number of households, gross employment income is equal to or even above the diagonal line, irrespective of the amount of gross employment income they earn.

The effective average net tax rates for the self-employed are much below the statutory rates. The top decile of the self-employed pay 7% of their gross employment income as net taxes, as shown in Table 5, even though statutory rates suggest that they should be paying 24% (see Table 4). While statutory rates might be somewhat overstating taxes because of carried-forward losses, or other tax-relevant features imperfectly captured by the EU-SILC data, the difference is sufficiently large to be noteworthy. In fact, the main drivers are lower effective taxes and social contributions paid by the self-employed (whereas effective and statutory benefits received by the self-employed are similar). In contrast, the statutory and effective net tax rates for employees are similar. This results in a large effective net tax rate difference between the two groups: effective average net tax rates are up to five times lower for the self-employed as compared to employees. Additionally, self-employed average net tax rates are less progressive: effective average tax rates are flat, with some progressivity coming from public transfers. The lack of progressivity of effective tax rates for the self-employed can be seen in Figure 5.

As in the section of statutory net taxes, our concepts and definitions may influence the results, but moving to an individual analysis is problematic given our EU-SILC data. Nevertheless, the results for this restricted subsample of "single" households, as in Section 3, are in line with the fuller sample—if anything, the difference in net tax rates between employees and self-employed is larger, meaning that our results can be seen as conservative.[16]

16 We test a subsample of households with one household member aged 18–62, but allow older and younger members to coexist in Table A4 in Appendix. We label this as singles' households. Again, tax rates are similar as in the previous analysis, although the difference in tax rates between the employed and the self-employed becomes even more apparent. Public transfer rates (and subsequently net tax rates) become higher in the singles' households because of many co-inhabiting pensioners or dependents. Removing them and keeping only households with a single household member aged 18–62 brings public transfers closer to initial estimates found in Table 5, whereas tax rates are closer to those found in Table A4 in Appendix. These results are not presented due to the fact of there being too few observations.

Figure 5 Net taxes are higher and more progressive for employees than for the self-employed.



Standard errors were compiled taking into account survey design with the help of codes from Goedemé (2013) and Zardo Trindade and Goedemé (2016) and computed using Lumley (2018) R package as described in Lumley (2004). We use 95% confidence intervals.

5 Optimal Net Tax Schedule

In this section, we estimate an optimal net tax schedule for Lithuania. This allows us to evaluate whether the statutory net tax schedule described in Section 3 is in line with the economic fundamentals of the country. Additionally, it could shed light on whether the difference between the effective and statutory rates is likely due to overly high (economically unsustainable) statutory rates, or due to low tax compliance. We use a model developed by Saez (2002); it provides the whole optimal net tax schedule given a number of elasticities, government preferences for redistribution and its budget, and a preexisting income distribution.

5.1 The model

The model is taken from Saez (2002), with the exception that individuals are replaced with households (see Section 2). The model starts by indexing households by $m \in M$. The measure of households on M is denoted by $dv(m)$. The household's utility depends positively on net labor income c, and the chosen occupation $i \in 0,1,..,10$; thus, $u(c,i)$. $i = 0$ denotes unemployed or inactive households. The higher the i, the higher the gross employment income w associated with that occupation and the higher the net labor income. In our study, the i represents the same (pseudo) deciles used in Sections 3 and 4.

The fraction of households choosing i is denoted by $h(c_0, c_1, ..., c_I)$, meaning that households weight the net disposable income associated with each job before choosing the best one for them.

The government chooses the net taxes, T_i, that each household should pay or the benefits it should receive and maximizes welfare:

$$W = \int_M \mu^m u_m(w_i - T_i, i) dv(m),$$

where μ^m are positive weights and subject to a budget constraint (1) described below.

The rest of the derivations are found in Saez (2002), but they eventually lead to a system of three equations that show how the government chooses T to maximize W. In the equations provided below, we go through each of these separately.

$$\sum_{i=0}^{I} h_i T_i = H \tag{1}$$

$$\sum_{i=0}^{I} h_i g_i = 1 \tag{2}$$

$$\frac{T_i - T_{i-1}}{c_i - c_{i-1}} = \frac{1}{\zeta_i h_i} \sum_{j=i}^{10} h_j [1 - g_j - \eta_j \frac{T_j - T_0}{c_j - c_0}] \tag{3}$$

Eq. (1) is the government's budget constraint mentioned previously. H is the per capital government's budget net of redistribution. In the simulation, $h_i(c_i - c_0)$, meaning that each household considers the relative gain in net labor income of becoming employed, $c_i - c_0$.

Eq. (2) is a normalization of the welfare function expressed in terms of social welfare weights. Specifically, g_i denotes the value (in terms of public funds) of giving an additional dollar to a household in occupation i. In other words, the government is indifferent regarding giving one more dollar to a household in occupation i and getting g_i of public funds. The higher the g_i, the happier the government is to give money to this occupation and, assuming the government values redistribution, g_i decreases as i increases. Additionally, g depends on net labor income c, the marginal value of public funds p, and the distributional tastes of the government v as shown in Eq. (4). If c is already equally distributed, then there is less reason to further redistribute and so g should be equal across i's. The higher the p, the more the government values its public funds and the less keen it is to redistribute income. The higher the v, the keener the government is to give money to the poorest members of society instead of to the wealthiest.

$$g_i = \frac{1}{p c_i^v} \tag{4}$$

Eq. (3) defines the optimal net tax schedule of a change in net tax rate for occupation i by a small amount dT. Three effects are at work here, which have to be balanced to reach optimal net tax rates. First, there is the mechanical effect of a change in net tax rate. The rise in T_i causes the government to collect more revenue from all those in occupation i and all richer occupations $i+1, i+2, ..., 10$. This is represented by $\sum_{j=i}^{10} h_j$. Second, we include the effect of social weights, g_i attached to each occupation. This is done by stating that the government values each dollar collected by occupation i at $1 - g_i$, since the government may prefer not taking money from some groups in the first place (e.g., the very poor). Third, it includes two behavioral responses: the extensive response and the intensive response.

The extensive response is captured by the extensive labor supply elasticity (technically, the extensive mobility elasticity),

$$\eta_i = \frac{c_i - c_0}{h_i} \frac{\partial h_i}{\partial (c_i - c_0)} \tag{5},$$

which refers to T_i becoming so large that some people working in i may choose to become unemployed or inactive (i_0). It measures the percentage change in number of employed in occupation i when the difference between net labor incomes of employed in occupation i and unemployed/inactive changes by 1%. For example $\eta_i = 0.5$ means that if $c_i - c_0$ increases by 1%, employment in i will rise by 0.5%.

The intensive response is captured by the intensive mobility elasticity (akin to the intensive labor supply elasticity), as indicated by the following equation:

$$\zeta_i = \frac{c_i - c_{i-i}}{h_i} \frac{\partial h_i}{\partial (c_i - c_{i-i})} \tag{6},$$

which refers to people moving from one occupation to another in search of lower net taxes. It measures the percentage increase in supply of job i when $c_i - c_{i-1}$ is increased by 1%. This specification ignores income effects, or the effect of rising incomes for all occupations simultaneously. In the literature, however, income effects are, in any case, found to have a small impact, according to Saez (2002).

Finally, h_i represents the optimal i distribution given the empirically observed h_i^0 distribution

$$h_i = h_i^0 \left(\frac{c_i - c_0}{c_i^0 - c_0^0}\right)^{\eta_i} \tag{7},$$

where the h_i^0 are reconfigured to account for the extensive response to change in net taxes. Here, c_i^0 represents the actual net income and c_i represents the optimal net income, which is estimated simultaneously with Eqs (1–3). Whenever net taxes are lowered for households of occupation i, so that $c_i - c_0$ becomes bigger, more households should be working in i, given extensive elasticity η_i and actual net incomes $c_i^0 - c_0^0$.

5.2 **The parameters**

There are several parameters that need to be chosen for Lithuania: the labor supply elasticities (or, actually, long-run taxable income elasticities), societies' preferences, and others. We use taxable income elasticities, e_z, defined as

$$e_z = \frac{1-\tau}{z} \frac{\delta z}{\delta (1-\tau)},$$

namely the percent in reported income when the net-of-tax rate increases by 1%. The benefit of this "sufficient" elasticity is to capture directly all behavioral effects or raising taxes, including real responses (e.g., labor supply adjustments), tax avoidance (e.g., claiming deductions or (legal) income shifting between tax bases), and illegal tax evasion behavior (e.g., see Saez et al. 2012). Nevertheless, we also rely on the available labor supply elasticity estimates for Lithuania.

5.2.1 **Elasticities**

We start with choosing (uncompensated) intensive and extensive labor mobility elasticities for (5) and (6), respectively. Income effects are usually found to be small on aggregate (Saez 2002; Bargain et al. 2014), which justifies considering uncompensated labor supply elasticity instead of compensated labor supply elasticity. Additionally, we require different extensive and intensive mobility elasticities for high- and low-income households. If these differ, this should produce a kink in the optimal tax schedule: higher extensive elasticities for low incomes calls for subsidies to the poor.

First, it should be noted that ζ is not observed empirically, but can be calculated as

$$\zeta_i = \frac{\varepsilon_i w_i}{w_i - w_{i-1}}$$

by first estimating

$$\varepsilon_i = \frac{1-\tau}{w}\frac{\delta w}{\delta(1-\tau)}$$

where ε shows how much wage responds to the net-of-tax rate change.

Second, as the magnitude of elasticities is uncertain, Saez (2002) proposed a wider range of εs and ηs for the upper and lower tail of distribution based on the summary of literature (see Table 6). Unfortunately, the ranges are large, are based mainly on US data, are ambiguous about being short- or long-run elasticities, and refer to labor supply responses only (i.e. are not elasticities of taxable income). This has been partly remedied by newer studies.

Barrios et al. (2019) estimated Lithuania's short-run labor supply elasticity as

$$e_h = \frac{w}{h}\frac{\delta h}{\delta w} \quad (8),$$

denoting a percent change in net-wage on the number of hours worked, to be between 0.15 for high-skill individuals and 0.3 for low-skill individuals. This elasticity captures the main behavior effect: the real response of labor employment and work duration (the sum of ε and η). While there are no estimates for Lithuania's intensive, e_{hi}, and extensive, e_{he}, margins, Bargain et al. (2014) study these distributions across income quantiles countries largely comparable to Lithuania, such as Estonia, Hungary, Finland, and Poland. For the four countries, the extensive labor elasticities for the lower quantiles, e_{hel}, ranges between 0.08 and 0.26 (an exception is Finland, with 0.8). For the higher end, e_{heh} ranges between 0.05 and 0.23. For the same four countries, intensive labor elasticities range between 0 and 0.03 for the lower e_{hil}, and −0.04 and 0.03 for the higher e_{hih} deciles. The extensive elasticity was found to vary between 0.3 and 0.65 in Staehr (2008) for Estonia, while intensive elasticity was negligible. This suggests that for Lithuania,

Table 6 Intensive and extensive elasticities as proposed by Saez (2002)

	High income ($w \geq$ 20,000 $)	**Low income ($w <$ 20,000 $)**
η	0	[0–1]
ε	[0.25–0.5]	[0.25–0.5]

The table indicates a range of possible elasticities for the United States.

also, most of the labor supply would come from the extensive margin for both the lower and higher income households, even though there may not be large differences between the upper and bottom income distributions.[17]

Lithuania's long-run labor supply elasticity could be much higher, and long-run taxable income elasticities are larger still. We opt for long-run elasticities to capture long-run effects on the economy. Jäntti et al. (2015), who has access to long-term data for largely Scandinavian countries, find e_{he} to range between 0 and 0.4, while e_{hi} ranges between 0 and 0.28. This suggests that a fair long-run range for Lithuania's e_h is 0.1 to 0.7. It is expected that $e_z \geq e_h$. Empirical studies such as Jongen and Stoel (2019) for the Netherlands show that e_h is only 0.05, while e_z is 0.21 in the long run. Lithuania's long-run elasticity of taxable income should also have a similar range, but is more likely to be from 0.2 to 0.8, with the most likely elasticities at 0.5 at the top and the bottom of the income distribution (the intensive margin more relevant for the top and the extensive margin for the bottom). This falls within the range of e_z estimates, although it exceeds the average of 0.3 (Neisser 2017).

One reason for the larger e_z in Lithuania could be the tax system. The narrower the tax base, hence many tax avoidance possibilities, the higher is the elasticity (Saez et al. 2012). The statutory net tax of Lithuania shows that avoidance possibilities exist, especially for the self-employed. Another reason could be the low level of law enforcement. The large shadow economy in Lithuania suggests that tax rules there are not enforced sufficiently. The final list of e_z is presented in Table 7. We assumed that the high income corresponds to 12,000 Euro.

While elasticities in Table 7 apply to the general population, which is dominated by employed households, it does not necessarily apply to average self-employed households. For instance, tax evasion can be higher among the self-employed, since they are not subject to third-party reporting. Indeed, the elasticities for the self-employed are found to be up to three times larger in Spain (Almunia and Lopez-Rodriguez 2019) and in Poland (Zawisza 2019). Other studies also show that elasticities of self-employed income are roughly two times higher than for other types of income (Neisser 2017). However, since we have no available elasticities for Lithuania, we leave this for future work.

Table 7 Ranges of elasticities of taxable income for Lithuania

	High income ($w \geq$ 12,000 Euro**)**	**Low income ($w <$ 12,000** Euro**)**
η	[0.2, **0.3**, 0.5]	[0.2, **0.4**, 0.6]
ε	[0.1, **0.2**, 0.3]	[0.02, **0.1**, 0.2]

Note: non bolded values represent intervals while bolded variables represent the chosen point estimate.
The preferred taxable income elasticities for Lithuania are represented in bold characters whereas the range of possible elasticities is in brackets. *w* is equivalized employment income, which includes employer's and employee's social contributions.

17 The unresponsiveness of elasticities to income deciles was explained in a more recent study for Slovakia by Siebertová et al. (2015). There, e_{heh} falls to 0.06 e_{heh} from 0.16 e_{hel} when only prime age workers are considered, but not when a larger share of older workers are included. For Lithuania, then, where pensions are relatively low compared to the average wage, potential pensioners are also more likely to respond strongly to wages.

5.2.2 **Society's preferences and other parameters**

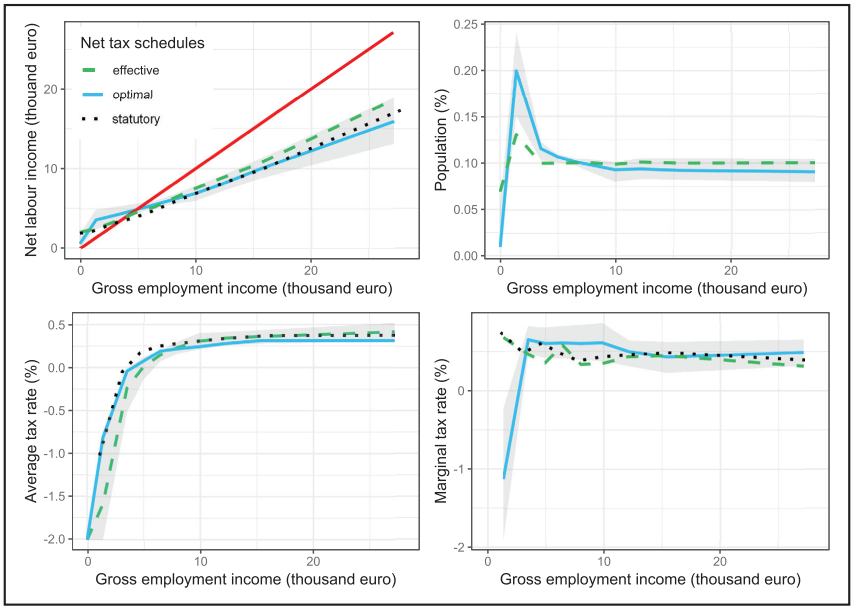
Another parameter is the society's preference parameter *v*. Saez (2002) in most cases used $v = 1$, which already has a high preference for redistribution, while $v = 0.25$ would be a lower point estimate. According to surveys, 92% of Lithuanians believe income inequality is too high, Lithuania being one of the leading countries in the EU. Additionally, Lithuania's government explicitly tries to reduce poverty and income inequality (LR Vyriausybe 2017). Therefore, v should be clearly positive and relatively high. We set $v = 1$ in the baseline and $v = 0.7$ as an alternative scenario.

The other parameters are derived from EU-SILC data itself. $H = 2{,}199$ as this was the sum of net transfers from the EU-SILC survey, c_i^0, and h_i^0 was taken from the EU-SILC survey as well. $i = 1, 2, ..., 10$ so that each occupation constitutes about 10% of population, although the first bin is smaller, so that $w_0 = 0$.

5.3 **The simulations**

Given the model and the parameters, we utilize an R-package by Hasselman (2018) to run the simulations for Lithuania. We obtain four key variables: net labor income, population distribution by income, and average and marginal net tax rates. Information about each variable is presented in four graphs in Figure 6 and Table 8. In each graph, the preferred parameter

Figure 6 Optimal, effective, and statutory net tax schedules.



In each graph, the optimal net tax schedule with the preferred parameter specification (see Table 7) is depicted by a blue line while alternative parameter choices are presented as a shaded area around the blue line. The green dashed line and the black dotted line represents the variables distributions in line with the effective and statutory tax schedules respectively. The diagonal red line on the top-left figure is a 45-degree line depicting zero net taxes.

Table 8 Effective and optimal variables for Lithuania

2*percentile	Gross employment income	Net labor income[a]	Net labor income[b]	Percent of households[a]	Percent of households[b]	Average tax rate[a]	Average tax rate[b]	Marginal tax rate[a]	Marginal tax rate[b]
0–7	0.0	2.1	0.6	6.9	0.9				
−20	1.4	2.5	3.6	13.0	20.0	−81.2	−158.5	67.7	−112.4
−30	3.5	3.7	4.4	10.0	11.5	−4.1	−23.5	46.4	64.8
−40	5.0	4.6	4.9	10.0	10.7	7.4	0.8	35.4	60.1
−50	6.5	5.2	5.5	10.0	10.1	19.1	14.7	57.9	60.7
−60	8.0	6.3	6.2	10.0	9.8	21.9	23.6	33.4	60.1
−70	9.9	7.5	6.9	9.9	9.3	24.2	30.5	34.6	60.9
−80	12.1	8.8	8.0	10.1	9.4	27.8	34.1	43.3	49.7
−90	15.4	10.6	9.9	10.0	9.2	31.3	36.0	44.5	42.9
−100	27.1	18.6	15.9	10.0	9.1	31.3	41.4	31.3	48.7

[a]Effective variable.

[b]Optimal variable.

Gross employment income and net labor income are in thousand Euro per equivalized household in Lithuania in the period 2014–2015. Share of households, average tax rates, and marginal tax rates are in percentages. Number of observations per decile is available in Table A2 in Appendix.

specification is depicted by a blue line, and alternative parameter choices are presented as a shaded area around the blue line. The green dashed line represents the effective net tax schedule, and the black dotted line is the statutory tax schedule. Let us go through what messages each graphs suggest in turn.

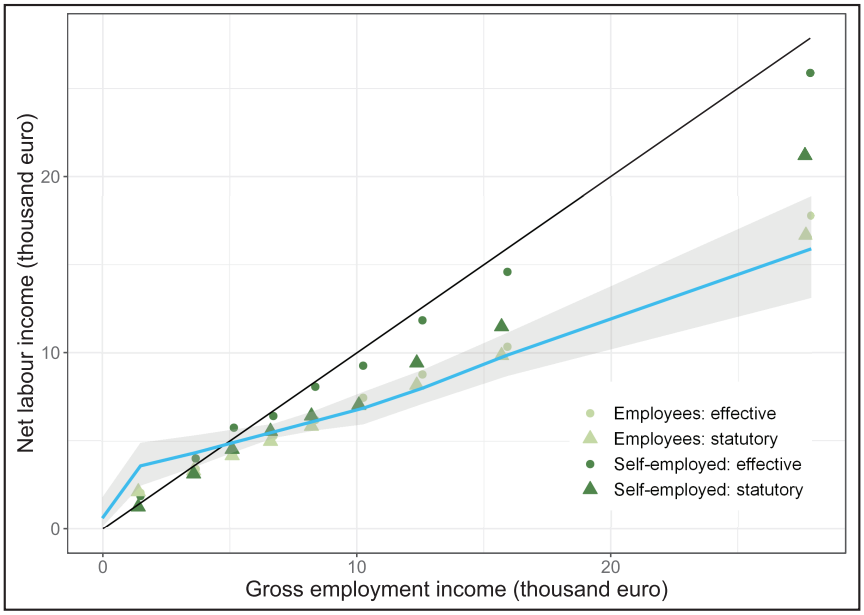
The effective and statutory net tax schedule coincides with the optimal net schedule for the middle of the income distribution, but less for the tails. The figure on the top-left holds the transformation from gross employment income to net labor income. Effective net labor income and statutory schedule coincides with the optimal net labor income for middle (gross employment) incomes, and, in most cases, falls within the range of optimal schedules. At higher incomes, the optimal net labor income is slightly below the net labor income of the statutory and well below the effective net tax schedules. For those earning little gross employment income, the optimal tax rates suggest that more can be done to increase labor market participation and reduce unemployment: less income should be directed to the very poorest and in-work credits should be provided. Unemployment and nonparticipation would then drop (from 6.9% to close to 0.9%) while the share of households employed at lower income levels rises (from 13.0% to 20.0%) because of more in-work credits, as illustrated by the top-right figure. The unemployment and nonparticipation drop should be taken with caution. The optimal net tax model does not distinguish between work capacity and household preferences. For example, some households may suffer from severe disability or wish to attend to their own children. In these cases, it may not make sense to fully remove benefits or expect that in-work incentives would encourage these people to work.

Effective/statutory average and marginal tax rates are close to their optimal levels in the middle of the income distribution, but not the tails. Optimal marginal tax rates for the bottom deciles are strongly negative: 112.4% of their gross employment income. This contrasts markedly with the effective positive 67.7% marginal tax rate for the bottom deciles. Additionally, the optimal marginal tax rate for the top of the gross employment income distribution is 48.7% while the effective marginal tax rate is 31.3% or about 11% below statutory. Empirical studies suggest that optimal tax rates tend to be much higher than statutory rates at top incomes. Saez (2002) shows that the majority of estimates of optimal tax rates for top incomes for the United States lie above 50%. Klemm et al. (2018) also find that the top optimal marginal tax rates exceed 50% and tend to be 10—but sometimes even 30—percent points above the statutory marginal tax rates in 27 countries. Therefore, the 11% difference is on the lower side of the estimates. Part of the reason for the gap is the large extensive labor elasticity in Lithuania for top incomes, which prevents taxing high incomes too high. Another reason is a large presence of self-employed individuals.

6 Statutory, Effective, and Optimal Net Tax Schedules for Employees and the Self-employed

Here, we compare statutory, effective, and optimal (equivalized) net tax schedules for employee and self-employed households. The three net tax schedules coincide more for employees than the self-employed. This can be seen in Figure 7, where the two groups are distinguished. The effective and statutory net tax schedules for employees lie close to the optimal tax schedule,

Figure 7 Statutory, effective, and optimal equivalized tax schedules for households grouped according to income.



The graph illustrates how household's equivalized gross employment income translates into equivalized net labor income for three tax schedules: effective, statutory, and optimal. The effective and statutory tax schedules are presented as points (for each decile) for employees and the self-employed. The optimal tax schedule is calculated for the total population aged 18–62, and excludes students. The shaded area around the dashed blue line illustrates a range of optimal tax schedules using a range of parameters as shown in Table 7. The data comes from EU-SILC, the simulation was carried out with the help of EUROMOD, and the optimal tax schedule was computed along the lines of Saez (2002).

while the self-employed are further away—in most cases, outside of the optimal net tax schedules range. The self-employed are subject to lower statutory net tax rates, which exceed the range of optimal net tax schedules for higher gross employment income deciles. Additionally, the self-employed effectively pay even lower effective tax rates than they are required. This holds true for the whole gross employment income distribution. There is also a smaller difference between the two groups at the bottom. The self-employed face relatively higher net tax rates than employees due to lower public transfers (compare Tables 4 and 5).

There are several possible ways to explain the large gap between the effective and statutory net tax schedules for the self-employed. The most likely explanation is tax evasion. In Lithuania, there is a tendency to underreport self-employment income or not declare being self-employed at all, as previously noted by Navicke and Cižauskaite (2018). Assuming that survey respondents are more willing to reveal their true self-employed incomes in questionnaires, we can compare the effective and statutory tax rates to obtain an estimate for evaded taxes in Lithuania, as done in Table 9. Employee households may not pay up to about 5.6% to 14.4% of their taxes, while the self-employed may evade as much as 69.9%, depending on the gross employment income distribution. Assuming that missing taxes arise from underreported income, we see that these numbers are high, but plausible,

Table 9 Estimated difference between statutory and effective tax schedules per equivalized household per year

2*percentile	% of statutory tax		Th. Euro	
	Employees	Self-employed	Employees	Self-employed
−7				
−20	4.76	64.12	0.02	0.28
−30	10.30	57.10	0.14	0.62
−40	13.72	46.82	0.27	0.75
−50	10.89	50.87	0.28	0.96
−60	8.75	63.99	0.29	1.50
−70	9.00	61.01	0.36	1.96
−80	9.25	63.95	0.47	2.33
−90	6.07	67.33	0.39	3.17
−100	8.11	63.41	0.94	4.83

The figures are derived from the difference between statutory and effective average tax rates from Tables 4 and 5, respectively. Percentiles are sorted according to the equivalized household gross employment income of all nonstudents aged 16–62.

given the empirical literature. An estimate for Lithuania is found in Kukk et al. (2019), who estimated income underreporting of the self-employed in surveys to be around 25% to 30%, depending on the definition of "self-employed." The study, however, uses the consumption approach to estimate tax evasion, which should give a lower bound of underreporting estimates. In addition, income underreporting in surveys does not necessarily mean that people equally underreport income to authorities. For example, the same study estimated that, in Estonia, the self-employed underreport 22% of their income, while Paulus (2015) estimated that as much as 71% of self-employment income is unreported to authorities, which is what matters for tax collection. Estimates from other countries are generally in line with what we expect given our results. Paulus (2015) finds that, in Estonia, up to 20% of employees underreport income. Paulus (2015) also finds that underreporting is greatest at the tails of the income distribution, something also found by Johns and Slemrod (2010) for the United States. While there is greater underreporting at the lower percentiles for Lithuania, the message is less clear for the top. However, this may be due to the failure to capture top incomes in the survey for Lithuania. Many more studies find that the self-employed evade much more taxes than employees by underreporting income. Baldini et al. (2009) find that, in Italy, the self-employed tend to evade more income tax than employees do. Pissarides and Weber (1989) find that the self-employed in the UK actually have 1.55 times the reported income, meaning that they underreport income by 35% in the UK, while Slemrod (2016) cites International revenue service (IRS) studies in the United States, where 56% of income may be unreported for the self-employed. A study by Artavanis et al. (2016) in Greece shows that the self-employed in certain professions, such as doctors, lawyers, engineers, and scientists, as well as accountants and financial service agents, underreport more than half of their income.

Even though tax evasion is a likely explanation for the difference between effective and statutory tax rates for the self-employed, it is also reasonable to assume that some of this difference is due to measurement error. However, it is not clear if in aggregate the error under- or

overestimates the difference. First, EUROMOD does not model all taxes and contributions, which would result in lower tax evasion. Second, there might still be some income, particularly self-employment income, that is not reported to the authorities and not revealed in the questionnaire, which would mean greater tax evasion.

The difference in statutory rates between employees and the self-employed could be accounted for in several ways. For example, the government may perceive the self-employed more favorably than employees. There could be at least two reasons for this. One is that the self-employed would not be able to become employees, and this scenario is better than being unemployed. A second reason is that the government believes that the self-employed tend to contribute more to society, either by themselves producing significantly more earnings due to lower taxes, by supporting the rest of the economy by being entrepreneurs and eventually hiring more labor, or by producing other positive externalities (see Scheuer and Slemrod 2019). However, the first theory does not stand up to the data and the literature, while the second lacks credible evidence. Regarding the first reason, the self-employed are bunched at the top of the income distribution. If these households tend to earn high incomes, it is not clear why they could not become employees or pay higher taxes as self-employed. Regarding the second reason, a minority of the self-employed, according to EU-SILC, could be considered entrepreneurs and less than 10% of self-employed at the top of income distribution have employees of their own.

This leaves the possibility that the self-employed are especially responsive to tax rate changes or bring about large positive externalities—something that has not yet been tested for Lithuania. At the same time, a review of the literature suggests that a major reason for becoming self-employed is not entrepreneurship, but greater tax evasion/avoidance opportunities (Baliamoune-Lutz and Garello 2014). Additionally, the empirical literature is mixed concerning whether the self-employed respond to tax changes, thereby placing lower statutory rates into question (Baliamoune-Lutz and Garello 2014). For example, Bruce (2002) shows that higher statutory tax rates on self-employed income in the United States did not lead to the closing of small businesses. On the contrary, higher proportional taxes on the self-employed, together with the possibility of offsetting losses, actually encourages entrepreneurship via a risk-sharing channel, as first explained by Domar and Musgrave (1944) and later found in empirical work (e.g., Baliamoune-Lutz and Garello 2014). Apparently, it is progressive self-employment taxes that seem to deter self-employment, as shown by Gentry and Hubbard (2000) for the United States and by Baliamoune-Lutz and Garello (2014) in Europe.

7 Conclusions, Limitations, and Recommendations

We compared the statutory, effective, and optimal net tax schedules for Lithuania for the period 2014–2015. We did this for all Lithuanian households and then looked at employee and self-employed households separately to investigate different forms of employment.

We found that the three schedules largely coincide for the middle of the income distribution for all households. The three diverge, however, at the tails of the income distribution. At the bottom of the income distribution, the optimal net tax schedule suggests that

more in-work benefits should be provided for the least paid, to encourage employment. At the top of the income distribution, more effort could be made to extract tax revenue to improve tax compliance. The results for employee households were similar to that of all households.

We found that the three net tax schedules coincide more for employee households than for self-employed households. Except for those at the very bottom of the income distribution, the self-employed are subject to lower statutory net tax rates and very low progressivity, as compared to employees. Unfortunately, using the same elasticities for the employed and the self-employed does not allow us to draw strong conclusions about optimal taxes for the self-employed. Nevertheless, the self-employed do effectively pay much lower taxes than the statutory tax schedule would suggest. This holds throughout the income distribution and could mean that as much as 70% of self-employed taxes are not paid.

Our conclusion can be viewed as a conservative one. If we were to exclude pension contributions or consider all social contributions as generating actuarially fair benefits, the inadequacy in taxation levels would likely be even larger. The divergence would be greater still if we were to consider income taxes only, and not social contributions or benefits. Additionally, we considered a budget-neutral tax schedule. Finally, the fact that statutory rates differ substantially can explain why optimal taxes are also relatively low. Were there fewer opportunities to avoid taxes by having a broader tax base, measured elasticities would be smaller and optimal taxes would be higher.

As this is an initial step in comparing the three schedules, there are ways to improve the estimates. First, the EU-SILC is known to poorly capture top incomes; greater access to administrative data could help solve this problem. Second, the fact that the statutory tax schedule differs from the effective tax schedule for the self-employed means that the household misreport their employment status and incomes to the authorities, EU-SILC, or both. Third, we were not able to find Lithuania-specific long-run estimated elasticities, meaning that the current ones had to be taken from other studies. Nonetheless, such elasticities can be eventually estimated, particularly as a large income tax reform took effect in 2019. Obtaining taxable income elasticities for the self-employed and the employed separately would be especially beneficial. Fourth, one may consider a different set of elasticities or/and preferences for the optimal net tax schedules of employees and the self-employed. For example, society could value the self-employed more, or they themselves could be more responsive to wages.

The findings presented in this article point to several recommendations.

First, the effective net tax schedule indicates that less taxes and social contributions are collected than households are statutorily required to provide. Therefore, more efforts can be placed on the auditing of households, especially at the upper tail of the income distribution, to extract more government revenue. Before doing so, the marginal cost of the audit and the marginal value of public funds should be estimated.

Second, the optimal net tax schedule recommends providing tax credits to those who receive low wages. Upon obtaining better estimates of the bottom of the distribution, this policy could be considered further. This is especially relevant with the resurgence of discussions on universal incomes, which counters in-work credit suggestions.

Third, the optimal tax schedule recommends fewer benefits to unemployed and non-active households. With the combination of lower out of work benefits and higher tax credits, households would be more inclined to seek employment. However, one would first have to consider at least the health and preferences of households, as many benefits relate to health, disability, and children.

Fourth, the benefits of the current lower statutory taxes for the self-employed should be closely weighted alongside the associated costs of lower tax revenue. As the majority of the self-employed are found at the upper tail of the income distribution, a great deal of tax revenue is not collected. Furthermore, international evidence shows that some companies start hiring and individuals start choosing self-employment purely for the purpose of paying less tax. In such cases, it may be in the general interest to raise statutory tax rates for the self-employed closer to, or even above, the tax rates of employees.

Declarations

Availability of Data and Materials

The results presented here are based on EUROMOD (version I1.0+). EUROMOD is maintained, developed, and managed by the Institute for Social and Economic Research (ISER) at the University of Essex, in collaboration with national teams from the EU member states. We are indebted to the many people who have contributed to the development of EUROMOD. The process of extending and updating EUROMOD is financially supported by the European Union Programme for Employment and Social Innovation "EaSI" (2014-2020). The results and their interpretation are the author's(') responsibility. This paper is based on data from Eurostat, EU Statistics on Income and Living Conditions [2015:2016], and RPP 128/2019-EU-SILC. The responsibility for all conclusions drawn from the data lies entirely with the author(s). Data are available from the authors upon reasonable request and with permission of Eurostat and ISER only.

Competing Interests

The authors declare that they have no competing interests.

Funding

Jousten acknowledges financial support from FNRS through grant J.0177.18. Work on this paper began when Alain Jousten was on sabbatical leave with ISM (Vilnius) and the IMF (Washington, DC) while Nerijus Cerniauskas was working for the Bank of Lithuania and studying and working in Vilnius University.

Authors' Contributions

AJ supervised the project, and NC carried out the computations. Both authors analysed the data and have co-written the draft and the final version of the paper. Both authors agreed on the drafts and the final version of the paper.

Acknowledgements

We are grateful to Emmanuel Saez for sharing the optimal tax codes, for Benjamin Hemingway's help with the modeling, for the patience of the Bank of Lithuania's Economics Department, and for comments provided by the anonymous referees, the participants at Bank of Lithuania's internal research seminar, and Vilnius University's PhD seminar. The views expressed in this paper are those of the authors and do not necessarily represent those of the Bank of Lithuania or the Eurosystem.

References

Almunia, Miguel; David Lopez-Rodriguez (2019): The elasticity of taxable income in Spain: 1999–2014. SERIEs, 10(3), 1–40.

Alstadsæter, Annette; Niels Johannesen; Gabriel Zucman (2017): Tax Evasion and Inequality. NBER Working Paper 23772. Cambridge: National Bureau of Economic Research.

Artavanis, Nikolaos; Adair Morse; Margarita Tsoutsoura (2016): Measuring income tax evasion using bank credit: Evidence from Greece. The Quarterly Journal of Economics, 131(2), 739–798.

Baldini, Massimo; Paolo Bosi; Michele Lalla (2009): Tax evasion and misreporting in income tax returns and household income surveys. Politica Economica, 25(3), 333–348.

Baliamoune-Lutz, Mina; Pierre Garello (2014): Tax structure and entrepreneurship. Small Business Economics, 42(1), 165–190.

Bargain, Olivier; Kristian Orsini; Andreas Peichl (2014): Comparing labor supply elasticities in Europe and the United States New Results. Journal of Human Resources, 49(3), 723–838.

Barrios, Salvador; Viginta Ivaškaite˙-Tamošiu˜ne˙; Anamaria Maftei; Edlira Narazani; Janos Varga (2019): Progressive tax reforms in flat tax Countries. EUROMOD Working Paper Series 2. Essex: University of Essex.

Bruce, Donald (2002): Taxes and entrepreneurial endurance: Evidence from the self-employed. National Tax Journal, 5–24.

Callan, Tim; Karina Doorley; Alyvia McTague (2020): Top incomes in Ireland: Reconciling evidence from tax records and household survey data.

Chetty, Raj; Michael Stepner; Sarah Abraham; Shelby Lin; Benjamin Scuderi; Nicholas Turner; Augustin Bergeron; David Cutler (2016): The association between income and life expectancy in the United States, 2001–2014. Jama, 315(16), 1750–1766.

Chiappori, Pierre-André; Costas Meghir (2015): Intrahousehold inequality. In *Handbook of income distribution*, 2, 1369–1418. Elsevier.

Domar, Evsey D; Richard A Musgrave (1944): Proportional income taxation and risk-taking. The Quarterly Journal of Economics, 58(3), 388–422.

Eurostat (2018): EU Statistics on Income and Living Conditions (EU-SILC) Methodology – Concepts and Contents, http://ec.europa.eu/eurostat/statistics- explained/index.php?title=EU_statistics_on_income_and_living_conditions_(EU-SILC)_methodology_%E2%80%93_concepts_and_contents#Adjusted_cross_sectional_weight_.28RB050a.29.

Frölich, Markus; David Kaplan; Carmen Pagés; Jamele Rigolini; David Robalino (2014): Social insurance, informality, and labour markets: How to protect workers while creating good jobs. No. 7879. Oxford University Press.

Gentry, William M; R Glenn Hubbard (2000): Tax policy and entry into entrepreneurship. American Economic Review, 90(2), 283–287.

Goedemé, Tim (2013): How much confidence can we have in EU-SILC? Complex sample designs and the standard error of the Europe 2020 Poverty Indicators. Social Indicators Research, 110(1), 89–110. doi: 10.1007/s11205-011-9918-2

Guner, Nezih; Remzi Kaygusuz; Gustavo Ventura (2014): Income taxation of US households: Facts and parametric estimates. Review of Economic Dynamics, 17(4), 559–581.

Hasselman, Berend (2018): Nleqslv: Solve Systems of Nonlinear Equations. R package version 3.3.2, https://CRAN.R-project.org/package=nleqslv.

Heathcote, Jonathan; Kjetil Storesletten; Giovanni L Violante (2017): Optimal tax progressivity: An analytical framework. The Quarterly Journal of Economics, 132(4), 1693–1754.

Hlasny, Vladimir; Paolo Verme (2018): Top incomes and inequality measurement: A comparative analysis of correction methods using the EU SILC data. Econometrics, 6(2), 30.

Huang, Jason; Juan Rios (2016): Optimal tax mix with income tax non-compliance. Journal of Public Economics, 144, 52–63.

Immervoll, Herwig; Henrik Jacobsen Kleven; Claus Thustrup Kreiner; Nicolaj Verdelin (2011): Optimal Tax and Transfer Programs for couples with extensive labor supply responses. Journal of Public Economics, 95(11–12), 1485–1500.

Jäntti, Markus; Jukka Pirttilä; Håkan Selin (2015): Estimating labour supply elasticities based on cross country micro data: A bridge between micro and macro estimates? Journal of Public Economics, 127, 87–99.

Johns, Andrew; Joel Slemrod (2010): The distribution of income tax noncompliance. National Tax Journal, 63(3), 397.

Jongen, Egbert L. W; Maaike Stoel (2019): The elasticity of taxable labour income in the Netherlands. De Economist, 167(4), 359–386. doi: 10.1007/s10645-019-09349-7

Klemm, Alexander D; Li Liu; Victor Mylonas; Philippe Wingender (2018): Are elasticities of taxable income rising? IMF Working Paper 18/132. International Monetary Fund.

Kukk, Merike; Alari Paulus; Karsten Staehr (2019): Cheating in Europe: Underreporting of self-employment income in comparative perspective. International Tax and Public Finance, 27(2), 363–390.

LR Vyriausybe (2017): Nutarimas de˙l Lietuvos Respublikos Vyriausybe˙s Programos ¸gyvendinimo Plano Patvirtinimo [Resolution on the Approval of the Program Implementation Plan of the Government of the Republic of Lithuania], https://www.e-tar.lt/portal/lt/legalAct/2389544007bf11e79ba1ee3112ade9bc.

Lumley, Thomas (2004): Analysis of complex survey samples. R package verson 2.2, Journal of Statistical Software, 9(1), 1–19.

Lumley, Thomas (2018): Survey: Analysis of complex survey samples. R package version 3.34.

Leventi Chrysa; Manos Matsaganis; Maria Flevotomou (2013): Distributional implications of tax evasion and the crisis in Greece. EUROMOD Working Paper EM17/13. EUROMOD.

Meriküll, Jaanika; Tairi Room (2019): Are survey data underestimating the inequality of wealth? Technical report. Bank of Estonia.

Milanez, Anna; Barbara Bratta (2019): Taxation and the future of work. OECD Taxation Working Papers 41. Paris. doi: 10.1787/20f7164a-en

Mirrlees, James (1971): An exploration into the theory of optimal taxation. Review of Economic Studies, 38.

Navicke˙, Jekaterina; Aušra Čižauskaite˙ (2018): EUROMOD Country Report – LITHUANIA (LT) 2015–2018. https://www.euromod.ac.uk/sites/default/files/country-reports/year9/Y9_CR_LT_Final.pdf.

Navicke˙, Jekaterina; Romas Lazutka (2016): Functional and personal income distribution in the Baltics: Comparison of National and Households Accounts. In Business and Management 2016: the 9th International Scientific Conference, May 12–13, 2016, Vilnius, Lithuania. Technika.

Neisser, Carina (2017): The elasticity of taxable income: A meta-regression analysis. IZA Discussion Paper 11958. Bonn: IZA Institute of Labor Economics.

OECD (2019): Taxing wages 2019. OECD Publishing. Paris. doi: 10.1787/taxwages-2019-en.

Pahl, Jan (1995): His money, her money: Recent research on financial organisation in marriage. Journal of Economic Psychology, 16(3), 361–376.

Paulus, Alari (2015): Tax evasion and measurement error: An econometric analysis of survey data linked with tax records. ISER Working Paper Series 2015-10. Essex: University of Essex.

Pissarides, Christopher A; Guglielmo Weber (1989): An expenditure-based estimate of Britain's black economy. Journal of Public Economics, 39(1), 17–32.

Saez, Emmanuel (2001): Using elasticities to derive optimal income tax rates. The Review of Economic Studies, 68(1), 205–229.

Saez, Emmanuel (2002): Optimal income transfer programs: Intensive versus extensive labor supply responses. The Quarterly Journal of Economics, 117(3), 1039–1073.

Saez, Emmanuel; Joel Slemrod; Seth H Giertz (2012): The elasticity of taxable income with respect to marginal tax rates: A critical review. Journal of Economic Literature, 50(1), 3–50.

Scheuer, Florian; Joel Slemrod (2019): Taxation and the Superrich. Technical report. National Bureau of Economic Research.

Siebertová, Zuzana; Matúš Senaj, ; Norbert Švarda; Jana Valachyová (2015): To work of not to work? Updated Estimates of Labour Supply Elasticities. CELSI Discussion Paper 32. Bratislava: Central European Labour Studies Institute.

Slemrod, Joel (2016): Tax compliance and enforcement: New research and its policy implications. Ross School of Business Working Paper 1302. Michigan: University of Michigan.

Staehr, Karsten (2008): Microfoundations of Economic Success: Lessons from Estonia. Chap. Estimates of Employment and Welfare Effects of Personal Labour Income Taxation in a Flat-Tax Country: The Case of Estonia, edited by David Mayes, 242–290. Cheltenham: Edward Elgar Publishing.

Törmälehto, VM (2017): High income and affluence: Evidence from the European Union Statistics on Income and Living Conditions (EU-SILC). Statistical Working Papers. Luxembourg: Eurostat.

Vogler, Carolyn; Jan Pahl (1994): Money, power and inequality within marriage. The Sociological Review, 42(2), 263–288.

Zardo Trindade, Lorena; Tim Goedemé (2016): Notes on Updating the EU-SILC UDB Sample Design variables, 2012–2014. CSB Working Paper. Antwerp: University of Antwerp.

Zawisza, Tomasz (2019): Optimal taxation of employment and self-employment: Evidence from Poland and implications.

Appendix

Table A1 Number of observations per decile from EUROMOD output

Percentile	Gross employment income	Total observations	Employed	Self-employed
0–7	0	690	0	0
−20	1,413	784	174	20–49
−30	3,588	649	385	59
−40	5,102	649	462	20–49
−50	6,609	641	543	20–49
−60	8,219	635	559	20–49
−70	10,080	626	557	20–49
−80	12,357	615	545	52
−90	15,690	635	578	20–49
−100	27,651	572	461	107

Data is sorted according to equivalized gross employment income (includes social contributions that are evaluated by EUROMOD). All figures are taken from EUROMOD and are weighted to include only those households with at least one member who is 18–62 years old and is not a student. Deciles are based on weighted observations, which results in different number of observations per quantile; 20–49 indicates that there are between 20 and 49 (inclusive) number of observations, although the number is not publishable due to confidentially reasons. The first 7 percentiles do not have any gross employment income.

Table A2 Number of observations per decile from EU-SILC

Percentile	Gross employment income	Total observations	Employed	Self-employed
0–7	0	680	0	0
−20	1,395	787	300	20–49
−30	3,525	629	501	60
−40	4,972	636	557	20–49
−50	6,470	653	602	20–49
−60	8,047	631	576	20–49
−70	9,888	623	569	20–49
−80	12,141	619	557	55
−90	15,425	625	569	20–49
−100	27,143	576	467	107

Data is sorted according to equivalized gross employment income (includes social contributions). All figures are taken from EUROMOD and are weighted to include only those households with at least 1 member who is 18–62 years old and is not a student. Deciles are based on weighted observations, which results in different number of observations per quantile; 20–49 indicates that there are between 20 and 49 (inclusive) number of observations, although the number is not publishable due to confidentially reasons. The first 7 percentiles do not have any gross employment income.

Table A3 Household statutory average net tax rates in Lithuania, net taxes as a share of gross employment income.

Percentile	Net taxes			Taxes		Public transfers	
	All	Employees	Self-employed	Employees	Self-employed	Employees	Self-employed
0–24							
−50	−0.946	−0.006	−0.012	0.354	0.327	0.361	0.339
	[996]	[281]	[60]	[281]	[60]	[281]	[60]
−75	0.116	0.170	0.003	0.398	0.303	0.229	0.300
	[720]	[599]	[53]	[599]	[53]	[599]	[53]
−100	0.293	0.312	0.207	0.419	0.295	0.106	0.088
	[653]	[577]	[68]	[577]	[68]	[577]	[68]

Our sample is restricted to households with one household member aged 18–62, but can include older and younger household members as well. Percentiles are sorted by gross employment income (includes social contributions). Taxes include income tax and social contributions. Public transfers include old-age, disability, unemployment, and other benefits. Net taxes are taxes minus public benefits. Gross employment income is taken from EU-SILC, whereas all other figures are estimated by EUROMOD, which takes into account various individual and household characteristics (e.g., age, health status). All figures are taken from EU-SILC and are weighted to include only those households with one member aged 18–62 and who is not a student, but older and younger household members may be present. The number of observations per quantile is in []. EU-SILC, European Union Statistics on Income and Living Conditions.

Table A4 Household average effective net tax rates in Lithuania, net taxes as a share of gross employment income. Sample restricted to households with 1 household member aged 18–62, but can include older and younger household members as well

Percentile	Net taxes			Taxes		Public transfers	
	All	Employees	Self-employed	Employees	self-employed	Employees	Self-employed
0–24							
−50	−0.522	−0.064	−0.190	0.316	0.117	0.380	0.307
	[996]	[287]	[55]	[287]	[55]	[287]	[55]
−75	0.142	0.173	−0.115	0.362	0.105	0.188	0.221
	[721]	[605]	[51]	[605]	[51]	[605]	[51]
−100	0.285	0.323	0.018	0.399	0.082	0.076	0.064
	[653]	[576]	[68]	[576]	[68]	[576]	[68]

Percentiles are sorted by gross employment income (which includes social contributions). Taxes include income tax and social contributions. Public transfers include old-age, disability, unemployment, and other public benefits. Net taxes are taxes minus public benefits. All figures are taken from EU-SILC and are weighted to include only those households with one member aged 18–62 and is not a student. The number of observations per quantile is in []. EU-SILC, European Union Statistics on Income and Living Conditions.

3rd publication

Income inequality and redistribution in Lithuania: the role of policy, labour market, income and demographics

N. Černiauskas, C. O'Donoghue, D. Sologon, L. Tarasonis.

Review of Income and Wealth (in press)

NOTES

NOTES

Vilnius University Press
Saulėtekio al. 9, III rūmai, LT-10222 Vilnius
E-mail. info@leidykla.vu.lt, www.leidykla.vu.lt
Print run copies 15.