2019

IARIW-HSE

Special IARIW-HSE Conference "Experiences and Future Challenges in Measuring Income and Wealth in CIS Countries and Eastern Europe" Moscow, Russia, September 17-18, 2019

Changes in Income Inequality in Lithuania: The Role of Policy, Labour Market Structure, Returns and Demographics

Nerijus Černiauskas
Denisa M. Sologon
Cathal O'Donoghue
Linas Tarasonis

Paper Prepared for the IARIW-HSE Conference
Moscow, Russia, September 17-18, 2019
Session 7: Poverty and Wealth
Time: 13:15 – 15:15, September 18

Changes in income inequality in Lithuania: the role of policy, labour market structure, returns and demographics*

Nerijus Černiauskas[†], Denisa M. Sologon[‡], Cathal O'Donoghue[§], Linas Tarasonis[¶]

31st August 2019

Preliminary draft - Please do not distribute

Abstract

This paper uses a method developed by Sologon et al. (2018) to model the household disposable income distribution in Lithuania and explore the drivers of the increase in income inequality between 2007 and 2015. During this period, Lithuania faced the global financial crisis, over which household disposable income fluctuated severely and a series of tax and benefit reforms were implemented. Net emigration accelerated the ageing problem, putting additional strain on the income distribution.The method allows for the quantification of the contributions of four main factors to changes in the disposable income distribution: (i) labour market structure; (ii) returns; (iii) demographic composition; and (iv) tax-benefit system. Our empirical exercise accounts for the distributional effects of the global financial crisis and the subsequent rapid economic recovery. Results show that these effects were substantial and reflected markedly different developments over two periods: changes in the tax and benefit system successfully accommodated a rapid rise in market income inequality due to the crisis during the 2007-2011, but failed during the subsequent years when the rising returns in the labour market significantly increased the disposable income inequality. We also find the demographic effects contribute significantly to the rising inequalities in Lithuania.

Keywords: income distribution, inequality, decompositions, microsimulation, tax-benefit policies, crisis, austerity, overtime comparison

JEL Codes: D31, H23, J21, J31, I38

*This research is part of the SimDeco project (*Tax-benefit systems, employment structures and cross-country differences in income inequality in Europe: a micro-simulation approach*) supported by the National Research Fund, Luxembourg (grant C13/SC/5937475). We are indebted to the many people who have contributed to the development of EUROMOD, which is maintained, developed and managed by the ISER at the University of Essex, in collaboration with national teams from the EU member states. The results and their interpretation are our responsibility.

[†]nerijus.cerniauskas@evaf.vu.lt, Bank of Lithuania and Vilnius University

[‡]denisa.sologon@liser.lu, Luxembourg Institute of Socio-Economic Research (LISER)

[§]cathal.odonoghue@nuigalway.ie, The National University of Ireland, Galway

[¶]linas.tarasonis@evaf.vu.lt, Vilnius University and Bank of Lithuania

Contents

1 Introduction

The 2011 World Economic Forum has identified income inequality as one of the "two most serious challenges the world is facing today". Income inequality has increased since the 1980s in most advanced economies as well as emerging markets including post-Soviet countries. According to Eurostat, the Gini index of household equivalized disposable income for Lithuania was 37.6% in 2017, the second highest across the whole of the the European Union (EU) in 2017. Understanding what drives changes in income distributions over time is a central issue in economic research and policy analysis.

We study changes in the income distribution in Lithuania between 2007 and 2015. The case study of Lithuania is particularly interesting, given country's experienced transition from planed to market economy and its ongoing convergence to the EU and large fluctuations in disposable income. During this period, Lithuania faced the global financial crisis, over which household disposable income fluctuated severely and a series of tax and benefit reforms were implemented. Net emigration accelerated the ageing problem, putting additional strain on the income distribution. To effectively combat these income inequality fluctuations, one must first understand to what extent each of these factors are contributing to income inequality and see whether the tax and benefit system in place is able to tackle such challenges.

Various strategies have been used in the literature to analyse changes in the distribution of income over time. Traditional approaches compute one particular inequality summary index in two different moments in time and then use decomposition methods to break down the observed changes into the contribution of a number of components (these are typically based on the methodologies proposed in Reynolds and Smolensky 1977, Shorrocks 1980, Shorrocks 1982, Shorrocks 1984 and Lerman and Yitzhaki 1985). A second strand of literature focuses on modelling the market income distribution using parametric and semi-parametric econometric techniques and building counterfactual scenarios that allow for an assessment of the contributions of various factors to the overall evolution of the distribution (see for example Juhn et al. 1993, DiNardo et al. 1996 and Bourgignon et al. 2008). Finally, there is a sizeable literature that departs from an observed fixed market income distribution and focuses on assessing the role of the tax-benefit system in determining changes in the disposable income distribution, through the use of tax-benefit microsimulation models (see for example Bourgignon and Spadaro 2006 and Bargain 2014). These strategies, while interesting and useful in their own right, are limited in their scope, as they refer to the analysis of either a summary measure or only one part of the income distribution, be it market incomes or taxes and benefits.

We build on the approach developed in Sologon et al. (2018), adapting it to study changes in income distributions over time for one single country instead of differences in income distributions across countries in one given moment.[1] The method integrates both a micro-econometric and microsimulation approaches, combining a flexible parametric modelling of the distribution of household market income with the EUROMOD model to simulate the value of taxes and benefits. It generates a multitude of counterfactual income distributions, obtained by "swapping" the characteristics of the country in two different moments in time along four main dimensions: (i) labour market structure; (ii) returns; (iii) demographic composition; and (iv) tax-benefit system. The comparison of these counterfactual distributions then allows to quantify the contribution of each dimension to the changes in the income distribution (and functionals) observed between any two moments in time. By applying this approach we provide a more detailed decomposition than existing studies that seek to unpack the drivers of inequality changes. Most existing studies on the topic follow the approached proposed by Bargain and Callan (2010) and Bargain (2012), which uses two "swaps": gross income and tax-benefit. We engage in a higher level of disaggregation by breaking up market income into the institutional structures in terms of employment rates, the number of people with income sources, the distribution of the market, the distribution of the returns and the demographics. 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 remainder of the paper is organised as follows: Section 2 introduces the methodology used to model the household disposable income distribution and decompose changes in this distribution over time; Section 3 presents the context and results of the application done to Lithuania between 2007 and 2015; Section 4 concludes and discusses some policy implications.

2 A method of household disposable income distribution

In this section we present the method used for modelling the household disposable income distribution and an analysis of the anatomy and determinants of changes in this distribution (or functionals such as inequality measures) over time. Our method builds on the approach developed in Sologon et al.

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

(2018), adapting it to study changes in income distributions over time for one country instead of differences in income distributions across countries (see also Sologon et al. (2019) for an adaptation for Portugal).

We start by modelling three sources of market income (labour, capital and other), estimating separately the probabillity of receipt and the level. Having obtained estimates for market incomes we then feed these values into EUROMOD to estimate the value of different types of taxes and benefits. Adding benefits to market income and subtracting taxes we obtain an estimated value of disposable income for each household from which we can construct the full household disposable income distribution and compute any functional of interest. Changes in this distribution (or functionals) are then decomposed into the contributions of four main factors: (i) labour market structure; (ii) returns; (iii) demographic composition; and (iv) tax-benefit system. These contributions are estimated by simulating and comparing a sequence of counterfactual distributions obtained by swapping the characteristics of the economy in each year along the four main dimensions considered.

Below we describe these steps in detail.

2.1 Household disposable income components

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
- public benefits, y_h^B ,
- household direct taxes, y_h^T

Summing, we define household disposable income as:

$$y_h = \underbrace{y_h^L + y_h^K + y_h^O}_{Market} + \underbrace{y_h^B - y_h^T}_{Non-market} \tag{1}$$

Most of these five components are themselves aggregates of smaller components of income (notably contributions of individuals to overall household income), which we model separately in order to have a representation that is defined at a fine level of disaggregation. We provide the main aspects of this disaggregation below, leaving the details for the next two sections.

Market incomes

For each component of market income, income is estimated at the individual level, and then for each household the incomes of all individual members are added to obtain the household's income. Each component is disaggregated into two sources: *Labour income* into employment (emp) and self-employment (semp) income*; Capital income* into investment (inv) and property (prop) income; and *Other non-benefit pre-tax income* into private pensions (pripen) and a catch-all concept that aggregates all other non-benefit individual incomes (mainly private transfers such as alimonies) (other). For each income source, we first estimate a binary participation indicator equal to one if the individual receives that type of income and zero otherwise and then, for the individuals receiving it, we estimate the level. For labour income, we first estimate a binary indicator equal to one if the individual is working and zero otherwise and then, for those individuals working, we assign the estimated income from employment and self-employment. We then have:

$$y_h^L = \sum_{i=1}^{n_h} I_{hi}^{lab} \left(I_{hi}^{emp} y_{hi}^{emp} + I_{hi}^{semp} y_{hi}^{semp}\right) \tag{2}$$

$$y_h^K = \sum_{i=1}^{n_h} \left(I_{hi}^{inv} y_{hi}^{inv} + I_{hi}^{prop} y_{hi}^{prop}\right) \tag{3}$$

$$y_h^O = \sum_{i=1}^{n_h} \left(I_{hi}^{pripen} y_{hi}^{pripen} + I_{hi}^{other} y_{hi}^{other}\right) \tag{4}$$

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 ∈{emp, semp, inv, prop, pripen, 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 amount of income received from that source by individual hi.

Non-market incomes

Non-market incomes are obtained by feeding EUROMOD with the individual-level estimates of market incomes, together with the corresponding socio-demographic characteristics*. Benefits* are composed of a range of individual-level replacement incomes (including retirement and survivor pensions and disability, sickness and unemployment benefits), household-level means-tested social assistance (including housing support) and universal (non means-tested) transfers (including child support). For simplicity we will refer to three broad household-level aggregates: public pensions (pens), means-tested benefits (mtb) and non-means-tested benefits (nmtb):

$$y_h^B = y_h^{pens} + y_h^{mtb} + y_h^{nmtb} \tag{5}$$

Direct taxes are given by the sum of income taxes (tax) paid at the household level and social security contributions (ssc) paid at the individual level. They are determined by the tax schedule in place as a function of the vector of gross incomes (i.e. market income plus benefits) and household characteristics and composition:

$$y_h^T = y_h^{tax} + \sum_{i=1}^{n_h} y_{hi}^{ssc} \tag{6}$$

2.2 Parametric modelling of market incomes

The modelling of market income builds heavily on the method developed in Bourgignon et al. (2008), relying on hierarchically structured, parametric, multiple equation specifications for each detailed source of market income. Each income source is represented as a combination of three types of elements: (i) a set of basic household and individual observable characteristics; (ii) a vector of parameters describing how the receipt and level of income vary with the observable characteristics; and (iii) a vector of household-specific residuals, linking the predictions from the model to the observed values of income. We give special attention to the modelling of labour income, in order to be able to capture in detail the rich contribution provided by different aspects of the labour market structure and wage compensation to the evolution of market income.

An important note should be made at this point. In our setting, model parameters are not meant to capture causal relationships between the various endogenous and exogenous variables considered. The parametric relationships are reduced-form projections that aim to describe the empirical associations between basic conditioning variables and various components of income. The objective is to use these estimated projections to pinpoint the key sources of changes in income distributions over time, whether they mainly arise from differences in population characteristics, from how these characteristics determine income differences or from residual heterogeneity.

Participation indicators

The modelling of each market income source starts by the estimation of a participation indicator for each individual, capturing whether the individual is receiving income from that particular source. In addition, in the case of labour income, a labour market participation indicator is considered, to capture whether the individual is working and therefore receiving income from one or both sources of labour income. All these indicators are modelled in an analogous way, using a binary logistic regression model, and therefore we start by giving a general description of this procedure.

Consider any income source s for which we wish to estimate a participation indicator for individual hi, I^s_{hi}, equal to one if the individual receives any amount of income from that source and zero otherwise. It is assumed that the outcome of this binary variable depends on the value of a continuous latent variable, I^{s*}_{hi}, being equal to one when this value is positive and zero otherwise, i.e. $I^s_{hi} = 1$ if $I^{s*}_{hi} > 0$ and $I^s_{hi} = 0$ otherwise. The latent variable in turn depends linearly on a set of observable characteristics, x_{hi}, and on an error term, ϵ^s_{hi}, such that $I^{s*}_{hi} = x_{hi}\gamma^s + \epsilon^s_{hi}$. We then have $I^s_{hi} = \mathbf{1}\left[\epsilon^s_{hi} > -x_{hi}\gamma^s\right]$, where $\mathbf{1}\left[cond\right]$ is equal to one if *cond* is true and zero otherwise. The logistic regression model assumes that ϵ^s_{hi} has a logistic distribution such that the (conditional) probability of earning any income from source s is given by:

$$\Pr(I^s_{hi} = 1|x_{hi}) = \Pr(x_{hi}\gamma^s + \epsilon^s_{hi} > 0) = \Pr(-\epsilon_{hi} < x_{hi}\gamma^s) = \frac{\exp(x_{hi}\gamma^s)}{1 + \exp(x_{hi}\gamma^s)} \tag{7}$$

The characteristics included in x_{hi} are: age (and age squared); academic achievement (whether holds a university degree); marital status; number of own children in the household (separating children under 4, children between 4 and 11 and children between 12 and 15); and citizenship. Separate sets of parameters are allowed for men, single women, and women in couple.

This way, we obtain estimates for: I^{lab}_{hi}, I^{emp}_{hi} and I^{semp}_{hi}appearing in equation (2); I^{inv}_{hi} and I^{prop}_{hi}appearing in equation (3); and I^{pripen}_{hi} and I^{other}_{hi}appearing in equation (4).

Levels of labour income sources

We start by modelling the earnings of self-employed workers, the variable y^{semp}_{hi} appearing in Equation (2), according to the following log-linear regression model:

$$y^{semp}_{hi} = \exp(x_{hi}\beta^{semp} + \upsilon^{semp}_{hi}) \tag{8}$$

where υ^{semp}_{hi} is a zero-mean residual with homoscedastic variance $\sigma^{2,semp}$.[2]

We then model the earnings of salaried employed workers, the variable y^{emp}_{hi} appearing in Equation (2). This is given by the multiplication of the individual's hourly wage, w_{hi}, by the individual's total number of hours worked, s_{hi}:

[2]The variances of residuals are part of the parameter vector. They are set to unity in logistic regression models.

$$y_{hi}^{emp} = w_{hi} s_{hi} \tag{9}$$

We model hours worked using a basic linear regression model:

$$s_{hi} = x_{hi} \gamma^{hrs} + \epsilon_{hi}^{hrs} \tag{10}$$

where ϵ_{hi}^{hrs} is a zero-mean residual with homoscedastic variance $\sigma^{2,hrs}$.

The final step in the modelling of labour incomes is a specification for wages. Given the central importance of wages in the distribution of household income, we adopt a specification that connects individual characteristics to the whole conditional wage distribution and not only to the conditional mean as in the regressions used for other sources of income. To do so, we assume that wages follow a Singh-Maddala distribution, F_X, represented by:

$$F_{X=z}(w) = \mathrm{SM}(w; a(z), b(z), q(\mathrm{z})) = 1 - \left[1 + \left(\frac{w}{b(z)}\right)^{a(z)}\right]^{-q(z)} \tag{11}$$

where the X indicates that the distribution is conditional on a vector of characteristics z. The Singh-Maddala distribution is a flexible unimodal three-parameter distribution that has been shown to provide good fit to wage distributions (Van Kerm et al. 2016). The parameter $q(z)$ 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. Each of these parameters is allowed to vary log-linearly with individual characteristics $\theta(z) = \exp(z\beta^{\theta,emp})$, as in Biewen and Jenkins (2005) or Van Kerm (2013). Individual wage is then given by:

$$w_{hi} = F_{X=z}^{-1}(\upsilon_{hi}^{emp}) = b(z)[(1 - \upsilon_{hi}^{emp})^{-\frac{1}{q(z)}} - 1]^{\frac{1}{a(z)}} \tag{12}$$

where υ_{hi}^{emp} is a random term uniformly distributed. The model is estimated for men and women separately. For women, we estimate a participation-corrected model as in Van Kerm (2013).

Besides the previously introduced conditioning variables, x_{hi}, z contains three additional variables that are particularly relevant to the modelling of wages: occupation, occ_{hi}, industry, ind_{hi}, and sector, pub_{hi}, of main job. We model these three variables only for people in salaried employment (i.e.with $I_{hi}^{emp} = 1$), who are the ones receiving a wage.

Occupation is a categorical variable with 8 categories, based on the ISCO-08 classification. The categories considered are: managers; professionals; technicians and associate professionals;

clerical support workers; services and sales workers; craft and related trades workers; plant and machine operators and assemblers; and unskilled [3]. We model it using a multinomial logistic regression model. A latent variable $I_{hi}^{k,occ*} = x_{hi}\delta^{k,occ} + \epsilon_{hi}^{k,occ}$ is associated to each of the $k \in m^{occ}$ alternative occupations with $\epsilon_{hi}^{k,occ}$ following an extreme value distribution. The observed occupation for individual hi, say j ($I_{hi}^{j,occ} = 1$ and $I_{hi}^{k,occ} = 0$ for $k \neq j$), is such that $I_{hi}^{j,occ*} > I_{hi}^{k,occ*}$. Under an extreme value distribution for the residuals, the probability of being in occupation j is given by:

$$\Pr(I_{hi}^{j,occ} = 1|x_{hi}) = \frac{\exp(x_{hi}\delta^{j,occ})}{\sum_{k=1}^{m^{occ}} \exp(x_{hi}\delta^{k,occ})} \tag{13}$$

with the parameter vector for the first alternative normalized to $\delta^{1,occ} = 0$. When only two choices are available, this is equivalent to the binary logistic model.

Industry of employment can be primary, secondary, or tertiary and is modelled similarly to occupation using a multinomial logistic model, with $m^{ind} = 3$. Sector of employment is either public or private (public sector includes public administration jobs but also army, health and education) and it is modelled using a binary logistic regression model, like the one described for the participation indicators, with one corresponding to being in the public sector. We add occupation as a conditioning variable in the models for industry and sector of employment, which are thus determined by (x_{hi}, occ_{hi}). Parameters are estimated using maximum likelihood.

Levels of other market income sources

We adopt a much simpler parametrisation for the levels of all other sources of market incomes, a log-linear regression model, similarly to the modelling of self-employment income. For each source p with $p \in${inv, prop, pripen, other}, we have:

$$y_{hi}^{p} = \exp(x_{hi}\beta^{p} + \upsilon_{hi}^{p}) \tag{14}$$

where υ_{hi}^{p} is a zero-mean residual with homoscedastic variance $\sigma^{2,p}$.

This way, we obtain estimates for: y_{hi}^{inv} and y_{hi}^{prop}appearing in equation (3); and y_{hi}^{pripen} and y_{hi}^{other}appearing in equation (4).

[3]The original ISCO classification considers one more category, "skilled agricultural, forestry and fishery workers", which we merge into the "managers" category. For more information on the ISCO-08 classification see http://www.ilo.org/public/english/bureau/stat/isco/isco08/index.htm.

2.3 Simulation of benefits, taxes and social security contributions

The final two components of household disposable income are benefits (or public transfers) received, y_h^B, and direct taxes paid, y_h^T. We derive the bulk of these components using EUROMOD, a pan-European tax-benefit static microsimulation engine (see Sutherland and Figari (2013) for a presentation of the model). This large-scale income calculator incorporates the tax-benefit schemes of EU member countries and uses harmonised input datasets, allowing for the estimation of benefit and tax (both direct taxes and social insurance contributions) entitlements as a function of pre-tax pre-benefit income sources, household characteristics and other variables that may influence the benefit eligibility and tax liabilities according to the rules in place (see Figari et al. (2015) for a discussion of the modelling of taxes and benefits using microsimulation models). It also makes it possible to implement 'policy swaps' in which particular tax or benefit policies from one reference country or year are applied to other countries or years (see for e.g. Levy et al. (2007), Bargain and Callan (2010) and Bargain (2012)).

EUROMOD simulates a wide range of benefits including family benefits, housing benefits, social assistance, and other income-related benefits. Not all benefits are however, evaluated by EUROMOD. Two main sources of benefits are not simulated (or are only partially simulated): contributory benefits and retirement and disability pensions, which generally depend on past employment histories or other information (e.g., about the severity of a disability) that is usually not observed in the household survey data that inputs the tax-benefit simulator. For these components of y_h^B, the benefits measured at the individual level are modelled like non-labour incomes (with a logistic regression model for receipt and a log-linear regression model for the amount received), while benefits measured at the household level are modelled similarly except that only one household level equation is specified for each model and the exogenous characteristics x_h are composed of household-level demographic composition and of the individual characteristics of the 'household head' (where household head is defined as the individual with the highest individual income or the eldest in the case of equal income). We rely entirely on EUROMOD for the computation of direct taxes, which include income and property taxes and social security contributions.

A few additional variables that are not a part of household income can influence the amount of taxes and benefits, such as mortgages, rents paid, and contributions for private pensions. These variables are also modelled, using the same strategy as for non-labour incomes and benefits not evaluated by EUROMOD. The estimates obtained do not determine household income directly, but are fed into the tax-benefit microsimulation engine to calculate taxes and benefits of household h.

2.4 Counterfactual distributions and decomposition of changes over time

We describe next the method used to generate counterfactual distributions and perform the decomposition of changes in the income distribution between any two years.

The Income Generation Process

We start by introducing a generic representation of the household income generation process (IGP):

$$Y = m(X, \Upsilon) \tag{15}$$

where Y is household disposable income, X is a vector of exogenous characteristics and Υ is a vector of unobserved heterogeneity (residual) terms (see Matzkin (2003) and Rothe (2010)). The function m describes jointly the relationship between X and Y and the heterogeneity in Y that is not 'explained' by X. The derivative of m with respect to its first argument reflects variations in Y across households that can be attributed to differences in observable characteristics while the derivative of m with respect to its second argument reflects variations in Y across households with identical observable characteristics.

The parametric functional forms adopted for the different income components imply a particular parametric shape for m, such that:

$$Y = m^{\xi}(X, \Upsilon; \xi) \tag{16}$$

where m^{ξ} represents the specific parametric structure adopted for the income generation model and ξ is the vector of parameter values. Equation (16) has no 'structural' interpretation but it should be viewed as a set of reduced form equations linking household characteristics and income (a relationship that may arise from an unknown, broader structural model) through earnings functions, equations for employment and occupational and industrial structure, equations for non-labour income and replacement incomes and through tax-benefit rules.

We are interested in studying the distribution F of the random variable Y as well as any functional of interest $\theta(F)$ (such as inequality indices). In particular, we want to examine why F or $\theta(F)$ may differ between two periods. This will depend on the (joint) distribution of X and Υ in the population through m^{ξ} and ξ. Therefore, differences in F and $\theta(F)$ over time will be a result of differences in the distributions of observable characteristics and unobservable residual heterogeneity and differences in the model's parametric structure and parameter values. For tractability reasons, we assume that all years can be represented by a common parametric model of the form m^{ξ} but that years differ in the values taken by the parameters ξ.

In order to quantify the relative contributions of these factors to changes in F and $\theta(F)$, we define a number of 'transformations' that, when applied to the model, allow us to build counterfactual distributions which, when compared to the baseline distribution, capture how sensitive F and $\theta(F)$ are to specific dimensions of the model. The transformations are then calibrated to reflect actual differences between periods in the factors concerned, leading to a decomposition of over time differences into specific factors of interest. We describe the several steps of this process below.

Four transformations of the Income Generation Process

We focus on four 'transformations' of the IGP that allow us to capture the relative contributions of four main factors (or subsets thereof): (i) a *labour market structure* transformation; (ii) a *returns* transformation; (iii) a *demographic composition* transformation; and (iv) a *tax-benefit system* transformation. These transformations follow naturally from the characteristics of our model, but it should be noted that they are specific choices among other possibilities that could be explored.

The *labour market structure* transformation consists of changing the values of parameters that define crucial aspects of the labour market structure such as employment probabilities and occupational, industrial and sectoral structures. This involves modifying certain elements of the parameter vector ξ, including the ones characterising employment probabilities (γ^{lab}, γ^{emp}), hours worked (γ^{hrs}) and occupational, industrial and sectoral structures ($\delta^{j,occ}$, $\delta^{f,ind}$, δ^{pub}). This produces an alternative parameter vector, $\tilde{l}(\xi)$, based on which we obtain new outcomes for income, Y^l (which leads to a new counterfactual distribution of income, F^l):

$$Y^l = m^{\xi}(X, \Upsilon; \tilde{l}(\xi)) \tag{17}$$

The *returns* transformation again acts through the parameter vector ξ. Specifcally, it involves changing the parameters of the equations characterising the levels of labour earnings (($\beta^{semp}, \sigma^{semp}$), ($\beta^{a,emp}, \beta^{b,emp}, \beta^{q,emp}$)) and of all other pre-tax incomes (($\beta^{inv}, \sigma^{inv}$), ($\beta^{prop}, \sigma^{prop}$), ($\beta^{pripen}, \sigma^{pripen}$), ($\beta^{other}, \sigma^{other}$)). This produces an alternative parameter vector, $\tilde{r}(\xi)$, based on which we obtain new outcomes for income, Y^r (which leads to a new counterfactual distribution of income, F^r):

$$Y^r = m^{\xi}(X, \Upsilon; \tilde{r}(\xi)) \tag{18}$$

This transformation is analogous, albeit in a multiple equations setup, to the manipulation of the vector of coefficients in Mincerian earnings regressions in order to capture 'price' effects (as distinct from 'composition' effects) in traditional Oaxaca-Blinder decomposition exercises. It closely resembles the decomposition of Juhn et al. (1993) in the way residual variances are accounted for.

The *demographic composition* transformation consists of changing the values of variables that reflect basic socio-demographic characteristics of the population, such as: age, gender, marital status, education level and number of children. This involves a modification of the distribution of the random variables in X, and obtaining new outcomes for income, Y^d, based on this alternative distribution of X, $\tilde{X}(X)$ (which leads to a new counterfactual distribution of income, F^d) :

$$Y^d = m^{\xi}(\tilde{X}(X), \Upsilon; \xi) \tag{19}$$

The fourth and last transformation that we apply is the *tax-benefit system* transformation. This works as a particular transformation of the parameter vector ξ which modifies (i) the regression parameters determining the level and eligibility of benefits that are not (or only partially) simulated by EUROMOD and (ii) the rules and parameters of the tax-benefit calculator for tax liabilities and those benefits that are determined directly by EUROMOD. This produces an alternative parameter vector, $\tilde{tb}(\xi)$, based on which we obtain new outcomes for income, Y^{tb} (which leads to a new counterfactual distribution of income, F^{tb}):

$$Y^{tb} = m^{\xi}(X, \Upsilon; \tilde{tb}(\xi)) \tag{20}$$

For each of the four transformations, we can compute the impact on any distribution functional of interest, θ. 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:

$$\Delta_{\theta}^{k}(F) = \theta(F^k) - \theta(F) \tag{21}$$

Swapping of components of the Income Generation Process between years

The transformations just described can then be used to create counterfactual distributions that allow to answer the question: *'What would the income distribution of year t be if its IGP was the one of year s along one or more of the dimensions considere*d?*'*. This is done by estimating the IGP for each year separately and calibrating transformations so as to replace components of the IGP of year t with components of the IGP of year s. We call this process a 'swapping' of components between the IGP of years t and s. Once again this procedure is analogous to standard Oaxaca-Blinder decompositions (swapping regression coefficients across earnings equations for alternative groups) but implemented in a multiple equations model and in an over time instead of cross country framework.

For the *labour market structure* transformation applied to period t, the parameter vector is transformed such that we obtain a new vector composed by: the subset of parameters from period t that are not affected by the transformation ξ^t; and the subset of parameters that are 'imported' from period s through the swapping procedure, ξ^s. The transformed IGP for year t thereby corresponds to a simulated distribution for year t as if it had the labour market structure of year s and all other components of the model unchanged.

The *returns* transformation involves a similar procedure, i.e. an 'importation' or swaping of the returns-related year s parameters onto year t's IGP,. Unlike the labour market structure transformation, however, the returns transformation also involves swapping variance terms, σ^2. This is achieved as in Juhn et al. (1993), by rescaling the residuals of period t by the ratio $\frac{\sigma_{\frac{K}{s}}}{\sigma_t^K}$ for each of the five income components K that are affected by the transformation. This procedure scales the distribution of residuals heterogeneity terms, but preserves the rank correlation of the residuals across the different equations of the IGP.

The *demographic* transformation involves modifying the distribution of population characteristics, X, of year t in such a way that it has the distribution of year s. The distribution of X is modified but the *conditional* distribution of the residuals,Υ given X, must not be affected and remain as it is in t. As shown in DiNardo et al. (1996) and Barsky et al. (2002), this can be achieved semi-pametrically by reweighting population t households by a factor given by:

$$\omega(X) = \frac{\Pr(X|s)}{\Pr(X|t)} = \frac{\Pr(s|X)}{\Pr(t|X)} \frac{\Pr(t)}{\Pr(s)} \tag{22}$$

The probabilities in equation (22) can be estimated by standard techniques for binary responses (see for e.g. Biewen and Juhasz (2012) for a recent application of this approach).

Finally, the calibration of the *tax-benefit system* transformation combines both 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. Such swapping of tax-benefit policy rules and parameters has already been done in other studies for the analysis of trends in income distributions (see Bargain and Callan (2010), Bargain (2012), Herault and Azpitarte (2016), Paulus and Tasseva (2017)) and cross-country differences (see Dardanoni and Lambert (2002) and Levy et al. (2007)).

Decomposition of changes in the income distribution over time

Next, we proceed to the decomposition of the observed differences between the income distributions and corresponding functionals in years t and s into the contributions of a set of determinants. Suppose for example that we compute a certain functional $\theta(F)$ for each of the two years, obtaining $\theta(F^t)$ and $\theta(F^s)$. A decomposition procedure aims at (additively) decomposing the total observed difference, $\theta(F^s) - \theta(F^t)$, into the contributions of each of the individual determinants k of a set K:

$$\Delta_\theta(F^t, F^s) = \theta(F^s) - \theta(F^t) = \sum_{k=1}^{K} \Delta_\theta^k(F^t, F^s) \tag{23}$$

A common way to build such a decomposition is by applying each determinant sequentially, one after the other, from the original distribution, F^t, to the target distribution, F^s, and taking the difference between two consecutive steps of the sequence.

The drawback of such a sequential decomposition is its path-dependence, i.e. the dependence of the estimated contribution of each factor on the precise sequence of transformations chosen[4]. To tackle this issue, we follow Biewen and Juhasz (2012) and Biewen (2014) and examine 'direct effects', which assess the impact of each determinant from the same initial benchmark distribution and therefore avoid composing transformations. Each direct effect is defined as:

$$D_\theta^k(F^t, F^s) = \theta(F_t^k) - \theta(F^t) \tag{24}$$

where F_t^k is the counterfactual distribution obtained by applying one single particular transformation k to the initial distribution F^t.

As Biewen and Juhasz (2012) argue, comparing 'direct effects' is a natural way to assess the effects of alternative transformations. However, the sum of all direct effects and residuals (unexplained factors) does not add up to the overall difference between income distributions. The discrepancy captures the interactions between the different transformations. In the context of our

[4]Some authors have proposed to calculate the contribution of each factor in all possible sequence of introduction of factors and average across sequences (Devicienti (2010), Chantreuil and Trannoy (2013), Shorrocks (2013)). This approach can however be computationally prohibitive for complex models and does not necessarily improve the economic interpretation of the estimated components.

model, we then have six components: the four direct effects of each transformation, the residuals, and an interaction term:

$$D^l_\theta(F^t, F^s) = \theta(F^l_t) - \theta(F^t) \tag{25}$$

$$D^r_\theta(F^t, F^s) = \theta(F^r_t) - \theta(F^t) \tag{26}$$

$$D^d_\theta(F^t, F^s) = \theta(F^d_t) - \theta(F^t) \tag{27}$$

$$D^{tb}_\theta(F^t, F^s) = \theta(F^{tb}_t) - \theta(F^t) \tag{28}$$

$$\Delta^\Upsilon_\theta(F^t, F^s) = \theta(F^{d,l,r,tb}_s) - \theta(F^t) \tag{29}$$

$$I_\theta(F^t, F^s) = \Big(\theta(F^s) - \theta(F^t)\Big) - \left(\sum_{k\in\{l,r,d,tb,\Upsilon\}} D^k_\theta(F^t, F^s)\right). \tag{30}$$

The term $D^\Upsilon_\theta(F^t, F^s)$ captures the contribution of differences in the distribution of scaled residual or unobserved heterogenity terms Υ.[5] $I_\theta(F^t, F^s)$ is an interaction term equal to the total difference in θ and the sum of direct effects, accounting for all two-way and three-way interactions between the four components in the model (Biewen (2014)).

The total observed difference is decomposed as:

$$\Delta_\theta(F^t, F^s) = D^l_\theta(F^t, F^s) + D^r_\theta(F^t, F^s) + D^d_\theta(F^t, F^s) + D^{tb}_\theta(F^t, F^s) + \Delta^\Upsilon_\theta(F^t, F^s) + I_\theta(F^t, F^s) \tag{31}$$

3 An application to Lithuania between 2007 and 2015

We now apply the method presented in the previous section to study the changes in the income distribution in Lithuania between 2007 and 2015.

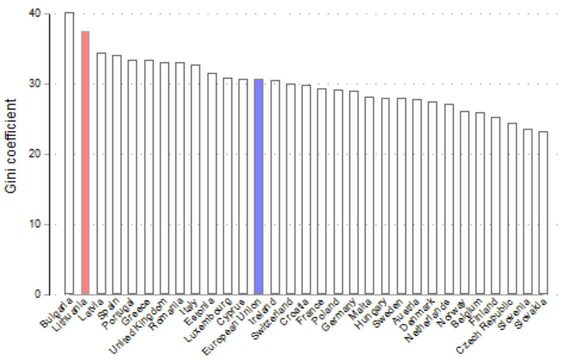
3.1 Evolution of income inequality in Lithuania

Changes in the income distribution in Lithuania is an interesting case study since Lithuania displayed one of the highest levels of income inequality across the European Union (EU) in 2017. According to the European Union Statistics on Income and Living Conditions (EU-SILC), the most reliable data on income inequality currently available to Eurostat, the Gini index of household equivalized disposable income was 37.6% in Lithuania in 2017 (see Figure 1). This made Lithuania the second most unequal country in the EU ranking 7 gini points higher than the EU average and

[5] $D^\Upsilon_\theta(F^t, F^s)$ is obtained by transplanting residuals across periods. This is achieved, in reverse, by starting from time s and jointly applying all four transformations calibrated to period t parameters. The difference between this construct and time t's orginal distribution reflects the 'direct' effect of transplanting residuals from t to s.

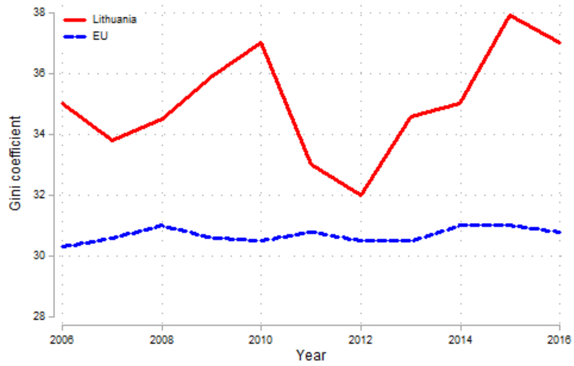
14.4 gini points higher than Slovakia, a country with the most equal income distribution in the European Union and another former member of the Soviet Union.

Figure 1: Gini coefficient, European Union, 2017



Source: Eurostat data.

Figure 2: Gini coefficient, Lithuania, 2006-2016



Source: Eurostat data.

Income inequality in Lithuania has been on the rise over the past two decades. Figure 2 portrays the dynamics of the Gini coefficient for Lithuania and the European Union as a reference from 2007 to 2015. Income inequality in Lithuania has consistently exceeded income inequality

in the European Union. In what follows we discuss potential drivers of changes in the Lithuanian income distribution: business cycle, demographics, structural changes in the labor market and changes in the tax and benefit systems.

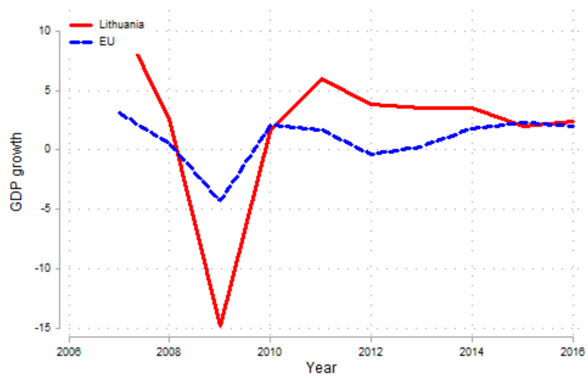
Business cycle

Looking at Figure 2, the Gini coefficient in Lithuania appears to be strongly procyclical, much more so than the average in European Union, which looks very stable of the period under discussion. The Gini in Lithuania grew somewhat during 2006-2009, it peaked at 37% in 2010 and then fell to a low of 32% in 2012, before starting to rise again reaching an all time high of 38% 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 outbreak of the 2007-2008 crisis in the US first hit Lithuania particularly hard. Lithuania, along with the other three countries in the Baltics, experienced one of the highest levels of GDP contraction globally. This 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, accelerating wage growth—that exceeded productivity growth.

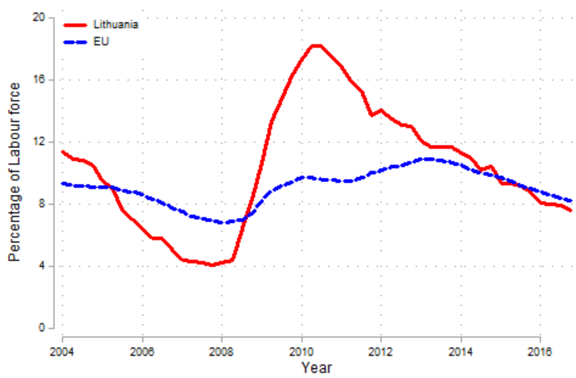
The domestic bubbles burst in early 2008, when the credit supply decelerated and banks started tightening credit conditions. The downturn was further exacerbated by negative developments in the external economic environment after the Lehman Brothers' bankruptcy. Figure 3 portrays GDP growth of the Lithuanian economy versus the average in the European Union. During the peak of the crisis in 2009, the Lithuanian economy has contracted by almost 15%. It has bounced back from the global financial crisis just as fasting recording growth rates significantly above the EU average in the early 2010s.

Figure 3: GDP growth



Source: Eurostat data.

Figure 4: Unemployment rate



Source: Eurostat data.

The crisis affected the labor market painfully. As can be seen in Figure 4, the unemployment rate rose steadily between 2008 and 2011, from 4% to almost 18%. For comparison, the fluctuations in the average unemployment rate in European Union was pronounced but significantly less so. Again, the labor market bounced back rather quickly during the expansion period: the unemployment rate fell below the EU average in 2015.

Demographics

The demographic situation of Lithuania has been affected by two important trends over this period: migration and ageing. The latter had a sizeable effect to the total size of the population, which has shrunk throughout this period. Specifically, the population of Lithuania decreased by 18% from 2004 to 2016 and most of it was due to the permanent annual 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, life expectancy has been on the rise like in most of the Europe. 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 has risen to 28 by 2016. This shift might have had important consequence for the income distribution, since a greater fraction of the population became dependent on pension income.

Structural changes in the labour market

The labour market has also experienced several important structural changes common to most developed countries. In particular, there has been a 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 number of reforms of the the tax and benefit system during this period. First, immediately following the crisis, 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 only provisional - between January 1st, 2010 and December 31st, 2011. The new law essentially 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 lowered. Maternity and paternity benefits were reduced during this period also. While most of these temporary provisions expired at the end of 2011, several, such as reduced state pensions for officers, soldiers or academic workers, persisted until the end of 2013.

There has been important changes in the 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. Since the 1st of January 2012 the state pension age gradually increases 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 has been reformed to allow for an opportunity to accumulate and invest a part of the funds in the private sector. Since 2004 every person insured for full pension insurance (basic and supplementary parts of pension) may voluntarily choose either to stay only in the public social insurance system or 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. This cumulative part of the pension adds as the supplementary part of the old-age pension. At the same time because of lower contributions to the state's social insurance fund, the old-age pension will be respectively lower.

In addition, there has 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 by 20% rate. Since the 1st of January 2014 this tax rate was lowered to 15%.

There were also changes in one the largest component of labour costs - social insurance contributions. Those are flat-rate without ceilings, but they differ for employees and self-employed. Employee's contribute 3% of gross wages and salaries as contributions to pension social insurance and, since 2009, an additional 6% to health social insurance. Employers, on the other hand, 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 to pension social insurance depending on their income. Since 2009 self-employed persons additionally contribute to sickness and maternity social insurance. Starting 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 shall focus on the period between 2007 and 2015, which was a very intense period for the Lithuanian economy, and we are going to divide this period into two sub-periods, 2007-2011 and 2011 - 2015. The first period is dominated by the effects of the 2007-2008 crisis, the later comprises the recovery period taken from 2011 onwards.

3.2 Data

Our database is taken from the European Union Statistics of Income and Living Conditions (EU-SILC) survey for Lithuania, for the years 2007, 2011 and 2015. The EU-SILC is a nationally representative household survey, which contains detailed information about income as well as about the socio-economic characteristics of households and their members. It is currently the key source of official statistics on income distribution for most European countries, including Lithuania.

Given that a central component of our model is the tax-benefit microsimulation engine EUR-

OMOD, we use the 'EUROMOD input data' versions of the EU-SILC datasets, which have been standardized for common definitions of income variables and household characteristics. The definition of disposable household income in EUROMOD includes the sum across all household members of market incomes and public pensions plus cash benefit minus taxes and social insurance contributions. Cash benefits, taxes and social insurance contributions are not reported by survey respondents but are calculated by EUROMOD. EUROMOD assumes away any tax evasion and assumes full take-up of benefits. All income measures are expressed in 'single adult equivalent' by dividing total household income by the square root of household size and attributing that value to each member of the household. Furthermore, all income measures are CPI adjusted. Sample sizes are around 10 thousand individuals, which is just under 5 thousand households in each year.

Table 1 shows a number of population socio-economic characteristics for each of the three years, based on the samples in our database. We can see that there were some significant changes over the whole period of 2007-2015. In terms of socio-demographic characteristics, there was a considerable increase in the share of people with tertiary education among the population aged 25-64, by more than 7 p.p., with the most of the increase concentrated in the period between 2007 and 2011. There was also ageing effect of the population, with the share of people aged 65+ increasing by more than 3 p.p. over the whole period. Two reasons are potentially behind this. First, life expectancy in Lithuania rose rapidly, meaning that people could live longer lives. Second, as discussed above this period was marked by high emigration that exceeded immigration flows. Since largely working age population emigrated in the hope of getting higher earnings abroad, the share of elderly in Lithuania increased. The changes in the labour market structure are more nuanced. There is a relatively small dip of in work population in 2011 as compared to 2007, despite Lithuania having experienced one of the highest output falls and unemployment increases in Europe. This can be partly explained by the fact that Lithuanian workers managed to earn at least some income during the year. Additionally, there was a long term trend in activity rates. This may explain why not only did the share of people earning income barely declined, but in 2015 exceeded the 2007 level. This is in line with official statistics, which showed that activity (participation) rates, especially of the elderly, has increased over the whole period. What the crisis seemed to have changed, however, is the share of employees and self-employed among those who were working. 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, with the main ones being the increase in the share of professionals and the decrease in the share of associate

Table 1: Population socio-economic characteristics (shares of total population)

	2007	2011	2015
Demographic			
Tertiary Education	0.287	0.332	0.358
People 16-65	0.684	0.670	0.665
People >65	0.148	0.173	0.179
Child 0-3	0.038	0.037	0.039
Child 4-11	0.080	0.073	0.081
Child 12-15	0.049	0.047	0.036
Married	0.578	0.530	0.469
Citizen	0.995	0.995	0.992
Male	0.444	0.450	0.451
Labour market			
In-work	0.615	0.606	0.653
Employee/Self-Employed	0.897	0.942	0.910
Occupation			
Managers	0.139	0.115	0.115
Professionals	0.168	0.233	0.229
Associate Prof.	0.104	0.084	0.071
Clerks	0.041	0.038	0.043
Service	0.118	0.125	0.122
Craft	0.204	0.193	0.189
Plant	0.112	0.103	0.103
Unskilled	0.113	0.110	0.129
Industry			
Agriculture	0.078	0.058	0.052
Industry	0.246	0.155	0.151
Services	0.676	0.788	0.797
Business certificate	0.262	0.191	0.215
Other market factors			
With capital income	0.085	0.075	0.164
With other income	0.018	0.023	0.020

Notes: The estimates are weighted. The shares for education refer to age-group 25-64; for married, sex to age >= 16; for in-work to ages 16 to 80; for employees, occupation, industry and sector to those in work aged [16, 80); for citizen to the entire sample; for buisness sertificates to self-employed. The shares for capital refer to age>= 16.

professionals. This may relate with an increase of people with tertiary education, who were able to meet higher working criteria. There was also a large shift towards the service sector at the expense of the agricultural and industry sectors. Finally, considering other market factors, there was a significant increase in the share of people with capital income, of 8 p.p..

3.3 Changes in the income distribution between 2007 and 2015

We start by characterizing the changes in the distribution of household disposable incomes in Lithuania between 2007 and 2015, considering both the period 2007-2015 as a whole and the sub-periods 2007-2011 and 2011-2015. The three distributions (2007, 2011 and 2015) were obtained using the methodology presented in Subsections 2.1 to 2.3.

Table 2 shows the mean and median monthly income and Gini index associated with each of these distributions. The mean and median income were almost identical in 2007 and 2011. This is not surprising, as we capture a segment of the upturn (2007-2008), the downturn (2008-2010) and the start of the recovery (2010-2011). However, we see that over this period, people's incomes barely changed. In contrast, the mean and median income rose to 618 and 508 in 2015, roughly a 40% increase since 2011. Interestingly, the Gini responded differently. It slightly fell from 2007 to 2011, but then increased by 2.9 Gini points in 2015 despite the rising incomes.

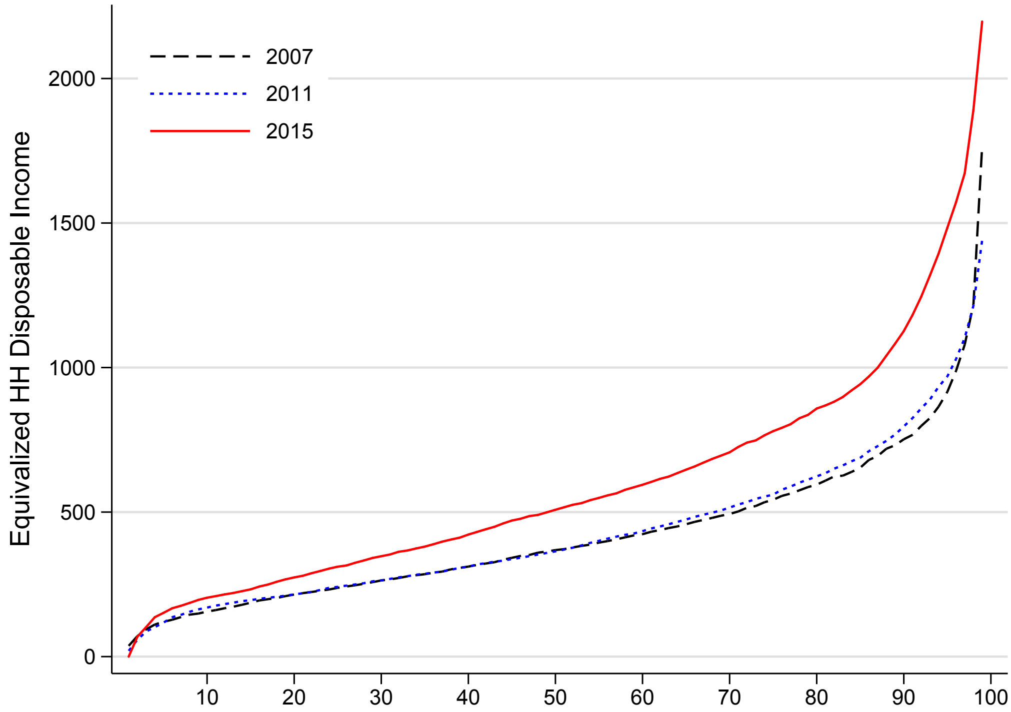
Table 2: Summary measures of equivalised household disposable income distribution (monthly, in euros)

	Mean	Median	Gini
2007	440	370	0.339
2011	441	364	0.331
2015	618	508	0.360

The rise of the Gini alongside rising mean and median incomes suggests that incomes rose very unevenly for the population, particularly from 2011 to 2015. For that, we need to consider the full distributions of incomes, which are given in Figure 5 in the form of Pen's parades. When comparing the distributions of 2007 and 2015, it can be seen that almost all quintiles experienced an income increase. Even the quantiles at the very bottom of income distribution, which are less visible due to the scale used, also exhibited an increase. It is also immediately clear that the greatest increase in income came since 2011, while the income distributions in 2007 and 2011 were very similar. What we also see is that the income of different quantiles increased by different absolute amounts - with those at the top gaining significantly more than those at the bottom. Nevertheless, one should consider the relative increase in income to relate back to the changes in the Gini index.

The relative increase in income is presented in Figure 6, which shows the pairwise differences between the three distributions shown in Figure 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. We can clearly see that the 2007-2015 period comprised two very distinct sub-periods in what 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

Figure 5: Distribution of equivalised household disposable income (Pen's parades)



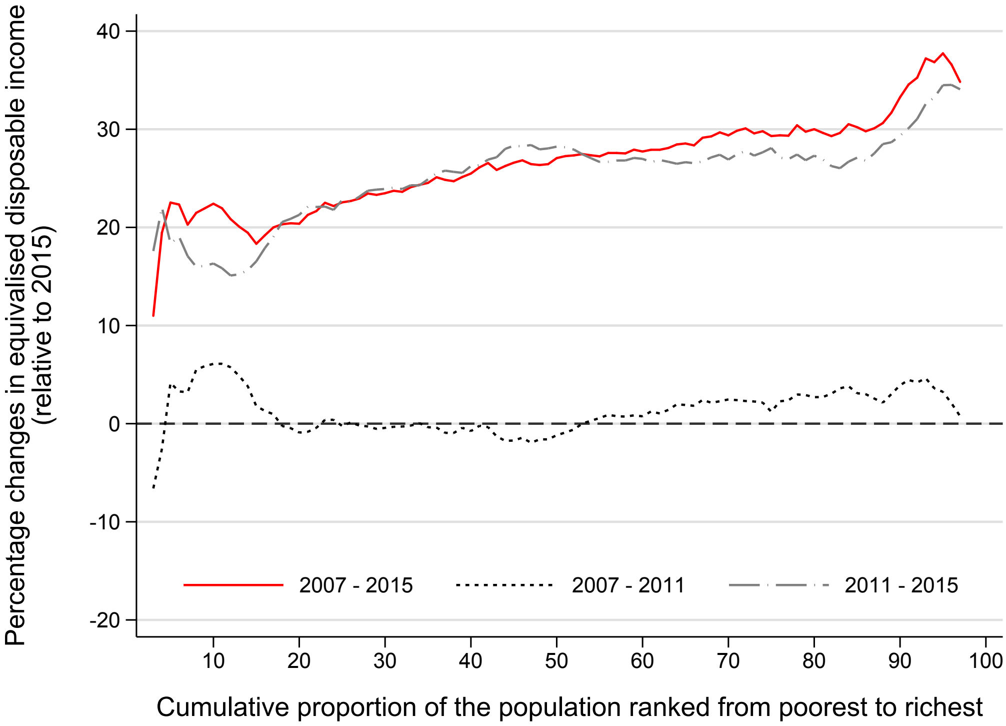
40-50% actually lost incomes. This contrasts with the 2011-2015 period, where income of the entire distribution rose. However, the rise in income in 2011-2015 period differs along the distribution: it rose by around 20% for the bottom 20% of the population and around 30% for the top the population and even more for the top 5%. Therefore, the economic upturn increased the inequality between the tails of the distribution.

Considering the evidence presented in Figures 5 and 6, together with the context described in Subsection 3.1, it seems reasonable to conclude that the measures that were implemented before the crisis (2007-2008) were able to help those at the bottom of distribution not only cushion the effects of the crisis, but also ensure greater income equality between households in 2011. However, when incomes were rising and the economy was growing, the lack of substantial changes to the measures and freezing of benefit payouts meant that these (or similar households) got left behind.

3.4 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 3 we provide some summary indicators of the effect of the system as a whole as well as the partial effects of taxes and transfers. Specifically, we present measures of: absolute

Figure 6: Changes in the distribution of equivalised household disposable income



redistribution, given by the Reynolds-Smolensky index; relative redistribution, which is equal to absolute redistribution as a share of the Gini of market income; size effect, measured by 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 progressivity/regressivity effect, measured by the Kakwani index[6].

[6]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.

Table 3: The redistributive effect of the tax and transfer system

	2007	2011	2015	2007-2011	2011-2015	2007-2015
Gini Gross Income	0.473	0.513	0.515	0.040	0.002	0.042
Gini Gross Income (incl. benefits)	0.369	0.364	0.391	-0.005	0.026	0.021
Averate transfer rate	0.186	0.252	0.223	0.066	-0.029	0.037
Benefit Regressivity (K)	0.768	0.845	0.801	0.078	-0.044	0.034
Benefit Redistribution (RS)	0.104	0.148	0.124	0.045	-0.024	0.021
Gini (gross + benefits - income taxes)	0.341	0.343	0.372	0.002	0.030	0.032
Averate tax rate	0.177	0.100	0.107	-0.077	0.007	-0.070
Tax Progressivity (K)	0.144	0.199	0.161	0.055	-0.038	0.017
Tax Redistribution (RS)	0.029	0.022	0.019	-0.007	-0.003	-0.010
Gini Disposable Income	0.339	0.331	0.360	-0.008	0.029	0.021
Net Redistributive Effect	0.134	0.182	0.155	0.048	-0.026	0.021

Notes: K = Kakwani; RS = Reynolds-Smolensky.

The analysis of these indicators suggests several findings. First, in terms of overall redistribution, we can see that the tax and transfer system as a whole was a crucial determinant of the level of disposable income inequality in Lithuania in the period under analysis. Indeed, in each of the three years considered, the Gini of disposable income is about 30% smaller than the Gini of market income. However, the system was not equally redistributive throughout the period of 2007-2015. The system became more redistributive in 2011 as compared to 2007, as seen by the net redistributive effect. Even though market income inequality rose in the period (the Gini of gross income rose by 4 p.p.), the system responded by higher redistribution and the resulting disposable income inequality was actually smaller than in 2007. The system, however, became less redistributive in 2015 as compared to 2011: disposable income inequality fell, even though market income inequality did not change in 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 reduced inequality by about a fifth. This compares to about a 6% reduction in inequality due to the tax system. Additionally, benefits became even more important in 2011: the average transfer rate and the benefit regressivity increased as compared to 2007 while average taxes fell. The increase in the importance of benefits in 2011 was partly undone by 2015, however. In 2015, average transfer rates and regressivity decreased.

3.5 Drivers of changes in the income distribution between 2007 and 2015

We now present the results from decomposing the changes presented in Subsection 3.3 into the contributions of the main factors considered in our model, as described in Subsection 2.4. We therefore move from a purely descriptive analysis of the evolution of the income distribution, focused on the 'what', to a thorough understanding of its driving forces, focused on the 'why'.

Decomposing changes in incomes

Figure 7 shows the contribution of each factor to the total changes in income distributions shown in Figure 6. Analogously to the results presented in Figure 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 6 is equal to the sum of the factor contributions given in Figure 7.

Over the 2007-2015 period, three main conclusions can be drawn regarding the changes in income distribution.

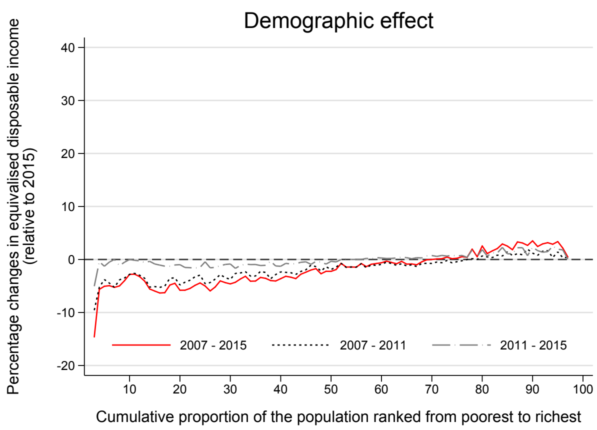
First, all four effects played a role in changing the level of household income. The effect of the tax and benefit system as well as the effect of price and returns overshadowed the two other effects (labour market effect and demographics) in terms of the incomes that people received. Because of the tax and benefit effect, disposable income in 2015 was about 35% higher than if the 2007 tax and benefit policy system would have been retained. Accordingly, because of the price and returns effect disposable income was 20% higher. This contrasts sharply with the up to 5% rise in income due to the labour market structure effect and a reduction in income due to the demographic effect.

Second, the tax and benefit effect as well as the price and returns effect affected different segments of the income distribution. The tax and benefit effect increased the income of the bottom deciles more while the price and returns effect affected much more the top of the income distribution. The timing of the effects differed too. The biggest gain for the bottom of the income distribution was due to the changes in measures over the 2007-2011 period. In fact, the effect of the system in 2011-2015 benefited the bottom 20% significantly less than the rest of the population. This contrasts sharply with the price and returns effect, where the majority of the effect was felt in 2011-2015. In essence, the measures adapted by the system could not deliver sufficient redistribution at a time when incomes were rising rapidly.

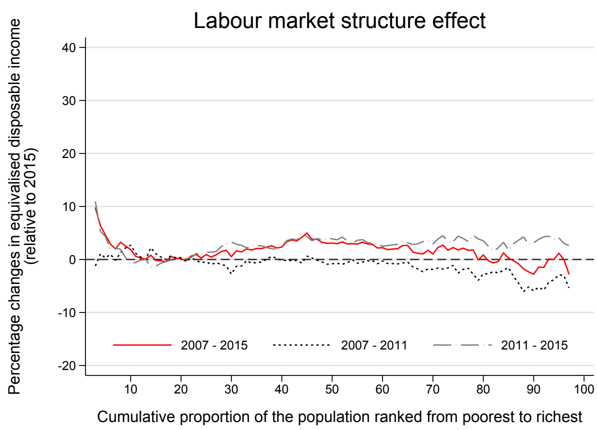
Third, although the demographic effect of disposable income is smaller than the other two, it appears to have persistently increased inequality over the analysed period. Because of the demographic effect, the bottom 30% of the population were 5% to 15% poorer in 2015 than if the demography would not have changed. At the same time, those at the top 30% became up to 5% richer. This represents up to 20 p.p. difference between the top and bottom quantiles, which is comparable to the tax and benefit effect on inequality. Additionally, the demographic effect increased inequality in both periods, although the effect on the lower quantiles was more pronounced in 2007-2011. It is likely that ageing, emigration and higher education achievement have all contributed to the demographic effect. At the time of writing, these three trends still persist and could mean still rising inequality.

Fourth, the labour market structure also affects different parts of the income distribution. The strongest effect was on the bottom 5% of households. Their incomes increased by about 10% of 2015 level over the period of 2007-2015, with most change happening in 2011-2015. The incomes of households in the middle of the income distribution also gained - about 5%. Interestingly, the top of the income distribution either did not gain or lost income because of changes in the labour market structure.

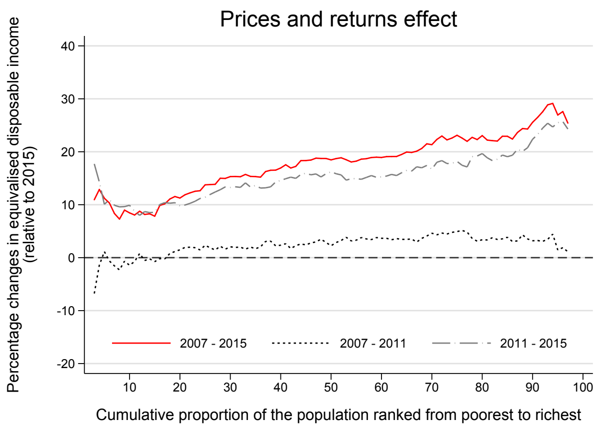
Figure 8 in the Appendix portrays the joint effect of the interaction effects and the residuals.



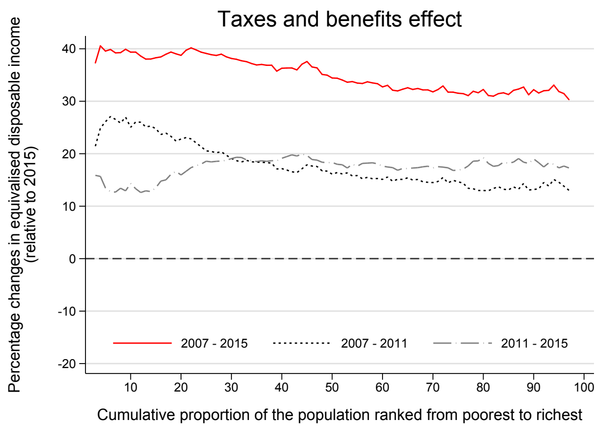
(a) Demographic effect



(b) Labour market structure effect



(c) Prices and returns effect



(d) Taxes and benefits effect

Figure 7: Decomposition of changes in the distribution of equivalised household disposable income

Decomposing changes in inequality and redistribution

Having investigated the contribution of the different factors to the evolution of households' incomes across the whole of the income distribution, we now turn to the analysis of how these factors contributed to the changes in market and disposable income inequality as well as absolute redistribution presented in Table 3. Results are given in Table 4. To better understand the drivers of the contributions of the labour market structure and returns factors, we provide some detail into their main components. As before, for each indicator, the total change in one period is equal to the sum of the contributions by all factors and the changes in the 2007-2015 period are equal to the sum of the changes in the 2007-2011 and 2011-2015 periods.

Market income inequality has grown significantly from 2007 to 2015. The Gini of market income grew by 4.2 p.p.. Interestingly, demographics contributed about half of this increase, particularly in the earlier sub-period of 2007-2011. Looking back to1, this was a period when the share of the elderly and those with tertiary education increased, suggesting that education, migration and ageing was behind this rise in inequality. Higher returns in the labor market also contributed. Although the effect of returns on the total period is much smaller (0.6 p.p.), returns was the largest contributor to market income inequality in the post-crisis sub-period (1.3 p.p.). This is not surprising, since the economy and wages were growing rapidly. Additionally, much of the increase to inequality over the whole period and pre-crisis period in particular is unexplained by the four considered factors, meaning that some other factors also played a role. Therefore, the only market factor that was inequality reducing were the changes in the labour market structure. However, they only partly offset the increase in income inequality.

The tax and benefit system managed to counter half of the increase in market income inequality. Specifically, the Gini of disposable income grew by 2.1 p.p.. As mentioned before, this is largely because of the tax benefit measures applied betwewen 2007-2011, over which disposable income inequality fell. As discussed in section 3.1, no additional measures were implemented in the second sub-period and many payouts, such as pensions, were frozen. This meant that tax and benefit system did not have any effect on the disposable Gini. As in the case when analysing market inequality, demographic changes resulted in greater inequality, while changes in the labour market structure resulted in lower disposable income inequality. What is different, however, is the much larger contribution to income inequality that is due to the returns effect. In particular, over the period after the crisis, the Gini of disposable income rose by 3.2 p.p. due to higher returns, which, as seen from Figure 7c, benefited the upper tail of the disposable income distribution. Again, we see a misalignment between rising market income in the latter period and stagnant tax and benefit system.

3.6 Summary and discussion of main findings

Results illustrate how the tax and benefit system responded to the rising market income inequality in Lithuania from 2007 to 2015. The tax and benefit system helped to reduce disposable income inequality during the financial crisis, which increased market income inequality. It was less able to do so during the recovery period of 2011-2015: disposable income inequality rose even though market income inequality remained at a similar, elevated level.

Looking at the two periods individually reveals that market incomes may change very rapidly, making it difficult to apply timely discretionary measures. During the recovery period, wages and capital income (referred to as prices and returns within the analysis) increased disposable income by 25% for the top of income distribution over 2011-2015, roughly 6 percent per year, while wages rose by just 10% for the bottom of income distribution.

Although the price and returns effect was the main contributor to market returns, especially during 2011-2015 period, other important factors play a role in increasing market income inequality in Lithuania. Specifically, the demographic effect persistently increased income inequality over the whole 2007-2015 period. Because of the demographic effect, the bottom of the income distribution became poorer, while the top - richer. It is likely that a shift towards higher education, emigration and ageing all relate to rising inequality in Lithuania.

With positive price and returns effects, a strong response from the tax and benefit system was wanted in order to redistribute these income changes. However, tax and benefit system did not reduce inequality in 2011-2015 but it may have actually reduced it. It is possible that the market incentives that the government enacted to encourage business and employment (increased retirement age and reduced income taxes) may have made the tax and benefit system less able to respond.

Table 4: Decomposition of changes in inequality and redistribution

	2007-2015	2007-2011	2011-2015
Gini Market Income			
Total change	0.042	0.040	0.002
Labour Market Structure	-0.015	-0.015	-0.001
Returns	0.006	-0.007	0.013
Tax-benfit system	0.002	0.004	-0.001
Demographics	0.020	0.017	0.003
Interactions	0.003	0.009	-0.006
Unexplained	0.026	0.033	-0.007
Gini Disposable Income			
Total change	0.021	-0.008	0.029
Labour Market Structure	-0.012	-0.017	0.005
Returns	0.030	-0.002	0.032
Tax-benefit system	-0.020	-0.021	0.000
Demographics	0.013	0.008	0.006
Interactions	0.017	0.002	0.016
Unexplained	-0.008	0.022	-0.030
Net Redistribution			
Total change	0.021	0.048	-0.026
Labour Market Structure	-0.003	0.003	-0.006
Returns	-0.024	-0.005	-0.019
Tax-benefit system	0.023	0.024	-0.001
Demographics	0.006	0.009	-0.003
Interactions	-0.014	0.007	-0.021
Unexplained	0.034	0.010	0.023

Notes: LMS: labour market structure; TB: tax-benefit system. Columns indicate the time period over which statistics where calculated (e.g. 2007-2011 refers to the change from 2007 to 2011).

4 Concluding remarks

In this paper we study the drivers of the changes in the distribution of income in Lithuania from 2007 to 2015 by adapting a model developed by Sologon et al. (2018) to assess the role played by changes in the labour market structure, in economic returns, in demographics and in tax-benefit rules. The case study of Lithuania is particularly interesting, given country's experienced transition from planed to market economy and its ongoing convergence to the EU and large fluctuations in disposable income. During this period, Lithuania faced the global financial crisis, over which household disposable income fluctuated severely, a series of tax and benefit reforms and net emigration and ageing. To effectively combat these income inequality fluctuations, one must first understand to what extent each of these factors are contributing to income inequality and see whether the tax and benefit system in place is able to tackle such challenges. Our research suggests that the tax and benefit system can only partly do so. It effectively reduced income inequality in 2007-2011 when market income inequality was rising, but not able to hold the rising inequality in the subsequent period, when market incomes were rising rapidly.

In 2007-2011, the tax and benefit system was well targeted at the very bottom of the income distribution. Therefore, those who lost work had access to relatively high unemployment benefits, maternity/paternity benefits, sick leaves, old age pension and other benefits. This was not the case in 2011-2015. On the contrary, disposable income inequality marginally increased over the period, perhaps due to market friendly reforms (prolong the pension age and reductions in income tax rates).

Several lessons can be drawn from the post 2007-2008 crisis Lithuanian experience. First, effective tax and benefits systems are able to reduce market income inequality even at a time when the economy is facing an economic crisis. Second, governments in transition countries should anticipate economic expansion following a crisis by carefully planning the redistributive system to tackle rapidly rising income inequalities.

Bibliography

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

Bargain, O. (2014). Decomposing changes in income distribution. *Handbook of Microsimulation Modelling, Contributions to Economic Analysis*, 293.

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

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

Biewen, M. (2014). Additive decompositions and interaction effects. *Applied Economics Letters*, 21(9).

Biewen, M. and Jenkins, S. P. (2005). A framework for the decomposition of poverty differences with an application to poverty differences between countries. *Empirical Economics*, 30(2).

Biewen, M. and Juhasz, A. (2012). Understanding rising income inequality in Germany, 1999/2000-2005/2006. *Review of Income and Wealth*, 58(4).

Bourgignon, F., Ferreira, F., and Leite, P. (2008). Beyond Oaxaca-Blinder: accounting for differences in household income distributions. *Journal of Economic Inequality*, 6(2).

Bourgignon, F. and Spadaro, A. (2006). Microsimulation as a tool for evaluating redistribution policies. *Journal of Economic Inequality*, 4(1).

Chantreuil, F. and Trannoy, A. (2013). Inequality decomposition values: the trade-off between marginality and efficiency. *The Journal of Economic Inequality*, 11(1).

Dardanoni, V. and Lambert, P. (2002). Progressivity comparisons. *Journal of Public Economics*, 86(1).

Devicienti, F. (2010). Shapley-value decompositions of changes in wage distributions: a note. *Journal of Economic Inequality*, 8(1).

DiNardo, J., Fortin, N., and Lemieux, T. (1996). Labor market institutions and the distribution of wages, 1973-1992: a semiparametric approach. *Econometrica*, 64(5).

Figari, F., Paulus, A., and Sutherland, H. (2015). Microsimulation and policy analysis. In Atkinson, A. B. and Bourguignon, F., editors, *Handbook of Income Distribution*, volume 2B, chapter 24. North-Holland.

Firpo, S., Fortin, N., and Lemieux, T. (2009). Unconditional quantile regressions. *Econometrica*, 77(3).

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

Juhn, C., Murphy, K., and Pierce, B. (1993). Wage inequality and the rise in returns to skill. *Journal of Political Economy*, 101(3).

Lerman, R. and 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).

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

Matzkin, R. (2003). Nonparametric estimation of nonadditive random functions. *Econometrica*, 71(5).

Paulus, A. and Tasseva, I. (2017). Decomposition of changes in the EU income distribution in 2007-2011. *EUROMOD Working Papers*, (9/17).

Reynolds, M. and Smolensky, E. (1977). Public expenditures, taxes and the distribution of income: the United States, 1950, 1961, 1970. *Academic Press, New York.*

Rothe, C. (2010). Nonparametric estimation of distributional policy effects. *Journal of Econometrics*, 155(1).

Rothe, C. (2012). Partial distributional policy effects. *Econometrica*, 80(5).

Shorrocks, A. (1980). The class of additively decomposable inequality measures. *Econometrica*, 48(3).

Shorrocks, A. (1982). Inequality decomposition by factor components. *Econometrica*, 50(1).

Shorrocks, A. (1984). Inequality decomposition by population subgroups. *Econometrica*, 52(1).

Shorrocks, A. (2013). Decomposition procedures for distributional analysis: a unified framework based on the Shapley value. *Journal of Economic Inequality*, 11(1).

Sologon, D., Van Kerm, P., Li, J., and O'Donoghue, C. (2018). Accounting for differences in income inequality across countries: Ireland and the United Kingdom. *LISER Working Papers*, No. 2018-01.

Sologon, D. M., Almeida, V., and Kerm, P. V. (2019). Accounting for the distributional effects of the 2007-2008 crisis and the economic adjustment program in portugal. *LISER Working Papers n° 2019-05, 52 p.*

Sutherland, H. and Figari, F. (2013). EUROMOD: the European Union tax-benefit microsimulation model. *International Journal of Microsimulation*, 6(1).

Van Kerm, P. (2013). Generalized measures of wage differentials. *Empirical Economics*, 45(1).

Van Kerm, P., Choe, C., and Seunghee, Y. (2016). Decomposing quantile wage gaps: a conditional likelihood approach. *Journal of the Royal Statistical Society (Series C)*, 65(4).

Additional Figures

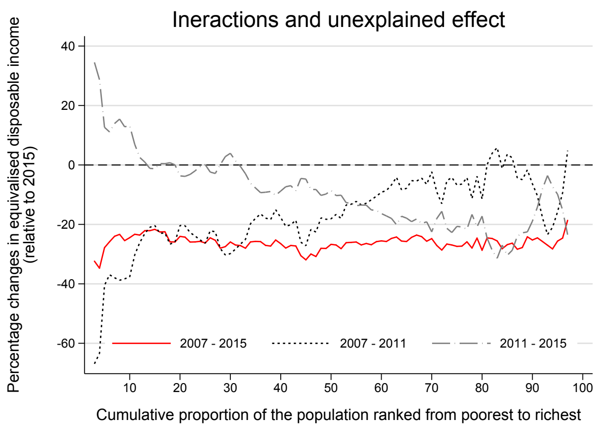


Figure 8: Interactions and unexplained effect