Joonas Ollonqvist

Accounting for the role of tax-benefit changes in shaping income inequality: A new method, with application to income inequality in Finland

Aboa Centre for Economics

Discussion paper No. 122

Turku

December 2021 (previous version December 2018)

The Aboa Centre for Economics is a joint initiative of the economics departments of the University of Turku and Åbo Akademi University.

ISSN 1796-3133

Printed in Uniprint
Turku
December 2021 (previous version December 2018)

Joonas Ollonqvist

Accounting for the role of tax-benefit changes in shaping income inequality: A new method, with application to income inequality in Finland

Aboa Centre for Economics
Discussion paper No. 122
December 2021 (previous version December 2018)

ABSTRACT

This paper introduces a new method of analysing how the changes in the tax-benefit-system have been reflected in income inequality. This method is a combination of microsimulation based decomposition (Bargain and Callan, 2010) and a multivariate regression based decomposition (Fields, 2003; Yun, 2006). It allows analysis of how the policy changes have affected the importance of different individual characteristics in income inequality. With the variance of log of incomes, the decomposition can be made further to separate the changes directly related to policy decisions from the overall price- and residual effects. This method is applied to analyse the evolution of income inequality in Finland from 1993 to 2014.

JEL Classification: D31, H24

Keywords: income inequality, microsimulation, regression, tax-benefit-system, decomposition

Contact information

Joonas Ollonqvist Department of Economics
University of Turku
FI-20014, Finland
Email: joonas.ollonqvist (at) utu.fi

Acknowledgements

I sincerely thank Kaisa Kotakorpi, Matti Viren, Markus Jäntti, Momi Dahan, Ilpo Suoniemi, Oskari Vähämaa and Erik Mäkelä for valuable comments and discussions. I thank the participants at the NORFACE WSF Final Conference (2018), D06: Inequality session in IIPF conference (2018), FDPE Public Economics and Labour Economics workshop II/2017 and ACE-workshop (2017) for their helpful comments and suggestions. I thank NORFACE for funding this research. All errors remain my own.

1 Introduction

The rise in inequality is one of the main concerns in the modern world. During the past decades, income inequality has risen substantially among the developed countries (OECD, 2011; Atkinson and Bourguignon, 2015) and Finland is not an exception in this matter (see **Figure 1**). There are many different reasons behind this evolution, but the question is not completely solved. Crudely speaking there are four different issues which affect income distribution: 1) changes in socio-demographic characteristics, 2) changes in the importance of different characteristics on an individual's income, 3) changes in politics and 4) business cycles. What makes the analysis difficult is that these factors may also interact with each other. Raising labour taxes, for example, lowers the value of being employed. From a policy maker's perspective especially, it is important to be able to separate the factors driving the evolution of economic inequality, since the policy actions will vary depending on the reason. Currently there is no method for isolating the effect of policy changes (changes in income inequality that are due to the changes in the tax-benefit-system) from the overall evolution of the importance of individual characteristics in income inequality. To fill this gap in the literature, I propose a new method, which is a synthesis of two decomposition methods used in the literature.

Usually, the distributional effects of policy changes are analysed by using microsimulation based methods (Bargain and Callan, 2010; Bargain, 2012a,b; Herault and Azpitarte, 2016) and the contributions of different characteristics to income inequality can be analysed by a multivariate regression based decomposition (Fields, 2003; Yun, 2006).[1] The focus in the microsimulation literature is on questions like '*How have changes in the tax-benefit-systems affected income distribution?*'. With the microsimulation methods, it is possible to isolate the policy effect (effect that is due to changes in tax-benefit legislation) on income inequality from the other effects and these methods can also be extended to analyse the behavioural effects of the policy changes; however, they do not reveal anything about how these changes affect the importance of different characteristics in income inequality. In contrast, the multivariate regression based decomposition reveals answers to the question '*How do individual/household characteristics contribute to income inequality?*'. Moreover, it is possible to further analyse whether the evolution has been driven by *price effect* (part of the change in income inequality that is explained by the change in the importance of a variable on income), *quantity effect* (the part of the change in income inequality caused by the change in the distribution of a characteristic among the

[1] For an extensive overview of other decomposition methods used analysing distribution of incomes see Fortin et al. (2011).

population) or residual effect. The disadvantage of this method is that it cannot distinguish whether or not the changes are driven by the changes in politics.

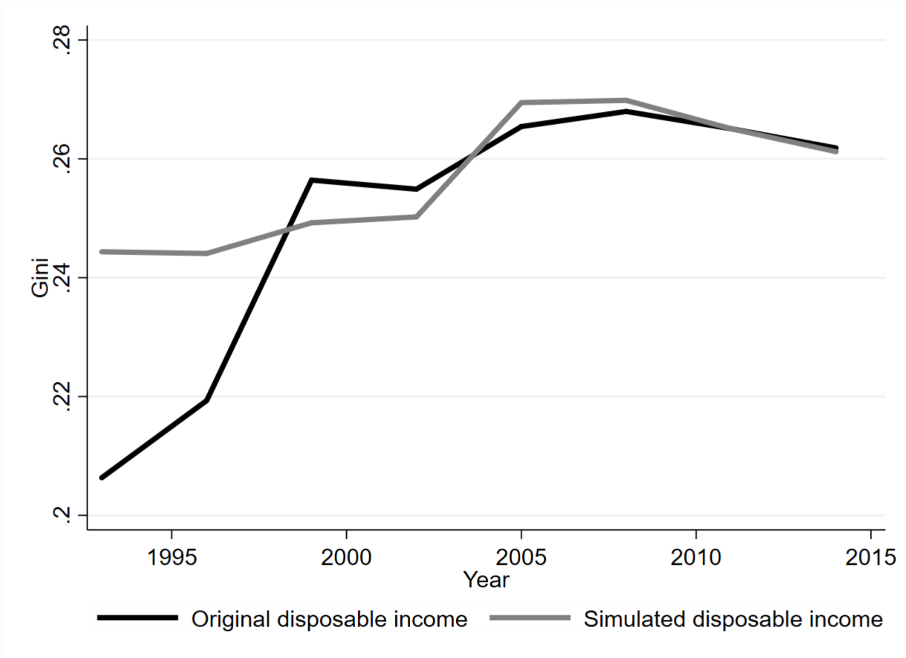
My proposed method combines the benefits of both methods. This new method conveys how the changes in the tax-benefit-system have affected the importance of each characteristic as regards income inequality. With the variance of log of incomes, the policy effect can be isolated from the total price- and residual effects, providing more information about the reasons behind the evolution of these two items. For instance, it is possible to analyse how much of the change in the education premium is explained by the tax-benefit changes and how it has affected the income distribution. However, as is typical for the decomposition methods, the causality of the results cannot be guaranteed.

The proposed method is applied to study the evolution of income inequality in Finland and **Figure 1** shows how the Gini coefficient of disposable income has evolved in Finland from 1993 to 2014. The evolution is shown in two situations: 1) Using the actual data values (black line) and 2) using simulated (counterfactual) disposable incomes (grey line). The latter one is formed by using the 2011 data and simulating the disposable income of the households according to each year's tax-benefit legislation using Statistics Finland's SISU-microsimulation model (Statistics Finland, 2014). In other words, it describes what would be the level of income inequality if only the tax-benefit legislation had changed conditional on the population of 2011. Meaning that it shows the contribution of tax-benefit changes on income inequality.

From **Figure 1** can be seen that, during the 1990s inequality rose rapidly in Finland and since 2005 only small changes have happened. It also shows that policy changes have increased income inequality, which has been observed also in earlier studies conducted with Finnish data (Bargain and Callan, 2010; Honkanen and Tervola, 2014).[2] However, as the actual change in the Gini coefficient is larger than the one with the simulated incomes, only a part of the change in income inequality can be explained by the policy changes. One explanation presented in the literature is the rise in capital incomes among the top earners (Riihelä et al., 2010). Additionally, the collapse of Soviet Union produced a large unexpected shock at the Finnish economy and changed its composition (Gorodnichenko et al., 2012). Moreover, the macroeconomic situation in Finland has changed and varied considerably since 1993 as can be seen from **Figure 2**. Real GDP per capita, employment and unemployment rates developed into a favourable direction until the 2008 economic

[2]This is somewhat different from the observations found in the international context (see Hills et al., 2014 and Figari et al., 2015).

crises, but since 2008 the direction has reversed and negated part of the earlier development. Finally, the distribution of socio-demographic characteristics in Finland have changed considerably since the early 1990's. For example, in 1993 only 4.6% of over 15-year-olds had master's degree or higher while in 2014 the percentage had become 9.5% (Statistics Finland, 2021b). These changes in the distribution of the characteristics may have altered the income distribution, but also the importance of these characteristics in individual incomes may have changed as well.

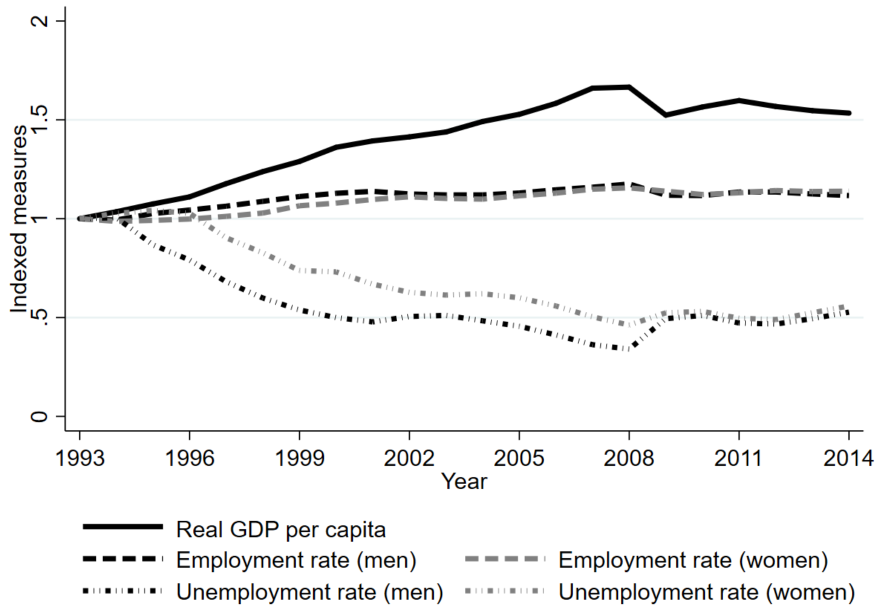


Note: Majority of the simulations are done using data from year 2011. For the year 2011 actual data is used.
Source: Author's own calculation based on service data of income distribution and the SISU-model

Figure 1: Gini coefficient in Finland

For these reasons, I am using this proposed method to investigate the evolution of income inequality in Finland.[3] I start by analysing how different individual/household characteristics contribute to income inequality and how the contribution have been altered by the changes in tax-benefit-system. The second aim is to study whether the changes in income inequality can be explained by the changes in the distribution of the characteristics (quantity effect) or by the changes in the importance of the characteristics on income (price effect). Furthermore, the role of the residual term and the effect of tax-benefit changes are investigated as well. My final

[3]One reason to concentrate on total income inequality than, for instance, poverty is that the proposed method can fully be used only with the variance of log of incomes as an inequality measure. Nonetheless, it would be important to analyse how different characteristics affect poverty, but for now this is left for the future.



Note: All the values have been scaled by the base year values, thus in year 1993 every measure receives the value 1. Employment and unemployment rates are calculated among 15 to 64 years old.
Source: Statistics Finland (2021a,c) and author's own calculation. Values used in the calculations are shown in **Table H.1**.

Figure 2: Macroeconomic developments in Finland

aim is to analyse the channel by which policy changes have altered income distribution. Since the evolution of income inequality and the macroeconomic development differ substantially before and after 2005, the analysis is conducted separately for both periods as well. This allows me to investigate the possible reasons why the trends in income inequality differ before and after 2005.

I find that price, quantity, policy and residual effects all explain a major share of the evolution in income inequality from 1993 to 2014. Before 2005 the price, policy and residual effects were the key drivers of the income inequality, but since 2005 those have equalised the income distribution. Quantity effects have, however, increased the income inequality during the whole time period and since 2005 quantity effects have generated more inequality than before 2005. I also find that policy changes have mostly affected the income inequality by changing the importance of individual characteristics and prior 2005 around 2/3 of the price effects can be traced to policy changes. Nevertheless, policy changes have also affected income distribution in a way that cannot be explained.

This paper is organised as follows: Section 2 presents both the microsimulation- and multivariate regression based decompositions and the synthesis of these two.

Section 3 discusses the data and the empirical strategy and shows the decomposition results for Finland. Then Section 4 discusses the robustness of the findings and the final section presents the conclusions.

2 Unified framework

2.1 Decomposition of policy effect with microsimulation

First, defining some notation and terminology. The socio-economic characteristics of households in year j are described by vector $\mathbf{X}_j$ and their original (gross) incomes (in year j) are denoted by vector $\mathbf{Y}_j$. Following Bargain and Callan (2010) and Figari et al. (2015), I distinguish between the tax-benefit function (e.g. the rules of the taxation and benefits) and the monetary parameters (e.g. tax brackets). Tax-benefit-system k is a function defined as $f_k(\mathbf{X}, \mathbf{Y}, m_k)$, where parameters m_k are the monetary parameters used in the tax-benefit-system. Household disposable income in year k with the tax-benefit-system from year j is then

$$\gamma_j(\mathbf{X}_k, \mathbf{Y}_k, m_j) = \mathbf{Y}_k + f_j(\mathbf{X}_k, \mathbf{Y}_k, m_j)$$

In this paper, I restrict attention to the static (i.e. non-behavioural) effects of policy changes, but it is also possible to take into account the behavioural (indirect) effects of policy changes as is done by Bargain (2012a,b). Therefore, the direct effect of policy changes from A to B on household disposable income, while keeping the original incomes and characteristic unchanged, is

$$\Delta\gamma = \gamma_B(\mathbf{X}_A, \mathbf{Y}_A, m_B) - \gamma_A(\mathbf{X}_A, \mathbf{Y}_A, m_A)$$

This is the so called "morning after" policy effect. These kinds of calculations are usually used to form counterfactuals in a so called "what if" setup. For example, an analysis of what household's disposable income in year B would be if we had the tax-benefit-system from year A, is denoted as γ_B^A (To shorten notation $\gamma_A^A = \gamma_A$). However, in these kinds of studies the policy parameters may not be directly adapted to other years, since prices and incomes may have changed.[4] Therefore some kind of adjustment to policy parameters is needed and α is used as this adjusting factor.[5]

[4]Without the adjustments of policy parameters, for instance, the tax-brackets may not correspond with the level of incomes and therefore simulated taxes might be unrealistically low or high. If the tax-brackets are unrealistically low the tax burden for individuals with earnings is higher than it realistically should be.

[5]How to choose the parameter α is discussed in Hills et al. (2014)

When using end period B data (denoting the initial period with A) the change in disposable income is:

$$\begin{aligned}\Delta\gamma =&\gamma_B(\mathbf{X}_B, \mathbf{Y}_B, m_B) - \gamma_A(\mathbf{X}_B, \mathbf{Y}_B, \alpha m_A)\\ =&\gamma_B - \gamma_B^A\end{aligned}$$

and with the initial period data A is:

$$\begin{aligned}\Delta\gamma =&\gamma_B(\mathbf{X}_A, \mathbf{Y}_A, \alpha^{-1} m_B) - \gamma_A(\mathbf{X}_A, \mathbf{Y}_A, m_A)\\ =&\gamma_A^B - \gamma_A\end{aligned}$$

Where the subscripts of γ denote the year of original incomes and characteristics (i.e. population) and the superscripts of γ indicate the year of the policy-parameters and the tax-benefit function.

Then using similar notation, the effect of these policy changes to some inequality measure I is

$$\begin{aligned}\Delta I =&I\left[\gamma_B(\mathbf{X}_B, \mathbf{Y}_B, m_B)\right] - I\left[\gamma_A(\mathbf{X}_B, \mathbf{Y}_B, \alpha m_A)\right]\\ =&I_B - I_B^A \qquad (1)\end{aligned}$$

and

$$\begin{aligned}\Delta I =&I\left[\gamma_B(\mathbf{X}_A, \mathbf{Y}_A, \alpha^{-1} m_B)\right] - I\left[\gamma_A(\mathbf{X}_A, \mathbf{Y}_A, \alpha m_A)\right]\\ =&I_A^B - I_A \qquad (2)\end{aligned}$$

As Bargain and Callan (2010) point out, this decomposition gives the absolute policy effect on income distribution and it is possible to conduct the decompositions with either end or base year data. They also argue, that if there is access to both the end and base year data, the relative policy effect can be formed by using the Shorrocks-Shapley decomposition.[6] The Shorrocks-Shapley decomposition is in this case just the average of these two effects:

$$\begin{aligned}\Delta_P =&\frac{1}{2}\left[I_B - I_B^A\right] + \frac{1}{2}\left[I_A^B - I_A\right]\\ =&\frac{1}{2}\left[I_B - I_B^A + I_A^B - I_A\right] \qquad (3)\end{aligned}$$

[6] See for example Shorrocks (2013).

In other words, the above equations capture the direct policy effect on income inequality when the other factors are kept constant.

2.2 Multivariate regression based decomposition

Following the work of Fields (2003), the observed inequality can be decomposed to the contributions accounted for by each household/individual characteristics. This decomposition can be formed with virtually any inequality measure. However, according to Yun (2006), with the variance of log of incomes the decomposition can be taken further.

The procedure of the decomposition is simple. First, the income generating function is estimated by using the OLS:

$$y_i = \sum_{c=0}^{N} \beta_c X_{ci} + \varepsilon_i$$

where $y_i = \ln(\gamma_i)$ is the log of household disposable income, β:s are the regression coefficients, X:s represent the set of household/individual characteristics ($X_i \in \mathbf{X} \quad \forall i$) and ε is the error term. To ease the notation, I suppress the individual subscripts in the equations.

Then the fitted values of the estimation are used to form the relative characteristics inequality weights:

$$s_c = \frac{\operatorname{cov}(\beta_c X_c, y)}{\sigma^2(y)}, \tag{4}$$

These weights are invariant of the choice of inequality measure and the share of the residual can be calculated the same way. These weights give the relative contribution of each characteristic and the residual term on income inequality.

The absolute contribution of each characteristic on income inequality is just the product of the calculated weight and the value of the inequality measure: $S_c = s_c I$.[7] Then, the absolute change in income inequality explained by some characteristic c is

$$\Delta S_c = s_{cB} I_B - s_{cA} I_A$$

[7]Here I indicates an inequality measure that is calculated using disposable income.

And the total change can be expressed as:

$$\Delta I = \sum_{c=1}^{N} (s_{cB} I_B - s_{cA} I_A) \tag{5}$$

Where subscript A denotes the base period and B denotes the end period.

Yun (2006) shows that it is possible to take the decomposition further in order to separate the quantity-, price- and residual effects, when using the variance of the log of incomes as the inequality measure.[8] Then the decomposition takes the form:

$$\begin{aligned}\Delta\sigma_y^2 =& \sum_{c=1}^{N} \left(s_{cB}\sigma_{y_B}^2 - s_c^* \sigma_{y^*}^2\right) + \sum_{c=1}^{N} \left(s_c^* \sigma_{y^*}^2 - s_{cA}\sigma_{y_A}^2\right) + \sigma_{\varepsilon_B}^2 - \sigma_{\varepsilon_A}^2 \\ =& \Delta^Q + \Delta^P + \Delta^\varepsilon \end{aligned} \tag{6}$$

where A and B are defined as before, σ_y^2 is the variance of the log of incomes, superscript $*$ refers to values that are formed using an auxiliary income distribution, where the coefficients of characteristics are replaced while keeping the characteristics intact. Formally defined as:

$$y^* = \sum_c \beta_{cB} X_{cA} + \varepsilon_A \tag{7}$$

The first terms in **equation** (6) capture the *quantity effects*, the second ones are the *price effects* and the last terms present the change in inequality given by the changes in the residual.[9] Price effect is the change in income inequality that is explained by a change in the importance of a variable on income (the rise (decrease) in the price effect means that the particular variable has become the more (less) important determinant of an individual's income). Whereas the quantity effect of income inequality is caused by the change in the distribution of a characteristic among the population (the rise (decrease) in the quantity effect is caused by the particular variable becoming more (less) unequally distributed among the population).

There is also another possible way to form the auxiliary income distribution, it can be formed by replacing the characteristics while keeping the coefficients intact:

$$y^{**} = \sum_c \beta_{cA} X_{cB} + \varepsilon_B \tag{8}$$

[8]One problem with the variance of log of incomes is that it cannot be guaranteed to satisfy the Pigou-Dalton principle of transfers.

[9]Due to construction of OLS, $s_\varepsilon \sigma_y^2 = \sigma_\varepsilon^2$.

With the auxiliary income distribution defined in **equation** (8) the decompositions of the price-, quantity and residual effects takes the form:

$$\begin{aligned}\Delta\sigma_y^2 =& \sum_{c=1}^{N}\left(s_{cB}\sigma_{y_B}^2 - s_c^{**}\sigma_{y^{**}}^2\right) + \sum_{c=1}^{N}\left(s_c^{**}\sigma_{y^{**}}^2 - s_{cA}\sigma_{y_A}^2\right) + \sigma_{\varepsilon_B}^2 - \sigma_{\varepsilon_A}^2 \\ =& \Delta^P + \Delta^Q + \Delta^\varepsilon \end{aligned} \tag{9}$$

Therefore the price- and quantity effects presented can be calculated in three ways: 1) using **equation** (6), 2) using **equation** (9) or 3) taking the average of these two. The last one corresponds to the Shorrocks-Shapley decomposition and it uses more information than the other two. In addition, there is no particular reason to prefer the first or the latter decomposition and therefore in this paper the price- and quantity effects are formed using the average of **equations** (6) and (9).[10]

The Fields' method can be extended similar to the factor source decomposition (presented in Shorrocks (1982)) to distinguish between the pure- and interaction effects of characteristic c in income inequality.[11] For this reason, it is possible to decompose the price- and quantity effects in a similar way.[12]

2.3 Synthesis

In this paper, I only consider the static effects of policy changes and therefore a policy change has two ways of affecting the income distribution: i) it may alter the importance of the characteristics on individual income (i.e. price effect) and/or ii) change the explanatory power of the characteristics (i.e. change the residual).[13] For this reason, combining these two methods is straight forward to do to form the price- and residual effects of the policy change on each characteristic. As before, there are three possible combinations that can be analysed: 1) using end period data, 2) using initial period data or 3) using both.

First, the income generating function is estimated and the relative characteristic inequality weights are formed according to **equation** (4) with the simulated and original datasets. Then, accordingly combining **equations** (5) and (1) gives the

[10] In **Appendix G** are shown the price- and quantity effects formed according to **equations** (6) and (9).

[11] Details are shown in **Appendix B**.

[12] Proof is shown in **Appendix C**

[13] In a behavioural setting, it could also have an impact on the gross incomes and distribution of characteristics, namely labour market status.

decomposition with the end period data.

$$
\begin{aligned}
\Delta I_{PB} =& \sum_{c=1}^{N} \left[s_{cB} I_B - s_{c_B^A} I_B^A \right] + s_{\varepsilon_B} I_B - s_{\varepsilon_B^A} I_B^A \\
=& \Delta_{PB}^{P} + \Delta_{PB}^{\varepsilon},
\end{aligned}
\tag{10}
$$

where subscript B indicates which year's population is used in the decomposition. Similarly, for the initial period data the following decomposition holds:

$$
\begin{aligned}
\Delta I_{PA} =& \sum_{c=1}^{N} \left[s_{c_A^B} I_A^B - s_{c_A} I_A \right] + s_{\varepsilon_A^B} I_A^B - s_{\varepsilon_A} I_A \\
=& \Delta_{PA}^{P} + \Delta_{PA}^{\varepsilon}
\end{aligned}
\tag{11}
$$

and when having access to both initial and end period data we arrive at the following:

$$
\begin{aligned}
\Delta I_{PAB} =& \frac{1}{2} \left[\sum_{c=1}^{N} \left[s_{c_B} I_B - s_{c_B^A} I_B^A + s_{c_A^B} I_A^B - s_{c_A} I_A \right] \right] \\
&+ \frac{1}{2} \left[s_{\varepsilon_B} I_B - s_{\varepsilon_B^A} I_B^A + s_{\varepsilon_A^B} I_A^B - s_{\varepsilon_A} I_A \right] \\
=& \Delta_{PAB}^{P} + \Delta_{PAB}^{\varepsilon}
\end{aligned}
\tag{12}
$$

In all of the above decompositions, the term inside the sum operator is the change in the absolute contribution of each characteristic c accounted for by the policy change. This change can be interpreted as the price effect of the policy change on income inequality, since the population and original incomes are kept intact.[14] The second terms capture the change in the contribution of the residual term explained by the policy change.

Before this, I have not made any specific assumptions about the inequality measure. It is possible to conduct the above decompositions with virtually any well behaving inequality measure as is the case with Fields' decomposition. When using the variance of the log of incomes as the inequality measure it is possible to compare the total price effect obtained from **equation** (6) with the price effect explained by the change in the tax-benefit-system (**equations** (10), (11) and (12)). The same applies to the residual effect.

The price effect calculated by the **equation** (6) or (9) is the total price effect between periods A and B conditional on the base or end year data. These can be expressed as a sum of the change explained by the policy changes and the change

[14]This is shown formally with the variance of logs in **Appendix A**.

accounted for by other factors and the same holds true for their average as well. Therefore, with the variance of the log of incomes the following holds true:

$$\Delta^P = \Delta_P^P + \Delta_O^P, \tag{13}$$

Where Δ^P is the total price effect, Δ_P^P is the part of price effect explained by the direct policy effects and Δ_O^P is the price effect explained by other factors than the direct policy effects. Δ^P and Δ_P^P are obtained from the earlier decompositions and, thus, it is possible to calculate the Δ_O^P. The same can be done with the residual effect:

$$\Delta^\varepsilon = \Delta_P^\varepsilon + \Delta_O^\varepsilon, \tag{14}$$

As before, Δ_P^P and Δ_P^ε can be calculated either with initial or end period data or using both. However, both initial and end period data is required to form the total price- and residual effects. Moreover, the price effect can be formed in three possible ways (**equation** (6), (9) or both). Therefore, there are 9 possible combinations to form Δ_O^P and three possible combinations to form Δ_O^ε.

Even though, I am only studying the total effects of policy changes, the same equations can be applied when studying the effects of some of the changes in the tax-benefit-system have. For example, it could be in our interest to study separately the effects of the changes on the benefit side or taxation side.

3 Application to income inequality in Finland

3.1 Data and empirical strategy

The analysis is performed with triennial cross-sectional data from 1993 to 2014. The data used is the service data of income distribution collected by Statistics Finland, which can be used (from year 2011 onwards) with the Finnish microsimulation model SISU (Statistics Finland, 2014). The yearly sample size is around 25 000 individuals in approx. 10 000 households and it includes a large amount of information about the individual/household characteristics and their incomes.

The analysis is conducted in two steps. First, the counterfactual datasets are formed using microsimulation and after that the simulated and original datasets are analysed using the regression methods. In the following paragraphs, I will explain the assumptions made in the analysis in more details.

Starting with the microsimulation. Counterfactuals are formed by using the SISU microsimulation model. The majority of the Finnish tax-benefit-system is encoded in the SISU model from the year 1993 onwards.[15] However, as the SISU model is not compatible with the data before 2011 I cannot calculate the policy effect of each sub period according to **equation** (1) or (2) as it would require simulating either end or initial period data. Therefore, I have to approximate the policy effect. This is done by carrying out all the simulations using data from the year 2011. The following example illustrates how the approximation is done.

First keeping the same notation as before and thinking of three years: 1993, 1996 and 2011, where the year 2011 is the only one applicative to microsimulation. The policy changes from the years 1993 and 1996 to 2011 are then:

$$\Delta_{P1} = I_{2011} - I_{2011}^{1993}$$
$$\Delta_{P2} = I_{2011} - I_{2011}^{1996}$$

Then the difference between these two give the approximation of the effect of policy change from the year 1993 to 1996 conditional on the data from the year 2011:

$$\Delta_P = I_{2011}^{1996} - I_{2011}^{1993}$$

In this paper, the above calculation is used to approximate the effect of policy change from 1993 to 1996. The same approach is used to approximate the policy effect of each subsequent period before the year 2008.

As mentioned earlier, the value of money and incomes change over time and for this reason the monetary parameters of tax-benefit legislation from one year are not necessarily comparable with data from other year. Therefore, the policy parameters are typically adjusted according to some adjusting factor (denoted as α in **equations** (1) and (2)). There are at least three different possibilities how to choose the adjusting factor α: 1) the Consumer price index (CPI), 2) the Market income index (MII) or 3) no indexation (i.e. using the nominal values). The differences between these three and how they affect the counterfactuals are discussed more details in Hills et al. (2014).

In the main specification the CPI is used, for several reasons. First, benefits in Finland, if they are tied, are tied to the CPI. Second, using CPI means that every family can afford to buy the same basket of goods over time. Subsequently, the policy changes give a better reflection of the welfare changes associated with the

[15] One limitation of the SISU-model is that earnings-related pensions cannot be simulated as it would require information about the individuals' past earnings. **Appendix D** shows more details about the SISU-model and how the simulations are carried out.

changes in legislation. Third, CPI is a middle option of these three, as it is typically smaller than MII. However, it has its downsides. First, not every benefit is tied to an index. Meaning that, adjusted benefit levels may be higher than they really should be. The second problem is related to growth in incomes. If the incomes grow at a higher rate than the CPI, then the tax-brackets grow at a slower pace than earnings.[16] Indicating that the tax revenues grow faster than the expenditures of benefits. However, since the economic theory does not say what the correct indexation strategy is, I conduct the analysis also with the other two indexations as well, in order to test the robustness of the findings.

In the second step, both the actual data years and counterfactual datasets are analysed according to Fields' and Yun's methods. In the analysis I make several assumptions where I mostly follow the example of Brewer and Wren-Lewis (2016). First, all the characteristics in the analysis are transformed to a set of indicator variables. Finally, the relative shares of indicator variables are summed up together to form the total effect of each characteristic. This is possible, because of the additive nature of the method. Second, the decompositions are conducted with the log of household's disposable income scaled by the modified OECD equivalence scale. Third, the household heads are used in the analysis, but the household head is given the information about any spouses. Fourth, the decompositions are done at the individual level by using household level weights multiplied with the number of people in the household. Fifth, only variables that are found in each set of data are used in the analysis. Sixth, there have been changes in the variables during the time span of the analysis and therefore some variables are recoded to make them comparable between each data year. The most crucial change occurred in 1997 when the definition of the level of education changed. This change mostly affected the education listed in either secondary schooling or the lowest/lower tertiary level. It is impossible to identify the field of education before 1997 and therefore in the analysis secondary schooling is combined together with the lowest/lower tertiary level.

Table 1 provides some descriptive statistics about selected variables from the sample. It can be seen that there are changes in the distribution of characteristics. The percentage of single households, well-educated individuals and pensioners have increased from 1993. In addition, the mean age of both genders have increased. The share of unemployed individuals for both genders is smaller in the year 2014 than it was in 1993. For both genders, the share of employed individuals was higher in 2005 than in 1993 or in 2014. For men, however, the share of employed individuals

[16] In Finland tax-brackets are not juridically tied to any index. Those are, however, increased 'manually' each year more or less according to market income index.

was smaller in 2014 than in 1993 whereas for women the share was larger in 2014.

Table 1: Descriptive statistics about the sample

	Female	Single hhs	Mean age		Employed		Unemployed		Pensioners		Education	
Year			Male	Female	Male	Female	Male	Female	Male	Female	Male	Female
1993	52.8%	36.8%	46.1%	48.2%	58.4%	48.7%	14.0%	16.3%	24.7%	30.1%	6.6%	4.4%
2005	52.7%	38.6%	48.6%	50.0%	62.0%	52.8%	8.0%	11.5%	26.7%	31.2%	9.0%	8.2%
2014	52.2%	41.0%	50.3%	51.8%	55.9%	50.5%	9.3%	11.0%	30.3%	33.0%	11.4%	12.8%

Note: Well-educated refers to those with upper tertiary level degree or higher. Individual level characteristics are calculated separately for men and women including all age groups. Employed refers to both employed and self-employed individuals. Single hhs refers to single households. Author's own calculations based on the service data of income distribution.

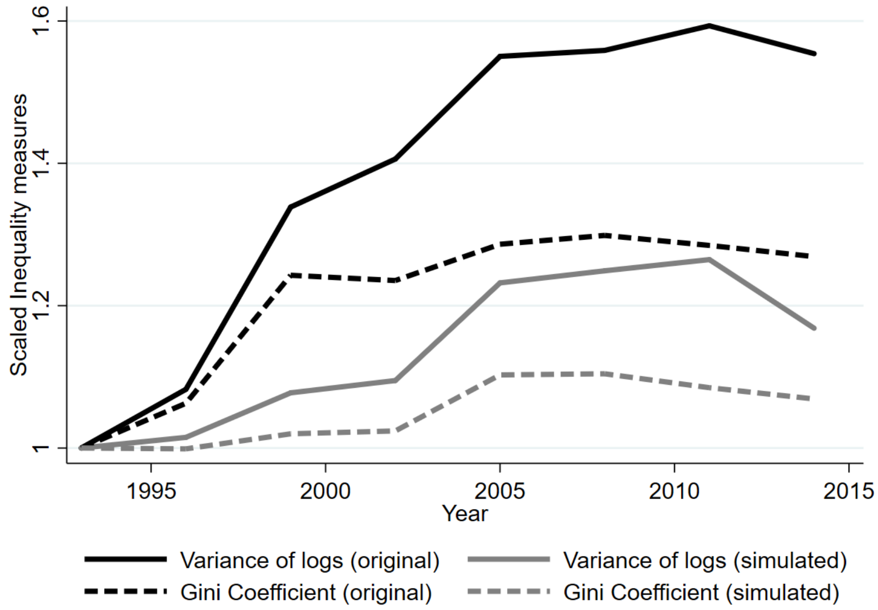
3.2 Results

To better illustrate the evolution of income inequality **Figure 3** presents the indexed evolutions of the variance of log of disposable income and the Gini coefficient for both the actual (black lines) and simulated data (grey lines). Again, the simulated data series display the role of tax-benefit changes on income inequality conditional on the 2011 data (i.e. those show what would be the level of income inequality with the 2011 population if only the tax-benefit legislation had changed).

According to the **Figure 3** inequality has risen with every specification from 1993 to 2014 and the majority of the increase happened before 2005. Since 2005 the level of income inequality has remained the same or slightly decreased. Still, there can be seen differences between the specifications. First, the relative increase in income inequality is larger with the actual data set than with the simulated one. This means that the policy changes have increased income inequality (with both measures), but those can explain only part of the change in income inequality. Second, the relative increase in income inequality is larger when inequality is measured with the variance of log of incomes than with the Gini coefficient, which can be seen with both datasets. Despite this difference the choice of the inequality measure should not alter the results drastically as the overall trends with both measures and datasets are fairly similar.

Now inequality is decomposed by six individual (age, employment status and education by both genders) and two household characteristics (region and household type). The decomposition is done by using the characteristics of the household heads and the characteristics of any partner of the household head. For each of the characteristics, indicator variables are created according to which subgroup the individual belongs to. Then, the total contribution by each group is formed as a sum of the shares of the indicator variables belonging to that particular group.[17]

[17]See **Appendix E** for full information of the variables.



Note: Simulations are conducted using the 2011 data. All the values have been scaled by the base year values, thus in year 1993 every measure receives the value 1.
Source: Author's own calculation based on service data of income distribution and the SISU-model

Figure 3: Scaled inequality measures for actual and simuated data in Finland

Table 2 presents the share of inequality explained by each characteristic and the residual with the actual and simulated data (formed according to **equation** (4)) for the years 1993, 2005 and 2014.[18] These years were chosen mainly since the inequality rose rapidly before the year 2005 and has remained roughly the same since. These shares are invariant for the choice of the inequality measure. However, those do not take into account the changes in the level of income inequality. In other words, the characteristics may generate more (or less) income inequality even though the share explained has remained the same. In **Appendix G** are shown the 95% confidence intervals of these shares for each year.

For both datasets, the error term, employment status, education and the age of males explain large fractions of the observed income inequality, whereas the region, household type and the age of females explain a smaller part. Actual changes in the relative contributions of household type, education, age of males and male employment status are also statistically significant (shown in **Appendix G**). Whereas the changes in the shares calculated using the simulated data are rather modest and not statistically significant. Indicating that the policy changes have altered the relative importance of the characteristics on income inequality only mildly if at all.

[18]The shares and absolute contributions for each year are shown in **Appendix F**.

According to **Table 2** the largest share with both datasets is explained by the residual meaning that there is considerable variation that cannot be explained using these characteristics. However, the share of the residual has substantially decreased meaning that the explanatory power of the variables used has increased. In 1993 only 39% could be explained, but in the year 2014 the same variables explained around 49% of the income inequality. This finding is also statistically significant. With the simulated data the relative contribution of the residual has decreased only by around 2.7 percentage points and since 2005 it has remained nearly the same. However, the findings are not found to be statistical significant. Still, the results suggest that before 2005 part of the change in the explanatory power of the variables may be related to the policy changes, but since 2005 policy changes have not affected the explanatory power of the characteristics.

Table 2: Shares of characteristics in income inequality (%)

Actual data									
			Household	Age		Employment status		Education	
Years	Residual	Region	type	Male	Female	Male	Female	Male	Female
1993	61.0	3.2	4.4	5.8	3.8	5.3	7.1	5.5	4.0
2005	56.9	1.8	3.0	5.6	3.1	7.5	7.8	8.6	5.7
2014	51.1	2.2	1.9	9.3	4.4	9.4	7.2	7.9	6.6
Simulated data									
			Household	Age		Employment status		Education	
Years	Residual	Region	type	Male	Female	Male	Female	Male	Female
1993	57.7	1.9	2.3	8.1	3.1	7.7	7.8	6.4	5.0
2005	55.1	1.9	1.9	7.7	2.8	9.2	8.4	7.6	5.4
2014	55.0	2.0	2.4	7.9	3.1	8.6	8.3	7.4	5.3

Note: Simulated years are formed using data from the year 2011. 95% confidence intervals are shown in **Appendix G**. Author's own calculations based on the service data of income distribution.

Employment status makes the highest relative contribution to income inequality among the characteristics and it has become even a more important determinant in income inequality since 1993. In 1993, it accounted for around 12.4% (5.3% among men and 7.1% among women) of the income inequality and in 2014 it was increased to 16.2% (9.4% among men and 7.2% among women). Its role among men has increased considerably over time, while the contribution of women employment status to inequality has remained roughly the same since 1993. With the simulated data, the pattern is similar before 2005, indicating that part of the rise in the relative contribution of employment status may be associated with the policy changes. However, after 2005 policy changes may have slightly decreased the relative contribution of employment status of men to income inequality.

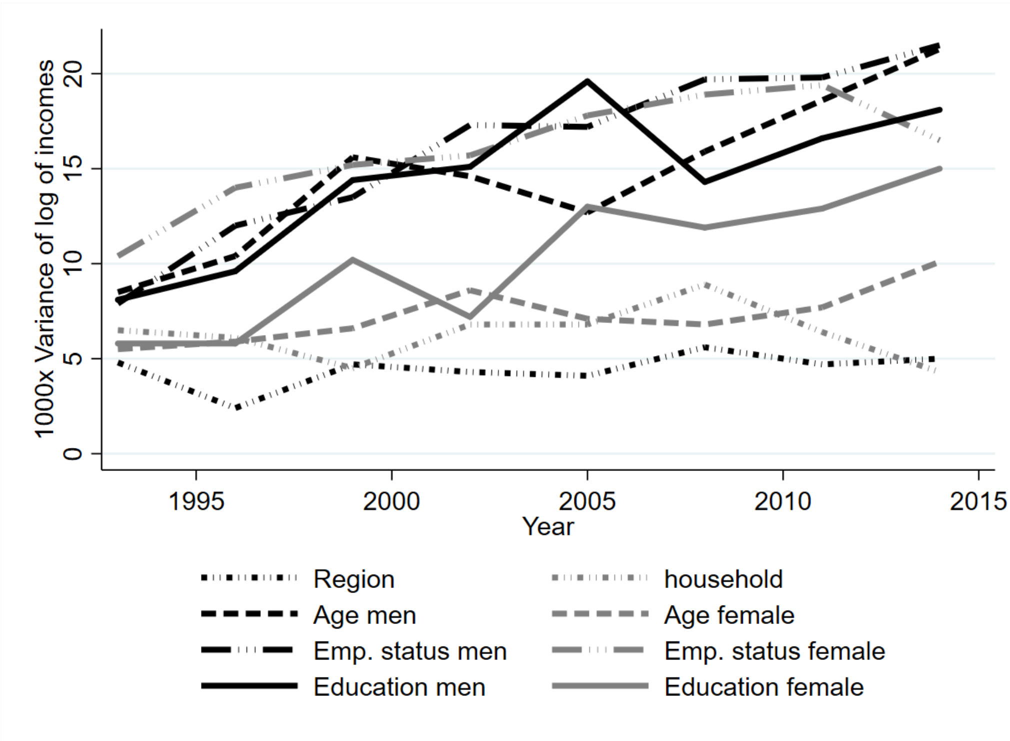
For both genders, the relative contribution of education to income inequality has increased, but the patterns of the evolution partly differ between genders. In 1993, education accounted for around 9.5% of the income inequality and in 2014 the share was around 14.5%. The relative contribution of male education to income inequality increased rapidly from 1993 (5.5%) to 2005 (8.6%), but have slightly decreased since that (7.9% in 2014). For women, the relative contribution of education to income has been rising more steadily since 1993 from 4.0% to 6.6%. The shares obtained using simulated data vary only mildly. For males the share has slightly increased from 1993 to 2005, but otherwise there is no indication that the policy changes have altered the relative importance of education in income inequality.

The relative contribution of age has increased substantially since 2005, but it was mildly decreasing before 2005. In 1993, age explained around 9.6% of the income inequality and in 2014 the share was increased to 13.7%. This change is mostly driven by men as their share increased from 5.8% to 9.3% whereas the contribution of age of women increased only by 0.6 percentage points from 1993 to 2005. Interestingly the age of men contributes to income inequality as much as the employment status of men and it is contributing more to inequality than the education of men. Whereas among women it is the opposite as education and employment status are both contributing more to income inequality than age. With the simulated data, the shares have remained the same indicating that the changes in tax-benefit system have not altered the relative importance of age in income inequality. However, the SISU-model does not simulate earnings related pensions, which potentially underestimates the role of policy changes in shaping the importance of age in income inequality.

Household type accounts for a significantly smaller share in inequality in 2014 (1.9%) than in 1993 (4.4%). The annual decompositions reveal (shown in the **Appendix F**), that at the beginning of the 2008 crisis its contribution peaked. This indicates that the recession potentially had a different effects according to the type of the household. Still, the relative contribution of household type has decreased significantly and in 2008 the relative contribution was no higher than in 1993. With the simulated data, the role of the household type has not changed since 1993, indicating that the policy changes seem to not have affected the relative contribution of household type in income inequality.

The smallest share is accounted for the region, which explained around 3.2% of the income inequality in 1993. In 2005, the share was decreased to 1.8% but has slightly increased since 2005 (2.2% in 2014). These changes cannot be traced to policy changes as there is no variation in the shares calculated using the simulated data.

In **Figure 4** is shown the absolute contribution of each characteristic on income inequality with the 1000× the variance of log of incomes as the inequality measure using actual data.[19] For the simulated data the changes in the absolute contribution of each characteristic are shown in **Table 4** and the absolute contributions are shown in **Table F.4**. **Figure 4** tells the similar story as **Table 2** with the difference that it takes into account the change in the inequality measure as well. As the income inequality, measured with the variance of log of incomes, has increased since 1993, the absolute contributions of the characteristics have a clearer upward trend compared with the relative contributions. Still, the same characteristics are the most important as were in **Table 2**. The employment status, education and the age of both genders are contributing more to income inequality in 2014 than in 1993 and these increases are also statistically significant (shown in **Table G.5**). Furthermore, the decrease in the absolute contribution of household type is also statistically significant.



Source: Author's own calculation based on service data of income distribution.

Figure 4: Absolute contribution of characteristics on income inequality

Next, the absolute contributions of each characteristic are decomposed to the part caused by the changes in the importance of the characteristics on income (*price effect*), part caused by the changes in the distribution of the characteristics among

[19]These are shown also in **Table F.3**.

population (*quantity effect*) and *residual effect* according to the average of **equations** (6) and (9). These results together with the total policy effect are shown in **Table 3**. After that, the price- and residual effects are further decomposed to separate the part explained by the policy changes (see **equations** (13) and (14)) from the part explained by the other effects. These findings are reported in **Table 4**. The analysis is also conducted separately for the years before and after 2005, since the trends in income inequality and macroeconomic development are different in those periods. However, it should be noted that the choice of the time period will affect the results. To be transparent in this matter, the annual decompositions are shown in **Appendix F**.[20]

I will first go through the total effects on both tables and then discuss more detailed about the findings related to each characteristic. According to the **Table 3**, approximately 42% of the change in income inequality from 1993 to 2014 is explained by the changes in the importance of the characteristics on income (i.e. price effects) and it is solely traced to the pre 2005 era. This means that the variables used have become more important determinants in individual incomes and therefore are generating more income inequality in 2005 than in 1993. However, since 2005 price effects have had nearly no effect on income inequality. The changes in the distribution of the e characteristics among the population (i.e. quantity effects) have also had an important role in the change of income inequality. The quantity effects account around 25% for the total change, but opposite to the price effects those have increased income inequality both before and after 2005. Moreover, the changes in the distribution of characteristics generated more inequality after 2005, indicating that the distributional changes of the characteristics have affected income inequality more drastically since 2005. Around one-third of the total change in income inequality could not be explained (i.e. residual effect). Before 2005 around half of the increase in income inequality could not be explained, but since 2005 the role has reversed. Since 2005, the changes in the error term have equalised the income distribution.

Comparison with the macroeconomic development (shown in **Figure 2**) reveals that the changes in real GDP per capita may mostly be associated with the price effects. At least those have evolved rather similarly, as before 2005 the real GDP per capita rose substantially and at the same time price effects made substantial increase to income inequality. Furthermore, in 2014 the real GDP per capita was around the same as in 2005 and during that period price effects have not affected the income inequality. Of course, the changes in real GDP per capita may have been

[20]One thing to be noted is that, price- and quantity effects are not time additive, whereas the other effects calculated are. Therefore, for these two effects, the yearly decompositions presented in Appendix F will not sum up to the results shown in **Table 3**.

Table 3: Changes in income inequality, price-, quantity and residual effects (1000x variance of log of incomes)

Years	Total Change	Price Effect	Quantity Effect	Residual Effect	Policy Effect	Region		Household type		Age				Employment status				Education			
										Male		Female		Male		Female		Male		Female	
						P	Q	P	Q	P	Q	P	Q	P	Q	P	Q	P	Q	P	Q
1993 - 2005	81	31.5	9.4	40.2	32.7	-0.5	-0.2	0.3	0.1	1.8	2.5	-0.4	1.9	12.2	-2.9	6.8	0.7	7.9	3.6	3.6	3.7
2005 - 2014	0.6	-1	14.5	-13	-9.9	0.9	0.2	-2.8	0.3	7	1.6	1.5	1.6	0.3	4.1	-2.6	1.2	-4	2.5	-1.2	3.2
1993 - 2014	81.6	33.9	20.6	27.2	22.8	0.2	0	-1.8	-0.4	8.4	4.4	1.2	3.4	14.1	-0.5	4.5	1.6	4.8	5.3	2.5	6.7

Note: Negative values indicate that on average the effect has negative contribution to income inequality. The price and quantity effects are not time additive and therefore the effects of the sub-periods do not sum up to the total effect. Annual decompositions are presented in **Appendix F**.
Source: Author's own calculations based on the service data of income distribution.

Table 4: Price and residual effects decomposed to policy and other effects (1000x variance of log of incomes)

Years	Price effect		Residual effect		Region		Household type		Age				Employment status				Education			
	Total								Male		Female		Male		Female		Male		Female	
	Δ_O^P	Δ_P^P	Δ_O^ε	Δ_P^ε	Δ_O^P	Δ_P^P	Δ_O^P	Δ_P^P	Δ_O^P	Δ_P^P	Δ_O^P	Δ_P^P	Δ_O^P	Δ_P^P	Δ_O^P	Δ_P^P	Δ_O^P	Δ_P^P	Δ_O^P	Δ_P^P
1993 - 2005	11.7	19.8	27.2	12.9	-1.2	0.7	0.4	-0.1	0.1	1.7	-0.8	0.4	6.3	5.9	3	3.8	3	4.9	1	2.6
2005 - 2014	3.3	-4.3	-7.4	-5.6	0.9	-0.1	-3.7	0.9	7.2	-0.3	1.1	0.3	2.5	-2.2	-1.4	-1.1	-2.7	-1.2	-0.5	-0.7
1993 - 2014	18.4	15.5	19.8	7.3	-0.4	0.6	-2.6	0.8	7	1.4	0.5	0.7	10.4	3.7	1.8	2.7	1.1	3.6	0.6	1.9

Note: Negative values indicate that on average the effect has negative contribution to income inequality. Subscript P refers to the part accounted to direct policy changes and subscript O refers to other effects. Annual decompositions are presented in **Appendix F**. Simulations conducted using the 2011 data.
Source: Author's own calculations based on the service data of income distribution and SISU-model.

associated with the residual- and quantity effects (or with employment rates), but the connection does not seem to be as strong as with the price effects. Changes in employment and unemployment rates may at least partly be linked to quantity effects as those directly influence the quantity effects of employment status. This will be discussed more when concentrating on the findings related to employment status.

The policy effects behave similarly to the residual effect. In total, it accounts for about 28% of the change, but these are already included in the price- and residual effects. Policy changes increased inequality before 2005 and since then those have reduced income inequality. The same is also true with the Gini coefficient (shown in **Appendix G**).

When looking at the channels by which policy changes have affected the inequality (shown in **Table 4**) some interesting aspects arise. First of all, during the whole time span, the policy changes have mostly affected the income distribution by changing the importance of different characteristics (around two-thirds of the policy effect). Nevertheless, around one-third of the policy effect has been affecting the income distribution in a way that cannot be explained (i.e. through residual effect). Since 2005, the channels have changed a little and over half of the equalising effect of the policy changes has happened through the residual effect.

Before 2005, the policy changes also accounted for around 2/3 of the total price effect. Since 2005, policy changes have reduced the price effects, but other effects have continued to increase the price effects. Overall, these two can almost be seen to cancel each other out. Related to the residual effect the numbers are somewhat different. Prior 2005 policy changes explained around one-third of the total residual effect and after 2005 the percentage has increased to around 43%.

The following concentrates more on the different characteristics. Overall the same variables as in **Table 2** are the most important, but there are some additional features related to them. From 1993 to 2005 education accounted for around 23% of the total change in inequality but since 2005 education has had nearly a zero effect. The results are also qualitatively similar to both genders. The majority of the effect of education on income inequality is explained by the changes in the distribution of education and during the whole time period quantity effects of education have generated more inequality. This is not surprising since the share of well-educated individuals in both genders has increased substantially from 1993 (shown in **Table 1**). Still, before 2005 education also became a more important determinant in income, which, especially among men, generated more inequality. Since 2005, however, the price effects of education have decreased income inequality, which almost negated the effect of quantity effects on income inequality.

The results for education may also be linked to the fact that the field of education and the working sector differ between genders and that well-educated females may have a shorter work history compared with male counterparts.[21] These, however, cannot be reliably analysed with the data at hand. Furthermore, the larger price effects of education of men and the direction of price effects are consistent with the changes in real GDP per capita. This is because men tend to work more on the private sector, where earnings usually grow faster. Furthermore, recessions typically have larger impact on the private sector than the public sector. Finally, the change in the classification of education that occurred in 1997 may affect the results for the first half of the analysis.

Employment status accounts for around 21% of the change in inequality before 2005. Since 2005, it has continued to increase the inequality, but the size of the effect has become substantially smaller. The changes in the employment status of men are solely due to the price effect and it alone accounts for 42% of the total price effect. For women, the price effect also drives the contribution, but the magnitude is substantially smaller and since 2005 those have mildly reduced inequality. Furthermore, these findings are consistent with the changes in real GDP per capita.

Quantity effect of the male employment status decreased inequality before 2005, but the sign of the contribution has reversed since it. This is no surprise as the employment rate increased and unemployment rate decreased rapidly before 2005 and those have evolved in an opposite direction since 2005. However, the magnitude of the quantity effect is surprisingly small before 2005 given the large changes in employment and unemployment rates. Among women the quantity effect of employment status is only slightly positive in both periods. This is little surprising, but it may be that the other changes in the distribution of employment status have partly offset the effect of the changes in employment and unemployment rates.

Prior 2005, the policy changes accounted about half of the price effect of employment status among both genders. Since 2005 policy changes have slightly decreased the price effect and for women it has amplified the other effects. For men, the policy changes nearly completely offset the other effects.

Here also, the role of age has several interesting features that explain the changes in inequality. The age of men has clearly affected the evolution whereas the contribution of the age of females is substantially lower. Nonetheless, the age of both genders can be seen to be generating more inequality in 2014 than in 1993. The dif-

[21]The last one is due to fact that the education level has increased more rapidly among women than among men.

ference between genders is mostly due to changes occurring after the 2005. Before 2005, ages of both genders together accounted for only around 7% of the change in inequality and the majority of the change was explained by the changes in the age distribution among population. Since 2005, the age of men has become a more important factor in the income of individuals and this is generating more inequality, but the same cannot be said about the age of women. This solely explains the gender differences in the contribution of age since the quantity effect is almost identical between the genders. Moreover, according to **Table 1** the mean ages of both genders have increased since 1993.

The difference in price effect cannot be explained by the policy changes, but one possible explanation may be that income composition of an individual is correlated with age and gender. For instance, men receive more income from self-employment and capital, which usually tend to increase over age. In addition, older age groups are more likely receiving pensions whereas younger age groups, especially middle-aged, are more likely to receive employment income. This may also explain why the price effect of age of men increased income inequality while the economy was in a recession or stagnated as recessions have varying effects on different income sources. Pensions, for instance, are less likely to be affected by recessions compared with market incomes. Furthermore, the earnings of men might increase faster with age than the earnings of women due to differences in working histories, profiles and sectors. For example, women typically stay at home to take care of a new born child and tend to work more at the public sector whereas men are working more at the private sector.

Household type is the only variable that has equalised the income distribution. This is mostly due that household type has become a less important factor in determining individuals' income. Policy changes seem not to have altered the way household type is generating income inequality. In addition, changes in the household structure among population have had nearly no effect on income inequality. Therefore, it seems that the increase in the proportion of single households has not increased income inequality at the aggregate level.

As expected from the earlier results, the region have not contributed to the change in income inequality at all.

4 Robustness of the results

I conduct several robustness checks to check the sensitivity of the results. First, I study how sensitive the results are for the tails of the income distribution. To test it, I conduct the analysis by excluding the top and bottom one percent of the income distribution. These results are shown in **Table 5**. Not surprisingly the changes in income inequality are smaller, but despite that the results are extremely similar to the main specification. All the effects account for roughly the same percentage of the change and the overall patterns are similar to the main specification. Furthermore, it seems that the difference in the total change is rather uniformly distributed to each of the characteristics.

As discussed earlier, the classification of education changed in the middle of time span of this study (in 1997), which mostly affected the education categorization in either secondary schooling or a lower tertiary level. To make the data years comparable with each other secondary schooling was combined with the lowest/lower tertiary level in the main specification. Therefore, the changes in the distribution of the characteristics are smaller than those should be. This potentially decreases the quantity effect and increases the price effect.

When focusing more on the policy changes, there are at least four reasons why the results of the policy effect may be sensitive. First, as Bargain and Callan (2010) mention, the longer the time span between the data and the legislation year the more inaccurate the results will be. This inaccuracy can be reduced by using more data to minimize the distance between the legislation and data year. Additionally, using the both initial and period data will reduce inaccuracy, but, unfortunately, neither one is possible with the available microsimulation model.

Second, the choice of the indexation will affect the results. To test how sensitive the results are for the choice of the indexation of policy parameters, the analysis is conducted using two other indexations: 1) the market income index (MII) and 2) no adjustment (i.e. using the nominal values). These results are shown in **Table 6**. Overall the policy effects calculated using the MII are larger. This is as expected since it is usual for wages to increase faster than prices and therefore the policy parameters are adjusted with larger coefficients. Therefore, individuals with earnings are better off than individuals with benefits. The opposite can be seen with the nominal values, which yields clearly smaller policy effects than the CPI. In addition, the total policy effect with nominal values from 1993 to 2014 had a negative contribution to income inequality. However, with all three different indexations the policy changes can be seen to have increased income inequality before 2005 and decreased it since. Similar kind of findings can be seen when comparing the channels of the policy changes. With the MII the price effect is almost entirely traced to policy changes whereas non can be traced with the nominal values. Again, the results with the CPI is midway between these two. These findings together with the fact that

Table 5: Changes in income inequality, excluded top and bottom 1% (1000x variance of log of incomes)

Years	Total Effect	Price Effect	Quant Effect	Residual Effect	Policy Effect	Region		Household type		Age				Employment status				Education			
										Male		Female		Male		Female		Male		Female	
						P	Q	P	Q	P	Q	P	Q	P	Q	P	Q	P	Q	P	Q
1993 - 2005	56.5	27	6.4	23.1	28.8	-0.3	-0.3	0.1	0	2.5	2.1	-0.3	1.4	10.4	-3.3	6.6	0.6	5.1	2.8	3	3.1
2005 - 2014	14.7	0.3	11.3	3	-7.3	1	0.2	-2.3	0.2	4.4	1.4	0.6	1.2	-0.1	3	-1	0.9	-1.6	1.9	-0.6	2.7
1993 - 2014	71.1	29.6	15.4	26.1	21.4	0.6	0.1	-1.8	-0.4	6.9	3.4	0.6	2.5	11.3	-1.3	5.9	1.1	3.8	4.4	2.4	5.8

Note: Top and bottom 1% of the income distribution are excluded from the sample. Negative values indicate that on average the effect has negative contribution to income inequality. Author's own calculations based on the service data of income distribution.

Table 6: Price and residual effects decomposed to policy and other effects using different indexation (1000x variance of log of incomes)

Years	Policy effect Total	Price effect Total		Residual effect		Region		Household type		Age				Employment status				Education			
										Male		Female		Male		Female		Male		Female	
		Δ_O^P	Δ_P^P	Δ_O^ε	Δ_P^ε	Δ_O^P	Δ_P^P	Δ_O^P	Δ_P^P	Δ_O^P	Δ_P^P	Δ_O^P	Δ_P^P	Δ_O^P	Δ_P^P	Δ_O^P	Δ_P^P	Δ_O^P	Δ_P^P	Δ_O^P	Δ_P^P
Market income index (MII)																					
1993 - 2005	45.2	-0.1	31.6	26.6	13.6	-1.6	1.1	0.3	0	-1.6	3.3	-0.9	0.5	3.3	8.9	-0.4	7.2	1.3	6.5	-0.4	4
2005 - 2014	-1.9	-1	0.1	-11	-2	0.7	0.1	-3.9	1.1	6.7	0.3	1	0.5	1.4	-1.1	-2.5	0	-3.3	-0.6	-1.1	-0.1
1993 - 2014	43.3	2.2	31.7	15.5	11.6	-1	1.3	-2.9	1.1	4.8	3.6	0.2	1	6.3	7.8	-2.7	7.2	-1.2	5.9	-1.4	3.8
Nominal values (i.e. $\alpha = 1$)																					
1993 - 2005	22.4	17.7	13.8	31.6	8.6	-0.9	0.4	0.7	-0.4	0.8	0.9	-0.5	0.1	7.8	4.4	4.7	2	3.6	4.3	1.6	2
2005 - 2014	-29.6	14.2	-15.1	1.6	-14.5	1.3	-0.5	-3.2	0.5	8.7	-1.7	1.4	0	5.3	-5	1.3	-3.8	-1.3	-2.6	0.8	-1.9
1993 - 2014	-7.2	35.2	-1.3	33.1	-5.9	0.3	-0.1	-1.9	0.1	9.2	-0.8	1	0.1	14.7	-0.6	6.3	-1.8	3.1	1.6	2.4	0

Note: In the upper part of the table policy parameters are indexed using market income index and in the second half nominal values are used. Negative values indicate that on average the effect has negative contribution to income inequality. Author's own calculations based on the service data of income distribution.

most of the benefits are juridically tied to the CPI and that the tax-brackets follow the MII indicate that the policy effect calculated with the CPI indexation should offer a rather conservative estimate for the policy effect.

Third, the Pigou-Dalton principle of transfers cannot be guaranteed to hold with the variance of log of incomes. Using more frequent data cannot solve this problem and it is an undesirable feature of the method proposed here. Fortunately, it seems that the calculated policy effects are fairly similar to the Gini coefficient and the variance of log of incomes (see **Figure 3** and **Appendices F** and **G**) and this should, therefore, not drastically change the results.

The fourth issue is the measurement error, which can be due to several issues. First, the modelling is done using annual incomes, which will underestimate the eligibility to certain benefits. This is because the eligibility to some benefits, e.g. social assistance, is determined according to monthly incomes. Second, the non-take-up of the benefits are not taken into account. In the simulations, it is assumed that all the eligible individuals would apply for the benefit, but in reality this is not the case. This will, therefore, overestimate the benefits. Finally, there might be inaccuracy in the data or in the microsimulation model. Since there are possible reasons for both under and overestimation, it is difficult to say in which direction the possible measurement error drives the results.

Finally, I estimate the policy effect using the Gini-coefficient and calculate the 95% confidence intervals for the shares explaining income inequality and absolute contributions of each characteristics in income inequality. Because of the sheer volume of the robustness checks, these are only displayed in the Appendices.

5 Conclusions

In this paper, a new method was introduced for analysing how policy changes affect the income distribution. This new method combines the benefits of microsimulation- and regression based decompositions. It allows for an analysis of how policy changes have affected the importance of socio-demographic characteristics in income inequality. With the variance of log of incomes, the decomposition can be made further to isolate which part of the price- and residual effects are explained by the policy changes and which part is not. This new method was applied to study how different characteristics have affected the income distribution in Finland since 1993. The first aim was to quantify the contributions to income inequality of different individual/household characteristics and how policy changes have affected their contribution. The second aim was to analyse whether the changes in the importance of the characteristics in individual incomes (i.e. price effects) or the changes in the distribution of the characteristics among population (i.e. quantity effects) have driven the evolution of income inequality. In addition, the aim was to investigate how the error term and policy changes in a static setting have influenced income inequality. The final aim was to investigate how much policy changes have altered the price effect of individual/household characteristics.

The price-, quantity- and residual effects made an important contribution to the total change in inequality from 1993 to 2014. The price effects accounted for the largest percentage of the change (about 42%), quantity effect accounted for around 25% and around one-third of the change could not be explained. This indicates that both the changes in the distribution of the socio-demographic characteristics among population and the changes in the importance of these characteristics explain a significant part of the rise in inequality. However, the magnitudes and directions of the effects varied over time. The price and residual effects were the main drivers of the inequality before 2005, but after 2005 the price effects had almost no impact on inequality while residual effect has substantially lowered the inequality. The quantity effects, on the other hand, continued to increase the inequality in the second half and it even became the dominant driver of the inequality.

The results for policy changes were in line with the earlier findings of Honkanen and Tervola (2014). Policy changes accounted for about 28% of the increase in income inequality in Finland since 1993. However, this result is very sensitive and the policy changes have not increased inequality during the whole time span. Since 2005, policy changes have equalized the income distribution. About two-thirds of the policy effects were affecting trough the price effects of different characteristics and the remainder of policy changes were affecting the income inequality through the residual.

The most interesting finding was that before 2005 around 2/3 of the total price effect could be explained by the policy changes. After 2005, policy changes were reducing the price effect, but other effects continued to increase the price effect. The same does not apply to the residual effect, where the majority of the changes remained unexplained.

I also found interesting features in the contributions of different characteristics on inequality. For instance, the most important characteristic was the employment status of men and it was solely driven by the price effect; whereas the changes in the distribution of level of education made a substantial increase in income inequality. However, the results for education may have been affected by the change in classification which occurred in 1997. Furthermore, the price effects were mostly following the changes in real GDP per capita and the changes in employment and unemployment rates were partly consistent with the quantity effect of employment status. It should be, however, noted that, the results obtained in this paper are more descriptive in nature, as the causality of the results cannot be guaranteed.

In the analysis, additional sensitivity checks were also made. Firstly, the tails of the income distribution were excluded and the result remained qualitatively the same. The overall change in inequality were smaller, but it did not alter much the relative importance of different effects or characteristics. Secondly, two different indexations were used in the simulations. Market income indexation yielded larger policy effect whereas nominal values yielded for smaller policy effect than the main specification.

This new method does not, at least not yet, allow for the studying how of policy changes have affected the quantity effect. It can be the case that policy changes have had indirect effects that may have altered the distribution of the characteristics, namely the employment status, and it would be interesting to analyse these effects. Moreover, the focus of this paper was on the total income inequality and it would be beneficial to study also how the characteristics have influenced on poverty. For the time being these must be left for future research.

References

Atkinson, A. B. & Bourguignon, F. (2015) Introduction: Income distribution today. In *Handbook of Income Distribution*, eds. A. B. Atkinson & F. Bourguignon, vol. 2 of *Handbook of Income Distribution*, xvii – lxiv, Elsevier, URL http://www.sciencedirect.com/science/article/pii/B9780444594280099896.

Bargain, O. (2012a) Decomposition analysis of distributive policies using behavioural simulations. *International Tax and Public Finance*, vol. 19 (5), 708–731.

Bargain, O. (2012b) The Distributional Effects of Tax-benefit Policies under New Labour: A Decomposition Approach. *Oxford Bulletin of Economics and Statistics*, vol. 74 (6), 856–874.

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

Brewer, M. & Wren-Lewis, L. (2016) Accounting for changes in income inequality: Decomposition analyses for the uk, 1978–2008. *Oxford Bulletin of Economics and Statistics*, vol. 78 (3), 289–322, URL http://dx.doi.org/10.1111/obes.12113.

Fields, G. S. (2003) *Worker Well-Being and Public Policy (Research in Labor Economics, Volume 22)*, chap. ACCOUNTING FOR INCOME INEQUALITY AND ITS CHANGE: A NEW METHOD, WITH APPLICATION TO THE DISTRIBUTION OF EARNINGS IN THE UNITED STATES, 1–38. Emerald Group Publishing Limited.

Figari, F., Paulus, A. & Sutherland, H. (2015) Chapter 24 - microsimulation and policy analysis. In *Handbook of Income Distribution*, eds. A. B. Atkinson & F. Bourguignon, vol. 2 of *Handbook of Income Distribution*, 2141 – 2221, Elsevier, URL http://www.sciencedirect.com/science/article/pii/B978044459429700025X.

Fortin, N., Lemieux, T. & Firpo, S. (2011) Decomposition methods in economics. In *Handbook of Labor Economics*, eds. O. Ashenfelter & D. Card, vol. 4A, chap. 01, 1–102, Elsevier, 1 edn., URL https://EconPapers.repec.org/RePEc:eee:labchp:4-01.

Gorodnichenko, Y., Mendoza, E. G. & Tesar, L. L. (2012) The finnish great depression: From russia with love. *American Economic Review*, vol. 102 (4), 1619–44.

Herault, N. & Azpitarte, F. (2016) Understanding changes in the distribution and redistribution of income: A unifying decomposition framework. *Review of Income and Wealth*, vol. 62 (2), 266–282, URL http://dx.doi.org/10.1111/roiw.12160.

Hills, J., Paulus, A., Sutherland, H. & Tasseva, I. (2014) A lost decade? Decomposing the effect of 2001-11 tax-benefit policy changes on the income distribution in EU countries. ImPRovE Working Papers 14/03, Herman Deleeck Centre for Social Policy, University of Antwerp.

Honkanen, P. & Tervola, J. (2014) Vero- ja tulonsiirtojräjestelmän vaikutus tulonjakoon suomessa 1995-2013. *Yhteiskuntapolitiikka*, vol. 79, 306–317, in Finnish.

OECD (2011) Growing income inequality in oecd countries: What drives it and how can policy tackle?

Riihelä, M., Sullström, R. & Tuomala, M. (2010) Trends in top income shares in finland 1966–2007. Research Reports 157, Government Institute for Economic Research Finland (VATT).

Shorrocks, A. (2013) Decomposition procedures for distributional analysis: a unified framework based on the shapley value. *The Journal of Economic Inequality*, vol. 11 (1), 99–126.

Shorrocks, A. F. (1982) Inequality decomposition by factor components. *Econometrica*, vol. 50 (1), pp. 193–211.

Statistics Finland (2014) *SISU-malli - Käyttöopas tulonsiirtojen ja verotuksen mikrosimulointiin*. Official Statistics of Finland (OSF), in Finnish.

Statistics Finland (2021a) *Official Statistics of Finland (OSF): Annual national accounts [e-publication]*. Helsinki, issn=1798-0623 edn., URL https://www.stat.fi/til/vtp/tau_en.html, referred: 27.10.2021.

Statistics Finland (2021b) *Official Statistics of Finland (OSF): Educational structure of population [e-publication]*. Helsinki, issn=2242-2919 edn., URL http://www.stat.fi/til/vkour/index_en.html, referred: 27.4.2021.

Statistics Finland (2021c) *Official Statistics of Finland (OSF): Labour force survey [e-publication]*. Helsinki, issn=1798-7857 edn., URL http://www.stat.fi/til/tyti/tau_en.html, referred: 27.10.2021.

Yun, M.-S. (2006) Earnings inequality in usa, 1969–99: Comparing inequality using earnings equations. *Review of Income and Wealth*, vol. 52 (1), 127–144, URL http://dx.doi.org/10.1111/j.1475-4991.2006.00179.x.

Appendices

A Effect of the policy change in price- and residual effects

To show that policy effect really is either price- or/and residual effect, the counterfactuals need to be formed first. Then the change between counterfactual and actual data is decomposed according **equation** (6). With the end, period prices it takes the form:

$$\begin{aligned}\Delta\gamma =& \gamma_B(\mathbf{X}_B, \mathbf{Y}_B, m_B) - \gamma_A(\mathbf{X}_B, \mathbf{Y}_B, \alpha m_A) \Rightarrow \\ \Delta I(\gamma) =& I[\gamma_B(\mathbf{X}_B, \mathbf{Y}_B, m_B)] - I[\gamma_A(\mathbf{X}_B, \mathbf{Y}_B, \alpha m_A)]\end{aligned}$$

Now the change can be decomposed like before to the price-, quantity-, and residual effects:

$$\Delta\sigma_y^2 = \sum_{c=1}^{N}\left(s_{cB}\sigma_{y_B}^2 - s_c^*\sigma_{y^*}^2\right) + \sum_{c=1}^{N}\left(s_c^*\sigma_{y^*}^2 - s_{c_B^A}\sigma_{y_B^A}^2\right) + \sigma_{\varepsilon_B}^2 - \sigma_{\varepsilon_B^A}^2$$
$$= \Delta^Q + \Delta^P + \Delta^\varepsilon$$

Where the income generating functions are estimated with OLS as:

$$\begin{aligned}y_B &= \sum_c \beta_{cB}X_{cB} + \varepsilon_B \\ y_B^A &= \sum_c \beta_{cB}^A X_{cB}^A + \varepsilon_B^A\end{aligned}$$

And y^* is defined as:

$$y^* = \sum_c \beta_{cB}X_{cB}^A + \varepsilon_B^A$$

Since X_{cB}^A is equal to X_{cB} (the same data set used in the analysis), the above equations take the forms:

$$y_B = \sum_c \beta_{cB} X_{cB} + \varepsilon_B \tag{15}$$

$$y_B^A = \sum_c \beta_{cB}^A X_{cB} + \varepsilon_B^A \tag{16}$$

$$y^* = \sum_c \beta_{cB} X_B + \varepsilon_B^A \tag{17}$$

Therefore,

$$\begin{aligned}
s_{cB}\sigma_{y_B}^2 =& \text{cov}(\beta_{cB}X_{cB}, y_B) = \text{cov}(\beta_{cB}X_{cB}, \sum_c \beta_{cB}X_{cB} + \varepsilon_B) \\
=& \text{cov}(\beta_{cB}X_{cB}, \sum_c \beta_{cB}X_{cB}) + \beta_{cB} \overbrace{\text{cov}(X_{cB}, \varepsilon_B)}^{=0 \text{ eq. (15)}} \\
=& \text{cov}(\beta_{cB}X_{cB}, \sum_c \beta_{cB}X_{cB}) + \beta_{cB} \overbrace{\text{cov}(X_{cB}, \varepsilon_B^A)}^{=0 \text{ eq. (16)}} \\
=& s_c^*\sigma_{y^*}^2, \quad \forall c \Rightarrow \\
\Delta^Q =& \sum_{c=1}^N \left(s_{cB}\sigma_{y_B}^2 - s_c^*\sigma_{y^*}^2 \right) = 0
\end{aligned}$$

Whereas the price- and residual effects won't (necessarily) vanish since:

$$\begin{aligned}
s_c^*\sigma_{y^*}^2 =& \text{cov}(\beta_{cB}X_{cB}, y_B) = \text{cov}(\beta_{cB}X_{cB}, \sum_c \beta_{cB}X_{cB} + \varepsilon_A) \\
=& \beta_{cB}^2 \text{cov}(X_{cB}, X_{cB}) + \sum_{i \neq c} \beta_{cB}\beta_{iB} \text{cov}(X_{cB}, X_{iB}) + \beta_{cB} \overbrace{\text{cov}(X_{cB}, \varepsilon_B^A)}^{=0 \text{ eq. (16)}} \\
\neq& \beta_{c_B^A}^2 \text{cov}(X_{cB}, X_{cB}) + \sum_{i \neq c} \beta_{c_B^A}\beta_{i_B^A} \text{cov}(X_{cB}, X_B) + \beta_{c_B^A} \overbrace{\text{cov}(X_{cB}, \varepsilon_B^A)}^{=0 \text{ eq. (16)}} \\
=& s_{c_B^A}\sigma_{y_B^A}^2 \Rightarrow \\
\Delta^P \neq& 0
\end{aligned}$$

and

$$\varepsilon_B^A \neq \varepsilon_B \quad \Rightarrow \quad \sigma_{\varepsilon_B^A}^2 \neq \sigma_{\varepsilon_B}^2 \quad \Rightarrow \quad \Delta^\varepsilon \neq 0$$

The same can be applied to the case where initial period data is used in the decomposition.

B Pure and interaction effects

Shorrocks (1982) showed that the contribution of income source k can be decomposed to two parts: A) the pure contribution of income source k on inequality and B) the interaction effect of income source k on inequality. Fields' (2003) decomposition method is very close to the factor source decomposition and thus Shorrocks example is easy to apply in this set up.

Following Shorrocks' example, the contribution of characteristic c can be regarded in two ways: A) the inequality which would be observed if characteristic c was the only characteristics affecting income and B) the amount by which inequality would fall if differences in characteristics c were eliminated. Formally, these can be expressed as:

$$\begin{aligned} C_c^A =& I\left(\beta_c X_c + (\mu - \overline{\beta_c X_c})\right) \\ C_c^B =& I(y) - I\left(y - \beta_c X_c + \overline{\beta_c X_c}\right) \end{aligned}$$

Where I is some inequality measure, y log of income, μ is the mean of log of income and over line presents the mean as well.

Shorrocks (1982) showed that this can consistently be done with the variance of incomes and square of the coefficient of variation as an inequality measure. With the variance of log of incomes C_c^A becomes:

$$C_c^A = \sigma^2\left(\beta_c X_c + (\mu - \overline{\beta_c X_c})\right) = \sigma^2\left(\beta_c X_c\right)$$

And similarly C_c^B is

$$\begin{aligned} C_c^B =& \sigma^2(y) - \sigma^2\left(y - \beta_c X_c + \overline{\beta_c X_c}\right) \\ =& \sigma^2(y) - \sigma^2\left(y - \beta_c X_c\right) = \sigma^2(y) - \sigma^2(y) - \sigma^2(\beta_c X_c) + 2\mathrm{cov}(y, \beta_c X_c) \\ =& -\sigma^2(\beta_c X_c) + 2\mathrm{cov}(\beta_c X_c, \beta_c X_c) + 2\mathrm{cov}(\beta_c X_c, y - \beta_c X_c) \\ =& \sigma^2(\beta_c X_c) + 2\mathrm{cov}(\beta_c X_c, y - \beta_c X_c) \end{aligned}$$

And the contributions derived from the **equation** (4) are then:

$$S_c = s_c \sigma_y^2 = \mathrm{cov}(\beta_c X_c, y) = \frac{1}{2}\left(C_c^A + C_c^B\right)$$

C Price- and quantity effects decomposed to pure- and interaction effects

The income generating functions are defined as:

$$
\begin{aligned}
y_A &= \sum_c \beta_{cA} X_{cA} + \varepsilon_A \\
y_B &= \sum_c \beta_{cB} X_{cB} + \varepsilon_B \\
y^* &= \sum_c \beta_{cB} X_{cA} + \varepsilon_A
\end{aligned}
$$

And the pure- and interaction effects are defined as:

$$
\begin{aligned}
C_c^A =& \sigma^2 \left(\beta_c X_c\right) = \beta_c^2 \sigma^2 \left(X_c\right) \\
C_c^B =& \sigma^2(\beta_c X_c) + 2\mathrm{cov}(\beta_c X_c, y - \beta_c X_c)
\end{aligned}
$$

Given C_c^A and C_c^B, the contributions derived from the **equation** (4) are then:

$$
S_c = s_c \sigma_y^2 = \mathrm{cov}(\beta_c X_c, y) = \frac{1}{2}\left(C_c^A + C_c^B\right)
$$

The share of C_c^A and C_c^B in the total change of inequality is then:

$$
\Delta S_c = \frac{1}{2}\left[\Delta C_c^A + \Delta C_c^B\right]
$$

From the **equation** (6) we know that:

$$
\Delta I = \Delta^Q + \Delta^P + \Delta^\varepsilon = \sum_c \left[\Delta_c^Q + \Delta_c^P + \Delta_c^\varepsilon\right]
$$

Where Δ_c^Q is

$$\begin{aligned}
\Delta_c^Q =& s_{c_B}\sigma_{y_B}^2 - s_c^*\sigma_{y^*}^2 \\
=& \mathrm{cov}(\beta_{cB}X_{cB}, y_B) - \mathrm{cov}(\beta_{cB}X_{cA}, y^*) \\
=& \frac{1}{2}\left(C_{cB}^A + C_{cB}^B\right) - \frac{1}{2}\left(C_{c^*}^A + C_{c^*}^B\right) \\
=& \frac{1}{2}\left(C_{cB}^A - C_{c^*}^A\right) + \frac{1}{2}\left(C_{cB}^B - C_{c^*}^B\right) \\
=& \frac{1}{2}\left[\beta_{cB}^2 \mathrm{Var}(X_{cB}) - \beta_{cB}^2 \mathrm{Var}(X_{cA})\right] \\
& + \frac{1}{2}\left[\beta_{cB}^2 \mathrm{Var}(X_{cB}) - \beta_{cB}^2 \mathrm{Var}(X_{cA})\right. \\
& \left. + 2\mathrm{cov}(\beta_{cB}X_{cB}, y_B - \beta_{cB}X_{cB}) - 2\mathrm{cov}(\beta_{cB}X_{cA}, y^* - \beta_{cB}X_{cA})\right] \\
=& \frac{1}{2}\left[\beta_{cB}^2 \Delta\mathrm{Var}(X_c)\right] + \frac{1}{2}\left[\beta_{cB}^2 \Delta\mathrm{Var}(X_c) + 2\sum_{i\neq c}\beta_{cB}\beta_{iB}[\Delta\mathrm{cov}(X_c, X_i)]\right]
\end{aligned}$$

And similarly Δ_c^P is

$$\Delta_c^P = \frac{1}{2}\left[\mathrm{Var}(X_c)\Delta\beta_c^2\right] + \frac{1}{2}\left[\mathrm{Var}(X_c)\Delta\beta_c^2 + 2\sum_{i\neq c}\mathrm{cov}(X_{cA}, X_{iA})\Delta\beta_c\beta_i\right]$$

D Data and microsimulation

Data used in this analysis is the service data of income distribution collected by Statistics Finland, which is basis of EU-SILC data for Finland. Years covered in the analyses are 1993, 1996, 1999, 2002, 2005, 2008, 2011 and 2014.

Simulations are conducted using Statistics Finland's microsimulation model SISU. The majority of the Finnish tax-benefit-system is encoded in SISU model from the year 1993 onwards. The SISU model, however, is not compatible with the data before the year 2011. Therefore the simulations are carried out by using data from the year 2011. For year 2011 actual data from the year 2011 is used.

The SISU model has 12 different sub-models: Health care insurances; Unemployment benefits; Home care allowance; National pensions; Student allowance; Taxation; Estate taxation; Child allowance; Pensioner's housing benefit; Housing benefit; Day care fees and Social assistance. All of these sub-models are used in the simulations except the estate taxation.

All the tax-benefit changes, however, cannot be simulated. Earnings-related pensions are not simulated at all since it would require information of past earnings from the entire work history. In addition, some changes, for instance, for eligibility or changes in the duration of benefit payments cannot be simulated. For example, the duration of unemployment insurance was decreased from 500 days to 400 days, which effect cannot be simulated with SISU model.

To make different years' tax-benefit-system more comparative to other years data, the monetary parameters are in the main specification indexed by the consumer price index. In the robustness checks also the market income index and nominal values (i.e. no indexation) are used.

Disposable incomes are at the household level and thus those are scaled by the modified OECD-equivalence scale. The analyses are conducted by using only the households' heads, which are given the information about the characteristics of any spouses. Weight used is the household level weight multiplied with the number of members in the household. This way the analyses are done at the individual level.

E Definition of population subgroups

Full details of the characteristics are presented here. *Education:* Elementary school or no education or education unknown; secondary school or lower/lowest tertiary level; upper tertiary level or higher. *Region:* Uusimaa; Varsinais-Suomi; Satakunta; Kanta-Häme; Pirkanmaa; Päijät-Häme; Kymenlaakso; Etelä-Karjala; Etelä-Savo; Pohjois-Savo; Pohjois-Karjala; Keski-Suomi; Etelä-Pohjanmaa; Pohjanmaa; Keski-Pohjanmaa; Pohjois-Pohjanmaa; Kainuu; Lappi; Ahvenanmaa. *Household type:* 1 adult, no children; 2 adults, no children; 3 or more adults, no children; single parents, youngest child under 7 years old; two adults, youngest children under 7 years old; 3 or more adults, youngest children under 7 years old; single parents, youngest child over 6 years old; two adults, youngest child over 6 years old; 3 or more adults, youngest child over 6 years old; 1 adult, household head over 64 years old; other families with household head over 64 years old. *Age:* under 25 years old; 25–34; 35–44; 45–54; 55–64; 65–74; over 74 years old. *Employment status:* employed; unemployed and others; self-employed; pensioner; student; schoolchild; unknown.

F Annual decompositions

In this section are presented the annual decompositions of the results.

Table F.1: Shares of characteristics in income inequality (%)

			Household	Age		Employment status		Education	
Years	Residual	Region	type	Male	Female	Male	Female	Male	Female
1993	61.0	3.2	4.4	5.8	3.8	5.3	7.1	5.5	4.0
1996	58.5	1.5	3.8	6.5	3.7	7.5	8.8	6.1	3.6
1999	57	2.4	2.3	7.9	3.4	6.8	7.7	7.3	5.2
2002	56.7	2.1	3.3	7.1	4.2	8.3	7.6	7.3	3.5
2005	56.9	1.8	3.0	5.6	3.1	7.5	7.8	8.6	5.7
2008	55.6	2.4	3.9	6.9	3.0	8.6	8.2	6.2	5.2
2011	54.8	2.0	2.7	7.9	3.3	8.4	8.3	7.1	5.5
2014	51.1	2.2	1.9	9.3	4.4	9.4	7.2	7.9	6.6

Note: Author's own calculations based on the service data of income distribution.

Table F.2: Shares of characteristics in income inequality (Simulated data, %)

			Household	Age		Employment status		Education	
Years	Residual	Region	type	Male	Female	Male	Female	Male	Female
1993	57.7	1.9	2.3	8.1	3.1	7.7	7.8	6.4	5.0
1996	57.3	1.8	2.2	7.7	2.9	8.0	8.2	6.7	5.2
1999	55.0	1.8	2.2	7.8	2.9	9.0	9.0	6.9	5.4
2002	54.5	1.8	2.4	7.8	3	8.9	9.1	6.9	5.4
2005	55.1	1.9	1.9	7.7	2.8	9.2	8.4	7.6	5.4
2008	54.8	1.9	2.1	7.7	2.9	9.2	8.6	7.5	5.4
2011	54.8	2.0	2.7	7.9	3.3	8.4	8.3	7.1	5.5
2014	55.0	2.0	2.4	7.9	3.1	8.6	8.3	7.4	5.3

Note: All years besides the year 2011 are simulated using data from year 2011. For the year 2011 actual data is used. Author's own calculations based on the service data of income distribution.

Table F.3: Absolute contribution of characteristics in income inequality (1000 $\times$ variance of logs)

				Household	Age		Employment status		Education	
Years	Total	Residual	Region	type	Male	Female	Male	Female	Male	Female
1993	147.2	89.8	4.8	6.5	8.5	5.5	7.9	10.4	8.1	5.8
1996	159.4	93.3	2.4	6.1	10.4	5.9	12.0	14.0	9.6	5.8
1999	197.1	112.3	4.7	4.5	15.6	6.6	13.5	15.2	14.4	10.2
2002	207.0	117.4	4.3	6.8	14.6	8.6	17.3	15.7	15.1	7.2
2005	228.3	129.9	4.1	6.8	12.7	7.1	17.2	17.8	19.6	13.0
2008	229.5	127.7	5.6	8.9	15.9	6.8	19.7	18.9	14.3	11.9
2011	234.6	128.5	4.7	6.4	18.6	7.7	19.8	19.4	16.6	12.9
2014	228.8	116.9	5.0	4.3	21.3	10.1	21.5	16.5	18.1	15

Note: Author's own calculations based on the service data of income distribution.

Table F.4: Absolute contribution of characteristics in income inequality (Simulated data, 1000 × variance of logs)

				Household	Age		Employment status		Education	
Years	Total	Residual	Region	type	Male	Female	Male	Female	Male	Female
1993	193.9	111.9	3.6	4.4	15.7	6.0	15.0	15.2	12.4	9.6
1996	193.4	110.9	3.4	4.2	14.9	5.7	15.5	15.9	12.9	10.0
1999	201.6	110.9	3.6	4.4	15.7	5.9	18.2	18.0	13.9	10.9
2002	203.9	111.2	3.7	5.0	16.0	6.1	18.2	18.6	14.1	11.1
2005	226.6	124.8	4.3	4.3	17.3	6.4	20.9	19.0	17.3	12.2
2008	229.0	125.4	4.4	4.7	17.6	6.6	21.1	19.6	17.3	12.3
2011	234.6	128.5	4.7	6.4	18.6	7.7	19.8	19.4	16.6	12.9
2014	216.7	119.3	4.2	5.2	17.1	6.7	18.7	17.9	16.1	11.5

Note: All years besides the year 2011 are simulated using data from year 2011. For the year 2011 actual data is used. Author's own calculations based on the service data of income distribution.

Table F.5: Change in the absolute contribution of characteristics in income inequality (1000 × variance of logs)

				Household	Age		Employment status		Education	
Years	Total	Residual	Region	type	Male	Female	Male	Female	Male	Female
1993 to 1996	12.2	3.5	-2.4	-0.4	1.9	0.4	4.1	3.6	1.5	0.0
1996 to 1999	37.7	19	2.3	-1.6	5.2	0.7	1.5	1.2	4.8	4.4
1999 to 2002	9.9	5.1	-0.4	2.3	-1.0	2.0	3.8	0.5	0.7	-3
2002 to 2005	21.3	12.5	-0.2	0.0	-1.9	-1.5	-0.1	2.1	4.5	5.8
2005 to 2008	1.2	-2.2	1.5	2.1	3.2	-0.3	2.5	1.1	-5.3	-1.1
2008 to 2011	5.1	0.8	-0.9	-2.5	2.7	0.9	0.1	0.5	2.3	1
2011 to 2014	-5.8	-11.6	0.3	-2.1	2.7	2.4	1.7	-2.9	1.5	2.1

Note: Author's own calculations based on the service data of income distribution.

Table F.6: Absolute policy effect of characteristics in income inequality (1000 × variance of logs)

				Household	Age		Employment status		Education	
Years	Total	Residual	Region	type	Male	Female	Male	Female	Male	Female
1993 to 1996	-0.5	-1.0	-0.2	-0.2	-0.8	-0.3	0.5	0.7	0.5	0.4
1996 to 1999	8.2	0.0	0.2	0.2	0.8	0.2	2.7	2.1	1.0	0.9
1999 to 2002	2.3	0.3	0.1	0.6	0.3	0.2	0.0	0.6	0.2	0.2
2002 to 2005	22.7	13.6	0.6	-0.7	1.3	0.3	2.7	0.4	3.2	1.1
2005 to 2008	2.4	0.6	0.1	0.4	0.3	0.2	0.2	0.6	0.0	0.1
2008 to 2011	5.6	3.1	0.3	1.7	1.0	1.1	-1.3	-0.2	-0.7	0.6
2011 to 2014	-17.9	-9.2	-0.5	-1.2	-1.5	-1.0	-1.1	-1.5	-0.5	-1.4

Note: All years besides the year 2011 are simulated using data from year 2011. For the year 2011 actual data is used. Author's own calculations based on the service data of income distribution.

Table F.7: Changes in income inequality, price-, quantity and residual effects (1000x variance of log of incomes)

Years	Total Change	Price Effect	Quantity Effect	Residual Effect	Region		Household type		Age				Employment status				Education			
									Male		Female		Male		Female		Male		Female	
					P	Q	P	Q	P	Q	P	Q	P	Q	P	Q	P	Q	P	Q
1993 to 1996	12.2	3.5	7.4	1.3	-2.1	-0.3	-0.3	-0.2	1.4	0.5	0.4	0.1	4.9	-0.8	3.3	0.4	0.5	1.1	-0.7	0.7
1996 to 1999	37.7	19.0	15.7	3.0	2.5	-0.1	-1.8	0.2	3.9	1.4	0.3	0.5	2.5	-1.1	-0.2	1.4	5.0	-0.3	3.6	1.0
1999 to 2002	9.9	5.1	-1.7	6.6	-1.0	0.5	2.0	0.3	-2.2	1.3	0.4	1.6	2.8	1.0	0.4	0.2	-0.5	1.1	-3.6	0.6
2002 to 2005	21.2	12.6	9.5	-0.9	0.1	-0.4	0.5	-0.4	-2.1	0.2	-1.9	0.3	2.1	-2.2	3.7	-1.6	2.7	1.9	4.5	1.4
2005 to 2008	1.2	-2.2	-1.0	4.5	1.4	0.1	1.1	1.0	2.8	0.4	-0.7	0.5	2.5	0.0	-0.3	1.3	-5.7	0.4	-2.1	0.9
2008 to 2011	5.1	0.8	1.6	2.7	-1.0	0.1	-2.2	-0.3	2.9	-0.2	0.7	0.3	-2.0	2.0	0.5	0.1	2.0	0.4	0.8	0.3
2011 to 2014	-5.8	-11.6	-2.1	8.0	0.3	0.0	-1.9	-0.3	1.0	1.8	1.2	1.3	-0.1	1.9	-3.0	0.1	0.2	1.3	0.2	2.0

Note: Price- and quantity effects are formed as an average of **equations** (6) and (9). Author's own calculations based on the service data of income distribution.

Table F.8: Price and residual effects decomposed to policy and other effects (1000x variance of log of incomes)

Years	Price Effect		Residual Effect		Region		Household type		Age				Employment status				Education			
									Male		Female		Male		Female		Male		Female	
	Δ_O^P	Δ_P^P	Δ_O^ε	Δ_P^ε	Δ_O^P	Δ_P^P	Δ_O^P	Δ_P^P	Δ_O^P	Δ_P^P	Δ_O^P	Δ_P^P	Δ_O^P	Δ_P^P	Δ_O^P	Δ_P^P	Δ_O^P	Δ_P^P	Δ_O^P	Δ_P^P
1993 to 1996	7.2	0.6	4.5	-1.0	-2.0	-0.2	-0.1	-0.2	2.1	-0.8	0.7	-0.3	4.8	0.5	2.5	0.8	0.0	0.4	-1.0	0.4
1996 to 1999	8.1	8.2	19.0	0.0	2.4	0.2	-1.7	0.2	2.9	0.8	0.2	0.2	0.0	2.7	-2.2	2.1	4.0	1.0	2.5	0.9
1999 to 2002	-3.4	2.0	4.7	0.3	-1.0	0.1	1.3	0.5	-2.5	0.3	0.2	0.3	2.9	0.0	0.0	0.5	-0.6	0.2	-3.7	0.2
2002 to 2005	0.5	9.0	-1.1	13.6	-0.5	0.6	1.0	-0.6	-3.4	1.3	-1.9	0.2	-0.6	2.7	3.4	0.4	-0.7	3.2	3.2	1.1
2005 to 2008	-2.6	1.8	-2.8	0.6	1.3	0.1	0.8	0.4	2.3	0.3	-1.0	0.2	2.5	0.2	-0.8	0.6	-5.6	0.0	-2.1	0.1
2008 to 2011	-1.0	2.5	-2.3	3.1	-1.3	0.4	-3.9	1.7	1.7	1.0	-0.5	1.1	-0.4	-1.3	0.6	-0.2	2.6	-0.7	0.1	0.6
2011 to 2014	6.5	-8.6	-2.3	-9.3	0.7	-0.5	-0.7	-1.1	2.4	-1.5	2.2	-1.0	1.0	-1.1	-1.4	-1.5	0.7	-0.5	1.6	-1.4

Note: Simulations done using the year 2011 data. Author's own calculations based on the service data of income distribution.

G Robustness checks

In this section is presented the robustness checks of the results obtained earlier. In the first table is shown the change in absolute contribution of characteristics in income inequality measured with the 100 $\times$ Gini coefficient. Then is shown the policy effects by using 100 $\times$ Gini coefficient as an inequality measure. In the following tables are shown the price- and quantity effects formed using **equations** (6) and (9). Then the final tables presents the bootstrapped 95% confidence intervals of the results. Confidence intervals are formed using 1 000 replications.

Table G.1: Change in absolute contribution of characteristics in income inequality (100 $\times$ Gini coefficient)

				Household	Age		Employment status		Education	
Years	Total	Residual	Region	type	Male	Female	Male	Female	Male	Female
1993 to 1996	1.3	0.26	-0.34	-0.07	0.23	0.04	0.55	0.46	0.19	-0.02
1996 to 1999	3.71	1.77	0.29	-0.25	0.6	0.05	0.1	0.05	0.55	0.54
1999 to 2002	-0.15	-0.16	-0.08	0.25	-0.22	0.2	0.37	-0.04	-0.02	-0.44
2002 to 2005	1.05	0.66	-0.06	-0.04	-0.32	-0.24	-0.13	0.14	0.42	0.63
2005 to 2008	0.25	-0.2	0.17	0.24	0.37	-0.03	0.3	0.13	-0.61	-0.13
2008 to 2011	-0.29	-0.39	-0.11	-0.31	0.24	0.08	-0.07	-0.01	0.21	0.07
2011 to 2014	-0.32	-1.14	0.04	-0.22	0.34	0.28	0.23	-0.3	0.19	0.26
1993 - 2005	5.91	2.53	-0.19	-0.11	0.29	0.05	0.89	0.61	1.14	0.71
2005 - 2014	-0.36	-1.73	0.1	-0.29	0.95	0.33	0.46	-0.18	-0.21	0.2
1993 - 2014	5.55	0.8	-0.09	-0.4	1.24	0.38	1.35	0.43	0.93	0.91

Note: Author's own calculations based on the service data of income distribution.

Table G.2: Policy effect of characteristics in income inequality (100 $\times$ Gini coefficient)

				Household	Age		Employment status		Education	
Years	Total	Residual	Region	type	Male	Female	Male	Female	Male	Female
1993 to 1996	-0.14	-0.18	-0.02	-0.03	-0.1	-0.04	0.06	0.09	0.05	0.04
1996 to 1999	0.53	-0.27	0.01	0.02	0.06	0.01	0.29	0.22	0.09	0.09
1999 to 2002	0.06	-0.09	0.01	0.06	0.01	0.03	-0.02	0.04	0	0.01
2002 to 2005	2.1	1.3	0.06	-0.09	0.11	0.01	0.27	0	0.34	0.1
2005 to 2008	0.03	-0.07	0	0.04	0.01	0.02	0	0.05	-0.02	0
2008 to 2011	-0.3	-0.16	0.02	0.17	0.04	0.1	-0.24	-0.1	-0.15	0.01
2011 to 2014	-0.39	-0.15	-0.02	-0.09	-0.04	-0.06	0.02	-0.03	0.06	-0.07
1993 - 2005	2.54	0.76	0.06	-0.04	0.09	0	0.6	0.35	0.49	0.24
2005 - 2014	-0.66	-0.38	0	0.12	0.01	0.06	-0.22	-0.09	-0.11	-0.05
1993 - 2014	1.88	0.38	0.06	0.08	0.1	0.06	0.38	0.26	0.38	0.18

Note: All years besides the year 2011 are simulated using data from year 2011. For the year 2011 actual data is used. Author's own calculations based on the service data of income distribution.

Table G.3: Changes in income inequality, price-, quantity and residual effects (1000x variance of log of incomes)

Years	Total Change	Price Effect	Quantity Effect	Residual Effect	Region		Household type		Age				Employment status				Education			
									Male		Female		Male		Female		Male		Female	
					P	Q	P	Q	P	Q	P	Q	P	Q	P	Q	P	Q	P	Q
1993 to 1996	12.2	3.5	7.8	0.8	-2.1	-0.3	-0.3	-0.1	1.4	0.5	0.4	0	5.3	-1.2	3.3	0.3	0.4	1.1	-0.6	0.6
1996 to 1999	37.7	19	16.3	2.4	2.6	-0.2	-1.5	-0.1	3.7	1.5	0.4	0.4	2.7	-1.3	-0.1	1.3	5	-0.3	3.4	1.1
1999 to 2002	9.9	5.1	-1.4	6.3	-0.9	0.4	1.8	0.4	-2.2	1.3	0.5	1.5	2.9	0.9	0.5	0.1	-0.5	1.1	-3.5	0.5
2002 to 2005	21.2	12.6	9.5	-0.9	0.1	-0.3	0.4	-0.3	-2	0.1	-1.7	0.2	2.1	-2.2	3.9	-1.8	2.6	1.9	4.3	1.5
2005 to 2008	1.2	-2.2	-0.8	4.3	1.4	0.1	1.2	0.9	2.6	0.6	-0.7	0.5	2.7	-0.2	-0.3	1.3	-5.6	0.3	-2.0	0.8
2008 to 2011	5.1	0.8	1.5	2.8	-1.0	0.1	-2.3	-0.2	2.7	0.0	0.7	0.2	-1.7	1.7	0.4	0.1	1.9	0.5	0.7	0.3
2011 to 2014	-5.8	-11.6	-2.1	8.0	0.3	0.0	-1.8	-0.3	0.9	1.8	1.2	1.2	-0.1	1.9	-2.9	0.0	0.2	1.3	0.2	1.9

Note: Price- and quantity effects are formed according to **equation** (6). Author's own calculations based on the service data of income distribution.

Table G.4: Changes in income inequality, price-, quantity and residual effects (1000x variance of log of incomes)

Years	Total Change	Price Effect	Quantity Effect	Residual Effect	Region		Household type		Age				Employment status				Education			
									Male		Female		Male		Female		Male		Female	
					P	Q	P	Q	P	Q	P	Q	P	Q	P	Q	P	Q	P	Q
1993 to 1996	12.2	3.5	6.9	1.8	-2.1	-0.2	-0.2	-0.2	1.4	0.5	0.3	0.1	4.5	-0.3	3.2	0.4	0.6	1.0	-0.8	0.7
1996 to 1999	37.7	19	15.1	3.6	2.3	0.0	-2.0	0.5	4.0	1.2	0.1	0.6	2.3	-0.9	-0.3	1.5	5.0	-0.2	3.7	0.8
1999 to 2002	9.9	5.1	-2.0	6.9	-1.0	0.6	2.1	0.2	-2.1	1.2	0.2	1.7	2.7	1.1	0.3	0.3	-0.5	1.1	-3.7	0.7
2002 to 2005	21.2	12.6	9.5	-0.8	0.1	-0.4	0.5	-0.4	-2.1	0.2	-2	0.4	2.1	-2.2	3.5	-1.4	2.7	1.8	4.6	1.2
2005 to 2008	1.2	-2.2	-1.1	4.6	1.4	0.1	1.0	1.0	3.0	0.1	-0.7	0.4	2.3	0.2	-0.2	1.2	-5.8	0.5	-2.2	1.0
2008 to 2011	5.1	0.8	1.7	2.5	-0.9	0.1	-2.1	-0.3	3.0	-0.3	0.6	0.3	-2.2	2.3	0.6	0.0	2.0	0.3	0.8	0.2
2011 to 2014	-5.8	-11.6	-2.1	7.9	0.3	0.0	-1.9	-0.2	1.0	1.7	1.1	1.3	0.0	1.8	-3.0	0.1	0.2	1.3	0.2	2.0

Note: Price- and quantity effects are formed according to **equation** (9). Author's own calculations based on the service data of income distribution.

Table G.5: Shares of characteristics, 95% confidence intervals for actual data

Years	Residual	Region	Household type	Age		Employment status		Education	
				Male	Female	Male	Female	Male	Female
1993	(58.7%–63.3%)	(2.5%–3.9%)	(3.4%–5.4%)	(4.6%–6.9%)	(2.9%–4.7%)	(4.4%–6.3%)	(6%–8.1%)	(4.4%–6.6%)	(3.2%–4.7%)
1996	(56.4%–60.6%)	(1%–2%)	(2.7%–4.9%)	(5.4%–7.7%)	(2.8%–4.6%)	(6.3%–8.8%)	(7.6%–10%)	(4.9%–7.2%)	(2.8%–4.4%)
1999	(54.8%–59.2%)	(1.8%–3%)	(1.2%–3.3%)	(6.7%–9.1%)	(2.5%–4.3%)	(5.6%–8.1%)	(6.4%–9%)	(6.2%–8.4%)	(4.2%–6.1%)
2002	(54.1%–59.3%)	(1.4%–2.8%)	(2.2%–4.3%)	(5.9%–8.3%)	(3.2%–5.1%)	(7.1%–9.6%)	(6.4%–8.8%)	(6.2%–8.4%)	(2.5%–4.5%)
2005	(54%–59.8%)	(1.2%–2.3%)	(2%–4%)	(4.3%–6.8%)	(2.1%–4.1%)	(6.3%–8.8%)	(6.6%–9%)	(7.4%–9.8%)	(4.7%–6.7%)
2008	(53.3%–57.9%)	(1.8%–3%)	(2.7%–5%)	(5.7%–8.2%)	(2%–3.9%)	(7.3%–9.9%)	(7%–9.4%)	(5.1%–7.4%)	(4.2%–6.2%)
2011	(52.1%–57.5%)	(1.5%–2.6%)	(1.7%–3.7%)	(6.6%–9.3%)	(2.3%–4.2%)	(6.9%–10%)	(7.1%–9.5%)	(6%–8.2%)	(4.4%–6.6%)
2014	(49.2%–53%)	(1.6%–2.8%)	(0.8%–3%)	(7.8%–10.8%)	(3.3%–5.5%)	(8.1%–10.7%)	(6%–8.4%)	(6.7%–9.1%)	(5.6%–7.6%)

Note: Author's own calculations based on the service data of income distribution.

Table G.6: Shares of characteristics, 95% confidence intervals for simulated data

Years	Residual	Region	Household type	Age		Employment status		Education	
				Male	Female	Male	Female	Male	Female
1993	(54.5%–60.9%)	(1.3%–2.4%)	(1.3%–3.3%)	(6.6%–9.6%)	(2.1%–4%)	(6.3%–9.1%)	(6.5%–9.1%)	(5.2%–7.6%)	(3.9%–6.1%)
1996	(53.8%–60.9%)	(1.3%–2.3%)	(1.1%–3.2%)	(6.2%–9.2%)	(2%–3.9%)	(6.6%–9.4%)	(6.9%–9.6%)	(5.4%–7.9%)	(4.1%–6.3%)
1999	(52%–58%)	(1.3%–2.3%)	(1.2%–3.2%)	(6.4%–9.2%)	(2%–3.9%)	(7.6%–10.5%)	(7.6%–10.3%)	(5.7%–8.1%)	(4.3%–6.5%)
2002	(51.5%–57.6%)	(1.3%–2.3%)	(1.4%–3.5%)	(6.4%–9.3%)	(2.1%–4%)	(7.5%–10.3%)	(7.8%–10.5%)	(5.7%–8.1%)	(4.3%–6.6%)
2005	(52.4%–57.8%)	(1.4%–2.4%)	(0.9%–2.9%)	(6.3%–9%)	(1.9%–3.8%)	(7.8%–10.6%)	(7.1%–9.7%)	(6.4%–8.9%)	(4.3%–6.5%)
2008	(52%–57.5%)	(1.4%–2.4%)	(1%–3.1%)	(6.3%–9%)	(1.9%–3.8%)	(7.8%–10.6%)	(7.2%–9.9%)	(6.3%–8.8%)	(4.3%–6.5%)
2011	(52.1%–57.5%)	(1.5%–2.6%)	(1.7%–3.7%)	(6.6%–9.3%)	(2.3%–4.2%)	(6.9%–10%)	(7.1%–9.5%)	(6%–8.2%)	(4.4%–6.6%)
2014	(52.2%–57.9%)	(1.4%–2.5%)	(1.4%–3.5%)	(6.5%–9.3%)	(2.1%–4%)	(7.2%–10%)	(6.9%–9.6%)	(6.2%–8.6%)	(4.2%–6.4%)

Note: All years besides the year 2011 are simulated using data from year 2011. For the year 2011 actual data is used. Author's own calculations based on the service data of income distribution.

Table G.7: Absolute contribution of characteristics in income inequality, 95% confidence intervals for actual data (1000 × variance of logs)

				Household	Age		Employment status		Education	
Years	Total	Residual	Region	type	Male	Female	Male	Female	Male	Female
1993	(137.7–156.8)	(81.9–97.6)	(3.7–5.8)	(4.9–8)	(6.8–10.2)	(4.2–6.9)	(6.5–9.2)	(8.8–12)	(6.6–9.7)	(4.6–7)
1996	(151.1–167.7)	(87–99.6)	(1.5–3.2)	(4.3–7.8)	(8.5–12.3)	(4.4–7.4)	(10–14.1)	(11.9–16)	(7.9–11.4)	(4.4–7.1)
1999	(185.4–208.8)	(103.7–120.9)	(3.5–6)	(2.5–6.5)	(13–18.1)	(4.8–8.5)	(11–16)	(12.8–17.6)	(12–16.9)	(8.3–12.2)
2002	(195.4–218.7)	(107–127.7)	(3–5.7)	(4.7–8.8)	(12.1–17.2)	(6.5–10.7)	(14.7–19.8)	(13.3–18.1)	(12.6–17.5)	(5.3–9.1)
2005	(209.8–246.8)	(113.5–146.3)	(2.9–5.3)	(4.5–9.1)	(10.3–15.2)	(4.8–9.3)	(14.3–20.1)	(15.2–20.4)	(16.4–22.7)	(10.6–15.4)
2008	(217.6–241.5)	(117.6–137.7)	(4.1–7)	(6.4–11.3)	(12.9–18.9)	(4.7–9)	(16.8–22.7)	(16–21.8)	(11.6–16.9)	(9.5–14.2)
2011	(214.1–255.1)	(113–144.1)	(3.4–6)	(4–8.8)	(15.4–21.8)	(5.2–10.2)	(16.6–22.9)	(16–22.8)	(14–19.2)	(10.6–15.1)
2014	(218.2–239.5)	(109.8–124.1)	(3.7–6.4)	(1.9–6.8)	(17.6–25)	(7.5–12.7)	(18.5–24.6)	(13.9–19.1)	(15.4–20.8)	(12.6–17.4)

Note: Author's own calculations based on the service data of income distribution.

Table G.8: Absolute contribution of characteristics in income inequality, 95% confidence intervals for simulated data (1000 × variance of logs)

				Household	Age		Employment status		Education	
Years	Total	Residual	Region	type	Male	Female	Male	Female	Male	Female
1993	(179.5–208.2)	(98.6–125.3)	(2.5–4.7)	(2.4–6.5)	(13.1–18.2)	(4.1–7.8)	(12.5–17.4)	(12.5–17.9)	(10.1–14.8)	(7.7–11.6)
1996	(177.4–209.5)	(95.8–126)	(2.4–4.5)	(2.2–6.3)	(12.4–17.4)	(3.8–7.6)	(13.1–17.9)	(13.2–18.7)	(10.5–15.2)	(8–11.9)
1999	(188.7–214.5)	(99.1–122.7)	(2.5–4.7)	(2.3–6.6)	(13.1–18.3)	(3.9–7.8)	(15.5–20.9)	(15–21)	(11.4–16.4)	(8.9–13)
2002	(190.9–217)	(99.2–123.2)	(2.6–4.8)	(2.8–7.1)	(13.4–18.6)	(4.2–8.1)	(15.5–20.9)	(15.5–21.7)	(11.5–16.6)	(9–13.2)
2005	(213.4–239.7)	(113.2–136.5)	(3.1–5.5)	(2–6.6)	(14.5–20.2)	(4.2–8.5)	(17.8–24)	(15.7–22.3)	(14.3–20.3)	(9.8–14.6)
2008	(215.7–242.2)	(113.7–137.1)	(3.1–5.6)	(2.4–7)	(14.7–20.5)	(4.4–8.7)	(18–24.2)	(16.2–22.9)	(14.3–20.3)	(9.9–14.7)
2011	(214.1–255.1)	(113–144.1)	(3.4–6)	(4–8.8)	(15.4–21.8)	(5.2–10.2)	(16.6–22.9)	(16–22.8)	(14–19.2)	(10.6–15.1)
2014	(203.7–229.7)	(107.6–130.9)	(3–5.5)	(3–7.5)	(14.3–19.9)	(4.7–8.7)	(15.8–21.5)	(14.7–21.1)	(13.3–18.9)	(9.3–13.8)

Note: All years besides the year 2011 are simulated using data from year 2011. For the year 2011 actual data is used. Author's own calculations based on the service data of income distribution.

H Additional tables

Table H.1: Macroeconomic developement in Finland

Year	Real GDP per capita	Employment rate Total	Employment rate Men	Employment rate Women	Unemployment rate Total	Unemployment rate Men	Unemployment rate Women
1993	22413	60.6	61.5	59.6	16.5	18.2	14.5
1994	23201	59.9	61.1	58.8	16.7	18.3	14.9
1995	24087	61.1	63.1	59.1	15.5	15.8	15.1
1996	24889	61.9	64.2	59.5	14.6	14.4	14.9
1997	26387	62.9	65.4	60.3	12.7	12.4	13.1
1998	27753	64.1	66.9	61.3	11.4	10.9	12.0
1999	28901	66.0	68.4	63.5	10.3	9.8	10.7
2000	30506	66.9	69.4	64.3	9.8	9.1	10.6
2001	31231	67.7	70.0	65.4	9.2	8.7	9.7
2002	31688	67.7	69.2	66.2	9.2	9.2	9.1
2003	32246	67.3	68.9	65.7	9.1	9.3	8.9
2004	33435	67.2	68.9	65.5	8.9	8.8	9.0
2005	34248	68.0	69.5	66.5	8.5	8.3	8.7
2006	35490	68.9	70.5	67.3	7.8	7.5	8.1
2007	37213	69.9	71.3	68.5	6.9	6.6	7.3
2008	37330	70.6	72.3	68.9	6.4	6.2	6.7
2009	34152	68.3	68.8	67.9	8.4	9.0	7.6
2010	35079	67.8	68.7	66.9	8.5	9.3	7.7
2011	35806	68.6	69.8	67.4	7.9	8.6	7.2
2012	35138	69.0	69.8	68.1	7.8	8.5	7.1
2013	34662	68.5	69.2	67.8	8.3	9.0	7.6
2014	34386	68.3	68.7	67.9	8.8	9.6	8.1

Note: Employment and unemployment rates are calculated among 15 to 64 years old.
Source: Statistics Finland (2021a,c)

The **Aboa Centre for Economics (ACE)** is a joint initiative of the economics departments of the Turku School of Economics at the University of Turku and the School of Business and Economics at Åbo Akademi University. ACE was founded in 1998. The aim of the Centre is to coordinate research and education related to economics.

Contact information: Aboa Centre for Economics, Department of Economics, Rehtorinpellonkatu 3, FI-20500 Turku, Finland.

www.ace-economics.fi

ISSN 1796-3133