Who benefits from economic growth?

An analysis of income micro data over a growth period in Norway 1993-2015

Elisabeth Helgesen

Thesis for the Master's Degree Programme in Economics

Department of Economics

University of Oslo

May 2018

Abstract

In this thesis I deal with the relation between economic growth and income inequality to study how a long period with substantial economic growth affect income inequality. I investigate what effect economic growth has had on income inequality through the use of income distributions, and explore how household characteristics can explain the differences uncovered.

To perform these analyses I use two different approaches. Firstly, the growth incidence curve and the individual growth incidence curve are calculated to explore the direct effect of economic growth on different parts of the income distribution. Secondly, household characteristics are used to further explore the underlying factors shaping the income distribution. This is accomplished by doing a RIF regression combined with the logic of an Oaxaca-Blinder decomposition to decompose the effects into the composition and the structural effect.

The results point the economic growth as having a negative effect on income inequality which has increased. Even though the growth is relatively fairly distributed with an almost equal growth rate for large parts of the distribution, the top 10 percent of the income distribution experienced a higher growth throughout the period. It is this larger growth for the top percentile that leads to the conclusion of increased inequality. Despite of the increased inequality, there is little evidence in the results indicating prolonged poverty. Individuals whom are poor in one year are not necessarily poor in the next year, pointing to a high social mobility.

Preface

This thesis is written for the completion of the Master in Economic at the University of Oslo. Coming to the end of my time as a student there are several people that deserve gratitude for their help and support.

First I wish to thank my supervisor Thor Olav Thoresen for all his time and effort. For his excellent supervision and help with everything from choice of topic to his invaluable feedback and pointers on the language.

I would also like to thank Statistic Norway for the dataset used in this thesis and the great work space. In addition I would like to thank Oslo Fiscal Studies for granting me their scholarship.

Thank to all my fellow master students at Statistics Norway for help with the thesis, regular long lunch breaks, discussions about the little things and of course the weekly Wednesday cake.

I need to thank my wonderful partner Stian for all of his support and help. Thanks for calming me when I was stressing out and motivating me when I struggled with motivation, and thanks for proof reading the entire thesis.

Lastly I want to note that any errors or inaccuracies in this paper are mine and mine alone.

Contents

1. Introduction

Redistribution, inequality and income growth are core topics within economics. Several studies discuss the relationships, however no clear conclusions exists on the interactions between them. Some studies, such as Ostry et al. (2014), find that lower income inequality can be a driving force for growth, while others find the complete opposite. Other studies have looked into the effect of redistribution on growth, with no clear conclusion on how it works. In this thesis I look at the reverse relationship, how economic growth affects income distribution and inequality. The impact of growth on the income distribution can be thought of as a function of two factors: the magnitude of the growth measured by the change in mean income, and how these income gains are distributed across the population. In this thesis I use a large period of growth in Norway to look at how economic growth affects the income distribution and income inequality.

One of the theories established for the relationship between economic growth and income inequality is the Kuznets curve, which predicts that as the economy develops, it will first experience an increased income inequality, followed by a turning point after which it experiences a decrease in income inequality (Kuznets, 1955). This would imply that an economy experiencing growth past the turning point would see a lower inequality. Evidence from an investigation into the top income earners in the United States showed that there was a decline in the share of income held by the top percentile in a period after the second world war, but this decline did not last. After a brief period of time the decline was followed by an increase in the income shares of the top percentile (Piketty and Saez, 2003). This means that the gap between the richest and the middle class had increased recently, which does not comply with Kuznets curve. Others, like Dollar and Kraay (2002), found no overall effect of growth on income inequality.

The fact that no clear conclusion can be made on the effect of growth on income distribution and inequality is an important motivation for this thesis. If the Kuznets curve is confirmed for the Norwegian economy implying that economic growth leads to more equality, that could be an argument for economic growth. Next, one must decide on which method to use when studying the effect of growth on inequality and income distribution. Studying averages tell us nothing about how the rise in average income is distributed. It could be unequally distributed between different social groups, leaving winners and losers. I want to identify if there are any such winners and losers. In this thesis focus on the effect on income growth on the entire income distribution, using household data for Norway.

Norway has been chosen due to the steady economic growth the country experienced between 1993 and 2015, where mainland GDP per capita in nominal terms almost tripled from approximately NOK 200, 000 to approximately NOK 600, 000. The recent financial crisis occurred within this time span, but did hardly affect the Norwegian economy, whereas other countries struggled with low or even negative growth and high unemployment.

In this thesis I will investigate whether the steady income growth in Norway has benefited everyone equally, or if some parts of the population has gained more than others. To evaluate Norway's growth I use Norwegian micro register data from 1993 to 2015. In addition to information from the tax return registry the data contain information from other registries like the education registry and social insurance data

Using this registry containes the information on all individual income that is suitable for thesis because it provides longitudinal and cross-sectional information on household income and household characteristics.

To study the effect of growth on the income distribution I will use two different approaches. Firstly, the pattern of growth will be plotted with the use of the method proposed by Ravallion and Chen (2003), the Growth Incidence Curve(GIC). In brief the GIC is a graph that indicates the growth rate in income between two points in time at each percentile of the distribution. The GIC displays whether income growth has been greater for some part of population. It allows us to say something about if the growth is pro-rich or pro-poor. This kind of graph has been used to study growth for several countries, such as Australia and England. In this thesis I investigate the patterns for Norway because it is interesting to see how "steady growth" affects the income distribution and inequality in a Western economy.

Secondly, I want to go deeper into possible explanations for the growth pattern by using the recentered influence function (RIF), proposed by Firpo et al. (2009b), with the logic of a Oaxaca-Blinder decomposition. This allows us to separate the effects that come from changes in observed characteristics, and the changes in the way these characteristics affect income. The separation is done by decomposing the effects into the composition effect and the structural effect.

I find that the recent economic growth in Norway has increased the income inequality. Most of the distribution have experienced an almost equal growth rate, but the highest percentile of the income distribution have experienced a higher growth rate than the average growth rate throughout the period. This is the main reason for the increase in inequality. On the other hand the results also point to the fact that individuals at the low end of the income distribution experienced a much higher growth rate than the mean growth rate, indicating that individuals who are poor in one year do not necessarily remain poor.

The remainder of this thesis is organized as follows. In Section 2 I present some background information on the economic development and inequality measures for Norway, and discusses some of the literature relevant for the relation between economic growth and income inequality. Section 3 outlines the methods and Section 4 the data used in this thesis. The results of the different approaches used are presented in Section 5. Section 6 summarizes the main conclusion.

2. Background and previous findings

2.1 Economic and inequality in Norway

At the beginning of the 1990s the Norwegian economy experienced the strongest recession since the second world war. This recession was largely due to domestic conditions, like the extensive debt for households that had been building up during the 1970s. The large debt, coupled with increased unemployment and falling housing prices lead to a downturn in household demand, again leading the economy into a recession. The international downturn in 1990 extended the downturn for the Norwegian economy til 1992, but in 1993 the economy saw a turning point.

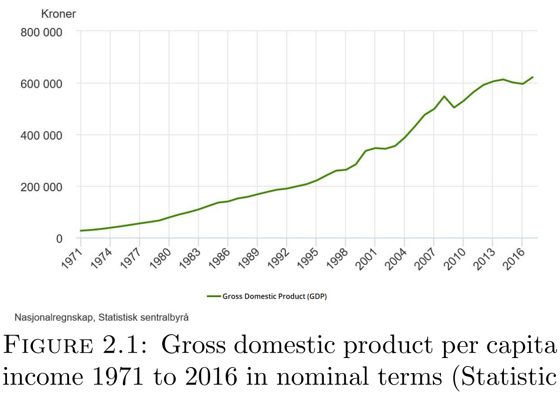

Figure 2.1: Gross domestic product per capita income 1971 to 2016 in nominal terms (Statistic Norway, 2017a)

By the start of 1993, the Norwegian economy entered into a prolonged upturn. Since then, the economy has experience almost uninterrupted growth. One natural factor explaining this growth is the petroleum industry and the management of the petroleum incomes. As shown in Figure 2.1 shows the per capita GDP in nominal terms has been growing steadily since 1970, however the growth was slower during the 70s and 80s, and after 1993 the per capita GDP has been growing more rapidly and almost tripled.

During the financial crisis the Norwegian economy experienced a small downturn, but the economy did comparatively well during the financial crisis due to active fiscal and monetary policies that moderated the decline in 2009.

The focus in this thesis is how this growth has been distributed in the population and thereby how it has affected income inequality. Let us also take a look at recent developments in inequality. Inequality is measured by the like Gini coefficient and the P90/P10 gap.

The Gini coefficient is the most commonly used measure for inequality, since it offers an opportunity to compare income inequality across countries and across time. In Figure 2.2 the Gini coefficient and the P90/P10 gap is shown for the period 1990 through 2015. The Gini coefficients varies between zero and one where zero means perfect equality while one equals maximum inequality, meaning that all the income goes to one individual.

The ratio of income to the person who has an income just over 90 percent of the population and the income of the person who has an income just over 10 percent of the population.

The P90/P10 income inequality gap is the ratio of the income of the top 90^{th} percentile and the 10^{th} percentile. Meaning that a ratio between the person who has an income just over 90 percent of the population, meaning that 90 percent have a lower income than the individual and 10 percent have a higher income, and the income of the person who has an income just over 10 percent of the population, meaning that it is the point where 90 percent have a higher income and 10 percent have a lower income. As a measurement for inequality it measures the gap between the top and the bottom income earners of all individuals, and the income of the individuals in the 10^{th} percentile, meaning individuals with the lowest 10 percent in income. As a measurement for inequality it measures the gap between the top income earners and the bottom.

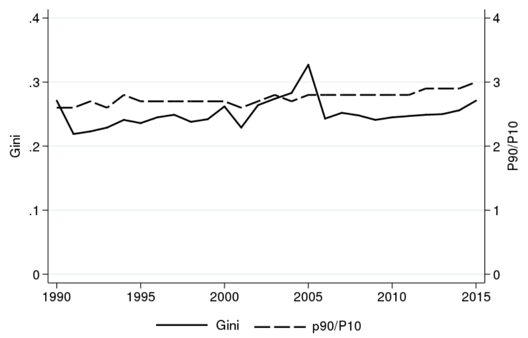

Figure 2.2: Change in the inequality measures Gini and P90/P10 from 1990 to 2015Statistic Norway (2017b)

Figure 2.2 shows that inequality has not changed a lot over this time period, by both measures. Inequality is somewhat higher in 2015 than in 1993 with a Gini coefficient increase from around 0.2 to about 0.25, and the same trend is observed for the P90/P10 ratio. This may indicate that the income growth has not been fairly distributed. At the same time the main takeaway from this figure is that income inequality overall has not changed much even though the growth has been substantial, indicating that growth might have little effect on income inequality.

The Gini coefficient cannot tell us anything about how the growth process is distributed amongst the population or which part of the distribution that benefits the most from economic growth, because the coefficient can change due to changes in the number of individuals in high-income or low-income groups and not because of changes in the incomes of these groups. Thus, the P90/P10 ratio might give a better picture, but at the same time this measurement does not take into account the income changes for the 80 percent in the middle of the distribution. Simply looking at inequality measures such as the Gini coefficient or the P10/P10 measurement is not sufficient to conclude on how the economic growth has been distributed amongst the population.

2.2 Relevant literature

In this section I present some of the literature and previous findings related to the effect of growth on income inequality and the distribution of income. The relations between income inequality, redistribution and economic growth have been investigated without any clear conclusion for cause and direction of effects.

Theory is ambiguous when it comes to the effect of inequality and poverty on economic growth. Some theories and empirical research point to inequality as having a positive effect on growth. Ostry et al. (2014) conclude that lower inequality is correlated with faster and more stable growth. While on the other hand, Aghion et al. (1999) find that greater inequality can have a negative effect on growth when capital markets are imperfect. Deininger and Squire (1998) suggest a strong negative relationship between initial inequality and long-term growth.

Another interesting relation is what effect redistribution have on economic growth. There are several arguments for the fact that redistribution has an effect on growth, but no clear indication for whether the effect is positive or negative. One of the arguments for redistribution having a positive effect on growth is human capital accumulation theory Galor and Zeira (1993), where the basic idea is that inequality can be harmful to growth due to reduced chances and opportunities for the poor. The poor often have less ability to stay healthy and have the resources to accumulate human capital, so the argument for redistribution as positive for growth comes from prospects for the income-poor. They will get more opportunities for education and by that increase their productivity.

One conceptualization for the negative effect of redistribution is the "leaky bucket" theory. It was proposed by Okun (1975) based on the fact that governments would need to prioritize and choose between equality or efficiency. The theory is that in the transfer from the income-rich to the income-poor some of the money would simply disappear, like water out of a leaky bucket. Okun attributed these losses to less incentives to work and invest due to transfers and taxes, and the administrative costs associated with the redistribution. The thought is that redistribution will have a negative effect on growth since it removed the incentives to work.

As illustrated with these examples the relations between redistribution, inequality and growth are quite interconnected and clear results are difficult to make. In this thesis I shall instead look at the "other side of the coin"- how growth influence income distributions and income inequality. More precisely, I provide descriptions on how growth is distributed. In the rest of this section I refer to some of the empirical evidence regarding this relationship.

When it comes to how growth affect the income distribution the dominating theory is the Kuznets curve. As mentioned previously, this is a graphical depiction of how income equality develops with growth. The curve has an inverse U-shape pattern, with income inequality increasing with GDP up to a turning point where inequality decreases as GDP increases. According to this theory growth can have a positive or negative effect on income inequality depending on whether the economy has reached the turning point or not. One would expect developed economies to be past the turning point and be on the right-hand side of the curve, meaning that growth would lead to lower income inequality.

When Kuznets published this theory he relied on historical data from the beginning of the nineteenth century from only three countries: the United States, England and Germany. The historical argument for the shape of the curve is that the societies became more unequal in the first stage of the industrialization, and then inequality decreased as societies matured. Regardless of this it has been used as a starting point for many empirical investigation into the relation between income growth and income inequality.

Deininger and Squire (1998) is amongst the ones who discuss the empirical support of the Kuznets curve. They find that there is virtually no support for an increase in inequality at low levels of income and decrease at higher levels. This leads them to the conclusion that this theory is not relevant for developing countries, or that the curve is too flat to be noticeable.

Further, Palma (2011) finds by doing an empirical investigation that the upward side of the U-shape has evaporated, and with it the theory that inequality has to increase before it can decrease. He finds that for some countries inequality has increased with the growth in recent years, while for other countries economic growth has lead to less inequality. Another investigation find that only about 10 percent of countries experience the invert U-shape and that most countries exhibit no statistically significant tendency (Fields, 2007).

One would expect that Norway would be on the right-hand side of the Kuznets curve, implying that the results in this thesis should show that recent growth in the Norwegian economy has lead to a reduction in income inequality, i.e, the growth has benefited the poor more than the rich.

Piketty and Saez (2003) look at the shares of top income earners over a long time period in a developed economy (US). They find that the shares of top income earners follows a U-shape, with top income earner having a lower income share in the time after the second world war, but that this was only temporary. The shares of the top income earners increased again, meaning that the income inequality increased, contrary to Kuznets hypothesis. The increase in the shares of the top income earners have continued after 1998 (Saez and Zucman, 2016).

The results from these papers show that the Kuznets curve cannot accurately explain the relation between income growth and income inequality. From previous research one can see that no simple conclusion on the effect of economic growth on the income distribution can be drawn, since growth can have both positive and negative effects on income inequality. It has proven to be difficult to disentangle cause and effect definitively in the relationships between income growth and income inequality. The literature which has focused on whether higher level of income is associated with higher or lower inequality, seems to have searched a consensus that there is no overall net effect (Ostry et al., 2014).

This same conclusion is found by Dollar and Kraay (2002) which compared the average income of the poorest 20 percent of the income distribution with the average income for the entire distribution, to see whether income for the income-poor increased or decreased by the same rate as average income. The investigation is done by using a large sample of countries spanning four decades, with the result that the income of the poorest rise and fall at the same rate as average income, meaning that income inequality would be unchanged with growth. Dollar et al. (2016) find the same results when they do a similar test in 2016. They use a dataset containing 121 countries and find that the data suggest that the income shares of the bottom 20 percent and bottom 40 percent show no systematic tendency to decline over time, concluding with no worldwide trend towards a greater inequality within countries.

The book Capital in the Twenty-First Century (Piketty, 2014) has since it was published attributed substantially to the public discussion on the possible linkages between inequality, redistribution and growth. Piketty is stressing how accumulation of capital among the few is making the relative value of work and effort less important. The argument for this is that capital under normal circumstances will pay a higher interest than the overall growth level in the economy, that returns on capital is higher the the growth rate in the economy. Individuals who accumulates capital will over time will get relatively richer than wage earners by doing nothing else than holding capital, increasing inequality between individuals holding a lot of capital and individuals with little or no capital. Furthermore, Piketty points out that it has become increasingly important to be born by rich parents. The question here is if the same pattern is found in Norway: has the case that the economic growth in Norway has mainly benefited the richest parts of the income distribution.

2.3 Literature using methods as in this present study

This thesis focus on the effects of growth on the income distribution. Research focusing on the effect of growth on inequality often focus on the poorest, so-called pro-poor growth. Pro-poor growth refers to growth that, that provide the poor with opportunities to improve their economic situation. There are three main definitions of pro-poor growth. The first by Datt and Ravallion (1992) and Ravallion and Chen (2003) simply defines pro-poor growth as any growth that leads to a reduction in poverty. This is often pointed to as the weakest definition since it means that almost all growth is pro-poor. The second definition is often referred to as absolute pro-poor growth. Growth is pro-poor if the poor enjoy a greater absolute benefit than the non-poor (Grosse et al., 2008). The final definition is the relative definition. This says that growth is pro-poor if the incomes of poor people grow faster than the rest of the population, the poor benefit proportionally more than the non-poor. Or put another way growth results in a redistribution of income in favor of the poor (Kakwani et al., 2000).

For this thesis the focus is how the economic growth has been distributed across the whole population. One way this has been done is by placing every person in the economy with their height proportional to their income and then order them from highest to lowest, with the highest, the giants being the richest and the dwarfs being the poor. Income growth over time corresponds to changes in the heights of Parade participants. This is something called Pen's parade, introduced by Pen (1974). The logic of Pen's parade is taken one step further in the calculation of the growth incidence curve (GIC). A GIC is a graphical illustration that shows the quantile specific rate of economic growth between two points in time as a function of each percentile.

In the paper by Lakner and Milanovic (2016) the global GIC for the period 1988-2008 was piloted using national household surveys. By choosing a global perspective the authors focused on the development of international growth on global inequality. They find something that has been refereed to as the elephant graph, a graph that has a distinct supine S shape with the highest growth rate at the middle of the distribution peaking between the 50^{th} and the 60^{th} percentile. The top one percent of the distribution also experienced growth higher than the mean. The accumulation of wealth at the middle of the distribution can be explained by the economic expansion in Asia, especially from China.

Using the strong economic growth that Australia experienced at the start of the 2000 Azpitarte (2014) investigated the pro-poorness of the growth. Using household income data from the HILDA survey from the years 2001-2008 he finds that the growth could be called pro-poor, in terms of the first definition (see above). However,f was the highest percentile that experience the highest growth rate meaning that the income growth did not lead to a reduction in income inequality.

Futher Essama-Nssah et al. (2013) do not just want to investigate pro-poor growth and simple growth incidence analysis, they wish to account for the heterogeneity. After they use the recentered influence function (RIF) regression to link the GIC to household characteristics, they are able to perform a detailed decomposition to try and explain how different household characteristics shape the distribution. They apply this method to consumption expenditure in Cameroon in 2001-2007, to identify the sources of the variations. Their main conclusion is that the structural effect is the main driver of the observed pattern of growth, while composition effects accounts for the largest share of the observed variation in the social impact of growth. Their methodological approach shows similarities with the framework used here.

One of the key assumptions behind the GIC is the anonymity axiom. Anonymity axiom implies that the growth assessment is based on a cross-sectional comparison of the marginal distribution before and after the economic growth, not taking income mobility into account. Mobility means that the people who are poor this year might not be the same people who were poor last year. Poverty becomes more problematic if it is the same individuals how are poor year after year. If one want to evaluate the full impact of growth on inequality it is important to have an understanding of this, how growth affects the re-ranking of individuals. Breaking the anonymity axiom would imply that individual growth is assessed on the basis of there placement in the initial distribution. To assess whether the individuals who are poor (or rich) this year are gainers or losers, one has to track the fortunes of individuals not the fortunes of income groups such as the poor or the rich whose composition may change from one year to the next. Taking there initial status in the distribution into account also allows us to se the effects on lifetime income.

By not taking the initial status of the individuals into account when evaluating the effect of growth, it can be hard to identify who the winners and losers from the growth is. If the initially poor remain poor throughout a longer period of growth that makes the poverty more problematic. Information on this can be especially useful when evaluating the efficiency of policy reformsPalmisano (2018). Basically it is of interest to get a sense of how income growth is beneficial for those who were initially most disadvantaged. It is of policy interest to check if economic growth is beneficially to the initially poor, helping them get out of poverty, while at the same time maybe push some originally non-poor into poverty. One can think that by ignoring income mobility is almost the same as saying that only the post-growth income distribution matters in the social evaluation Bourguignon (2011). There is an increasing number of contribution that in recent years have proposed alternative models to evaluate and rank growth processes.

Bourguignon are amongst them that have proposed different methods for calculating the effect of growth rate over the distribution, taking the initial distribution into account. One of these is the non-anonymous growth incidence curve which plots the rates of growth in income against quantiles for the initial distribution of income (Bourguignon, 2011). This is done by fixing the rank of income units according to the initial distribution, linking each of these units to the corresponding in the post-growth distribution. Thus at a given point, the non-anonymous GIC measures the mean income growth of all units located at the p^{th} quantile of the initial distribution.

Another method proposed by Van Kerm (2009) and Jenkins and Van Kerm (2011) is to calculate an income mobility profile or income mobility curve. This curve contrasts the GIC by summarizing income changes for the same individual and depends on the bi-variate distribution of income. As the name states the income mobility profile is a tool that can be used to say something about the income mobiltiy and identify teh association between individual movements and the individuals initial status. The income mobility profile (cumulative mobility profiles) are plots of the average income growth for people with an initial income at or below a given percentile x(p) in the base-year distribution. By doing this one can illustrate the effect of growth among the poorest part of the population and not just the growth experienced by the poorest quantile. They argue that this gives a better understanding on the effect of growth since the members of the poorest quantile might not be the same from one period to the next.

Jenkins and Van Kerm (2011) point out that studying income growth in the distribution is commonly done using income values (quantiles), but do not take into account that income groups and quantile compositions changes over time due to income mobility. Those that are poor in one period might not be it in the next period. So what they do is introduce something they call the income mobility profile. This allows them to summarize the pattern of income growth while at the same time tracking the fortunes of the same individual. They apply this method to The British Household Panel Survey in the period 1991-2005, where they find that income growth was more pro-poor under Labor government (1997 to 2007) especially during the early years, than under the early years of the Conservative government.

Grimm (2007) uses the 'individual growth incidence curve'(IGIC), arguing that the GIC might give an incomplete picture. The IGIC curve is found with much of the same logic as in the previously mentioned papers. Individuals are ordered according to their position in the initial distribution and quantile specific growth rates are computed. This method I will come more back to later since this is the method breaking with anonymity that I will apply in this thesis. Comparing the results from the GIC and the IGIC using data from Peru and Indonesia, he shows that the IGIC gives a different results compared to the GIC, and by this concludes that the GIC gives an incomplete picture of the effect of growth and that the best would be to include both types og methods

3. Empirical investigation

3.1 GIC and IGIC

In this section I will present the basics for the growth incidence curve, henceforward GIC, introduced by Ravallion and Chen (2003) as a method for measuring pro-poor growth. Other researchers have expanded the method as to account for non-anonymous distribution, and by that allowing us to study whether there is persistence in poverty for the same people. Since I have access to panel data, I am able to to calculated both the GIC and the IGIC.

Growth Incidence Curve

The GIC is an instrument that can be used to illustrate how the growth rate could be different across various parts of the distribution. Simply put the GIC shows the growth rate of each quantile for the income distribution between two points in time as a function of each percentile and can thus be considered as an indication for the pattern of growth.

One important assumption for the standard GIC is the anonymity axiom, which means that we assume that the original income level does not matter for the final distribution. This implies that growth assessments are based on a cross-sectional of the marginal distribution of income before and after the economic growth. This kind of measurement will not consider any personal characteristics of the individuals that occupy the different positions in the distribution.

Income is assumed continuously distributed over the entire population of interest. $F_t(y)$ denotes the cumulative distribution function (CDF) of income, showing the proportion of the population with less than y income at period t. The invert of the CDF at the p^{th} quantile gives the income of that quantile: $F_t^{-1}(p) = y_t(p)$. The GIC is defined as:

$$g_t(p) = \frac{y_t(p)}{y_{t-1}(p)} - 1 \tag{3.1}$$

Where $q_t(p)$ shows the growth rate of the p^{th} quantile. p can vary between 0 and 1, 1 being the top end of the income distribution, and $p = 0.5$ that means the 50^{th} percentile.

The GIC, as defined in equation (3.1), compares the two periods t and $t-1$ and evaluates the impact of the growth rate in income of the p^{th} quantile, where $t-1$ is the initial period and t is the period after the growth. In this thesis I will look at the growth rate in income from 1993 to 2015.

A downward sloping GIC indicates that growth contributes to equalize the distribution of income; the growth rate decreases as p increases. An upward sloping GIC indicates a non-equalizing growth; the growth rate increases with p and the top part of the distribution experience the highest level of growth. The in follows that if GIC is a horizontal line it indicates that all parts of the distribution experience the same rate of growth and the inequality does not change over time, the GIC is horizontal.

When studying the effect of growth on the distribution, the effect can be split into two components. The first is the size effect. This is essentially the growth rate experienced by the entire distribution, if the growth has been distributed neutrally then everyone experience the same income growth. This would mean that the growth rate for mean income is the same as the growth rate for every percentile of the distribution. The second component is the redistribution effect of growth: a purely redistributional effect would mean that only relative inequality would change with growth and not the mean income.

The GIC can be used to test whether the final distribution first-order stochastically dominates the initial distribution.

Individual Growth Incidence Curve

The main difference between GIC and IGIC is that in the IGIC individuals are ordered according to their position in the initial income distribution, holding membership to the initial quantiles fixed. The outcomes for individuals in each such quantile are identified in the final distribution and used to compute the quantile specific means and growth rates. The individual growth incidence curve (IGIC) is defined as:

$$g_t(p) = \frac{y_t(p_{t-1})}{y_{t-1}(p_{t-1})} - 1 \tag{3.2}$$

The construction of the IGIC is closely related to the standard GIC, with the main difference that the final period income y_t is defined by an individual rank in the initial period p_{t-1}, which for this thesis means that the distribution of income in 1993 defines the ranking of individuals that is used for calculating the differences to 2015. This curve plots the rates of growth in income quantiles for the initial distribution of income, so that at any given point the IGIC measures the mean income growth of all units located at the p^{th} quantile of the initial distribution, which in this case is the distribution in 1993. This calculation allows us to take the mobility and the re-ranking into account for the overall growth process.

Calculating the IGIC requires access to panel data, so that one can follow the individual over time. If this is not available and the researcher only have access to cross-sectional data the results will reflect this by the fact that only GIC can be calculated.

Even though the GIC and IGIC are defined quite similarly they measure completely different things. With the GIC I find the growth rate in income for the different percentiles of the distribution, not taking into account that individuals could have changed percentile membership during the growth period. This measures how growth is distributed along the income distribution, while the IGIC measures the individual growth experienced by holding percentile membership fixed during the growth period. Thus, the growth rate calculated by the IGIC can also give indications on income mobility in Norway.

3.2 Further information about explanations

3.2.1 Recentered influence function

In this section I will go into how RIF regression can be used to study the underlying factor shaping the pattern of growth identified by the GIC and identify how the returns to various characteristics might vary for different parts of the distribution. Recentered Influence Function (RIF) regression was introduced by Firpo et al. (2009b), and is a method to evaluate the impact of changes in the distribution of the explanatory variable on the quantiles of the unconditional distribution of an outcome variable. In this case the method is used to evaluate the impact of different household characteristics on income. The RIF regression method is a quantile regression method, and it allows us to appreciate the extent of heterogeneity in the impact across different quantile. Where OLS only gives the average impact of the different characteristics, a quantile regression give information about effect in different parts of the distribution. One can think that the returns to the various characteristics vary along the distribution, and using quantiles allows us to zoom in on the different parts and study these variations. The idea is that these variations might allow us to identify what shapes the GIC, why some have experienced a different economic growth than others.

Using RIF for this thesis is relevant since it offers a simple way of establishing a direct link between the income distribution and the household characteristics. The link offers a simple way of performing a detailed decomposition for any distributional statistic for which an influence function (IF) can be defined. The RIF regression is similar to a standard regression model except that the dependent variable Y is replaced by the recentered influence function for the statistic of interest, in this case the dependent variable is replaced by a RIF of the income variable.

The recentered influence function can be defined as:

$$RIF(y, q) = q(F_Y) + IF(y, q) \tag{3.3}$$

$IF(y, q)$ is the influence function corresponding to the observed income y for the distributional statistic of interest, q(F_Y). An influence function is basically an analytical tool that can be used to asses the effect on the value of a statistic by removing an observation without having to re-calculate that statistic. An influence function taking quantiles into consideration can be defined as:

$$IF(Y, q_p) = \frac{p - I(y \leq q_p)}{f_Y(q_p)} \tag{3.4}$$

where $I(.)$ is an indicator function for whether the income variable is less than or equal to the pth quantile, and $f_Y(q_p)$ is the density function of the income variable y evaluated at the pth quantile of the distribution. The RIF equation then becomes:

$$RIF(y; q_p) = q_p + IF(y; q_p) = q_p + \frac{p - I(Y \leq q_p)}{f_Y(q_p)} \tag{3.5}$$

Using equation (3.5) repeatedly for each of the different percentiles of interest indicates how the household characteristics might have different impact on income for various percentiles along the income distribution.

3.2.2 Decomposition

The next step is to use the results of the RIF regression to preform an Oaxaca-Blinder decomposition. This enables us to summarize what has happened between 1993 and 2015 in a specific way. It is useful to go into what might explain the changes in income distributions over time, separating the effect that comes from changes in observed characteristics and changes in the way these characteristics affect income. This is done through a decomposition where we try to distinguish between the underlying factors, namely the composition effect (endowment effect) and the structural effect (price effect).

As the name suggest, the method was introduced by Blinder (1973) and Oaxaca (1973). It is a statistical procedure that often is used for explaining the difference in the means of dependent variables between two groups. The decomposition is done by dividing effect of two groups ($t = 0, 1$) into the composition and structural effect.

$$\triangle_O^\mu = \mathbb{E}(Y_1) - \mathbb{E}(Y_0) = \mathbb{E}(X_1)'\boldsymbol{\beta}_1 - \mathbb{E}(X_0)'\boldsymbol{\beta_0}$$

In this thesis instead of looking at the difference in the mean, I look at the difference across the distributions and between the two years in question 1993($t = 0$) and 2015($t = 1$).

This method has often been used for studies in labor economics, for example identifying wage discrimination between men and women. However, it can also be employed to study group differences in any continuous and unbounded outcome variables, with clearly defined groups.

Here I present some of the basics for the calculation of the classical Oaxaca-Blinder decomposition. The goal of the Oaxaca-Blinder decomposition is to decompose differences in mean income growth into two groups $t = 0, 1$ with an outcome variable Y, $\triangle_O^\mu = \bar{Y_1} - \bar{Y_0}$ into the composition and structural effect.

$$\triangle_O^\mu = \triangle_S^\mu + \triangle_X^\mu \tag{3.6}$$

The overall difference between the two groups is defined as $\triangle_O^\mu$, divided into the composition effect ($\triangle_X^\mu$) and the structural effect ($\triangle_S^\mu$). By the composition effect we mean looking at changes due to the distribution of observed characteristics. The structural effect is the change in the returns to different characteristics, whether the returns vary along the distribution.

The standard Oaxaca-Blinder decomposition assumes a linear model. The conditional expectation of Y given by X is assumed to be linear:

$$\mathbb{E}[Y_{ti}|X] = X_i'\boldsymbol{\beta_t} + \epsilon_{ti} \qquad (t \in 0, 1) \tag{3.7}$$

X_i' is the vector containing the predictors, in this case the different household characteristics and a constant. β contains the slope parameters and the intercept, and ϵ is the error term. It is this β coefficient I want to identify in the RIF regression and use to perform the decomposition. From the linearity assumption $\mathbb{E}[\epsilon|T = t] = 0$, thus the error term can be removed:

$$\begin{gathered}\triangle_O^\mu = \mathbb{E}[X|T = 1]\boldsymbol{\beta_1} - \mathbb{E}[X|T = 0]\boldsymbol{\beta_0} \\ \triangle_O^\mu = \mathbb{E}[X|T = 1]\boldsymbol{(\beta_1 - \beta_0)} + (\mathbb{E}[X|T = 1] - \mathbb{E}[X|T = 0])\boldsymbol{\beta_0}\end{gathered} \tag{3.8}$$

Equation (3.8) is the simple Oaxaca-Blinder decomposition where the first part of the equation is the composition effect and the second part is the corresponding structural effect and where β_0 and β_1 are returns on observable characteristics in year $t = 0$ (1993) and $t = 1$ (2015), respectively.

For the decomposition in this thesis I use the recentered influence function (RIF) with the logic of Oaxaca-Blinder decomposition to decompose the composition and structural effects, instead of just looking at the mean. One thing to note about the influence function of a distributional statistic is the fact that the expected value is equal to zero, which implies that the expected value of the RIF is equal to the corresponding distributional statistic: $q(F_y) = E[RIF(y; q)]$.

By the law of iterated expectations this means that the distribution statistic of interest can be written as the conditional expectation of the recentered influence function, given the observable characteristics (X). This is the RIF regression that, for $q(F_y)$, can be expressed as: $E[RIF(y; q)|X]$. The distributional statistic $q(F_y)$ can therefore be expressed in terms of this conditional expectation as follows:

$$q(F_Y) = \int \mathbb{E}[RIF(y; q)|XdF(X)] \tag{3.9}$$

In its simplest form, the approach assumes that the conditional expectation of the $RIF(Y; q_p)$ can be modelled as a linear function of the explanatory variable, then by the OLS one gets the following equation in order to estimate β:

$$RIF(Y; q) = X_i' \boldsymbol{\beta_t} + \epsilon_i \tag{3.10}$$

The expected value of this linear approximation of the RIF regression will be equal to the expected value of the true conditional expectation, so $E[RIF(Y; q)|X] = X\beta$. This fact makes the extension of the Oaxaca-Blinder decomposition to RIF regression both simple and meaningful.

Applying the standard Oaxaca-Blinder approach to equation (3.10), the decomposition can be defined as it is defined in equation (3.8).

There are a number of limitations on the decomposition method, and the goal of the decomposition is often quite ambitious, which means that strong assumptions typically underlie this method. If the linear model is misspecified, it results in misleading classification into the structural and composition effects. Furthermore, Oaxaca-Blinder decomposition assumes that the groups are mutually exclusive, meaning that no person can be identified in both groups. In my analysis I do not follow individuals, so an individual appearing in both groups will be defined as different individuals due to changes in income. The focus in the classical Oaxaca-Blinder is only on the mean, which limits the possibility to address complex development in the income distribution. To account for these complexities, I will use this method in combination with the RIF regression as explained above. The aim is that using this decomposition method might explain some of the underlying factors shaping the pattern of growth.

4. Data Description

4.1 Dataset and definitions

The point of departure for this thesis is an ambition to discuss the distribution of income growth. To do this I use register data for the period from 1993 to 2015, based on income and wealth statistics(Statistic Norway, 2017b). The register data contains information spanning the whole population as of the 31st of December of the income year. The registry is created by linking different administrative registers and statistical data sources, amongst others it contains information from: income tax returns, education statistics, and data from the Norwegian Labour and Welfare Organisation (NAV), this means that in addition to information on income, the data contains information about a large number of variables, such as age, gender and education.

Among the variables listed above, the income variable is the most central one in my analysis. I use an income variable for disposable income defined as the sum of wages and salaries, capital income and government transfers, minus tax payments. By using disposable income one can get a better picture of the actual consumption opportunities of the individuals. To avoid any effect of inflation, all monetary values have been adjusted to 2015-levels.

I aggregate income over household membership when establishing the income concept used in the analysis. The purchasing power of a family or household is determine by the shared income, which means that it is the total income of the household that determines whether a family is poor or rich. Using household characteristics is also similar to what is done by other researchers of similar studies, like Azpitarte (2014) and Essama-Nssah et al. (2013).

Since the data set contains information on the individual level, households needs to be constructed. This is done by using Statistics Norway's definition of families, which are identified by a reference persons identification number becoming the identification number for the family. Statistics Norway defines families as spouses/registered partners with or without children, single parents with children, and cohabitants with at least one common child have the same family number. Cohabitants without children have different family numbers, which is a problem. The age of the children is not taken into account so adult children living with their parents are treated as belonging to the same household as their parents. In 2005 Statistics Norway started to register households, where a household is defined as people with the same address in the National Registry. This definition of household is not used in this thesis since using different definitions for the household would make the comparison difficult. Therefor I keep using the family definition, defined the same way as in 1993. Eventhoug I use Statistics Norways family definition I refere to them as households in this thesis, where each household in the analysis will be represented by a "head of household". The head of household is defined as the individual who is the reference person for the family number. This means that each family/household is represented by one person, but the income cosidered is based on aggregate income.

The income variable used is defined by the size of the household using equivalence scales. Equivalence scales can vary between 0 and 1, in this thesis I use the square root equivalence scale.

$$Y = \frac{\sum_i y_i}{\sqrt{n}} \tag{4.1}$$

Y is the income variables used to represent the household income, defined by summing together the individual incomes of the members of a household, y, divided by the square root of the size of the household, n. Using this method takes the household size into consideration, while also reflecting the fact that the income need for a household does not grow proportionally with the number of members.

Furthermore, some restrictions has been put on the data. Firstly, any individuals with missing information on gender, age, etc., has been removed. This also includes individuals for which a household can not be identified. Secondly, households where the head of household is under the age of 18 has been removed, and finally individuals with negative or extremely small incomes has been removed. After these adjustments, I am left with 2,069,714 households in 1993.

Here I will use the fact that the data contains identification numbers for the individuals to track the individuals over time. As mentioned in Section 3, the access to panel data is a prerequisite for calculating the IGIC. For this calculation an additional restriction is put on the data, only individuals that are present (or registered) in all the years which the analysis span, is included in the analysis.

This means that households not present in the initial year, or have left the data set (for example people die).

4.2 Household characteristics

The household characteristics are used for the RIF regression and the Oaxaca-Blinder decomposition. Most of the variables used are dummy variables calculated on the basis of the whole household, but some variable like age, gender and education are only based on the head of the household.

The choice of dummy variables implies that the reference household lives in Oslo, has a male head of household with an education level lower than high school, is not currently a student and does not have immigrant background or parents with immigrant background. The coefficients for the dummy variables are for households that are different from the reference household in one or more accounts.

TABLE 4.1: Household characteristics

	1993		2015	
Variable	Mean	std.dev	Mean	std.dev
Age of head	50.2	19.17	49.9	18.1
No. of children	0.47	0.89	0.4	0.83
	Percent		Percent	
Female	34.1		35.9	
University/College	10.4		31.3	
High School	56.7		41.5	
Student	4.7		6.1	
Immigrant	5.7		17.0	
Immigrant parent(s)	2.2		4.1	
Eastern Norway	37.3		37.1	
Southern Norway	5.3		5.4	
Western Norway	24.0		25.0	
Trøndelag	8.8		8.8	
Northern Norway	10.8		9.4	
Observations	2 069 714		2 761 643	

There are five regional dummies explaining where the family lives. Each of these take the value one if the household live in the specific region. The definition of the regions are based on the regions in Norway

5. Results

In this section I present the results. In section 5.1 I present the results from the calculation of the growth incidence curve and the individual growth incidence curve before section 5.2 presents the results from the RIF regression and in the last section 5.3 the results from the decomposition is presented.

5.1 Results GIC and IGIC

Table 5.1 suggest that mean and median income grew with an average of 2.43 percent over the entire period. The results in the table are calculated on the basis of cross-sectional data, based om the anonymity axiom. When splitting the period up into smaller periods of six to seven years, one can see that the growth rate has some variations through the period, with the largest average income growth happening during the period 1998-2004. This is also the period where the GDP per capita grew at the highest rate.

TABLE 5.1: Annual income growth in Norway 1993-2015

	Mean		Median	
Period	Variation (NOK)	Growth rate	Variation (NOK)	Growth rate
1993-2015	6 878	2.43	5 573	2.19
1993-1998	5 205	2.29	4 290	2.06
1998-2004	6 969	2.62	4 233	1.79
2004-2010	5 230	1.69	6 495	2.39
2010-2015	6 929	1.99	4 558	1.47

[a] Note: All incomes are adjusted to 2015 levels

Growth Incidence Curve

I have calculated five different GICs. Figure 5.1 shows the growth incidence curve for the income growth between 1993 and 2015. Appendix A shows curves for shorter periods of time.

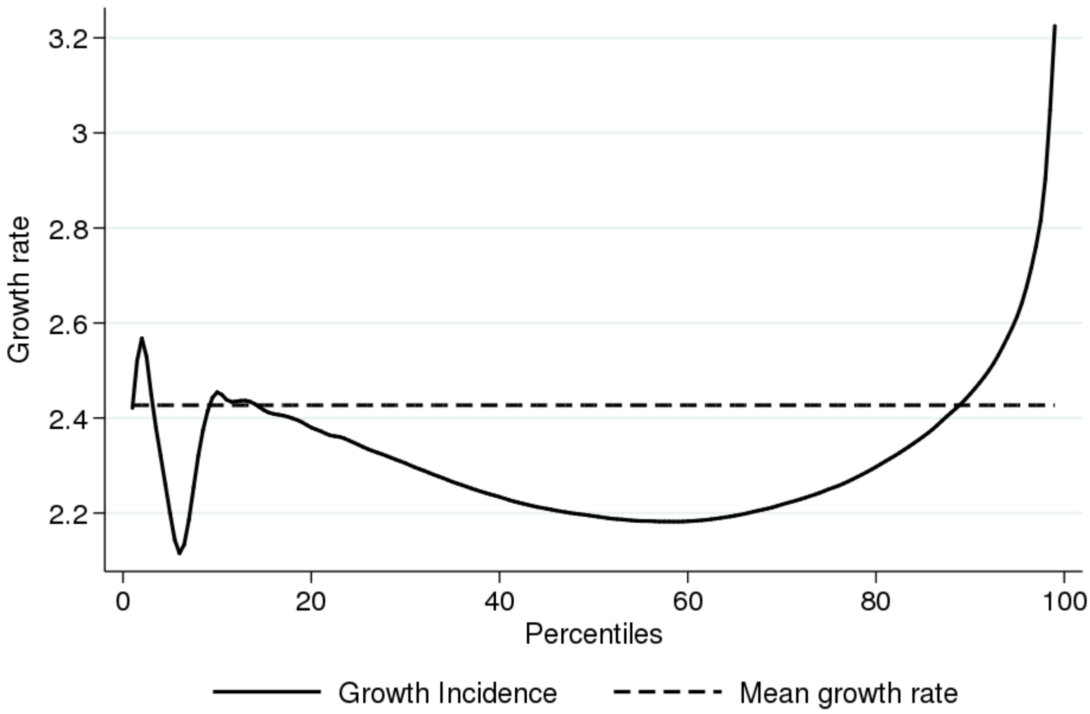

FIGURE 5.1: Growth incidence curve for Norway, 1993-2015

From Figure 5.1 it is clear that the winners are the top 10 percent of the population with the highest annual growth rate for the absolute top end of the distribution. A large part of the distribution is located below the average income growth rate, but the majority has an annual growth rate above 2.2 percent. The most interesting part of this distribution is the shape of annual income growth at the bottom 5^{th} percentile, where the absolute poorest have an annual growth rate higher than the mean, but was causes this effect is unclear. No part of the distribution has experienced negative growth, meaning that every part of the distribution is richer in 2015 than in 1993.

In Appendix A, Figure A.1 the GIC is displayed for shorter periods of time. One showing the results from 1993 to 1998, one from 1998-2004, from 2004-2010 and the finale graph is from 2010 to 2015. The same pattern is found for most of the shorter time periods as in the period 1993 to 2015, with the middle of the income distribution experience approximately the same growth rate in all periods. The main variations is at the top end and bottom end of the distribution. The graph spanning the period from 2010 to 2015 is the only graph that is strictly increasing, where the bottom part of the distribution has the lowest growth rate while the top end has the highest.

In the period spanning from 2004 to 2010, the top end of the income distribution experiences significantly lower growth rates than the rest of the population. One possible explanation for this is that during the financial crisis, capital investments yielded lower income, and a possible theory for the rich getting richer is that they mainly benefit from capital investments. To obtain further understanding of this, one could look at the growth in different mean components. But that is beyond the scope of this analysis.

By just looking at Figure 5.1 it is difficult draw any conclusion on what effect the growth has had on the overall income equality. The figure does not point to any higher growth rate, among the income poor (more than other parts of the distribution). If this graph shows us anything it is that the economic growth has lead to higher income inequality, in that the growth rate of the richest is substantially higher than the rest, confirming the development in the Gini coefficient as illustrated in Figure 2.2. The fact that economic growth has lead to an increase in income inequality is the opposite of what one might expect if the Kuznet hypothesis was true. The next thing to look at here is to identify if it is the same people that are poor in year after year. This is done by the calculation of the IGIC.

Individual Growth Incidence Curve

Recall that for the calculation of the IGIC a smaller sample is used than for the GIC, due to the restriction that only individuals that are present in both years are included. The dataset is based on tracing out the individuals income growth, meaning that only individuals who are 18 years or older in 1993 are included in the dataset.

The results from the IGIC spanning the entire period is presented in Figure 5.2. The first thing to notice is that the curve has the opposite trend of the GIC, the IGIC for the entire period being strictly decreasing. The bottom end of the distribution experienced the highest growth rate, which for the bottom 1 percent is above 10 percent. The obvious reason for this is that the bottom might consist of many young individuals who in 1993 are in the beginnings of their career and therefore experience a massive increase in income.

The parts of the distribution that experience a growth larger than zero are richer in 2015 than in 1993, leaving only the top end of the distribution poorer in 2015 than in 1993 according to the IGIC. People at the top end of the distribution might have left a high paying job to go into retirement and by that also reducing

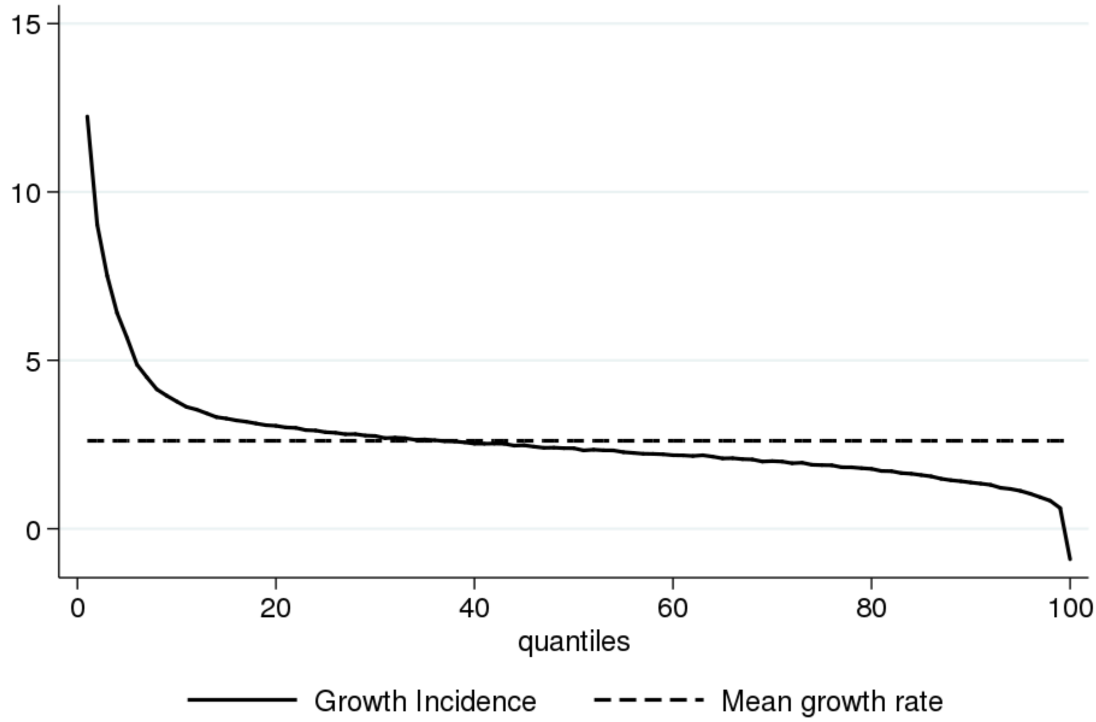

FIGURE 5.2: Individual Growth incidence curve for Norway, 1993-2015

their income, which may explain why I see that the rich have experienced negative income growth.

The IGIC from 1993 to 2015 use the distribution is defined by individuals income in 1993. To test whether this pattern holds for other years, with other initial years I have calculated more detailed results presented in Appendix A, Figure A.2. They are calculated for the same time periods as the GIC. In each of the shorter periods the initial yeas of that period defines the distribution for that particular graph, so for the graph showing the IGIC from 1998 to 2004, the distribution is defined from the distribution in 2004.

All the results in Figure A.2 show the same pattern: a decreasing slope where the poorest experience the highest growth and the richest experience the lowest growth, with the absolute top part experiencing negative income growth. Again, this is likely explained by an age-earnings profile. According to an age-earnings profile found by Bhuller et al. (2017) individuals will experience a sharp increase in income while young, peak at around 50 years and then experience a slightly lower income. So, if the lowest percentiles of the population is dominated by young people then the large increase in their income can be explained by going from being a student to start working and being at the start of the age-earnings profile.

IGIC can provide us with information about the economic and social mobility in Norway. The shape in Figure 5.2 point to there being a high income mobility in Norway and the results shows that individuals are not stuck in one part of the distribution. It can also be though of as a curve showing income mobility (Jenkins and Van Kerm, 2011). The shape of curves defined here is similarly to the results found by Palmisano (2018) when he evaluates the na-GIC (similarly defined as IGIC) for other countries like Switzerland and Germany. The same is found by Van Kerm (2009) when studying income mobility profiles for amongst other Greece and Ireland.

The GIC and the IGIC provide different insights into how economic growth is distributed. The GIC point to economic growth as leading to increased inequality in Norway since the top percentile experience the highest growth rates. At the same time the IGIC point to these top income earners as not being the same in both periods. According to the IGIC the top percentile have experienced a negative income growth. At the same time this indicates that the poor are not the same for both years. So, even though the inequality has increased with growth, the poor are not the same people for both periods.

5.2 RIF results

In this section I will present some of the RIF results and discuss how the coefficients of the various characteristics might vary for the different parts of the distribution and how these characteristics might be a driving force for the growth pattern presented above. The RIF regressions[1] are carried out for the years 1993 and 2015 using the same variables for both years. Table 5.2 displays the regression results from 1993 and Table 5.3 the results from 2015. The first column in each table shows results from a simple OLS regression. The remaining five columns are RIF regressions for the specific quantiles 10, 25, 50, 75 and 90. These tables show the coefficients and the associated standard error in the parenthesis. The first thing to notice about the results for both periods is that most of the results are significant, with only a few exceptions. For a graphical depiction of the estimated RIF coefficients for the different characteristics see Appendix B, Figures B.2 and B.1. For these graphs the RIF regression procedure has been used on all quantiles from the 5^{th} to the 95^{th} as well as results from the first quantile and the 99^{th} quantile.

[1]calculated using Firpo et al. (2009a) providing a RIF regression procedure.

Using the quantile regression as mentioned, allows me to look at how the impact of different characteristics might vary along the income distribution, and by comparing the results from 1993 to the results in 2015 one can also study the change over time. Overall these results gives an impression on how different household characteristics affect the distribution of income showing that different characteristics affect the income in very different ways, with some variables having large variations across the distribution.

For the most part the results are as one might expect, with a positive effect of higher education and higher age, while other characteristics such as being a student or the effect of one additional child in the family has a negative effect on income. For most characteristics there is not a large change from year to year, with some exceptions: having a female head of household, the effect of one additional child and the effect of living in Western Norway. Below I will go more into depth on some of the results presented in Table 5.2 and Table 5.3 and discuss what might explain these differences.

In 1993 having a female head of household is associated with a lower income than households with a male head of household. Having a female head of household indicates in most cases that there is no male provider in the household, the fact that this coefficient is negative, can be explained by the gender pay gap. Another important finding suggest that the negative effect varies along the distribution, with the largest negative effect being around the 25^{th} percentile, and the lowest negative effect is at the highest percentile. This RIF coefficient is one of the coefficients showing the largest variations across the distribution. Some of this variation might be explained by variations in the distribution of female households, with fewer of female lead household at the top end of the income distribution than at the middle. Both 1993 and 2015 show the same pattern, but the negative effect associated with a female has increased. The largest negative effect can in 2015 be found a bit further out in the distribution at around the 50^{th} percentile. There is no large change in the size of this variable from 1993 to 2015, so the larger negative effect can not just be explained by changes in the distribution of female heads of household, which might indicate that the gender pay gap has increased from 1993 to 2015.

In 1993 additional children results in a decrease in expected income. This may be due to the definition of individual income, as it takes the size of the family into consideration. A child increases the household size, but will in most cases not increase the total household income directly apart from the governmental child

Table 5.2: OLS and RIF regression coefficient on ln income, 1993

Eq Name	OLS	Quantile 10	Quantile 25	Quantile 50	Quantile 75	Quantile 90
Dep. Var:	ln income	RIFQT_10	RIFQT_25	RIFQT_50	RIFQT_75	RIFQT_90
Constant	10.857	10.027	10.409	11.145	11.707	11.986
	(0.003)**	(0.006)**	(0.004)**	(0.003)**	(0.003)**	(0.003)**
Female	-0.262	-0.190	-0.326	-0.308	-0.226	-0.120
	(0.001)**	(0.001)**	(0.001)**	(0.001)**	(0.001)**	(0.001)**
Age of Head	0.057	0.063	0.067	0.051	0.036	0.033
	(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.000)**
Age of Headˆ2	-0.001	-0.001	-0.001	-0.001	-0.000	-0.000
	(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.000)**
No. of children	-0.039	0.008	-0.020	-0.052	-0.066	-0.074
	(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.001)**
University/	0.481	0.287	0.405	0.500	0.437	0.475
College	(0.001)**	(0.001)**	(0.001)**	(0.001)**	(0.001)**	(0.002)**
High School	0.231	0.200	0.271	0.232	0.147	0.136
	(0.001)**	(0.001)**	(0.001)**	(0.001)**	(0.001)**	(0.001)**
Student	-0.362	-0.487	-0.398	-0.293	-0.153	-0.121
	(0.002)**	(0.003)**	(0.003)**	(0.002)**	(0.001)**	(0.001)**
Immigrant	-0.195	-0.261	-0.204	-0.145	-0.060	-0.006
	(0.001)**	(0.003)**	(0.002)**	(0.002)**	(0.001)**	(0.002)**
Immigrant	-0.026	-0.066	-0.050	-0.017	0.009	0.026
parent(s)	(0.002)**	(0.004)**	(0.003)**	(0.002)**	(0.002)**	(0.002)**
Eastern Norway	-0.063	-0.041	-0.067	-0.056	-0.050	-0.061
	(0.001)**	(0.002)**	(0.001)**	(0.001)**	(0.001)**	(0.001)**
Southern Norway	-0.092	-0.050	-0.080	-0.085	-0.089	-0.106
	(0.002)**	(0.003)	(0.002)**	(0.002)**	(0.002)**	(0.002)**
Western Norway	-0.049	-0.048	-0.067	-0.031	-0.024	-0.036
	(0.001)**	(0.002)**	(0.002)**	(0.001)**	(0.001)**	(0.001)**
Trøndelag	-0.11	-0.073	-0.109	-0.086	-0.083	-0.108
	(0.001)**	(0.002)**	(0.002)**	(0.002)**	(0.001)**	(0.001)**
Northern Norway	-0.074	-0.055	-0.087	-0.058	-0.058	-0.090
	(0.001)**	(0.002)**	(0.001)**	(0.002)**	(0.001)**	(0.002)**
Observations:	2 069 714	2 069 714	2 069 714	2 069 714	2 069 714	2 069 714
R-squared:	0.334	0.1381	0.285	0.294	0.199	0.106
F-statistic:	74088	13604	60769	95113	44666	14341

[a] Estimated coefficients calculated based on registry data on income and wealth from 1993. Income is defined as in Section 4, adjusted to 2015, her as ln income. The first column displays results from an OLS regression, the next five displays the results from a RIF regression for different quantiles. Standard errors in parentheses, $*p < 0.05$ and $**p < 0.01$.

Table 5.3: OLS and RIF regression coefficient on ln income, 2015

Eq Name:	OLS	Quantile 10	Quantile 25	Quantile 50	Quantile 75	Quantile 90
Dep. Var:	*ln income*	*RIFQT_10*	*RIFQT_25*	*RIFQT_50*	*RIFQT_75*	*RIFQT_90*
Constant	11.475	10.724	11.226	11.785	12.131	12.408
	(0.002)**	(0.005)**	(0.003)**	(0.003)**	(0.002)**	(0.003)**
Female	-0.340	-0.193	-0.357	-0.399	-0.337	-0.306
	(0.001)**	(0.001)**	(0.001)**	(0.001)**	(0.001)**	(0.001)**
Age of Head	0.047	0.047	0.045	0.038	0.038	0.037
	(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.000)**
Age of Head^2	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000
	(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.000)**
No. of children	-0.048	-0.035	-0.073	-0.061	-0.052	-0.058
	(0.000)**	(0.001)**	(0.000)**	(0.000)**	(0.000)**	(0.001)**
University/	0.384	0.258	0.384	0.417	0.338	0.342
College	(0.001)**	(0.001)**	(0.001)**	(0.001)**	(0.001)**	(0.001)**
High School	0.188	0.152	0.254	0.219	0.129	0.087
	(0.001)**	(0.001)**	(0.001)**	(0.001)**	(0.001)**	(0.001)**
Student	-0.391	-0.733	-0.395	-0.209	-0.106	-0.072
	(0.001)**	(0.003)**	(0.003)**	(0.001)**	(0.001)**	(0.001)**
Immigrant	-0.220	-0.239	-0.171	-0.174	-0.159	-0.137
	(0.001)**	(0.002)**	(0.001)**	(0.001)**	(0.001)**	(0.001)**
Immigrant	-0.044	-0.076	-0.051	-0.030	-0.002	0.021
parent(s)	(0.001)**	(0.002)**	(0.002)**	(0.002)**	(0.002)	(0.002)**
Eastern Norway	-0.038	0.001	-0.019	-0.030	-0.047	-0.072
	(0.001)**	(0.001)	(0.001)**	(0.001)**	(0.001)**	(0.001)**
Southern Norway	-0.059	-0.013	-0.039	-0.044	-0.065	-0.098
	(0.001)**	(0.002)**	(0.002)**	(0.002)**	(0.002)**	(0.002)**
Western Norway	0.005	0.025	0.007	0.014	0.007	-0.012
	(0.001)**	(0.002)**	(0.001)**	(0.001)**	(0.001)**	(0.001)**
Trøndelag	-0.074	-0.020	-0.044	-0.054	-0.074	-0.115
	(0.001)**	(0.002)**	(0.002)**	(0.001)**	(0.001)**	(0.001)**
Northern Norway	-0.048	0.020	-0.010	-0.029	-0.063	-0.112
	(0.001)**	(0.002)**	(0.002)**	(0.001)**	(0.001)**	(0.001)**
Observations:	2 761 643	2 761 643	2 761 643	2 761 643	2 761 643	2 761 643
R-squared:	0.343	0.199	0.261	0.293	0.206	0.104
F-statistic:	99999	23859	74509	1.3e+05	60911	19589

[a] Estimated coefficients calculated based on registry data on income and wealth from 2015. Income is defined as in Section 4, her as ln income. The first column displays results from an OLS regression, the next five displays the results from a RIF regression for different quantiles. Standard errors in parentheses, $*p < 0.05$ and $**p < 0.01$.

support included in the disposable income variable. The most interesting thing to notice about these results are that for the lowest part of the distribution, an increase in the number of children can have a positive effect on income. One possible reason for this is that low income families may get more financial support from the government than others, and for some of the households at the lowest percentiles financial aid might be the key income. This positive effect for the lowest percentile would not be picked up by a standard OLS regression. By comparing the results from 1993 to the results from 2015 one can see that the negative income effect of an additional child has increased, especially for the lower end of the income distribution. This may follow from child benefits not being increased since mid 90's, so with inflation that means that the child allowance is a smaller share of the parents income in 2015.

As one might expect, higher age is associated with higher income. Work experience often means higher wages. There is not that much variation in the returns to age across the distribution except for somewhat higher return at the lower end of the income distribution. This can be explained by more young individuals, students and individuals starting their careers, indicating that they are at the beginning of the age-earnings profile, as mentioned under IGIC. These results holds for both 1993 and 2015.

Not surprisingly more education is associated with higher income. The reference household has not passed high school and just passing high school has a positive effect on income, and university education results in an even larger positive effect. In 1993 the highest return to a university education is found at the middle of the distribution, where a university degree is associated with 50 percent higher income than not finishing high school. The return to education of a university degree compared to high school is as one might expect twice as large, except for at the 10^{th} percentile. Perhaps there are individuals with university degrees located at this point?

There is a lower expected return to education in 2015 than in 1993, and the year to year difference here is large compared to other coefficients. This change can likely be explained by the change in the number of individuals taking a higher education. The number of head of households with higher education has increased with 20 percent from 1993 to 2015.

Migration has a negative effect on income, either being an immigrant or having immigrant parents, and this effect is the same for both years. The effect of immigrant background is almost equal for both periods and follows the same pattern.

It is also quite stable across the distribution. As one might expect, there is a negative effect associated with an immigrant background or having immigrant parents, the effect of being an immigrant is larger than the effect of having immigrant parents. The effect is largest for the lowest end of the income distribution and then only increase. The number of immigrants increases from 1993 to 2015, but even thought this happens the the coefficient is relatively similar for both years.

In 1993 all of the regional variables are associated with a lower expected income than for households living in Oslo, which implies living in Oslo gives a higher income. The size of the effect does not vary a lot across the income distribution, but varies a lot across different regions with the largest difference for Trøndelag. The pattern is mostly similar for both time periods, however the negative effect is slightly lower across the entire distribution in 2015 than i 1993. This indicates that the difference between those living in Oslo and those who does not, has decreased from 1993 to 2015. The largest negative effect is related to living in Trøndelag and Northern Norway in both time periods. What is most interesting is to see that there is a positive effect associated with Western Norway in 2015, the only region with a larger expected income than Oslo. The positive effect is largest for the middle of the distribution, something that might be explained by the oil industry mainly being located in this region.

Ideally doing the same RIF regression for more years could give a better picture of the development in the returns to the different characteristics.

5.3 Decomposition results

The Oaxaca-Blinder decomposition[2] is used to study the differences between two groups, in this case the groups are represented by the two years 1993 and 2015. In the RIF results some of the differences across time were pointed out, and the decomposition is done on the basis of results from the RIF regressions. With the decomposition I should be able to identify which of the year to year differences can be explained by the composition effect and which can be explained by the structural effect, and by this explain why the GIC has the shape that it does, with high income earners experiencing the largest annual growth rate

The actual decomposition is done by calculating the Oaxaca-Blinder decomposition for the percentiles 1, 5 then every 5 until 95. The results are displayed in

[2]For the calculation of the Oaxaca-Blinder decomposition A STATA procedure developed by Jann (2008) is used

Figure 5.3, showing the overall change at each percentile, and the decomposed changes based on the composition and structural effects.

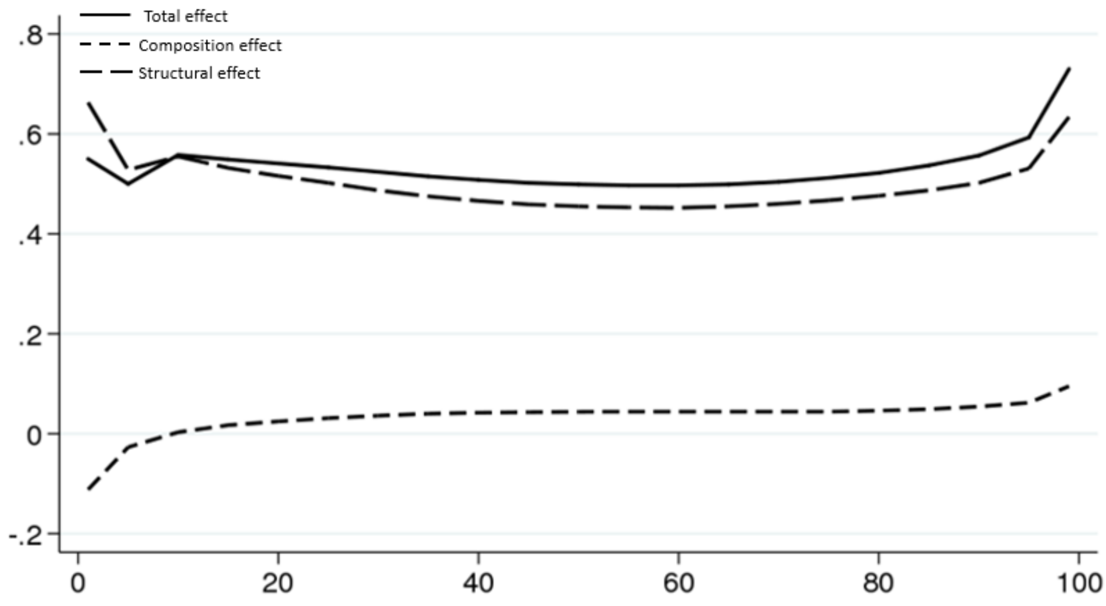

Figure 5.3: Decomposition into composition and structural effect

When studying the graph it is obvious that the structural effect dominates the entire distribution, with the composition effect having very little to no effect. This indicates that the effect of growth across the distribution is mainly affected by changes in the returns to the characteristics, which in turn might indicate that other factors are more important for the pattern of growth than the household characteristics investigated in this thesis.

Next, I want to investigate which factors that drives the composition and the structural effect. To make the decomposition a bit easier I have divided the effects into different groups: one accounting for household demographics: age, number of children and female; one group for the household as a student, this group consist only of the student variable; one for education consisting of the education variable; one for immigrant background accounting for both immigrant background and immigrant parents; and the last is the regional variable consisting of all areas. The resulting graphs are displayed in Figure 5.4.

The structural effect follows the total incidence quite closely and both of the graphs have quite a similar shape to the growth incidence curve. The results show that the demographic variable varies the most across the distribution indicating that different returns to the demographic household characteristics can vary across the population. The figure further reveals that the main contribution for the composition effect is the changes in the level of education.

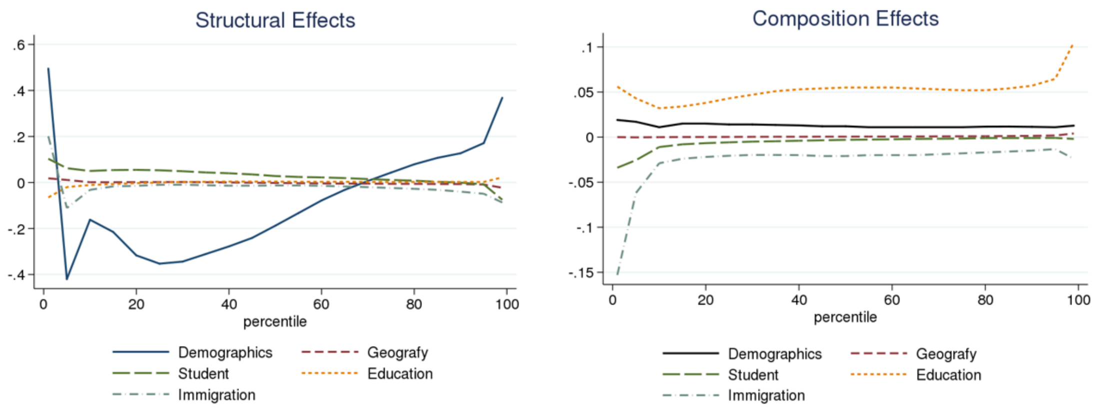

FIGURE 5.4: Detailed decomposition results

The main takeaway from the decomposition is that structural effects are the main contributions to the shape of the growth pattern, implying the returns to characteristics as being the main reason for the shape. This is something that is quite vague, indicating that there are possibly other underlying factors not included in this thesis that shape the growth pattern. Maybe the answer lies in going into the source of the income, the effect of sector of employment for example, working in the public vs. private, main source of income, capital income, changes in governments distribution programs etc.

From looking at the RIF results one can see that for most of the characteristics tested, there is not a large difference in the returns in 1993 and 2015, which might explain why the decomposition results yields so limited insight into what actually shapes the GIC. This points to other factors not included in this analysis as more important in determining the shape of the GIC, for example investigating the specific income sources for the households. As Piketty (2014) points out, individuals holding large amounts of capital might get a larger return on these than the growth rate in the economy. So even though the economic growth in Norway has been very large, the return on capital might have been even higher, and this might explain why the richest part of the population has experienced a higher annual growth rate than the rest of the distribution.

6. Conclusion

The aim of this thesis is to take a deeper look into the relation between economic growth, income distribution and income inequality, taking advantage of the steady growth experienced in Norway in the period 1993 to 2015. The goal was to explore how growth affects the income distribution and identify if there has been any "winners and losers" from the growth. Is income inequality related to income growth, in the sense that growth periods lead to higher income inequality. Further, I wanted to explore possible explanations to what is observed.

The first method I utilized was the growth incidence curve which indicated that all in all, large parts of the distribution saw a growth that was relatively fairly distributed with an almost equal growth rate. Despite of this somewhat equal growth, the top income earners experienced the highest growth rate throughout the period, on the expense of major parts of the distribution which had a growth below the mean growth. Since the top income earners experienced the highest growth in the period it indicates the the growth has lead to a small increase in income inequality. From the GIC alone, the conclusion is that the overall income inequality has increased slightly, but the individuals have not necessarily suffered.

However, this relation is completely reversed when examining the individual growth incidence curve. Here I held the initial distribution constant, and saw that the top income earners experienced low to negative growth, while at the same time the lowest percentile experience the highest growth rate. The IGIC points to high income mobility in Norway, meaning that individuals whom are poor in one period may not be so in the following period, shown by the lowest part of the distribution experiencing the highest growth rate in the IGIC.

To study which underlying factors might shape the income distribution, I looked at the effect of household characteristics. The RIF regression combined with the Oaxaca-Blinder decomposition showed that the structural effect dominate the entire distribution. Results from the RIF regression indicated that the returns varies a bit along the distribution, and this variation might explain the growth pattern. The results did not give a clear conclusion on what has shaped the distribution over this growth period, which points in the direction of other factors also being relevant such as source of income, like capital income, social benefits or labor income. Investigating this is beyond the scope of this thesis.

A. Growth Incidence Curve

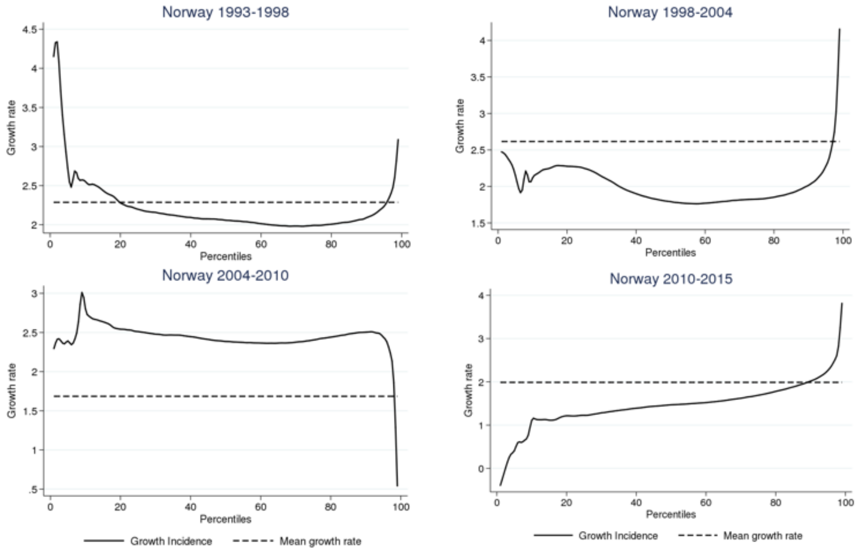

FIGURE A.1: Growth incidence curves for Norway for the periods, 1993-1998, 1998-2004, 2004-2010, 2010-2015.*Notes*:Estimates computed using cross-sectional data

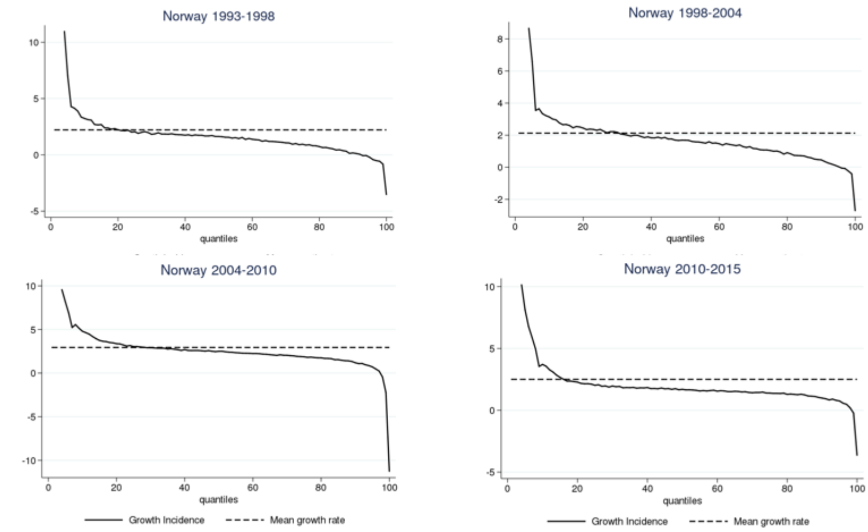

FIGURE A.2: Individual Growth incidence curves for Norway for the periods, 1993-1998, 1998-2004, 2004-2010, 2010-2015.*Notes*:Estimates computed using longitudinal data

B. Returns to Household Characteristics

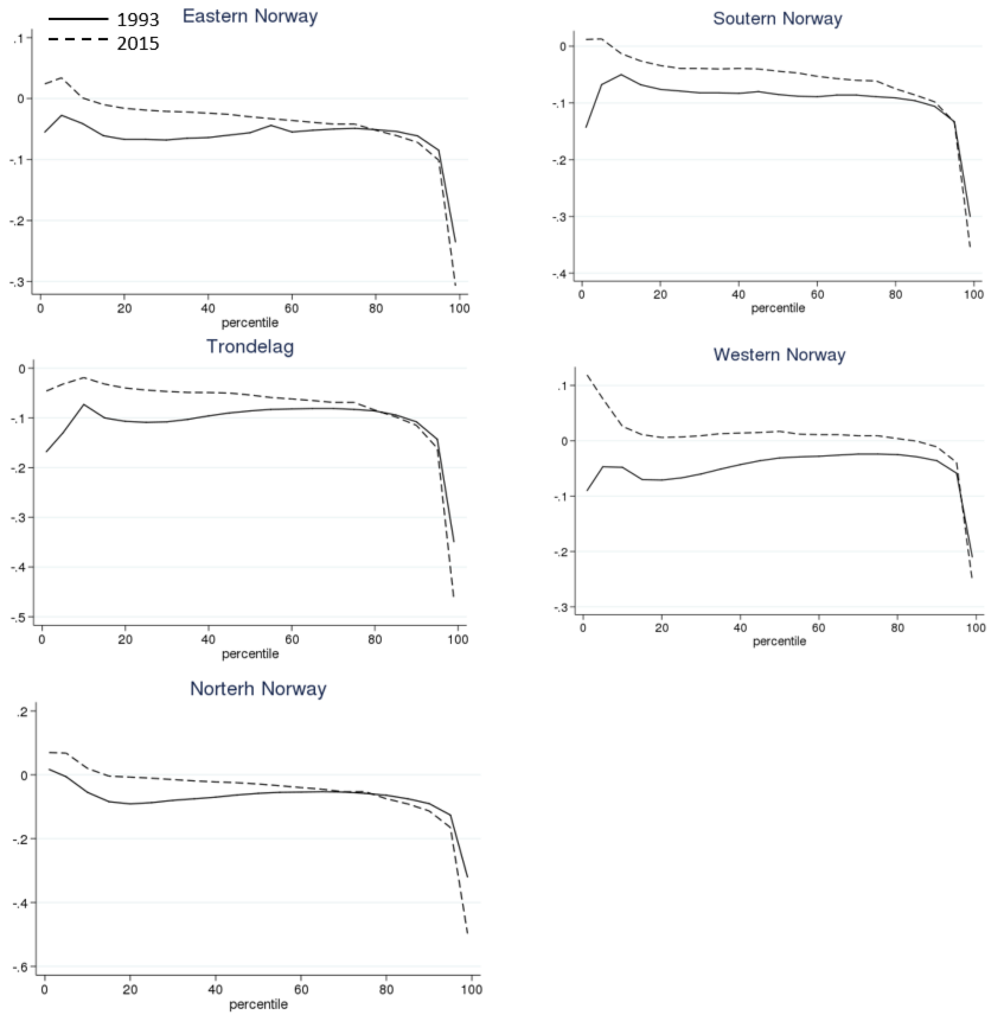

Figure B.1: Recentered influence function coefficient results for regional dummies

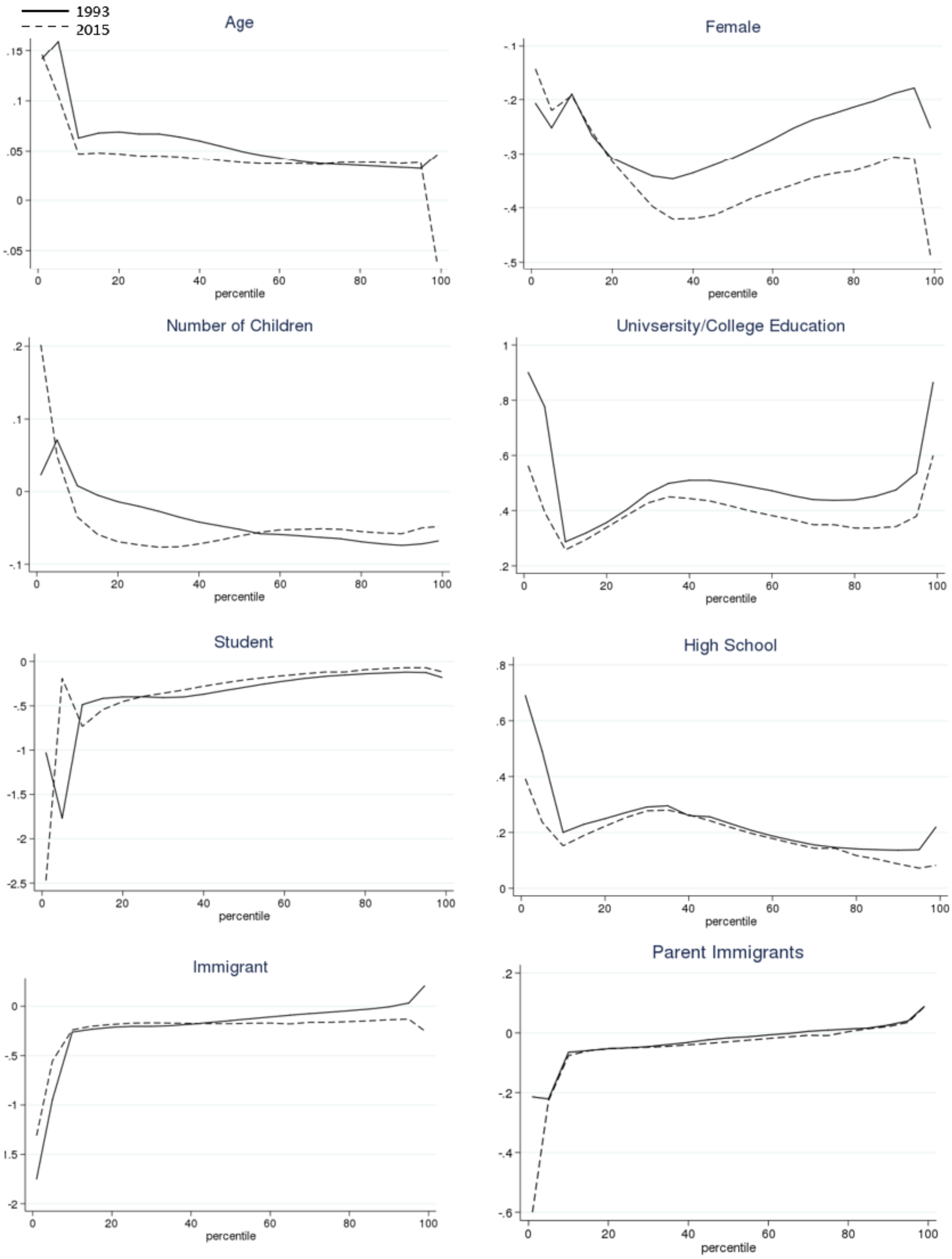

FIGURE B.2: Recentered influence function coefficient results

References

Aghion, P., Caroli, E., and Garcia-Penalosa, C. (1999). Inequality and economic growth: the perspective of the new growth theories. *Journal of Economic literature*, 37(4):1615–1660.

Azpitarte, F. (2014). Was pro-poor economic growth in australia for the income-poor? and for the multidimensionally-poor? *Social indicators research*, 117(3):871–905.

Bhuller, M., Mogstad, M., and Salvanes, K. G. (2017). Life-cycle earnings, education premiums, and internal rates of return. *Journal of Labor Economics*, 35(4):993–1030.

Blinder, A. S. (1973). Wage discrimination: reduced form and structural estimates. *Journal of Human resources*, pages 436–455.

Bourguignon, F. (2011). Non-anonymous growth incidence curves, income mobility and social welfare dominance. *The Journal of Economic Inequality*, 9(4):605–627.

Datt, G. and Ravallion, M. (1992). Growth and redistribution components of changes in poverty measures: A decomposition with applications to brazil and india in the 1980s. *Journal of development economics*, 38(2):275–295.

Deininger, K. and Squire, L. (1998). New ways of looking at old issues: inequality and growth. *Journal of development economics*, 57(2):259–287.

Dollar, D., Kleineberg, T., and Kraay, A. (2016). Growth still is good for the poor. *European Economic Review*, 81:68–85.

Dollar, D. and Kraay, A. (2002). Growth is good for the poor. *Journal of economic growth*, 7(3):195–225.

Essama-Nssah, B., Paul, S., and Bassolé, L. (2013). Accounting for heterogeneity in growth incidence in cameroon using recentered influence function regression. *Journal of African Economies*, 22(5):757–795.

Fields, G. S. (2007). How much should we care about changing income inequality in the course of economic growth? *Journal of Policy Modeling*, 29(4):577–585.

Firpo, S., Fortin, N. M., and Lemieux, T. (2009a). RIF-Regression STATA. http://faculty.arts.ubc.ca/nfortin/datahead.html.

Firpo, S., Fortin, N. M., and Lemieux, T. (2009b). Unconditional quantile regressions. *Econometrica*, 77(3):953–973.

Galor, O. and Zeira, J. (1993). Income distribution and macroeconomics. *The review of economic studies*, 60(1):35–52.

Grimm, M. (2007). Removing the anonymity axiom in assessing pro-poor growth. *The Journal of Economic Inequality*, 5(2):179–197.

Grosse, M., Harttgen, K., and Klasen, S. (2008). Measuring pro-poor growth in non-income dimensions. *World Development*, 36(6):1021–1047.

Jann, B. (2008). OAXACA: Stata module to compute the Blinder-Oaxaca decomposition. Statistical Software Components, Boston College Department of Economics.

Jenkins, S. P. and Van Kerm, P. (2011). Trends in individual income growth: measurement methods and british evidence.

Kakwani, N., Pernia, E. M., et al. (2000). What is pro-poor growth? *Asian development review*, 18(1):1–16.

Kuznets, S. (1955). Economic growth and income inequality. *The American economic review*, pages 1–28.

Lakner, C. and Milanovic, B. (2016). Global income distribution: From the fall of the berlin wall to the great recession. *The World Bank Economic Review*, 30(2):203–232.

Oaxaca, R. (1973). Male-female wage differentials in urban labor markets. *International economic review*, pages 693–709.

Okun, A. M. (1975). *Equality and efficiency: The big tradeoff.* Brookings Institution Press.

Ostry, M. J. D., Berg, M. A., and Tsangarides, M. C. G. (2014). *Redistribution, inequality, and growth.* International Monetary Fund.

Palma, J. G. (2011). Homogeneous middles vs. heterogeneous tails, and the end of the 'inverted-u': It's all about the share of the rich. *Development and Change*, 42(1):87–153.

Palmisano, F. (2018). Evaluating patterns of income growth when status matters: a robust approach. *Review of Income and Wealth*, 64(1):147–169.

Pen, J. (1974). *Income distribution.* Penguin (Non-Classics).

Piketty, T. (2014). Capital in the 21 century. *Trans. Arthur Goldhammer. Belknap Press.*

Piketty, T. and Saez, E. (2003). Income inequality in the united states, 1913–1998. *The Quarterly journal of economics*, 118(1):1–41.

Ravallion, M. and Chen, S. (2003). Measuring pro-poor growth. *Economics letters*, 78(1):93–99.

Saez, E. and Zucman, G. (2016). Wealth inequality in the united states since 1913: Evidence from capitalized income tax data. *The Quarterly Journal of Economics*, 131(2):519–578.

Statistic Norway (2017a). Fakta om norsk økonomi. https://www.ssb.no/nasjonalregnskap-og-konjunkturer/faktaside/norsk-okonomi. Accessed: 30.04.2018.

Statistic Norway (2017b). Income and wealth statistics for households. https://www.ssb.no/en/inntekt-og-forbruk/statistikker/ifhus. Accessed: 01.05.2018.

Van Kerm, P. (2009). Income mobility profiles. *Economics Letters*, 102(2):93–95.