The Evolution of Regional Income Inequality and Social Welfare in Norway: 1875-2015

Jarle Kvile

Master of Philosophy in Economics

Department of Economics
UNIVERSITY OF OSLO

December 2017

The evolution of Regional Income Inequality and Social Welfare in Norway: 1875-2015

Jarle Kvile

http://www.duo.uio.no/

Printed: Reprosentralen, University of Oslo

Abstract

This thesis describes the regional development of social welfare and income inequality in Norway from 1875 to 2015 with an urban/rural divide. It is a descriptive analysis that combines theory of economic development with estimates of inequality. Focusing on five measures; the Gini coefficient, the mean income, social welfare, poverty, and affluence.

Kuznets famous inverted U-hypothesis from 1955, describes a relationship between regional inequality and economic growth. The theory has been the basis for much research, and in many countries, regional income inequality has followed the inverted U-pattern. By contrast, Norway does not exhibit a inverted U-pattern. The regional inequality declined from 1875 to 1980, and increased after 1980.

Since 1875, Norway's economy has developed from a pre-industrial economy to a modernized economy with a significant welfare state. Social welfare and mean income have increased, while the poverty rate has declined.

This thesis demonstrates that the mean income, income inequality and social welfare have converged between regions from 1875 to 1980. After 1980, however, income inequality and the mean income have diverged, while social welfare has continued to converge.

Before the Second World War, the convergence in social welfare between regions was mainly driven by a decline in inequality above the median in the regions with relative high inequality. The upper tail inequality continued to decline until 1980. From the Second World War until 1980, the correlation between mean income and income inequality was negative, implying that the regions with high mean incomes, also had low income inequality. After 1980, the correlation becomes positive.The recent convergence of social welfare is solely driven by a shift in the correlation between the mean income and income inequality.

Acknowledgements

The thesis is written through my job as a research assistant in the Research Department at Statistics Norway. All the data material I have used is also provided by Statistics Norway.

First and foremost, I would like to express my deepest, warmest and most humble feelings to my main supervisor Rolf Aaberge. Under his guidance, I have learned more than I could possibly imagine before starting on this exciting journey. His engagement has been far beyond what one can expect from a supervisor. Constantly working hard and methodically precise on important issues have benefited me greatly, not only as a student, but also as a person. I will always hold this experience in high regards. In addition, I would like to thank my co-supervisor Jørgen Modalsli for always having an open door when discussing problems concerning missing data, as well as an introduction to the micro data. He has been an inspiration in every aspect of this process.

I would also like to express the greatest of gratitudes to Edda Solbakken, who, as a PHD student of Rolf Aaberge and Kalle Moene, has taken her time, patience and all of her knowledge to guide me trough the process of learning STATA. Edda has been more patient than anyone could ask of her. I cherish this time greatly. She has been a role-model in terms of the co-worker I want to be. I would consider whomever encounters her in the future to be lucky.

Finally, thank you to all of my friends and fellow students who have engaged in meaningful discussions about my topic and my process together with their own processes. This has been of great inspiration, both for valuable discussions and insights. Thank you for politely laughing at my silly jokes.

Any inaccuracies or errors in this thesis are my responsibility alone.

STATA do files, and files containing all estimates [1] are available upon request.

Jarle Kvile

December 2017

[1] Microsoft Excel or STATA format

Contents

List of Tables

List of Figures

1 Introduction

The focus on the historical development of income inequality has risen since the work of Atkinson and Piketty (2007), who studied changes in the income shares of the richest. Increases in income shares of the richest will generally lead to higher inequality, but the development of the income distribution will also depend on how incomes are distributed for the rest of the society. However, with a few exceptions such as Atkinson and Søgaard (2015) and Aaberge et al. (2017), little research has been devoted to studying the development of the entire distribution of income. While their work focus on the whole country, this thesis goes one step further and describes the entire income distribution in Norway from 1875-2015 by regions. Furthermore, by using estimates of regional inequality and mean income, the evolvement of social welfare is determined.

This thesis will quantify the developments of inequalities by regions and investigate to what extent demographical and other economic developments have had an effect on the developments of inequality across regions. This will provide a broader angle on the development of income inequality in Norway.

A significant difficulty faced by researchers who want to analyze the historical development of inequality is the lack of relevant data. The approach of Aaberge et al. (2017) provides a framework for estimating historical series of income inequality. They do not only implement tax data, but explores the possibilities of the measurement of income inequalities, which creates a more a solid base for their analysis on the development of income inequality. Following their approach enables us to estimate the distribution of income in the period between 1875-2015 for regional and rural/urban differences. This thesis examines the advantages and challenges of following such a framework when the availability of data is less consistent over time. Furthermore, illustrating how this methodology can be used to build on measures of social welfare and its evolution.

The purpose of the thesis is threefold.

(i) provide evidence of the development of regional inequality.
(ii) provide a new framework to estimate historical regional poverty.

(iii) provide evidence of the development of regional social welfare.

The thesis is structured in the following way. Chapter 2 surveys Norway's economic history from the beginning of the 19th century until today. It reviews the theory of regional income inequality. Chapter 3 describes the methodology of estimation, including a discussion of the strengths and weaknesses of the available data. Chapter 4 is dedicated to the urban/rural divide, where we are interested in similarities and differences between the development of urban and rural areas, compared to the national level. Chapter 5 is devoted to the historical development of regions in Norway. We compare estimates of inequality within a region, and compare regions to the national average, to determine if regions have converged regarding inequality, social welfare and mean income. In chapter 6 analyze the affluence measure, to provide more insight on the upper part of the distribution, and provide further justifications for changes in inequality.

2 Background Theory

This chapter gives a review of the background for this thesis. First, some theories of on the long run development of income is discussed. In particular, the inverted U hypothesis is described, which focuses on the relationship between regional income inequality and economic development. In addition, a review of the empirical literature concerning the relationship is included. Further, a survey of some of the literature on the long run development in Norway, and other studies on regional differences is presented. This chapter concludes with a presentation of the Norwegian economic history.

2.1 Theory of economic development

An important question regarding the development of income inequality is if there is a causal relationship between income inequality and economic growth. If there is such a relationship, how does this development evolve? Kuznets (1955) theory of the inverted U-curve is highly cited.

Kuznets describes this development as following. At first, in the early stages of economic development, the income inequality will increase. After a restructuring of the labor force due to changes in demand, the inequality will then fall. Kuznets also states that part of this development is because of migration from rural to urban areas.

Kuznets predicts the changes in inequality in the following way. At first, there is a small, wealthy group with the possibility to save and invest. Starting economic growth, followed by the starting-phase of industrialization. The rich earn more due to increased earnings on savings, which the workers do not yet benefit from. Higher capital accumulation leads to higher incomes, explaining the upward sloping of the curve. As industrialization increases, demand for labor increases as well, and therefore more people are employed in the sector. Since the industrial sector is more profitable than the primary sector, earnings and living standard for workers will rise. Allowing the working class to start education, and becoming a more substantial part of society, leading to inequality.

Figure 1 shows the relationship described as the Inverted U-curve.

Figure 1: An example of the Inverted U Curve

Kuznets (1955) also emphasizes the impact of urbanization on income inequality. Kuznets argues that both income per capita and income inequality is usually higher in cities than in rural areas. In a development process of industrialization and urbanization, income inequality will rise for two reasons. First, the share of the population living in urban areas increases, which is relatively more unequal income distribution of the two. Second, the relative difference in income per capita between the rural and the urban population might increase. The reason is that the productivity in urban industries rises faster than the agricultural productivity in rural areas. Eventually, income inequality will start to decline as low-income groups gain more influence over political decisions in the cities. The political pressure on governments to ensure redistribution and taxation of the richest will increase, driving down income inequality.

Standard neoclassical theory suggests that in a well–functioning economy regional inequalities should be eliminated through factor mobility, trade, or arbitrage. Wages will be equalized across all sectors. The article by Harris and Todaro (1970) describes a theoretical framework where this is not the case. Formal and informal work is considered, where formal work has

bounds or tariffs, while informal labor and agricultural work is flexible. The Harris-Todaro model predicts a view of a migration equilibrium with an excess supply of labor in the expected high-wage formal sector, spilling over to the informal urban sector or unemployment. Wages do not equalize across sectors, but rather the *expectation* of wages. The expectation of wages is derived from the probability of employment in the formal and the informal sector that equalizes. It follows that if the productivity in the agricultural sector relative to other sectors decreases, there will be a migration towards the urban areas.

Acemoglu and Robinson (2002) claims that the fall in inequality is not a consequence of economic development in itself, as a "law-life-relationship", but due to political changes influenced by a large mass of people. The share of people being allowed to vote has increased since the beginning of the 20th century, changing political power from the rich to the poor. Acemoglu and Robinson (2002) claims that this democratization has been the primary driver towards redistribution and lower levels of inequality.

Milanovic (2016) shows the overall trajectory of global inequality, before decomposing inequality into what happens within and between countries, using a long historical view. He then pulls it together and discusses how inequality is likely to evolve in the 21st century. He casts doubt on the Kuznets-curve and rather discuss Kuznets waves, arguing that the Kuznets-curve does not explain the rebound of inequality in much of the West since 1980. Milanovic describes this as more a long wave, with contending forces alternating over time. He identifies growth, differentiation and elite capture as drivers that increase inequality, while war, welfare, and progressive politics decrease it. On inequality between countries, he argues that the convergence between rich and developing nations is returning the world to levels seen around 1820. In 1820, the main source of inequality was due to class, rather than location. In other words, where you stood in the social pecking order of your country was more important than where you were born. The general finding from Milanovic is a quite stable inequality between countries.

2.2 Empirical review

This section gives an overview of the empirical literature of the development of income inequality. Economists do not agree on how to explain the development of income inequality over time nor the main drivers. To be better able to understand why, and to create a foundation for analysis,

it is important to understand why there are somewhat conflicting findings in the literature.

2.2.1 Empirical review on Income Inequality

Following Kuznets hypothesis, researchers have tried to explore whether there is such a relationship between industrialized countries. Williamson (1991) found that income inequality in England increased with Gini coefficients of 0,40 to 0,63 in 1877, decreasing to 0,44 in 1901. Williamson attributed this to the relative development of high-skilled and low-skilled workers. Higher capital intensive industry increases demand for high-skilled workers, and with a surplus of low-skilled workers, the difference in real wages increases. When human capital rises, more workers turn to high competence jobs, reducing inequalities.

Aaberge and Atkinson (2010) finds a development of the top income share consistent with the inverted U curve described by Kuznets. Atkinson and Piketty (2007) investigates the top income shares in the US and several European countries, examines the top income shares the US and in several European countries, using comparable income sources and methods for all nations, and relate their findings to the inverted U hypothesis. They explain that income inequality decreased between 1915 and 1948, but not because of the gradual, structural changes as proposed by Kuznets. Income inequality declined during a politically chaotic period; notably, it declined during the two world wars and in the early 1930s, during the Great Depression. The decline was entirely due to a fall in top capital incomes. The explanation for the decrease in income inequality is thus that capital owners incurred severe shocks to their capital holdings, as destruction, inflation, and bankruptcies, during this period. Confirmed by available wealth and estate data. But it seems like everything else, for instance, wage incomes, has been stable over the same period. It is also true, as proposed by Kuznets (1955), that the number of low wage workers in rural areas has decreased over this time, but only to be replaced by low wage workers in cities, meaning that this is not what caused income inequality to decline. Piketty also explains that the top capital incomes did not increase again after 1948 because of the introduction of progressive taxes.

Several other researchers have also tried to test the inverted U hypothesis.Thomas (1991) argues that there is more evidence supporting the declining part of the inverted U-curve, than the increasing part. Also Lindhert and Williamson (1985) show a decline in the top income

shares in Great Britain, USA, West-Germany, Preussen, The Netherlands, Sweden, and Denmark in the period from 1867 to 1979. One reason that the increasing part of the curve is more difficult to find evidence for is that there are no available income data that far back in time.

Other comprehensive studies, find little evidence of a Kuznets curve in the data. In analyzing their new data set on inequality, Deininger and Squire. (1996) find that in 88 spells of decade-long growth, inequality improved in 45 of them and worsened in 43 of them. In most of the cases, the changes in inequality were relatively small, and they were uncorrelated with initial income. As a result, according to their findings, the poorest fifth of the population saw their incomes improve in 77 of these 88 spells.

Gottschalk and Smeeding (2000) investigate the development in income inequality between 1980 and 1995. They find an increasing trend in income inequality over this period for most countries included in the study, also for Norway. Furthermore, they find that the Gini coefficient 1 increases by more than 2 per cent per year in the UK, and by 1 per cent per year in Sweden, the Netherlands, and Australia. In Japan, Taiwan, the US, Switzerland, France, Germany, and Norway it increases by 0,5 to 1 per cent per year. They find no change in Israel, Canada, Finland and Ireland, and only a modest decline in Italy.

Autor (2014) studies the effects of returns to skills, not to education, and how it matters for the changes in inequality for the US. He explores the role of the college premium and how it is a driver for inequality among the 'other 99 per cent'. He finds that rising inequality is mostly about skills, with a rising skill return after the 1980s. Firpo et al. (2011) claim that the change in 'college premium' can explain for 95 per cent of the increase in inequality in the 90-10 group.

2.2.2 Literature review on income development and inequality in Norway

To study economic changes over time long-run time series of data are needed. Concerning income inequality, the data will have to cover the whole income distribution. Such data is limited in most countries for the period before the Second World War. Also, the definitions of these measures change over time, making comparisons of figures over the whole period challenging. Still, some studies of the economic development over the last 150 years have been done, also for Norway.

Aaberge and Atkinson (2010) analyze the historical development of top income shares in the Norwegian adult population2 from 1875 to 2006. They use tabulations of the distribution of income as assessed for tax purposes before 1967, and micro-data from the tax register files available to Statistics Norway from 1967 to 2006. Their findings suggest that the income shares of the 10 per cent to 0,5 per cent of individuals with the highest income both rises and falls for short periods, but that the top income shares declined steadily from 1875 to 1948. At the same time, the Norwegian society developed from a pre-industrial to an industrial society, experiencing economic growth and a shift of the population from rural to urban areas. But as explained in Atkinson and Piketty (2007) the decrease could be caused by shocks to the capital owners capital holdings during the world wars and the Great Depression. From the post-war period until the late 1980s the top income shares continued to decline. During this period taxes gradually increased, and the Norwegian welfare state expanded.

From the early 1990s, the top income shares rose again, primarily because of an increase in the income shares of the top 1 per cent. Partly explained by the financial deregulation in 1984 and the tax reform in 1992 when taxes on capital income was reduced, making the distribution of capital income more uneven. This is similar to the findings for many other European countries in Atkinson and Piketty (2007). In addition, the Norwegian banking crisis ended in 1992, and there was a change in business cycles towards an economic upturn with lower unemployment and stronger growth. Also, there was a structural change from traditional manufacturing to services and technology during this period. A temporary tax reform on dividends in 2001 and a permanent dividend tax in 2006 gave an increase in the top income shares in 2000 and 2005, and the subsequent decline, however smaller than the jump, the year the tax was introduced. Creating some interpretation problems, but still, Aaberge and Atkinson (2010) show an increasing trend in the top income shares overall from the 1990s and onwards.

Solbu (2009) bases his analysis on the same data material as Aaberge and Atkinson (2010), but he investigates the whole income distribution for the population covered by the tax statistics from 1858 to 2006. He uses "Ginis nuclear family," three measures where one of them is the Gini coefficient. As in Aaberge and Atkinson (2010) the data material is based on tabulations of the income distribution in certain years before 1966, and microdata from tax register files is used from 1967 to 2006. Before 1966, he uses state tax statistics in some years, while he uses municipal tax statistics in other years to create a time-series over the whole period. In

1858 the figures were based only on cities in Norway. Because a more substantial part of the population, in general, was covered by the municipal tax statistics, and since the development in cities most likely differed from other areas in Norway, this gives some inconsistency to the results. In addition, there are only a few data points before 1966. His findings are in general similar to those for the top income shares from 1875 to 1945 and from 1990 to 2006, and are also in line with the results of Soltow (1965) before 1960.

To be better able to provide for interpretations of what is a fair income distribution, Almaas et al. (2014) studies the impact of what they call "unfair inequality", where they assume that some inequality is fair, because some are more productive, and make a higher effort than others. They then try to remove the effect of effort into the equation and measure the "unfair Gini coefficient" for Norway between 1985 to 2005. They find that even though including effort and skill into the equation lowers the Gini coefficient, there is still an overall increase in inequality.

Aaberge et al. (2017) describes in their seminal paper the development of inequality in Norway for the same period investigated in this thesis. They present a new methodology constructing income indices from tabulated income data, poverty statistics and population censuses. Their findings cast doubt on the idea that Norway was a egalitarian society in the nineteenth century. Income inequality did not exhibit a downward trend over time, rather episodic shifts in a series of episodes. Their methodology will be the framework for analysis, and their findings will be used as a reference of analysis in this thesis.

2.2.3 Literature review on regional differences

Soltow (1965) did an analysis of the income inequality of cities in Østfold and Vest-Agder from the mid-1800s until 1960 based on tax tabulations for each year. Soltow found that the income inequality has decreased for both regions. Soltow attributes the decrease in inequality to four different factors, decreased monopolies, equalization effects of «property income», «alleviation of economic misfortune» and increasing educational opportunities for the many. Soltow emphasizes these factors affect inequality, not necessarily average income or economic growth

Mjelve (1998) builds on Soltow's study, but includes newly available data from 1960 to 1990. Her findings suggest that income inequality has been both falling and converging over time in the eight selected cities. This can be explained by a more heterogeneous economic basis

in the earlier period. A large part of the economy was typically based on a few industries, and an economic shock affecting one of these would significantly affect both the income inequality and the economic development in a city. For the later decades, the economy is more broad and homogeneous, making the economy less prone to economic shocks striking one or two industries.

Lund (2012) studies the development in income inequality measured by the same three income inequality measures as Solbu (2009) in Norway from 1894 to 2010. Her estimation method is, however, different from that of Solbu (2009), before 1967. Lund (2012) takes advantage of both state tax statistics, municipal tax statistics and poverty statistics to estimate four income groups. These groups are used to estimate income inequality from 1894 to 1966. She also distinguishes between urban and rural areas. She finds that the income inequality in urban and rural areas develop somewhat differently in the period before 1967. In rural regions, income inequality follows an inverted U pattern as proposed by Kuznets (1955), while only the declining part of this curve is visible for urban areas. In other words, it seems like income inequality started falling earlier in urban areas. Kuznets (1955) argues that income inequality would increase at the beginning of a period of urbanization and industrialization.

The results of Lund (2012) shows the opposite, income inequality was lower in urban areas and started decreasing earlier in urban areas as well. From the 1960s until today urban and rural regions follow a similar pattern of declining income inequality until the early 1990s. The results in this period are in line with those of Mjelve (1998), who emphasizes the importance of the development of the welfare state, securing redistribution and more equal opportunities for everyone.

From 1991 to 2010 inequality rises, and it rises more in urban than in rural areas. However, the income inequality Lund (2012) finds evidence for rises less than the top income shares in the study of Aaberge and Atkinson (2010). This suggests that the increase in income inequality over the last decades is mostly due to an increase in the top income shares.

Strøm Fjære (2014) describes the development in 16 municipalities in relation to the inverted U-hypothesis. She focus on population growth, industry structure, mean income, poverty and income inequality. She finds that mean income has increased, poverty rates have declined. In addition, she finds a relation between economic growth and inequality to be more similar to an actual U than the inverted U found by Kuznets. She finds that since 1990, after a long period of convergence, income inequality and mean income have diverged between municipalities.

Modalsli (2017) find that urban incomes were 4,5 as high as rural incomes, with income inequality higher in high-income municipalities. Enflo and Rosès (2015) finds for Sweden that regional inequality declined from 1860 to 1980, but did not exhibit the Kuznets-curve pattern. They identify three important periods for the development of regional inequality in Sweden. Between 1860 to 1940, the unrestricted access of market forces such as expanding markets, high rates of immigration led to compression of regional income differentials. Between 1940 and 1980, the regions converged even more. The institutional arrangements favored the reduction of productivity differentials, and public policy makers aided the reallocation of the labor force from unproductive to productive regions and economic sectors. During the third period, from 1980 until the present, the regional incomes have diverged. The authors claimed that the development of the knowledge-intensive industries favored economic growth in the main urban areas. Felice (2011) finds that regions in Italy does show the pattern of the inverted U-function, as does Combes et al. (2011) for France, Martínez-Galarraga et al. (2013) for Spain and Badia-Miró et al. (2012) for Portugal.

(Rodriguez et al., 2010) studies the relationship between social mobility and income from a regional perspective, and finds a positive relationship between these two variables, and claims that greater social mobility makes greater inequality values more tolerable.

It is therefore attractive to investigate different regions in Norway and to observe whether the development is equal or different for the various regions, to examine whether it is possible to identify driving forces of regional inequality. Norway is quite similar to Sweden in many ways, but as we know, we participated in the Second World War, where Sweden was neutral and benefited economically from both wars.

2.3 Norwegian Economic history

The following section briefly describes the economic history in Norway from 1830 until today. This section will allow us to gain a better understanding of the developments discussed later in the thesis. This section follows the descriptions made by Hodne (1981), Hodne and Grytten (1992) and Bergh et al. (1983).

Before 1830, Norway was a traditional agricultural country. Almost 3/4 of the population

were employed in the primary sector. Norway was considered a developing country, with a significant share of the people living on the bare minimum. Urbanization was modest, and there were only a few large cities. From around 1830, the development started to change, with increased production in the agricultural sector. This allowed farmers to not only produce what they needed themselves but for trade as well.

In the period between 1835 and 1855, the industrialization began in Norway. The textile industry started to grow in production, helped with growth in the import of cotton. The industry developed in large because of imported technology and competence from neighboring countries. Productivity in the agricultural sector continued to grow, with a doubled production in grains and potatoes in the period between 1835 to 1855. In the fishery sector, the productivity more than doubled during this time.

After 1850, the growth started to move towards a different direction with the wooden industry increasing from 12 000 employees in 1850 to 48 000 employees in 1875. Fishery also continued to grow, helped by increased demand from Europe, but the main driver for the growth in the economy was the shipping sector, with Øst-Agder, in the South, with Arendal as Norway's shipping capital. The shipping sector increased fivefold in just 25 years, making Norway the third largest shipping nation in the world.

Between 1875 and 1905, the GDP grew only by 55 per cent, less than half the increase seen in neighboring Sweden and Germany. The lower growth rates were followed by increased emigration to North America since the 1880s. The construction of railroads, which had experienced a boom in the 1870s, stopped entirely between 1882 and 1892. However, a fall in prices allowed for new industries to arise within Norway. High economic growth characterized the 1890s, a substantial expansion in the wooden sector helped with improved technology to produce paper. Stavanger, a large city in South-West, increased its production of canned fish. The construction of railroads started again, which all lead to increased growth. The growth continued until the "Kristiania"-crash in 1899, leading to falls in property values and stagnation for several years. This growth was further stagnated by problems in the shipping industry, in the technological change from sails to steam.

From 1905 until 1914, the economic growth was high and stable, with GDP increasing by 4 per cent yearly. All sectors increased their productivity, and also the shipping industry started to grow, following a boom in the international economy. The main driver behind the economic

growth, however, was the introduction of the new hydro-power industry. This new technology, to turn water into electricity expanded massively. Norway had a comparative advantage in this industry because it has more accessible waterfalls than other European countries. This was important for the export sector, and together with growth in shipping, agriculture and industrial sector this period is characterized by a high, steady economic growth.

During the first world war, Norway as a neutral country, had good profit opportunities, due to increased demand for goods. Norway experienced an economic boom with a reallocation of the labor force, full capacity, increased wages, expansive monetary policy and an increase in the export sector. Following the boom, the share of the labor force also rose, and mean income increased.[1] The public sector started to grow, although being a small part of the economy at the time. At the end of the First World War, Norway experienced the downsides of the war, with the destruction of the trade-fleet and a lack of coal and food.

In the 1920s the depression started. The growth in GDP was steady in the 1920s in spite of increasing unemployment. Unemployment rose from 2 percent to 17 percent of union workers, which presumably had an adverse impact on inequality. The consensus explanation of why the increase in GDP was steady is that capital-intensive sectors such as shipping, hydro-power, and whaling were the sectors with economic growth. The industry sector experienced a 25 per cent real wage growth of in the period between 1920 and 1934, while the agricultural sector stagnated. The fall in GDP during the economic recession in 1929 was also relatively low, with 8 per cent, contra 31 per cent in the US.

During the 1930s, the productivity started to rise again. Increasing prices for exporting goods from 1937 was an important factor. A softer monetary policy with decreasing interest rates allowed for increased investment. An increase in employment followed increased production in the industry sector. The industry sector spread to districts in Norway, previously only with workers in the primary sector. Norway restructured the production, where the food and clothing industry expanded, with a higher productivity than before.

From 1940 to 1945, the Second World War Norway suffered a loss of human lives, a destruction of the trade fleet and lower supply of goods. This caused a recession, but not as

[1]The Inflation also increased, but as this was at the time derived from the "Levekårsundersøkelsen" made by Statistics Norway, it is possible that it did not reflect all of the regional differences. This was based on interviews conducted in Oslo, and may not capture all regional differences

hard as for other European countries involved in the Second World War. Already in 1947, GDP per capita was higher than in 1939, partly due to a significant demand for export goods from Norway. The Second World War also had other implications for the Norwegian economy. Norges Bank [2] was no longer independent and put under the Ministry of Finance. In the agricultural sector, a monopoly was established in 1941, that lasted until the 1990s.

The years following the second world war saw the largest economic growth in the history of Norway. Political agreement concerning rebuilding the country after the war allowed for economic stability. In the 1960s, wealth increased in Norway aided with structural changes shrinking the primary sector and expanding the secondary sector. The employment of the public sector started to grow more rapidly in the 1970s, and public services have had a steady growth since the second world war. The foundation of growth in the service industry was fueled by technological changes in the traditional industry sector. Increased earnings changed the pattern of demand in Norway, and in the 1990s the largest share of the labor force was in the service sector. The fraction of the population in the labor force also increased, helped with the introduction of women to the labor market, classifying labor done by women privately and out of the labor market, now to the public sector and other areas of the labor force. The main driver behind the economic boom in Norway has been the oil sector. The oil sector also fueled an increase in the shipping sector, followed by increased demand for oil in shipping.

Since the Second World War until today there has been a political motivation for a low-income inequality. The motivation has resulted in an increased welfare state since the 1960s. In addition to an expanded welfare state, the tax system is designed to have an equalizing effect on inequality. The number of students attending higher education has increased the last decades sharply motivated by increased unemployment and increased demand for high-skilled workers. In addition, the welfare state is designed to ensure that social mobility is high, meaning that choice of future career should be independent of parents career or income. The increased economic growth has enhanced the migration from rural to urban areas, from the North to the South of Norway. Cities with universities get a more significant share of the higher educated labor force relative to their population size, creating an environment for innovation and further growth. The percentage of the population living close to the coast has also increased massively since the beginning of the 1800s.

[2]The Norwegian Central Bank

3 Methodological Approach

This chapter covers the methodological approach in this thesis. The motivation of the thesis is to provide information on the development of regional income inequality and social welfare in Norway. Following the work of Strøm Fjære (2014); Lund (2012) and Aaberge et al. (2017), I will consider the entire population. First, the definition of regions is explained. Second, the measure of inequalility is introduced, followed by a description of the data avaliable for the analysis. Also, the estimation process of each variable, and the challenged when comparing data over a 140 year long period is examined. Further, the estimation of social welfare is discussed. Finally, the affluence measure, providing more information on those who possess income above median is explained.

3.1 About the definition of regions

The heterogeneity in the economic structure across regions is a basic feature of the Norwegian economy. Norway as a country is 1790 kilometers long [1]. The distance to neighboring Denmark and Sweden, relatively more developed economies in the 1800s and early 1900s were very different for the different regions in Norway.

The merging of counties to regions is presented in table 1. We want to use as much data available at the different periods of time, but at the same time have comparable data. [2] We still separate every region into rural and urban areas.

3.1.1 Changes in Norwegian municipalities

This thesis follows the extensive work of Juvkvam (1999) to estimate the changes between rural and urban municipalities. The underlying assumption here is that municipalities that are urban now are treated as urban throughout the series, and vice versa. When changes are made within

[1] From Lista in Rogaland to Kavringen near Vadsø in Finnmark

[2] Income tabulations for 1906 and 1910 were summarized into the regions used in this thesis

Table 1: From Counties to Regions

From County	To Region
Oslo, Akershus and Østfold	Oslofjord
Hedmark and Oppland	Oplandene
Buskerud	Buskerud
Vestfold	Vestfold
Telemark	Telemark
Vest- and Øst Agder	Agder
Hordaland and Rogaland	South-West
Sogn and Møre	North-West
Sør - and Nord Trøndelag	Trondelagene
Nordland, Troms and Finnmark	North

a region from one rural municipality to another rural, these changes are disregarded, as they do not change the unit of analysis.

The transfers between urban and rural municipalities do cause some problems when comparing data material over time, as for the period between 1910 and 1934, a lot of data on income is collected into "Rural Vestfold" for instance. When we transfer municipalities from rural areas to the cities, we lose the information in this period. In this case, we do not interpolate the income for the rural and urban areas, but rather estimate a ratio of how large the income of the transferred municipalities are relative to both the rural and the urban municipalities. We use this ratio to estimate the share of income that is overstated in for instance "Rural Østfold," and transfers this income to "Urban Østfold," before merging the counties into regions.

A detailed overview of the changes made to Norwegian municipalities to their changes in status from urban to rural or vice versa is found in Appendix C.

3.2 The Lorenz-curve and the Gini coefficient

When we want to compare the distribution of income over time, we want comparable measures that are independent of scale[3], satisfies anonymity[4], independent of population size and satisfy the Pigou-Dalton Principle of transfers[5]. Two estimates that follow these principles are the

[3]The measures are not affected by the size of the economy, which tend to change over time

[4]Does not matter who the rich or poor are

[5]In other words, a transfer of utility from the rich to the poor is desired, as long as it does not bring the rich to a poorer situation than the poor.

Lorenz curve and the Gini coefficient. The Lorenz curve is a graphical representation of the distribution of income or wealth. It was developed by Max O. Lorenz in 1905 for representing inequality of the income or wealth distribution. The Gini coefficient is a measure of statistical dispersion to describe the income distribution of a population. Corrado Gini developed it in 1912.

Let X be an income variable with cumulative distribution function F($\cdot$), and mean μ. Let $[0, \infty >$ be the domain of F where $F^{-1}(0) \equiv 0$. The Lorenz curve L($\cdot$) for F is defined by

$$L(u) = \frac{1}{\mu} \int_0^u F^{-1}(t)dt, 0 \leq u \leq 1 \tag{3.1}$$

where F^{-1} is the left inverse of F.

The Lorenz curve L(u) gives the share of total income that the poorest u proportion of the population possesses. Note that L(0) = 0 and L(1) = 1, since zero per cent of the population will always possess zero per cent of total income, and the entire population will always hold all of the total income. The linear line from the point (0,0) to (1,1) represents a perfectly equal income distribution. For any other distribution, the Lorenz curve will be an increasing convex function.

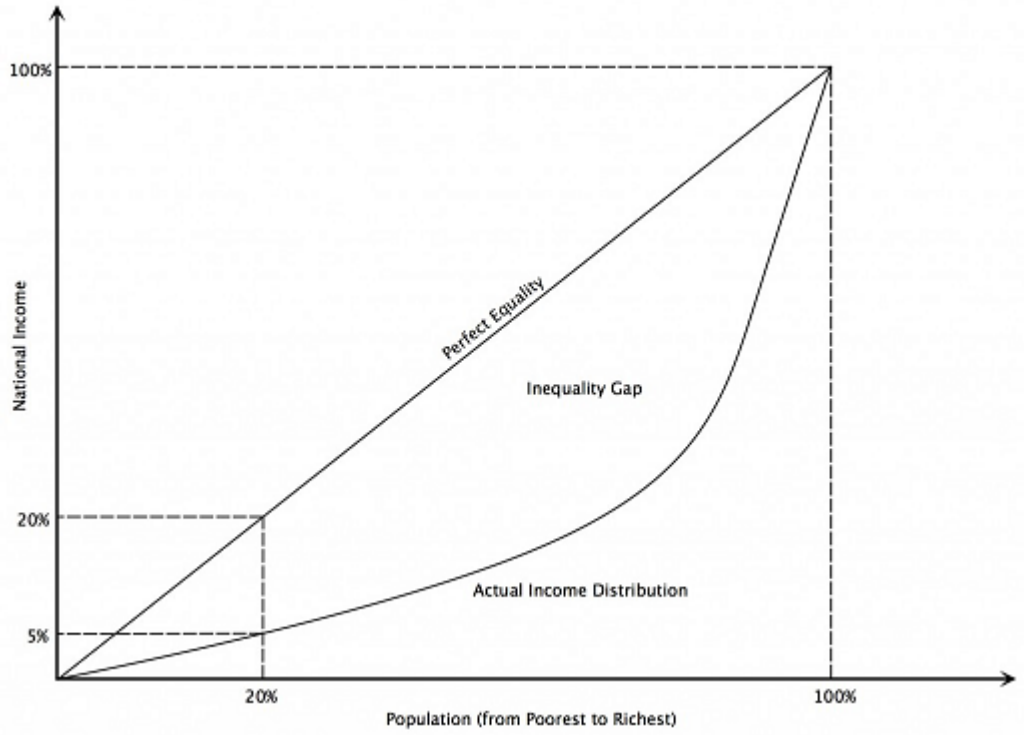
Following Atkinson (1970), Lorenz curves are commonly ranked by applying the criterion of Lorenz-dominance[6]. Which means that when Lorenz curves intersect, which is a common phenomenon, then the Lorenz-dominance criterion is not valid. Aaberge (2000) discusses two different solutions. One is to find a weaker ranking criterion, the other to apply summary measures of inequality. The most commonly used summary measure is the Gini coefficient. The Gini coefficient has the following relation to the Lorenz-curve.

$$G = 1 - 2 \int_0^1 L(u)du, \tag{3.2}$$

This means that the Gini-coefficient is 1 minus 2 times the area below the Lorenz curve, or equivalently twice the area between the Lorenz curve and the perfect equality line. The

[6]A Lorenz curve L_1 first-degree dominates a Lorenz curve L_2 iff $L_1(u) \geq L_2(u)$ for all u

Figure 2: Illustration of the Lorenz curve



Source: Intelligent Economist, 2017

Gini-coefficient will always be between 0 and 1 and is increasing in inequality.

Figure 2 is an example of a Lorenz curve where we have ideal data. The ideal data used for a precise historical analysis of income inequality would be based on data where we have data on all individuals in the society. This means that we would have all information on income levels for all individuals, including income in kind. We would like the data to include the entire adult population, giving a comprehensive representation of how income is distributed. The definition of income should also be consistent over time, allowing for an analysis free of disturbances such as changes in tax regulations, unidentified economic shocks, and measurement errors. If such data were available for the entire time period, we would be able to analyze all changes in inequality as actual changes. As an example, Aaberge and Atkinson (2010) find that the tax reforms on dividends gave an increase in the top income shares in 2000 and 2005, and the following decline the year the tax was introduced. These changes are not attributed to changes in inequality, but rather a disturbance in the data.

However, since individual data on incomes in Norway is only available from 1967, this thesis uses different sources of data to estimate the Gini coefficient, affluence, and social welfare. The

following section describes the data available, and how we can use this information to create Lorenz-curves and estimate Gini coefficients, even though the ideal data is not available.

3.3 Data on Income tax in Norway

This section is structured in the following way. First, we describe the available data on income tax. The next section is devoted to how we can use the data to estimate income groups, by using information on two different tax statistics. The municipality tax and the state tax. The threshold income required for paying municipality tax is lower than the threshold for paying state tax, implying that the municipality tax gives us information on a larger share of the population.

3.3.1 Access to data

First, when studying the long run trends of economic inequality, there are several issues. One of the main ones is access to data. Typical data used is some income survey, but the problems with this are that these types of surveys are quite recent. Meaning they do not give us a clear overview of the development over time nor cover the super-rich. Solutions to this issue are to use administrative time data by tax records. The quality and the historical period of these vary a lot between countries. Also, tax systems tend to change over time. What is taxed is not a constant over time, and the levels of tax avoidance and tax planning are important factors to consider. For instance, in 2005 in Norway, a new tax system was introduced that increased payments in 2004. These are important issues to consider when analyzing trends, and ups and downs in income inequality.

Few countries have data on income inequality providing information for the whole population covering a century or more in a continuous time series. The access to income data will be an essential predictor of the quality of the analysis. Historical data may be hard to access and analyze, because they are often missing, of poor quality, or only covers a fraction of the population. Before 1966 the data on taxpayers is extracted from the sources of municipal and state tax records[7], and information on the poor from poverty statistics. From 1967, I have

[7]Statistics Norway (2017b).

access to data on an individual level[8], with details of the income of every core household. The definition of core households follows the work of Aaberge et al. (2017). The unit of analysis is defined as a single person, couple or any dependent children above 16.

Lund (2012) and Strøm Fjære (2014) have used similar sources. I also have access to income tabulations used in the work of Mjelve (1998) and Solbu (2009). This thesis will use all available data, unlike previous work on a rural/urban divide. As this chapter will illustrate, this will provide a more precise estimate of the development of inequality.

Although tax income data is publicly available through Statistics Norway, they are not readily available for estimation and analysis. The gathering of data before 1967 has been a meticulous process of punching tax records for each municipality in every year where tax records are available. The same process for poverty statistics between 1920 to 1964. This, in itself, is a contribution to the literature of income inequality in Norway, because it allows for a broader range of analysis concerning the historical development of income inequality. The reason for not analyzing all municipalities in this thesis is a matter of priority. When constructing a study based on social welfare, affluence on regions, there was just not enough time to perform this analysis on a municipality level as well.

In those periods where the tax data is not available, such as in 1936 when there was no municipality tax, will be disregarded from the analysis. For periods between population censuses, I will use linear interpolation to estimate the changes between periods. When information is not available for the poor relief given to poor households, I will construct a ratio based on information on a national and urban/rural level to add poor relief for missing periods. The importance and practice are discussed further in section 3.3.2.

When analyzing income inequality, it is vital to be clear and precise concerning the definition of inequality. This thesis considers income inequality between core households. The income of a household is used because before 1959; married women were taxed together with their husbands. Children under 16 are not included in the analysis.

[8]Statistics Norway (2017c)

3.3.2 Estimating the poor

This section describes some of the difficulties that arise when estimating the poor, both concerning the size of the population, but also to the poor relief given to those who are poor. This section then moves forward and introduces a ratio that uses the relevant information available from historical statistics. This rate described in this section, allows us to include the development of poverty in the analysis.

The data covering the number of poor in the different regions are of varying quality. From 1875 to 1895, they are available for every year, and from 1895 until 1920, they are available every five years. [9] Between 1920 and 1935, there are no data available on the number of poor for the different regions.[10]. The thesis follows the following assumption. In this period, during the several crises in the Norwegian economy, described in section 2.3, the number of poor cannot be linearly interpolated, because doing so would not capture the changes in different regions in this period.

The income possessed by the poor in regions before 1935 is not available from the poverty statistics. The poverty statistics only covers urban and rural average income of the poor. There are multiple ways to proceed; I have evaluated four. The first is to assume that national averages are a representation of every district, but then we will we loose between district variation on how much income the poor possesses. A second approach is to create a ratio of total poor relief from one district in one year relative to total poor assistance in one year; we assume that the share of poor relief from each region is constant over time. In other words, every region experiences the same business cycles. A third approach is to assume that the poor relief given by a region is a constant fraction of total income. By believing this, we would lose the following mechanical effect; when the economy is booming, a higher share of the population is employed, and tax income rises. In this boom, there is a lower share of the population on poor relief, which would all else being equal, give a too high poor relief per person relative to the actual figures available on a national level.

A fourth alternative is to create a ratio in a base year where we have all available data. This ratio gives information on how much a district pays in poor relief relative to the national

[9] NSD Kommunedatabase (2017b).
[10] Statistics Norway (2017a)

average. This ratio can be multiplied with the data available on the national level for the years missing on the regional level. Say that z^*_{ij} is the total poor relief given to county i and $j = 1$ if urban, 0 if rural. '*' Denotes the base year 1935, where all information previously described is available. n^*_{ij} is the number of poor. This, combined with the ratio between poor relief and number of poor on an urban-rural level, allows us to calculate α_i. α_i is a ratio that allows for variation in poor relief between counties. We are therefore able to estimate a ratio that allows for variation across time in average poor relief.

Formally,

$$\alpha_i = (\frac{z^*_{ij}}{n^*_{ij}}) \cdot \left(\frac{\frac{z_{+j}}{n_{+j}}}{\frac{z^*_{+j}}{n^*_{+j}}} \right) \tag{3.3}$$

Where + denotes urban or rural. This ratio is then used to calculate the poor relief in the counties which have missing years.

$$\widehat{Z_{ij}} = n_{ij} \cdot \alpha_i \tag{3.4}$$

The assumptions for calculating the number of poor in each region are equivalent. I estimate a ratio between the number of poor, by the same method as in equation 3.5. Here we have better data available because the number of poor is available on a regional level 1920 and in 1935.

What about the 'small cities' [11] noted in the poverty statistics? These are assumed to be, small, and therefore are treated as rural within their region concerning the calculation of their ratio. Those small cities spend their money on poor relief, in relative terms, equal to neighboring rural districts, is a reasonable assumption. As Strøm Fjære (2014) argues in her thesis, there could have been a lot of work done over time to treat these as individual cities, but since a lot of these towns, such as Drøbak, are considered rural. The many changes in the structure of Norwegian municipalities would make this analysis very time-consuming, without giving any more consistency to the analysis.

[11] For certain years, cities regarded as 'small' is compiled into one group

3.3.3 Estimating Income groups

A key feature when comparing data over time is to find terms that fit the purpose of the analysis. A desirable feature is to find a measure of income that is comparable trough the time period. From the tax statistics, before 1967, we get information on tax units, those who are eligible to pay tax. However, not everyone paid income tax. The fraction of the population not paying income tax has varied over time. In 1895, this was 54 per cent of the rural population, and 42 per cent of the urban population. In 1955, in comparison, this was 85 and 79 per cent for rural and urban areas respectively. This section describes how to use the estimated income groups and Gini coefficients, including those who are not covered by the tax statistics. Following this discussion, section 3.4 elaborates on what implications these income groups have for the Lorenz curve and the Gini coefficient.

In the years before 1967[12], the thesis follows the work of Lund (2012) and Strøm Fjære (2014) to estimate three interior points of the Lorenz curve based on four income groups. I will also use the methodology from Aaberge et al. (2017) to estimate upper and lower bounds of the Gini coefficient.

In addition to the information provided by tax statistics, the poverty statistics contains data used to define the population that possesses the least income. This group is assumed to have such low income that they are not subject to taxation. This provides the first group of estimating the Lorenz curve.

In all years, the tax statistics, cover only a fraction of the total population. The second group is the part of the population not registered in neither the tax or the poverty statistics. Following the notation in Lund (2012), this population is called NAP [13].

The third group is the MUN. This category refers to the population covered in the municipality tax statistics but not the state tax statistics. The group is created by subtracting the aggregate income and taxpayers from each region in the state tax statistics from the corresponding group in the municipality tax statistics. The lower limits of taxation are generally lower in the municipality tax, meaning that the municipality tax includes a larger part of the population. The underlying assumption is here that those paying state tax are part of those who

[12]Where personal tax records are available
[13]NAP is defined as Non-Assisted Poor

pay municipality tax and therefore are the highest earners. The years where this assumption does not hold is disregarded in the analysis. When the aggregated income is higher in the state tax statistics than in municipality tax statistics, such as in 1913, it is pointless to assume they are all a part of MUN. These years are therefore ignored in the analysis.

The fourth group is the aggregated income and number of taxpayers registered in the state tax statistics. This group is called CG.[14]

3.3.4 Control totals

To arrive at an estimate of income inequality across the entire population, rather than only among the taxpayers, the tax data have to be combined with independent estimates of the total number of tax units. The first step is to estimate the total of the population above 16 years, and then to subtract married women, as they are taxed together with their spouses until the beginning of the 1960s. In the period between 1959 and 1967, we do not have sufficient data to control for married couples being taxed separately. From 1959 and onwards, married women could be taxed independently [15]. From 1967, we solve this using information on married couples, adding their income together, and treat them as one core household.

As for the population, the income tax tabulations do not cover the total income. To estimate a total income, Strøm Fjære (2014) follows the work of Gerdrup (1998) to determine the total income to be 25 per cent higher than the income of the municipality tax. This could lead to underestimation of the Gini coefficient because it adds more income to the MUN group than recorded in the tax statistics. Strøm Fjære (2014) corrects for these estimate by assuming that the poor had zero income, due to lack of data or estimation methodology available. The thesis would ideally continue to follow the work made by Aaberge and Atkinson (2010) and Aaberge et al. (2017) to estimate the total control income. However, as the regional data on the total income coverage is poor, it is not possible to use the national income as an estimate of total income for each region relative to total income. The thesis, therefore, assumes that the internal data from the tax records are sufficient to provide a comparable inequality measure. The problems occur when the CG becomes very large[16], and the average income, therefore, falls

[14] Notation by Lund (2012)

[15] Married women were allowed to either be taxed individually or together with their spouse.

[16] As the economy develops, a more significant share of the population is included in the state tax statistics, resulting in a large 'top earner group'

relative to the MUN group. It is therefore essential when providing meaningful estimates of inequality, to supplement within-group inequality. Further estimation of control totals could be done, but this would be very time consuming, and yield little return.

3.3.5 Using detailed tabulations of incomes

The main issue when using only four incomes groups is that we do not get information of the inequality within the different groups. As mentioned in section 3.3.1, regarding the access to data, we can use income tabulations where individuals are grouped into intervals, allowing us to estimate a Gini coefficient within taxpayers.

For certain years, the coverage of tax tabulations is a lot better than others. We have detailed tabulations in certain years on the number of people within an interval for both the MUN taxpayers and the CG taxpayers. The income tabulations are a lot better on the larger cities. [17]

Aaberge et al. (2017) have more information on the national income tabulations than the less aggregated rural-urban perspective. The income tabulations are the following, in a decreasing order of completeness of the period up to 1957:

Table 2: Income Tabulations available for Urban/Rural divide

Data	Year
MUN and CG distributional data:	1906, 1913 and 1929
MUN distributional data:	1875 and 1888
MUN distributional data:	1952-1957
CG distributional data:	1892-1903, 1938, 1948-1951
CG distributional data, individuals only	1952-1957

Whereas for the regional levels, no income tabulations are available between the period of 1930 to 1967. After 1967, we have access to personal tax statistics[18], which is close to the ideal data, described in section 3.2. It may seem as if I am 'grasping at straws' [19], but the method of estimating changes based on income tabulations gives information on how many individuals possessing different incomes. The information provided by the income tabulations are added to

[17] When the summation in the tables is available only for the largest cities, with medium and small cities summarized in the respective two groups, we use the information we have in previous years, together with information on the average income, to estimate changes for cities noted as 'other large cities' or 'small cities'.

[18] Statistics Norway (2017c)

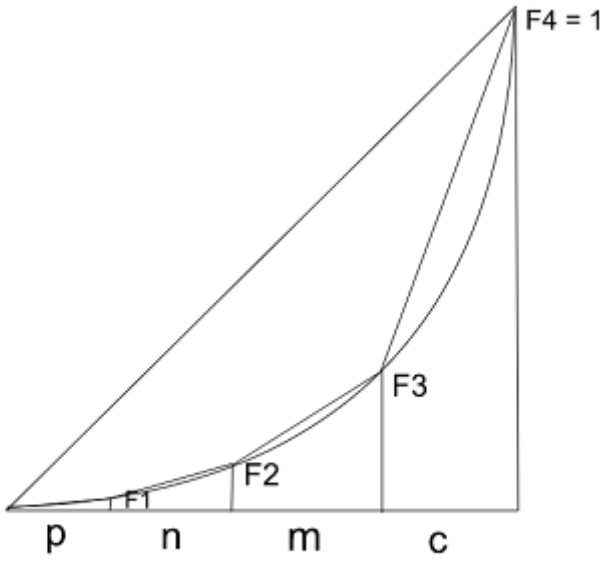
[19] Aaberge et al. (2017)

the data on mean income within groups, which is available for almost every year. This allows for a curved Lorenz-curve, rather than straight lines, which obviously adds more information than the work done by Lund (2012). In other words, we use the aggregation of many straws, or sources of information and together, they combine to a solid foundation for analysis.

3.4 Implications for the Lorenz curve and the Gini coefficient

The implications for the Gini coefficient with four income groups relative to three [20] are most easily seen regarding the area under the Lorenz curve since the Gini is equal to 1 minus twice the area under the Lorenz curve. The introduction of the missing population as seen in figure 3 has two effects. It squeezes the Lorenz curve for the taxpayers, MUN and CG, to the right. Simultaneously, it adds additional area under the first segment. In other words, it increases the Gini coefficient relative to the Gini for taxpayers alone.

Figure 3: The Lorenz Curve with four income groups, 1875 - 1967



This measure of inequality based merely on the four income groups is inaccurate, as seen in the figure. It will underestimate income inequality as inequality within-group is not considered,

[20] Disregarding the 'missing population' (NAP)

illustrated by the distances from the straight lines to the curves below. Using the information on within-group inequality will give a more convex shape of the Lorenz curve, providing a larger area between the curve and the line of perfect equality, which increases the Gini Coefficient.

The thesis extends the work of Strøm Fjære (2014) providing a more accurate measurement of inequality. This is achieved by including income groups available for years before 1938, and microdata from 1968. When estimating within-group distribution for the taxpayers for certain years, the thesis creates a Pareto distribution to be, following the approaches of Aaberge et al. (2017) to be constant within periods of time. This may be a strong assumption, but this is part of the motivation behind the upper and lower bounds. To use all the information available from the tax records, we can build knowledge upon a new framework.

Strøm Fjære (2014) estimated inequality between the four income groups using the following formula

$$G = \frac{1}{\mu} \int_0^1 L(u)du = 1 - [pF_1 + n(F_1 + F_2) + m(F_2 + F_3) + c(F_3 + 1)] \tag{3.5}$$

where G is the Gini coefficient.

The points in Equation (3.5) corresponding to Figure 3 are the cumulative population shares and cumulative income shares of the four groups.

p: The share of population registered as poor

n: The share of population of NAP[21]

m: The share of population paying municipal tax, not state tax

c: The share of population paying both municipal tax and state tax

F_1: The share of income to the poor

F_{2u}: The share of income to NAP and poor, NAP assumed to be 33 per cent of MUN income.

F_{2l}: The share of income to NAP and poor, NAP assumed to be equal to poor.

[21]Defined as Non-Assisted Poor

F_3: The share of income to the population paying municipal tax, but not state tax.[22]

F_4: The share of income to the top earners. Equal to 1. [23]

When the tabulations are available, the estimation of inequality will follow the work of Aaberge et al. (2017) use the following formula to estimate the Gini coefficient.

$$G = 1 - (pF_1 + (F_2 + F_1)n + F_2(1 - p - n) + (1 - f2) * (1 - p - n)) + [(1 - F_2)(1 - p - n)G^*] \tag{3.6}$$

where G^* is the Gini coefficient among taxpayers.

What we observe from equation (3.6) is that, relative to equation (3.5), the extra information from income tabulations, increases convexity of the Lorenz curve, and increases the Gini Coefficient. This can be seen in Figure 3, where we now use the curves under the straight lines for taxpayers, in m and c, respectively. This implies that comparisons with Lund (2012) are not relevant, because we use different sources for estimation of the Gini coefficient.

3.4.1 Estimating the changes in population

As previously mentioned, the population is measured in core households. The choice to analyze core households reflect the constraints given in the tax data. In the historical tax data [24], married couples are taxed as one unit. The information on population is collected from population censuses [25]. As there were no population censuses every year before 1950, the thesis follows the approach of Strøm Fjære (2014) and Aaberge et al. (2017) to linearly interpolate the changes in population, meaning that the population development between these points cannot be covered in the analysis. After 1950, we have population censuses available for every year.

[22] MUN group, defined in section 3.3.3
[23] CG group, defined in section 3.3.3
[24] Statistics Norway (2017b)
[25] NSD Kommunedatabase (2017a)

3.4.2 Creating Upper and lower bounds on the Gini Coefficient

We have, until this point, defined income and fraction of the population eligible for state tax, municipal tax but not state tax, and poor relief. In this section, we determine the income possessed by those not covered by tax statistics or poverty statistics, the NAP group. Furthermore, this section describes how we use information from tabulations of income groups between periods where they are not available [26].

The incomplete coverage of the population in the tax data means that there are challenges when estimating overall income inequality. The thesis follows the approach of Aaberge et al. (2017) when determining the income possessed by the missing population group, denoted as the NAP.

These bounds are no bounds in the mathematical sense, instead bounds based on judgments. The authors make two assumptions to create upper and lower bounds. They assume that the NAP group earns the same amount as the poor, creating an upper bound on inequality. The lower bound is based on the assumption that the NAP group has the average income of one-third of the MUN group. Higher compared to Gerdrup (1998), where he estimated that the income of the NAP group was at 25 per cent of the income of the MUN group.

In addition, from 1967 and onwards, the upper bound of the Gini coefficient is on the group receiving 50 per cent of the minimum pension for a single person. The lower bound is based on those not covered by the tax tabulations receiving mean income equal to 150 per cent of the mean income assumed for the upper bound. [27]

Furthermore, between periods where income tabulations are not available, we use the information from the two closest reference points in time. As an example, between 1938 and 1948 we do not have income tabulations for the rural/urban divide. For the lower bound estimate, we use the lowest of the two within taxpayers Gini coefficient and assume that the inequality between taxpayers is constant between the two periods. For the upper bound case, we use the highest of the two estimates. When estimating upper and lower bounds for the regions, we expect to see larger differences between the upper and lower bound of the Gini coefficient

[26] See table 2 for more information.
[27] This assumptions follows the work of Aaberge et al. (2017)

between 1930 and 1967, compared to before and after. [28]

Also, it may be argued that from 1967 and onwards, that the personal tax statistics understates the income, equivalently for all individuals. Since pension payments are subject to taxation, we should therefore not transfer income to the population with zero income in the tax statistics. Figure 17 and 18 in Appendix E estimates for the Gini coefficient, by following this argument. However, as we believe that no one can live on zero income, we do not find this too relevant.

3.4.3 Estimating the Average Income

To get a measure of the income development over time that provides information on the development of welfare, figures on the income relative to the population is needed. Two distinct options are available. One is to measure the average income relative to the taxpayers available, and the other is to measure the average income relative to the adult population. The obvious argument against the first alternative is that over time, this population and tax regulations change over time, which gives an inconsistent time series as well as overestimating the mean income.

The second alternative is then to measure income per core household [29] from 1875-2015. This will give a more consistent measure of the income development, as income will be measured relative to the population over the same period.

We have now created a foundation which includes the estimation of the Gini coefficient and the average/mean income. The next section is committed to how we can use these estimates, move forward and estimate social welfare.

[28] As a reminder, we do not have income tabulations during this period. Within taxpayers, inequality has decreased on a national level in this period. The upper bound use constant estimates of within taxpayers Gini from 1930 up until 1966, while the lower bound use the estimate from 1967, and assumes this is constant between 1931-1967. (see (Aaberge et al., 2017) for further justification)

[29] Measured in core households

3.5 Estimating Social Welfare

Traditionally, economics has treated efficiency and equity as separable. The theoretical basis for their separation is the Second Fundamental Theorem of Welfare Economics[30], which holds that any Pareto efficient outcome can be implemented as a competitive equilibrium given the appropriate lump-sum taxes and transfers.

This thesis follows the work made started by Sen (1974) ranking of welfare functions that represents a trade-off between efficiency and equality. The general family of rank-dependent measures of social welfare introduced by Yaari (1988) is defined by

$$W_p(F) = \int_0^1 P'(t)F^{-1}(t)dt, t \in (0,1), P(0) = P'(1) = 0, P(1) = 1 \tag{3.7}$$

Since $W_p(F) \leq \mu$ can be interpreted as the equally distributed equivalent income (EDE) [31] , Yaari (1988) defines the dual family of inequality measures as

$$J_p(F) = 1 - \frac{W_p(F)}{\mu} \tag{3.8}$$

Where $\mu = EX = \int xdF(X)$. The product $\mu J_p(F)$ is a measure of loss in the social welfare due to inequality in the distribution F.

The weighting function of P' is the derivative of a planners preference function. Ebert (1987) and Aaberge (2001) provides a normative justification of the family of welfare functions W_p where equation (3.9) shows the trade-off directly between the mean and the inequality in the distribution of income.

$$W_p(F) = \mu[1 - J_p(F)] \tag{3.9}$$

[30]We use pre-tax data in this thesis. This theorem is still valid to make the argument for estimating social welfare. Also, we do consider transfers such as poor relief.

[31] Defined as follows by Atkinson (1970) EDE is that level of income that, if obtained by every individual in the income distribution, would enable the society to reach the same level of welfare as actual incomes. If the marginal utility of income falls with income, then EDE is less than the mean level of income, as long as $G > 0$.

To be able to quantify social welfare, it is required to specify the weight function. By choosing $P'(t) = 2 - 2t$ J_p is equal to Gini-coefficient, and $W_p(F)$ represents the Gini social welfare function, see Sen (1979).

Donaldson and Weymark (1980) showed that if we assume that the weight function P(t) is defined by

$$P_{1k}(t) = 1 - (1-t)^{k-1}, k > 2 \tag{3.10}$$

Then J_p becomes equal to the extended Gini family of social welfare.

$$W_G = (1-G)\mu \tag{3.11}$$

The welfare function shows that the larger the inequality, the lower is the welfare, and the lower the mean income, the lower is the welfare.

To further estimate the evolution of social welfare, this measure could be applied to other estimates of inequality than the Gini coefficient. As Aaberge (2007) state, the Lorenz-curve may give a 'family' of inequality measures. This is further discussed in Appendix D, but by applying the Bonferroni coefficient [32] or the C_3 measure [33] to estimate social welfare, we could be able to provide more insight on its development. Compared to the Gini[34], the Bonferroni gives higher estimates and C_3 provides lower. Implying, if we are more concerned with the lower part of the distribution, our estimates for Social Welfare would be lower, and higher for the upper part. As this is interesting from a social planners perspective, this could be a basis for further research.

3.6 Affluence

Before we judge a measurement of inequality, we need to ensure that it satisfies the necessary criteria such as dominance and anonymity. Why do we now depart from the characterizations of the Lorenz-curve that evaluates the entire population, and now use the median as a point of

[32]Weight changes in the lower part of the distribution relatively more than the rest.
[33]Most concerned with changes at the top
[34]Gini C_2 ,Bonferroni C_1

reference? With the rise of inequality at the top in many countries, the mean has been a less adequate measure of the overall progress, which turns the attention to the median. As pointed out in Stiglitz et al. (2009), 'median income' provides a better measure of the 'typical' household than average consumption. In the literature of the 'middle class,' the group is typically defined regarding a range around the median income.

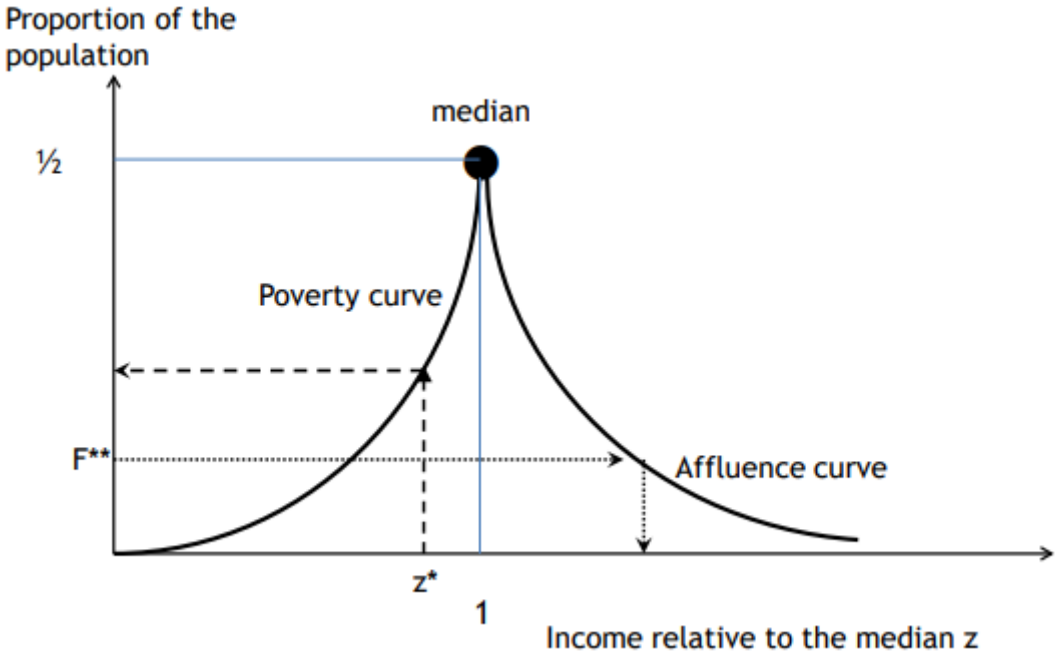
As the proportion of taxpayers comprise between 40 and 83 per cent of the annual tax unit population in regions such as urban Oslofjord before the Second World War, the information basis is better for the upper tail than the lower tail of the income distribution. Thus, it will be useful to complement estimates of the Gini coefficient for the entire income distribution with estimates of affluence and the upper tail Gini coefficient. This relies on the following affluence measure introduced by Aaberge et al. (2015). The affluence measure is defined by

$$A = \frac{1}{3}\left(\frac{\mu_u}{\mu}G_u + \frac{\mu_u - \mu}{\mu}\right) \tag{3.12}$$

where μ is the overall mean income, and μ_u and G_u are the mean and the Gini coefficient of the conditional distribution of income given that the income is larger than the mean. Definition (3.12) demonstrates that the measure of affluence, A, increases when either the income gap increases or the inequality in the distribution of incomes among the wealthiest 50 per cent increases.

To provide further insight into this measure, Figure 4 is provided and shows the relation between the well-known poverty curve and the affluence curve.

Figure 4: The relation between Poverty and affluence curves



Source: Aaberge et al. (2015)

Reading note: The left hand ('poverty') curve shows the proportion of the population with income less than or equal to a poverty defined relative to the median; the right hand ('affluence') curve shows the proportion of the population with income equal to or above an affluence threshold defined relative to the median.

4 The development between rural and urban Norway

In the chapter the development in the urban/rural divide is discussed. The chapter follows the structure from chapter 3, and starts by estimating the population development. Further, the poverty rate is provided. Here, we are able to provide further insight into the development of the poverty rate due to the ratio of poor, as discussed in section 3.3.2. We move forward and present the development of income inequality and the mean income. These measures of inequality and economic development are then used to provide estimates on the development of social welfare. The chapter concludes by considering estimates of the mean income and social welfare relative to the national average, and make comparisons with economic theory and literature.

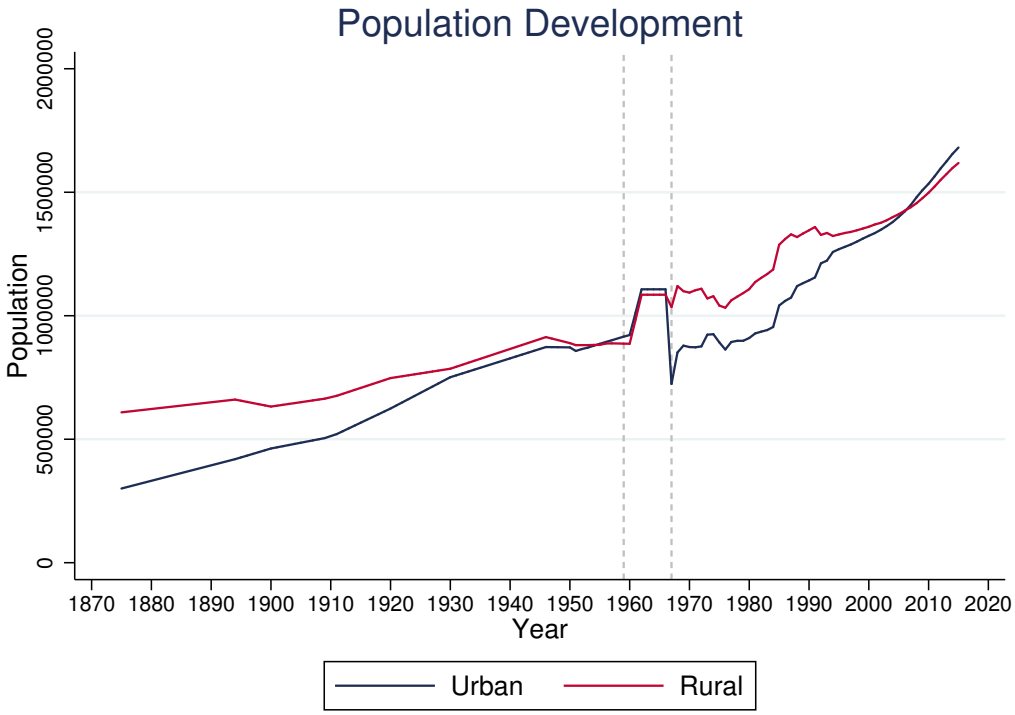
4.1 Population Development, core households, 1875 - present

Kuznets (1955) theory implies that as a nation undergoes industrialization, the center of the nation's economy will shift to the cities. He predicts that farmers, looking for better-paying jobs in the cities, will migrate from rural to urban areas. This chapter, therefore, begins to describe the development of the population, to investigate whether this theory holds for Norway.

There are some similarities between the data and the theory by Kuznets. We observe that the population rose by 256 per cent for the rural areas, and 559 per cent of the urban population, as predicted by the Kuznets hypothesis. During the first phase of industrialization, with the subsequent economic development, the population migrates towards the urban areas.

In the first part of the period, according to Hodne (1981), the population growth was driven by decreasing mortality rates due to better living standards, combined with high fertility rates. The population increase has later slowed down, due to lower fertility rates and increased female participation. Also, the increased share of the labor force has risen sharply since 1980, especially in the urban areas. Sobye (2012) argues that immigration is one of the main drivers behind the population growth in the last decades.

Figure 5: Population 1875 - present. Measured in Core Households



Reading notes: 1. Population, aged above 16, is measured in core households. Meaning that married couples are counted as one. 2. In the year 1959 taxing rules changes, which allowed for married women to be taxed separately from their husbands. The changes in the tax units should, therefore, be interpreted with caution for this period, marked with the dotted lines.

4.1.1 Development of the Poverty rate 1875-1967

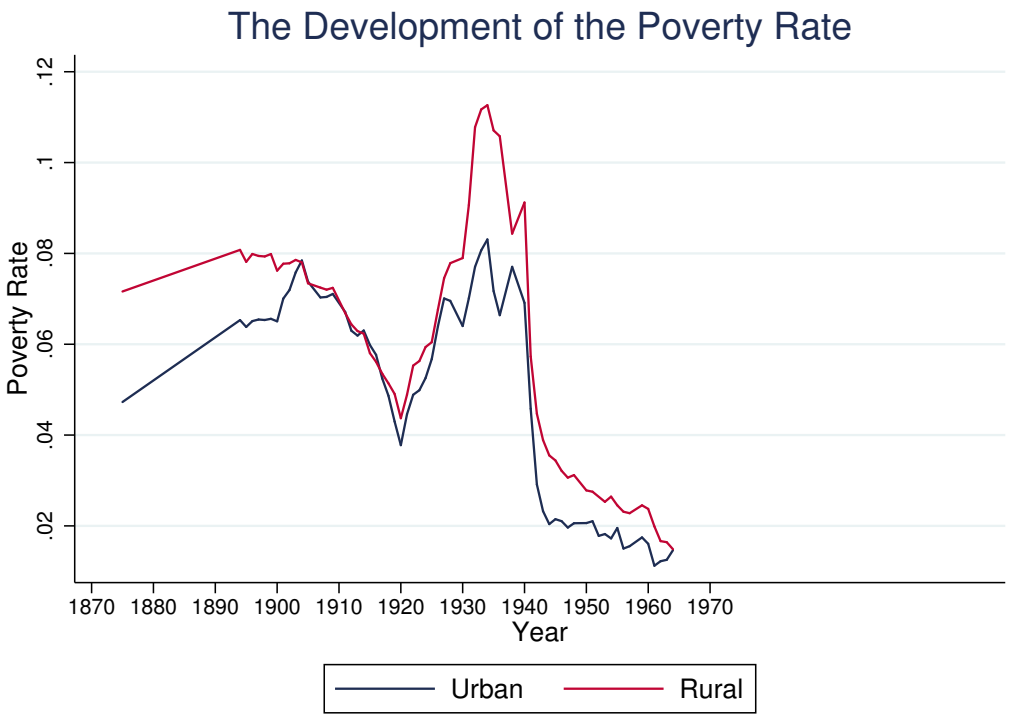
This section describes the development of the poverty rate from 1875 to 1967.

The lives of the poor have been of great concern to economists for centuries, and a significant focus concerning the development of inequality by economists such as Banerjee and Duflo (2007), Stiglitz (2012) and Milanovic (2016). How the level of the poverty rate has developed has a high impact on the Gini coefficient, but it is also a measure of inequality in itself. Figure 6 describes the development of the poverty rate in Norway before 1964 [1].

For both the urban and rural areas we observe a similar pattern. The poverty rate declined from We see a decline in the poverty rate from 1905 until the end First World War. As discussed in chapter 2, this was a time of significant economic growth in Norway. The poverty rate again rose during the Great Depression until the just before the Second World War. Following the crisis in 1929, the poverty rate again increased. The increase in the poverty rate may have been

[1] When the poverty statistics ends. After 1967, we use information provided by 'Folketrygden'.

Figure 6: The Poverty Rate 1875-1964



driven by increased unemployment, and the fishery crisis that lasted until 1938. [2] During the Second World War, it fell sharply and remained low and stable until 1967.

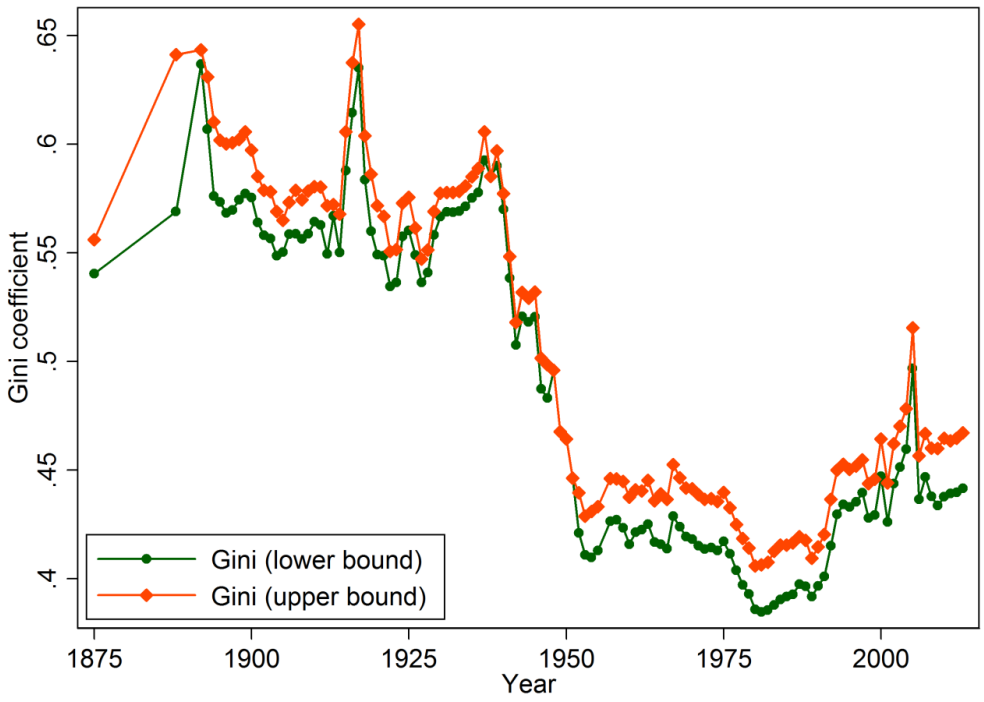
4.2 Inequality development over time

This section covers the development of inequality over time. When evaluating the development of income inequality over time, we stress that it is important to consider the upper and lower bounds. This importance stems from the fact that, we just do not know how the NAP group was treated. In order to move forward, we make estimates based on logic.

Figure 7 displays the Gini coefficient with upper and lower bounds for the entire country from 1875 to 2013. Note for further comparison, that this scale is different than on figures later in the chapter.

[2] See Chapter 2 for a more elaborate description.

Figure 7: Gini Coefficient for Norway 1875-2013. Upper and Lower bound



Source: Aaberge et al. (2017)

4.2.1 Comparing the Urban and Rural areas

Figure 8 shows the development of the Gini coefficient for the urban and rural areas in Norway. We see a sharp increase in the rural Gini coefficient after 1892. This may be attributed to the depression from 1876 and onwards and the following emigration to North America. Also, the economic growth in the 1890s may have affected income inequality relatively more in rural areas.

The Gini coefficient fell after a sharp increase at the beginning of the First World War. The cities grew in wealth, as seen in the development of the mean income. After this boom, there were several recessions with deflation and the depression that caused inequality to again rise during the 1920s and the beginning of the 1930s. We see a rise in inequality again just before the war, when GDP started to rise. This is similar to the increase in the national level by Aaberge et al. (2017), but in comparasion, the increase is stronger for both the rural and the urban areas.

During the Second World War, inequality fell sharply until 1953. Since the mid-1950s, we observe a relatively low rate of inequality until the mid-1980s. The increase shown in the graph

for the upper bound may be since women were taxed separately, which could lead to increased inequality for the data if they are included as persons with a low income. We see, however, with improved data from 1967, it declines and remains stable until the 1980s. Suggesting that we should interpret this increased inequality between the mid 1950s and 1967 with caution.

Figure 8: Rural vs Urban development of the Gini coefficient 1875 - present

Urban Gini Coefficient Rural Gini Coefficient

Reading note: The dotted line mark the introduction of personal tax data in 1967.

Figure 8 displays a convergence between the rural and urban areas, especially after the Second World War. Urban areas were relatively more unequal measured by the Gini coefficient before the First World War, and during the Great Depression, they diverged. It also seems as if the fluctuations in terms of the Gini coefficient is higher in urban areas, before 1940. Not surprisingly, as it is reasonable to assume that cities experience economic booms in a higher sense than the rural do. Recall the Harris-Todaro model described in chapter 2.3. It predicts that when the expected wage in the cities is higher than the salary in the rural areas, rural workers will migrate towards the cities. This fits well with the assumptions made in the Harris-Todaro model, as it also states that this may lead to unemployment, as labor in the cities are assumed to be bounded with a minimum wage.

Furthermore, these results show a pattern consistent with the findings of Lund (2012). She found that the rise in inequality took place later in the rural areas compared to the urban areas.

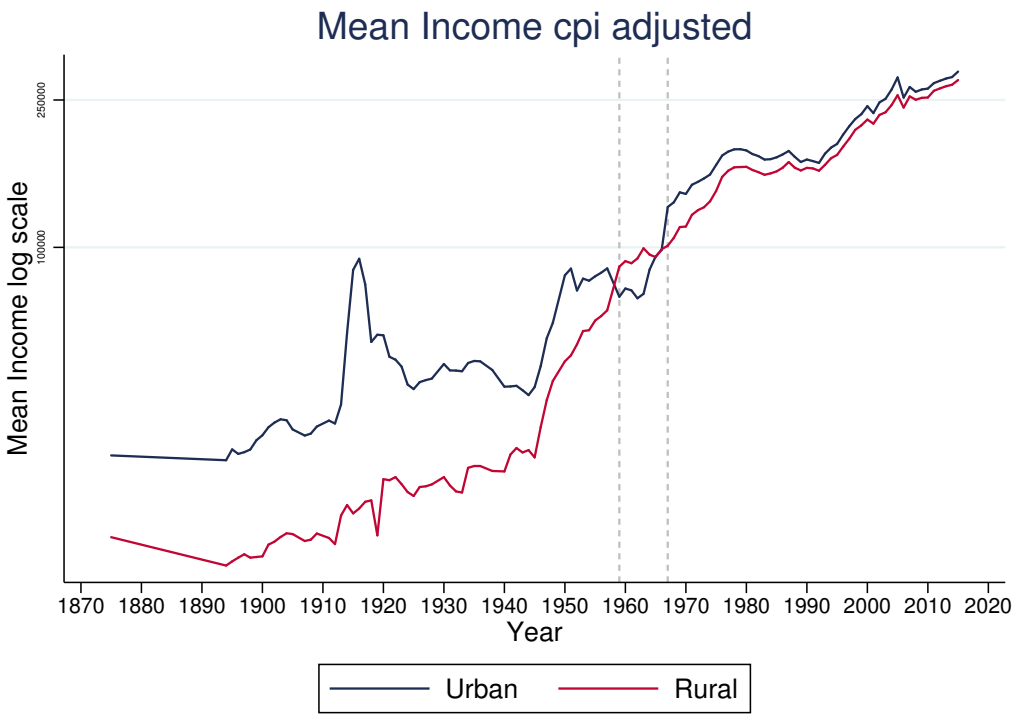
4.3 The Development of the Mean Income 1875-2015

This section describes the development of the mean income. The income figures are at constant prices, meaning they are CPI adjusted. The growth of the mean income in Norway has been

closer to exponential than linear, especially since after the Second World War, and is therefore on a logarithmic scale in figure 9.

Before the Second World War, the rural and urban mean income was less equal. The mean income increased substantially during the First World War, with a subsequent decline during the Great Depression. During the 1920s, we observe that the mean income decreases in the urban areas while it increases in the rural. As explained in Atkinson and Piketty (2007) the decrease could be caused by shocks to the capital owners' capital holdings during the world wars and the Great Depression. Since 1967, there has been a similar evolution of the mean income in rural and urban areas. Both areas have seen exponential growth since the 1960s, and both experienced a dip in mean income during the 1980s. The declined mean income corresponds to the banking crisis, that was followed by the deregulation of financial markets in 1984.

Figure 9: Mean Income development for Norway 1875- present. CPI adjusted



Reading note: The dotted lines represent the period where tax rules changed, and women were allowed to be taxed on their own. From 1967, they are again taxed as one core household.

4.4 The Evolution of Social Welfare between Urban and Rural areas

This section describes the evolution of social welfare from 1875 to 2015. Remember from section 3.5, that social welfare is estimated using the following equation. $SocialWelfare = (1 - G) \cdot \mu$.

Figure 10 shows the development of social welfare, introduced by Sen (1979). We observe that the social welfare in Norway has increased heavily since the beginning of the period. What is also interesting, is the level of social welfare measured in 1916. [3] These levels of social welfare were as high in the urban areas, as at the beginning of 1950 in Norway. For the rural areas, levels of social welfare remained low and stable until the end of the Second World War.

We do observe a decline in the 1980s that follows the banking crisis in Norway ended in 1992. Inflation grew in this period, which seem to cancel out the boom at the beginning of the 1980s. The fall of the oil price in 1984-85 and the collapse of banks in 1987 may explain the fall in the levels of social welfare for both rural and urban areas. It took nearly 20 years until we reached a level seen in 1980. The mean income started to stagnate, and together with the increased inequality, this explains why social welfare levels decline.

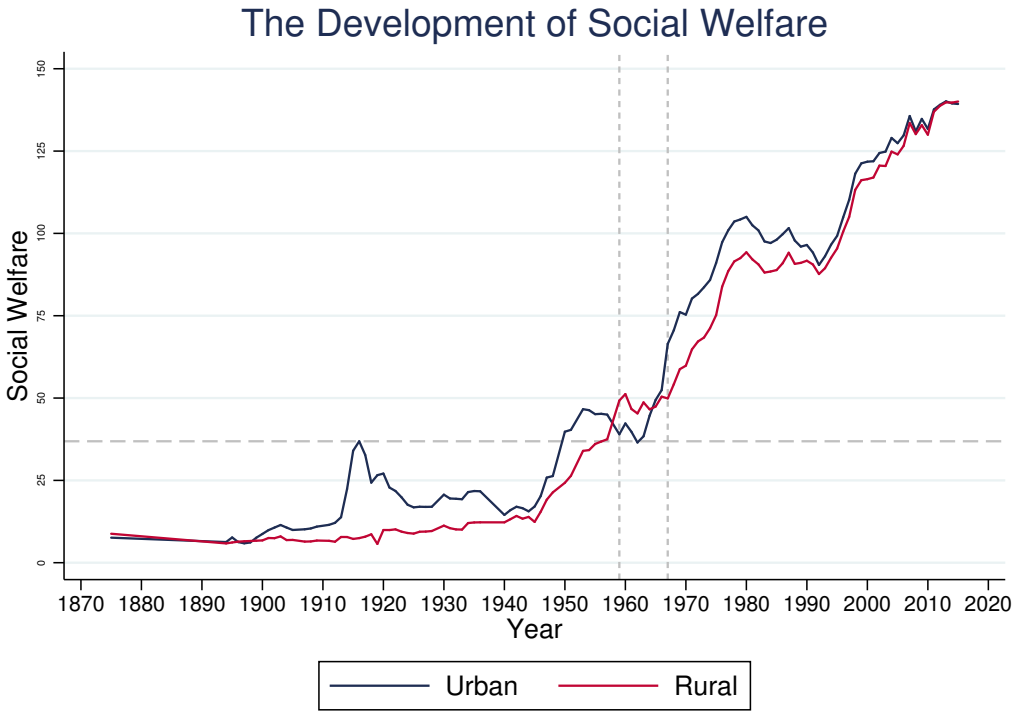
In general, Figure 10 tells a positive story from an economic perspective. Declined inequality combined with an exponential growth in the mean incomes has lead to a considerable rise in the Social Welfare measure since the end of the Second World War. The surge has been helped by an increased welfare state since the 1960s, which lowers the risk of individuals subject to economic recessions.

4.5 Possible explanations and comparisons with literature

The following section is devoted to analysis economic data and combine these analysis with historical evidence. Before we go through the episodic shifts in inequality, mean income and social welfare, it is useful to summarize them. Table 3 shows the Gini coefficient, mean income and social welfare in the rural and urban areas relative to the national average for selected years.

[3] When Norway's profit opportunities was at its highest during the First World War. This was also the year before the NKR started to deflate.

Figure 10: Social Welfare 1875-present



Reading note: The Scale of Social Welfare is the equally distributed income, described in section 3.5. To get a better reading of the scale, it has been divided by 1000.

Table 3: The development of Gini, mean income and Social Welfare for selected years

Area	1875			1914			1938			1953			1980			2015			Increase	
	μ	SW	Gini	μ	SW	Gini	μ	SW	Gini	μ	SW	Gini	μ	SW	Gini	μ	SW	Gini	μ	SW
Rural	0,75	1,07	0,79	0,51	0,52	1,00	0,69	0,78	0,90	0,84	0,84	0,99	0,95	0,95	1,00	0,97	1,00	0,97	1615 %	1492 %
Urban	1,25	0,93	1,22	1,49	1,48	1,01	1,30	1,22	1,10	1,16	1,16	1,01	1,05	1,05	1,00	1,03	1,00	1,03	987 %	1730 %
St Dev	0,25	0,07	0,22	0,49	0,48	0,01	0,30	0,22	0,10	0,16	0,16	0,01	0,05	0,05	0,00	0,03	0,00	0,03		

Reading note: Here μ represents the mean income, and SW represents social welfare. The increase shows the increase of CPI adjusted mean income and social welfare from 1875 to 2015. The standard deviation included.

The years are selected with the following motivation. 1875 is the starting point of our analysis. Between 1905 and 1914, the economic growth was high and stable. We are interested in determining how the growth influenced the development of urban and rural areas. Kuznets (1955) predicts that this growth would lead to migration from urban and rural areas, higher wages in the urban areas and increased inequality. Between 1914 and 1938, Norway had good profit opportunities during the First World War, unemployment during the Great Depression and experienced an increase in productivity for capital-intensive sectors. In 1938, the fishing industry started to grow, following a crisis. [4]. Furthermore, it is based on practicalities. 1938 is also the year before the Second World War we have income tabulations available. 1953 marks the year where inequality had declined after the Second World War. 1980 is a useful reference

[4] Accumulating in 'råfiskloven'

because this is before the period of the YAP [5] and the banking crisis, before increased inequality on a national level [6]. The year 2015 is the last year where we have available data.

Figure 3 shows the following development. Mean income was lower in rural areas compared to urban areas in 1875, but lower estimates of income inequality resulting in higher levels of social welfare. This is consistent with historical evidence, following from Kuznets hypothesis during the period of industrialization. The export industry had started to grow, with the shipping industry becoming the most important one. They were typically located in urban areas. As the productivity increases in urban areas, the mean income increases as well.

Between the years from 1875 to 1914, table 3 shows that mean income levels had diverged in this period. Although the country was relatively equal regarding the Gini coefficient, social welfare levels were significantly higher in urban areas. Furthermore, the variation between urban and rural areas was higher in 1914, relative to 1875. [7] To further provide insight, in chapter 2 we discuss the Kuznets U-curve and the Harris-Todaro framework. In a development process of the industrialization, which 1875-1914 was in Norway, Kuznets predicts [8] firstly, the share of the urban population increases, consistent with Figure 5 in section 4.1. Second, the relative difference in income per capita increases between urban and rural areas. As the Harris-Todaro model [9] predicts, when the expected wage is higher in urban than in rural areas, there will be urban migration, which may cause unemployment in urban areas.

During the First World War and the Great Depression until 1938, we observe that the Gini coefficient has diverged, but the mean income levels have converged. The convergence in mean income is consistent with the historical evidence, pointing out that capital owners suffered significantly as well. As discussed in chapter 2, Aaberge and Atkinson (2010) shows that the top income shares declined steadily from 1875-1948 in Norway, although unemployment increased during this period. But regarding social welfare, we observe a divergence. This corresponds to what we have found in the development of mean income in section 4.4.

In 1953, we observed that there had been a convergence regarding inequality and the mean income. Social welfare levels are higher in urban areas than in rural areas. Also, Norway

[5] Young Aspiring Professionals. (In Norwegian: 'Jappetiden')
[6] Aaberge et al. (2017)
[7] Measured regarding standard deviation (σ)
[8] See (Kuznets, 1955)
[9] See (Harris and Todaro, 1970)

experienced an export boom following the Second World War, and with a political motivation to rebuild the entire country, which likely has been a driver for the convergence of mean income, during a period of economic growth. As discussed in section 3 of this chapter, the Gini coefficient has fallen in this period, and to a level where rural and urban are relatively similar regarding within-group income inequality. In addition, there was a clear shift in the covariance between mean income and inequality. From being positive in 1938, it is now negative, implying a negative relationship between mean income and income inequality. This correlation may be attributed to the downside seen in Kuznets inverted U-hypothesis.

During the period from 1953 until 1980, the regions have converged regarding both mean income, income inequality, and social welfare. During this time, taxes started to increase gradually, and the welfare state expanded and increased considerably. An example of this is the introduction of the social security plan called 'Folketrygden' i 1967. This policy contributed to transfers to unemployed, senior citizens, farmers, fishermen, and students.

From 1980 until 2015, mean income and social welfare have converged. In fact, social welfare is estimated to be 0.53 per cent higher in the rural areas compared to the urban areas in Norway 2015. Although the mean income is slightly lower, the Gini coefficient is also lower, resulting in higher estimates of social welfare. The increase in the mean income is 64 per cent higher for rural areas than the increase seen in urban areas, while the rural areas show a 16 per cent lower increase regarding social welfare, compared to the urban.

Although we do see a convergence between rural and urban areas, it is hard to analyze the drivers behind the convergence. Is it so that some rural regions are richer regarding mean income than urban? It is natural to assume that for instance rural areas in the South-West, with relative prosperous farmers in Jæren for example, possesses higher income than small farmers in the North.

To summarize, we are already able to identify a few drivers in the evolution of regions in Norway. While, rural and urban areas have converged over time, the development over time has been quite different. Urban areas displayed high income inequality at the end of the 19th century, while the rural was relatively more egalitarian. Income inequality declined sharply from the end of the 19th century until the First World War, while rural income inequality increased in this period. This suggests that the decline in income inequality in Norway as a whole, was driven by the urban decline. This prompted social welfare to diverge between urban and rural

areas, before converging during the Great Depression and until the start of the Second World War.

The next chapter investigate the development of the urban/rural divide further, by estimating poverty, income inequality, the mean income and social welfare in the defined ten regions. This allows for further investigation on what role the different regions played in the episodic shifts described in this chapter.

5 The development of the Ten Regions

This chapter describes the evolution of regional income inequality and social welfare. This chapter follows the same structure as chapter 3 and 4, and compares the estimates of this chapter on the rural/urban divide, in order to determine regional differences. Further, the chapter continues by determining if the mean income has diverged or converged over time. Also, we estimate regional social welfare, and how the development of income inequality and mean income have affected the evolution of social welfare. Section 5.5 discusses possible explanations for trends in our estimates and relates these to the Norwegian economic history described in section 2.3. This section aims to elaborate further, adopting a more holistic view of the estimates combined with historical evidence.

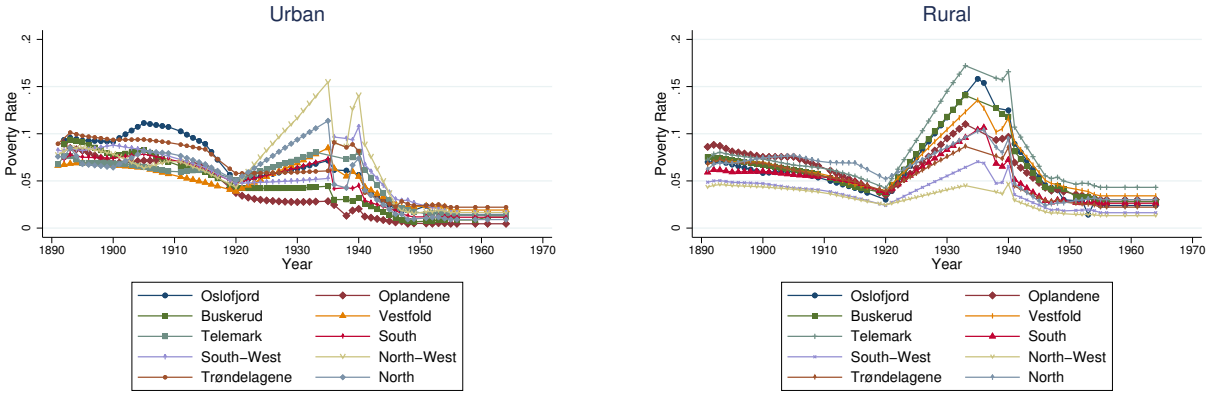
5.1 Development of the Poverty rate

Figure 11 displays the development of the poverty rate in the period until 1964, separated into rural and urban areas. Section F.1 of the appendix describes the development of the poverty rate in figures by region.

There are a few similarities between regions. The fraction of the population described as being poor declined from 1891 until 1920. Poverty rate increased during the Great Depression, with a peak during the 1930s. During the Second World War, poverty rates declined sharply, and remained stable until the 1960s. There are, however, also a few differences within regions. Urban North-West had the highest poverty rate before the Second World War, while the rural areas had the lowest poverty rates. During the Great Depression, the fishing sector which was an essential sector in urban North-West, particularly Aalesund, the most significant city of the region, was in a crisis.[1]. This may have been a contributing factor for the development of the poverty rate in the North-West.

[1]This crisis lasted until 1938 when 'råfiskloven' was introduced

Figure 11: Development of the Poverty rate for regions. Divided into urban and rural. 1891-1964



5.2 Income inequality over time

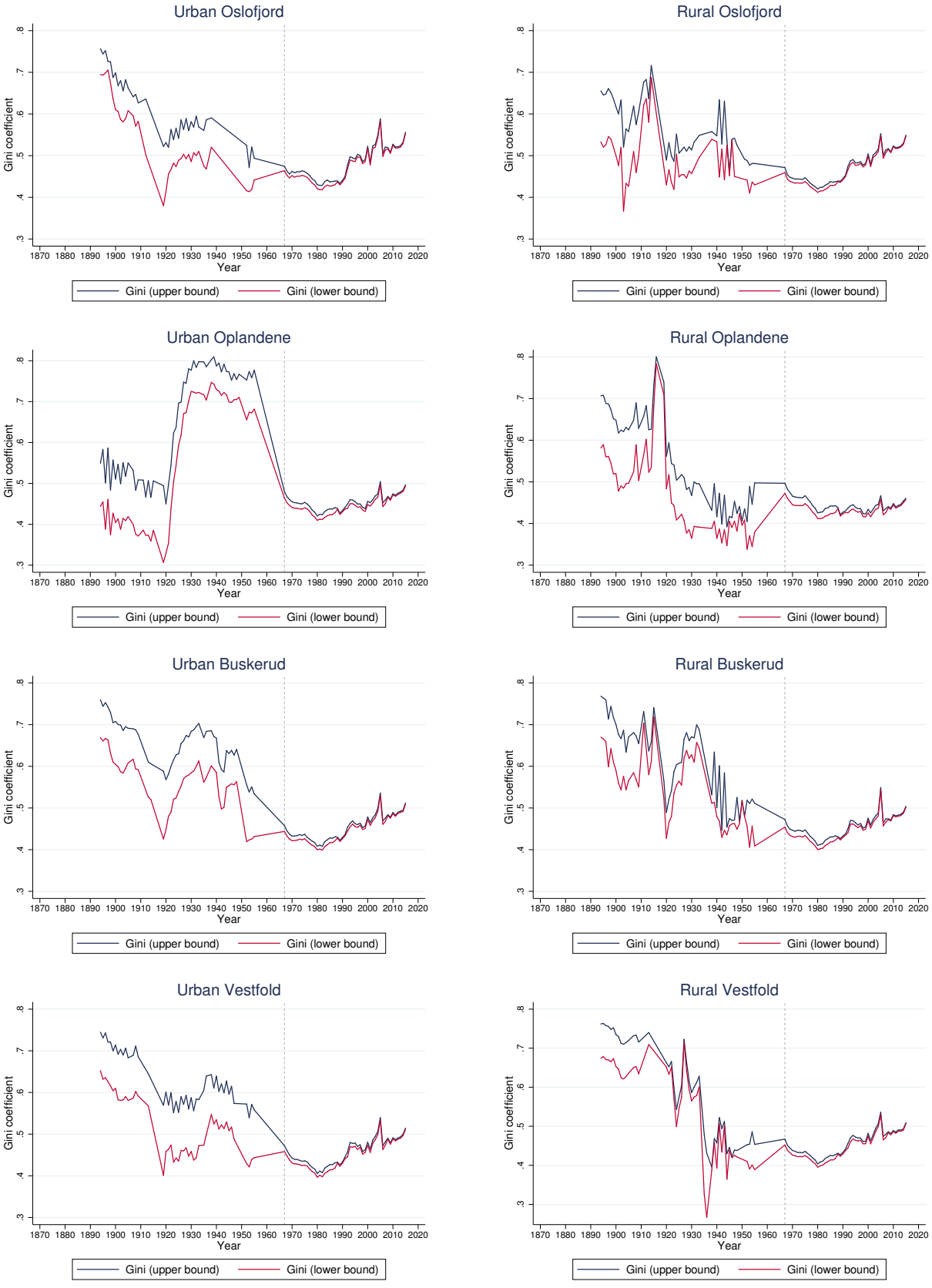
This section describes the development of income inequality over time, by estimating upper and lower bounds on the Gini coefficient, as discussed in chapter 3.4.2.

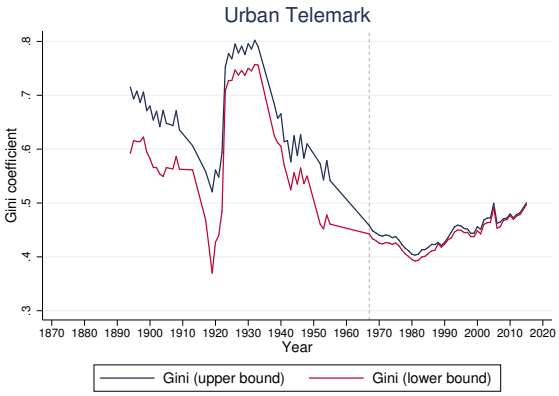
Figure 12 shows the development of income inequality. There are a few similarities with the urban and rural division. Income inequality has decreased over time. Although there are similarities between regions with an overall decline in income inequality, the regions have developed differently over time. Urban Oslofjord follows the pattern of the national level Gini coefficient, in spite of higher levels of inequality at the end of the 19th century, and with the a larger increase in inequality since 1988[2].

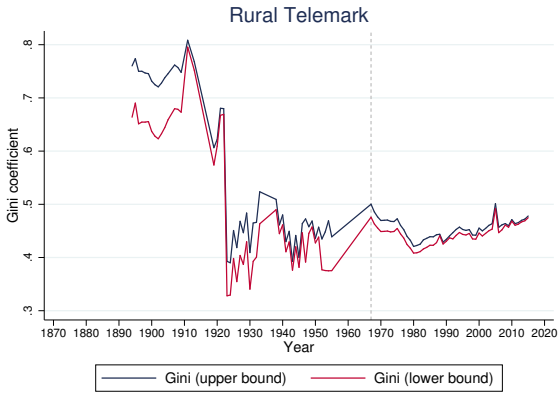
In urban areas of Oplandene, there seems to be a development of income inequality consistent with the Kuznets-curve. When analyzing the data, we observe an increase in the NAP[3] group, to almost 50 per cent of the population. Since 1967[4] inequality decreased close a level of the years before the Great Depression. This may indicate that the high levels of income inequality is partly driven by disturbances in the data, and the estimate might therefore be overstated. [5]

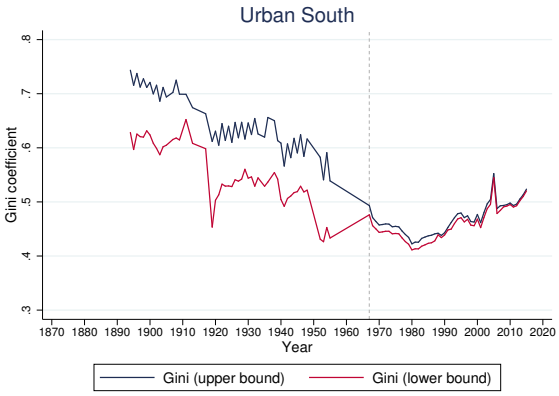
[2]Following the start of the Banking crisis
[3]Defined as Non-Assisted Poor (missing population)
[4]Personal tax data available Statistics Norway (2017c)
[5]This has been a topic for an ongoing debate, and has been tested using different sources of data, such as tax statistics by industry sector. As the estimates are robust, we chose to proceed. This could, however, be the topic for further research.

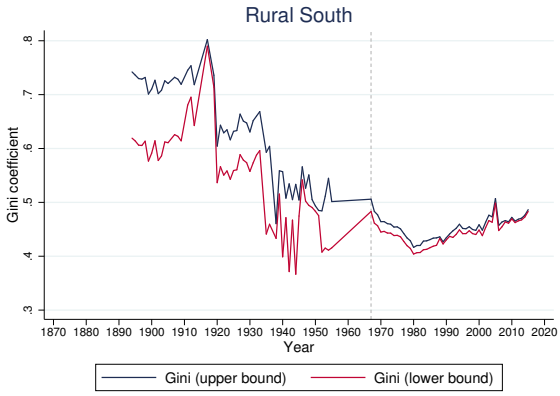
Figure 12: The Gini Coefficient, upper and lower bounds. 1891 - present.

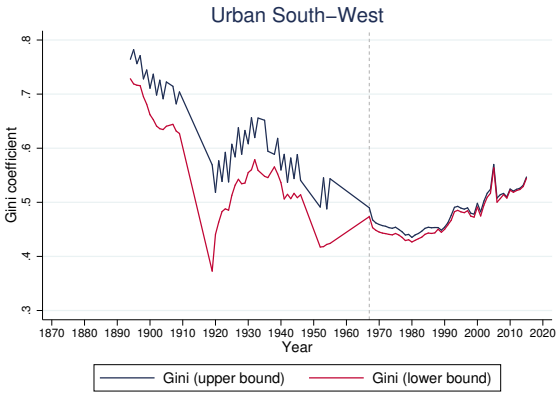


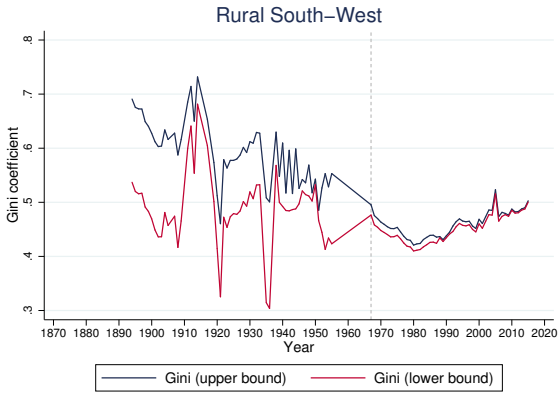

Urban Telemark
Gini coefficient
.3
.4
.5
.6
.7
.8
1870 1880 1890 1900 1910 1920 1930 1940 1950 1960 1970 1980 1990 2000 2010 2020
Year
Gini (upper bound)
Gini (lower bound)

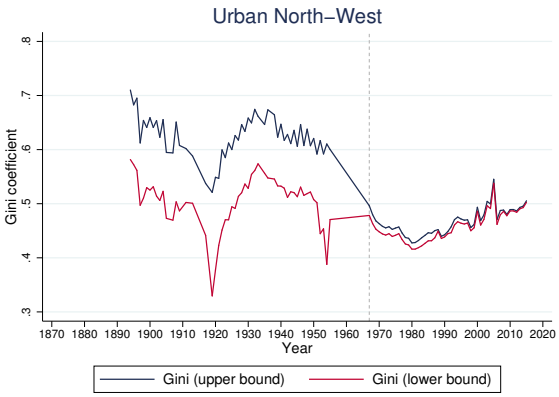

Rural Telemark
Gini coefficient
.3
.4
.5
.6
.7
.8
1870 1880 1890 1900 1910 1920 1930 1940 1950 1960 1970 1980 1990 2000 2010 2020
Year
Gini (upper bound)
Gini (lower bound)

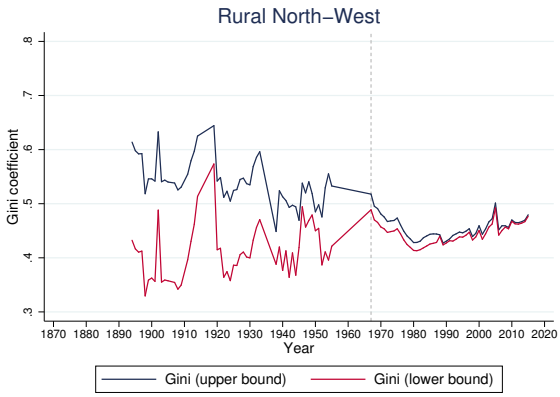

Urban South
Gini coefficient
.3
.4
.5
.6
.7
.8
1870 1880 1890 1900 1910 1920 1930 1940 1950 1960 1970 1980 1990 2000 2010 2020
Year
Gini (upper bound)
Gini (lower bound)

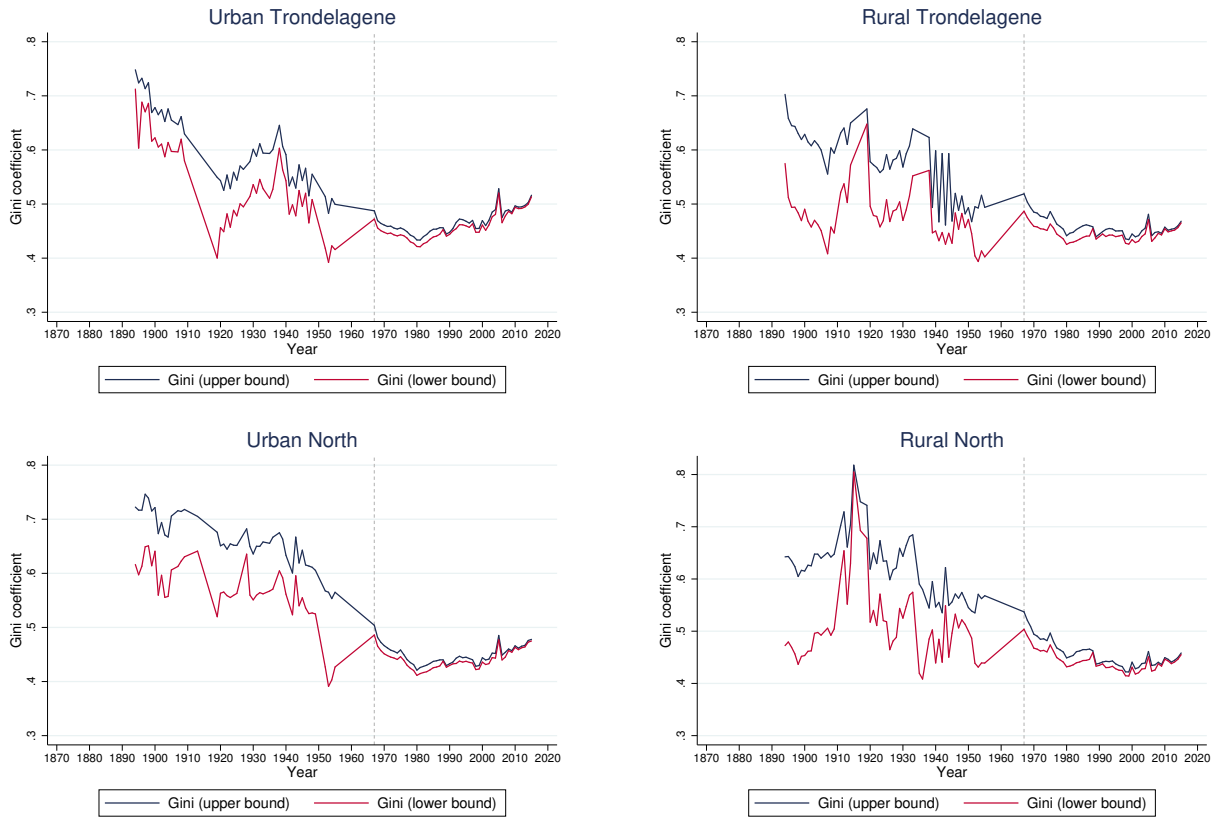

Rural South
Gini coefficient
.3
.4
.5
.6
.7
.8
1870 1880 1890 1900 1910 1920 1930 1940 1950 1960 1970 1980 1990 2000 2010 2020
Year
Gini (upper bound)
Gini (lower bound)


Urban South–West
Gini coefficient
.3
.4
.5
.6
.7
.8
1870 1880 1890 1900 1910 1920 1930 1940 1950 1960 1970 1980 1990 2000 2010 2020
Year
Gini (upper bound)
Gini (lower bound)


Rural South–West
Gini coefficient
.3
.4
.5
.6
.7
.8
1870 1880 1890 1900 1910 1920 1930 1940 1950 1960 1970 1980 1990 2000 2010 2020
Year
Gini (upper bound)
Gini (lower bound)


Urban North–West
Gini coefficient
.3
.4
.5
.6
.7
.8
1870 1880 1890 1900 1910 1920 1930 1940 1950 1960 1970 1980 1990 2000 2010 2020
Year
Gini (upper bound)
Gini (lower bound)


Rural North–West
Gini coefficient
.3
.4
.5
.6
.7
.8
1870 1880 1890 1900 1910 1920 1930 1940 1950 1960 1970 1980 1990 2000 2010 2020
Year
Gini (upper bound)
Gini (lower bound)



5.3 The development of Mean Income

This section describes the development of the mean income. Table 4 shows the development of the mean income, relative to the national average at the time. Comparing with the national average allows us to analyze how the regions have developed relative to each other. Figures on a regional level, are available in appendix F.2

The central insight from table 4 is that regions have converged regarding the mean income from 1892 until today. There are fewer differences between regions, and between the urban and rural divide. Rural areas in the North in 2015 has the lowest fraction of mean income, being 85 per cent of the national average, while the highest fraction in rural regions in Oslofjord is 122 per cent of the national average. Also, there has been a clear converge within regions, especially since 1914. However, since 1980, there has been a divergence in the share of mean income.

While there are similarities between the regions, there are also quite a few differences. Before the Second World war, the mean income in Buskerud, Vestfold, and Oslofjord was above the national average for both the rural and urban areas, while rural Oplandene regions, the

Table 4: The Development of Mean Income relative to national average for the Regions.Increase from 1892-2015. σ as a measure of convergence.

Region	1892	1914	1938	1953	1980	2015	Increase
Urban Oslofjord	1,59	1,84	1,48	1,47	1,12	1,15	903 %
Rural Oslofjord	1,19	0,72	0,97	1,25	1,16	1,22	1484 %
Urban Oplandene	1,59	1,66	0,88	0,78	0,99	0,94	428 %
Rural Oplandene	0,68	0,57	0,67	0,96	0,88	0,85	1378 %
Urban Buskerud	1,33	1,19	1,22	1,27	1,07	1,04	842 %
Rural Buskerud	1,07	0,84	0,97	1,07	1,02	1,03	1192 %
Urban Vestfold	1,09	1,31	1,42	1,17	1,02	0,98	1031 %
Rural Vestfold	1,28	1,16	1,72	1,25	1,05	1,02	1068 %
Urban Telemark	0,84	1,16	1,17	1,10	1,05	0,94	647 %
Rural Telemark	1,18	0,90	1,01	0,98	0,92	0,91	976 %
Urban South	0,79	1,04	1,15	0,95	1,01	1,01	929 %
Rural South	1,06	0,57	0,88	0,80	0,94	0,95	1285 %
Urban South-West	1,02	1,93	1,11	1,08	1,10	1,18	1129 %
Rural South-West	0,58	0,50	0,65	0,82	1,01	1,09	2143 %
Urban North-West	1,23	1,18	1,02	0,92	0,99	1,02	839 %
Rural North-West	0,53	0,39	0,50	0,71	0,89	0,96	1975 %
Urban Trøndelagene	0,81	1,16	1,15	1,06	1,03	0,99	925 %
Rural Trøndelagene	0,62	0,51	0,56	0,70	0,87	0,89	1660 %
Urban North	0,87	0,84	1,02	0,99	1,02	0,96	973 %
Rural North	0,66	0,51	0,46	0,68	0,85	0,86	1483 %
σ (Std. Dev.)	0,32	0,45	0,33	0,22	0,09	0,10	

West and the North were below the national average. In the rural areas of Oplandene, the mean income have increased by 1378 per cent [6], which relative to areas with similar mean income before the First World War, such as Rural Trøndelagene, Rural North - and South West, was relatively modest.

The mean income across the regions has evolved differently over time. While the mean income of urban areas of Oplandene was equal to 1.58 of the national average, it decreased to 0.88 of the national average in 1938. This decline in mean income could partly be driven by the closing of factories in Oplandene, such as 'Hunton' which was a central part of the economy in Gjøvik. Before the crisis in the 1920s, the construction of railroads around Hamar created stable jobs helped see the mean income lie above average. The creation of railroads also employed workers in Vestfold, Telemark, Buskerud, the South-West and the South. The differences between rural and urban areas of the regions are largest in Oplandene, where urban mean incomes were high at a stable level from the 1890s until the beginning of the Great Depression.

[6] CPI adjusted

In figure 22 in appendix F.2, we observe a lot of fluctuations in the mean incomes for regions located by the coast such as Vestfold, South West, and North West. The mean income in the South-West cities was as high in 1916 as it was at the beginning of the 1970s. This may be caused by the boom in Norway during the First World War [7]. As we discussed in section 2.3, the shipping industry and woodworking industry experienced an increase in demand from Europe. In chapter 4, this boom was present in both rural and urban areas, however, they are not for every region. This suggests, that some regions in Norway were more affected than others by the profits during the First World War. Rural regions show fewer fluctuations compared to urban in the development of mean incomes. Assuming that unemployment shifts the mean wage, this may be explaied by the description of rural wages by Harris and Todaro (1970). They suggest that rural wages are not bound by a minimum wage. This will lead to a low unemployment rate in the rural areas. Following the Second World War, however, the rural areas has seen a steady growth in terms of the mean wage. Also, this was a period where the government heavily regulated markets, formalizing the more informal, rural sectors.

5.4 The Development of Regional Social Welfare

The following section describes the development of social welfare in the ten regions in Norway. The estimate of social welfare follows from section 3.5 and is calculated as follows, $SocialWelfare = (1 - G) \cdot \mu$, where μ is the mean income and G is the Gini coefficient.

Table 5 displays perhaps the most interesting findings of this thesis. There is a clear pattern of convergence from 1892 to 2015. When comparing with the convergence of the mean income, we observe that social welfare has converged significantly more than mean income. Also, even after 1980, where, intriguingly, there was a divergence in the mean income. Between 1980 and 2015, the Gini coefficient has also diverged across regions.

Why has social welfare converged between regions? In order to build some insight, we estimate the correlation between the mean income and income inequality. In 1980, the correlation was negative, implying that regions with a low mean income, also had higher income inequality, resulting in low estimates of social welfare. Regions with high mean income had a relatively

[7] The measure of inflation at the time should be regarded with caution, as the measure was constructed using "Levekårsundersøkelsen" by Statistics Norway, which mainly focused on the living standard of people living in Oslo. It is not clear how the national inflation affected the different regions.

lower income inequality, leading to higher social welfare. [8] During the period between 1980 and 2015, the correlation had shifted from negative to positive. In 2015, those regions possessing the highest mean income also were the regions with highest income inequality. In other words, regions such as Oplandene, North, and Trøndelagene are now relatively more equal, while Oslofjord and the South-West are now relatively more unevenly distributed. This implies that the effect of the tax reforms during this period has affected the wealthiest regions the most. This shift in the correlation between the mean income and income inequality is solely the driver behind the convergence in social welfare. The interaction between income inequality and the mean income is displayed in figure 22 in appendix F.4.

Table 5: The Development of Social Welfare relative to national average. Increase from 1892-2015. σ as a measure of convergence.

Region	1892	1914	1938	1953	1980	2015	Increase
Urban Oslofjord	1,05	1,27	1,51	1,56	1,10	1,03	1528 %
Rural Oslofjord	1,04	0,96	1,01	1,32	1,16	1,11	1665 %
Urban Oplandene	2,98	2,68	0,45	0,41	1,01	0,95	430 %
Rural Oplandene	0,68	0,71	0,91	1,04	0,88	0,92	2145 %
Urban Buskerud	1,05	1,22	0,99	1,25	1,09	1,02	1512 %
Rural Buskerud	0,75	0,99	1,06	1,10	1,04	1,03	2187 %
Urban Vestfold	0,87	1,15	1,32	1,16	1,05	0,96	1728 %
Rural Vestfold	0,82	0,83	2,40	1,37	1,08	1,01	1936 %
Urban Telemark	1,46	1,22	0,93	1,05	1,08	0,95	982 %
Rural Telemark	0,81	0,64	1,14	1,09	0,92	0,96	1855 %
Urban South	1,03	0,81	1,04	0,93	1,01	0,97	1465 %
Rural South	0,73	0,62	1,12	0,82	0,95	0,98	2134 %
Urban South-West	0,81	0,89	1,07	1,06	1,07	1,08	2102 %
Rural South-West	0,63	0,64	0,59	0,81	1,01	1,10	2796 %
Urban North-West	1,28	1,41	0,93	0,83	0,99	1,02	1216 %
Rural North-West	0,74	0,70	0,66	0,71	0,89	1,01	2171 %
Urban Trøndelag	0,87	1,17	0,99	1,13	1,01	0,97	1744 %
Rural Trøndelag	0,61	0,69	0,52	0,74	0,84	0,96	2498 %
Urban North	0,99	0,78	0,84	0,98	1,02	1,01	1599 %
Rural North	0,80	0,61	0,52	0,64	0,81	0,94	1844 %
σ (Std. Dev.)	0,50	0,46	0,42	0,27	0,09	0,05	

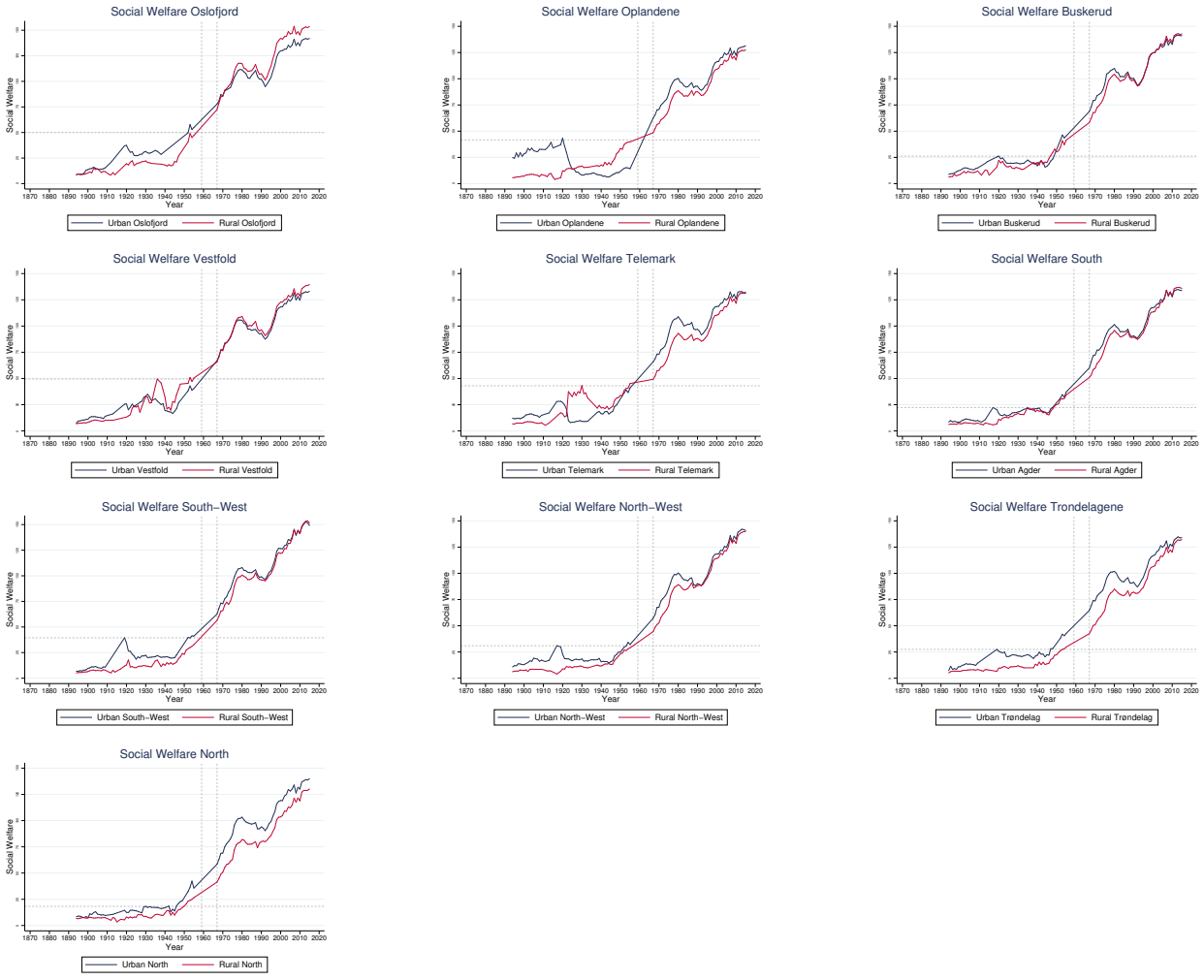
Furthermore, table 5 shows the development of the social welfare over time, as an increase from 1892 to 2015. We do find that regions have experienced differences in the development over time. Rural South-West and rural Trøndelag had the lowest levels of Social Welfare in 1892. Their levels has increased significantly more than regions such as urban Oplandene and

[8] This is in line with a common perception of Norway. We have experienced a large economic growth, with great natural resources and a high mean income, and also a low income inequality.

urban North-West. Table 5 shows that the differences are quite small in 2015, with the greatest relative difference is between urban and rural North, namely 7,6 per cent.

Figure 13 displays the development of social welfare by region, following the development of different regions over time. The dotted lines on the y-axis indicate a point in time where social welfare was as high pre-World War II, as after. Firstly, the overall increase in social welfare in chapter 4 is still present, but not all regions follow the same development. The levels of social welfare after 1980 dips for all regions. Even more so in the urban areas than in the rural areas, and more for the regions located by the coast. Moreover, we observe a decline in social welfare levels due to the introduction of the dividend tax in 2006. The timing effect on increased dividend payments in 2005 by the 10 per cent richest, resulting in increased inequality and decreased social welfare, is consistent with the findings of Aaberge and Atkinson (2010) and Alstadsæter and Fjærli (2009).

Figure 13: Social Welfare in the ten regions



5.5 Possible explanations and comparisons with literature

This section is devoted to possible explanations for the episodic shifts in income inequality and social welfare. Our focus is on the principal changes in the economy, and to further elaborate on the findings in chapter 4. We are interested in how the shifts in the Norwegian economy affect the different regions. We do not investigate in detail every region in detail through the period.

As discussed in chapter 2, Soltow (1965) and Mjelve (1998) state that, the effect of historical events which affected one or a few industries across the regions, varied more before 1967, where we observe more shifts in the data. The industry structure between regions has become more similar as we move forward in time. This thesis does not identify all drivers causing shifting inequality between 1875 and 2015 but merely adds pieces to a puzzle by combining historical evidence with economic theory and data.

Before the First World War, there were both booms and recessions in the economy. From appendix F.1 figure 19, we find that for Trøndelagene, Oslofjord, Telemark, South and South West, in particular, experienced urban migration during this period. Inequality in the regions with the most apparent urban migration, also declined significantly than other regions. Between 1891 and 1914, the mean income had diverged between regions, suggesting that the industrialization did lead to higher inequality in mean income between regions. This is in line with the Kuznets hypothesis. Social welfare levels converged for two reasons. Firstly, the Gini coefficient between regions converged, and secondly, the covariance between mean income and the Gini coefficient was less positive. Implying that high mean income did not necessarily imply high inequality or vice versa.

From 1914 to 1938, the poverty rate rose significantly. While social welfare was relatively higher in the urban compared to rural areas in 1938, this is not the case for all regions. In fact, we only observe this in five of the ten regions. The main driver behind the higher levels of social welfare in urban areas seems to be Oslofjord, South- and North-West, Trøndelagene and the North. Implying that in these regions, the rural areas were more adversely affected by the Great Depression, relative to its urban areas.

During the Second World War, income inequality steadily declined to a low and stable level until the mid-1950s, lasting until the 1980s. Following the Second World War, tax rates

became more progressive. A progressive tax rate affects the higher earners more than the lower ones, in order to redistribute income. Also, between 1945 and 1948, GDP increased by 50 per cent, due to significant demand from Europe on Norwegian export goods. Historical evidence tend to regard Norway as an egalitarian society in the years following the Second World War. Supported by the variation in the levels of social welfare between regions in 1953, it was not as egalitarian as maybe the historical literature suggests.

The period between the Second World War and the 1960s is often referred to as "the golden years," because Norway experienced a long-lasting economic boom. The export sector increased 2.5 more than the GDP in the period between 1948 to 1973. The two primary industries driving this boom were the shipping industry and the energy sector. Also, Norway benefited greatly from the Marshall-aid program. Furthermore, social welfare programs such as 'Folkepensjonen' provided transfers for individuals possessing the lowest wages. [9]. From 1953 until 1980, the mean income had also converged, and together with a relatively low and stable income inequality, the social welfare levels had converged significantly.

Since 1980, the average growth in Norway's economy has been higher than our trade partners'. This is primarily due to income from the oil sector and its complementary industries. After the mid-1980s, however, the income inequality has both diverged between regions and started to increase. In this recent period, income inequality have experienced two upward jumps, identified by Aaberge and Atkinson (2010). Both rises are related to tax reforms. In 2001, there was a temporary tax reform on dividends, and in 2006 there was a permanent dividend tax. The authors find that the tax on dividends gave an increase in the years before the reforms, declining the following year, which corresponds to the findings of the thesis. In addition, these jumps have affected the social welfare levels differently for the regions. The correlation between the mean income and income inequality shifted from negative to positive during the banking crisis, and became further positive following the tax reform in 1992.

This section has tried to summarize the regional development of inequality, the mean income, and social welfare, by comparing estimates with economic theory and historical evidence. The evolution over time may have been more complicated than this section suggests. Further research topics could combine these findings with a broader range of economic theory or/and investigate these periods further combined with more detailed data on different industry groups.

[9] such as students, the poor and senior citizens

6 Affluence

This chapter describes the development of the affluence measure. Since the information basis is better for the upper tail than the lower tail of the income distribution, it will be useful to complement estimates of the Gini coefficient for the entire income distribution with estimates of affluence and the upper tail Gini coefficient. Affluence is a weighted average of top income shares, which gives us more information about the population possessing income above the mean.

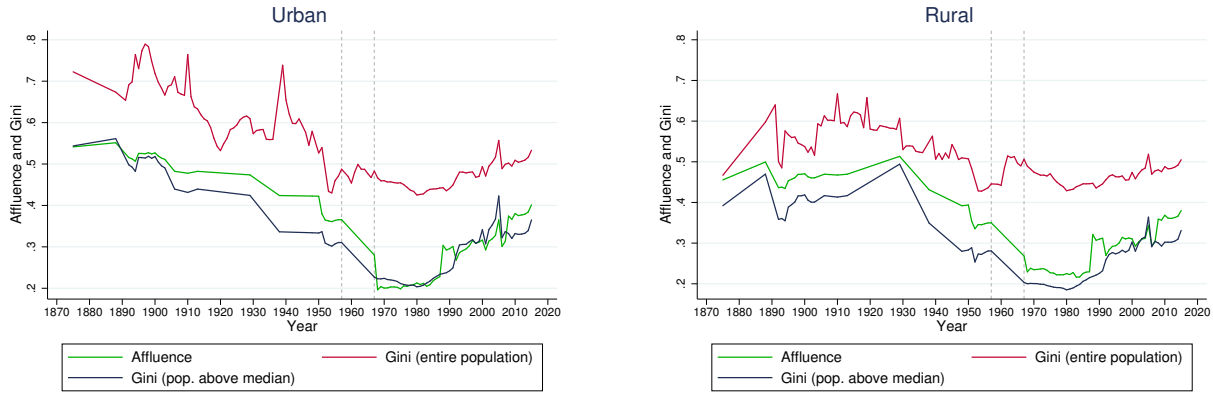
The measure is defined as[1],

$$A = \frac{1}{3}\left(\frac{\mu_u}{\mu}G_u + \frac{\mu_u - \mu}{\mu}\right) \tag{6.1}$$

Note that 3A can also be interpreted as the richness gap, if individuals that possess a higher income than the median all possess the same income.

6.1 Rural-Urban Norway

Figure 14: Gini for entire population, Gini above median and affluence. 1875 - present



Reading note: Dotted lines represents period where married women are taxed separately, and this period should therefore be interpreted with caution.

[1] As described in section 3.6, but since this is a relatively long thesis, it could be useful to get a reminder. For further justification, See (Aaberge et al., 2015)

The Gini coefficient for the population above median does follow the same trend as the Gini coefficient for the entire population, except for the spike in inequality before the Second World War. The fall in inequality from 1875 to 1980 is even larger for the population earning above the median, suggesting that the decline in overall inequality is driven by a drop in upper tail inequality for the urban areas. There is an increased divergence between the upper tail Gini coefficient and the Gini coefficient from 1953 to 1984. In 1984, there was a deregulation of financial markets, consistent with increased upper tail Gini.

In rural areas, the decline in inequality after the Second World War was primarily driven by the decline in upper tail inequality, since 1929. The ratio between upper tail mean income and overall mean income was also decreased. Poverty rates, as seen in Figure 11, increased during the Great Depression and remained high until 1940, but inequality was relatively stable. This leads us to believe that in rural areas, the Great Depression affected the rich as much as the poor.

Table 6: Urban : Overview of the evolution of overall inequality, upper tail inequality, and affluence(Changes in percentage points in parentheses).

Period	Overall Gini coefficient	Upper tail Gini Coefficient	Affluence	Upper mean relative to overall mean
1875-1892	Decrease (-4)	Decrease (-9)	Decrease (-5)	Decrease
1892-1914	Decrease (-8)	Decrease (-12)	Decrease (-6)	Decrease
1914-1929	Decrease (-4)	Decrease (-3)	Decrease (-2)	Decrease
1929-1938	Decrease (-7)	Decrease (-21)	Decrease (-11)	Decrease
1938-1953	Decrease(-24)	Decrease (-9)	Decrease (-14)	Decrease
1953-1980	Decrease (-2)	Decrease (-33)	Decrease (-41)	Decrease
1980-2015	Increase (+25)	Increase (+79)	Increase (+89)	Increase

Table 7: Rural : Overview of the evolution of overall inequality, upper tail inequality, and affluence(Changes in percentage points in parentheses).

Period	Overall Gini coefficient	Upper tail Gini Coefficient	Affluence	Upper mean relative to overall mean
1875-1892	Increase (+7)	Decrease (-9)	Decrease (-4)	Decrease
1892-1914	Increase (+17)	Increase (+16)	Increase (+8)	Increase
1914-1929	Increase (+4)	Increase (+19)	Increase (+9)	Increase
1929-1938	Decrease (-11)	Decrease (-29)	Decrease (-16)	Decrease
1938-1953	Decrease (-21)	Decrease (-22)	Decrease (-20)	Decrease
1953-1980	Unchanged	Decrease (-32)	Decrease (-35)	Decrease
1980-2015	Increase (+18)	Increase (+79)	Increase (+69)	Increase

The evolution of inequality in rural/urban Norway is best characterized as a series of episodes identified with sub-periods, summarized in tables 6 and 7. As demonstrated by the change in percentage points, the evolution of the overall Gini coefficient is closely related to the development of the above median Gini, and the Gini-based affluence measure. In the same way, the relative income of the upper group typically moves in the same direction.

The contribution regarding magnitudes does however differ. Taken together, the periods from 1875 to 1938, the affluence measure for rural areas decreased two percentage points, whereas upper tail Gini coefficient declined by 7 per cent. For urban areas, the changes for the equivalent period are a decrease in affluence by 9 per cent and a reduction of 21 percentage points for the upper tail Gini coefficient. The different evolution of upper tail and overall inequality corresponds to a significant decrease in the ratio between mean incomes for the upper and lower part of the distribution for the urban areas. While an increase in the same ratio for rural areas corresponds to the evolution of the two Gini coefficients. From 1938 to 1953, we observe a decrease in the overall Gini coefficient by 12 and ten percentage points for the urban and rural areas respectively. These falls do not seem to be driven by a decreased upper tail Gini coefficient, as described in Aaberge et al. (2017), but rather a decrease in the ratio between upper mean income and overall mean income. From 1953 to 1980, we observe a slight decrease in the overall Gini coefficients for both urban and rural areas, while both upper tail Gini and affluence have decreased by a more significant magnitude.

When comparing with the national levels by Aaberge et al. (2017), there are several differences. Firstly, the decrease in the Gini coefficient in 39 per cent for urban areas are higher than the 15 per cent found for Norway, while the 8 per cent decrease in the Gini coefficient is less when comparing to the national level. The differences between urban and rural areas support the findings of Lund (2012), which the decline in inequality happened later in the rural areas. The upper tail Gini coefficient reaches a lower level for both urban and rural areas compared to the national level before 1980. Since 1980, the increase of 79 per cent in the upper tail Gini coefficient evident in both urban and rural areas. This increase in upper tail Gini coefficient is significantly higher compared to the national level [2] findings of 12 per cent, further underlining the importance of dividing Norway into rural and urban areas. A possible explanation could be that the deregulation of financial markets in 1984, followed by the banking crisis in 1987, had a larger impact within rural and urban areas, than when analyzing Norway as a whole.

To gain more insight on this development, we disaggregate the country even further and estimate affluence and Gini coefficients for both the upper tail and the average between upper and lower tail for the entire population.

[2] See (Aaberge et al., 2017)

6.2 Regional

This section describes the combined regional development of affluence, the upper tail Gini, and the Gini coefficient. As stated at the end of chapter 4, although we can identify specific drivers, and gain some insight in the urban/rural divide, it is difficult to determine the effect each region has on the estimates evaluated.

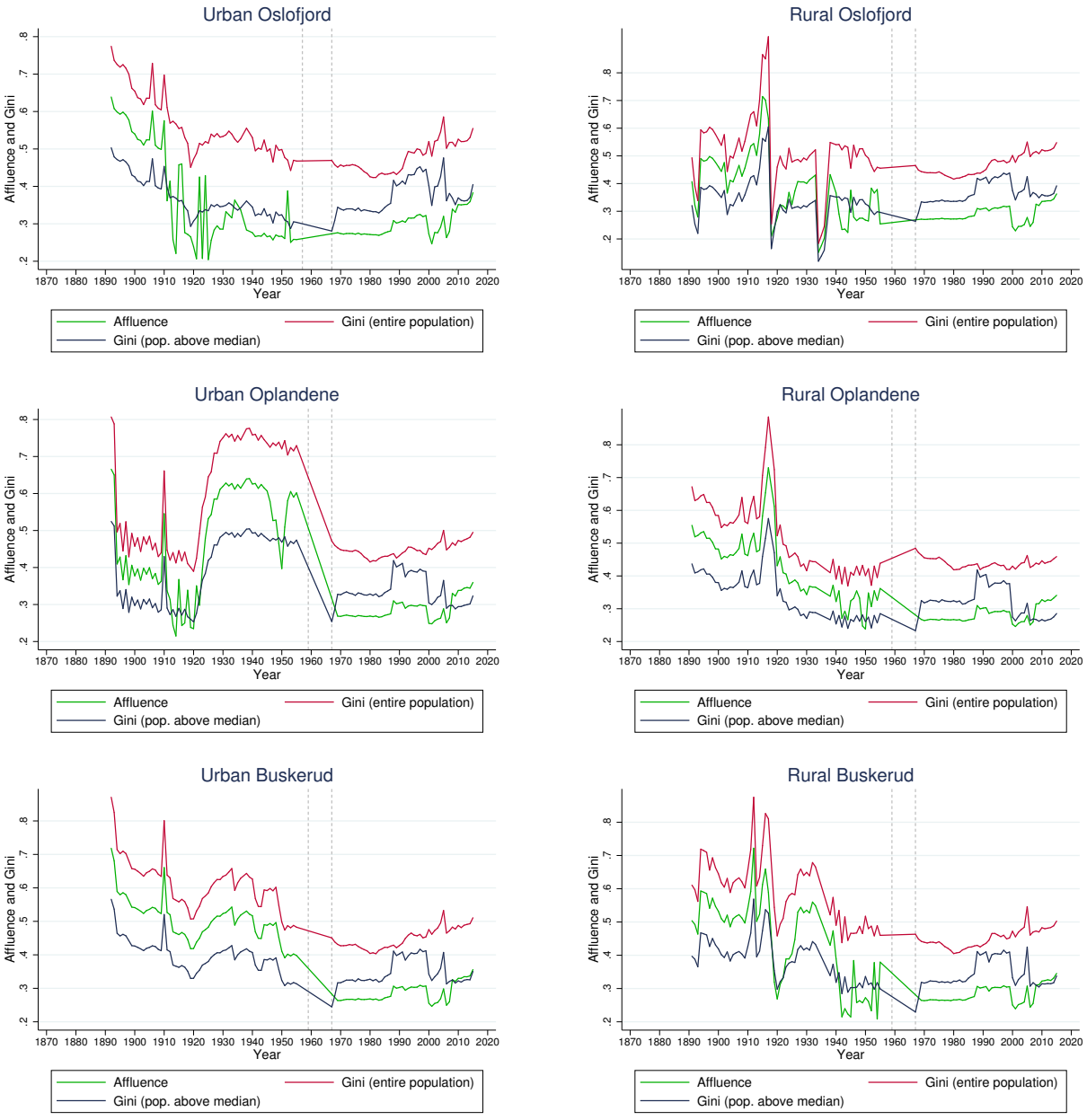
Figure 15 displays the development of three estimates. The Gini coefficient for the entire population, taking the average between upper and lower bounds, the Gini coefficient for the population of individuals who possess income above the median, and the affluence measure.

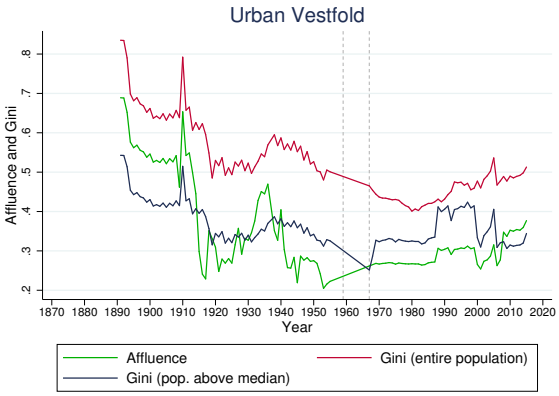
There are similarities between all regions. Between 1938 and 1953, the affluence measures decrease substantially more than the overall Gini coefficient and the upper tail Gini coefficient. The decreased affluence is driven mainly by the decrease in the ratio between the upper tail mean income and the overall mean income. From 1967 until the present, the development of the upper tail Gini coefficient is consistent for all regions. In 2000, the affluence measure increased for all regions, while the upper tail Gini coefficient declines. In 2005, the upper tail Gini coefficient and the overall Gini coefficient increases. Both jumps are related to tax reforms. In 2001 there was a temporary tax reform on dividends, and in 2006 there was a permanent dividend tax. Aaberge and Atkinson (2010) find that the tax reforms on dividends gave an increase in the top income shares in 2000 and 2005, and the subsequent decline the year the tax was introduced, which corresponds to these findings. Also, the increase in the upper tail Gini coefficient in 1988 corresponds to the banking crisis in Norway, followed by the crash in the Wall Street stock market in 1987.

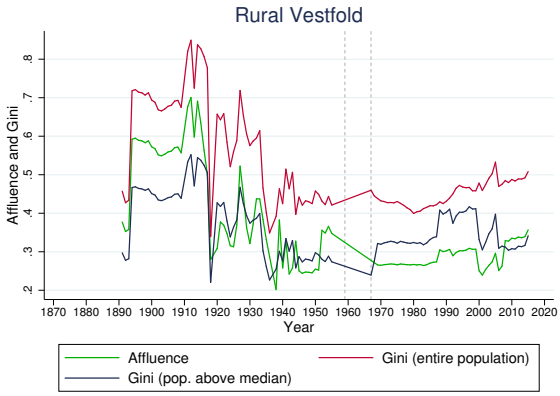
Magnitudes of the different measures still differ on a regional level. In urban areas of Oslofjord, Vestfold, South, and Trondelagene, the affluence measure falls below the upper tail Gini coefficient during the Second World War, driven by a declining ratio between the two mean incomes. The trend for rural regions are consistent with earlier findings concerning this rate; they seem to experience a decrease in inequality later compared to the cities. We observe that the jump in the upper tail Gini coefficient in 2005, although visible at all regions, differs regarding magnitude. In urban North-West, the increase was 19 per cent, while only 12 per cent for the rural areas.

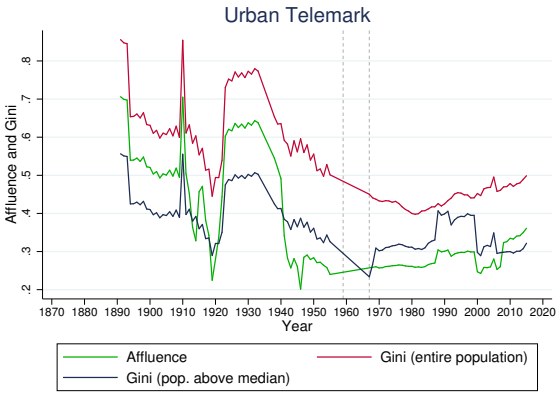
The jump in upper tail Gini followed by the stock market crash in 1987 also differs regarding magnitude. In Trondelagene, upper tail Gini increased by 19 per cent in urban areas and 30 per cent for rural areas. In Oslofjord, the increase was 16 per cent both for rural and urban areas. The two jumps, in 1988 and 2005, represent two different aspects of the economy. In the banking crises, the ratio between the above median mean income and the increased, caused by a fall in the mean income, as seen in Figure 20 in appendix F.2. In 2005, the year before the permanent divided tax, the ratio did not increase, but increased dividends caused the jump in inequality, typically benefiting the richest.

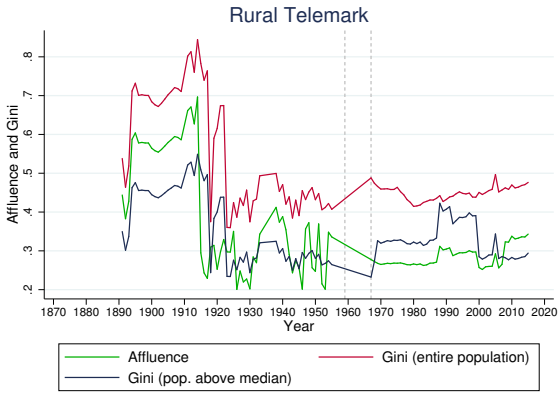
Figure 15: Gini for entire population, Gini above median and affluence. 1891 - present

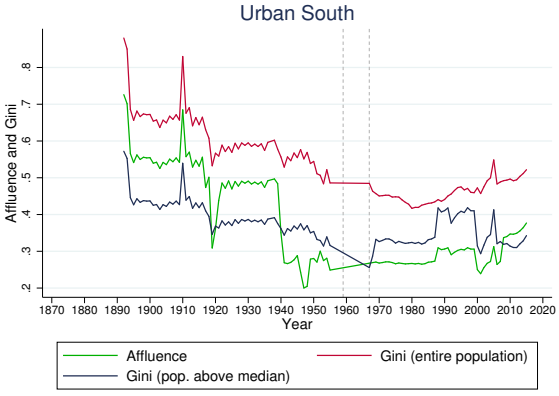


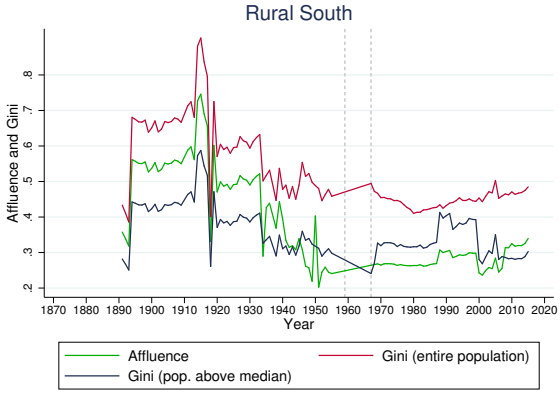

Urban Vestfold
Affluence and Gini
.2
.3
.4
.5
.6
.7
.8
1870 1880 1890 1900 1910 1920 1930 1940 1950 1960 1970 1980 1990 2000 2010 2020
Year
Affluence
Gini (entire population)
Gini (pop. above median)

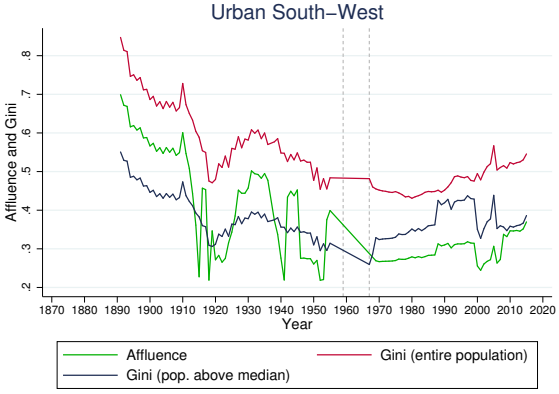

Rural Vestfold
Affluence and Gini
.2
.3
.4
.5
.6
.7
.8
1870 1880 1890 1900 1910 1920 1930 1940 1950 1960 1970 1980 1990 2000 2010 2020
Year
Affluence
Gini (entire population)
Gini (pop. above median)

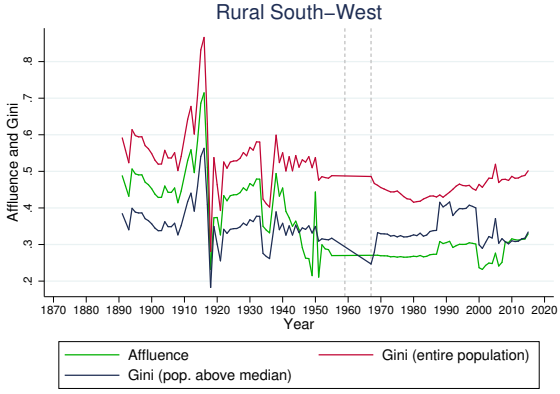

Urban Telemark
Affluence and Gini
.2
.3
.4
.5
.6
.7
.8
1870 1880 1890 1900 1910 1920 1930 1940 1950 1960 1970 1980 1990 2000 2010 2020
Year
Affluence
Gini (entire population)
Gini (pop. above median)

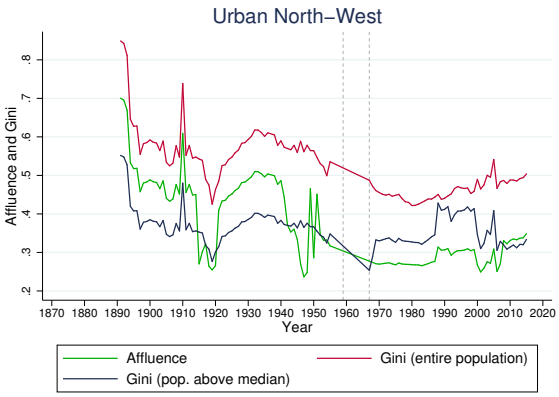

Rural Telemark
Affluence and Gini
.2
.3
.4
.5
.6
.7
.8
1870 1880 1890 1900 1910 1920 1930 1940 1950 1960 1970 1980 1990 2000 2010 2020
Year
Affluence
Gini (entire population)
Gini (pop. above median)

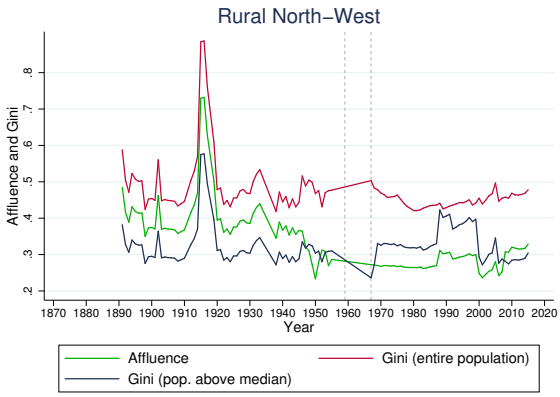

Urban South
Affluence and Gini
.2
.3
.4
.5
.6
.7
.8
1870 1880 1890 1900 1910 1920 1930 1940 1950 1960 1970 1980 1990 2000 2010 2020
Year
Affluence
Gini (entire population)
Gini (pop. above median)

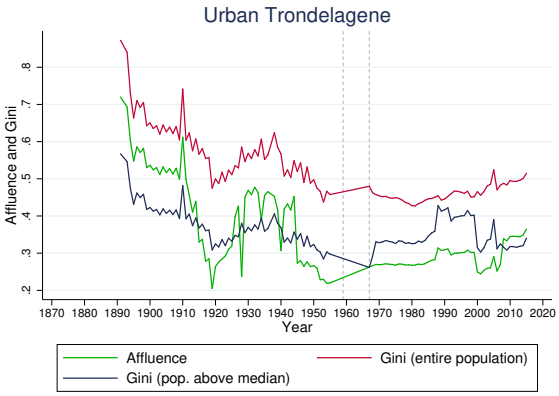

Rural South
Affluence and Gini
.2
.3
.4
.5
.6
.7
.8
1870 1880 1890 1900 1910 1920 1930 1940 1950 1960 1970 1980 1990 2000 2010 2020
Year
Affluence
Gini (entire population)
Gini (pop. above median)

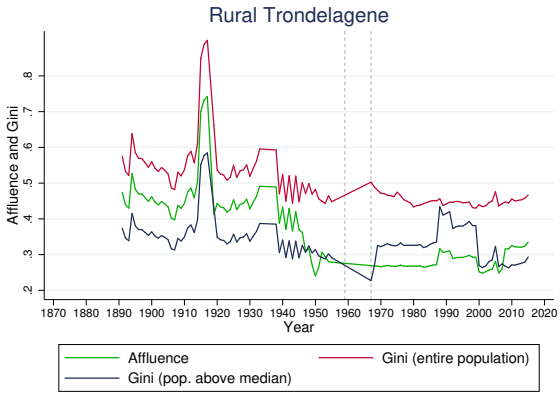

Urban South−West
Affluence and Gini
.2
.3
.4
.5
.6
.7
.8
1870 1880 1890 1900 1910 1920 1930 1940 1950 1960 1970 1980 1990 2000 2010 2020
Year
Affluence
Gini (entire population)
Gini (pop. above median)

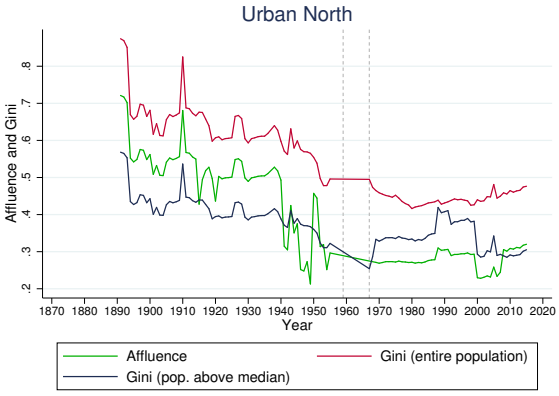

Rural South−West
Affluence and Gini
.2
.3
.4
.5
.6
.7
.8
1870 1880 1890 1900 1910 1920 1930 1940 1950 1960 1970 1980 1990 2000 2010 2020
Year
Affluence
Gini (entire population)
Gini (pop. above median)

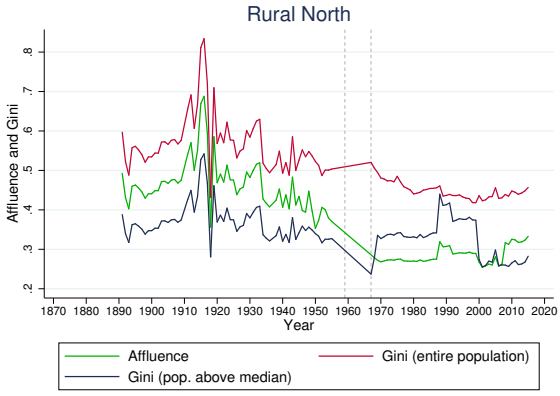

Urban North–West
Affluence and Gini
.2 .3 .4 .5 .6 .7 .8
1870 1880 1890 1900 1910 1920 1930 1940 1950 1960 1970 1980 1990 2000 2010 2020
Year
Affluence
Gini (entire population)
Gini (pop. above median)


Rural North–West
Affluence and Gini
.2 .3 .4 .5 .6 .7 .8
1870 1880 1890 1900 1910 1920 1930 1940 1950 1960 1970 1980 1990 2000 2010 2020
Year
Affluence
Gini (entire population)
Gini (pop. above median)


Urban Trondelagene
Affluence and Gini
.2 .3 .4 .5 .6 .7 .8
1870 1880 1890 1900 1910 1920 1930 1940 1950 1960 1970 1980 1990 2000 2010 2020
Year
Affluence
Gini (entire population)
Gini (pop. above median)


Rural Trondelagene
Affluence and Gini
.2 .3 .4 .5 .6 .7 .8
1870 1880 1890 1900 1910 1920 1930 1940 1950 1960 1970 1980 1990 2000 2010 2020
Year
Affluence
Gini (entire population)
Gini (pop. above median)


Urban North
Affluence and Gini
.2 .3 .4 .5 .6 .7 .8
1870 1880 1890 1900 1910 1920 1930 1940 1950 1960 1970 1980 1990 2000 2010 2020
Year
Affluence
Gini (entire population)
Gini (pop. above median)


Rural North
Affluence and Gini
.2 .3 .4 .5 .6 .7 .8
1870 1880 1890 1900 1910 1920 1930 1940 1950 1960 1970 1980 1990 2000 2010 2020
Year
Affluence
Gini (entire population)
Gini (pop. above median)

7 Conclusion

This thesis has shown a convergence of social welfare between regions in Norway from 1875-2015. This conclusion is derived from estimating the poor, the population not registered in poverty or tax statistics, inequality within taxpayers for historical data, the mean income and the Gini coefficient for the entire population. By first determining the development of the rural/urban divide, we have been able to build on previous research and different sources of data to investigate the development of regional inequality and social welfare.

The empirical findings may be summarized in three major conclusions. Firstly, there is a convergence in welfare levels when moving from the past towards the present time when comparing Norwegian regions. The second conclusion is that Norway was not an egalitarian society at the end of the 19th century, as commonly perceived. Regional inequality was as high as 0.8 measured in the Gini coefficient. Declining inequality from the end of the 19th century until 1980 was mainly driven by a decrease in the upper tail Gini coefficient. The third conclusion is that the development in Norwegian regions have been widely different over the last 140 years, but at present are relatively similar. This pattern may be attributed to the expansion of the welfare state.

In contrast to much of the existing work on the evolution of regional income inequality, this thesis suggests that Norwegian regions did not show a regional inverted U pattern proposed by Williamson (1965). The development of the different regions did not follow a similar nor specific pattern, but instead evolved differently due to several factors, consistent with the findings of Aaberge et al. (2017). To determine these factors, I have used historical evidence and economic theory, combined them with estimates of inequality, to provide estimates of the development of regional inequality and social welfare.

Bibliography

Aaberge, R. (1982). On the measurement of inequality. in norwegian (om måling av ulikskap). *Rapporter 82/9*.

Aaberge, R. (2000). Characterizations of lorenz curves and income distributions. *Social Choice and Welfare 17*.

Aaberge, R. (2001). Axiomatic characterization of the gini coefficient and lorenz curve orderings. *Journal of Economic Theory, 101*.

Aaberge, R. (2007). Gini's nuclear family. *J. Econ Inequal*.

Aaberge, R. and Atkinson, A. (2010). Top incomes in norway. *Top incomes in a global perspective, Oxford University Press*.

Aaberge, R., Atkinson, A., and Modalsli, J. (2017). On the measurement of long-run income inequality: Empirical evidence from norway, 1875-2013. *IZA Discussion Paper*.

Aaberge, R., Atkinson, A., and Sigstad, H. (2015). Income poverty, affluence and polarisation viewed from the median. *Centre for Analysis of Social Exclusion, London School of Economics*.

Acemoglu, D. and Robinson, J. (2002). The political economy of the kuznets curve. *Review of Development Economics*.

Almaas, I., Cappelen, A. W., Lind, J. T., Sorensen, E. O., and Tungodden, B. (2014). Skills, education, and the rise of earnings inequality among the "other 99 percent". *Science vol 344 no.6186*.

Alstadsæter, A. and Fjærli, E. (2009). Neutral taxation of shareholder income? corporate responses to an announced dividend tax. *Int Tax Public Finance 16:571*.

Atkinson, A. (1970). On the measurement of inequality. *Journal of Economic Theory, 2 (3)*.

Atkinson, A. B. and Piketty, T. (2007). Top incomes over the 20th century. *Oxford University Press Inc.*

Atkinson, A. B. and Søgaard, J. E. (2015). The long-run history of income inequality in denmark. *The Scandinavian Journal of Economics*.

Atkinson, A. and Bourguignon, F. (1989). The design of direct taxation and family benefits. *The Journal of Public Economics 41*.

Autor, D. H. (2014). Skills, education, and the rise of earnings inequality among the "other 99 percent". *Science vol 344 no.6186*.

Badia-Miró, M., Guilera, J., and Lains, P. (2012). Regional incomes in portugal: industrialisation, integration and inequality, 1890-1980. *Revista de Historia Económica, 30(2)*.

Banerjee, A. and Duflo, E. (2007). The economic lives of the poor. *Journal of Economic Perspectives 21(1)*.

Bergh, T., Hanisch, T., Lange, E., and Pharo, H. (1983). Norge fra u-land til i-land; vekst og utviklingslinjer 1830-1980. *Gyldendal Norsk Forlag A/s*.

Combes, P.-R., Lafourcade, M., Thisse, J., and Toutain., J.-C. (2011). The rise and fall of spatial inequalities in france: A long-run perspective. *Explorations in Economic History, 48, (2)*.

Deininger, K. and Squire., L. (1996). A new data set measuring income inequality. *World Bank Economic Review 10(3)*.

Donaldson, D. and Weymark, J. (1980). A single-parameter generalization of the gini indices of inequality. *Journal of Economic Theory, 22(1)*.

Ebert, U. (1987). Size and distribution of incomes as determinants of social welfare. *Journal of Economic Theory, 41(1)*.

Enflo, K. and Rosès, J. (2015). Coping with regional inequality in sweden: structural change, migrations, and policy, 1860–2000. *Economic History Review, 68*.

Felice, E. (2011). Regional value added in italy, 1891–2001, and the foundation of a long-term picture. *Economic History Review*.

Firpo, S., Fortin, N., and Lemieux, T. (2011). Decomposition methods in economics. *Handbook of Labor Economics*.

Gerdrup, K. (1998). Skattesystem og skattestatistikk i et historisk perspektiv. *Reports Statistics Norway.*

Gottschalk, P. and Smeeding, T. (2000). Empirical evidence on income inequality in industrialized countries. *Handbook of Income Distribution, 1:261–307.*

Harris, J. R. and Todaro, M. P. (1970). Migration, unemployment and development: A two-sector analysis. *The American Economic Review, Vol. 60, No. 1, pp. 126-142.*

Hodne, F. (1981). Norges økonomiske historie 1815-1970. *JW Cappelens forlag AS.*

Hodne, F. and Grytten, O. H. (1992). Norsk økonomi 1900-1990. *Tano Aschehoug.*

Juvkvam, D. (1999). Historisk oversikt over endringer i kommune- og fylkesinndelingen. *Reports Statistics Norway 99/13.*

Kuznets, S. (1955). Economic growth and income inequality. *The American Economic Review.*

Lindhert, P. and Williamson, J. (1985). Growth, equality and history. *Harvard Institute for Economic Research.*

Lund, M. (2012). Inntektsfordelingen i norge i et historisk perspektiv. *Master's thesis, UIO.*

Martínez-Galarraga, J., Roses, J. R., and Tirado, D. A. (2013). The long-term patterns of regional income inequality in spain, 1860–2000. *Regional Studies. pp. 1-16.*

Milanovic, B. (2016). Global inequality. a new approach for the age of globalization. *Harvard University Press.*

Mjelve, S. (1998). Økonomisk vekst og fordeling av inntekt i byene i vest-agder og Østfold, 1840-1990. *Oslo-Kongsvinger: Statistics Norway.*

Modalsli, J. (2017). The regional dispersion of income inequality in nineteenth-century norway. *Explorations in Economic History, forthcoming.*

NSD Kommunedatabase (2017a). Befolkning 1769-2015. (in norwegian, population 1769-2015). *https://trygg.nsd.uib.no/kdbbin/kategorier.exe?katalog=ok,.*

NSD Kommunedatabase (2017b). Fattigstatistikk 1875-1920. (in norwegian, poverty statistics 1875-1920). *https://trygg.nsd.uib.no/kdbbin/kategorier.exe?katalog=ok,.*

Rodriguez, J., Rodriguez, J., and Salas, R. (2010). Income mobility and economic inequality from a regional perspective. *Journal of Applied Economics*.

Sen, A. (1974). Informational bases of alternative welfare approaches. *Journal of Public Economics 3*.

Sen, A. (1979). Real national income. *The Review of Economic Studies, Vol. 43, No. 1*.

Sobye, E. (2012). Historien bak aldrende befolkning. *Samfunnsspeilet, 2/2012:6–13*.

Solbu, P. (2009). Inntektsulikhet i norge 1858-2006. *Master's thesis, UIO*.

Soltow, L. (1965). Toward income inequality in norway. *The University of Wisconsin Press, Madison-Milwaukee*.

Statistics Norway (2017a). Historisk statistikk; fattigvesen og sosialhjelpstatistikk 1866-1974. *https://www.ssb.no/a/histstat/publikasjoner/*.

Statistics Norway (2017b). Historisk statistikk; skattestatistikk, 1875-1966. *https://www.ssb.no/a/histstat/publikasjoner/*.

Statistics Norway (2017c). Personal tax data, 1967-2015. *https://www.ssb.no/a/histstat/publikasjoner/*.

Stiglitz, J., Sen, A., and Fitoussi, J. (2009). Report of the sen-stiglitz-fitoussi commission. *Report by the Commission on the Measurement of Economic Performance and Social Progress*.

Stiglitz, J. E. (2012). The price of inequality. *Penguin Books Ltd*.

Strøm Fjære, J. (2014). Income development in norwegian municipalities. *Master's thesis, UIO*.

Thomas, M. (1991). Income distribution in historical perspective, the evolution of inequality in australia in the nineteenth century. *Cambridge University Press*.

Williamson, J. (1965). Regional inequality and the process of national development: A description of the patterns. *Economic Development and Cultural Change 13(4)*.

Williamson, J. (1991). Income distribution in historical perspective, bristish inequality during the industrial revolution: accounting for the kuznets curve. *Cambridge University Press*.

Yaari, M. (1988). A controversial proposal concerning inequality measurement. *Journal of Economic Theory 44*.

A An overview of data sources

A.1 An overview of data sources for the Urban and Rural division

MT : Municipal Tax Statistics
ST : State Tax Statistics
PC : Population Census
PS : Poverty Statistics. Both number and income of poor.
NP : Number of poor only.
PTD : Personal Tax Data
PR : Population Register
SB : Statistic Bank

Table 8: Sources of data from 1875-1900

Year	Mean Income	Income Inequality	Poverty	Population	Income Tabulations
1875			NP 1875	PC 1875	MT
1876			NP 1876		
1877			NP 1877		
1878			NP 1878		
1879			NP 1879		
1880			NP 1880		
1881			NP 1881		
1882			NP 1882		
1883			NP 1883		
1884	MT		PS 1884		
1885	MT		PS 1885		
1886	MT		PS 1886		
1887	MT		PS 1887		
1888	MT		PS 1888		MT
1889	MT		PS 1889		
1890	MT		PS 1890		
1891	MT		PS 1891	PC 1891	
1892	MT	MT, ST, PS, PC	PS 1892		ST
1893	MT	MT, ST, PS, PC	PS 1893		ST
1894	MT	MT, ST, PS, PC			ST
1895	MT	MT, ST, PS, PC	PS 1895		ST
1896	MT	MT, ST, PS, PC			ST
1897	MT	MT, ST, PS, PC			ST
1898	MT	MT, ST, PS, PC			ST
1899	MT	MT, ST, PS, PC			ST
1900	MT	MT, ST, PS, PC	PS 1900	PC 1900	ST

Table 9: Sources of data from 1900-1966

Year	Mean Income	Income Inequality	Poverty	Population	Income Tabulations
1901	MT	MT, ST, PS, PC			ST
1902	MT	MT, ST, PS, PC			ST
1903	MT	MT, ST, PS, PC			ST
1904	MT	MT, ST, PS, PC			
1905	MT	MT, ST, PS, PC	PS 1905		
1906	MT	MT, ST, PS, PC			MT,ST
1907	MT	MT, ST, PS, PC			
1908	MT	MT, ST, PS, PC			
1909	MT	MT, ST, PS, PC			
1910	MT	MT, ST, PS, PC	PS 1910	PC 1910	
1911	MT	MT, ST, PS, PC			ST
1912	MT	MT, ST, PS, PC			
1913	MT	MT, ST, PS, PC			MT,ST
1914	MT	MT, ST, PS, PC			
1915	MT	MT, ST, PS, PC	PS 1915		
1916	MT	MT, ST, PS, PC			
1917	MT	MT, ST, PS, PC			
1918	MT	MT, ST, PS, PC			
1919	MT	MT, ST, PS, PC			
1920	MT	MT, ST, PS, PC	PS 1920	PC 1920	
1921	MT	MT, ST, PS, PC	PS 1921		
1922	MT	MT, ST, PS, PC	PS 1922		
1923	MT	MT, ST, PS, PC	PS 1923		
1924	MT	MT, ST, PS, PC	PS 1924		
1925	MT	MT, ST, PS, PC	PS 1925		
1926	MT	MT, ST, PS, PC	PS 1926		
1927	MT	MT, ST, PS, PC	PS 1927		
1928	MT	MT, ST, PS, PC	PS 1928		
1929	MT	MT, ST, PS, PC	PS 1929		MT,ST
1930	MT	MT, ST, PS, PC	PS 1930	PC 1930	
1931	MT	MT, ST, PS, PC	PS 1931		
1932	MT	MT, ST, PS, PC	PS 1932		
1933	MT	MT, ST, PS, PC	PS 1933		
1934	MT	MT, ST, PS, PC	PS 1934		
1935			PS 1935		
1936			PS 1936		
1937	MT	MT, ST, PS, PC	PS 1937		
1938	MT	MT, ST, PS, PC	PS 1938		MT*
1939	MT	MT, ST, PS, PC	PS 1939		
1940	MT	MT, ST, PS, PC	PS 1940		
1941	MT	MT, ST, PS, PC			
1942	MT	MT, ST, PS, PC			
1943	MT	MT, ST, PS, PC			
1944	MT	MT, ST, PS, PC			
1945	MT	MT, ST, PS, PC			
1946	MT	MT, ST, PS, PC	PS 1946	PC 1946	
1947	MT	MT, ST, PS, PC	PS 1947		
1948	MT	MT, ST, PS, PC	PS 1948		ST
1949			PS 1949		ST
1950	MT	MT, ST, PS, PC	PS 1950	PC 1950	ST
1951	MT	MT, ST, PS, PC	PS 1951	PC 1951	ST
1952	MT	MT, ST, PS, PC	PS 1952	PC 1952	MT
1953	MT	MT, ST, PS, PC	PS 1953	PC 1953	MT
1954	MT	MT, ST, PS, PC	PS 1954	PC 1954	MT
1955	MT	MT, PS, PC	PS 1955	PC 1955	MT
1956			PS 1956	PC 1956	
1957	MT	MT, PS, PC	PS 1957	PC 1957	
1958			PS 1958	PC 1958	
1959			PS 1959	PC 1959	
1960	MT	MT, PS, PC	PS 1960	PC 1960	MT,ST
1961	MT	MT, PS, PC	PS 1961	PC 1961	MT,ST
1962			PS 1962	PC 1962	MT,ST
1963			PS 1963	PC 1963	MT,ST
1964			PS 1964	PC 1964	MT,ST
1965	MT	MT, PS, PC		PC 1965	MT,ST
1966					MT,ST

A.2 Overview of data sources for regions

MT : Municipal Tax Statistics
ST : State Tax Statistics
PC : Population Census
PS : Poverty Statistics. Both number and income of poor.
NP : Number of poor only.
PTD : Personal Tax Data
PR : Population Register
SB : Statistic Bank

Table 10: Sources of data from 1875-1900

Year	Mean Income	Income Inequality	Poverty	Population	Income Tabulations
1875			NP 1875	PC 1875	
1876			NP 1876		
1877			NP 1877		
1878			NP 1878		
1879			NP 1879		
1880			NP 1880		
1881			NP 1881		
1882			NP 1882		
1883			NP 1883		
1884	MT		NP 1884		
1885	MT		NP 1885		
1886	MT		NP 1886		
1887	MT		NP 1887		
1888	MT		NP 1888		
1889	MT		NP 1889		
1890	MT		NP 1890		
1891	MT		NP 1891	PC 1891	
1892	MT	MT, ST, PS, PC	NP 1892		ST
1893	MT	MT, ST, PS, PC	NP 1893		ST
1894	MT	MT, ST, PS, PC			ST
1895	MT	MT, ST, PS, PC	NP 1895		ST
1896	MT	MT, ST, PS, PC			ST
1897	MT	MT, ST, PS, PC			ST
1898	MT	MT, ST, PS, PC			ST
1899	MT	MT, ST, PS, PC			ST
1900	MT	MT, ST, PS, PC	PS 1900	PC 1900	ST
1901	MT	MT, ST, PS, PC			ST
1902	MT	MT, ST, PS, PC			ST
1903	MT	MT, ST, PS, PC			ST
1904	MT	MT, ST, PS, PC			
1905	MT	MT, ST, PS, PC	NP 1905		
1906	MT	MT, ST, PS, PC			
1907	MT	MT, ST, PS, PC			
1908	MT	MT, ST, PS, PC			
1909	MT	MT, ST, PS, PC			
1910	MT	MT, ST, PS, PC	NP 1910	PC 1910	

Table 11: Sources of data from 1900-1950

Year	Mean Income	Income Inequality	Poverty	Population	Income Tabulations
1911	MT	MT, ST, PS, PC			ST
1912	MT	MT, ST, PS, PC			
1913	MT	MT, ST, PS, PC			
1914	MT	MT, ST, PS, PC			
1915	MT	MT, ST, PS, PC	NP 1915		
1916	MT	MT, ST, PS, PC			
1917	MT	MT, ST, PS, PC			
1918	MT	MT, ST, PS, PC			
1919	MT	MT, ST, PS, PC			
1920	MT	MT, ST, PS, PC	NP 1920	PC 1920	
1921	MT	MT, ST, PS, PC			
1922	MT	MT, ST, PS, PC			
1923	MT	MT, ST, PS, PC			
1924	MT	MT, ST, PS, PC			
1925	MT	MT, ST, PS, PC			
1926	MT	MT, ST, PS, PC			
1927	MT	MT, ST, PS, PC			
1928	MT	MT, ST, PS, PC			
1929	MT	MT, ST, PS, PC			ST
1930	MT	MT, ST, PS, PC		PC 1930	
1931	MT	MT, ST, PS, PC			
1932	MT	MT, ST, PS, PC			
1933	MT	MT, ST, PS, PC			
1934	MT	MT, ST, PS, PC			
1935			PS 1935		
1936			PS 1936		
1937	MT	MT, ST, PS, PC			
1938	MT	MT, ST, PS, PC	PS 1938		
1939	MT	MT, ST, PS, PC	PS 1939		
1940	MT	MT, ST, PS, PC	NP 1940		
1941	MT	MT, ST, PS, PC	NP 1941		
1942	MT	MT, ST, PS, PC			
1943	MT	MT, ST, PS, PC			
1944	MT	MT, ST, PS, PC			
1945	MT	MT, ST, PS, PC			
1946	MT	MT, ST, PS, PC	NP 1946	PC 1946	
1947	MT	MT, ST, PS, PC	NP 1947		
1948	MT	MT, ST, PS, PC	PS 1948		
1949			PS 1949		
1950	MT	MT, ST, PS, PC	PS 1950	PC 1950	
1951	MT	MT, ST, PS, PC	PS 1951	PC 1951	
1952	MT	MT, ST, PS, PC	PS 1952	PC 1952	
1953	MT	MT, ST, PS, PC	PS 1953	PC 1953	
1954	MT	MT, ST, PS, PC	PS 1954	PC 1954	
1955	MT	MT, PS, PC	PS 1955	PC 1955	
1956			PS 1956	PC 1956	
1957	MT	MT, PS, PC	PS 1957	PC 1957	
1958			PS 1958	PC 1958	
1959			PS 1959	PC 1959	
1960	MT	MT, PS, PC	PS 1960	PC 1960	
1961	MT	MT, PS, PC	PS 1961	PC 1961	
1962			PS 1962	PC 1962	
1963			PS 1963	PC 1963	
1964			PS 1964	PC 1964	
1965	MT	MT, PS, PC		PC 1965	
1966					
1967					
1968	MT	MT, PS, PC			

B A detailed overview of the changes in Norwegian Municipalities

RTU : Rural municipality changed to Urban

UTR : Urban municipality changed to Rural

Table 12: Changes made in Municipality status

From Municipality	To Municipality	Change in Status
Borge	Fredrikstad	RTU
Varteig	Fredrikstad	RTU
Onsøy	Fredrikstad	RTU
Rolvsøy	Fredrikstad	RTU
Torsnes	Fredrikstad	RTU
Glemmen	Fredrikstad	RTU
Varteig	Sarpsborg	RTU
Skjeberg	Sarpsborg	RTU
Tune	Sarpsborg	RTU
Jeløy	Moss	RTU
Berg	Halden	RTU
Idd	Halden	RTU
Aker	Oslo	RTU
Drøbak	Frogn	UTR
Son	Vestby	UTR
Hølen	Vestby	UTR
Vang	Hamar	RTU
Biri	Gjøvik	RTU
Snertingdal	Gjøvik	RTU
Fåberg	Sollien	RTU
Hønefoss	Ringerike	
Tyristrand	Ringerike	RTU

Table 12: Changes made in Municipality status

From Municipality	To Municipality	Change in Status
Hole	Ringerike	RTU
Nordrehov	Ringerike	RTU
Ådal	Ringerike	RTU
Åsgårdsstand	Borre	
Horten	Borre	
Svelvik	Svelvik	UTR
Andebu	Sandefjord	RTU
Stokke	Sandefjord	RTU
Sandar	Sandefjord	RTU
Skoger	Drammen	RTU
Stavern	Larvik	RTU
Tjølling	Larvik	RTU
Botne	Holestrand	RTU
Kragerø	Kragerø	UTR
Langesund	Bamble	UTR
Stathelle	Bamble	UTR
Brevik	Porsgrunn	RTU
Eidanger	Porsgrunn	RTU
Gransherad	Kongsberg	RTU
Lillesand	Lillesand	UTR
Søndeled	Risør	RTU
Landvik	Grimstad	RTU
Fjære	Grimstad	RTU
Tromøy	Arendal	RTU
Hisøy	Arendal	RTU
Moland	Arendal	RTU
Øyestad	Arendal	RTU
Hidra	Flekkefjord	RTU
Nes	Flekkefjord	RTU
Gyland	Flekkefjord	RTU
Bakke	Flekkefjord	RTU
Herad	Farsund	RTU

Table 12: Changes made in Municipality status

From Municipality	To Municipality	Change in Status
Spind	Farsund	RTU
Lista	Mandal	RTU
Holum	Mandal	RTU
Randesund	Kristiansand	RTU
Oddernes	Kristiansand	RTU
Tveit	Kristiansand	RTU
Helleland	Egersund	RTU
Heksestad	Egersund	RTU
Høyland	Sandnes	RTU
Hetland	Sandnes	RTU
Hølen	Sandnes	RTU
Madla	Stavanger	RTU
Skåre	Haugesund	RTU
Torvastad	Haugesund	RTU
Sogndal	Sokndal	UTR
Kopervik	Karmøy	UTR
Skudeneshavn	Karmøy	UTR
Laksevåg	Bergen	RTU
Fana	Bergen	RTU
Arna	Bergen	RTU
Åsane	Bergen	RTU
Årstad	Bergen	RTU
Eikefjord	Flora	RTU
Kinn	Flora	RTU
Borgund	Ålesund	RTU
Bolsøy	Molde	RTU
Veøy	Molde	RTU
Grip	Kristiansund	RTU
Bremsnes	Kristiansund	RTU
Frei	Kristiansund	RTU
Leinstrand	Trondheim	RTU
Bryneset	Trondheim	RTU

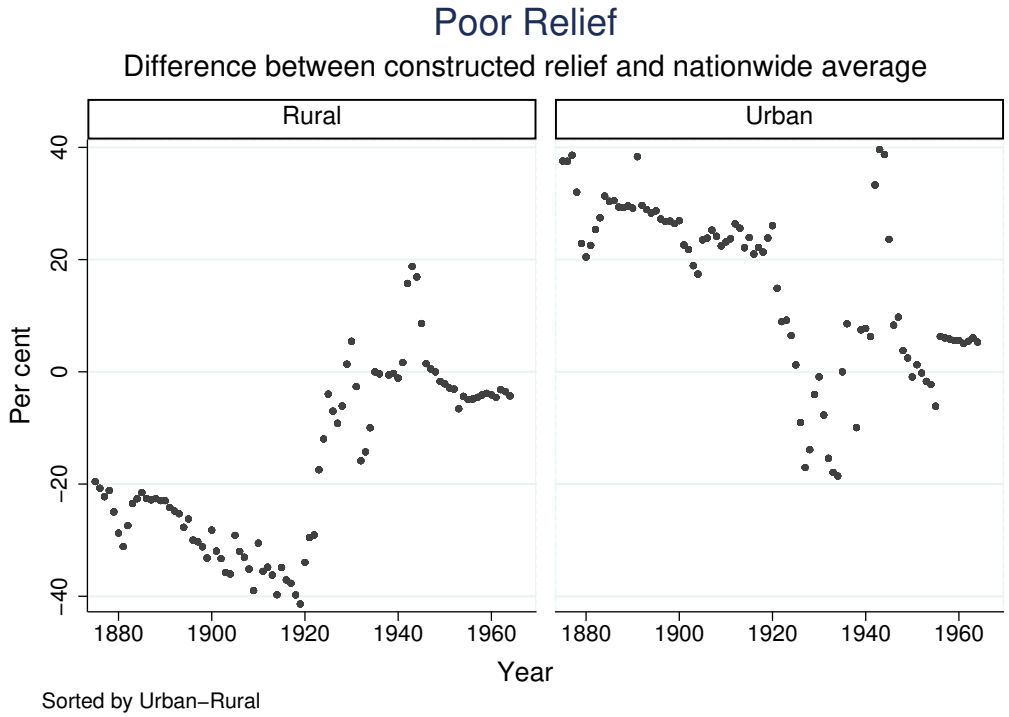
Table 12: Changes made in Municipality status

From Municipality	To Municipality	Change in Status
Strinda	Trondheim	RTU
Tiller	Trondheim	RTU
Beitstad	Steinkjer	RTU
Sparbu	Steinkjer	RTU
Ogndal	Steinkjer	RTU
Egge	Steinkjer	RTU
Stod	Steinkjer	RTU
Kvam	Steinkjer	RTU
Vedmundvik	Namsos	RTU
Klinga	Namsos	RTU
Mo	Rana	UTR
Skjerstad	Bodø	RTU
Bodin	Bodø	RTU
Ankenes	Narvik	RTU
Brønnøysund	Brønnøy	UTR
Mosjøen	Vefsn	UTR
Svolvær	Vågan	UTR
Bjarkøy	Harstad	RTU
Ullsfjord	Tromsø	RTU
Hillesøy	Tromsø	RTU
Tromsøysund	Tromsø	RTU
Sandtorg	Harstad	RTU
Trondenes	Harstad	RTU
Sørøysund	Hammerfest	RTU
Nord-Varanger	Vadsø	RTU

C Constructed Poor Relief

These scatter plots shows the difference between the constructed poor relief for each year and the national average for each year. It shows, quite naturally, that the poor relief is higher for the urban areas, and lower for the rural areas than the national average.

Figure 16: Poor Relief. Constructed vs National Average



There is an argument to be made that the poor relief is underestimated at the beginning of the 1920s. This thesis understands this argument, but at the same time, we believe that the difference between some rural and urban areas captures the variation between regions. This should be further tested on a municipality level to check for robustness.

D Ginis Nuclear Family

Aaberge (2007) introduces a wider range of measures for the Gini family. It could be interesting to see how the development of social welfare followed with introducing a wider range of Gini measures.

The upside of the measure used in this thesis is that the first two measures of the Lorenz curves gives us information about the shape of the curve and is, therefore, a measurement of inequality. The three first moments summarizes the essentials of the distribution function and, added together, will give information on the shape of the distribution of income.

The downside is that none of these moments are sensitive to changes in the lower part of the distribution. This is due to the convexity of the Lorenz curve. When comparing data over time, it is, therefore, preferable to supplement with other measures more sensitive to changes in the lower part of the distribution. This can be done by transforming the Lorenz curve to the M-curve, defined by Aaberge (1982) to the scaled conditional mean curve.

$$M(u) = \frac{E[X|X \leq F^{-1}(u)]}{\mu} = \begin{cases} \frac{1}{u\mu} \int_0^u F^{-1}(t)dt & \text{if } 0 < u \leq 1 \\ 0 & \text{if } u = 0 \end{cases} \tag{D.1}$$

When inserting for (3.2) in (3.4) the following simple relationship between the M-curve and the Lorenz-curve emerges,

$$M(u) = \begin{cases} \frac{L(u)}{u} & \text{if } 0 < u \leq 1 \\ 0 & \text{if } u = 0 \end{cases} \tag{D.2}$$

where M(1) = 1, and $\lim_{u\to\infty}(L(u)/u) = M(0)$. Thus, the M-curve is a representation of inequality that is equivalent to the Lorenz curve.

The M-curve possesses several attractive properties. It is always increasing, and it can be both concave and convex. It also satisfies the axioms required for a measurement of inequality. For a fixed u, M(u) is the ratio between the mean income of the poorest 100u per cent of the population and the overall mean. Thus, it yields essential information on poverty provided we know the poverty rate. The M-curve to a uniform distribution proves to be a diagonal line from (0,0) to (1,1), and thus represents

a useful reference. If a curve crosses the diagonal line one time from above, then this distribution will have a lower inequality compared to a uniform distribution (0, a) below the intersection and a higher inequality above the intersection. The M-curves is bounded by the unit square. Therefore visually, there is a sharper distinction between the two different M-curves than the corresponding Lorenz-curves. This distinction, brought forward by Atkinson, A. and Bourguignon, F. (1989), appears to be particularly visible at the lower parts of the income distributions.

As Aaberge (2007) writes, the Lorenz curve can give the following "family" of measures of inequality.

$$D_k(F) = \frac{1}{k}\left((k+1)\int_0^1 u^k dL(u) - 1]\right), k = 1, 2, \ldots, \tag{D.3}$$

This "family" included an infinite number of members, but the first three will give a good summary of the information given by the Lorenz-curve. However, these coefficients are not optimal as they focus on changes in the middle or the top end of the distribution. A more desirable option is, therefore, to use the first three moments of the scaled conditional mean curve as primary quantities for measuring inequality distribution. The k^{th} order moment of the scaled conditional mean curve is given by

$$C_k(F) = \int_0^1 u^k dM(u), \tag{D.4}$$

Equation (3.6) can also be written as

$$C_k(F) = k\int_0^1 u^{k-1}(1 - M(u))du, k = 1, 2, \ldots, \tag{D.5}$$

Equation (3.8) shows that the k^{th} moment is a sum, for all u, of weighted differences between the line of perfect equality and the M-curve. M-curves dominates other M-curves, by having the same mean but $M_1 > M_2$, will also give a lower C_k for all k. The first three moments, will reflect three different measures of inequality which all weight the different parts of the distribution differently. These three measures are described as following in Solbu (2009)

1. C_1 - Bonferroni-coefficient which add more weights to the lower part of the distribution
2. C_2 - Gini-coefficient which adds equal weight to the distribution

3. C_3 - which adds more weight to the upper part of the distribution

It is useful that both C_1 and C_3 adds information to the Gini coefficient by emphasizing changes in different parts of the distribution. This adds to the entire picture of the changes made in the historical perspective.

Comparing with the Lorenz-curve and its moments, the D_1 is the equivalent to the C_2. It is therefore clear that the Bonferroni is not included in the family of Lorenz curves. The reason for this is the convexity of the Lorenz curve, while the M-curve can be both concave and convex, allowing for a more nuanced interpretation.

E Gini with or without transfers

The thesis assumes that no one can live with zero income. the upper bound of the Gini coefficient is on the group receiving 50 per cent of the minimum pension for a single person. The lower bound is based on those not covered by the tax tabulations receiving mean income equal to 150 per cent of the mean income assumed for the upper bound.

How does this change if we assume that, yes, no one can live with zero income, but those who have zero income have a *relatively equal* underestimated income from the tax statistics as the rest of the population, meaning we should note them as zero.

The following figures show the changes in urban and rural Gini coefficients when changing the assumptions from 1967. The Gini coefficient with transfers is based on taking the mean of the upper and lower bound.

Figure 17: Transfers or no transfers for urban and rural areas. 1967 - present.

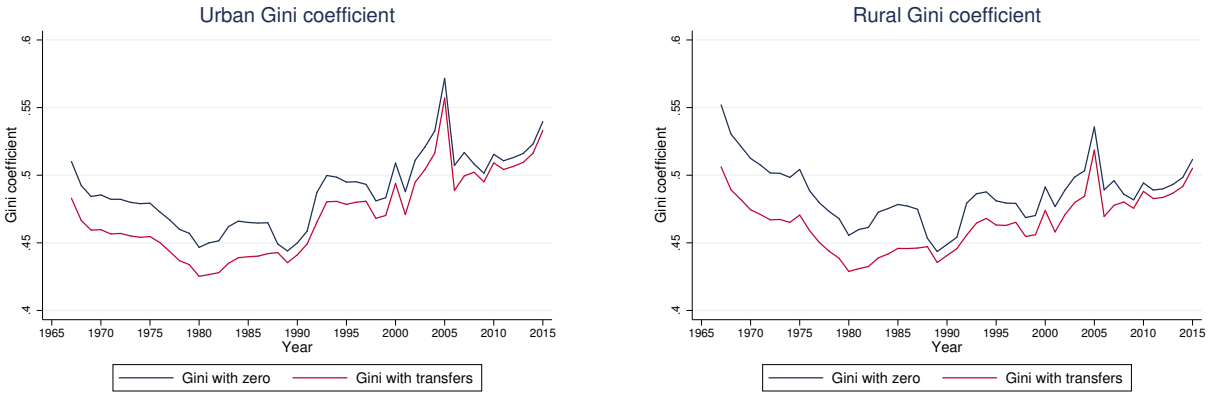
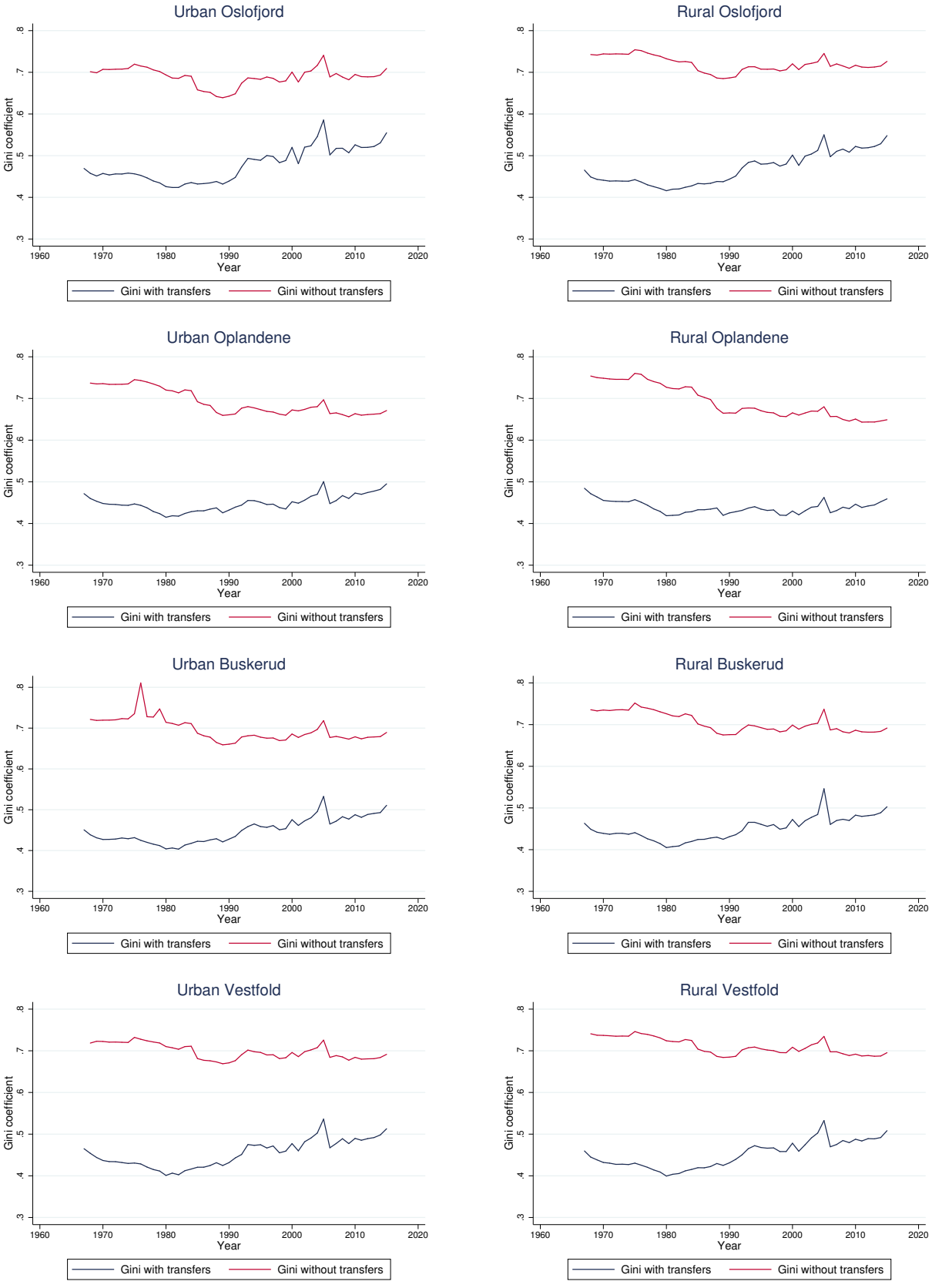
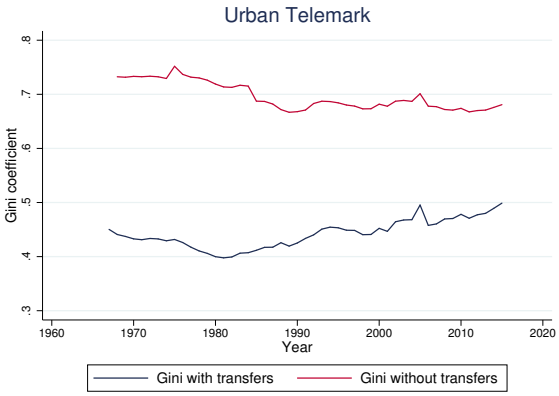
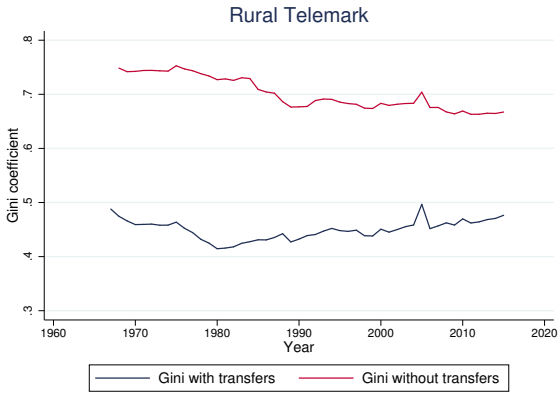
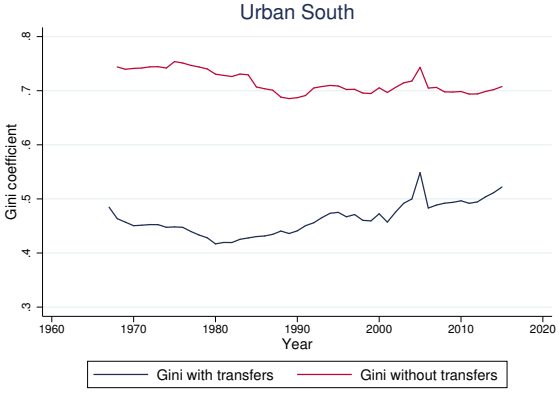


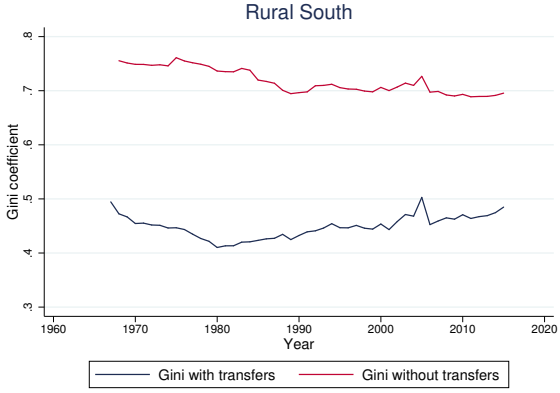
Figure 18: The Gini Coefficient with zero income or transfered income

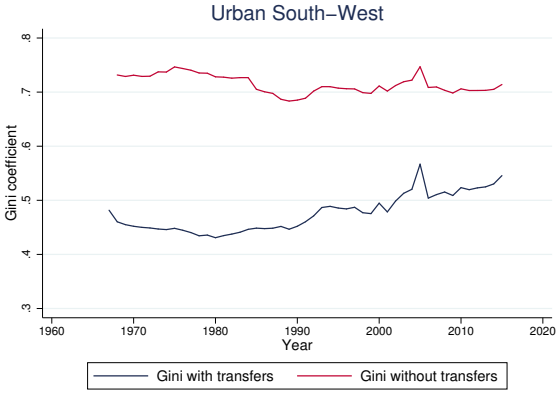


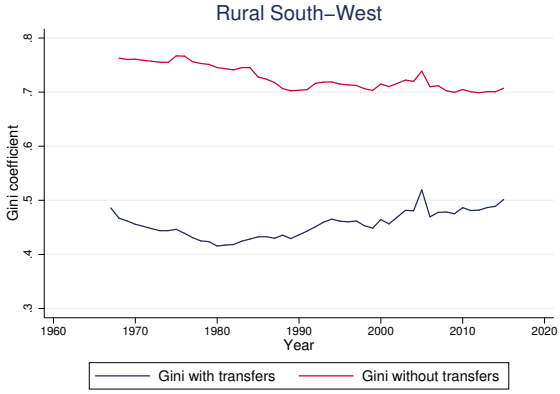

Urban Telemark
Gini coefficient
.3
.4
.5
.6
.7
.8
1960
1970
1980
1990
2000
2010
2020
Year
Gini with transfers
Gini without transfers

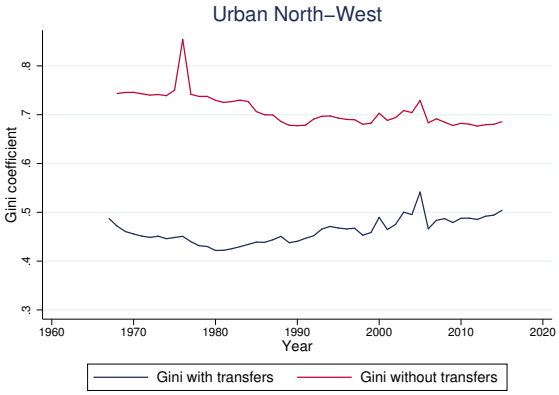

Rural Telemark
Gini coefficient
.3
.4
.5
.6
.7
.8
1960
1970
1980
1990
2000
2010
2020
Year
Gini with transfers
Gini without transfers

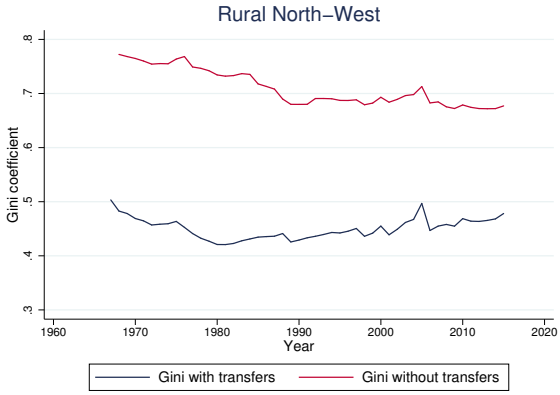

Urban South
Gini coefficient
.3
.4
.5
.6
.7
.8
1960
1970
1980
1990
2000
2010
2020
Year
Gini with transfers
Gini without transfers

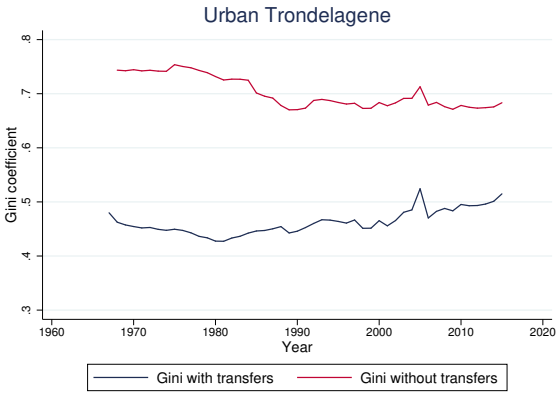

Rural South
Gini coefficient
.3
.4
.5
.6
.7
.8
1960
1970
1980
1990
2000
2010
2020
Year
Gini with transfers
Gini without transfers

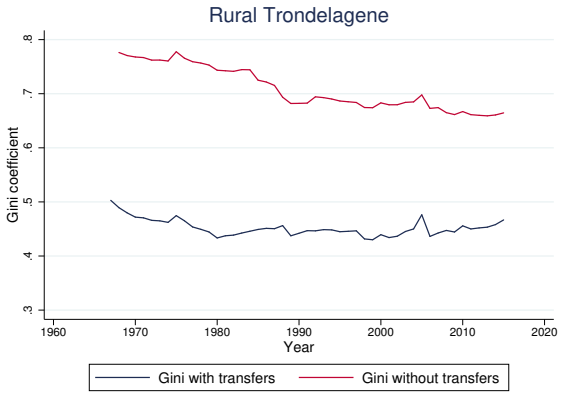

Urban South–West
Gini coefficient
.3
.4
.5
.6
.7
.8
1960
1970
1980
1990
2000
2010
2020
Year
Gini with transfers
Gini without transfers

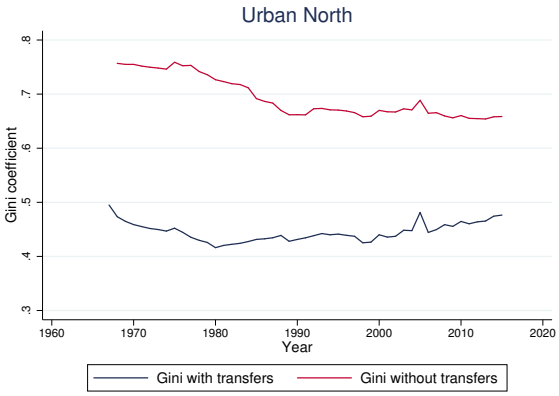

Rural South–West
Gini coefficient
.3
.4
.5
.6
.7
.8
1960
1970
1980
1990
2000
2010
2020
Year
Gini with transfers
Gini without transfers

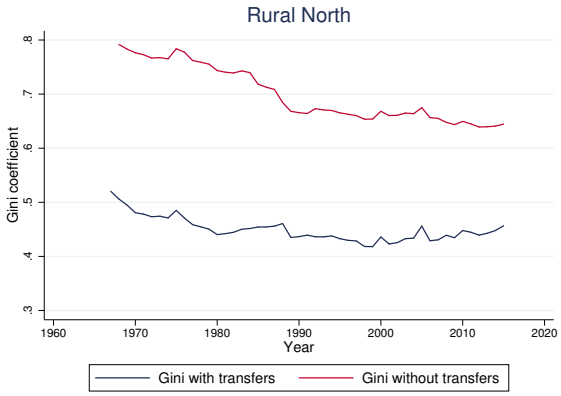

Urban North–West
Gini coefficient
.3
.4
.5
.6
.7
.8
1960
1970
1980
1990
2000
2010
2020
Year
Gini with transfers
Gini without transfers


Rural North–West
Gini coefficient
.3
.4
.5
.6
.7
.8
1960
1970
1980
1990
2000
2010
2020
Year
Gini with transfers
Gini without transfers


Urban Trondelagene
Gini coefficient
.3
.4
.5
.6
.7
.8
1960
1970
1980
1990
2000
2010
2020
Year
Gini with transfers
Gini without transfers


Rural Trondelagene
Gini coefficient
.3
.4
.5
.6
.7
.8
1960
1970
1980
1990
2000
2010
2020
Year
Gini with transfers
Gini without transfers


Urban North
Gini coefficient
.3
.4
.5
.6
.7
.8
1960
1970
1980
1990
2000
2010
2020
Year
Gini with transfers
Gini without transfers

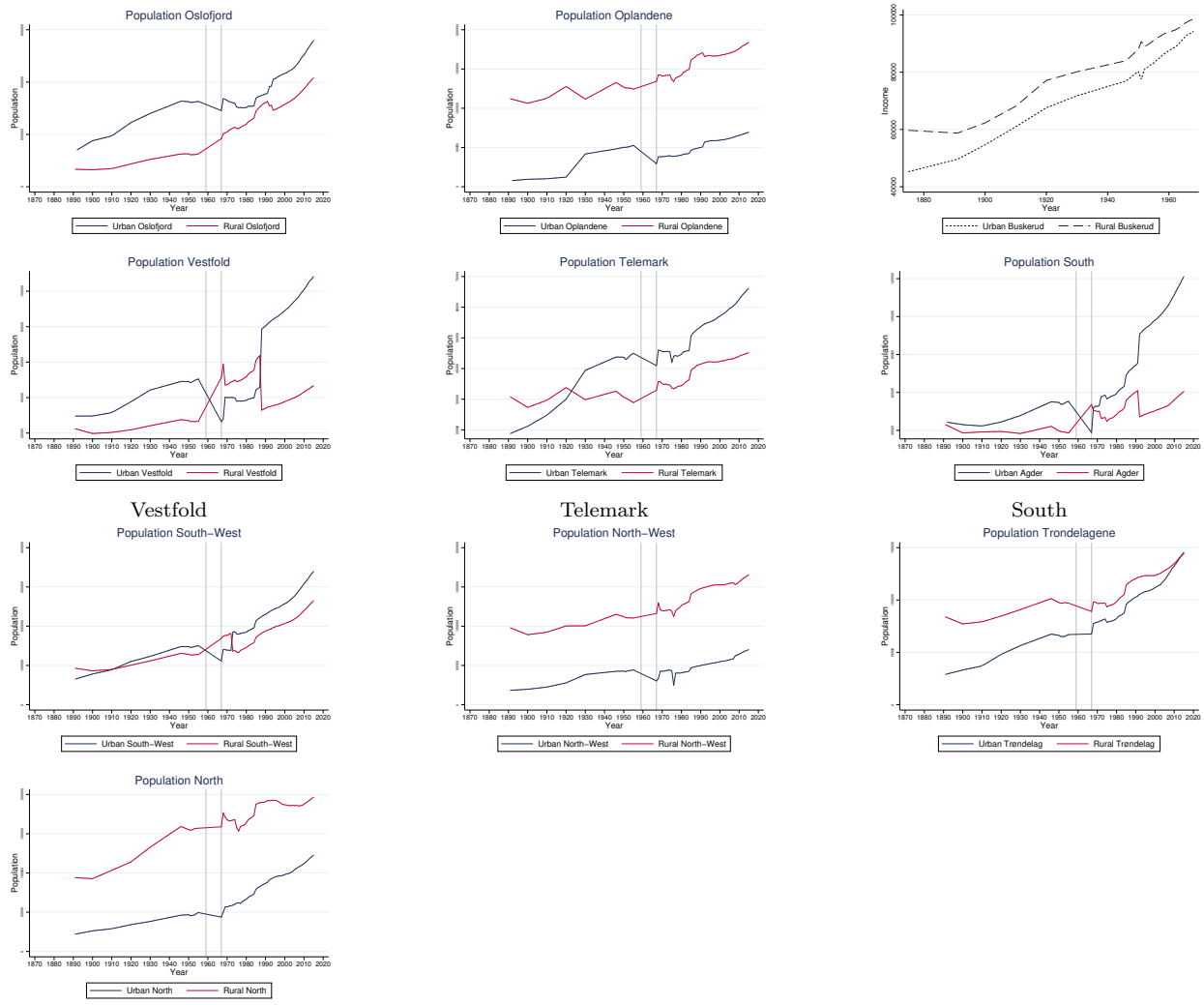

Rural North
Gini coefficient
.3
.4
.5
.6
.7
.8
1960
1970
1980
1990
2000
2010
2020
Year
Gini with transfers
Gini without transfers

F Figures describing Regional Development

F.1 Regional Population Development

These figures show the development of the population within regions. It is noteworthy that there was an increase in population in Oplandene region, coinciding with the increase in the Gini coefficient.

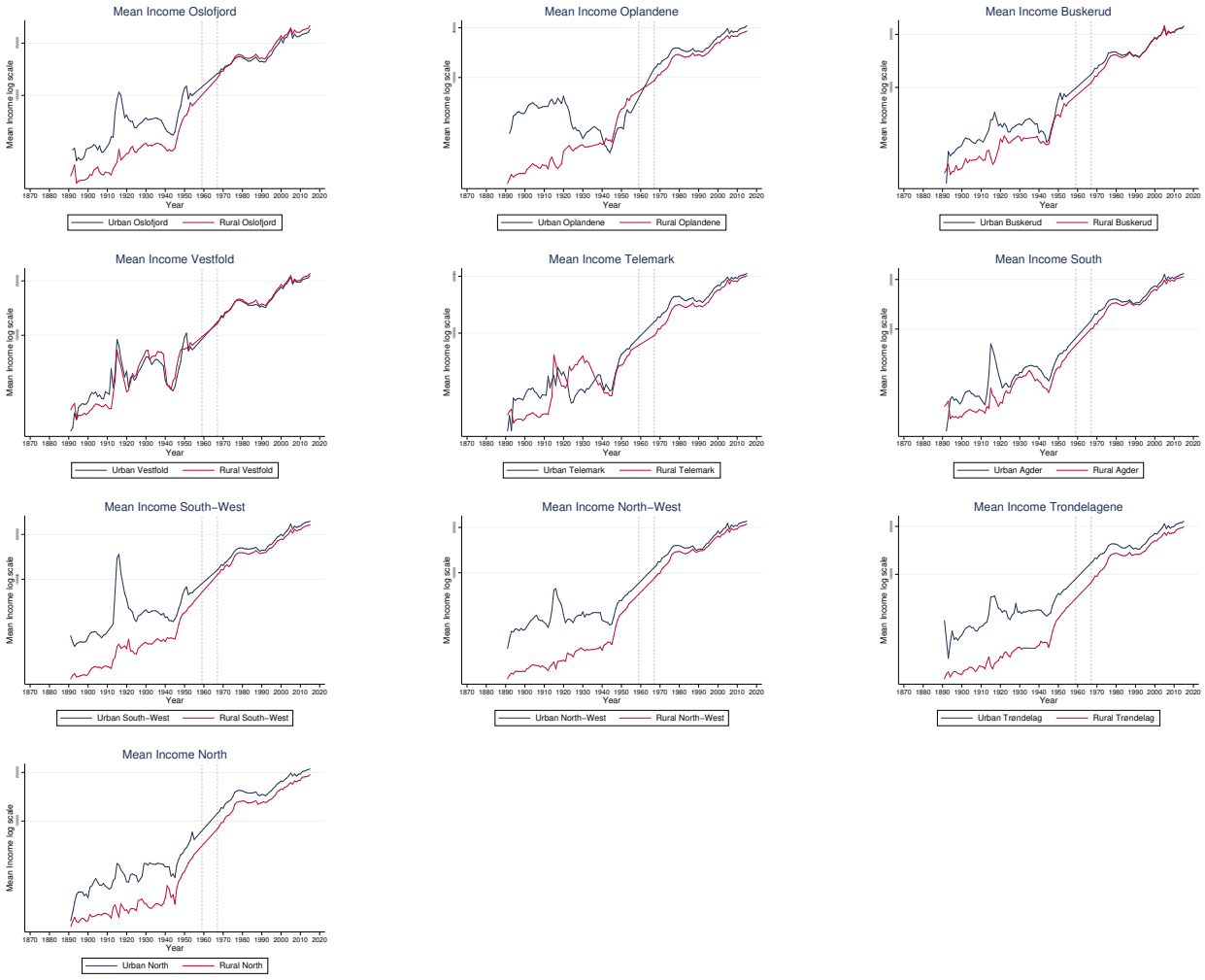
Figure 19: Regional Population Development 1891 - present. Core Households



F.2 The Development of Mean Income in the ten regions

The following figure shows the development of the CPI-adjusted mean income over time for the different ten regions. This should be viewed as a complement to the story told by figure XX in chapter 5. The main similarities are that the regions have experienced an exponential growth in the mean income over time. This is consistent with the national average, and also for several western countries since the beginning of the 20th century.

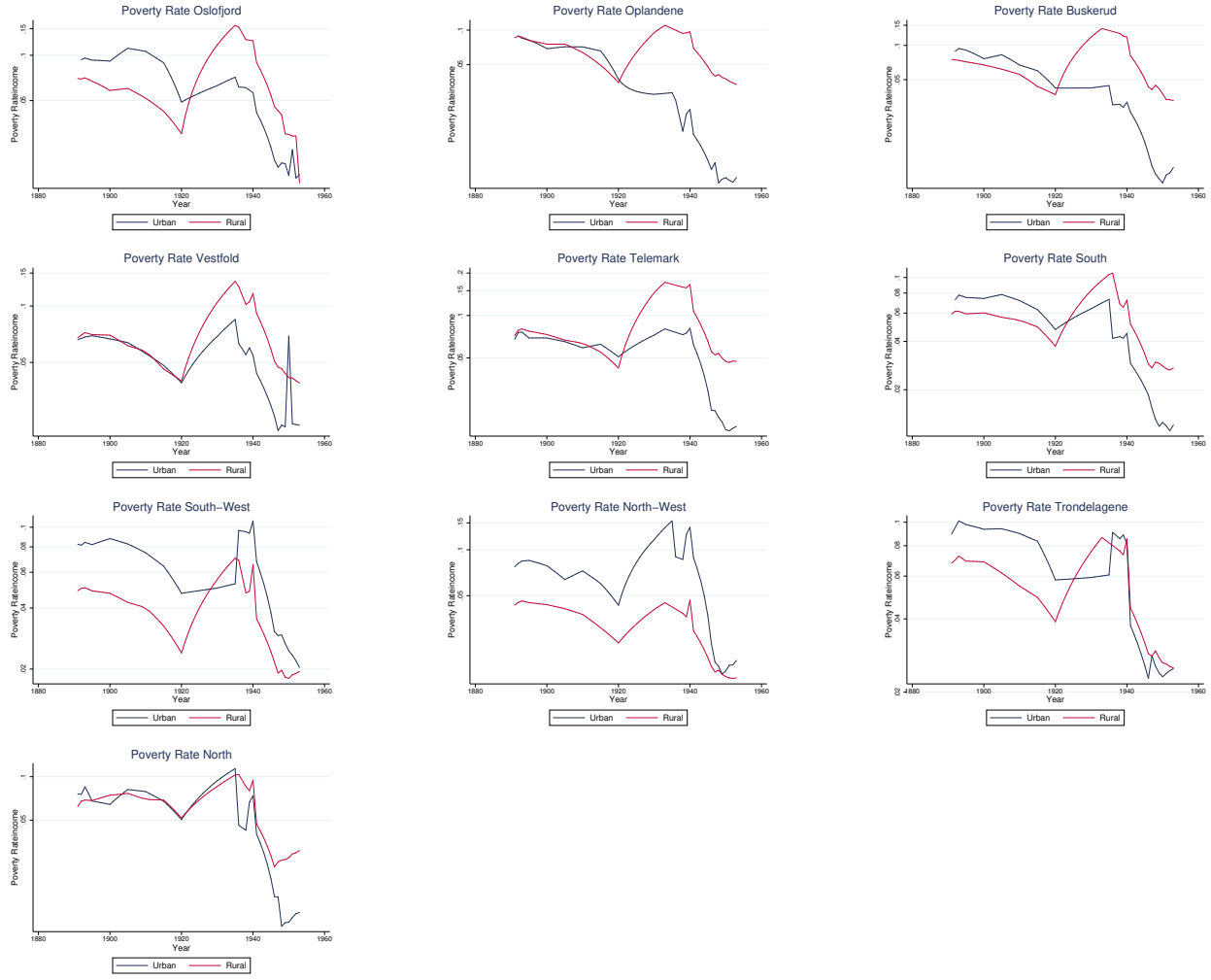
Figure 20: The Development of the Mean Income by regions 1891 - present. CPI adjusted



F.3 Poverty Rate in Regions

The following supplement the analysis in section 5.1, when discussing the development of the poverty rate in regions.

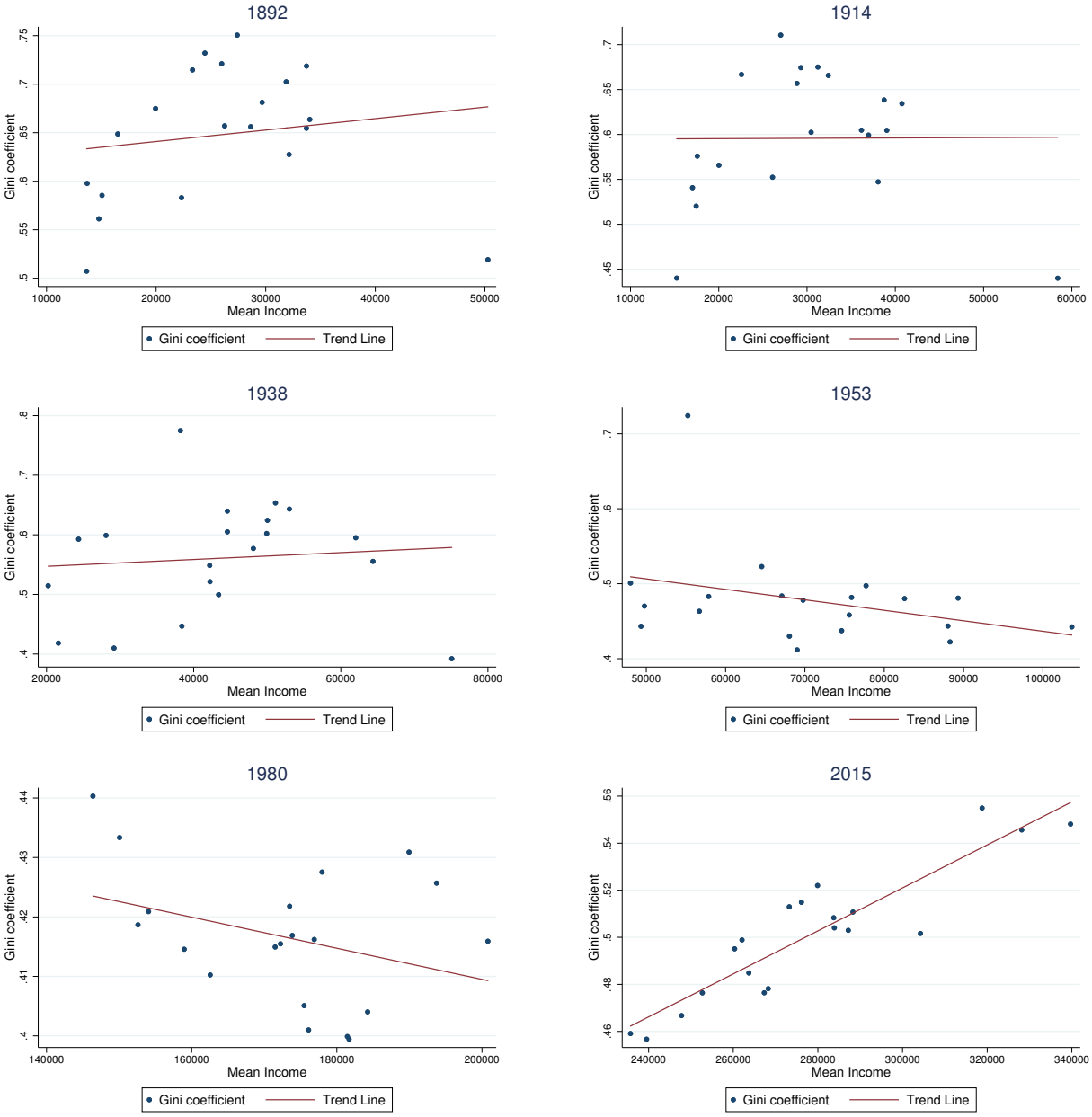
Figure 21: The Development of the Poverty rate by regions 1891 - 1964



F.4 Mean Income and Gini coefficient.

As we have discussed in Chapter 5, we see a convergence in social welfare between regions from 1980 to 2015, while the mean income has diverged for the same period. Since inequality has increased in this period as well, it is interesting to see how the two variables that estimate social welfare interacts. In the following figure, we see a weak negative interaction between the Gini coefficient and mean income in 1980. This suggests that regions with relatively low mean income were more unequal than regions with higher mean income. In 2015, this changed. Now, the regions with the highest mean income also have the highest levels of inequality. This is the main driver behind the convergence in social welfare. Note that the interaction is negative in 1953 and 1980, but positive for the other selected years.

Figure 22: Interaction between Mean Income and the Gini coefficient for selected years



Reading note: A trend line added to be able to observe the trends. Selected years follow from tables in chapter 5. Scales are different for all figures. Note therefore that the negative trend is stronger in 1953 than in 1980.