Journal of Public Economics 187 (2020) 104196

Contents lists available at ScienceDirect

Journal of Public Economics

journal homepage: www.elsevier.com/locate/jpube

Estimating long-run income inequality from mixed tabular data: Empirical evidence from Norway, 1875–2017

Rolf Aaberge [a,*], Anthony B. Atkinson [b], Jørgen Modalsli [a,c]

[a] Statistics Norway, Norway
[b] Nuffield College, Oxford, United Kingdom of Great Britain and Northern Ireland
[c] Oslo Business School, Oslo Metropolitan University, Norway

ARTICLE INFO

Article history:
Received 26 June 2018
Received in revised form 7 April 2020
Accepted 17 April 2020
Available online 5 June 2020

Keywords:
Distribution of income
Long-run inequality
The Gini coefficient
Norway

JEL classification:
D31
D63
N33
N34

ABSTRACT

Most evidence on the long-run evolution of income inequality is restricted to top income shares. While this evidence is relevant and important for studying the concentration of economic power, it is incomplete as an informational basis for analysing inequality in the income distribution as a whole. This paper proposes a non-parametric approach for estimating inequality in the overall distribution of income on the basis of tabular data from different sources, some in a highly aggregated form. The proposed approach is applied to Norway, for which rich historical data exist. We find evidence of very high income inequality from the late nineteenth century until the eve of World War II, followed by a rapid equalization until the 1950s. Income inequality remained low during the post-war period but has increased steadily since the 1980s. Estimates of a measure of affluence demonstrate that overall inequality has largely been governed by changes in the top half of the distribution and in the ratio between the mean incomes of the lower and upper halves of the population.

1. Introduction: inequality in the long-run

The extensive country-specific top income studies initiated by Piketty (2001) and provided by Atkinson and Piketty (2007, 2010) gave rise to a broad public debate on the rising income inequality in OECD countries, although the results of these studies dealt exclusively with the evolution of top income shares. Indeed, most of the discussion on long-run income inequality concerns the increasing share of total income received by the top 1 and top 10% of income recipients. This is a legitimate and important concern, as high top income shares reflect the fact that a disproportionate fraction of total economic resources is being controlled by a small minority of the population. However, an exclusive focus on the concentration of the top 1 and 10% ignores the distribution of income among the remaining 99 and 90% of the population and can contribute to misinterpretation of the long-run evolution in overall income inequality. The debate between Autor (2014) and Piketty and Saez (2014) on the driving forces behind the steep rise in income inequality in the US in recent decades underlines the importance of accounting for the rise in income inequality among "the other 99%".

The main objective of this paper is to propose a recipe for how the inequality of the income distribution as a whole can be estimated on the basis of different sources of tabular data from historical statistical income publications, which are available in many countries. This is possible without making assumptions about the distribution of incomes within wage groups (as in the "social tables" approach, e.g. Lindert and Williamson, 2016 for the United States) or relying exclusively on annual tabulations covering the majority of the population as is the case for Denmark (Atkinson and Søgaard, 2016). Complete detailed tabulations, like those for Denmark, appear to be an exception. Most developed countries have, however, collected income taxes on a regular basis and as a minimum published various aggregated quantities in some periods and detailed tabulations in other periods. This paper demonstrates that such combined data provide sufficient information to obtain reliable estimates of the Gini coefficient across time. The presence of various aggregated quantities of individual incomes for each year, allows estimation of points on the Lorenz curve and proves to provide a useful basis for estimating the Gini coefficient and any other rank-dependent

* Corresponding author.
E-mail address: rolf.aaberge@ssb.no (R. Aaberge).

https://doi.org/10.1016/j.jpubeco.2020.104196
0047-2727/

measure of inequality. Thus, in contrast to most previous studies, we do not limit the description of the evolution of long-run income inequality to years where complete tabulations are available, nor do we rely on interpolation of observations between years. Indeed, the paper demonstrates that much can be learned even in cases where detailed annual tabulations do not exist. By combining data from different historical statistical sources, a more complete picture of the distribution can be attained than that obtained from central government tax records alone. Similar procedures can probably be applied to other countries, to examine whether the development found for Norway also applies to other institutional and geographical settings.

The starting point is the information provided by the detailed tabulations of incomes by ranges as used by Aaberge and Atkinson (2010) in their study of top income shares in Norway. Section 2 of this paper describes the tabulated data available from the published income tax records from 1875 (annually from 1892) and in the form of micro-data from 1967 onwards. The tax information is a rich source, but it varies in form from year to year, and is limited in coverage, as it excludes non-taxpayers. The incomplete coverage of the population by the tax data means that there is a challenge involved in seeking to measure overall income inequality, as represented here by the Gini coefficient. We meet this challenge by creating "upper" and "lower" bounds on the Gini coefficient. Section 3 gives an account of the data and assumptions that are used to arrive at these bounds. In particular, we rely on aggregate information from the municipal and central government tax records, which are available annually for a long period. Combining these data with assumptions about the relative positions of different groups allow us to narrow the bounds on the estimated Gini coefficient. To this end, we add a further source of evidence about incomes at the bottom of the scale: administrative data on the number of recipients of public assistance, and the average amounts received.

The mixed tabulated data provide detailed information on the upper part of the Lorenz curve even in the 19th century, but less information on the lower part of the Lorenz curve. As is well known, the Lorenz curve is an increasing *convex* function taking values between 0 and 1. For distribution functions that are skewed to the right (heavy right tail), the Lorenz curve will exhibit weak (approximately linear) curvature in the lower part and strong convex curvature at the very top. Thus, to obtain a reliable estimate of the Lorenz curve for right-skewed distributions, it is necessary to have access to detailed tabular data for the top of the Lorenz curve, whereas it is sufficient to know a few points of the Lorenz curve below the median. As will be demonstrated in Sections 3 and 4, such mixed tabular data constitute an appropriate basis for using a non-parametric approach to estimate the Gini coefficient and any alternative measure of inequality that are explicitly expressed in terms of the Lorenz curve.

By deconstructing overall Gini inequality with respect to measures of affluence and poverty, we show in Section 5 that overall inequality is governed very much by what happens to inequality in the distribution of income in the top half of the distribution and to the ratio between the mean incomes of the lower and upper halves of the population, which means that the estimates of the overall Gini coefficient are less sensitive to assumptions made on how the income attributable to non-taxpayers is distributed. However, as demonstrated in Section 4 this does not mean that the evolution of the income shares of the top 1 or top 10% provides a complete picture of long-run income inequality in Norway.

A second objective of this paper is to provide new insight into long-run income inequality in Norway. The results presented in Sections 4 and 5 show that income inequality was high until the end of the 1930s, with substantial changes during the First World War. The turning point and origin of the low post-war inequality came with the German occupation between 1940 and 1945. The decline in inequality continued until the mid-1950s and remained stable at a low level between 1953 and 1980, but has increased steadily but moderately since 1980. Section 6 elaborates on how our results on the long-run evolution of income inequality contribute to an understanding of economic development in Norway since the late 19th century. It is shown to be a rich story that can be considered in terms of episodes of change.

1.1. Relationship with previous research

Our paper offers evidence of changes in inequality of the overall income distribution over a period of almost 150 years, and shows that changes in the ratio between the mean incomes of the upper and lower halves of the income distribution and in the inequality in the distribution of incomes among the richest 50% explain most of the changes in overall inequality. Apart from Atkinson and Søgaard (2016), who have had access to detailed annual income tabulations for the majority of the population in Denmark, previous research relies on less informative data and has mostly provided limited evidence on income inequality for selected years before 1945. Moreover, many of the scattered estimates of the overall distribution that do exist for earlier years are not comparable with modern series. The estimates for the United States provided by Spahr for 1890 and by King for 1910 (see Merwin, 1939) have been described in a review paper by Williamson and Lindert (1980, p. 91) as "eclectic size distribution guesses", with the conclusion that "it is better to pass over these". Williamson (1985) has produced estimates for the Gini coefficient for England and Wales, and Scotland, for selected years ranging from 1688 to 1915. None of these estimates can readily be linked to the modern series, but are made available in separate tables in a survey by Lindert (2000). Kuznets (1955) provided a comparative study by compiling income distribution estimates for a few scattered years for the United States, United Kingdom and Germany. Milanovic (2016, Chapter 2) collected evidence for several pre-industrial economies (based on social tables, wealth data and some income-based inequality series) and argues that inequality varies cyclically over time.

To our knowledge, there are three bodies of academic work that attempt to produce comparable estimates of overall inequality from the early twentieth century (or earlier) and onwards. First, Atkinson and Søgaard (2016) have estimated the Gini coefficient for Denmark for 1870 and from 1903 to 2010 based exclusively on annual detailed tax-based income tabulations, which emerge as an extraordinarily informative dataset compared to historic data from other countries. The Danish dataset suffers however, from a series break in 1970 when the unit of account changes from family to individual. Moreover, the Danish dataset only contains detailed tabulation for one year in the 19th century. Secondly, Vecchi (2017, p. 331) reports estimates of the income Gini coefficient for Italy between 1861 and 1931. The estimates of the early period are constructed by fitting a generalized beta distribution on household budget data, and these series has been supplemented with tax-based estimates for the later period. Thus, the overall series may suffer from weak comparability, whereas the estimates of the early period may depend heavily on the chosen parametric distribution. Thirdly, Garbinti et al. (2018) extend their previous series of top income shares in France by including estimates for the bottom 10% as well as the 10–50% from 1900 until 1985 (and a more complete income distribution after 1985).[1] Atkinson et al. (2017) provide a review of historic income inequality estimates, including how data points from separate studies can be merged to construct long-run series of income inequality for the United States (from 1918) and the United Kingdom (from 1938).[2] These estimates suggest that income inequality decreased in the early twentieth century and increased from the early 1980s, but reliable results are still not available for sufficiently many countries to

[1] Before 1970, Garbinti et al. (2018) assume constant income shares within the bottom 90%, e.g. the following shares are assumed to be constant for all years in the period 1900–1970 in France: the income of the lower 10% is 0.39% of the total for the bottom 90%, the next 40 (10–50) 26.30% and the next 40 again (50–90) 73.31%. See Appendix Table TD3 to Garbinti et al.

[2] Kopczuk et al. (2010) provide evidence of earnings inequality in the United States from 1937 onwards based on social security data and Kuhn et al. (2017) have produced estimates of income inequality for the United States starting in 1949.

justify a general trend. The longest previous series for income inequality in Norway were reported by Soltow (1965), who constructed a series of Gini coefficients based on samples of tax records for selected years between 1850 and 1960 for eight cities in southern Norway. The results of Soltow (1965) show decreasing income inequality among taxpayers living in these cities.

The methodological approach of the present paper extends previous analyses by combining detailed tabulations of income tax data for a limited proportion of the population with income data from other sources. For most years, we have access to detailed tabulations for the top of the income distribution. These data are supplemented by annual aggregate data from two different taxation schemes (municipal and central government taxes) and from poverty statistics. As is demonstrated in this paper, the shape of the Lorenz curves for right-skewed income distributions makes it feasible to use a non-parametric approach to estimate the Lorenz curve and the Gini coefficient when detailed tabulations are available at the top of the income distribution and aggregate data provide estimates of a few points of the Lorenz curve for the lower half of the income distribution.[3] By contrast, Garbinti et al. (2018) rely on the condition of constant income shares for the lower 90%, while Blanchet et al. (2017) and Vecchi (2017) rely on parametric distributions for broad intervals as a basis for estimating the overall income distribution.

The limited evidence on the evolution of overall inequality in the literature has been supplemented with useful information on the evolution of top income shares, not least owing to the top income books edited by Atkinson and Piketty (2007, 2010) and the extensive review provided by Roine and Waldenstrom (2015). Several studies combine results for the overall income distribution in modern times with estimates of top income shares for earlier years. For example, Piketty et al. (2018) report top income shares and estimates of the functional income distributions for the United States back to 1913, but provide no information on overall income distribution before 1964. Some studies suggest that top income shares could be a good proxy for overall inequality. The evidence provided by Leigh (2007), Roine and Waldenstrom (2015) and Morelli et al. (2015) is however mixed. The call for prudence made by Morelli et al. (2015) is supported by the results of the present paper. Actually, we demonstrate that top income shares might give a misleading picture of the evolution of overall income inequality, partly because changes in top income shares are normally accompanied by significant changes in the distribution of incomes in the upper half of the income distribution as well as by changes in the ratio between the mean incomes of the upper and lower halves of the population.

2. Income tax data in Norway

We begin with a brief account of Norwegian income tax data, and the way in which they can be used to produce results for income distribution as a whole. Since similar data are available for most countries, the method introduced below might be used as a recipe for estimating historical Gini series. This section is principally concerned with the years from 1875 up to 1951 when the published data are more fragmentary and vary in coverage. From 1952, the tabulations are more detailed, and from 1967 to the present we have access to micro-data. The income data originate directly from tax records (they are not inferred from taxes paid).

The same income definition, "antatt inntekt" (*assessed income*), is used over the entire period. It refers to income before tax (including capital income, taxable capital gains, taxable transfers and pension income), but after some pre-tax deductions. The pre-tax deductions represented relatively small amounts until the mid-1980s and were related to the expenses that were deemed necessary for the acquisition of income (Statistics Norway: Historical Statistics, 1994, p. 280). For the period where excluding such deductions from our income measure might create a bias - the late 1980s onwards - we show by using an alternative measure of income that the pattern of the historic series is similar to the pattern based on a more comprehensive income measure for this period.[4] Self-employment income was accounted for by assessing the productive capacity of farms (in particular smaller farms) and deriving figures from company accounts.

The tax unit (nuclear family), which is either a married couple or a single individual, defines the unit of analysis in this study. This choice is dictated by the tax statistics, as married couples were taxed jointly until 2018.

2.1. The income tax data from 1875

The income tax sources are municipal (MUN) and central government (CG) tax assessments: *Kommunenes skattelikning* and *Statsskattelikningen*.[5] The key feature here is that, for a number of years, the government has published tabulations of the distribution of income tax payers by income range. The sources are listed in Online Appendix B. In addition, we have data on the total number of MUN and CG taxpayers for all years, starting in 1892, as well as the total income earned by each group. As the MUN tax data are more extensive (tax thresholds are lower and more people pay MUN than CG tax), we assume that CG taxpayers are a subset of MUN taxpayers. Given the similar tax base and the way these sources are treated in the tax statistics, this is a reasonable assumption.

The coverage of the detailed income tax tabulations varies over the period. CG tax was introduced in 1892, so there is only distributional information on MUN tax for the years prior to that. The published tabulations for 1892 to 1903 only relate to CG tax, and the same applies to 1938 and 1948–1951. To summarize, in decreasing order of completeness, over the period up to 1951:

(i) MUN and CG distributional data: 1906, 1913 and 1929;
(ii) MUN distributional data: 1875 and 1888;
(iii) CG distributional data: 1892–1903, 1938, 1948–1951.

We supplement the distributional data with the data on the total number of taxpayers and their total income, which is available for nearly all years. This means that, in addition to the Lorenz curve from the distributional data, we have in case (iii) a further point corresponding to the total MUN taxpayers (and hence total taxpayers).

The tabulations of taxpayers by income range from 1952 to 1966, which precede the micro-data available from 1967, vary in their coverage (see below). Income is equal to income as assessed by municipal tax assessment for the years 1952–1955. In the tabulations for the years 1957 to 1966, income is defined as income as assessed by central government tax assessment if central government tax is levied. If not, income is defined as income as assessed by municipal tax assessment. There are no data for 1956 on account of the introduction of Pay-as-You Earn.

Since 1967, all individual incomes have been available on computer files at Statistics Norway. The income concept used is "antatt inntekt", income after some standard deductions, which is the same definition as that used in the pre-1967 tabulations. Using data from the Central Population Register, we merge married couples into single units, adding together the incomes of husband and wife to form the nuclear family.

[3] An illustration of the shape of Lorenz curves for Pareto distributions is displayed in Online Appendix D.

[4] Liberalization of the credit market in 1984–1985 combined with the right to deduct interest expenses and high marginal tax rates on capital incomes until the tax system was reformed in 1992 encouraged households to borrow, which led to a significant rise in interest deductions. However, although these reforms might have weakened the comparability of the historic income data, it should be noted that the evolution of the Gini coefficient for income after tax since the mid-1980s as displayed in Figure 4 in Section 3 is consistent with the evolution found for the historic Gini series over this period.

[5] This information, and further information below, comes from Gerdrup (1998) and the Introduction to Part XIII of Historisk Statistikk (HS) 1968.

An adjustment is required for the data from 1960 to 1967 to account for changes in tax unit definitions, as explained in Online Appendix E.

In the period 1921–1947, corporate incomes (as well as individual incomes) are included in the aggregates in the tax statistics publications. From 1937 onwards, we can obtain figures for individuals from other sources; between 1921 and 1936 we make adjustments to the totals based on observed rates from 1937 to 1947.

2.2. Control totals

The CG and MUN income tax tabulations for the late 19th and early 20th century do not cover significant proportions of the population as a whole. In order to arrive at an estimate of overall income inequality for the entire population, rather than only for the taxpayers, this study uses estimates of total number of tax units and total household income as starting points. The sources of these “control totals” are described in Online Appendix C. The first step in calculating total tax units is the adult population, defined here as those aged 16 and over. The second step is to subtract the number of married women. Defined in this way, the tax unit population (nuclear families) as reported in the population statistics increases from 847,000 in 1875 to 1.7 million in 1951 and 3.4 million in 2017. In 1875, 83% of nuclear families were covered by the tax statistics. In 1892, this figure had decreased to 52%. The share subsequently increased gradually to 80% by 1920. During the next period it decreased, to 66% in 1933, and subsequently increased again, reaching 86% in 1951. As explained below, after 1951 we rely on several different tabulations that together cover the entire population.

For total income, we use total household income for 1978 to 2017 from the National Accounts (NA) and extrapolate backwards using comparable historical series (see Online Appendix C). The resulting series for total household income as measured by the national accounts exceeds the total income recorded in the tax statistics (the internal total) in three main respects. First, the omission of the income of those not covered by the tax statistics. Second, understatement of income in the tax statistics. Third, differences in income definitions. Aaberge and Atkinson (2010) observed that the highest percentage for total NA household income recorded in the tax statistics was 72, and thus chose 72% of NA household income as control total. We use the same approach.[6] Total household income as measured by the NA is made up of (i) compensation of employees (not including employers' social security contributions), (ii) operating surplus of self-employed businesses, (iii) property income, (iv) transfers from government and from abroad, and (v) income not classified elsewhere. A comparison of the control total from the National Accounts and the total from the tax statistics is given in the left panel of Fig. 1 and in Online Appendix C (Table A5 and Figs. A3–A4).

The control totals provide estimates of the mean income per tax unit, which is displayed in real terms (as 2017 NOK) in the right-hand panel of Fig. 1. In the period since 1875, real income has risen by a factor of around 13.[7] But the growth has not been steady. Before 1914 there was an irregular pattern of recessions and recoveries. The inter-war period saw little improvement in real incomes. The post-World War II period, in contrast, experienced rapid growth up to the mid-1970s, which later slowed and was interrupted by the recession and banking crisis of the late 1980s and early 1990s.

3. Estimating the Lorenz curve and the Gini coefficient

We now move to the estimation of the Lorenz curve and Gini coefficient based on the data on MUN and CG taxpayers as well as the control total. Given that the data are typically incomplete, we have to make assumptions and work throughout with an upper and lower bound Gini coefficient. By consistently choosing assumptions that lead to higher inequality for the upper bound and lower inequality for the lower bound, we are able to efficiently bracket the true Gini coefficient that we would obtain if we had full information on the exact incomes of all nuclear families, and also to obtain a measure of the precision of our estimates.

The discussion in this section will be based on the available Norwegian historical data sources described in Section 2. However, the existence of several types of income tax as well as data on social assistance is by no means unique to Norway in this period. For this reason, the methods proposed here, utilizing tabular data to assess points on the Lorenz curve, are also applicable to other countries.

3.1. Estimation of Lorenz curves

The Lorenz curve plots cumulative income shares (on the vertical axis) against cumulative proportions of the population (on the horizontal axis), with the population ordered from low-income to high-income individuals. This means that the Lorenz curve will always be a convex function below the diagonal, as illustrated in panel (a) of Fig. 2. It is well-known that the Gini coefficient is defined by twice the area between the diagonal and the Lorenz curve. Hence, the bounds on the Lorenz curves constructed here correspond directly to bounds on the estimated Gini coefficients. A basic feature of the data used in this paper is that in all years, taxpayers amount to more than 50% of the population, and that the total number of taxpayers and their income are reported annually. The annual aggregates from the municipal and central government tax statistics provide accurate estimates for several points on the Lorenz curve every year.

Different formats of the overall Lorenz curves are shown in panel (b) of Fig. 2, which illustrates the case where we have distributional information on MUN taxpayers (with or without information on CG taxpayers) and panel (c) of Fig. 2, which illustrates the case where we have only aggregate information on MUN taxpayers. In our estimates, we assume that the total population of tax units is correctly measured by our control total. The difference between this total and the total recorded in the income tax tabulations is referred to as the “missing population”. Moreover, we assume that all individuals not represented in the statistics on MUN and CG taxpayers have incomes lower than those who pay tax. This means that the Lorenz curve for taxpayers is scaled down and connected with the final point for the missing population. In the case shown, the individuals in the missing population are all assumed to have identical incomes, so the first section of the Lorenz curve is a straight line. Further assumptions made about the distribution within the missing population are discussed below. Points H1 and H2 are points on the Lorenz curve constructed from MUN and CG taxpayer data. Panel (c) of Fig. 2 shows the case where there is no tabulated MUN data, only aggregates. On the assumption that those paying MUN tax but not CG tax all receive the mean income, the Lorenz curve for this group is represented by the dotted line.

The income attributable to the missing population is one element contributing to the difference between the income control total described above and the total income recorded in the tax statistics, where the latter is referred to as the “internal total”. In the period 1875 to 1951, there was a difference of around 20% between the internal and control totals (see Fig. A3), apart from during World War I. In our estimates, total income is taken as equal to the control total. This means that we can consider bounds on the Gini coefficient in terms of allocating the difference either to under-reporting in the tax data or to the missing population. Suppose that the amount by which the control total exceeds the internal total is equal to a proportion, α, of the internal

[6] Aaberge and Atkinson (2010, p. 476) provide further details and indicate that a similar approach has been used for Sweden and the United Kingdom.

[7] GDP per capita (in fixed prices) has grown by a factor of 18 over the same period. The discrepancy is largely due to the extensive demographic changes during this period; in 1985 Norway had a much younger population. The total population grew by a factor of 2.8 from 1875 to 2013, while total tax units (as defined here) grew by a factor of 3.6.

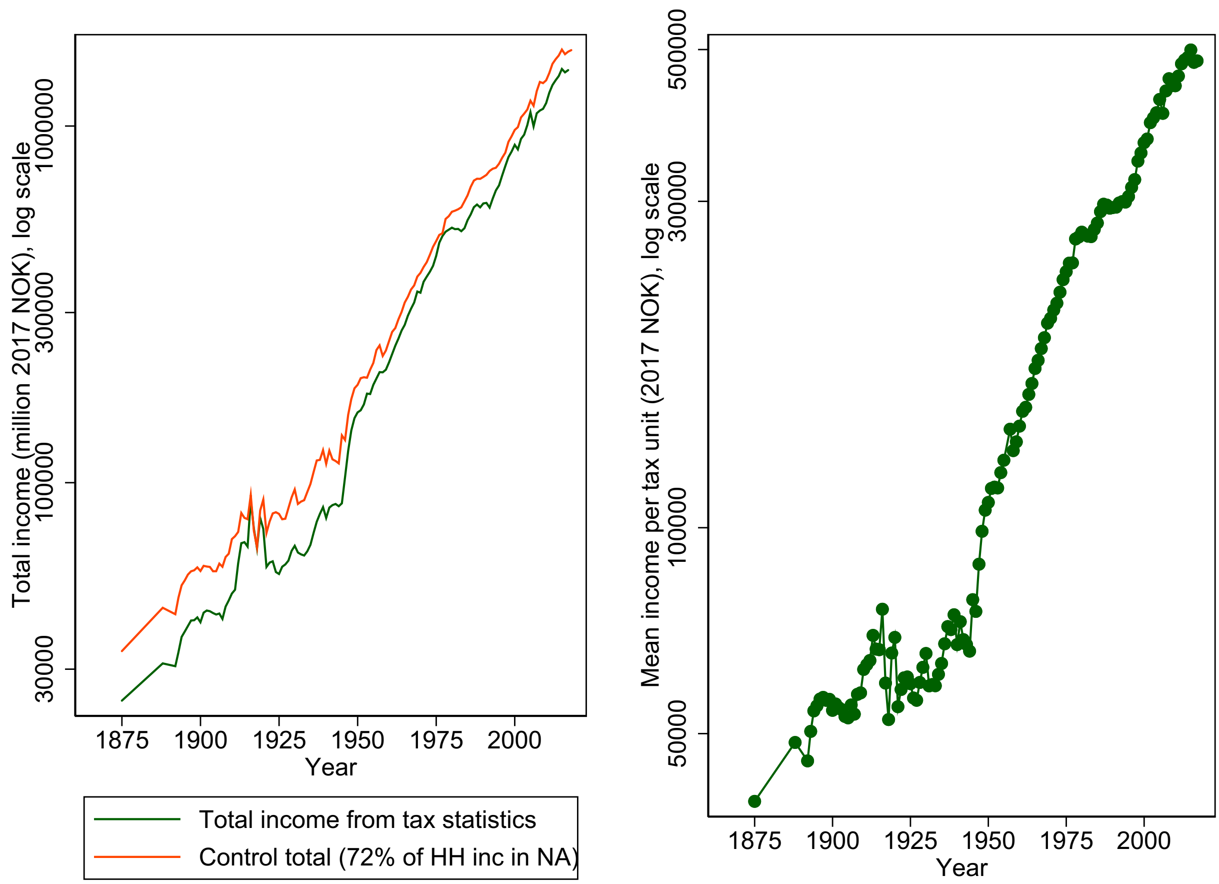

Fig. 1. Total income from tax statistics and control total (left panel) and mean income per tax unit (right panel). In 2017 NOK. Note: "Control total" refers to the total income one would obtain based on official (reconstructed) national accounts data, with the caveats explained in the main text. "Tax unit" refers to married couples and single individuals.

total, and that a proportion β of the internal total represents understatement in the tax data. This leaves $(\alpha-\beta)$ times internal total income to be allocated to the missing population, or $(\alpha-\beta)/(1+\alpha)$ times overall control income. If non-taxpayers constitute a fraction n of the total population, then the amount allocated per head to the missing population, expressed relative to the overall mean, is $(\alpha-\beta)\,/\,[n(1+\alpha)]$. This would be the overall slope of the first segment of the Lorenz curve in this example.

3.2. Implications for the Gini coefficient

The implications for the Gini coefficient are most easily seen in terms of the area under the Lorenz curve, since the Gini coefficient is equal to 1 minus twice the area under the Lorenz curve. For taxpayers alone, twice the area is equal to

$$B=\Delta F_1\,H_1+\Delta F_2\,\{H_1+H_2\}+\ldots+\Delta F_k\,\{H_{k-1}+1\} \tag{1}$$

where ΔF_i is the density in the range and H_i denotes the cumulative share of total income up to and including range i, where there are k ranges. It follows that the Gini coefficient for taxpayers alone is

$$G^{*}=1-B \tag{2}$$

The introduction of the missing population as in panel (b) of Fig. 2 has two effects. It squeezes the Lorenz curve for taxpayers to the right. In Eq. (1), this does not affect H_i but reduces ΔFi, and hence the area B, by a factor $(1-n)$. The second effect is that it adds additional area under the first segment. If it is assumed that all incomes are non-negative, then the least such addition is zero (i.e. β is set equal to α), in the case where the Lorenz curve in panel (c) of Fig. 2 initially follows the horizontal axis. Together, these two effects give an upper bound G_U for the overall Gini coefficient, which can be expressed

$$G_U=n+(1-n)G^{*}=G^{*}+n(1-G^{*}) \tag{3}$$

It is a weighted average of 1 and G^*. In 1875, for example, values of $n=16.8\%$ and $G^*=47.6\%$ imply that the upper bound is 56.4%.

Conversely, a lower bound might be sought by allocating all the difference to the missing population (β is set equal to 0), but this may violate the assumption that the missing population have incomes below the lowest income of taxpayers. Moreover, for some years there is contemporary evidence on which we can draw. For 1875, the tabulations published by Kiær (1893), which we are using, include an estimate of the numbers and income of the missing population.[8] The mean for the range NOK 0–NOK 400 was NOK 230, which was 40.9% of the overall mean. If, as an illustration, we attribute this amount per unit to the missing tax units, it means that, of the uplift moving from the NOK 345.5 million internal total to the NOK 475.8 million control total, 32.6 million NOK, or 28.3% of the uplift, is allocated to the missing population.

The lower bound adopted here is calculated by considering the area under the Lorenz curve, where the missing population is allocated a fraction h of total income. Twice the area under the Lorenz curve is therefore increased by h times n. At the same time, the Lorenz curve for taxpayers is squeezed vertically by a scale factor $(1-h)$, reducing its area but adding a rectangle, which adds $2h(1-n)$. The resulting

[8] Incomes below NOK 400 were exempt from taxation.

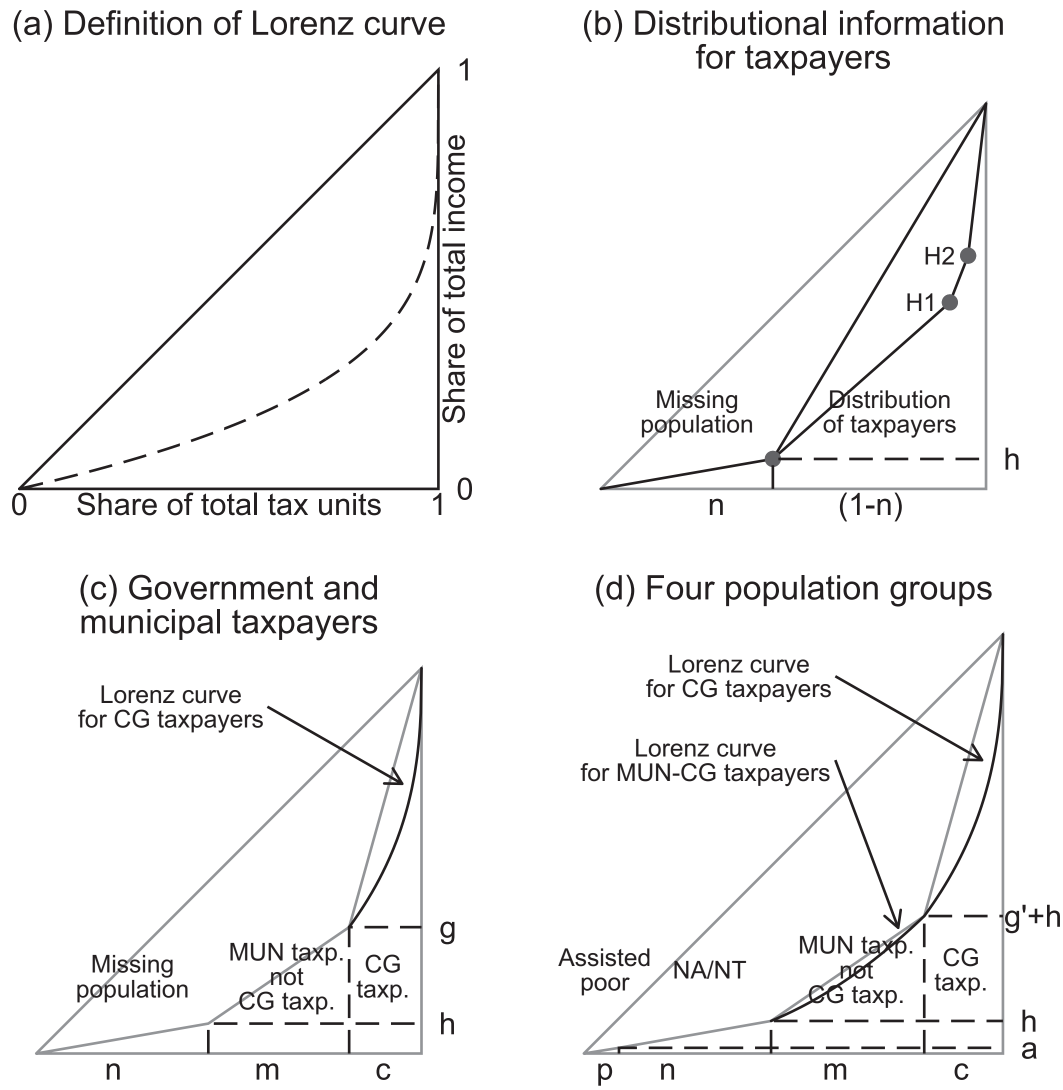

Fig. 2. Estimation of Lorenz curves from tabular data. Figure note: Fig. 2 shows how Lorenz curves are estimated on the basis of tabular data and assumptions of within-group distributions. Panel (a) shows the definition of the Lorenz curve as a plot of cumulative income shares against cumulative population proportions, where the diagonal line illustrates a hypothetical distribution with complete equality and the line tracing the lower and right corner of the figure illustrating extreme inequality (one tax unit owns the total income). Panel (b) illustrates how a Lorenz curve with distributional information for taxpayers (with two points H1 and H2 known from tabular data) is combined with the proportion of the population who pay tax to arrive at a Lorenz curve for the overall distribution. "Missing population" refers to the share of the population who are not covered by the income tax statistics. Panel (c) illustrates a Lorenz curve with two groups of taxpayers: those paying both central government (CG) and municipal (MUN) tax (denoted "CG taxpayers") and those paying only the municipal (MUN) tax. In this example, the distributions among the CG taxpayers as well as the proportion of the population not paying income tax are known. Panel (d) shows our baseline specification with four population groups, where the missing population is divided into those who receive poverty assistance (the assisted poor) and those who do not receive poverty assistance (the "non-assisted, non-taxed", NA/NT population). In this panel, we have applied within-group distributions for the central government taxpayers as well as for people who only pay municipal tax. For further details of estimation and definitions, see the text below.

lower bound Gini coefficient is

$$G_L = n + (1-n)G^* - h[1 + (1-n)G^*] = G_U - h[1 + (1-n)G^*] \quad (4)$$

The last term shows that the difference between the upper and lower bounds — a measure of our uncertainty about the extent of income inequality in that year — increases, as we would expect, by the value of h, magnified by a factor of $(1 + (1 - n)\, G^*)$. The 1875 value of $h = 8.6\%$, coupled with $n = 16.8$ and $G^* = 47.6\%$, generates a difference of 9.6 percentage points from the upper bound, or a value of 46.8% for the lower bound.

3.3. Using aggregate taxpayer data

For certain years, we have only the aggregate number and total income of the MUN taxpayers who are not liable for CG tax, and nothing is known about the distribution among this intermediate group. (We do however know the distribution among CG taxpayers.) This is the situation shown in panel (c) of Fig. 2.

Let us denote the proportion of the population in the MUN-CG group by m, the proportion of CG taxpayers by c, and the proportion of those in neither group by n (so $c + m + n = 1$). The contributions of the three groups to the overall Gini coefficient may be seen in panel (c) of Fig. 2. Denote the income share of the bottom group by h, and the combined share of the bottom two groups by g. Subtracting twice the area under the Lorenz curve from 1 gives the overall Gini coefficient:

$$G = 1 - \{hn + (g + h)m + c[1 + g - (1-g)G^*]\} \quad (5)$$

where G^* is the Gini coefficient for the CG taxpayers. This may be rewritten by introducing a new parameter $g' = g - h$ and replacing g

with $(g' + h)$ as

$$G = 1 + c(1-g')G^{*} - \{g'm + c(1+g')\} - h\{1 + m + cG^{*}\} \qquad (5a)$$

The upper bound is obtained by setting $h = 0$ and keeping the other parameters constant. The final term in Eq. (5a) shows that the difference between G and the upper bound is proportional to h, with a magnification factor that is less than 3, but which may nonetheless be substantial. In 1892, the first year for which there is CG data, $m =$ 36.6%, $c =$ 18.8% and $G^{*} =$ 44.8%, so that the magnification factor is 1.45.

What, if anything, can we say about years for which there are no detailed tabulations for CG taxpayers? Formula (5a) allows us to see the role played by inequality within the group of CG taxpayers when $h = 0$. The term $c(1 - g')G^{*}$ is an addition to the overall Gini coefficient. Suppose that we do not know G^{*}, but do know c and g'? Then the difference between the bounds would be widened to an extent that depends on the product of the population share and the income share of the CG taxpayers. Whereas the product may have been small in the nineteenth century, it was substantially higher in World War I and later. On the other hand, in the years for which we have tabulations, the Gini coefficient among taxpayers has rarely exceeded 50% or (apart from two exceptions) fallen below 30%.

3.4. Using data on the assisted poor

In order to provide a more solid basis for our treatment of the lower part of the distribution, we need additional information on the incomes of those below the tax threshold. In search of this, we explore one possible source: administrative data on the number of recipients of public assistance and the average amounts received. It is assumed that the recipient unit can be equated to the tax unit and that the poverty assistance is the same as subsistence market income. This means that all individuals/couples are assumed to have positive market income (where some could have zero); on the other hand, some of the recipients of assistance might also receive small amounts of market income.

In effect, using this additional administrative information means introducing into the three-group model a fourth group, by dividing those not paying tax into those who are assisted (the “assisted poor”) and those who are neither assisted nor taxed (NA/NT). The key assumption underlying our construction of the Lorenz curves and calculation of the Gini coefficient is that the groups can be ranked in order of increasing income, as shown in panel (d) of Fig. 2. As liability for taxation depends on both income and wealth, there could be cases where people are liable for MUN taxation on account of wealth (and hence are included in the tax authorities' calculations of the number and total income of MUN taxpayers) but have low incomes that would place them below people in the NA/NT group. But it seems a reasonable approximation.

When the proportion of assisted poor is denoted by p, the proportion in the NA/NT group by n, and the share of the first group by a, the Gini coefficient is now given by

$$G = 1 + c(1-g')G^{*} - a(n+p) - g'(c+m) - c - h\{1 + m - p + cG^{*}\} \qquad (6)$$

The population proportions, p, n, m and c are known. The total income received by the assisted poor and by the two groups of taxpayers is known. The total income of the NA/NT group is not reported in the tax statistics. Here we have to make assumptions regarding the upper and lower bounds, but with the advantage that this group — given our earlier assumption — is “sandwiched” between two groups about which we have information. The upper bound is calculated on the assumption that the NA/NT group has the same average income as the assisted poor, the lower bound on the assumption that the average income of the NA/NT group is equal to one third of the average income of the MUN-CG group.[9] For some years, the MUN-CG mean income turns out to be less than three times the mean poverty support. In these cases, the imputed income for the NA/NT group will be the same for the upper and lower bounds.[10]

Expression (6) for the Gini coefficient does not account for possible dispersion within any of the three groups with lowest incomes. However, the POOR and NA/NT groups are always relatively small and, given our assumption that groups are ranked by income, limited by the income ranges of neighbor groups (or zero, in the case of the poor). This puts a strict upper limit on the contribution to the overall Gini that could result from within-group dispersion in these groups. For example, the maximum consistent inequality in the poorest group would mean that the richest individuals in this group had the same income as the NA/NT mean income and the poorest individuals in this group had zero income. The effect of such a distribution would be largest in 1888, where the lower bound Gini measured in percentage points would increase only from 56.91 to 56.93.[11]

On the other hand, the MUN-CG group constitutes a relatively large proportion of the population, and the data show that the differences between the MUN-CG and CG mean incomes are substantial. For this reason, within-group dispersion is introduced for the MUN-CG group. Specifically, the incomes within this group are assumed to follow a uniform distribution. The details of this imputation are given in Online Appendix D, where the relationship between the dispersion parameter z and the within-group MUN-CG Gini coefficient $G^{**} = z/3$ is explained. As we maintain the assumption that there is no overlap between the income groups, there is a limit to the upper value of z. Overall, a value of $z = 0.4$ is consistent with the introduction of some dispersion without any MUN-CG taxpayers having either higher incomes than the lowest in the CG-group or lower incomes than the NA/NT group. Note, however, that the overall Gini coefficient proves to be insensitive to changes in z.

Finally, in 1875 and 1888 (the years before the introduction of CG tax in 1892) there was no state taxation, but instead detailed tabulations of the incomes of MUN taxpayers. We then assume that the lowest tabulated income group in the MUN tabulations is equivalent to the MUN-CG group in later years, and that the higher-income groups would have been subject to CG tax had that been in effect in these years.

[9] A number of further adjustments have to be made to the published tabulations in making these 4-group calculations. Assumptions are required when estimating the upper and lower bounds. For G^{*}, if the within-group Gini of the CG taxpayers is not available, the upper bound uses the maximum of the previous and the next observations of G^{*}. Similarly, the lower bound uses the minimum of the previous and the next observation if there are no data available. For the years 1875 to 1891, when there were no CG taxation, the average income of the NA/NT group for the upper bound Gini is taken as NOK 150. NOK 150 was 25% of the mean income of workers and 33% of the mean income of farmers (including cotters) in 1888/1889 (Sth. Prp. Nr 48, 1890).). Note that our “upper bound inequality” applies within the framework of assumptions outlined in this chapter. If, for example, we assume that the entire difference between total income from the tax statistics and national accounts was entirely “hidden income” that accrued exclusively to the rich, inequality would be higher. Given the nature of the tax system we do not find this assumption plausible. The lower bounds are assessed within the framework of the control total as described in Section 2.2.

[10] Alternatively, one could attribute zero income to recipients of poverty support on the grounds that one wanted the income definition to respect a strict “pre-tax” definition. A counter-argument is that the poverty support is likely to reflect the subsistence income received by these households. Changing the income level of the poor to zero (while maintaining the income levels for the NANT group) would increase the Gini coefficient by between 0.004 and 0.036. Results from this exercise are available on request.

[11] Graphically, we obtain the upper bound from 5 by extending the line for the NA/NT group (the slope of this group is the mean income of NA/NT relative to the population mean) down to zero. The resulting triangle (the contribution of the poor group to the overall Gini coefficient) is $\frac{a}{2}\cdot(p - \frac{an}{h-a})$. Introducing dispersion to the NA/NT group would decrease the maximum consistent contribution from the poor group.

3.5. Estimation of benchmark series from 1875 to 1951

The comprehensive approach described in Section 3.4 provides the basis for our analysis of the long-run evolution of inequality in Norway in this paper. The Gini coefficient for the years 1875, 1888 and 1892–1951 will then be given by

$$G = 1-pa-n(a+h)-m(2h+g\prime)-c(1+g\prime+h)+c(1-g\prime-h)G^{*} + g\prime mG^{**} \quad (7)$$

where
a = total income of the poor relative to control total,
h = total income of the poor and non-assisted/non-taxed (NA/NT) relative to control total,
g' = total income of MUN taxpayers who are not CG taxpayers,
p = the poor as proportion of total tax units,
n = NA/NT as proportion of total tax units,
m = MUN-CG taxpayers (those who pay municipal tax but not central government tax) as proportion of total tax units,
c = CG taxpayers as proportion of total tax units,
G^{**} = Gini coefficient among MUN-CG taxpayers,
G^{*} = Gini coefficient among CG taxpayers.

Expression (7) takes as its starting point extreme inequality where the Lorenz curve follows the horizontal axis between 0 and 1. The first four terms then subtract the areas of the triangles and parallelograms below the Lorenz curve as illustrated in Fig. 2. The latter two terms add in the within-group Gini coefficients for the two richest groups, scaled by group sizes and income shares.[12]

3.6. Bounds for 1952 to the present

The above discussion has described the bounds applied for the period 1875 to 1951. For the post-1951 period, when coverage was greater, relatively high numbers of tabulated intervals have been published by Statistics Norway (Statistics Norway: Historical Statistics, 1978). From 1967 onwards the incomes of the entire population of taxpayers are available as micro data. For this reason, the set of necessary assumptions for this period is smaller, similar to the situation shown in Fig. 2, where the assumptions relate only to the mean income of the missing population. These assumptions are designed to be comparable with those for the earlier period, while taking account of the changing role of assistance to the poor in the 1960s and later. In particular, there is a break in the poverty support series between 1964 and 1967, making mean payout per supported individual a less appropriate value for imputation at the lower end of the income distribution.

The upper bound of the Gini coefficient is based on the assumptions that (i) those not covered by the tax tabulations have a mean income equal to mean assistance (as before) for the years up to 1964 and (ii) from 1967 the group receives 50% of the minimum old-age pension for a single person.[13] The lower bound is based on the assumption that those not covered by the tax tabulations receive mean income equal to 150% of the mean income assumed for the upper bound.

We should emphasize at this point that the final series is based on a consistent population throughout the period. Despite the change from household-based to individual-based taxation, we can replicate pre-1960 nuclear families in the post-1966 microdata by merging spouses using personal ID numbers in the latter data that link taxpayers and the population recorded as individuals. The first year in which married women could choose to file taxes individually is 1960. We therefore transform individual data for the years 1960–1966 into household-based data using data from the 1960 census as well as the distribution of spouse's incomes, marriage and tax status in 1967. Similarly, adjustments are made to take account of a separate taxation system for sailors (1948–1966) and company taxation (1921–1947). These adjustments are all described in detail in Online Appendix D.

3.7. Long-run inequality in Norway

The results of the calculations discussed in Sections 3.5 and 3.6 are brought together in Fig. 3, which shows the upper and lower bounds for the Gini coefficient. The difference between the upper and lower bounds is largest for the pre-1914 period. The average difference over the period from 1892 to 1914 is 9.8 percentage points, whereas the average difference from 1915 to 1951 is 2.2 percentage points. While the difference represents potential error introduced at the stage of data analysis and is not comparable with the sampling error typically considered in distributional analysis, it is nonetheless interesting to compare their magnitudes. From that perspective, the 1892 to 1914 figure appears quite large, but the 1915 to 1951 average difference is not dissimilar to the confidence intervals obtained from the reported standard errors for the Gini coefficients obtained from household surveys: for example, the 95% confidence interval for the Gini coefficient of the distribution of disposable equivalent (household) income in Norway varied between 1.4 and 3.6 for the period 1986 to 1993.

In Appendix A, we perform an evaluation of the sensitivity of estimates of the Gini coefficient to the employment of additional data sources, and a robustness check of the Gini series based on two measures of inequality that complement the information provided by the Gini coefficient. The results displayed in Fig. A1 show how the estimated Gini coefficients depend on the choice of data for interpolating the Lorenz curve, starting with the simplest approach in Section 3.1 and increasing the sophistication of the method to arrive at our preferred estimates. It is shown that the Gini coefficients produced by the naive estimator are far too low.

Irrespective of whether we use a measure of inequality that are particularly sensitive to changes that take place in either the lower or the upper part of the income distribution the evolution of the associated inequality estimates shows, as demonstrated by Fig. A2 and Table A8, a

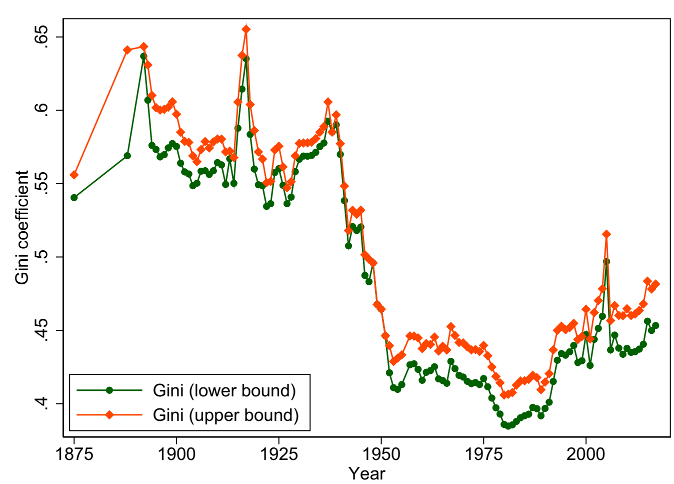

Fig. 3. Gini coefficient for the distribution of income in Norway, 1875–2017. Upper and lower bounds. Note: For sources, methods and assumptions, see text. "Lower bound" refers to inequality estimated using assumptions on average income level of the non-taxed and distribution among central government taxpayers that lead to lower inequality, while "upper bound" refers to inequality estimated using assumptions that lead to higher inequality. See Sections 3.1–3.6.

[12] While the Gini coefficient is calculated directly from Eq. (7), we can also construct Lorenz curves using the assumptions outlined here. These are available as an online appendix. In this case, a Pareto distribution is imposed for the richest (CG) group, with the dispersion and lower bound parameters set to match the mean income and Gini coefficients of this group. As long as these two conditions are satisfied, the choice of within-group dispersion has no impact on the estimated Gini coefficient for the entire population or any part of the population that includes the entire CG group.

[13] For the years 1965 and 1966, the minimum pension was projected back from 1967 (when it was introduced) in line with the growth of seamen's pensions, which were introduced in 1950. The same process applied to 1964 yielded a figure of NOK 2140, which was close to the poverty support level in that year of NOK 1975.

similar pattern as the estimates of the Gini coefficient. However, the measure that is most sensitive to changes in the upper tail of the income distribution shows significantly larger relative changes than the Gini coefficient during the post-war period. Moreover, the Gini coefficient shows significantly larger relative changes than the measure that are most sensitive to changes that occur in the lower part of the income distribution. As will be demonstrated in Section 4.2, these results prove to be consistent with the information obtained by comparing the evolution of the overall Gini coefficient with the evolution of the ratio of the upper and lower mean income, the upper tail Gini and the measure of affluence, which is discussed in Section 4.

3.8. Different income definitions

The standard "official" estimates of the Gini coefficient for the distribution of income in Norway accounts for taxation, public cash transfers as well as for the needs of household members and has been published since the mid-1980s. As indicated in the introduction our choice of definitions has been dictated by constraints in available historic data sources. This is why we have adopted a gross income definition, whereas statistical agencies today provide inequality estimates on the basis of disposable equivalent income.

The closest Norway gets to an official definition of income inequality is Statistics Norway's time series from 1986 onwards.[14] The construction of this series diverges from the approach used elsewhere in this paper in three ways. First, the household definition includes everyone living together with joint consumption except students not living at home. To account for scale economics the standard EU equivalence scale is used. Second, a somewhat larger set of income sources (various types of non-taxable transfers) is included than the "gross income" concept used in this paper. Third, the income basis is post-tax rather than pre-tax.

Fig. 4 compares the evolution of the Gini coefficient since 1986 for the two alternative definitions of income. As expected, inequality in the "official" series is much lower than the long-term series. This is largely due to the redistributive effects of public transfers and a progressive tax system, but it also reflects the treatment of the income unit. The use of a wider definition tends to reduce recorded inequality, since it assumes a greater degree of income-sharing. Taking account of economies of scale in larger households has also a significant effect on the measured level of inequality. However, since our focus is on the evolution of inequality, we find it reassuring that the pattern of the historic series captures the pattern of the official series from 1986 onwards. Most important here is that the development of inequality over time is similar for the two definitions. There was a significant increase from 1986 to around 2000, turbulence around the tax reforms of the early 2000s and a slight increase thereafter.

4. The relationship between overall inequality and inequality at the top

Although Eurostat, OECD and national statistical agencies publish top income shares, ratios of income quantiles and decile means on a regular basis, such quantities cannot be regarded as measures of inequality as they don't satisfy the Pigou-Dalton principle of transfers. Thus, in order to provide information on overall inequality, these institutions regularly publish estimates of the Gini coefficient. Since most of the discussion of the long-run evolution of inequality in OECD countries concerns the increasing top income shares, it is interesting to explore what we learn from the new series of overall inequality (as in Fig. 3) compared to

[14] See http://www.ssb.no/inntekt-og-forbruk/statistikker/ifhus

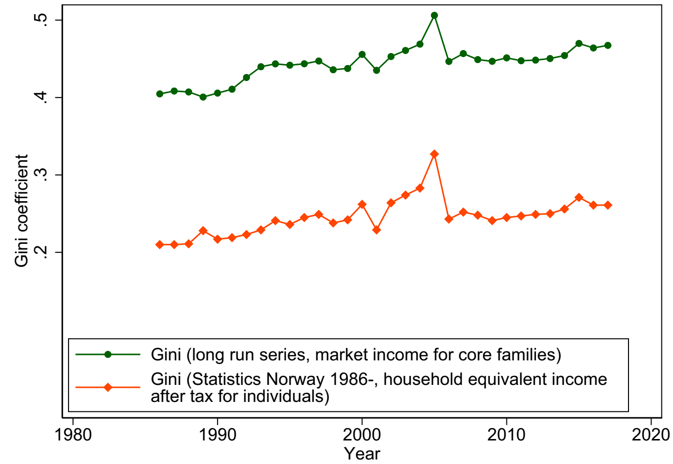

Fig. 4. Gini coefficient for the distribution of income in Norway based on two alternative income definitions, 1986–2017. Note: The long run series is the average of the upper and lower bound reported in Fig. 3. For sources, methods and assumptions, see text.

the top income series previously published by Aaberge and Atkinson (2010). To this end, we compare the evolution of the income shares of the top 1% and overall inequality, where we take the mean of the upper and lower bounds of Fig. 3 to give an "average series".[15]

4.1. The share of the top 1% and the Gini coefficient

When comparing top income shares and the Gini coefficient, it is useful to apply the following approximate decomposition proposed by Atkinson (2007, p. 19–20), and proved by Alvaredo (2011),

$$G \approx (1-S)\cdot G^{bottom99\%} + S \tag{8}$$

where S is the income share of the top 1% and $G^{bottom99\%}$ is the Gini coefficient of the bottom 99%.[16] The approximate decomposition of the Gini coefficient is shown in Fig. 5.

Fig. 5 demonstrates that the evolution of the share of the top 1% does not capture the evolution of income inequality in Norway, although overall inequality and top income shares have moved closely together in recent decades. Over the period 1882 to 1939 the Gini coefficient is seen to vary significantly, even though the Gini coefficient measured in percentage points only declined modestly, from 64% in 1882 to 59% in 1939. By contrast, the share of the top 1% decreased significantly, from 22% to 13%, over the same period, while the Gini of the 99% increased from 43 to 48%.

During the Second World War and the early post-war period, both overall inequality and top income shares showed a substantial decline. Between 1939 and 1953 the Gini coefficient fell from 59% to 42%,

[15] The averaging is done because we recognize that what many researchers require is a single series, and that if we do not provide an average ourselves, users will do so. At the same time, there is no evident justification for taking a simple average. It can be argued that the upper bound attributes an unreasonably low income to those recording zero. The appropriate weights may vary over the time period. But the simple average provides a point of reference.

[16] The exact decomposition is given by

$$G = 0.01\cdot S\cdot G^{top1\%} + (1-S)\cdot G^{bottom99\%} - 0.01\cdot(1-S)\cdot G^{bottom99\%} + S - 0.01,$$

where the first, third and fifth terms are relatively small compared to the second and fourth terms, which justifies the simplified expression in Eq. (8). The first term in Eq. (8) is always 0.001 or less in our data, while the third term is maximum 0.005. The last term is constant at 0.01 by definition. Note that the top 1% series shown here (and given in the Appendix) differs slightly from those published by Aaberge and Atkinson (2010), simply because the present series utilizes additional data sources, relies on different assumptions on the distribution of unmeasured income and uses interpolation of the top distribution across some years, as explained in Section 3.

while the share of the top 1% fell from 13% to 6%. The evolution over the next three decades was again rather different. There was a substantial decline in the share of the top 1%, from 6% in 1953 to 4% in 1980, whereas the Gini coefficient was fairly stable. Since 1990, the share of the top 1% has regained the lost ground, and was 10% in 2017 according to our estimates here, and the Gini coefficient too has risen — although only to around 47%. This difference between the time paths of the top shares and the Gini coefficient shows that, while the top share may have driven much of the recent increase in overall inequality, there have been other forces in operation as a result of which not all of the post-war equalization has been lost. Note that the evolution of the share of the top 10% parallels the evolution of the share of the top 1% (see Aaberge and Atkinson, 2010). We refer to Online Appendix G for a decomposition of the Gini coefficient by the income share of the top 10% and the Gini coefficient of the distribution of income among the bottom 90%. Furthermore, Appendix G provides results of the evolution of decile-specific income shares.

Since the evolution of top income shares and overall inequality differs in a number of periods, estimates on upper tail inequality and the ratio between the mean incomes of the lower and upper half of the population might provide essential information on whether changes in overall inequality are due to a widening of the income gap between the upper and lower half of the population and/or changes in the distribution of income among the richest 50% of the population. As will be demonstrated below, these distributional measures contribute to explain the driving forces behind the evolution of overall inequality.

4.2. Affluence

Before World War II, taxpayers comprised between 52 and 81% of the annual populations of tax units, which means that the data base for describing the upper half of the income distribution is richer than that for describing the lower half. This makes it particularly relevant to consider the evolution of the mean and the Gini coefficient for the most affluent 50% of the population and use the associated estimates as a basis for estimating "affluence", a measure introduced by Aaberge and Atkinson (2016). Affluence has been given an axiomatic justification and is defined by

$$A = \frac{1}{3}\left(\frac{\mu_U}{\mu}(G_U+1)-1\right) = \frac{2}{3(1+\gamma)}\left(G_U+\frac{1}{2}(1-\gamma)\right) \tag{9a}$$

where μ is the overall mean income, μ_U and G_U are the mean and the Gini coefficient, respectively, of the richest 50% of the population, $\gamma = \mu_L/\mu_U$ and μ_L is the mean of the poorest 50% of the population. Expression (9a) shows that affluence, A, increases with increasing inequality in the income distribution of the richest 50% and decreases with increasing mean income ratio γ.[17] Inserting the well-known expressions for μ_U and G_U into Eq. (9a) yields the following alternative expression for A,

$$A = \frac{4}{3}\int_{\frac{1}{2}}^{1}(2t-1)\left(\frac{F^{-1}(t)}{\mu}-1\right)dt \tag{9b}$$

where $F^{-1}(t)$ is the income of the individual with rank t in the income distribution F. Expression (9b) shows that A can be interpreted as a (normalized) weighted average of the income shares of the richest 50%, where the weight increases with increasing rank from 0 for the median income to 4/3 for the highest income. The affluence measure A has itself a range [0,1] and takes the value 0 if and only if all individuals receive the same income μ. At the other extreme, when total income is received by one sole individual, then A takes the value 1. Note that 3A becomes equal to the relative affluence gap $((\mu_U - \mu)/\mu = (1-\gamma)/(1+\gamma))$ if individuals with higher income than the median income receive the same income μ_U. The estimation results for affluence, A, and the upper tail (above median) Gini coefficient G_U are displayed in Fig. 6.

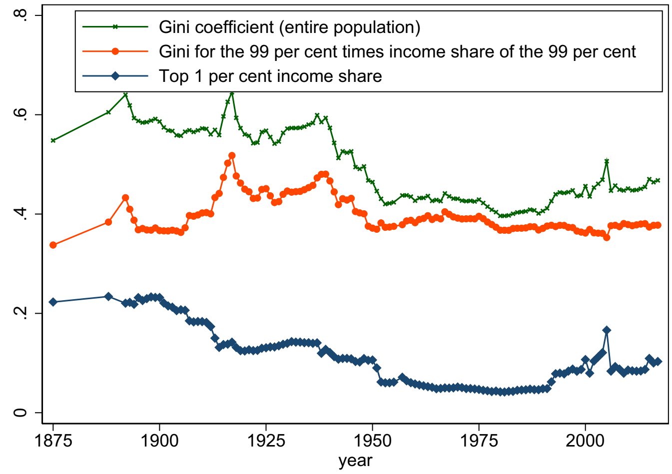

Fig. 5. Decomposition of the Gini coefficient by the income share of the top 1% and the Gini coefficient of the distribution of income among the bottom 99%, Norway 1875–2017. Note: The components are estimated using the same estimated Lorenz curves as were used for estimating the Gini coefficients in Fig. 3. The graphs display the mean of the upper and lower bound estimates. For sources, methods and assumptions, see text.

Since the available data provide a better basis for estimating affluence than overall inequality before World War II, it is reassuring that the affluence pattern largely captures the pattern of the overall Gini. Note also that the reliability of affluence (and the upper tail Gini coefficient) to a large extent carries over to the estimated overall Gini series. This is because income distributions are normally skewed to the right, which means that the upper tail Gini contributes a significantly larger share of the overall Gini than the lower tail Gini. Aaberge and Atkinson (2016) demonstrated that the overall Gini is equal to $3(A+P)/4$, where P is the poverty counterpart of affluence (A). Thus, in 1900, with $G = 0.586$ and $A = 0.515$ (see Table A7 in the Online Appendix), the contribution from affluence to the overall G was 66%, while the contribution of A had declined only marginally, to 62%, a hundred years later.

To complement the information on inequality given by the Gini coefficient, Appendix A provides estimates of two closely related rank-

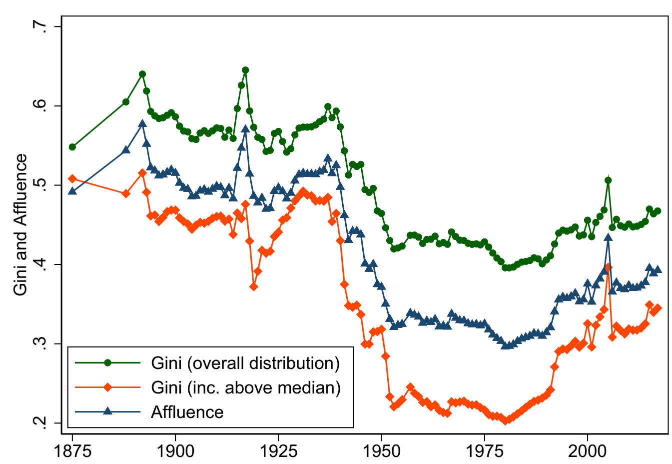

Fig. 6. Gini coefficients for the overall distribution of income and the distribution of income for those with income above the median, and a measure of affluence, Norway 1875–2017. Note: The components are estimated using the same distribution as the one used for estimating the Gini coefficients in Fig. 3. The graphs display the mean of upper and lower bound estimates. For sources, methods and assumptions, see text.

[17] As demonstrated by the following expression $\gamma = 2(\mu/\mu_U) - 1$ there is one to one correspondence between γ and μ/μ_U.

dependent measures of inequality. The results show to be in line with the evidence provided by this section, i.e. changes in inequality from 1939 to 1980 and from 1980 to 2017 largely concerned the upper part of the income distribution.

5. Long-run inequality in Norway: a series of episodes

The evolution of inequality in Norway is best characterized, we believe, as a series of episodes identified with sub-periods, which are summarized in Table 1. As demonstrated by the change in percentage points, the evolution of the overall Gini coefficient is closely related to the evolution of Gini-based affluence measure. In the same way, the upper tail Gini and the mean income ratio typically move in the opposite direction; rising (declining) upper tail Gini and declining (rising) mean income ratio. It is evident from Fig. 6 that three episodic changes in income inequality deserve special mention. First, inequality was turbulent during World War I, but analysis is complicated by price and wage fluctuations during this period. Leaving aside this period, the Gini coefficient in the four decades from the 1890s to the end of the 1930s was measured in the range of 0.60 plus or minus 0.05. Second, the decline during World War II was swift and extensive. Third, the post-1989 reversal took the Gini from around 0.40 to over 0.45 in two decades. We turn now to consider the individual sub-periods in more detail.

Taken as a whole, the period from 1875 to 1939 shows unchanged overall inequality and affluence, whereas the upper tail Gini decreased by 5 percentage points. The different evolution of upper tail inequality and overall inequality (and affluence) corresponds to a significant decrease in the ratio between the mean incomes of the lower and upper halves of the population (see Fig. 8 and Table A7 in the Online Appendix). The increase in the Gini coefficient from 0.55 in 1875 to 0.64 in 1892 reflects an increase of the share of total income accruing to the highest income group. The growth rates were low during this period, and emigration to North America increased sharply from 1880. This was followed by high economic growth in the 1890s, which ended in the so-called "Kristiania crash" in 1899 leading to substantial drops in property values and stagnation for several years. In particular, there appears to have been a downward tendency in overall inequality from the late-1890s to around 1905, followed by remarkable stability from 1905 to 1914. The most dramatic short-run event occurred during the First World War, where we observe the highest Gini coefficient of 0.65 and the smallest mean income ratio between the lower and upper half of the population occurred in 1917, when the mean income of the lower half was only 8.9% of that of the upper half. The low mean income ratio for this period reflects the significant income growth for ship owners and the high speculative profits for wealthy people during a significant economic boom, which was followed by a recession with high inflation, trade deficits and currency depreciation and hardships such as rationing that affected wage earners. As a result, the income of the rich declined and the mean income ratio doubled from 1917 to 1923. However, inequality quickly returned to its pre-war level in the early 1920s and increased slightly during the 1930s.

The simultaneous substantial growth of the mean income ratio and decline of the upper tail Gini coefficient led to a substantial fall in the overall Gini coefficient from 1939 to 1953. Since the mean income ratio stayed fairly flat at around 1/4 since the early 1950s, the rise in overall inequality and affluence after the turning point in 1980 was largely due to rising upper tail inequality. This means that the richest became richer, as is also confirmed by the rising top income shares during this period. The concentration in time of the sharp decrease in the Gini coefficient between 1939 and 1953 is likely a combination of several factors. First, the manner of operation of labor market institutions changed significantly during the 1930s, where collective bargaining was introduced at the national level. Economic turbulence may have postponed the immediate effects of these reforms. Second, more than 40% of the work force was still in agriculture in the 1930s, and rural-urban migration (and hence income equalization) was again constrained by high unemployment. Moreover, the Second World War was likely to have had an equalizing effect in itself, with more controls imposed on the economy where the German occupation led to increased labor demand for extensive construction projects and larger mean income for the bottom half of the population which resulted in increased mean income ratio. Moreover, the German command economy reduced the income opportunities of most capital owners, which might explain why the upper tail Gini coefficient sharply declined. The war experience might also have made Norwegians more receptive to the strict economic planning regime that was introduced during the early post-war period (Espeli, 2013).

Other sources support the finding of a significant fall in income inequality during this period. For example, the 1950 Wage statistics (NOS XI 092, p. 11, Table A) compares wages for various occupations in 1939 and 1950. While high-paid groups such as senior public servants had experienced nominal wage growth of 69%, the wage growth for sailors was 214%, for forestry workers 264% and for farm workers (servants) 380%. For the lower-income groups wage data are also available for 1944; they show that wage compression was well underway during the war.

We observe a stable income Gini coefficient from 1950 onwards, with a further slight decrease in the early 1970s. The fall in income inequality was reversed in the early 1980s. The turning point was largely due to increased wage inequality and came shortly after oil began to flow from the North Sea (Aaberge and Mogstad, 2011). By 1990 production had been at a high level for a number of years. The 1990s show a recovery of the shares of top incomes, probably as a result of expanded

Table 1
Changes in overall inequality, upper tail inequality, ratio of the mean income of the lower and upper 50% and affluence (changes in percentage points in parentheses).

Period	Overall Gini coefficient	Gini-based affluence	Upper tail Gini coefficient	Mean income ratio $\gamma = \mu_L/\mu_U$
1875–1892	Increase (+9)	Increase (+9)	Slight increase (+1)	Decrease (−11)
1892–1914	Decrease (−8)	Decrease (−9)	Decrease (−8)	Increase (+6)
1914–1917	Increase (+9)	Increase (+9)	Increase (+4)	Decrease (−8)
1917–1923	Decrease (−10)	Decrease (−10)	Decrease (−6)	Increase (+9)
1923–1939	Increase (+5)	Increase (+5)	Increase (+5)	Slight decrease (−4)
1939–1953	Decrease (−17)	Decrease (−20)	Decrease (−24)	Increase (+11)
1953–1980	Slight decrease (−2)	Slight decrease (−2)	Slight decrease (−2)	Slight increase (+3)
1980–2017	Increase (+7)	Increase (+10)	Increase (+14)	Slight decrease (−4)

Note: The components are estimated by using the estimated overall means jointly with the same estimated Lorenz curves as were used for estimating the Gini coefficients in Fig. 3. Changes are calculated on the basis of the average of upper and lower bound estimates. For sources, methods and assumptions, we refer to the text.

opportunities to earn and lose money created by the oil sector, a major financial market reform in the mid-1980s, and the 1992 tax reform whereby taxes on capital incomes were significantly reduced. Over the period from 1980 to 2017, the Gini coefficient increased by approximately 20%. The spike in income inequality in 2005 is largely due to the increased taxes on dividends in 2006. This tax reform gave owner-managers of closely held firms strong incentives to increase dividends in 2005. The effects of the reform discussed in further detail by Aaberge et al. (2016) and Alstadsæter et al. (2016) suggest that the level of inequality might have been larger after dividend taxation was implemented in 2006 than what has been captured by the standard income statistics data.

To get some sense of the magnitude of the changes in the Gini coefficient, note that the 22 percentage points fall in the Gini coefficient from 1892 to 1953 (see Table A8) corresponds to a 34% decrease in the Gini coefficient. This corresponds to the redistributive effect of the following hypothetical tax/transfer intervention in 1892 (see Aaberge, 1997): introduce a flat tax with tax rate 34% and allocate the collected tax as a fixed lump-sum transfer equal to the average tax of NOK 178. Then the 50% poorest increase their income on average from NOK 104 to NOK 247, while the 50% richest will get their mean income reduced from NOK 944 to NOK 801. Moreover, this hypothetical intervention would change the income of the poor from NOK 85 to NOK 234 and the 95% quantile from NOK 1630 to NOK 1254.

6. Summary

While data on top income shares provide valuable information on the concentration of economic power, this paper demonstrates that available historic data sources make it feasible to examine the evolution of the income distribution as a whole over long time periods. By combining detailed tabulations with aggregate information on the incomes of municipal and central government taxpayers, as well as administrative data on poverty support, we are able to provide an estimate of the income distribution in 1875 and annually from 1892 to 1951. This is then supplemented with detailed tax tabulations and micro data from 1952 onwards in order to provide income distributions through to the most recently available data for 2017. From these income distributions we can then estimate Gini coefficients, as well as other relevant measures of income inequality and affluence, for consistent definitions of population and income throughout the entire period in question. The proposed method is likely also to be of relevance for other countries.

The empirical results provide three novel insights into the long-run evolution of income inequality in Norway. First, our findings suggest that at the end of the nineteenth century, the Gini coefficient for gross family income in Norway varied between 0.50 and 0.60. Such an apparently Latin American value casts some doubt on the claim made in the official publication for the Paris Exhibition of 1900 that "among civilised states, there is scarcely any that is so fortunate with regard to the equality of its social conditions as Norway. There is no nobility with political or economic privilege, no large estates, no capitalist class" (Norway, 1900, page 203). While Norway has exhibited low inequality from the 1940s till 1980s, we find no indication that this represents a continuation of an earlier egalitarian society.

Second, the movement of income inequality over time appears to be driven by episodic changes rather than predictable, secular cycles. Overall gross income inequality among families in Norway fell from 1892 to 1914, largely due to a fall in inequality in the upper half of the income distribution. There was an upward spike during World War I, and a moderate rise between 1923 and 1939, again largely due to changes in inequality in the upper half. Inequality fell substantially between 1939 and 1953 as a result of a decline in both upper tail inequality and the gap between upper and lower tail means. Income inequality was low and stable between 1953 and 1980 and has risen again since 1980.

Expressed in this way, the history of Norwegian income inequality is better seen as a series of episodes than as the expression of some long-run pattern. It can neither be summarized by an inverse U nor by a U. Moreover, the series of 143 years of income inequality estimates does neither point in the direction of any regular cycles of increasing and decreasing inequality.

Third, it should be noted that the turning point and the origin of the low post-war inequality in Norway was the significant decline in inequality starting in 1940 and continuing during the German occupation. The war experience might also have made Norwegians more receptive to the strict economic planning regime that was introduced in the early post-war period (Espeli, 2013).

A comparison of the levels of income inequality in Norway with previous estimates for other countries is challenging for several reasons: No countries have complete micro data far back in time; there is no universally agreed definition of population or income (as these in turn are dependent on the available data); and there are often breaks even within series for comparable countries. Despite the break in the series for Denmark provided by Atkinson and Søgaard (2016) (based on tabulations from income taxes) these series makes an exception and shows, as for Norway, increased income inequality during World War I and a substantial decrease in income inequality during the mid-20th century. The turning point with increasing income inequality arose in the early 1980s for both countries, but inequality has risen more for Norway than for Denmark.

Declaration of competing interest

None.

Acknowledgement

We would like to thank Andrea Brandolini, Magne Mogstad, Agnar Sandmo, and Espen Søbye and two referees and the editor for helpful comments. Financial support from the Research Council of Norway, grant numbers 231625 and 261985, is gratefully acknowledged.

Appendix A. Robustness checks

A.1. Sensitivity of estimates to the use of additional data sources

Fig. A1 shows how sensitive the estimated Gini coefficients in the 1892–1951 period are to the inclusion of additional sources of data and methodological assumptions. The lowest curve in each panel, G1, displays estimates of the Gini coefficent when the population is defined by taxpayers (MUN and CG) and only one interior point of the Lorenz curve has been identified. This gives a Gini coefficient of around 0.4 before 1939 and less than 0.2 in 1950. By including non-taxpayers and assigning them zero income, we get G2, which shows that the Gini coefficient rose to 0.65 in 1892 and 0.3 in 1950. The G3 curve is obtained by assigning an income to the non-taxpaying population according to the procedure described in Section 3. As a transition from the distribution underlying G2 to the distribution underlying G3 could be obtained (with the appropriate scaling of mean incomes) by means of regressive transfers, G3 is always lower than G2. By taking account of within-group inequality in the richest group, i.e. the CG taxpayers, we get the G4 curve. For years where detailed CG tax tabulations are not available, the closest available earlier or later distribution has been used (the higher one for the upper bound and the lower for the lower bound). Finally, in G5, we also allow for inequality within the group of those who pay municipal tax but not central government tax. This increases inequality moderately, and more so in years when this group is large. It is evident from this exercise that the steps we propose in Section 3 are crucial for a correct estimation of income inequality, though the assumptions on within-group inequality in some groups have only a minor effect on the estimates.

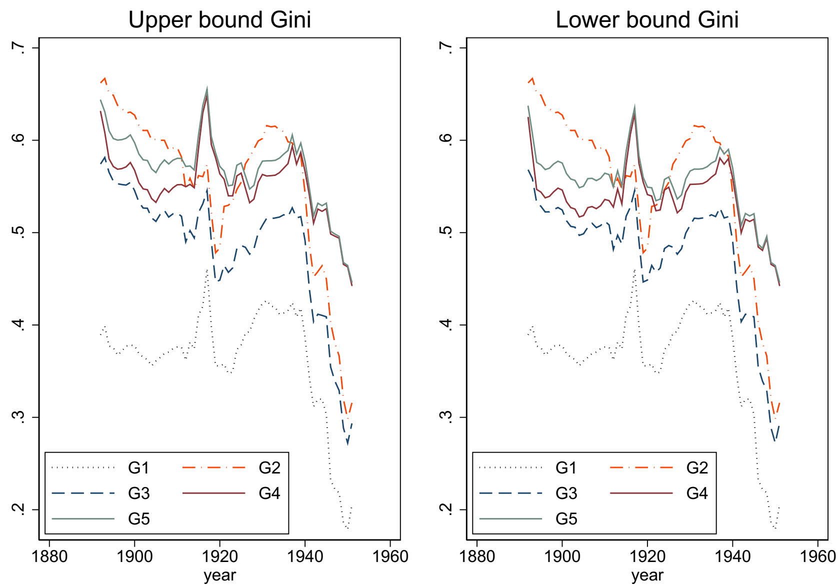

Fig. A1. Estimates of the bounds of the Gini coefficients using five different sets of assumptions, 1892–1951. Note: Definitions: G1: Inequality among taxpayers, no within-group dispersion. G2: Assuming zero income for non-taxpayers. G3: Baseline assumption about non-taxpayer income. G4: Within-group dispersion for CG taxpayers (using nearest year). G5: Within-group dispersion for MUN-CG taxpayers (our preferred estimate, as presented in Fig. 3).

A.2. Sensitivity to choice of inequality measure

To complement the information on inequality provided by the Gini coefficient, we employ two closely related rank-dependent measures of inequality (C_1 and C_3) discussed by Aaberge (2007) and defined by.

$$C_k = 1 - \frac{1}{\mu}\int_0^1 p_k(t)F^{-1}(t)dt, k = 1, 2, 3, \quad \text{(A1)}$$

where.

$$p_k(t) = \begin{cases} -\log t, & k = 1 \\ \frac{k}{k-1}\left(1 - t^{k-1}\right), & k = 2, 3. \end{cases} \quad \text{(A2)}$$

and μ and $F^{-1}(u)$ denote the mean and the left inverse of F. As demonstrated by Aaberge (2007) the measures C_1, C_2 and C_3, denoted Gini's nuclear family, jointly provide a good summary of the information provided by the Lorenz curve. While it can be shown that the Gini coefficient (C_2) tends to pay most attention to changes that occur in the middle part of a typical single peaked income distribution, the two other members of Gini's nuclear family are shown to be particularly sensitive to changes that occur in the lower part (C_1) and the upper part (C_3) of the income distribution.

Note that the ratio of the second term of Eq. (A1) can be interpreted as the ratio between the social welfare attained under the observed distribution F and that attained under complete equality. As a contribution to the interpretation of the inequality aversion profiles exhibited by C_1, C_2 and C_3, Table A1 displays ratios of the weights — as defined by Eq. (A2) — given to the median individual and the 5% poorest, the 30% poorest and the 5% richest individuals, respectively.

Table A1
Distributional weight profiles of the inequality measures C_1, C_2 and C_3.

Relative weights	C_1 (Bonferroni)	C_2 (Gini)	C_3
p(.05)/p(.5)	4.32	1.90	1.33
p(.30)/p(.5)	1.74	1.40	1.21
p(.95)/p(.5)	0.07	0.10	0.13

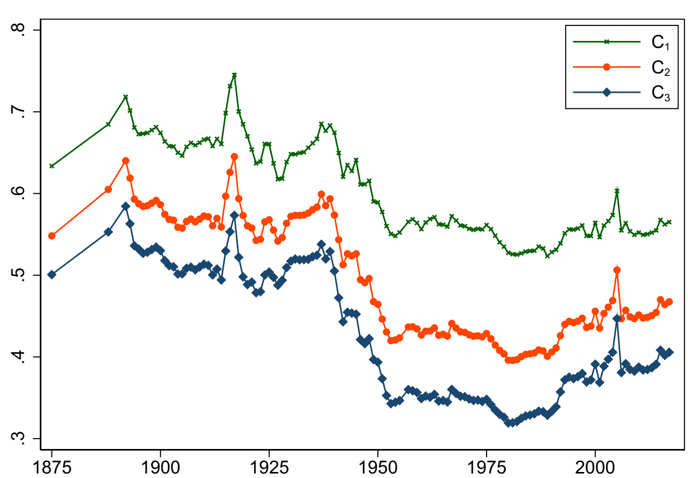

Fig. A2. Long-run evolution of income inequality described by three alternative measures of income inequality, Norway 1875–2017. Note: The components are estimated using the same distribution as the one used for estimating the Gini coefficients in Fig. 3. The graphs display the mean of the upper and lower bound estimates. For sources, methods and assumptions, see text.

As suggested by the table above, C_1 is more sensitive than C_2 to changes in the income distribution that concern the poor, whereas C_2 is more sensitive than C_3 to changes that occur in the lower part of the income distribution. For example, the weights in Table A1 demonstrate that the weight of an additional euro to a person located at the 5% decile is 4.3 times the weight for the median income earner when C_1 is used as a measure of inequality, whereas it is only 1.3 times the weight for the median earner when C_3 is used as a measure of inequality. This means that C_1 is particularly sensitive to changes that take place in the lower part of the income distribution, and C_3 to changes in the upper part of the income distribution.The results displayed in Fig. A2 show that the evolution of the Gini coefficient is largely reflected by the inequality measures that are particularly sensitive to changes in the lower and upper part of the income distribution, respectively. However, the magnitudes of the changes differ significantly for some periods. From 1939 to 1980, the C_1- coefficient decreased by 20%, the Gini coefficient by 33% and the C_3- coefficient by 40%, whereas the percentage changes were almost equal from 1875 to 1939. From 1980 to 2017 C_1, C_2 and C_3 (Atkinson et al., 2017) rose by 7, 20 and 27%, respectively. These results, which are in line with the evidence provided by Section 4.2, show that changes in inequality from 1939 to 1980 and from 1980 to 2017 largely concern the upper part of the income distribution.

Appendix B–G. Data, estimates and supplementary results

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jpubeco.2020.104196.

References

Aaberge, R., 1997. Interpretation of changes in rank-dependent measures of inequality. Econ. Lett. 55, 215–219.

Aaberge, R., 2007. Gini's nuclear family. J. Econ. Inequal. 5, 305–322.

Aaberge, R., Atkinson, A.B., 2010. Top incomes in Norway. In: Atkinson, A.B., Piketty, T. (Eds.), Top Incomes in a Global Perspective. Oxford University Press, Oxford.

Aaberge, R., Atkinson, A.B., 2016. Median as Watershed. Discussion Paper No. 749. Statistics Norway.

Aaberge, R., Mogstad, M., 2011. Robust inequality comparisons. J. Econ. Inequal. 9, 353–371.

Aaberge, R., Atkinson, A.B., Modalsli, J., 2016. The Ins and Outs of Top Income Mobility (Mimeo).

Alstadsæter, A., Jacob, M., Kopczuk, W., Telle, K., 2016. Accounting for Business Income in Measuring Top Income Shares: Integrated Accrual Approach Using Individual and Firm Data from Norway. Discussion Paper No. 837. Statistics Norway.

Alvaredo, F., 2011. A note on the relationship between top income shares and the Gini coefficient. Econ. Lett. 110, 274–277.

Atkinson, A., Piketty, T., 2007. Top Incomes Over the Twentieth Century: A Contrast Between Continental European and English-Speaking Countries. Oxford University Press, Oxford.

Atkinson, A., Piketty, T., 2010. Top Incomes in a Global Perspective. Oxford University Press, Oxford.

Atkinson, A.B., 2007. Measuring top incomes: methodological issues. In: Atkinson, A., Piketty, T. (Eds.), Top Incomes Over the Twentieth Century: A Contrast Between Continental European and English-Speaking Countries. Oxford University Press, Oxford.

Atkinson, A.B., Søgaard, J.E., 2016. The long run history of income inequality in Denmark. Scand. J. Econ. 118, 264–291.

Atkinson, A.B., Hasell, J., Morelli, S., Roser, M., 2017. The Chartbook of Economic Inequality. Institute for New Economic Thinking, Oxford.

Autor, D.H., 2014. Skills, education, and the rise of earnings inequality among the "other 99 percent". Science 344, 843–851.

Blanchet, T., Fournier, J., Piketty, T., 2017. Generalized Pareto Curves: Theory and Applications (Wid.world working paper 2017/3).

Espeli, H., 2013. Economic consequences of the German occupation of Norway, 1940–1945. Scand. J. Hist. 38 (4), 502–524.

Garbinti, B., Goupille-Lebret, J., Piketty, T., 2018. Income inequality in France, 1900–2014: evidence from Distributional National Accounts (DINA). J. Public Econ. 162, 63–77.

Gerdrup, K.R., 1998. Skattesystem og skattestatistikk i et historisk perspektiv. Statistics Norway (Report 98/6).

Kiær, A.N., 1893. Indtægtsforhold i Norge. Tillæg Statsøkonomisk tidsskrift. Aschehoug, Kristiania (1892–1893).

Kopczuk, W., Saez, E., Song, J., 2010. Earnings inequality and mobility in the United States: evidence from social security data since 1937. Q. J. Econ. 125, 91–128.

Kuhn, M., Schularick, M., Steins, U., 2017. Income and Wealth Inequality in America, 1949–2013. CEPR (Discussion Paper 12218).

Kuznets, S., 1955. Economic growth and income inequality. Am. Econ. Rev. 45, 1–28.

Leigh, A., 2007. How closely do top income shares track other measures of inequality? Econ. J. 117, 619–633.

Lindert, P.H., 2000. Three centuries of inequality in Britain and America. In: Atkinson, A.B., Bourguignon, F. (Eds.), Handbook of Income Distribution. Vol. 1. Elsevier, Amsterdam.

Lindert, P.H., Williamson, J.G., 2016. Unequal Gains: American Growth and Inequality since 1700. Princeton University Press, Princeton.

Merwin, C.L., 1939. American studies of the distribution of wealth and income by size. Studies in Income and Wealth. 3. National Bureau of Economic Research, New York.

Milanovic, B., 2016. Global Inequality: A New Approach for the Age of Globalization. The Belknap Press of Harvard University Press.

Morelli, S., Smeeding, T., Thompson, J., 2015. Post-1970 trends in within-country inequality and poverty: rich and middle-income countries. In: Atkinson, Bourguignon (Eds.), Handbook of Income Distribution. Vol. 2A. Elsevier, Amsterdam.

Norway, 1900. Official Publication for the Paris Exhibition. Kirke – og undervisningsdepartementet http://www.archive.org/details/norwayofficialp00konogoog.

Piketty, T., 2001. Les hauts revenus en France au 20ème siècle. Grasset, Paris.

Piketty, T., Saez, E., 2014. Inequality in the long run. Science 344, 838–843.

Piketty, T., Saez, E., Zucman, G., 2018. Distributional national accounts: methods and estimates for the United States. Q. J. Econ. 133 (2), 553–609.

Roine, J., Waldenstrom, D., 2015. Long-run trends in the distribution of income and wealth. In: Atkinson, Bourguignon (Eds.), Handbook of Income Distribution. Vol. 2A. Elsevier, Amsterdam.

Soltow, L., 1965. Toward Income Equality in Norway. The University of Wisconsin Press, Madison - Milwaukee.

Statistics Norway: Historical statistics 1968. NOS XII 245 https://www.ssb.no/a/histstat/hs1968.pdf.

Statistics Norway: Historical Statistics 1978. NOS XII 291 ISBN 82-537-0758-4 https://www.ssb.no/a/histstat/hs1978/hs1978.pdf.

Statistics Norway: Historical Statistics 1994. NOS C 188. ISBN 82-537-3965-6 https://www.ssb.no/a/histstat/nos/nos_c188.pdf.

Vecchi, G., 2017. Measuring Wellbeing: A History of Italian Living Standards. Oxford University Press.

Williamson, J.G., 1985. Did British Capitalism Breed Inequality? Allen and Unwin, London

Williamson, J.G., Lindert, P.H., 1980. American Inequality: A Macroeconomic History. Academic Press, New York.