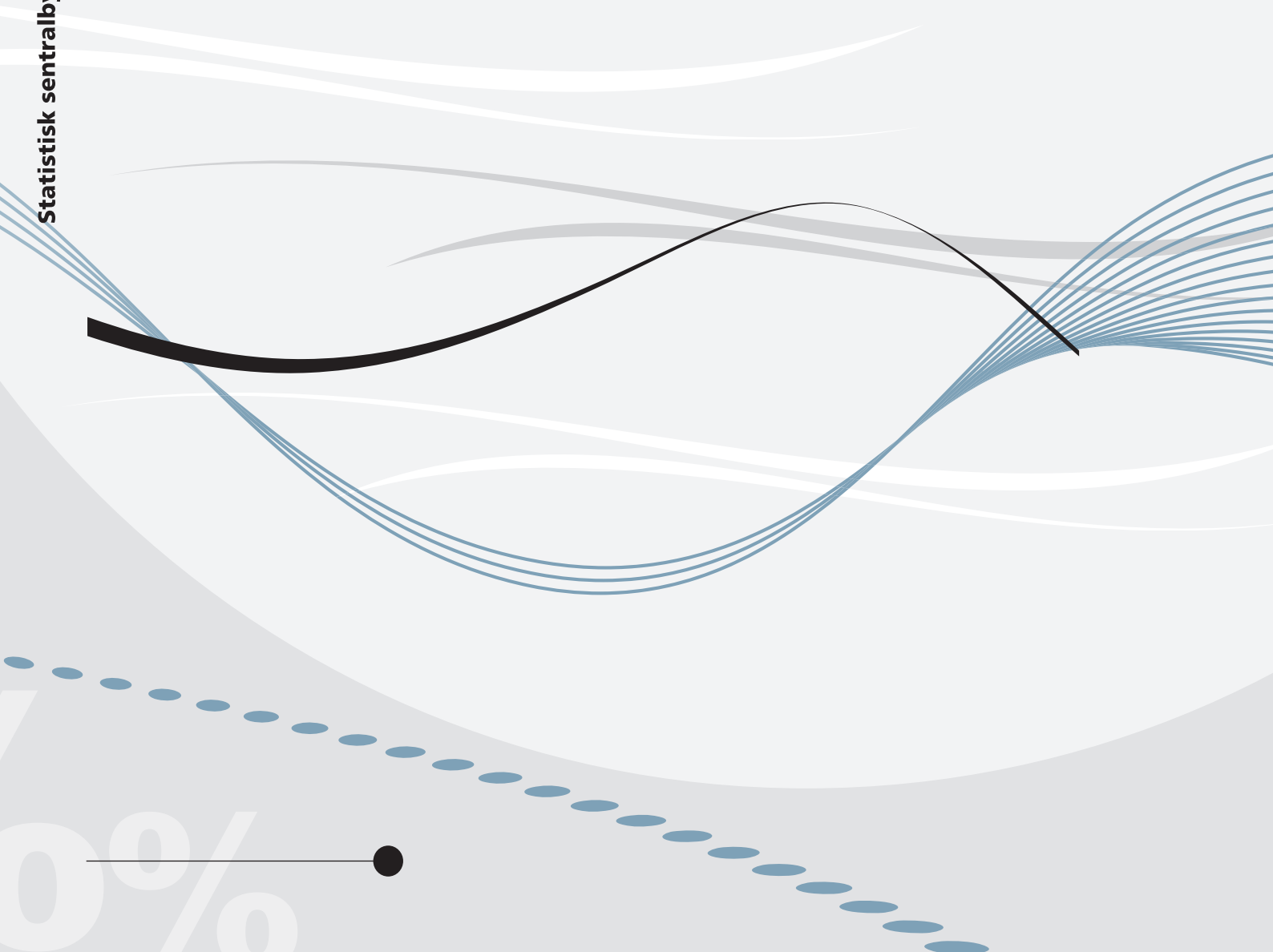
Discussion Papers
Statistics Norway
Research department

No. 847
October 2016

Statistics Norway
Statistisk sentralbyrå

Rolf Aaberge, Anthony B Atkinson and Jørgen Modalsli

On the measurement of long-run income inequality: Empirical evidence from Norway, 1875-2013



Discussion Papers No. 847, October 2016
Statistics Norway, Research Department

Rolf Aaberge, Anthony B Atkinson and Jørgen Modalsli

On the measurement of long-run income inequality: Empirical evidence from Norway, 1875-2013

Abstract:
In seeking to understand inequality today, a great deal can be learned from history. However, there are few countries for which the long-run development of income inequality has been charted. Many countries have records of incomes, taxes and social support. This paper presents a new methodology constructing income inequality indices from such tabular data.

The methodology is applied to Norway, for which rich historical data sources exist covering the period 1875 to 2013. Taking careful account of the definition of income and population and the availability of micro data starting in 1967, an upper and lower bound for the pre-tax income Gini coefficient for core households is produced.

Our findings cast doubt on the idea that Norway in the nineteenth century was an egalitarian society, supporting the view of de Tocqueville that the young United States exhibited less inequality than the states of Europe. We show that overall inequality of gross family incomes is lower today than a hundred years ago. At the same time, there has not been a consistent downward trend over time in inequality; rather, the fall in inequality took place in a series of episodes. Comparison to existing data for Denmark and the United States reveals remarkable commonalities, as well as distinct periods of difference. This supports the view that the evolution of income inequality is best studied, not in terms of an over-arching theory, but by studying episodes of rising and falling inequality, and the manifold forces in operation

Keywords: income, inequality, distribution, Norway, long-run changes

JEL classification: D31, D63, N33, N34

Acknowledgements: We thank Espen Søbye and Kjetil Telle for helpful comments. Financial support from the Norwegian Research Council, grant number 231625 ("People and their incomes in Norway"), is gratefully acknowledged

Address: Rolf Aaberge, Statistics Norway, Research Department. E-mail: roa@ssb.no

Jørgen Modalsli, Statistics Norway, Research Department. E-mail: mod@ssb.no

Anthony B Atkinson, Nuffield College, Oxford

Discussion Papers	comprise research papers intended for international journals or books. A preprint of a Discussion Paper may be longer and more elaborate than a standard journal article, as it may include intermediate calculations and background material etc.

Abstracts with downloadable Discussion Papers
in PDF are available on the Internet:
http://www.ssb.no/en/forskning/discussion-papers
http://ideas.repec.org/s/ssb/dispap.html

ISSN 1892-753X (electronic)

Samandrag

Det er viktig å vite mest mogleg om den historiske utviklinga av inntektsulikskap for å få ei breiare forståing av ulikskap i vår tid. Studiar som viser korleis inntektsulikskap har utvikla seg over eit langt tidsrom er det få av. Men mange land har historisk informasjon om inntekt, skatt og sosialstøtte for grupper i folket som kan brukast som grunnlag for å talfesta mål for inntektsulikskap. Denne artikkelen presenterer nye metodar for å talfesta inntektsulikskap frå slike data, og brukar desse metodane på tilgjengelege norske inntekts- og sosialstøttedata frå 1875 til 2013.

Basert på ein inntektsdefinisjon som er konsistent over tid gir vi ei øvre og nedre grense for ulikskap (målt ved Gini-koeffisienten) i fordelinga av markedsinntekt for kjernehushald (einslige og gifte par). I tråd med Toquevilles (1835) syn om høgre ulikskap i Europa enn i USA finn vi at Norge på 1800-talet ikkje var spesielt egalitært - inntektsulikskapen er vesentlig lægre i dag enn for 100 år sidan. Samstundes har det ikkje vore ein vedvarande nedgang i ulikskapen over tid; ulikskapen er stabil i nokre periodar, men fell og stig i andre. Den utviklar seg episodisk. Dette viser at utviklinga av ulikskap over tid ikkje kan forklarast ved enkle økonomiske teoriar, men heller som eit samspel av ei rekkje komplekse økonomiske krefter.

1 Introduction: Inequality in the long-run

Few countries have data on income inequality providing information for the whole population covering a century or more in a continuous time series. The studies initiated by Piketty (2001) provide rich information but this is typically limited to the top income groups, as in the study of top incomes in Norway from 1875 by Aaberge and Atkinson (2010). Modern series covering the whole income distribution were rarely available until after the Second World War. The reason is evident. The main instrument used today to record the whole income distribution is the household survey, and such surveys only came into regular use at the national level in the latter part of the twentieth century.

This means that there is only limited evidence about income inequality for years before 1945. The most widely-quoted Gini coefficients for the United States begin in 1929; the first in the modern series for the United Kingdom is for 1938 (see Atkinson and Morelli, 2014). While estimates certainly exist for earlier years, they are not comparable with modern series. For the United States, figures were given by Spahr for 1890 and by King for 1910 (see Merwin, 1939), but they are described by Williamson and Lindert as "eclectic size distribution guesses", and they conclude that "it is better to pass over these" (1980, page 91). Williamson (1985) has given figures for the Gini coefficient for England and Wales, and Scotland, for years ranging from 1688 to 1915. Again, these are not readily linked to the modern series: they are shown in separate tables in the survey by Lindert (2000). There are very few countries that have a continuous series on income inequality back to the nineteenth century.[1]

There is however considerable interest in viewing a continuous long-run series for overall income inequality. From Kuznets' Presidential Address (1955) onwards, long-run changes in inequality have been the basis for the development of theories of structural change, and there has been much discussion as to whether inequality first rises and then falls with industrialization. If incomes today are indeed less unequal than in the past, when did inequality fall? Have there been distinct periods when inequality has fallen or risen? What was the impact of shocks such as the First and Second World Wars? What was the effect of the Great Depression? The 1930s is a period where we know little about the development of the income distribution. In the United States, there has been a large literature on

[1]Atkinson and Søgaard (2016) calculate a wide bound on the Gini coefficient for Denmark in 1875 and then from 1903 onwards (though with some changes in data definitions over time). The longest existing series for income inequality in Norway were reported by Soltow (1965), who constructed a series of Gini coefficients based on samples of tax records. These are limited to eight Norwegian cities for selected years from 1850 to 1960.

the Great Depression, but "the role of income inequality before and during the Great Depression …has almost never been discussed thoroughly" (Belabed, 2016, Abstract). In the post-war period, are the recent changes in income inequality large or small in comparison with those in the past? Historical comparisons help place today's inequality in perspective.

The aim of this research is to demonstrate what can be said about the evolution of the overall income distribution in countries for which there are rich historical data. To this end, we present and explore a new series on the distribution of income in Norway as a whole spanning the period from 1875 to 2013. In constructing this series, we restrict attention to sources that provide information about the incomes of individuals and their families, where income is the total from all sources (earnings, investments and transfers) before deduction of tax, referred to here as "gross income". The information may be quite limited, such as the number of people who are below a particular income level (for example, the tax threshold). But it implies that we do not use, for example, data on the distribution of people by occupational groups, attributing to them an average income or wage (often referred to as the "social tables" approach – see Lindert, 2000, page 174).

Our starting point is the information provided by the detailed income tax tabulations of incomes by ranges as used by Aaberge and Atkinson (2010) in their study of top income shares in Norway. In Section 2, we describe the tabulated data available from the published income tax records from 1875 and in the form of micro-data from 1967. The tax information is a rich source, but it varies in form from year to year, and has to be used with caution for reasons explained in Section 2. It is also limited in coverage of the population, as it excludes non-taxpayers, who in the earlier years constituted the majority of the population. Using income tax data for periods when taxpayers were a minority may appear a triumph of hope over experience, but we show in the paper that overall inequality as measured by the Gini coefficient is governed very much by what happens to the top half of the distribution, and, as a consequence, the calculations are less sensitive to the assumptions made about the incomes of those in the lower half.

The incomplete coverage of the population in the tax data means that there is a challenge in seeking to measure overall income inequality, as represented here by the Gini coefficient. This challenge we meet by creating "upper" and "lower" bounds on the Gini coefficient. These are not bounds in a mathematical sense. The Gini coefficient can, for example, be raised to 100 per cent by assuming that enough of the missing population has large negative incomes. Rather the bounds are based on judgments. These judgments are open to debate, and we explain in detail in Section 3 the assumptions

made here. They involve introducing information from sources beyond the detailed income tax tabulations by ranges. In particular, we bring to bear aggregate information from the municipal and central government tax records, which is available annually for a long period. The additional information, coupled with assumptions about the relative position of different groups, allows us both to say more about the years before the Second World War and to narrow the bounds on the estimated Gini coefficient. To this, 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. Since the additional sources may be available for other countries, this methodological discussion is of wider interest than the application in this paper to the case of Norway.

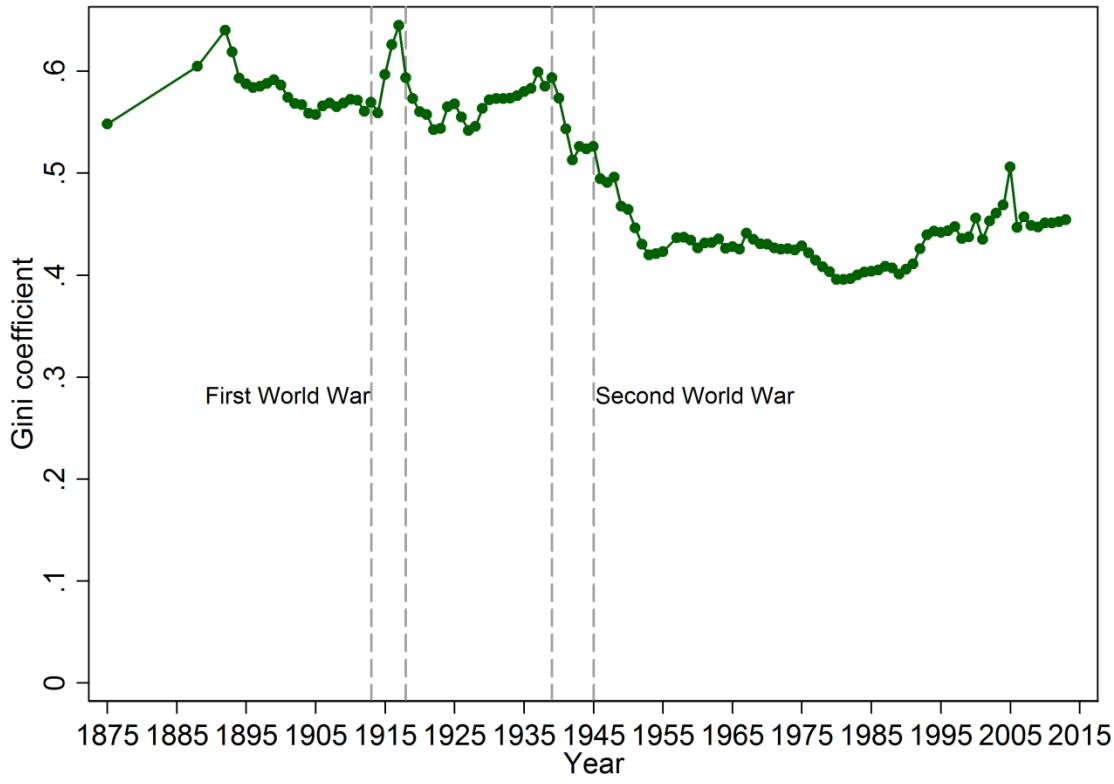
In this way, the paper demonstrates that much can be learned from administrative data even in cases where individual observations are not available (such as in Norway before 1967). By combining tabulations from different sources, a more complete picture of the distribution than what one obtains from (state) tax records alone can be attained. Similar procedures can likely be applied to other countries to examine whether the development found here for Norway carries over to other institutional and geographical settings.

The historical series is presented in Section 4, where we discuss the main features of the evolution of income inequality in Norway. Figure 1 provides a preview, showing the Gini coefficient since 1875 (this is an "average" series based on a simple mean of the upper and lower bounds). On the basis of the results in Section 4, we attempt to answer a number of questions. There is the long-run question as to how inequality today compares with that a century or more ago. How does the present day Gini coefficient compare with those found before 1914? The second set of questions concerns how we got from there to here. Have there been distinct periods when inequality has fallen or risen? Was it all a war-time phenomenon? Figure 1 already suggests the beginnings of an answer to this question: wartime is not the only explanation. A third set of questions concerns the underlying causes of peacetime equalization. Has there been an "inequality turn" in recent decades? Where does that leave Norway today?

In seeking to cover such a long period, we are constrained by the information contained in the available sources. As a result, the definition adopted here for the long-run series differs from the income distribution statistics produced today by Statistics Norway (website, "Income statistics for households"). Our series differs in three principal respects: (a) it relates to gross income including taxable transfers but before the subtraction of direct taxes, (b) family income is not adjusted for family

size, a tax unit being treated as a unity regardless of its composition, and (c) the unit of analysis is the inner family (defined as a single person, a couple, and any dependent children) rather than the household. Our choices reflect the constraints imposed by the material at our disposal, but there are intrinsic grounds for defending their use. From the standpoint of policy, there is now increasing concern about the determinants of gross incomes, in the face of the recognition of the limits to redistribution via taxation. As our results show, the Gini coefficient for gross incomes in Norway today is quite a lot higher than the disposable income figure that typically enters public discussion. Equally, the present-day focus on total household income is open to the objection that it assumes an unrealistic degree of income sharing within the household, where different family units may have different access to resources. And, finally, the need for equivalisation is lower in the case of inner family than for the more extensive household. This said, we recognize the interest in disposable income, and Section 5 shows the results – covering a shorter time period – for alternative definitions of income.

Figure 1: Preview of result: Gini coefficient in Norway 1875 to 2013 (average of upper and lower bounds)



In Section 6 we compare this series for Norway with the Gini coefficients for the same period for Denmark, and with the shorter period from 1918 for which overall inequality measures are available

for the United States. This comparison reveals a remarkable similarity in the movements of the two series for a long period: for some hundred years up to the mid-1970s.

The main conclusions are summarized in Section 7.

2 The income tax data in Norway

We begin with a brief account of the Norwegian income tax data, and the way in which they can be used to produce results for the income distribution as a whole. 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). This means that some sources of non-taxable welfare payments are not included. However, as the data for all income groups are obtained from official sources we expect the disconnect between observed and true income to be smaller than if one were to combine, say, average market wages with the distribution of occupations. Throughout the entire period, the same conceptual income definition is used, namely "alminnelig inntekt" (*common income*). This refers to income before tax, but after some pre-tax deductions.

Self-employment income is accounted for by a assessments of the productive capacity of farms (in particular for smaller farms) and derivation from company accounts. In aggregate statistics for some time periods, incomes for companies (in addition to individuals) are included; where this is the case, we adjust the calculations to account for this.

The income tax data from 1875

The income tax sources are municipal (MUN) and central government (CG) tax assessments: *Kommunenes skattelikning* and *Statsskattelikningen*.[2] The key feature here is that, for a number of years, the government has published tabulations of the distribution of income taxpayers by range of income. The sources are listed in Appendix B. As the MUN tax data are more extensive (tax thresholds are lower and there are more people paying MUN than CG tax), we assume that CG taxpayers are a subset of MUN taxpayers.

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

The coverage of these income tax tabulations varies over the period. The CG tax was introduced in 1892, so that the prior years have only distributional information from the MUN tax. The published tabulations for 1892 to 1903 only relate to the 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.

This may seem like lean pickings. However, we may supplement the distributional data with aggregate information on the total number of MUN 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 the case (iii) a further point corresponding to the total MUN taxpayers (and hence total taxpayers).

The tabulations of taxpayers by income ranges from 1952 to 1966, which precede the micro-data available from 1967, vary in their coverage (see Appendix B). Income is equal to assessed income by the municipal tax assessment for the years 1952-55. In the tabulations for the years 1957 to 1966, income is defined as assessed income by the central government tax assessment if central government tax is levied. If not, income is defined as assessed income by the municipal tax assessment. (There are no data for 1956 on account of the introduction of Pay-as-You Earn.)

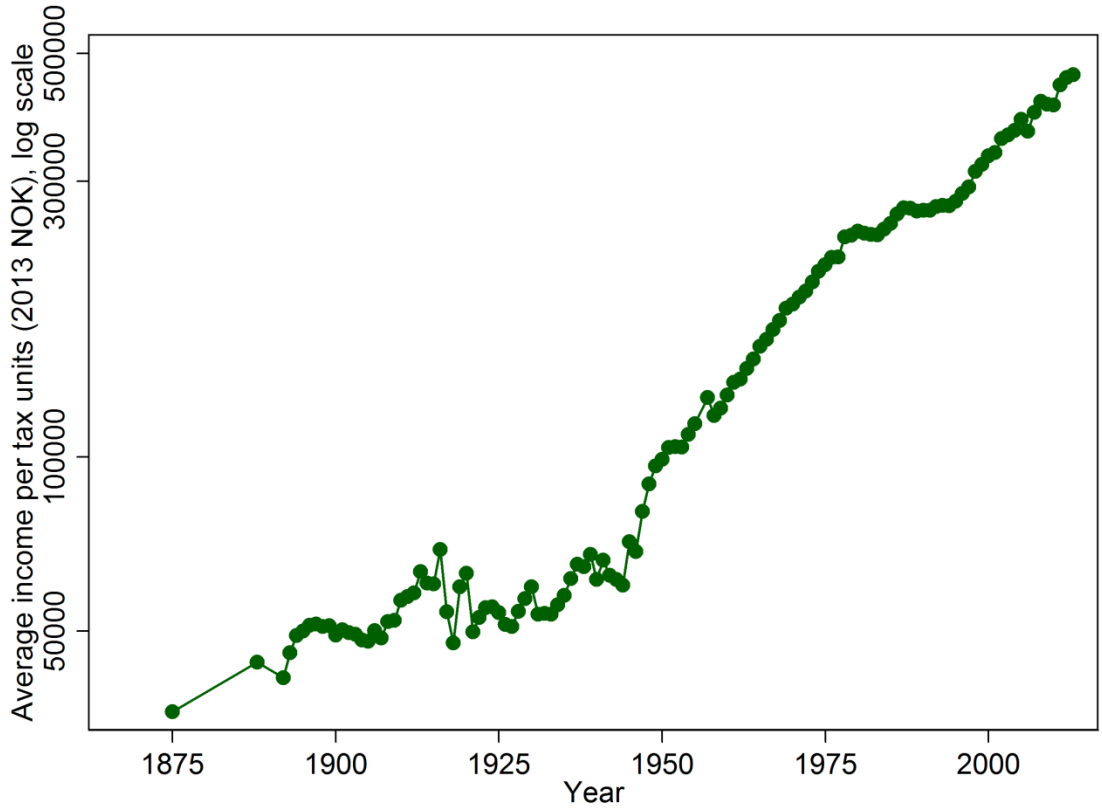
After 1967, all individual incomes are available on computer files at Statistics Norway. The income concept used is "alminnelig inntekt", income post some deductions, which is identical to the pre-1967 tabulations. Using data from the Central Population Register, we merge married couples into single units, adding the income of husband and wife to form the inner family.

Control totals

In all years, the CG and MUN income tax tabulations cover only a fraction of the total population. In order to arrive at an estimate of income inequality across the entire population, rather than only among the taxpayers, the tax data have to be combined with independent estimates of the total number of tax units and the total of household income. The sources for these “control totals” are described in 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. For total income, the starting point is a series for total household income provided for 1978 to 2013 by the National Accounts. Conceptually, total household income 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 elsewhere classified. In order to extrapolate this series backwards, we have made use of historical series that are as comparable as possible. As in the study of top incomes presented in Aaberge and Atkinson (2010), the control total is taken as a percentage (72 per cent) of the national accounts total household income, to allow for the more extensive coverage of the latter.[3]

Figure 2: Mean income per tax unit (CPI adjusted) in 2013 NOK



The control totals yield an estimate of the mean income per tax unit and this is expressed in Figure 2 in real terms (as 2013 NOK). Over the period since 1875, real income has risen by a factor of around 13.[4] But the growth has not been steady. Before 1914 there was an irregular pattern of growth and downturn. The inter-war period saw little improvement in real incomes. The post-Second World War period, in contrast, experienced rapid growth up to the mid-1970s, which later slowed and was

[3] A comparison of the National Accounts control total to the internal total using only taxpayer data is given in Appendix C.

[4] GDP per capita (in fixed prices) has grown by a factor of 18 over the same period. The discrepancy largely comes from the large demographic changes over this period; Norway in 1875 had a much younger population. 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.

interrupted by the banking crisis of 1988 to 1993. This macro-economic experience makes it all the more important to investigate what happened to the distribution of income.

3 Bounds on the Gini coefficient

We now move to an 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 will 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 core households and also to get a measure of the precision of our estimates.

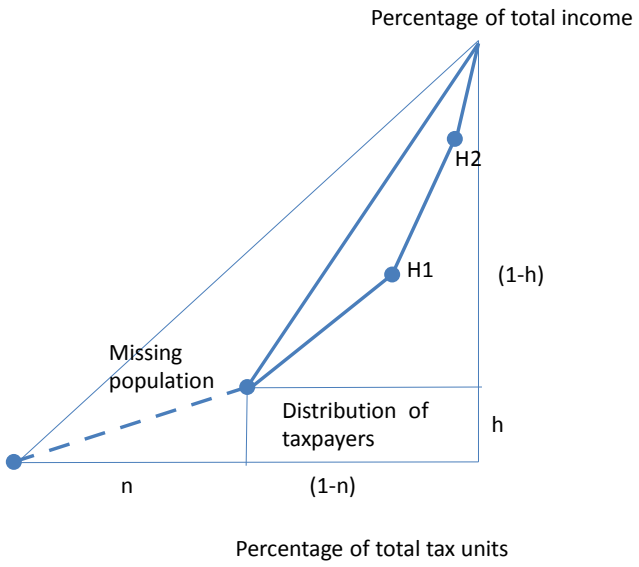
The discussion in this section will be based on the available Norwegian historical data sources. However, having several types of income tax as well as data on social assistance is by no means unique to Norway in this period. For this reason, the methods proposed here, utilizing tabular data to assess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

Different formats of the overall Lorenz curves are shown in Figure 3, which relates to the case where we have distributional information on MUN taxpayers (with or without information on CG taxpayers) and Figure 4, which relates to 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 of MUN and CG taxpayers have incomes lower than those who pay tax. This means that the Lorenz curve for taxpayers is scaled down and joined with the final point for the missing population. In the case shown, the missing population are all assumed to have identical incomes, so the first section of the Lorenz curve is a (dashed) straight line. Further assumptions made about the distribution within the missing population are discussed below. The points H1 and H2 indicate points on the Lorenz curve constructed from the MUN and CG taxpayer data. Figure 4 shows the case where there is no tabulated MUN data, only the aggregates. On the assumption that those paying the MUN tax but not the 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". Over the period 1875 to 1951, there was a difference of around 20 per cent between the internal and control totals (see Figure A 2), apart from during the First World War. 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 to either under-statement in the tax data or to the missing population. Suppose that the excess of the control total over the internal total is equal to a proportion, α, of the internal total, and that a proportion β of the internal total is assumed to represent under-statement in the tax data. This leaves (α-β) times internal total income to be allocated to the missing population, or (α-β)/(1+α) 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 (α-β)/[n(1+α)]. This is the overall slope of the first segment of the Lorenz curve.

Figure 3: Distribution of taxpayers and missing population



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 is equal to 1 minus twice the area under the Lorenz curve. For taxpayers alone, twice the area is equal to

$$B = \Delta F_1 \, \mathrm{H}_1 + \Delta F_2 \, \{H_1 + H_2\} + \ldots + \Delta F_k \, \{H_{k-1} + 1\} \qquad (1)$$

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

$$G^* = 1 - B \qquad (2)$$

The introduction of the missing population as in Figure 3 has two effects. It squeezes the Lorenz curve for taxpayers to the right. In equation (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 Figure 3 initially follows the horizontal axis. Together, these two effects give an upper bound G_U for the overall Gini coefficient, which can be calculated to be

$$G_U = n + (1-n)G^* = G^* + n(1-G^*) \qquad (3)$$

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

In the opposite direction, 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 (1892-3), which we are using, included an estimate of the numbers and income of the missing population[5]. The mean for the range NOK Norwegian kroner) 0 to 400 was NOK 230, which was 40.9 per cent 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 345.5 million NOK internal total to the 475.8 million NOK control total, 32.6 million NOK, or 28.3 per cent 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 increased therefore 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 2*h*(1-*n*). The resulting lower bound Gini is

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

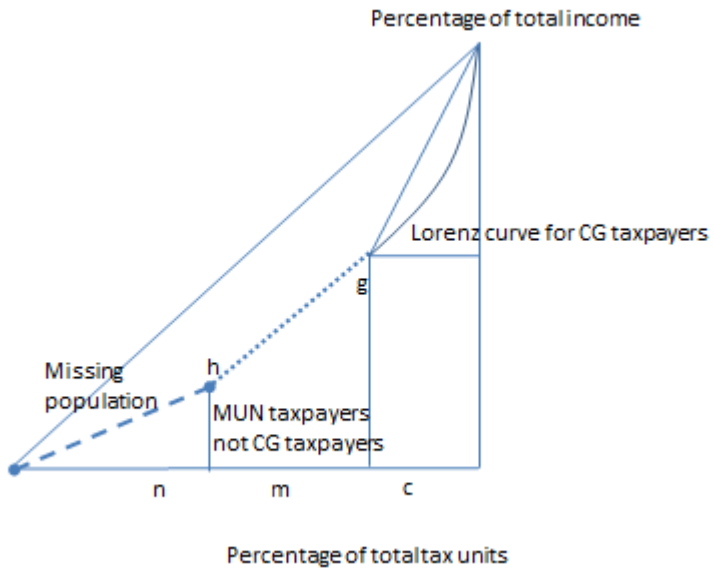
$$G_L = n + (1-n)G^* - h\left[1 + (1-n)G^*\right] = G_U - h\left[1 + (1-n)G^*\right] \qquad (4)$$

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

Using aggregate data on taxpayers

For certain years, we have only the aggregate number and total income of the MUN taxpayers who are not liable to 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 Figure 4.

Figure 4: Distribution from central government tax, aggregate of municipal taxpayers and missing population



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 is denoted by n (so *c*+*m*+*n* = 1). The contributions of the three groups to the overall Gini coefficient may be seen from Figure 4. 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 :

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

where G^* is the Gini coefficient among the CG taxpayers. This may be re-written by introducing a new parameter $g' = g\text{-}h$ and replacing g by $(g'+h)$ as

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

The upper bound is obtained by setting $h = 0$ and holding the other parameters constant. The final term in (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 are only CG data, m = 36.6 per cent, c = 18.8 per cent and G^* = 44.8 per cent, so that the magnification factor is 1.45.

What, if anything, can we say about years for which there are no detailed tabulations of the CG taxpayers? The formula (5a) allows us to see the role played by inequality within the group of CG taxpayers when $h = 0$. The term $c(1\text{-}g')G^*$ is an addition to the overall Gini coefficient. Suppose that we do not know G^*, but do know c and g'? So 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. While in the nineteenth century, the product may have been small, it was substantially higher in the First World War and later. On the other hand, in the years when we have tabulations, the Gini coefficient among taxpayers has rarely exceeded 50 per cent or fallen (apart from two exceptions) below 30 per cent. In what follows it does not seem unreasonable where G^* is not known to base the upper bound estimate of the Gini coefficient on an assumed 50 per cent and the lower bound on an assumed 30 per cent.

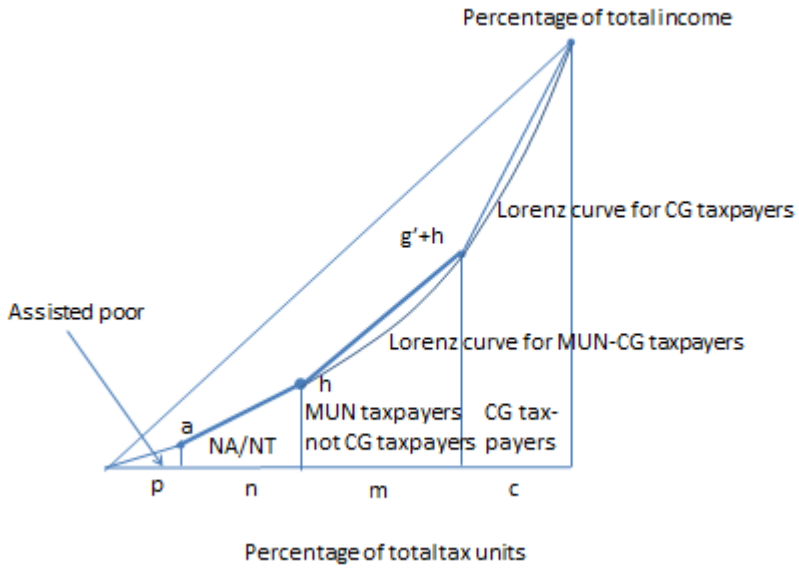
Using data on the assisted poor

In order to provide more foundation for the treatment of the lower part of the distribution, we need additional information on the incomes of those who are 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 recipients have no other source of income apart from the assistance received. To the extent that

they have other income, the degree of inequality is over-stated. Operating in the opposite direction is the assumption that the total paid in public assistance is divided equally among the assisted population.

In effect, use of this additional administrative information means introducing into the three-group model a fourth group: dividing those not paying tax into those who are assisted (the "assisted poor") and 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 Figure 5. This is can only be approximately correct. Liability for taxation depends on both income and wealth. People may be liable for MUN taxation on account of wealth but have low incomes that would place them below people in the NA/NT group. But it seems a reasonable first approximation.

Figure 5: Four groups (including assisted poor)



Where the proportion of assisted poor is denote by *p*, and the proportion in the NA/NT group by *n*, and the share of the first group is denoted by *a*, then 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. Total income received by the assisted poor, and by the two groups of taxpayers, is known. Unknown is the total income of the NA/NT group. Here we have to make assumptions regarding the upper and lower bound, 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.[6] 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 bound.

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 are "sandwiched" between other 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 dispersions in these groups. For example, the maximum consistent inequality in the poorest group would have the richest individuals in this group obtaining the same income as the NA/NT mean income and the poorest individuals in this group obtaining zero. The effect of such a distribution would be largest in 1888, where the lower bound Gini would increase only from 56.91 to 56.93.[7]

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 outlined in Appendix F, where the relationship between the dispersion parameter *z* and the within-group MUN-CG Gini coeffficient $G^{**} = z/3$ 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 introducing 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*.

---

[6] A number of further adjustments have to be made to the published tabulations in making these 4-group calculations. Assumptions are necessary when calculating 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 next observation of G^*. Similarly, the lower bound uses the minimum of the previous and next observation if there are no data. For the years 1875 to 1891, when there was no CG taxation, and the MUN-CG group is not defined, the average income of the NA/NT group in the calculation of the upper bound Gini is taken as NOK 150. NOK 150 was 25 per cent of the mean income of workers and 33 per cent of the mean income of farmers (including cotters) in 1888/89 (Sth. Prp. Nr 48, 1890).).

[7] 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 to the overall Gini from the poor group) is $\frac{a}{2} \cdot \left(p - \frac{an}{h-a}\right)$. Introducing dispersion to the NA/NT group would decrease the maximum consistent contribution from the poor group.

Finally, in 1875 and 1888 (the years before the introduction of the 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 groups in later years, and that the higher-income groups would have been subject to CG tax had that been in effect these years.

To sum up, this gives the Gini coefficient for the years 1875, 1888 and 1892-1951 as

$$G = 1 - pa - n(a+h) - m(2h+g') - c(1+g'+h) + c(1-g'-h)\ G^{*}\ ) + g'm\,G^{**} \qquad (7)$$

where

a = total income of poor relative to control total,

h = total income of poor and non-assisted/non-taxed (NA/NT) relative to control total,

g = total income of poor, NA/NT and MUN-CG (adjusted) relative to control total,

$g' = g-h$ = total income of MUN taxpayers who are not CG taxpayers,

p = 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 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 Figure 5. The latter two terms add in the within-group Gini coefficients for the two richest groups, scaled by group sizes and income shares. [8]

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

Bounds for 1952 to present

The above discussion has described the bounds applied for the period 1875 to 1951. For the period after 1951, when coverage was greater, relatively high numbers of tabulated intervals have been published by 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, corresponding to the situation shown in Figure 3, 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 series of poverty support 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 assuming (i) that those not covered by the tax tabulations have a mean income equal to the mean assistance (as before) for the years up to 1964 and (ii) that from 1967 it is based on the group receiving 50 per cent of the minimum pension for a single person.[9] The lower bound is based on those not covered by the tax tabulations receiving mean income equal to 150 per cent of the mean income assumed for the upper bound.

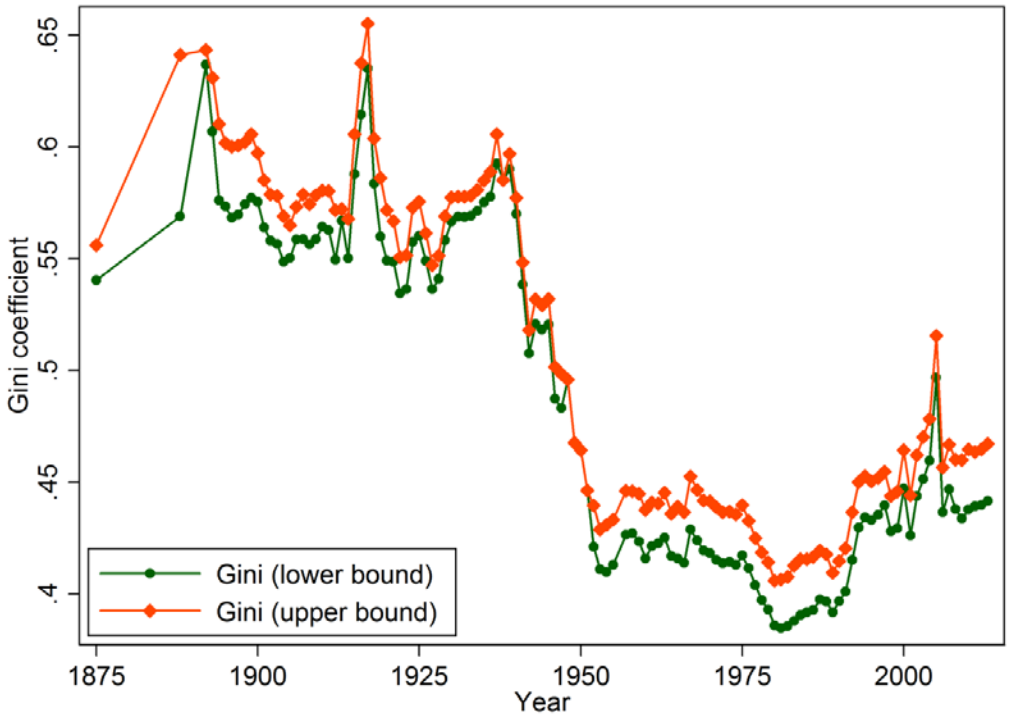
We should emphasize at this point that the final series is based on a consistent population throughout the period. Despite the change from household-based to individual-based taxation, we can replicate the core households from before 1960 on the post-1966 microdata by merging spouses using personal ID numbers in the latter that link taxpayers and the population recorded on an individual basis. The first year in which married women could choose to file taxes individually is 1960. For the years 1960-1966, we therefore transform the data to household basis using data from the 1960 Census as well as the joint distribution of income, marriage and tax status in 1967. Similarly, adjustments are applied to account for a separate taxation system for sailors (1948-1966) and company taxation (1921-1947). These adjustments are all described in detail in Appendix E.

[9] 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 2,140, which was close to the poverty support level in that year of NOK 1,975.

4 The long-run series: Income inequality in Norway 1875 to the present

The results of these calculations are brought together in Figure 6, which shows the upper and lower bounds for the Gini coefficient. , The difference between the upper and lower bounds is largest for the period before 1914. 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. The latter seems quite modest. 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. On that basis, the 1892 to 1914 figure appears quite large, but the 1915 to 1951 average difference is not dissimilar from the confidence intervals obtained from the reported standard errors for the Gini coefficients obtained from household surveys: for example, the 95 per cent 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.

Figure 6: Gini coefficient for Norway, 1875-present. Upper and lower bound



Sources: See text.

The series prior to 1915 must therefore be regarded with more caution than that for the past 100 years. This applies particularly in the earlier Figure 1, where we take the mean of the upper and lower bounds to give an “average series”. The averaging is done, since we recognize that a single series is

what many researchers require and that, if we do not ourselves provide an average, users will do so. At the same time, there is no evident justification for taking a simple average. A case can be made 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.
One immediate question is: what additionally do we learn from the new series in Figure 6 compared to the top income series previously published by Aaberge and Atkinson (2010)? In a number of periods, overall inequality and top income shares move closely together. Between 1939 and 1953, for instance, the Gini coefficient fell from 59 per cent to 41-43 per cent, and the share of the top 1 per cent fell from 13 per cent to 7 per cent. However, the experience of the next three decades was rather different. There was a significant decline in the share of the top 1 per cent, from 7.1 per cent in 1953 to 4.1 per cent in 1989, whereas over the period as a whole the Gini coefficient was little altered: in 1989 it was 39-41 per cent. Since 1989, the share of the top 1 per cent has regained the lost ground, being 7.8 per cent in 2011, and the Gini coefficient too has risen – although only to around 45 per cent. This difference between the time paths of the top shares and the Gini shows that, while the top share may have driven much of the recent increase in overall inequality, there were other forces in operation that mean that not all of the post-war equalization has been lost.

If we ask whether Norway is back where it was a hundred years ago, just before the First World War, then the evidence for top shares is limited to two observations (1910 and 1913). In our Gini series, we have constructed annual observations. These show that the Gini coefficient, averaged over 2000-2011, was 46 per cent, compared with 57 per cent in 1900-1914 averaged. Overall inequality has not returned to the levels of 100 years ago.

Second, there have been distinct periods of rise and fall in overall income inequality. There was a fall of some 6 points in the averaged Gini coefficient between the late 1880s/early 1890 and1914, followed by a volatile period that included both significant increase and decrease between 1914 and 1923 (with a net change of 2 points), a rise of 4 percentage points from 1923 to 1939, a large fall, as noted above, from 1939 to 1953 of 17 points, a fall of 2 percentage points from 1953 to 1989 and a rise of some 5 points since 1989. Taken together, -6, -2, +4, -17, -2 and +5 yields an overall change of -18 percentage points. Thirdly, it is evident from Figures 1 and 8 that the changes in overall income inequality have been quantitatively larger – in both directions – in the second part of the period. Leaving aside the World Wars (and 1892), the Gini coefficient in the four decades from the 1890s to the end of the 1930s was in the range of 60 per cent plus or minus 5 percentage points. The Second World War and

early postwar decline was much larger, and the post-1989 reversal took the Gini from around 40 per cent to around 45 per cent in two decades.

The long-run history of income inequality in Norway is indeed a rich story that needs to be considered in terms of episodes of change. We turn now to consider the individual sub-periods in more detail.

Before 1914

The wide bounds for this period limit what we can say, but this is a period of considerable intrinsic interest, and one for which few countries have data about the distribution of income on a regular basis covering the period of industrialization. For example, over the 20 years from 1894 to 1914, the upper bound fell by some 4 percentage points and the lower bound fell by 3 percentage points. There was a moderate convergence of the two bounds that tells us that inequality cannot have changed dramatically during this period. There is some evidence of increasing inequality between 1875 and 1888, but from 1888 onwards there is little sign of a rise in overall income inequality as would be associated with the Kuznets curve.

The nineteenth century data may tell more about change at a less grand scale. For example, Norway was hit by a depression from around 1876 onwards (Grytten 2008b). Growth rates were low, 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 over the years from the mid-1890s to around 1905, followed by remarkable stability from 1905 to 1914.

The World Wars

The Gini coefficients in Figure 6 show a sharp rise during the First World War, peaking in 1917. Norway was neutral, but its merchant fleet played a significant role (Grytten, 2008b). However, there was a severe recession that replaced a boom during the end of the war. Historical statistics provided by Statistics Norway show a drop in GDP per capita (at fixed prices) of 10 per cent from 1916 to1917 and 5 per cent from 1917 to 1918. There was also high inflation during this period. As demonstrated by Figure 6 the Gini coefficient was heavily affected by the boom in 1916 and the subsequent recession in 1917.

In contrast, the Second World War saw a marked fall in overall inequality and in the share of the top 1 per cent. The Gini coefficient was some 6 percentage points lower in 1945 than in 1939. The

circumstances were very different, with Norway, like Denmark, being invaded by Germany in 1940 and occupied until 1945. In this case, there is evidence on top shares for a range of countries, and this shows that declining top income shares during the Second World War was a quite widely-experienced phenomenon. Of the seventeen countries for which there is evidence on the share of the top 1 per cent, in all but two (South Africa and Southern Rhodesia) the top share fell between 1939 and 1945 (Atkinson, 2015, page 57). At the same time, reductions in top shares were not limited to the war years. In the case of the Norwegian Gini coefficient, the decline from 1939 to 1945 accounted for only 6 points out of the total decline of 17 points from 1939 to 1953.

The periods of World War were, we hope sui generis, but the Norwegian experience, with evidence from both, may serve to dispel the impression that it was during wartime that large reductions in inequality were secured. Overall inequality actually both rose and declined during the First World War. The fall in the Second World War only accounted for a fraction of the reduction that took place during the twentieth century.

Inequality in peacetime

What do we find if we turn to peacetime periods? For the inter-war period, from Figure 6, it may be seen that there was an initial fall and then a fairly steady rise in the Gini from 1923 onwards. The interwar period saw substantial economic hardship, including a banking crisis during the early 1920s where the five largest banks went into bankruptcy.

Both upper and lower bounds show a rise in the Gini coefficient of 4 percentage points between 1923 and 1939. No special significance attaches to 1929 (the US stock market collapse). We return below to the rather different experience of Norway, and other Nordic countries, during the Great Depression.

After the Second World War, inequality fell. Between 1946 and 1966, the bounds fell by between 6 and 7 percentage points. The fall continued according to the register data: between 1967 and 1980, the bounds fell by between 4 and 5 percentage points. This fall in inequality was reversed at the end of the 1980s. Over the period from 1989 to 2013, both the upper and lower bound of the Gini coefficient increased by some 5 percentage points. It has to be remembered that the graph shows bounds, not where we are located between these bounds. However, even if, in an unlikely event, the true value had been at the upper bound in 1989 and the lower bound in 2013, there would have been an increase of 3 percentage points. On the other hand, most of the increase in the Gini coefficient took place between 1989 and 1995: 71 per cent of the increase in the upper bound and 83 per cent of the increase in the

lower bound. It was more of a step up than a continuing upward trend. For this reason, it is better to talk about "increased inequality" than about "rising inequality".

Upper tail Gini, mean income gap and affluence

Before the Second World War, the detailed tabulated data by range cover only the highest-income part of the population, which means that we have a better informational basis for describing the upper than the lower tail of the income distribution. In view of this, we now provide estimates of the mean and Gini coefficient for the 50 per cent richest proportion of the population and use this information as a basis for estimating a measure of "affluence" introduced by Aaberge and Atkinson (2016). The affluence measure is defined by

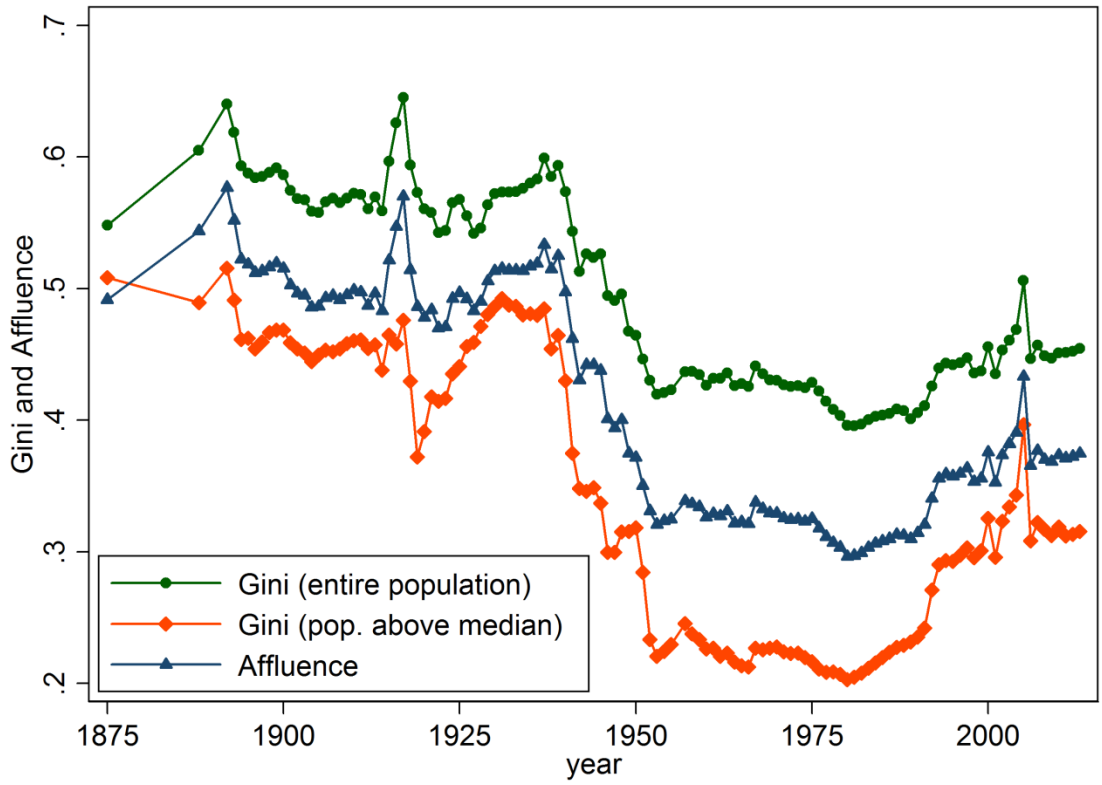
$$A = \frac{1}{3}\left(\frac{\mu_U}{\mu} G_U + \frac{\mu_U - \mu}{\mu} \right), \qquad (8a)$$

where μ is the overall mean income and μ_U and G_U are, respectively, the mean and the Gini coefficient of the conditional distribution of income given that the income is larger than the median. Inserting the well-known expressions for μ_U and G_U in (8a) yields the following alternative expression for *A*,

$$A = \frac{4}{3}\int_{\frac{1}{2}}^{1} (2t-1)\left(\frac{F^{-1}(t)}{\mu} - 1 \right) dt \ , \qquad (8b)$$

where $F^{-1}(t)$ is the income of the individual with rank t in the distribution of income F. Expression (8b) shows that *A* can be interpreted as a weighted average of top income shares, where the weight increases from 0 to 4/3 with increasing rank. The affluence measure, *A*, has itself 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 individual, then *A* takes the value 1. Note that 3*A* becomes equal to the richness gap (the second term within the parenthesis of (8a)) if individuals with higher income than the median income receive the same income μ_U.

Figure 7: Gini for entire population, Gini above median and affluence measure. Mean of upper and lower bound calculations



Sources: See text.

The estimation results for the affluence measure *A* and the upper tail (above median) Gini coefficient G_U are displayed in Figure 7. It is reassuring that the affluence measure and the upper tail Gini reproduce the time pattern of the overall Gini; not least since the available data provide more reliable estimates for the affluence measure and the upper tail Gini before the Second World War. However, note that the reliability of the affluence and the upper tail Gini partly carries over to the estimated overall Gini series. This is due to the fact that income distributions normally are skewed to the right, which means that the upper tail Gini contributes to a significantly larger proportion 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 the affluence measure *A*. So that, in 1900, with *G* = 0.586 and *A* = 0.515 (see Table A5), the contribution from the affluence term to the overall G was 66 per cent, while the affluence contribution had declined only marginally to 62 per cent one hundred years later.

Figure 7 shows the difference between the overall Gini and that for the upper half of the distribution. Although the pattern of the overall Gini broadly speaking is reflected by the upper tail Gini, the difference in magnitude has varied significantly over time. The largest difference is found after the Second World War, from late 1940s to late 1980s/early 1990s, when the gap was some 20 percentage points, approaching half of the total Gini. This was a period characterized by strict regulations and high marginal tax rates. The turning point came in the 1980s, when the difference between the overall Gini and the upper Gini fell to around 15 percentage points. This was accentuated by the banking crises and recession, the implementation of a major tax reform, a sharp increase in the dividends and capital gains among the richest people and a subsequent increase in the top income shares. By contrast, the rise in overall inequality during the 2000s was mainly driven by the sharp rise in upper tail inequality during this period, when the gap did not change. [10]

Another way of presenting the results is in terms of the affluence measure, shown in Figure 7. Since the sum of the upper (above median) mean and the lower (below median) mean is equal to twice the overall mean it follows from expression (8) that the affluence measure *A* is fully determined by the inequality in the distribution of income above the median income and by the relative gap between the upper and lower means. So that an affluence score of 0.296 in 1980 is generated by a Gini for the upper group equal to 20.3 per cent and a relative income level of 1.57. By 1995, the Gini for the upper group had risen to 29.3 per cent, but he relative upper mean had risen only to 1.6 (see Figure A 1), so that the rise in the affluence score to 0.357 was less dramatic than the change in the Gini. Thus, the increase in income inequality among the richest 50 per cent proportion of the population had a stronger effect on the change in overall inequality than the reduction in the (small) rise in the difference in average income between the 50 per cent richest and 50 per cent poorest proportions.

A series of episodes and their causes

The evolution of inequality in Norway is, we believe,best characterized 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 the upper group Gini and the Gini-based affluence measure. In the same way, the relative income of the upper group moves typically in the same direction.

The contribution in terms of magnitudes does however differ. Taken together, the periods from 1875 to 1939 show unchanged affluence, whereas the upper tail inequality decreased by 5 percentage points.

[10] See Aaberge and Atkinson (2010) for a further discussion of the development of top income shares in Norway.

The different evolution of upper tail and overall inequality (and affluence) corresponds to a significant rise in the ratio between mean incomes of the upper and lower half of the population (see Figure A 1). Since the average relative income gap between the upper and lower half stayed fairly flat during the last 60 years the rise in overall inequality and affluence after 1980 was largely due to rising upper tail inequality, the rich became richer as is also confirmed by the rising top income shares in this period.

Table 1: Overview of the evolution of overall inequality, upper tail inequality and affluence (Changes in percentage points in parentheses)

Period	Overall Gini coefficient	Gini-based measure of affluence	Upper tail Gini coefficient	Mean income of upper tail relative to overall mean
1875 – 1892	Increase (+7)	Increase (+8)	Slight increase (+1)	Increase
1892 - 1914	Decrease (-8)	Decrease (-10)	Decrease (-8)	Decrease
1914 – 1917	Increase (+9)	Increase (+7)	Increase (+4)	Increase
1917 - 1923	Decrease (-11)	Decrease (-10)	Decrease (-6)	Decrease
1923 - 1939	Increase (+5)	Increase (+5)	Increase (+4)	Increase
1939 – 1953	Decrease (-17)	Decrease (-20)	Decrease (-14)	Decrease
1953 – 1980	Slight decrease (-2)	Slight decrease (-2)	Slight decrease (-2)	Slight decrease
1980 - 2013	Increase (+5)	Increase (+5)	Increase (+12)	Slight increase

To get some intution about the magnitude of these changes, note that the 15 percentage points fall in overall inequality from 1875 to 1953 corresponds to a decrease in the Gini coefficient of 25 per cent. This is equal to the redistributive effect of the following hypothetical tax/transfer intervention (see Aaberge, 1997) in 1875: introduce a flat tax with tax rate equal to 25 per cent and allocate the collected tax as fixed lump-sum equal to the average tax 140 NOK. Then the 50 per cent poorest on average increase their income from 202 to 290 NOK, whereas the 50 per cent richest get their mean income reduced from 920 to 830 NOK. Similarly, this hypothetical intervention would change the income of the poor from 85 to 204 NOK and those at the 95th percentile from 1463 to 1237 NOK.

Turning to the macro-economic influences, what can we say about the extent to which the pattern of evolution of overall inequality is consistent with the evolution of the Norwegian economy? For the United States, as shown in Section 6, overall inequality rose during the “roaring 1920s”, followed by a collapse after 1929. On the other hand, it has been argued that the Nordic experience of the Great Depression was different: “the crisis was milder and shorter than in most other Western economies at the time, i.e. GDP growth rate and prices fell less and recovery was faster. However, despite the relatively rapid recovery in production, unemployment remained persistently high throughout the decade [the 1930s]” (Grytten, 2008a, page 370). The link between the macro-economy and income distribution has been a recurring theme. Aukrust (1957) sought to understand the difference between

the inter-war period and the immediate post-war period in terms of the relation with factor shares, where there had been an upward shift in the wage share (1957, Chart A). Aaberge et al (2000) investigated the impact of unemployment shocks on income distribution in the Nordic countries during the economic crises of the late 1980s and early 1990s. They concluded that the relation with unemployment is complex and may operate in the long-term rather than immediately. In a comparative study of OECD countries over the period 1970-1996, Checchi and Peñalosa (2010) examine the variation of the Gini coefficient, first with the labour share, and then with the simultaneous determinants of the share, the decile ratio for earnings and the unemployment rate.

The framework proposed by Checchi and Peñalosa provides a convenient organizing device, not least because it is based on an underlying theoretical model that it gives a role to labour market institutions that are relevant to the statistically significant coefficients (at 1 per cent level) in (Checchi and Peñalosa, 2010, Table 5). The omission of one variable in particular should be noted, in view of its prominence in the public debate: the unemployment benefit rate. Checchi and Peñalosa "find no evidence of a robust effect of the unemployment benefit on inequality" (2010, page 433). But there is a major obstacle to applying such a framework to the long-run series on inequality presented here: that we lack data for earlier years on key variables, such as unemployment.

Table 2: Changes in demographic and macroeconomic indicators in Norway 1875 to present

Period	Demographic changes. Proportion 20 years and over. Per cent	Employment by industry. Per cent			GDP by industries. Per cent			Average annual GDP per capita growth. Per cent
		Prim. Ind.	Second. Ind.	Service sector	Prim. Ind.	Second. Ind.	Service sector	
1875-1892	55 Overseas emigration in this period was 261 000.	50	21	29	33	25	42	0.7
1892 – 1914	55 Overseas emigration in this period was 312 000.	45	24	31	26	28	46	1.6
1914 – 1923	57 Total emigration 60 000.	43	24	33	23	29	48	1.1 (large year-to-year changes)
1923 – 1939	62 Total emigration 81 000.	41	25	34	17	33	50	3.1
1939 – 1953	69 Sharp decrease in overseas emigration.	32	29	39	15	35	50	2.3
1953-1980	68	17	35	48	10	35	55	3.6
1980 – 2013	73	5	27	68	5	40	55	1.8

Sources: Calculated from official population statistics (Column 1); Hansen and Skoglund (2008, 2009) (Columns 3-7); historical GDP statistics (Column 8).

Table 2 highlights changes in some basic demographic and economic conditions in Norway between 1875 and 2013 for which data are avilable. By comparing Tables 1 and 2 we find no clear relationship between the evolution of income inequality and changes in either the proportion of adults in the population, the emigration rate or the economic growth. For example, emigration was high both between 1875 and 1892 and between 1892 and1914, while inequality rose in the first period and declined in the last period. High economic growth was associated with decreasing inequality in the 1892-1914 period, but with increasing inequality in the 1923-1939 period. Between 1953 and 1980, when GDP per capita grew by an average of 3.6 per cent annually, the Gini coefficient stayed fairly stable.

Although we do not find a general relationship between economic growth and changes in inequality, we find strong association in volatility between these two indicators during the turbulent years of the First World War and the early 1920s. In this period, there were large fluctuations both in year-to-year GDP growth and in income inequality.

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 per cent of the work force was still in agriculture the 1930s, and rural-urban migration (and hence income equalization) was again held back by high unemployment. The Second World War is likely to have had an equalizing effect in itself, with more controls imposed on the economy and the German occupation leading to increased labor demand for construction projects.

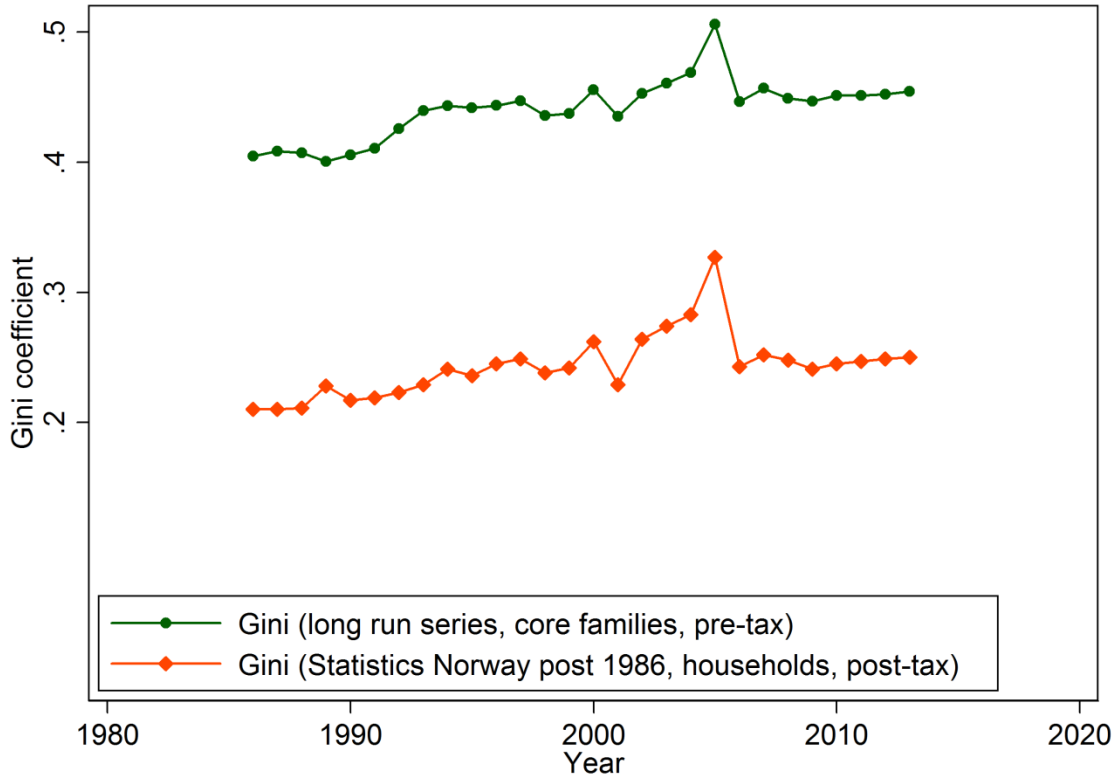
The evolution of inequality may also reflect quite specific events. The Gini coefficients for the overall distribution and for the upper half fell by 3 and 3.5 percentage points from 1945 to 1946 largely due to a nonrecurring tax on wealth increases that took place during the German occupation of Norway. The tax had highly progressive rates. The first 5000 NOK in wealth increase during the period 1940-1945 were tax-free, while the next 10,000 NOK were taxed at 30 percent. Above this level the tax rate rose stepwise up to 95 percent which was assessed for wealth increases beyond 70 000 NOK (the mean pre-tax income in 1945 was 3800 NOK). As a result the gap in the average income between the upper

half and the lower half of the income distribution and the inequality in the upper half incomes declined significantly during this period.

6 Different income definitions

This paper describes the evolution of inequality in Norway using the same population and definitions of household and income for the entire period 1875-2013. As indicated in the introduction our choice of definitions is dictated by constraints in available 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.

Figure 8: Gini coefficient for Norway, two definitions



Sources: See text.

The closest Norway gets to an official definition of income inequality is Statistics Norway's time series from 1986 onwards.[11] 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

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

equivalence scale is used. Second, a somewhat larger set of income sources (various types of non-taxable transfers) is included compared to the "gross income" concept used in this paper. Third, the income basis is post-tax rather than pre-tax.

In Figure 8, we compare the evolution of the Gini coefficient since 1986 for two alternative definitions of income. As expected the level of 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. Accounting for scale economics 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.

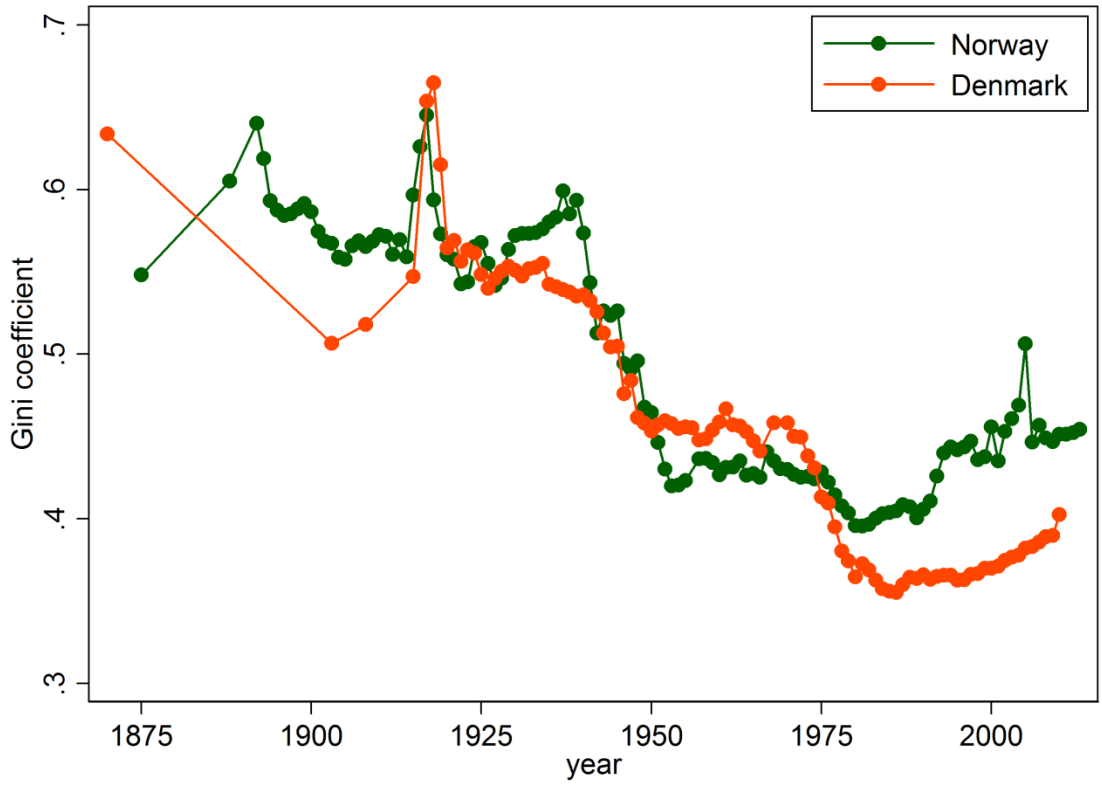
7 Comparison with Denmark and the United States

It is interesting to compare the estimates for Norway with the Gini coefficients for the same period for Denmark, one of the few countries for which such a series can be constructed back to the 1870s using income tax records[12], and with the shorter Gini series for the United States that, by linking, can be taken back to 1918. In Figure 9, the series for Norway are compared with the Gini coefficient based on the research of Atkinson and Søgaard (2016) for Denmark. The “continuous” Gini series for Denmark takes the average of the upper and lower bounds of the series given in that paper, (b) adjusts the estimates for 1970 to 1980 to a pre-1970 basis, by attributing all of the change between 1968 and 1970 to the switch to an individual basis, subtracting 7.9 percentage points, (c) linking to the micro-data series from 1980 onwards, and (d) putting the figures from 1994 onwards on to a pre-1994 basis by adding 4.26 percentage points (the difference between 1994 and 1993).[13] It is a Gini coefficient for family incomes and relates to gross income including transfers but before deduction of tax.

[12] Denmark had only one year with tabulated data for the 19th century (in 1870), but 26 observations for the first part of the 20th century. By contrast, Norway had 10 years with tabulated data for the 19th century and 7 observations with tabulated data between 1900 and 1938. Moreover, for Norway aggregate data on tax payers and assisted poor have been used for the entire period 1892 – 1951.

[13] This set of assumptions has been made by Atkinson; Jakob Søgaard should not be held in any way responsible.

Figure 9: Gini coefficients for Norway and Denmark compared



Sources: Figure 8 and calculations from Atkinson and Søgaard (2016)

The similarity in the movements in the series in Norway and Denmark up to the mid-1970s is remarkable. Over the period 1903 to the mid-1920s they move virtually in tandem. There was the rise in the early part of the twentieth century, followed by a steep increase in the First World War. The increase during the war is striking. While the rise in one country might be explained away as a statistical anomaly, the fact that it is found independently in both countries shows that something dramatic indeed happened during the First World War.

The fall in overall income inequality from before the Second World War to 1970 was large in both Denmark and Norway, but later there was a significant departure, as inequality continued to fall in Denmark until the mid-1980s. Both countries have seen a rise in recent years, but the differing experience during the 1970s leaves Denmark with a lower overall level at the end of the period. (It should be re-emphasized that these figures relate to gross incomes.)

Comparison with the United States

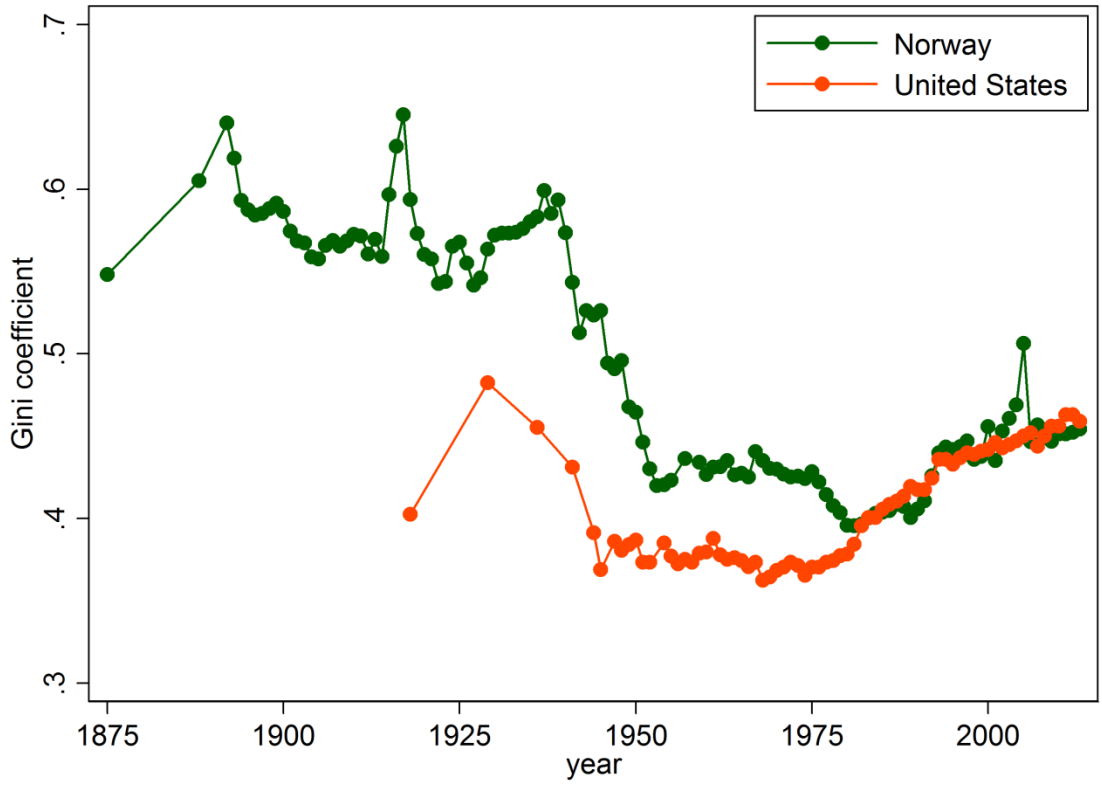
The comparison with the United States (US) is less easily made. Not only is the time period shorter by some 40 years, but the length of coverage in the US is only achieved by linking together different studies with different definitions. The Gini in Figure 10 for the US is that given by Atkinson and Morelli (2016). It is based initially on the Gini for the gross income of income recipients based on the NBER/Brookings synthetic estimates, calculated from the tabulations in Mitchell et al (1921, Table 25) and Leven, Moulton and Warburton (1934, Tables 27 and 29, excluding capital gains), linked to the BEA synthetic series for gross family incomes from Brandolini (2002, Table A1), who calculated the Gini coefficients from the original tabulations, which in turn is linked to the series from 1944 given by Budd (1970, Table 6), and linked at 1967 to the Gini coefficient for gross equivalised household income from the U.S. Bureau of the Census, *Income, Poverty, and Health Insurance Coverage in the United States: 2012*, (Table A-3, Selected measures of equivalence-adjusted income dispersion), where we have assumed that half of the recorded change between 1992 and 1993 was due to the change in methods (and therefore added 1.15 percentage points to the values from 1992 back to 1967.

The second issue concerns the definitions of income and income unit. The most recent US data relate to household income and are equivalised. As noted in the previous section, use of household income reduces measured inequality. On the other hand, the US data relate to gross income. The US income definition applicable in the most recent period is described as follows: “money income received (exclusive of certain money receipts such as capital gains) before payments for personal income taxes, social security, union dues, Medicare deductions, etc. Therefore, money income does not reflect the fact that some families receive noncash benefits, such as Supplemental Nutrition Assistance/food stamps, health benefits, subsidized housing, and goods produced and consumed on the farm (US Bureau of the Census, 2015, page 21).

The two series in Figure 10 should be viewed with these qualifications in mind. It does not follow that overall income inequality in 1918 was less in the US than in Norway. The relative changes over time may also not be fully comparable. Nevertheless, they suggest that there are interesting commonalities and interesting differences. In both countries there was a decline in income inequality in the Second World War, as there was in Denmark. The rise in inequality from 1980 to the present is not much smaller in Norway than in the US, which may come as a surprise to some readers. The Gini coefficient in Norway increased by 5.8 percentage points between 1980 and 2013; that for the US increased by 8.0 percentage points. But there are also contrasts. After the Great Crash, overall inequality appears to

have fallen in the US, but there was, if anything, a rise in Norway. The difference may reflect the US policies pursued in the New Deal, but this requires closer investigation. More recently, inequality appears to have been rising over the 1970s in the US, but falling in Norway.

Figure 10: Gini coefficients in Norway and US compared



Sources: See text

8 Summary of findings

This paper may be seen as an exercise in constructing bricks with a minimum of straw. But, while stressing the limitations of our materials and methods, we should begin this summary by saying that there is some straw. For every year in our series, there *is* some underlying information on individual incomes, even if of a highly aggregated form. We have not interpolated observations between years or extrapolated from comparisons with other countries. The position is perhaps best described in terms of a continuum with national income totals at one end and individual micro-data at the other end. For part of the earlier period we are close to, but not at, the national income end, using a small number of aggregated variables (such as the numbers of taxpayers). Over time, we move closer to the individual data that we ideally use.

Our findings – with attached qualifications - suggest that at the end of the nineteenth century, the Gini coefficient for gross family income in Norway was between 50-60 per cent. 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).

Today overall inequality of gross family incomes in Norway is lower than it was a hundred years ago. When did it change? It appears that overall gross income inequality among families in Norway

- Fell from 1892 to 1914, largely due to fall in inequality among the upper half of the income distribution;
- Reached an upward spike during the First World War (as in Denmark), but fell during the Second World War due to a decrease in upper income inequality as well as decline in the gap of the mean income between the upper and lower half of the population;
- Rose between 1923 and 1939, largely due to increase in inequality among the upper half of the population;
- Fell substantially between 1948 and 1953 as a result of decline in both upper tail inequality and the gap between upper and lower tail means;
- Stayed broadly flat between 1953 and 1980;
- Has risen again since 1980; largely due to increased upper tail inequality.

Expressed this way, the history of Norwegian income inequality is better seen as a series of episodes than as the working-out of some long-run pattern. It cannot be summarized as an inverse U or a U. In this respect, our conclusions for Norway run on similar lines to those of Lindert and Williamson for the United States over the period since 1700, where they say that "inequality movements are driven not by any fundamental law of capitalist development but instead by episodic shifts in six basic forces: politics, demography, education policy, trade competition, finance, and labor- saving technological change" (2016, page 12).

References

Aaberge, R, 1997, "Interpretation of Changes in Rank-Dependent Measures of Inequality", *Economics Letters*, 55, 215-219.

Aaberge, R and A B, Atkinson, 2010, "Top incomes in Norway" in A B Atkinson and T Piketty, editors, *Top incomes in a global perspective*, Oxford University Press, Oxford.

Aaberge, R and A B, Atkinson, 2016, "Deconstruction of inequality", Mimeo, Research Department, Statistics Norway.

Aaberge, R, Björklund, A, Jäntti, M, Pedersen, P J, Smith, N and T, Wennemo, 2000, "Unemployment shocks and income distribution: How did the Nordic countries fare during their crises?", *Scandinavian Journal of Economics*, **102**, 77-97.

Atkinson, A B, 2015, *Inequality: What can be done?*, Harvard University Press, London.

Atkinson, A B and Søgaard, J E, 2016, "The long run history of income inequality in Denmark", *Scandinavian Journal of Economics*, **118**, 264–291.

Atkinson, A B and Morelli, S, 2014: Chartbook of economic inequality. ECINEQ WP 2014-324.

Aukrust, O, 1957, "Trends and cycles in Norwegian income shares", *Review of Income and Wealth*, **6**, 283-305.

Belabed, C A, 2016, "Inequality and the New Deal", IMK Working Paper 166, Hans Boeckler Foundation in Düsseldorf.

Brandolini, A, 2002, "A bird's eye view of long-run changes in income inequality", Bank of Italy Research Department, Rome.

Budd, E C, 1970, "Postwar changes in the size distribution of income in the U.S.", American Economic Review, Papers and Proceedings, **60**, 247-260.

Checchi, D and Garcia-Peñalosa, C, 2010, "Labour market institutions and the personal distribution of income in the OECD", *Economica*, **77**: 413-450.

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

Grytten, O, 2008a, "Why was the Great Depression not so great in the Nordic Countries?" , *Journal of Economic History*, **37**: 369-393.

Grytten, O, 2008b, "The Economic History of Norway", EH.Net Encyclopedia, edited by Robert Whaples. March 16, 2008. URL http://eh.net/encyclopedia/the-economic-history-of-norway/

Hansen, S and T Skoglund, 2008: Syssselsetting og lønn i historisk nasjonalregnskap. Beregninger for 1930 – 1969. Notater 54/2008, Statistics Norway.

Hansen,S. and T. Skoglund, 2009: Syssselsetting og lønn i historisk nasjonalregnskap. Beregninger for 1900 – 1930. Notater 38/2009, Statistics Norway.

Kiær, A N, 1892-3, Indtægtsforhold i Norge, *Tillæg Statsøkonomisk tidsskrift*. Aschehoug, Kristiania
Kopczuk W, Saez E and Song J, 2010: Earnings inequality and mobility in the United States: Evidence from social security data since 1937. *Quarterly Journal of Economics*, **125**, 91-128.

Kuznets, S, 1955, "Economic growth and income inequality", *American Economic Review*, **45,** 1-28.

Leven, M, Moulton, H G and Warburton, C, 1934, America's capacity to consume, Brookings Institution, Washington, D.C.

Lindert, P H, 2000, Three centuries of inequality in Britain and America. In: A B Atkinson and F Bourguignon, editors, *Handbook of Income Distribution*, volume 1, Elsevier, Amsterdam.

Lindert, P H and 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.

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

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

US Bureau of the Census, 2015, *Income, Poverty, and Health Insurance Coverage in the United States: 2012*, US Government Printing Office, Washington, DC

Williamson, J G, 1985, *Did British capitalism breed inequality?,* Allen and Unwin, London.

Williamson, J G and Lindert, P H, 1980, *American inequality: A macroeconomic history*, Academic Press, New York.

Appendix A: Gini and other key variables for Norway 1875-2013

Below are shown three tables:

- Overall Gini coefficient and other measures of income dispersion
- Income and population totals for the four groups
- Parameters used for the calculation of the "four-group" Gini coefficient

The information is also available as an Online Appendix in Excel and format (on request).

Moreover, plots of Lorenz curves for all years are available as a PDF file.

Table A 1: Gini coefficient and affluence measures, 1875-2013 , Upper and lower bound measures

	Upper bound assumptions				Lower bound assumptions			
Year	Gini coefficient	$\frac{\mu_U}{\mu}$	Gini (above median)	Affluence	Gini coefficient	$\frac{\mu_U}{\mu}$	Gini (above median)	Affluence
1875	0.556	1.651	0.508	0.496	0.540	1.651	0.508	0.496
1888	0.641	1.828	0.489	0.574	0.569	1.763	0.489	0.542
1892	0.643	1.807	0.515	0.580	0.637	1.796	0.515	0.574
1893	0.631	1.801	0.491	0.562	0.607	1.761	0.491	0.542
1894	0.610	1.785	0.461	0.536	0.576	1.729	0.461	0.509
1895	0.602	1.770	0.462	0.530	0.573	1.724	0.462	0.507
1896	0.600	1.770	0.454	0.525	0.568	1.718	0.454	0.500
1897	0.601	1.765	0.460	0.525	0.570	1.716	0.460	0.501
1898	0.602	1.760	0.467	0.527	0.574	1.716	0.467	0.506
1899	0.606	1.764	0.469	0.530	0.577	1.720	0.469	0.508
1900	0.597	1.751	0.469	0.524	0.575	1.716	0.469	0.507
1901	0.585	1.736	0.459	0.511	0.564	1.703	0.459	0.495
1902	0.579	1.728	0.454	0.504	0.558	1.696	0.454	0.489
1903	0.578	1.729	0.451	0.503	0.557	1.696	0.451	0.487
1904	0.569	1.717	0.445	0.494	0.549	1.686	0.444	0.478
1905	0.565	1.707	0.450	0.492	0.550	1.686	0.449	0.481
1906	0.573	1.717	0.453	0.498	0.559	1.694	0.453	0.487
1907	0.579	1.721	0.458	0.504	0.559	1.699	0.446	0.486
1908	0.574	1.710	0.461	0.500	0.556	1.693	0.447	0.483
1909	0.578	1.714	0.466	0.504	0.559	1.695	0.451	0.486
1910	0.580	1.716	0.469	0.507	0.564	1.703	0.452	0.491
1911	0.580	1.713	0.470	0.506	0.563	1.698	0.452	0.489
1912	0.572	1.698	0.471	0.499	0.549	1.686	0.439	0.475
1913	0.572	1.712	0.457	0.498	0.567	1.704	0.457	0.495
1914	0.568	1.708	0.450	0.493	0.550	1.699	0.426	0.474
1915	0.606	1.754	0.480	0.532	0.588	1.748	0.449	0.511
1916	0.637	1.814	0.479	0.561	0.614	1.809	0.437	0.533
1917	0.655	1.836	0.498	0.583	0.635	1.836	0.454	0.557
1918	0.604	1.779	0.452	0.528	0.584	1.779	0.407	0.501
1919	0.586	1.792	0.401	0.504	0.560	1.792	0.343	0.469
1920	0.572	1.749	0.417	0.493	0.549	1.749	0.366	0.463
1921	0.567	1.729	0.439	0.496	0.549	1.729	0.397	0.472
1922	0.551	1.704	0.433	0.481	0.535	1.704	0.395	0.459
1923	0.552	1.703	0.435	0.481	0.536	1.703	0.399	0.461
1924	0.573	1.726	0.453	0.503	0.558	1.726	0.417	0.482
1925	0.576	1.729	0.459	0.507	0.560	1.729	0.423	0.487
1926	0.561	1.701	0.471	0.501	0.549	1.701	0.441	0.484
1927	0.547	1.679	0.472	0.490	0.536	1.679	0.446	0.476

1928	0.551	1.679	0.484	0.497	0.541	1.679	0.459	0.483
1929	0.569	1.700	0.493	0.513	0.558	1.700	0.468	0.499
1930	0.577	1.709	0.499	0.521	0.567	1.709	0.474	0.506
1931	0.578	1.706	0.503	0.521	0.569	1.706	0.482	0.509
1932	0.578	1.709	0.498	0.520	0.569	1.709	0.477	0.508
1933	0.578	1.710	0.497	0.520	0.569	1.710	0.476	0.508
1934	0.581	1.717	0.491	0.520	0.571	1.717	0.469	0.507
1935	0.585	1.722	0.492	0.523	0.575	1.722	0.469	0.510
1936	0.589	1.728	0.493	0.526	0.578	1.728	0.467	0.512
1937	0.606	1.751	0.500	0.542	0.593	1.751	0.470	0.525
1938	0.585	1.750	0.454	0.515	0.585	1.750	0.454	0.515
1939	0.597	1.758	0.472	0.530	0.590	1.758	0.457	0.520
1940	0.577	1.743	0.438	0.502	0.570	1.743	0.422	0.493
1941	0.548	1.735	0.386	0.469	0.538	1.735	0.364	0.455
1942	0.518	1.700	0.360	0.437	0.508	1.700	0.336	0.424
1943	0.532	1.728	0.359	0.450	0.521	1.728	0.333	0.435
1944	0.529	1.725	0.362	0.449	0.518	1.725	0.336	0.435
1945	0.532	1.730	0.350	0.445	0.520	1.730	0.324	0.430
1946	0.502	1.697	0.315	0.410	0.487	1.694	0.284	0.392
1947	0.498	1.682	0.315	0.404	0.483	1.676	0.284	0.384
1948	0.496	1.674	0.315	0.400	0.496	1.674	0.315	0.400
1949	0.468	1.615	0.315	0.375	0.468	1.615	0.315	0.375
1950	0.464	1.604	0.318	0.372	0.464	1.604	0.318	0.372
1951	0.446	1.597	0.284	0.350	0.446	1.597	0.284	0.350
1952	0.440	1.628	0.233	0.336	0.421	1.604	0.233	0.326
1953	0.429	1.619	0.221	0.326	0.411	1.596	0.221	0.316
1954	0.431	1.624	0.224	0.329	0.410	1.595	0.224	0.318
1955	0.433	1.619	0.230	0.330	0.413	1.592	0.230	0.319
1957	0.446	1.632	0.245	0.344	0.427	1.606	0.245	0.333
1958	0.446	1.636	0.238	0.342	0.427	1.611	0.238	0.331
1959	0.445	1.638	0.233	0.340	0.423	1.609	0.233	0.328
1960	0.438	1.628	0.226	0.332	0.416	1.599	0.226	0.320
1961	0.441	1.633	0.227	0.334	0.421	1.607	0.227	0.324
1962	0.440	1.635	0.221	0.332	0.423	1.612	0.221	0.322
1963	0.445	1.643	0.223	0.336	0.425	1.616	0.223	0.325
1964	0.436	1.628	0.216	0.327	0.417	1.603	0.216	0.317
1965	0.439	1.638	0.213	0.329	0.416	1.606	0.213	0.316
1966	0.437	1.636	0.212	0.328	0.414	1.604	0.212	0.315
1967	0.453	1.657	0.227	0.344	0.429	1.624	0.227	0.331
1968	0.447	1.646	0.226	0.339	0.424	1.615	0.226	0.326
1969	0.442	1.636	0.227	0.336	0.419	1.606	0.227	0.323
1970	0.441	1.635	0.228	0.336	0.418	1.603	0.228	0.323
1971	0.439	1.631	0.224	0.332	0.415	1.599	0.224	0.319

1972	0.437	1.629	0.223	0.331	0.414	1.598	0.223	0.318
1973	0.437	1.629	0.223	0.331	0.414	1.599	0.223	0.319
1974	0.436	1.630	0.220	0.329	0.413	1.599	0.220	0.317
1975	0.440	1.639	0.217	0.331	0.417	1.608	0.217	0.319
1976	0.433	1.628	0.211	0.324	0.412	1.599	0.211	0.312
1977	0.425	1.615	0.209	0.317	0.404	1.586	0.209	0.306
1978	0.419	1.604	0.209	0.313	0.397	1.575	0.209	0.301
1979	0.414	1.596	0.207	0.309	0.393	1.568	0.207	0.298
1980	0.406	1.584	0.203	0.302	0.386	1.558	0.203	0.291
1981	0.406	1.584	0.205	0.303	0.385	1.556	0.205	0.291
1982	0.408	1.585	0.208	0.305	0.386	1.556	0.208	0.293
1983	0.413	1.592	0.212	0.310	0.388	1.559	0.212	0.296
1984	0.415	1.595	0.216	0.313	0.390	1.562	0.216	0.300
1985	0.416	1.594	0.220	0.315	0.392	1.562	0.220	0.302
1986	0.417	1.593	0.224	0.316	0.393	1.561	0.224	0.303
1987	0.419	1.595	0.227	0.319	0.398	1.566	0.227	0.307
1988	0.418	1.591	0.229	0.318	0.397	1.563	0.229	0.307
1989	0.410	1.578	0.232	0.314	0.392	1.556	0.232	0.305
1990	0.415	1.584	0.235	0.319	0.397	1.562	0.235	0.310
1991	0.420	1.592	0.242	0.326	0.401	1.567	0.242	0.316
1992	0.437	1.605	0.271	0.346	0.415	1.577	0.271	0.335
1993	0.450	1.615	0.290	0.361	0.430	1.589	0.290	0.350
1994	0.453	1.618	0.294	0.364	0.434	1.594	0.294	0.354
1995	0.450	1.614	0.293	0.362	0.433	1.592	0.293	0.353
1996	0.452	1.612	0.297	0.364	0.435	1.592	0.297	0.355
1997	0.455	1.613	0.303	0.368	0.440	1.595	0.303	0.359
1998	0.444	1.600	0.296	0.358	0.428	1.580	0.296	0.349
1999	0.446	1.600	0.301	0.360	0.429	1.579	0.301	0.351
2000	0.464	1.615	0.325	0.380	0.447	1.594	0.325	0.371
2001	0.444	1.600	0.296	0.358	0.426	1.577	0.296	0.348
2002	0.462	1.614	0.323	0.378	0.444	1.591	0.323	0.368
2003	0.470	1.621	0.334	0.387	0.451	1.597	0.334	0.377
2004	0.478	1.629	0.343	0.396	0.460	1.605	0.343	0.385
2005	0.516	1.658	0.397	0.439	0.497	1.635	0.397	0.428
2006	0.457	1.615	0.308	0.371	0.437	1.590	0.308	0.360
2007	0.467	1.624	0.322	0.383	0.447	1.599	0.322	0.371
2008	0.460	1.616	0.317	0.376	0.438	1.588	0.317	0.364
2009	0.460	1.621	0.312	0.376	0.434	1.588	0.312	0.361
2010	0.465	1.626	0.319	0.381	0.438	1.591	0.319	0.366
2011	0.463	1.626	0.312	0.378	0.439	1.595	0.312	0.364
2012	0.465	1.628	0.313	0.379	0.440	1.596	0.313	0.365
2013	0.467	1.631	0.316	0.382	0.442	1.598	0.316	0.368

Table A 2: Number of individuals and mean incomes, by group, and control totals. Nominal values in NOK (not CPI adjusted)

			Number of individuals				Mean income				
									NA/NT		
Year	Population control	Income control (Mill NOK)	CG[14] taxpayers	MUN-CG taxpayers	NA/NT	Poor	CG taxpayers	MUN-CG taxpayers	(Upper Gini bound)	(Lower Gini bound)	Poor
1875	847 000	476	184 053	521 407	83 730	57 810	1 225	230	85	150	85
1888	919 000	442	65 845	410 098	365 502	77 555	2 537	345	76	150	76
1892	937 870	490	176 075	342 860	340 254	78 681	1 450	277	85	92	85
1893	944 840	500	102 542	421 050	340 318	80 930	2 039	338	83	113	83
1894	955 117	503	66 807	468 322	339 239	80 749	2 709	383	85	128	85
1895	968 945	514	68 227	476 723	345 533	78 462	2 712	387	93	129	93
1896	983 818	538	70 454	492 729	340 393	80 242	2 696	397	90	132	90
1897	999 315	559	75 578	505 845	337 693	80 199	2 700	401	91	134	91
1898	1 015 808	606	83 933	520 295	331 403	80 177	2 698	408	94	136	94
1899	1 031 501	639	91 422	521 528	337 821	80 730	2 669	419	95	140	95
1900	1 045 420	667	94 367	531 711	341 090	78 252	2 683	425	106	142	106
1901	1 058 452	657	95 767	548 176	332 368	82 141	2 625	427	108	142	108
1902	1 066 877	652	97 517	556 891	329 077	83 392	2 588	428	108	143	108
1903	1 071 397	648	96 431	557 972	330 361	86 634	2 574	428	107	143	107
1904	1 065 571	638	100 380	560 810	316 199	88 182	2 464	428	109	143	109
1905	1 070 722	654	99 463	569 071	318 628	83 560	2 537	426	118	142	118
1906	1 077 000	708	105 145	573 059	315 190	83 606	2 535	430	116	143	116
1907	1 088 673	749	113 288	577 144	315 810	82 431	2 519	433	116	144	116
1908	1 084 270	770	124 027	583 412	294 056	82 775	2 462	435	119	145	119
1909	1 102 688	784	130 278	589 821	298 630	83 959	2 472	437	117	146	117
1910	1 119 676	866	140 864	591 294	307 841	79 677	2 459	445	127	148	127
1911	1 122 989	920	153 321	599 031	290 504	80 133	2 462	457	127	152	127
1912	1 138 014	1 017	237 787	570 380	252 307	77 540	2 292	501	139	167	139
1913	1 181 740	1 130	256 299	586 611	261 662	77 168	2 486	498	149	166	149
1914	1 198 991	1 165	277 668	589 343	254 034	77 947	2 491	517	147	172	147
1915	1 191 118	1 590	315 126	570 701	230 296	74 995	3 241	546	159	182	159
1916	1 213 725	2 344	403 017	516 477	220 729	73 502	4 842	624	178	208	178
1917	1 234 220	2 785	414 844	561 408	188 332	69 636	6 040	574	224	224	224
1918	1 257 369	3 196	448 653	549 757	192 108	66 851	6 001	875	298	298	298
1919	1 274 625	3 890	556 348	476 176	179 495	62 606	5 959	819	395	395	395
1920	1 297 828	4 702	512 180	528 326	201 062	56 260	6 439	1 120	494	494	494
1921	1 320 416	3 512	448 155	521 486	285 654	65 122	5 161	974	507	507	507
1922	1 341 487	3 170	424 732	542 246	301 006	73 503	4 506	957	492	492	492
1923	1 359 382	3 140	412 585	561 104	309 925	75 768	4 317	933	469	469	469

[14] The Mun-CG divison for 1875 and 1888 has been discussed in the subsection “Using data on the assisted poor”.

1924	1 366 009	3 468	407 816	580 535	296 892	80 766	4 543	868	462	462	462
1925	1 382 259	3 491	409 671	581 312	305 945	85 331	4 578	848	480	480	480
1926	1 401 352	2 869	360 762	618 748	324 945	96 897	4 444	821	482	482	482
1927	1 416 889	2 587	332 276	644 336	332 594	107 683	4 248	797	485	485	485
1928	1 429 250	2 583	323 486	661 957	337 679	106 127	4 180	738	492	492	492
1929	1 455 069	2 656	330 210	668 110	348 444	108 305	4 252	713	451	451	451
1930	1 462 006	2 701	328 673	680 495	342 738	110 100	4 279	693	434	434	434
1931	1 484 265	2 331	290 127	704 092	364 673	125 373	4 220	682	371	371	371
1932	1 500 824	2 324	297 978	700 192	357 232	145 423	3 989	663	359	359	359
1933	1 520 458	2 323	300 982	706 177	360 450	152 850	3 878	649	352	352	352
1934	1 543 222	2 450	315 183	704 531	366 504	157 003	3 820	662	347	347	347
1935	1 565 806	2 627	330 193	712 185	365 219	158 209	3 922	673	354	354	354
1936	1 583 790	2 919	365 267	723 020	339 260	156 242	4 053	679	371	371	371
1937	1 610 577	3 372	409 369	717 636	334 786	148 786	4 366	672	378	378	378
1938	1 632 718	3 497	444 099	700 914	347 219	140 486	4 424	716	387	387	387
1939	1 654 129	3 755	471 654	712 892	331 743	137 840	4 498	670	406	406	406
1940	1 674 238	4 019	517 468	730 297	275 149	151 324	4 388	771	377	377	377
1941	1 688 313	5 134	664 652	653 608	278 337	91 716	4 530	828	523	523	523
1942	1 695 121	5 137	711 786	637 713	281 683	63 939	4 505	965	576	576	576
1943	1 704 634	5 223	739 956	618 432	291 888	54 358	4 623	868	555	555	555
1944	1 716 464	5 198	739 897	612 073	315 106	49 388	4 619	857	613	613	613
1945	1 730 001	6 330	768 327	596 787	315 131	49 756	4 599	918	576	576	576
1946	1 746 103	6 303	917 116	442 106	336 076	50 805	4 908	1 041	662	662	662
1947	1 752 946	7 456	979 409	416 528	308 940	48 069	5 572	1 081	733	733	733
1948	1 736 464	8 209	1 006 112	401 233	284 421	44 698	6 220	1 114	795	795	795
1949	1 733 690	8 800	1 076 360	372 694	243 250	41 386	6 266	1 382	835	835	835
1950	1 727 813	9 463	1 125 158	351 160	212 229	39 266	6 621	1 338	845	845	845
1951	1 721 099	11 472	1 026 214	455 814	199 979	39 092	8 156	2 091	908	908	908

Table A 3: Population and income control totals 1951 and later

Year	Population control	Income control	Year	Population control	Income control	Year	Population control	Income control
1951	1 721 099	11 472	1973	1 970 938	64 933	1994	2 553 029	475 796
1952	1 723 350	12 556	1974	1 989 957	74 603	1995	2 571 878	500 651
1953	1 723 163	12 760	1975	2 009 594	86 418	1996	2 590 583	526 145
1954	1 723 981	14 055	1976	2 032 203	98 343	1997	2 608 585	558 102
1955	1 725 450	14 826	1977	2 052 434	108 545	1998	2 629 277	612 113
1957	1 741 998	17 685	1978	2 076 830	128 727	1999	2 652 168	649 219
1958	1 748 932	17 301	1979	2 100 389	137 116	2000	2 668 561	697 332
1959	1 758 814	18 382	1980	2 126 458	156 663	2001	2 683 319	731 486
1960	1 771 109	19 601	1981	2 158 775	178 977	2002	2 705 535	789 216
1961	1 788 908	21 349	1982	2 190 717	201 213	2003	2 726 116	828 107
1962	1 809 911	22 996	1983	2 222 341	221 096	2004	2 752 110	854 120
1963	1 833 869	24 916	1984	2 254 414	244 354	2005	2 786 213	918 359
1964	1 854 113	27 564	1985	2 293 666	268 342	2006	2 825 535	908 676
1965	1 872 800	30 590	1986	2 330 892	303 474	2007	2 879 690	1 005 373
1966	1 889 704	32 847	1987	2 367 549	343 704	2008	2 933 108	1 110 046
1967	1 904 805	35 865	1988	2 402 329	370 905	2009	2 979 837	1 138 789
1968	1 885 438	38 272	1989	2 425 794	387 500	2010	3 035 119	1 184 684
1969	1 900 571	41 543	1990	2 450 457	408 447	2011	2 944 064	1 260 426
1970	1 914 912	47 014	1991	2 480 929	428 316	2012	2 999 539	1 331 274
1971	1 934 029	52 095	1992	2 508 283	449 394	2013	3 051 514	1 402 008
1972	1 950 723	57 432	1993	2 533 015	466 137			

Table A 4: Parameters used for the calculation of four-class Gini. For calculation, see text. Note: G=0.1333 (from z=0.4) for all years**

	Same for upper and lower bound					Upper bound				Lower bound			
Year	p	n	m	c	A	h	g	g'	G*	h	g	g'	G*
1875	0.068	0.099	0.616	0.217	0.010	0.025	0.364	0.338	0.494	0.037	0.371	0.335	0.494
1888	0.084	0.398	0.446	0.072	0.013	0.076	0.499	0.424	0.461	0.137	0.533	0.395	0.461
1892	0.084	0.363	0.366	0.188	0.014	0.072	0.324	0.252	0.448	0.078	0.328	0.250	0.448
1893	0.086	0.360	0.446	0.109	0.013	0.070	0.447	0.377	0.452	0.090	0.459	0.369	0.452
1894	0.085	0.355	0.490	0.070	0.014	0.071	0.533	0.463	0.447	0.100	0.548	0.448	0.447
1895	0.081	0.357	0.492	0.070	0.014	0.077	0.538	0.461	0.485	0.101	0.550	0.449	0.485
1896	0.082	0.346	0.501	0.072	0.013	0.070	0.542	0.472	0.480	0.097	0.555	0.458	0.480
1897	0.080	0.338	0.506	0.076	0.013	0.068	0.532	0.465	0.488	0.094	0.545	0.452	0.488
1898	0.079	0.326	0.512	0.083	0.012	0.064	0.517	0.453	0.494	0.087	0.529	0.442	0.494
1899	0.078	0.328	0.506	0.089	0.012	0.062	0.505	0.443	0.492	0.086	0.518	0.432	0.492
1900	0.075	0.326	0.509	0.090	0.012	0.067	0.507	0.440	0.495	0.085	0.517	0.432	0.495
1901	0.078	0.314	0.518	0.090	0.013	0.068	0.517	0.449	0.481	0.085	0.526	0.441	0.481
1902	0.078	0.308	0.522	0.091	0.014	0.069	0.521	0.453	0.473	0.086	0.530	0.444	0.473
1903	0.081	0.308	0.521	0.090	0.014	0.069	0.525	0.456	0.469	0.087	0.534	0.447	0.469
1904	0.083	0.297	0.526	0.094	0.015	0.069	0.527	0.458	0.469	0.086	0.536	0.450	0.456
1905	0.078	0.298	0.531	0.093	0.015	0.073	0.527	0.455	0.469	0.084	0.533	0.449	0.456
1906	0.078	0.293	0.532	0.098	0.014	0.065	0.514	0.449	0.456	0.077	0.520	0.443	0.456
1907	0.076	0.290	0.530	0.104	0.013	0.062	0.500	0.438	0.456	0.074	0.506	0.432	0.353
1908	0.076	0.271	0.538	0.114	0.013	0.058	0.486	0.428	0.456	0.068	0.491	0.423	0.353
1909	0.076	0.271	0.535	0.118	0.013	0.057	0.476	0.419	0.456	0.068	0.482	0.414	0.353
1910	0.071	0.275	0.528	0.126	0.012	0.057	0.464	0.407	0.456	0.064	0.468	0.404	0.353
1911	0.071	0.259	0.533	0.137	0.011	0.051	0.450	0.399	0.456	0.059	0.455	0.395	0.353
1912	0.068	0.222	0.501	0.209	0.011	0.045	0.374	0.328	0.456	0.052	0.378	0.326	0.353
1913	0.065	0.221	0.496	0.217	0.010	0.045	0.345	0.300	0.353	0.049	0.348	0.299	0.353
1914	0.065	0.212	0.492	0.232	0.010	0.042	0.335	0.293	0.353	0.047	0.338	0.291	0.284
1915	0.063	0.193	0.479	0.265	0.008	0.031	0.257	0.226	0.353	0.034	0.260	0.226	0.284
1916	0.061	0.182	0.426	0.332	0.006	0.022	0.161	0.138	0.353	0.025	0.163	0.138	0.284
1917	0.056	0.153	0.455	0.336	0.006	0.021	0.132	0.112	0.353	0.021	0.132	0.112	0.284
1918	0.053	0.153	0.437	0.357	0.006	0.024	0.172	0.148	0.353	0.024	0.172	0.148	0.284
1919	0.049	0.141	0.374	0.436	0.006	0.025	0.127	0.103	0.353	0.025	0.127	0.103	0.284
1920	0.043	0.155	0.407	0.395	0.006	0.027	0.175	0.148	0.353	0.027	0.175	0.148	0.284
1921	0.049	0.216	0.395	0.339	0.009	0.051	0.222	0.171	0.353	0.051	0.222	0.171	0.284
1922	0.055	0.224	0.404	0.317	0.011	0.058	0.259	0.201	0.353	0.058	0.259	0.201	0.284
1923	0.056	0.228	0.413	0.304	0.011	0.058	0.272	0.214	0.353	0.058	0.272	0.214	0.284
1924	0.059	0.217	0.425	0.299	0.011	0.050	0.253	0.203	0.353	0.050	0.253	0.203	0.284
1925	0.062	0.221	0.421	0.296	0.012	0.054	0.251	0.197	0.353	0.054	0.251	0.197	0.284
1926	0.069	0.232	0.442	0.257	0.016	0.071	0.294	0.224	0.353	0.071	0.294	0.224	0.284
1927	0.076	0.235	0.455	0.235	0.020	0.083	0.327	0.245	0.353	0.083	0.327	0.245	0.284
1928	0.074	0.236	0.463	0.226	0.020	0.084	0.327	0.243	0.353	0.084	0.327	0.243	0.284

1929	0.074	0.239	0.459	0.227	0.018	0.078	0.311	0.234	0.353	0.078	0.311	0.234	0.284
1930	0.075	0.234	0.465	0.225	0.018	0.073	0.306	0.233	0.353	0.073	0.306	0.233	0.284
1931	0.084	0.246	0.474	0.195	0.020	0.078	0.338	0.260	0.353	0.078	0.338	0.260	0.284
1932	0.097	0.238	0.467	0.199	0.022	0.078	0.337	0.259	0.353	0.078	0.337	0.259	0.284
1933	0.101	0.237	0.464	0.198	0.023	0.078	0.338	0.260	0.353	0.078	0.338	0.260	0.284
1934	0.102	0.237	0.457	0.204	0.022	0.074	0.333	0.258	0.353	0.074	0.333	0.258	0.284
1935	0.101	0.233	0.455	0.211	0.021	0.071	0.322	0.251	0.353	0.071	0.322	0.251	0.284
1936	0.099	0.214	0.457	0.231	0.020	0.063	0.296	0.233	0.353	0.063	0.296	0.233	0.284
1937	0.092	0.208	0.446	0.254	0.017	0.054	0.255	0.201	0.353	0.054	0.255	0.201	0.284
1938	0.086	0.213	0.429	0.272	0.016	0.054	0.246	0.192	0.284	0.054	0.246	0.192	0.284
1939	0.083	0.201	0.431	0.285	0.015	0.051	0.225	0.174	0.315	0.051	0.225	0.174	0.284
1940	0.090	0.164	0.436	0.309	0.014	0.040	0.231	0.191	0.315	0.040	0.231	0.191	0.284
1941	0.054	0.165	0.387	0.394	0.009	0.038	0.184	0.147	0.315	0.038	0.184	0.147	0.284
1942	0.038	0.166	0.376	0.420	0.007	0.039	0.193	0.155	0.315	0.039	0.193	0.155	0.284
1943	0.032	0.171	0.363	0.434	0.006	0.037	0.167	0.131	0.315	0.037	0.167	0.131	0.284
1944	0.029	0.184	0.357	0.431	0.006	0.043	0.170	0.127	0.315	0.043	0.170	0.127	0.284
1945	0.029	0.182	0.345	0.444	0.005	0.033	0.163	0.130	0.315	0.033	0.163	0.130	0.284
1946	0.029	0.192	0.253	0.525	0.005	0.041	0.130	0.089	0.315	0.041	0.130	0.089	0.284
1947	0.027	0.176	0.238	0.559	0.005	0.035	0.109	0.074	0.315	0.035	0.109	0.074	0.284
1948	0.026	0.164	0.231	0.579	0.004	0.032	0.096	0.065	0.315	0.032	0.096	0.065	0.315
1949	0.024	0.140	0.215	0.621	0.004	0.027	0.096	0.069	0.315	0.027	0.096	0.069	0.315
1950	0.023	0.123	0.203	0.651	0.004	0.022	0.080	0.058	0.318	0.022	0.080	0.058	0.318
1951	0.023	0.116	0.265	0.596	0.003	0.019	0.119	0.100	0.284	0.019	0.119	0.100	0.284

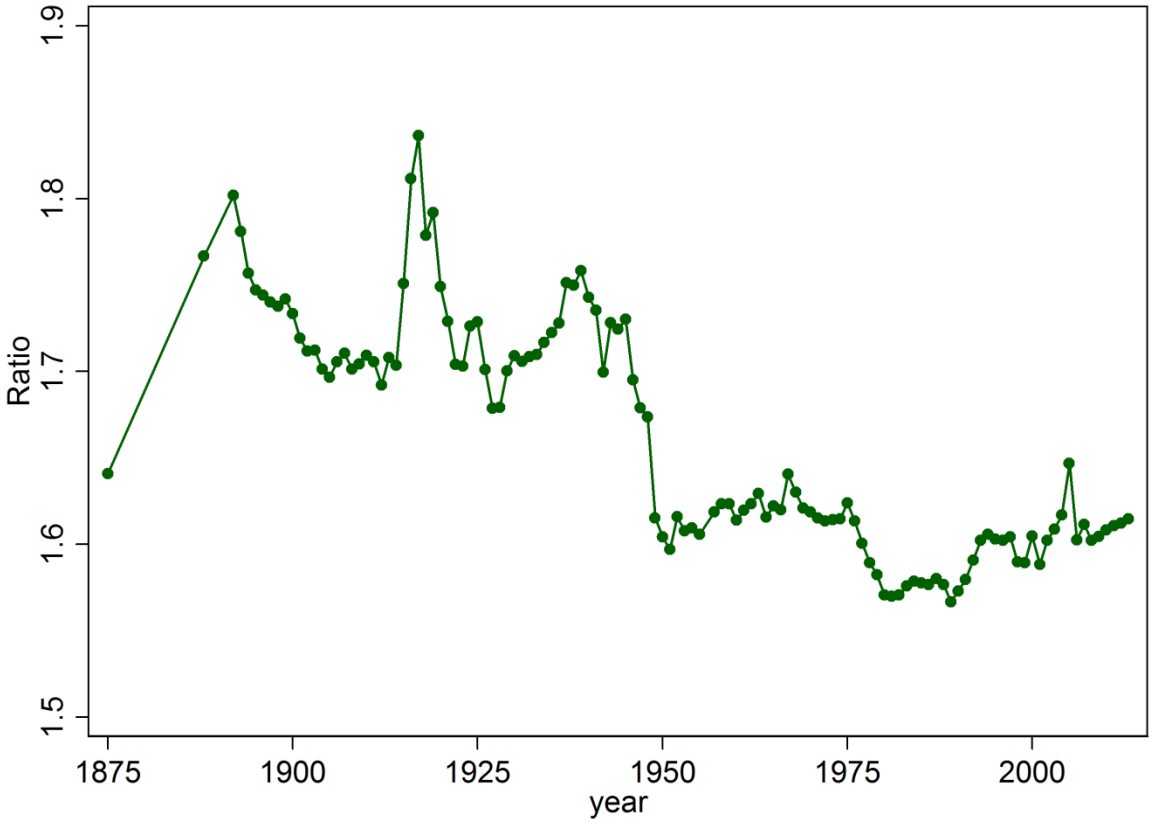
Table A 5: Mean of upper and lower bound and Gini (above median)/Affluence measures

year	Gini coefficient (average)	Gini above median)	$\frac{\mu_U}{\mu}$	Affluence
1892	0.640	0.515	1.802	0.577
1893	0.619	0.491	1.781	0.552
1894	0.593	0.461	1.757	0.522
1895	0.588	0.462	1.747	0.518
1896	0.584	0.454	1.744	0.512
1897	0.585	0.460	1.740	0.513
1898	0.588	0.467	1.738	0.516
1899	0.591	0.469	1.742	0.519
1900	0.586	0.469	1.734	0.515
1901	0.574	0.459	1.719	0.503
1902	0.568	0.454	1.712	0.496
1903	0.567	0.451	1.712	0.495
1904	0.559	0.444	1.701	0.486
1905	0.558	0.450	1.697	0.486
1906	0.566	0.453	1.706	0.493
1907	0.569	0.452	1.710	0.495
1908	0.565	0.454	1.701	0.492
1909	0.569	0.458	1.704	0.495
1910	0.572	0.461	1.709	0.499
1911	0.572	0.461	1.706	0.497
1912	0.561	0.455	1.692	0.487
1913	0.570	0.457	1.708	0.496
1914	0.559	0.438	1.704	0.483
1915	0.597	0.465	1.751	0.522
1916	0.626	0.458	1.812	0.547
1917	0.645	0.476	1.836	0.570
1918	0.594	0.430	1.779	0.514
1919	0.573	0.372	1.792	0.486
1920	0.560	0.391	1.749	0.478
1921	0.558	0.418	1.729	0.484
1922	0.543	0.414	1.704	0.470
1923	0.544	0.417	1.703	0.471
1924	0.565	0.435	1.726	0.492
1925	0.568	0.441	1.729	0.497
1926	0.555	0.456	1.701	0.492
1927	0.542	0.459	1.679	0.483
1928	0.546	0.471	1.679	0.490
1929	0.564	0.480	1.700	0.506
1930	0.572	0.486	1.709	0.513
1931	0.573	0.492	1.706	0.515
1932	0.573	0.488	1.709	0.514
1933	0.574	0.487	1.710	0.514
1934	0.576	0.480	1.717	0.514
1935	0.580	0.481	1.722	0.517
1936	0.583	0.480	1.728	0.519
1937	0.599	0.485	1.751	0.533
1938	0.585	0.454	1.750	0.515

1939	0.594	0.464	1.758	0.525
1940	0.574	0.430	1.743	0.498
1941	0.543	0.375	1.735	0.462
1942	0.513	0.348	1.700	0.431
1943	0.526	0.346	1.728	0.442
1944	0.524	0.349	1.725	0.442
1945	0.526	0.337	1.730	0.438
1946	0.494	0.300	1.695	0.401
1947	0.491	0.300	1.679	0.394
1948	0.496	0.315	1.674	0.400
1949	0.468	0.315	1.615	0.375
1950	0.464	0.318	1.604	0.372
1951	0.446	0.284	1.597	0.350
1952	0.430	0.233	1.616	0.331
1953	0.420	0.221	1.608	0.321
1954	0.421	0.224	1.609	0.324
1955	0.423	0.230	1.606	0.325
1957	0.437	0.245	1.619	0.339
1958	0.437	0.238	1.623	0.336
1959	0.434	0.233	1.623	0.334
1960	0.427	0.226	1.614	0.326
1961	0.432	0.227	1.620	0.329
1962	0.432	0.221	1.623	0.327
1963	0.436	0.223	1.629	0.331
1964	0.426	0.216	1.616	0.322
1965	0.428	0.213	1.622	0.323
1966	0.425	0.212	1.620	0.321
1967	0.441	0.227	1.641	0.338
1968	0.435	0.226	1.630	0.333
1969	0.431	0.227	1.621	0.330
1970	0.430	0.228	1.619	0.329
1971	0.427	0.224	1.615	0.326
1972	0.425	0.223	1.614	0.324
1973	0.426	0.223	1.614	0.325
1974	0.425	0.220	1.615	0.323
1975	0.429	0.217	1.624	0.325
1976	0.422	0.211	1.614	0.318
1977	0.414	0.209	1.600	0.311
1978	0.408	0.209	1.589	0.307
1979	0.404	0.207	1.582	0.303
1980	0.396	0.203	1.571	0.296
1981	0.396	0.205	1.570	0.297
1982	0.397	0.208	1.571	0.299
1983	0.400	0.212	1.576	0.303
1984	0.403	0.216	1.579	0.306
1985	0.404	0.220	1.578	0.308
1986	0.405	0.224	1.577	0.310
1987	0.409	0.227	1.580	0.313
1988	0.407	0.229	1.577	0.313
1989	0.401	0.232	1.567	0.310
1990	0.406	0.235	1.573	0.314

1991	0.411	0.242	1.580	0.321
1992	0.426	0.271	1.591	0.341
1993	0.440	0.290	1.602	0.356
1994	0.443	0.294	1.606	0.359
1995	0.442	0.293	1.603	0.357
1996	0.444	0.297	1.602	0.360
1997	0.447	0.303	1.604	0.364
1998	0.436	0.296	1.590	0.353
1999	0.437	0.301	1.589	0.356
2000	0.456	0.325	1.605	0.376
2001	0.435	0.296	1.588	0.353
2002	0.453	0.323	1.602	0.373
2003	0.461	0.334	1.609	0.382
2004	0.469	0.343	1.617	0.391
2005	0.506	0.397	1.647	0.433
2006	0.447	0.308	1.603	0.366
2007	0.457	0.322	1.611	0.377
2008	0.449	0.317	1.602	0.370
2009	0.447	0.312	1.604	0.369
2010	0.451	0.319	1.608	0.374
2011	0.451	0.312	1.611	0.371
2012	0.452	0.313	1.612	0.372
2013	0.454	0.316	1.615	0.375

Figure A 1: Mean income above the median relative to the overall mean



Appendix B: Sources of tabulated income tax data

The two income tax sources are the basis for the tabulations of taxpayers by income ranges from 1948 to 1966, which precede the micro-data available from 1967. As noted in the text, the number paying MUN tax exceeds that paying CG tax. In the tabulations (HS 1978, Tabell 314), income is equal to assessed income by the central government tax assessment for the years 1948-51, and assessed income by municipal tax assessment for the years 1952-55. This accounts for the jump in the number of taxpayers and the amount of assessed income in 1952, from 947,842 CG taxpayers in 1951 to 1,412,873 MUN taxpayers in 1952 (an increase of 49 per cent), and from 7,993 mNOK to 10,227 mNOK in 1952 (an increase of 28 per cent). The smaller percentage increase in total income reflects the fact that those paying MUN but not CG, a group referred to as (MUN-CG) here, have lower average incomes. In the tabulations for the years 1957 to 1966, income is defined as assessed income by the central government tax assessment if central government tax is levied. If not, income is defined as assessed income by the municipal tax assessment.

The sources for years before 1948 are listed in Table A 6.

Table A 6: Sources of tabulated income data

Year	Source	Taxpayer categories	Number of taxpayers	Number of groups
1875	Skattelikningen 1876 (A.N. Kiær - 1892-93, p. 110-113, included tax free incomes and Oth.Prp. nr. 11 for 1881 p. 20-25	MUN	705 460	33
1888	Sth. Prp. Nr. 48. (1890), p.42 and 122	MUN	472 104	9
1892	Oth. Prp. No. 39	CG	176 142	8
1893	Sth. Prp. No 91	CG	102 542	6
1894	Sth. Prp. No. 112	CG	66 807	5
1895	Sth. Prp. No. 104	CG	68 233	14
1896	Sth. Prp. No. 89	CG	70 454	14
1897	Statsskattens fordeling 1892/93-1898/99	CG	75 578	14
1898	Statsskattens fordeling 1899/00-1905/06	CG	94 587	15
1899	Statsskattens fordeling 1899/00-1905/06	CG	91 422	14
1900	Statsskattens fordeling 1899/00-1905/06	CG	94 367	14
1901	Statsskattens fordeling 1899/00-1905/06	CG	95 767	14
1902	Statsskattens fordeling 1899/00-1905/06	CG	97 517	14
1903	Statsskattens fordeling 1899/00-1905/06	CG	96 431	14
1906	Rygg, 1910, pages 50 and 69		677 487	17
1913	NOS VI.57, page 30*		774 308	12
1938	Stat Medd 1941, no 11 and 12, page 333		410 020	26
1948-1951	HS1978, Table 314, page 572-573	CG	Lowest: 954 524 Highest: 1 047 017	25
1952-1955	HS1978, Table 314, page 572-573	MUN	Lowest: 1 396 738 Highest: 1 439 770	25
1957-1966	HS1978, Table 314, page 572-573	MUN and CG	Lowest: 1 372 298 Highest: 1 543 022	25
1967 -	Administrative microdata			

Appendix C: Control Totals for total tax units and total income

Control totals: adults and tax units

The adult population is defined as those aged 16 and over. The data from 1948 onwards were supplied by Statistics Norway. For the period before 1948, data on the population by age is available from *Historical Statistics 1994*, Table 3.5, for 5 year intervals. We took the data for 31 December of year (t-1) as applying to year t, so that the data cover years ending in 1 or 6. From these, we calculated the proportion of the population aged 16 and over, and interpolated linearly for the intervening years. The percentages were then applied to the mean annual population figures given in *Historical Statistics 1978*, Table 9.

Total tax units are obtained by subtracting the number of married women. The number of married women is given at 5 year intervals in Historical Statistics 1994, Tabell 3.7. These are expressed as a percentage of the adult population and the percentages linearly interpolated. The results are shown in Table A2.

Control totals: household income

For total income, the starting point is a series for total household income as measured in the national accounts provided for 1978 to 2006 by Statistics Norway. Total household income 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 elsewhere classified. In order to extrapolate this series backwards, we have made use of series that are as comparable as possible, given the available materials from HS 1994 and earlier editions. In each case, the series have been linked at years where the estimates seem most comparable (for this reason we have started with 1979, rather than 1978). So that if the 1979 value from the Statistics Norway series is A1979, and first linked series is for 1975 to 1979, given by B1975, …, B1979, then for 1978 we take the value of B1978, multiplied by A1979/B1979.

Working backwards to 1950, we have used the *Nasjonalregnskap 1968-1979*, Tabell 33, pages 138-139 for the New Definition of Private Income for 1968 to 1978. For 1950 to 1968, we have used the Old Definition of Private Income from *Historisk statistikk 1978* (Statistics Norway, 1978), Tabell 59 (page 104) for 1965 to 1968 and from *Historisk statistikk 1968*, Tabell 70 (pages 110-111) for 1950 to 1964. In each case employers' social security contributions were subtracted from the total of private income; these were taken from *Nasjonalregnskap 1969-1980*, Tabell 30 (for 1969 to 1974),

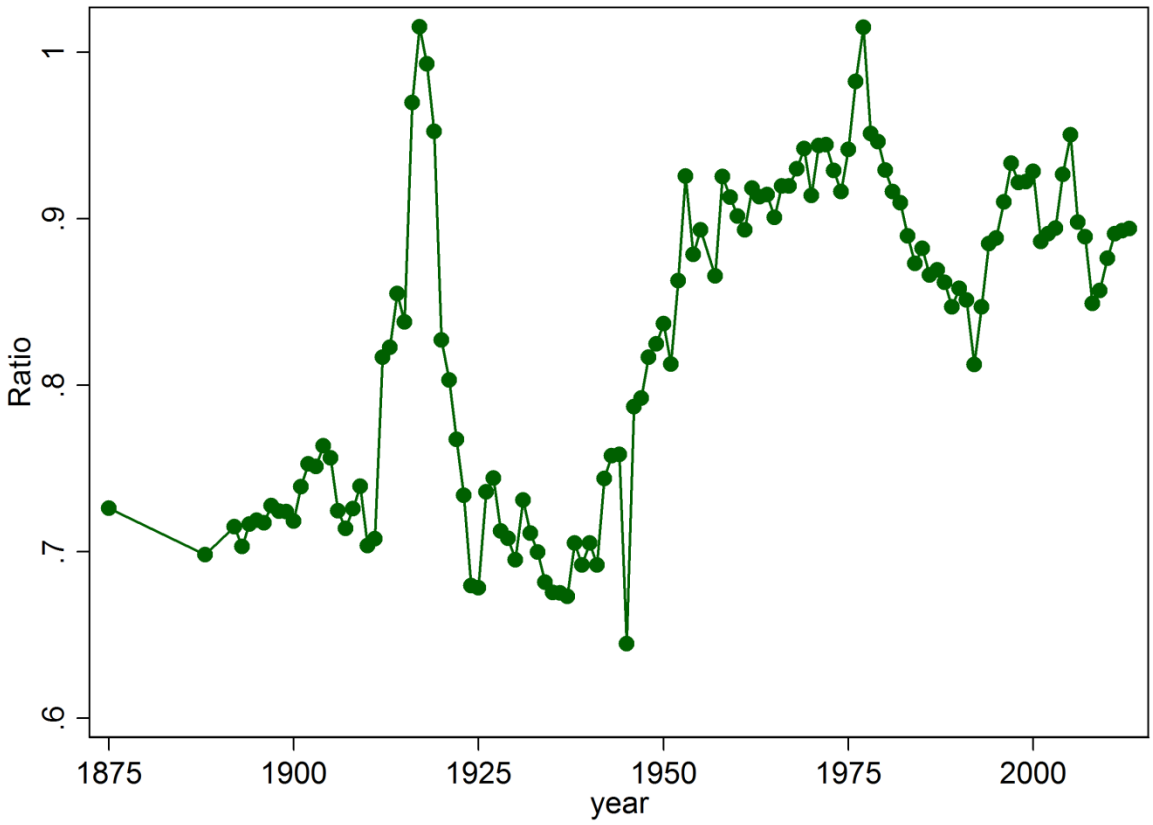
Nasjonalregnskap 1962-1978, Tabell 29 (for 1962 to 1968), *Nasjonalregnskap 1953-1969*, Tabell 14 (for 1953 to 1961), and *Nasjonalregnskap 1968-1979*, Tabell 14 (for 1950 to 1952).

For years prior to 1950, we use for 1930 to 1950 *Nasjonalregnskap 1865-1960* (NOS XII 163), Tabell 24, adding Direct taxes paid to Private disposable income. This source does not give figures for 1940 to 1945, and we have interpolated for 1940 to 1943 using the net real income figure in Tabell 35 of *Statistiske oversikter 1948* (NOS X 178). No figures are given for 1944 and 1945. For years prior to 1930, the main source is *Langtidslinjer i Norsk Økonomi 1865-1960*, Tabell VIII, where we have taken the sum of Private income from labour and capital and Transfers from government and Transfers from abroad. This source provides annual estimates from 1865 to 1900. For the period 1900 to 1930, the estimates are given at 5 yearly intervals. The figures for intermediate years have been interpolated using the series for "private gross income" from *Nasjonalregnskap 1900-1929* (NOS XI 143), Tabell 7.

The resulting series for total household income as measured in the national accounts exceeds the total income recorded in the tax statistics (the internal total) for three main reasons: (i) the omission of the income of those not covered by the tax statistics, (ii) understatement of income in the tax statistics, and (iii) differences in income definitions. In order to allow for the last of these, Aaberge and Atkinson (2010) took as the control total 72 per cent of the national accounts figure. We follow the same practice here. The resulting figures are given in Table A2.

Figure A 2 shows the internal total as a percentage of the control total over the period 1875 to 2011. As may be seen, the relationship differs between the periods before and after 1952. Before 1952, leaving aside the years of the First World War, the internal total was around 80 per cent of the control total. After 1952, the relation was much more variable and closer to 100 per cent on average.

Figure A 2: Internal total as share of Control total



Source: National accounts (Control total) and calculations from tax and poverty statistics (Internal total)

Appendix D: Source of aggregate statistics on taxpayers and poor

The source of the aggregate number of taxpayers and total assessed income (before the adjustment from all taxpayers to personal taxpayers) is displayed in Table A 7.

Table A 7: Sources of data on municipal and central government taxpayers

Years	Municipal tax aggregates	Central government tax aggregates
1875 and 1888	See Appendix B (Detailed sources)	Not applicable
1892-1899	Historical Statistics 1948, Table 220	Statistical Yearbook 1902, Table 99
1900	Historical Statistics 1948, Table 220	Statistical Yearbook 1906, Table 104
1901	Historical Statistics 1948, Table 220	Statistical Yearbook 1907, Table 104
1902	Historical Statistics 1948, Table 220	Statistical Yearbook 1908, Table 108
1903-1908	Historical Statistics 1948, Table 220	Statistical Yearbook 1909, Table 108
1909-1914	Historical Statistics 1948, Table 220	Statistical Yearbook 1915, Table 112A
1915-1916	Historical Statistics 1948, Table 220	Statistical Yearbook 1918, Table 124
1917-1919	Historical Statistics 1948, Table 220	Statistical Yearbook 1920, Table 143a
1920	Historical Statistics 1948, Table 220	Statistical Yearbook 1921, Table 160
1921	Historical Statistics 1948, Table 220	Statistical Yearbook 1924, Table 179
1922-1923	Historical Statistics 1948, Table 220	Statistical Yearbook 1926/1927, Table 178
1924-1926	Historical Statistics 1948, Table 220	Statistical Yearbook 1929, Table 199
1927-1936	Historical Statistics 1948, Table 220	Statistical Yearbook 1940, Table 267
1937-1945	Historical Statistics 1948, Table 220	NOS Tax Statistics *
1946-1953	Historical Statistics 1958, Table 200	NOS Tax Statistics *
* NOS tax statistics are annual publications; numbers for year t are reported in the publication with title "t+1/t+2" i.e. "NOS Skattestatistikk for budsjettåret 1938/39" have data for 1937 and so on.		

One problem in using these statistics is to restrict the coverage to personal taxpayers, by excluding non-personal taxpayers a group that “comprises joint-stock companies, co-operative societies and other corporations” (HS 1968, page 428). This applies to the tax data between 1921 and 1947 (from 1948 onwards we have separate reports on personal taxpayers and total taxpayers). For most years between 1937 and 1947, we have separate reports of the totals and interpolate the missing years using the ratio between personal and all taxpayers. There is little year-to-year variation in this ratio. For this reason, we use the 1937 ratio to impute the share of personal taxpayers (and their income) for the 1921-1936 period. For municipal taxpayers, this amounts to multiplying the total number of taxpayers by 0.937 and total income by 0.855. For central government taxpayers, the corresponding numbers are 0.973 and 0.848.

Data on the number of supported poor and the total poverty support 1875-1951 is obtained from the annually published poverty statistics. An overview of data for every fifth year is found in Historical Statistics 1994, Table 7.8.

Appendix E: Adjustments to tabulated data 1952 to 1966

From 1952 to 1966 the income distributions used in this paper are obtained from detailed tables in Historical Statistics 1978 (HS1978 henceforth). (There are also tables for 1948-1951 on the same pages, but these are CG taxpayers only and hence cover a lower share of the population. They are used to calculate G* for these years in the 4-class tables). Some adjustments to these data are required to make the time series consistent with the period up until 1951 and the micro data from 1967 onwards.

Adjustment for sailor taxation

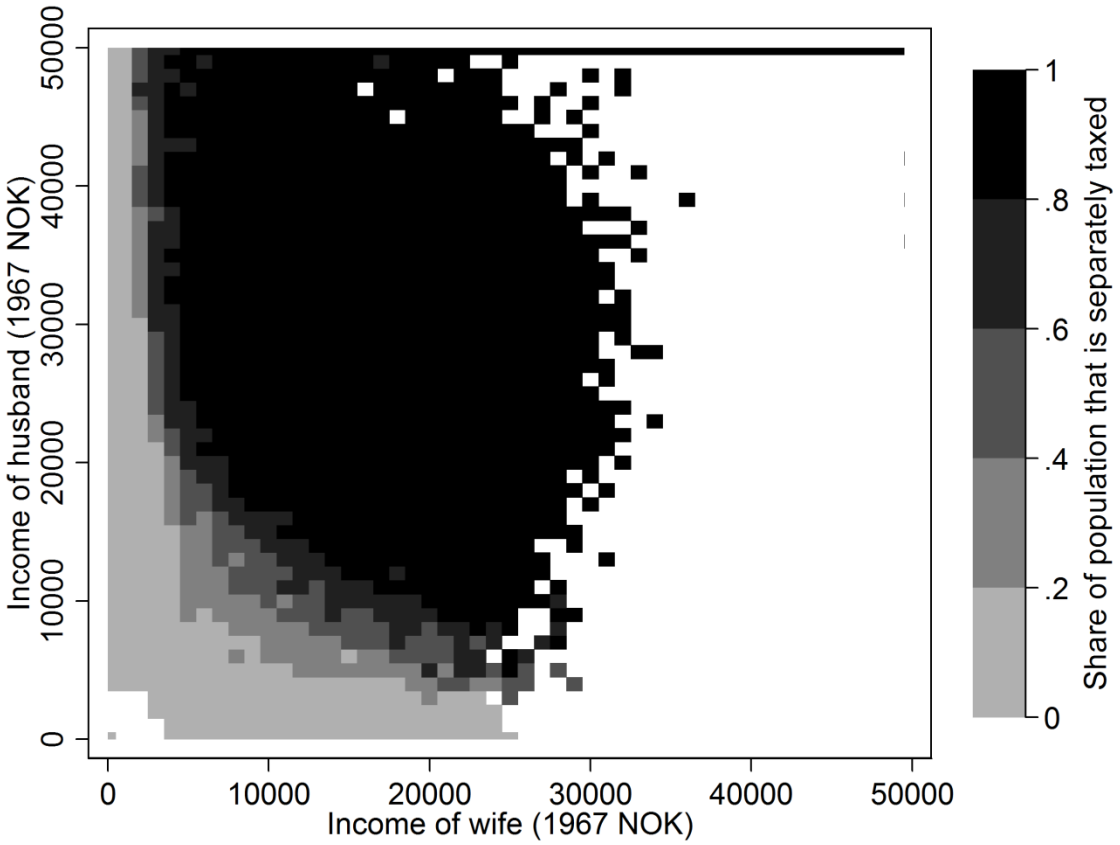
A separate sailor taxation was introduced in 1948 based on a law from 1947. Sailors are not included in the HS1978 detailed tables. We add sailors to these tables. From 1956 onward we have the number of sailors and their mean income from HS1978 table 308. Before 1956 we use the tax statistics, or HS1978 table 307 which shows total sailor taxes paid, and deduce the numbers from that. We use an uniform distribution on (0, 2*sailor mean inc) for sailor incomes and add these to the tables 1951-1966.

To apply the changes, the tabular data is re-grouped into 100 percentiles with mean incomes and population sizes. The same is done to the sailors; the tables are then added. For the spouses a transformation algorithm is applied based on registry data from 1967, where we observe spouses individually as well as a variable informing us about whether they chose to be taxed separately or not. This is described in detail in the next paragraph.

Treatment of married couples

Until 1960 married women are always taxed with their husband. From 1960, married couples could elect to be taxed separately. They are then included as two separate individuals in the tabulations. In the registry data (available from 1967) we can identify, on the individual level, which individuals were separately taxed. Hence, we can construct tabulation of units both by taxation status (as in HS1978) and by couples jointly (our preferred population, and the one used in tabulations before 1960). In this section we describe how we use information from the registry data to construct a conversion algorithm that we apply to the 1960-1966 tabulations, and in this way increase the comparability of the data.
In the 1967 income file, we observe 115,753 couples that are definitely separately taxed. These are mainly couples where both have high incomes, as shown in Figure A 3, which gives the share of couples that are separately taxed, by wife's and husband's income. The darkest shade denotes that more than 80% with this income combination is separately taxed, while the lightest shade denotes less than 20%. White means that there are few individuals with this income combination.

Figure A 3: Percentage of couples taxed separately in Norway, by husband and wife's income, 1967



In 1960, we have no registry data on incomes, but the Census of 1960 has information on the "main source of livelihood" for individuals and is available in registry form. The variable "main source of livelihood" has three possible values:

1. Income from own work
2. Pensions / transfers / income from wealth / loan / scholarships etc
3. Income from someone else's work (supported)

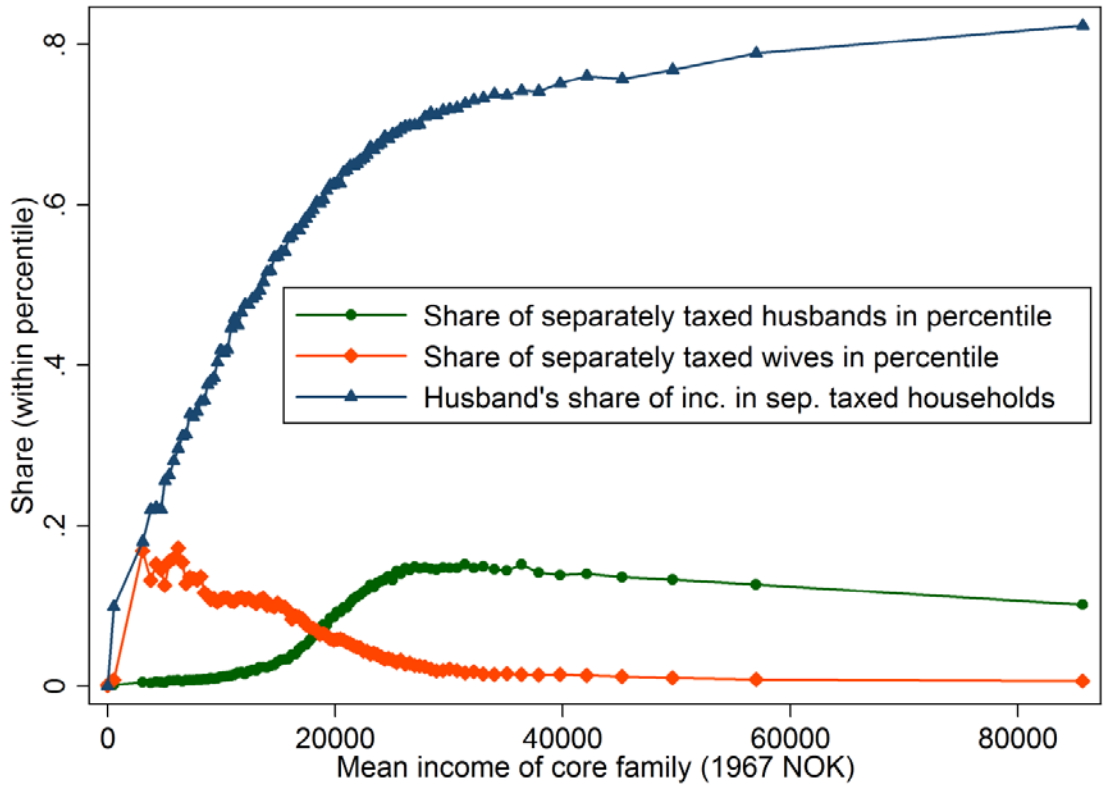
For married women, most are in category 3, while most married men are in category 1 (all combinations exist). There are around 45000 married couples where both the husband and wife are in category 1. Hence, we assume that 45000 couples were taxed separately in 1960. This corresponds well with the increase in the number of units in the tax statistics from 1959 to 1960 of 67509 (from 1 372 298 to 1 439 807), allowing for some growth in the general population in addition to the results of the tax law change. With no further data, we use a linear interpolation for the number of separately taxed between 1960 and 1967.

The approach adopted to transform the data from 1960 to 1966 is as follows:

- Construct a file of the 1967 population that corresponds to the tabular definitions from 1960 onwards. That is, merge married couples into one unit with income = (husband's income + wife's income) only if they are jointly taxed. If they are separately taxed, keep them as two units. Separate taxation usually takes place if both spouses have non-neglible incomes. In the file, each unit can be either
    - An unmarried (or widowed, etc) man
    - An unmarried (or widowed, etc) woman
    - A married couple with joint taxation (where at least one spouse is marked as not filing separately)
    - A married man with separate filing (whose wife also files separately)
    - A married woman with separate filing (whose husband also files separately)
- Divide this population into 100 percentiles, sorted by income.
- For each percentile, calculate
    - The share of units that is a married man with separate filing
    - The share of units that is a married woman with separate filing
    - Among units that are married men with separate filings, the income share of the husband in the marriage (ie (income of husband / (income of husband+income of wife))

These shares (from now on interpreted as probabilities contingent on income percentile) are shown in Figure A **4**.

Figure A 4: Prevalence of separate taxation in 1967, by income percentile



We now apply the transformation to each of the years 1960 to 1966 (year *t*) as follows:

- Adjust the "is husband" and "is wife" probabilities down by the factor (number of separately taxed couples in *t*)/(number of separately taxed couples in 1967)
- Divide the tabular population of year *t* into 100 percentiles, sorted by income. (Many of the percentiles will have equal incomes, as the tables have less than 100 categories. This is not a problem.)
- Divide each percentile group, with a given mean income y and population N, into three groups:
    - Separately taxed husbands: N*(probabilitiy of being separately taxed husband). We return to this group below
    - Separately taxed wives: N*(probability of being separately taxed wife). We delete these observations as we want to consider them together with their husbands
    - The remaining population N*(the sum of the two above probabilities) are either single individuals or jointly taxed couples and are left as is.

- For the separately taxed husbands, divide their income by the mean share of separately taxed husbands in the percentile. As we divide by a number between 0 and 1, these incomes are inflated. This step converts the separately taxed husband's incomes into couples' incomes.
- Finally, re-group the observations into 100 percentiles again. We will now have a smaller population as we have created "pseudo-couples" that closely resemble couples in the underlying population.

For 1967, the procedure gives near-perfect results. For earlier years, we cannot test the procedure directly, however, the sum of the imputed incomes are very close to the sums of original incomes (largest difference is 0.6 %), which is a sign that the interpolation is relatively precise.

Appendix F: Within-group distributions

The Gini coefficients discussed in the main paper are not dependent on assumptions on within-group inequality per se. Rather, they can be construed (in the years where there is no data on within-group dispersion) as interpolations based on within-group Gini coefficients. However, in some cases it is desirable to draw Lorenz curves for illustrative purposes or to estimate other inequality measures beside the Gini coefficient. In these cases, the following within-group distributions are one example of function forms that are consistent with the within-group Gini coefficients. Moreover, the calculations here verify that the within-group Gini coefficients are consistent across groups, that is, that the lowest-income individuals in the higher groups do not have lower incomes than the highest-income individuals in poorer groups.

CG group (highest incomes)

For the three-group case, consider a Pareto distribution for the CG group with the probability density function

$$f(y) = \frac{\alpha d^{\alpha}}{y^{\alpha-1}}$$

with mean income $\frac{\alpha d}{\alpha-1}$ and lower bound d. We set the parameter d to make the mean correspond to the mean income of the CG group, $\mu^{CG} = (1-g)/c$. This gives

$$d = \frac{\alpha - 1}{\alpha}\mu^{CG}$$

The within-group Gini coefficient of the GC group is

$$G^{*CG} = \frac{1}{2\alpha - 1}$$

Aaberge and Atkinson (2010) provide values of the Pareto coefficient α for the relevant period (1892-1903 and 1948-1951), which corresponds to within-group Gini coefficients between 0.33 and 0.5.

MUN-CG groups

For the individuals who pay municipal tax but not state tax, we use a uniform distribution with probability density function

$$f(y) = \frac{1}{b-a}, y \in [a, b]$$

with mean income $\mu = (a+b)/2$, lower bound a and upper bound b.

The Lorenz curve for a uniformly distributed population is

$$L(F) = \frac{1}{a+b}\left((b-a)F^2 + 2aF\right)$$

and the corresponding Gini coefficient is

$$G^* = 1 - 2\int_0^1 L(F)dF = \frac{1}{3}\frac{b-a}{a+b}$$

For our purposes, it is convenient to rephrase the uniform distribution using the mean m and a spread parameter z giving the relative distance of the lower and upper bound from the mean: $a = (1-z)\mu$ and $b = (1+z)\mu$. This gives a Gini coefficient of $G^{**} = z/3$.

To respect the assumption that the highest-income individual in the MUN-CG group should not have higher income than the lowest-income individual in the CG group, b^{MUN-CG} must be lower than or equal to d. Using the known means and inserting for the above equations, we get

$$(1 + z^{MUN-CG}) \leq \frac{\alpha - 1}{\alpha}\frac{\mu^{CG}}{\mu^{MUN-CG}}$$

Appendix G: Adjustments for years before 1892

Before 1892, some special adjustments are made, as no state tax was collected in this period. We do have the total number of poor and their total support, from which we get the mean income of the poor. For the NA/NT group, we use the income of the poor (85 NOK in 1875 and 76 NOK in 1888) for the upper bound Gini and 150 NOK as the upper bound Gini. We take the lowest income group in the tabulations, which contains 74 per cent of tabulated individuals in 1875 and 86 per cent of tabulated individuals in 1888, and treat these similarly to the MUN-CG groups in 1892 and thereafter. The remaining 26 and 14 per cent are treated similarly to the CG groups in later years, and the G* for these years are estimated on this population, with incomes above 400 NOK (in 1875) and 1000 NOK (in 1888).

Statistics Norway

Postal address:
PO Box 8131 Dept
NO-0033 Oslo

Office address:
Akersveien 26, Oslo
Oterveien 23, Kongsvinger

E-mail: ssb@ssb.no
Internet: www.ssb.no
Telephone: + 47 62 88 50 00

ISSN: 1892-753X

Statistisk sentralbyrå
Statistics Norway

Design: Siri Boquist