LUND UNIVERSITY

Incomes and Income Inequality in Stockholm, 1870–1970: Evidence from Micro Data

Bengtsson, Erik; Molinder, Jakob

2022

LUND UNIVERSITY

PO Box 117
221 00 Lund
+46 46-222 00 00

Incomes and Income Inequality in Stockholm, 1870–1970: Evidence from Micro Data

Erik Bengtsson &
Jakob Molinder

Incomes and Income Inequality in Stockholm, 1870–1970: Evidence from Micro Data[*]

Erik Bengtsson[§] and Jakob Molinder[a]

Abstract

This paper builds on a new dataset from the population register, comprising 38,022 randomly sampled Stockholm residents. The register was also the income tax list, with information about people's incomes of various types, age, and household composition, in the years 1870, 1880, 1900, 1910, 1920, 1940 and 1950. We use this dataset, along with the census of 1930, which uniquely included income information, and a Statistics Sweden random sample for 1960 and 1970, to calculate the growth and distribution of incomes in Stockholm over a hundred years. The Gini coefficient between 1870 and 1920 was high for both individuals and households, around 60-70 and with no change statistically significant at the 95 percent level. After 1920 inequality fell quite steadily for every benchmark year. The top decile's share of incomes (among households) fell from 50 percent or higher in 1870–1920 to less than 40 percent in 1930 and around 30 percent in 1950. The equalization was driven not only by gains for middle income groups, but also by gains for the bottom half of the distribution. Women constituted the larger share of the bottom half of income earners. Domestic servants, the single largest group in the city, earned very little but reduced their share of working-class jobs from 45 percent in 1870 to 10 percent in 1950. Generally, upgrading jobs was an important way of reducing income inequality. Decomposing the inequality decline from 1920 to 1950 between age, gender, class and sector of occupation shows that class was by far the largest determinant of inequality and of its decline.

Keywords: inequality, incomes, income distribution, Stockholm, gender gap, tax data

JEL codes: D31; N33; N34

[♣] Corresponding author. Department of Economic History, Lund University. erik.bengtsson@ekh.lu.se
[♠] Department of Economic History, Uppsala University. jakob.molinder@ekhist.uu.se

1. Introduction

A rich literature has documented the changes in income inequality since the early twentieth century. Beginning with tabulated tax data and calculating the shares of gross income accruing to the top percentile or the top decile (e.g. Piketty 2003), and continuing more recently with fuller estimates of the whole population gross and net (e.g. Piketty, Saez and Zucman 2018; Garbinti, Goupille-Lebret and Piketty 2018), it has been established that around 1900 income inequality was quite steep in industrial countries and fell sharply between 1920 and 1970, with slightly different timings in different countries. Sweden offers an interesting case for long-run inequality studies: the country was quite unequal around 1900 – indeed, one of the most unequal in the industrialized world (Roine and Waldenström (2008, 2009) – but experienced the greatest equalization until a low point of inequality was attained around 1980.

To provide a fuller account of income inequality in the late nineteenth century and the first 70 years of the twentieth century, we turn to micro data which allow us to explore various dimensions of income inequality such as variations along profession, gender and age lines. We have gone to the archives and randomly sampled Stockholm's archival population registers, which also served as the income tax lists, in 1870, 1880, 1900, 1910, 1920, 1940 and 1950; in total we have 38,022 individuals, the size per year varying over this period from 2,483 in 1880 to 8,822 in 1940. For every year the sample is random; it encompasses 1 percent of the taxpayers in our five first benchmark years, and 0.5 percent of the taxpayers in 1940 and 1950. For 1930 we use the Swedish census, the only census to have included income information, and for 1960 and 1970 we use a Statistics Sweden random sample of Swedes born on the 15th of any month, together with the income data also included. With our detailed individual-level data we can look at the role of occupation, gender, age, income composition and household makeup for the evolution of people's incomes. The contribution is at least four-fold.

For one thing, going back to 1870, we extend the time-frame of the analysis by a third of a century compared to 1903, the starting-point for Roine and Waldenström (2008, 2010). Over the early period, we find little "Kuznets" effect in Stockholm, with relatively stable inequality from 1870 to 1920: in fact, all the point estimates for these years have overlapping 95 percent confidence intervals. Our results indicate that income inequality was very high in Stockholm from 1870 to 1920, reinforcing the view from previous studies of a quite unequal country in the early 1900s (Roine and Waldenström 2008, 2009; Bengtsson et al

2018; Bengtsson 2019). We show that the results are robust whether one measures income inequality on the individual level or the household level, a matter of lively debate in the recent inequality literature (cf. Auten and Splinter 2019; Galbraith 2019; Larrimore, Mortenson and Splinter 2021).

Two, we show a declining trend in income inequality in the 1920s, 1930s and 1940s. Much of the fall in inequality is driven by relative gains by the middle-income groups (percentiles 50 to 89) at the expense of the top decile of income distribution. Unlike the movement in countries previously studied with broad income concepts (France by Garbinti et al, the US by Piketty et al), Sweden also saw some relative gains for the bottom half of the distribution before 1950, more specifically between 1920 and 1950.

Three, we break down the distribution by occupation and gender and show the importance of the changes in women's patterns of paid employment for overall distribution. In particular, the decline of domestic service and growing importance of more productive blue-collar jobs reduces income differences from the 1920s to the 1960s. Domestic service, which was lower paid than blue-collar jobs in offices or factories, shrank; from occupying 45 percent of all blue-collar workers in 1870 it supplied only 10 percent of the total in 1950. By using our micro data to reveal the composition of the bottom half of the income distribution, we show how important the varying fortunes of women are in the evolution of income inequality. On the individual level, women workers composed about a quarter of the bottom half from 1870 to 1950, while women workers were approximately three-quarters of it in 1870 and 1900, and half of this half in 1920 and 1950 (in which years men became more common in the bottom half). On the household measurement level, women-headed households were in the majority in the bottom half of households overall. We show that in 1870 and 1900 the most common occupational titles in the bottom half applied to maids and housekeepers, along with cleaners and seamstresses, and that household workers were key also in 1950, then alongside clerks. The continued decline of such low-paid jobs as maids played an important part in the overall income growth of the working class.

Four, we show the contribution of age to overall inequality and make a decomposition to show the relative weights of age, sex, sector and class for the reduction of inequality after 1920. Following Paglin (1975; cf. Harvey, Mierau and Rockey 2017), inequality over the life cycle may be seen as a less serious issue than inequality between people of the same age. We use the age information in our data to investigate the age component in overall income inequality in Stockholm and show that men had much more of an income curve over their life cycle, with incomes peaking in their forties and fifties, while

the income curve for women was much flatter. When we decompose the overall sources of inequality in 1920, 1940 and 1950 into sex, age, sector and class, it can be seen that class was by far the most important component in the high-inequality regime of 1920: the largest decrease from 1920 to the more equal 1940 and 1950 came from reducing the class differentials. Sex differentials decreased less from 1920 to 1940/1950. This decomposition result points to the crucial role of working class wage growth and jobs upgrading in the equalization of Swedish market incomes in the first half of the twentieth century; as well as the falling share of capital incomes as discussed by Roine and Waldenström (2008, 2010, pp. 309–314, 317–324).

On the whole, our comprehensive micro data approach gives a new, rounded picture of incomes and their distribution in Stockholm over a century: from a poor industrializing city in the 1870s and 1880s, to a mature industrial society in the 1930s and 1940s and the welfare society of the 1970s.

2. Background: inequality c. 1870–1970, in Stockholm and elsewhere

Why study inequality in Stockholm? We see at least four arguments for this choice, one methodological and three substantial. The methodological argument is simply that the data from Stockholm are uniquely rich, not only from a Swedish but from an international perspective. Since the income tax lists in Stockholm city were also the ongoing population register, these data provide a unique combination of information on profession and income, with demographic variables such as gender, age, and household composition. This is data of a quality which we would otherwise find only in a post-1970, register data context and the study of Stockholm is therefore uniquely placed to address the dimensions of income inequality in the 1870-1970 period.

Among the substantial arguments, the first is the outsize importance of the capital city in the national economy of Sweden. Piketty, Postel-Vinay and Rosenthal (2006, p. 240) in their study of wealth inequality in Paris from 1807 to 1994 baldly stated, "We chose Paris because a disproportionate share of the very rich lived there." The same might also be said of Stockholm for questions of Swedish inequality overall. While the city, as we will see, contained many low-income earners and poor people, the average income was 120 percent higher than the national average in 1860, and 50 percent higher in 1940 (Enflo and Roses 2015; cf. Bengtsson and Molinder 2021 on rich individuals). That incomes were high and top-income earners resided here is connected to the key role Stockholm has played in the Swedish economy, especially as a privileged trade city and as the city of finance and administration,

since the early modern period. Söderberg, Jonsson and Persson (1993) documented a relative decline of the city c. 1750–1850 but after the mid-nineteenth century Stockholm's role in the national economy recovered. Stockholm's population grew from about 100 000 in 1856 to 200 000 in 1884 and 390 000 in 1913, and its share of the country's population grew from 5.9 percent in 1900 to 10.6 percent in 1950 (Hammarström 1970, p. 5; SCB 1969). The second substantial reason for studying Stockholm is complementary to the first one: the economy of the city was quite varied. It was the financial and administrative center of the country, but also included important industrial firms such as the telephone company LM Ericsson and the engineering producer Atlas (cf. the discussion in Glete 1987, pp. 71–91). The third substantial reason is that studying a city allows one to avoid the complexities of structural transformation away from agriculture. We intend to address this important issue – so crucial to inequality studies at least since the time of Kuznets in the 1950s – in a further study, but here, we stick to the urban setting which gives consistency over time.

The original empirical contribution of our study is, then, that we can provide a consistent and detailed measure of incomes and income inequality over a hundred years, from 1870 to 1970. The leading study of income inequality in Sweden before the emergence of register data and household surveys in the late 1960s so far is the one by Roine and Waldenström (2008), based on tabulated state income taxes, from 1903 on. We will refer to Roine and Waldenström's study throughout the discussion to clarify our own contribution. Our study is also related to previous studies of incomes in Swedish cities: Järnek (1971) on Malmö from 1925–1964, Olsson (1972) on Gothenburg 1919–1960, and Gustafsson and Johansson (2003) on Gothenburg 1925–1958. The dissertations of Järnek (1971) and Olsson (1972) both focused on wage differentials and showed shrinking wage differentials between the late 1910s and the 1960s. Gustafsson and Johansson (2003), like Roine and Waldenström (2008), emphasized the part played by falling capital incomes in equalization; they also drew attention to the role of taxation, since the average tax payment for a Gothenburg resident grew from 5.7 percent in 1925 to 23.9 percent in 1958.

The economic historians Gustavsson, Husz, and Söderberg (2009) investigated wealth, not incomes, in Stockholm from 1914 to 1963 in a study with the evocative title "Collapse of a Bourgeoisie?". They studied the top 4 percent or so of wealth holders in 1914, 1928, and 1963. Mean wealth in the top groups fell by 50-60 percent in real terms over this period. They found that the periods 1928-30 and 1935-45 were crucial for wealth shrinkage in the long run: "The Depression and the Second World War, then, appear as more destructive phases to the wealthy of Stockholm than the growth of the welfare state." (p. 96) However,

policy did matter, especially increasing taxation in the 1930s and 1940s: falling capital incomes, the equalization of wages and salaries, and growing progressive taxation all "contributed to the eroding of wealth among the richest Stockholmers" (p. 105). We will not discuss the role of taxation here, as we focus on gross incomes, but we will come back to the timing of the inequality decline and drivers such as wage and class differentials.

Of course, our study is also related to an international literature. Given that a crucial contribution of the paper is to study the entire income distribution and not only top income earners as much of the previous long-run literature did (e.g. Piketty 2003; Roine and Waldenström 2008), it is related in spirit to the recent Distributional National Accounts (DINA) literature, and especially to the study by Garbinti, Goupille-Lebret, and Piketty (2018) on France from 1900 to 2014 and the study by Piketty, Saez and Zucman (2018) on the United States since 1913. Our study has greater historical depth because we have micro data for every year whereas the France study builds on tabulated data before 1970 and the US study builds on tabulated data before 1962.[1] However, our study does not encompass the national income as a whole as the DINA studies do; we study only the incomes received by households and not for example the corporate profits retained. Thus, we believe that our approach is complementary to the DINA approach in mapping the evolution of incomes and their distribution over time.

3. Data and method

This study builds on a new sample of income taxation records, with micro data at the individual and household level. The new taxation law of 1861 (*1861 års bevillningsförordning*) is said to have created the modern income tax system in Sweden (Lodin et al. 2017, p. 9). All non-agriculturalists were to pay 1 percent of their net income in tax, while farmers were to pay 1 percent of an assumed 5 percent net yield of the value of their land. This replaced a chaotic previous system of various taxes (Hedlund Nyström 1972; Cassel 1906). Everyone's income was estimated and if one earned at least 400 kr then one had to pay 1 percent in tax (with some rights for deductions). Until 1909, one's voting rights were allocated according to one's taxed income under the *bevillning* (cf. discussion in Piketty 2020), so people had good reasons not to shirk the tax. In 1902-1903 a new income tax was added, and the *bevillning* evaluation was now used solely for the municipal income tax, which was collected by the state but distributed to the municipalities. In 1928 the difference between

---

[1] For the US, see also the discussion in Fisher-Post (2020, pp. 5–14) who uses tabulated data, a household survey, and Pareto interpolation techniques to extend the comprehensive estimate for the entire population back to 1913.

the *bevillning* and the *inkomstskatt* was abolished; from this year our source is the municipal income tax (*kommunalskatt*).

We can conceive the strengths and disadvantages of the *bevillning* source, used for our years 1870–1920, by going back to the statistician Ludwig Widell's 1907 appeal for the calculation of Swedish income distribution statistics (an appeal which was not heeded). Widell thought it scandalous that Sweden, at a time of intense debate on living standards and inequality (the so-called "social question"), lacked comprehensive income distribution statistics, despite the availability of the necessary materials, the *bevillning* data, going back to 1862. They present comprehensive income statistics, but with some problems, said Widell. One problem is that the *bevillning*¸ before a tweak in 1920, did not tax paid dividends.[2] This gives a downward bias to our inequality estimates before 1920, albeit a limited one, since dividends were not a large part of income. In Appendix A, we simulate the size of our downward bias in inequality estimates by adding dividends to the top income earners. The robustness check of adding dividends merely confirms that inequality was indeed very high from 1870 to 1920 and then fell.

Another problem discussed by Widell is that people with incomes in several parishes were listed separately in each parish. This is not a problem in our study since the original taxation lists in Stockholm present briefly also a person's income in other places (even if they are not taxed in the same parish). A third problem is that those with very low incomes did not pay tax. The threshold for the *bevillning* of 400 kr was significantly lower than that for the *inkomstskatt* introduced in 1903 (1000 kr)[3], but still, earners below 400 kr did not tax. We will adjust for this problem, as discussed below. We think that, despite the non-representation of dividends before 1920, and the need to estimate the incomes of farmers and low-income earners from other sources, the *bevillning* is an excellent micro data source of light to shed on Swedish incomes.

We digitized the local taxation lists used by the state and local governments to collect taxes and duties. In Stockholm, the taxation lists were also the population register (*mantalslängder*), so we also included those who did not pay income tax (children and those

---

[2] For a discussion of the 1920 reform, see *Statistisk Årsbok för Stockholm* 1922, p. 127. The reason that dividends were not taxed under the *bevillning* before 1920 is that corporations were taxed alongside individuals in this period (as in the Finnish income tax used by Roikonen and Heikkinen 2020 for the years 1865-1880; see discussion on p. 50). In our estimates, we include people, but not corporations.

[3] This, together with the overall higher coverage in cities than in rural areas, explains the significantly higher coverage in our study than in that of Roine and Waldenström (2008, 2010). In the first year that they studied, 1903, 7.2 percent of households paid the state income tax that the study is built on (Roine and Waldenström 2010, Table 7C.1). Our coverage in Stockholm in 1900 is 35 percent (see Table 1 below) and for the country as a whole our coverage is 30 percent (Bengtsson, Molinder and Prado 2022).

with low income). The original population registers/taxation lists for Stockholm city are held at the Stockholm City Archive (*Stockholms Stadsarkiv*). The minimum income level needed to pay the tax was 400 kr until 1884, when the level was raised to 500 kr. In 1903 the level was further raised to 600 kr, and in 1958, 1200 kr (cf. Gustafsson and Johansson 2003, p. 194). We sampled 1 percent of the population in 1870, 1880, 1900, 1910 and 1920, and 0.5 percent in 1940 and 1950. See Appendix B for further information on the sampling.

The information given in the lists is job title (and sometimes workplace) or position in the household ("son", "wife", etc.), name, household structure, birth year, birthplace (in 1940 and 1950), and income of three kinds: from labor, from capital and from business. At the beginning of our period, the income tax was flat at 1 percent, and as pointed out above, until 1909 votes in local politics were allocated according to one's taxed income. Thus, there were no strong incentive to evade taxes.[4] Tax progressivity in the early 1900s was limited to 5 percent; the difference in the tax rate between a high-income earner and an average income earner did not become marked until the 1930s or even 1940s (cf. Henrekson and Stenkula 2015, pp. 17–19). Previous researchers have judged the tax returns to be quite correct, with the qualification that the self-employed with good accounting skills perhaps paid less than they might have (Järnek 1968, p. 19); Bentzel (1952, p. 29, cf. p. 61–63) believed that incomes in the 1930s and 1940s were understated by 2 to 7 percent. In the twentieth century, Sweden emerged as one of the most tax-compliant countries in the world (cf. Nistovskaya and d'Arcy 2018).

Table 1 presents our sample. As the taxation lists are also population registers, those without incomes (children, housewives and so on) are included. Women were a larger part of the population in Stockholm in these years (Matovic 1984), a feature which is also true of our sample; women's share of this sample varies from 55 to 60 percent.

---

[4] The requirement for individuals to file a tax return was legislated in 1902 (Malmer 2003, pp. 23–25). For a detailed discussion of how this was done, see the handbook by Bodin and Palmgren (1902).

Table 1. Sample

Year	1870	1880	1900	1910	1920	1930*	1940	1950	1960**	1970**
N	2,831	2,483	4,885	5,787	6,268	11,928	8,822	6,946		67,733
N, adults	2,215	1,875	3,844	4,468	5,060	10,538	8,804	6,476	21,195	52,731
N, taxpayers	617	706	1,318	1,947	3,362	7,377	6,277	5,225	17,204	41,940
					Panel B. Share of adults who paid income tax					
All	30 %	38 %	35 %	44 %	67 %	70 %	71 %	81 %	81 %	80 %
Men	62 %	82 %	76 %	78 %	82 %	80 %	86 %	93 %	91 %	86 %
Women	8 %	8 %	9 %	19 %	54 %	59 %	61 %	71 %	73 %	74 %

Note: Adult defined as 16 years or older.

* Data for 1930 is from the National Archive's (*Riksarkivet)* digitization of records from the 1930 census, which include income. This is an ongoing effort and we have included the observations that are available at this point.

** Data for 1960 and 1970 come from the born-on-the-15th sample. This is a Statistics Sweden dataset and does not build on the taxation archives. The 1960 file includes adults only.

The lower panel of Table 1 shows the share paying income taxes. It clearly conveys that coverage is a great deal better for men than for women. The late 1910s saw very large nominal wage increases for workers (cf. Bengtsson and Molinder 2017), and the share who were above the income tax threshold grew markedly. That the share of adults who paid taxes before 1920 was below 50 percent might be seen as a problem, but we should point out three things.

One, compared to the share covered in other tabulated tax sources used in the top incomes literature, the coverage in Stockholm is very good. While the share of tax-payers in the early years of top income studies, as discussed by Atkinson and Piketty (2007, Table 13.0) was initially under 5 percent in France, 0.1 percent in the UK, around 1 percent in the United States, and so on, our coverage is never lower than 30 percent This is because of the low threshold for paying the *bevillning* tax. Two, of course all studies built on tabulated taxed incomes suffer from the same problem as our study, which builds on individual tax returns. Three, a significant share of adults still do not pay income tax, so this problem is present even in contemporary studies. Splinter (2019) shows that in the US today, 36 percent of working-age adults pay no income tax, a share corresponding to our level in 1920 and 1940. While we would thus point out that this problem is far from unique to our study, we have of course taken care to adjust our data for the low incomes of non-filers. To ascertain that the sample

strategy is stable and precise, we also in Appendix B compare our estimates with those of tabulations of incomes from 1900 to 1950.[5]

The job titles show that a large share of those employed who did not pay the income tax were women textile workers and women working as domestic servants. Among men, metal workers, transport workers, domestic workers and unclassified workers were the most important non-covered groups. The important previous study of top income shares in Sweden, Roine and Waldenström (2010), made the uniform assumption that those below the threshold earned 80 % of the threshold, but in a previous study (Bengtsson 2021), we studied the incomes of the groups below the taxable level in a variety of previous papers and sources. Building on these data, in the present work we imputed incomes for the groups below the threshold for tax payment, such as textile workers, transport workers, widows, and widowers. See Appendix C for the details and robustness checks with different procedures for the treatment of non-filers.

An important discussion in the historical inequality literature concerns the unit of analysis. Should we measure inequality at the individual level or that of the household? And what level *can* we measure – in practice, studies often compare tax units, which are not identical with individuals nor households. This is not completely consistent in its use of individuals or households. It is dependent on the sources, the tax systems or household surveys used. Atkinson (2007, pp. 28–29) in a crucial methodological discussion pointed to the importance of the unit of analysis, a debate which is very much alive today (cf. Galbraith 2019, pp. 338–340; Larrimore, Mortenson and Splinter 2021). With regard to this discussion, a great advantage of the Stockholm data is that for all our years 1870 to 1950, we can calculate inequality both on the individual level and on that of the household. Table 2 shows the household composition of our Stockholm data from 1870 to 1950.

---

[5] A related concern is that our 1 percent random sampling strategy might not capture enough of the very top incomes. If the distribution is strongly right-skewed then this might underestimate inequality. In a companion paper (Bengtsson and Molinder 2021) we study the full population of the wealthy suburb of Djursholm, and nine city blocks in the wealthiest inner city area of Östermalm, in the years 1909, 1915, 1927, 1935 and 1950. In total the dataset encompasses 12,918 tax payers from Djursholm and 8 137 tax payers from Östermalm. The estimates made on the basis of this strategic sample cannot be compared with the estimates made here regarding levels of inequality, because we could not estimate the Ginis of top shares in the other paper. However, the trends can be compared. In the previous paper, we found that the median income relative to GDP per capita for the economic elite in our two elite areas was 14 times average income (GDP/capita) in 1909 and 1915, 10 times in 1927 and 1935, and 6 times in 1950. For the professional elite, the ratio was 12 in 1909, 9 in 1915 and 1927, 6 in 1935, and 4 in 1950. The types of estimate are of course very different from what we present here, but the overall picture is similar in that inequality was very high before 1920 and declined after 1920. Furthermore, in Appendix A we present alternative estimates of income inequality in 1900, 1920, 1940 and 1950 based on tabulations of *all* income taxpayers; these reinforce the view that our original estimates are reliable.

Table 2. Household composition

	Average						*Percent*
Year	Adults	Children	Total	Cons. units	No. of tax filers	No. of tax filers (minimum. one)	Any tax filer
1870	1.33	0.37	1.71	1.34	0.37	1.03	36%
1880	1.37	0.44	1.81	1.40	0.51	1.03	50%
1900	1.38	0.37	1.75	1.38	0.47	1.02	46%
1910	1.39	0.41	1.80	1.40	0.60	1.05	57 %
1920	1.34	0.32	1.66	1.33	0.89	1.10	81%
1930	1.37	0.18	1.55	1.28	0.96	1.24	78 %
1950	1.49	0.36	1.86	1.43	1.21	1.36	89%

Note: Consumption units are 1.00 for a single household, 1.51 for a couple, +0.60 for an extra adult, +0.52 for the first child, and +0.42 for any further child.

The number of children in the 1930 data is puzzlingly low. It may have to do with problematic coding of the census by the Riksarkivet.

Table 2 shows that the households become larger over time and that the number of income earners grew as well. In the analysis we will see how this affects inequality estimates.

4. Results

4.1 Income inequality in Stockholm, 1870–1950

Figure 1 shows the Gini coefficient for Stockholm from 1870 to 1970, measuring inequality at both the individual level and the household level. On both counts, the Gini coefficient in Stockholm was very high in the late nineteenth century, above 60. It remained mostly stable at this high level up to the 1920s, oscillating between 60 and 70. Given the underlying uncertainty and sensitivity of the estimates to limited samples, especially those of top income earners, the swings before the 1920s should be interpreted with caution, while the precision is much greater from this point on (see Figures A1 and A2 in Appendix A for versions of the series with bootstrapped confidence intervals). Recall also that before 1920 our data do not encompass stock dividends and so we slightly underestimate inequality in this period (see Appendix A). Overall, the verdict for the period up to 1920 is that inequality was very high, and essentially stable. This supports the view in Roine and Waldenström's (2008) study of national income distribution from 1903 on, and Bengtsson et al's (2018) study of wealth inequality up to 1900, of a very unequal economy in the late 1800s and early 1900s.

11

In light of the high and largely stable level of inequality up to WW1, the fall in inequality that took place after 1920 is remarkable. Both the between-individual and the household Gini fell by more than 15 points between 1920 and 1940, and went down by roughly further five points up to 1950. This can be compared to the contemporary Gini difference in household level incomes between Sweden and the United States of about 15 Gini-points, according to the Luxembourg Income Study.

Figure 1. Gini-coefficient for Stockholm, 1870–1970

Note: Consumption weights taken from modern estimates by Statistics Sweden (SCB).
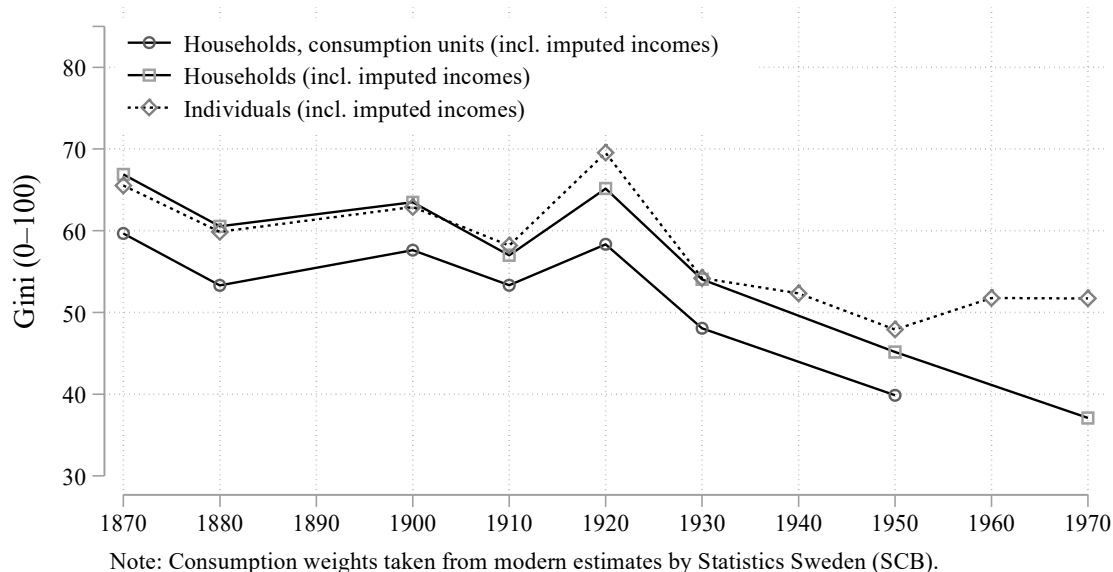
Since we study Stockholm over 80–100 years, from its early industrialization phase to its post-war welfare society, we might expect demographic changes to affect our measures. Piketty, Saez and Zucman (2018, p. 586) in their study of the US since the 1960s remind us that falling marriage rates among low-income earners have increased the inequality between households. Such compositional shifts may matter in our case too: in the relatively poor city of the late 1800s, women and children had to work and contribute to the household budget (even though women's wages were much lower than men's). With growing real wages, growing urbanization and increased numbers of marriages, the male breadwinner-housewife household became much more common in the 1930s and 1940s (Edvinsson and Nordlund Edvinsson 2017). For this reason, it is an advantage that with our Stockholm-based sources, we can actually measure incomes and their distribution both on the individual and the household level, and that we thus can also capture how much the household structure makes a

difference to the distribution of incomes. We see in Figure 1 that overall differences in estimates are surprisingly small, a couple of percentage points in every year.[6]

But there is an interesting pattern: that the relationship between individual-level measures and household-level measures invert over time. In 1870 and 1880, inequality was slightly higher if measured at the household level and not at the individual level, but in 1920 and 1950, inequality was higher at the individual level. This indicates that secondary incomes were concentrated in higher-income households in 1870 and 1880, while they were more common, or more substantial, in lower-income households in 1920 and 1950.[7] For Swedish cities in 1936, Nyberg (1989, pp. 151–155) found a U-curve in the relationship between husbands' incomes and the wives' propensity for wage labor; while low-income and high-income households were more likely to have wives working in the labor market, middle-income households were less likely. For Stockholm in 1930 she found that wives in low-income households (where men earned 2,000–3,000 kr) were more likely to engage in wage labor than wives in high-income households (where men earned 6,000–10,000 kr). It might be that increased access to good jobs (such as clerks – see Holmberg 2013 on jobs in banks, and Svensson 2003 for an aggregate view) increased the profitability of wage labor for lower- and middle-class women while the male breadwinner pattern was strengthened in upper-income families around 1950; at least, this is a hypothesis for further research.

After 1950 there is also a stark divergence between the individual-level and household-level measures, underlining the need to consider both measures for a complete appreciation of the dynamics of inequality. While household level inequality continued to fall further – by more than five points between 1950 and 1970 – the individual level Gini grew slightly. This was driven by the large-scale entry of married women, especially women from working class families, to the formal labor market, a process which began earlier in Stockholm than in the rest of the country (Molinder, 2021). While individual-level income inequality increased among women as more married women began to earn taxable incomes (see Figure A3 in the Appendix), their husbands' incomes placed them in the lower part of the

---

[6] Björklund and Palme (2000, pp. 126–127) decompose changes in income inequality from 1951 to 1973, using household surveys, and find that most, 87 percent, of the decline in inequality was driven by a decline in the inequality between men. A smaller share was explained by women's increasing wage labor.

[7] One might wonder if this result could be driven by the changing representation of women in the tax records. Recall (Table 1) that in 1870, 1880 and 1900 less than 10 percent of women earned enough to pay income tax. However, the estimates are adjusted with imputed incomes for women and men with job titles but no taxed income. A problem might still linger with women titled only "wife" (*fru*) but who nevertheless (a) undertook wage labor, but (b) did not earn enough to pay income tax. The pattern from 1920 (when 54 percent of women paid income tax) to 1950 (71 percent) is less likely to be affected by such issues.

household income distribution. The addition of a second income earner thus contributed to a compression of the household income distribution.

Included in the graph is also a measure of the inequality between households accounting for differences in household composition using modern consumption weights applied by Statistics Sweden. This uniformly lowered inequality levels, but the trend is very much the same. It indicates that high-income individuals had larger households than lower-income individuals in Stockholm between 1870 and 1950.[8]

4.2 What was happening at different points of the income distribution?

Figure 2 plots the share of income to different quantiles of the income distribution, following Piketty's (2014, pp. 246–268) terminology distinguishing between the bottom 50 percent (the "lower class"), the next 40 percent (the "middle class"), and the top 10 percent (the "upper class") of households. The share for the bottom half was very small until the post-1920 equalization: between 15 and 20 percent between 1870 and 1920. During the interwar period and the 1940s the share grew, however, and ended about 10 percentage points higher by the 1960s and 1970s. The shift was more dramatic for the "middle class", whose income share grew from around 30 percent in the 1870s and 35 percent in 1920 to more than 50 percent by 1970. The increase in the share of income accruing to the lower and middle class was mirrored by a fall in the share of the top 10 percent of households. Their share stood at a massive 50 percent before the 1920s and showed very little change over this period. Between 1920 and 1950 the share collapsed to 30 percent and continued to decline to less than 25 percent in 1970, when for the first time, the share for the lower class superseded that for the top 10 percent.

---

[8] Edin and Hutchison (1936) found a positive relationship between education and fertility among couples in Stockholm in the 1920s. See the review by Thomas (1937). Bernhardt (1972) discusses Edin and Hutchinson's findings on Stockholm in the 1920s as an exception to the general rule of a negative association between income and fertility. Molitoris and Dribe (2016) place the findings in a longer term perspective by studying Stockholm in the period 1878 to 1926. They find that the positive income-fertility association is only present in the 1920s.

Figure 2. Share of Income by Quantile, 1870–1950

Households (incl. imputed incomes)
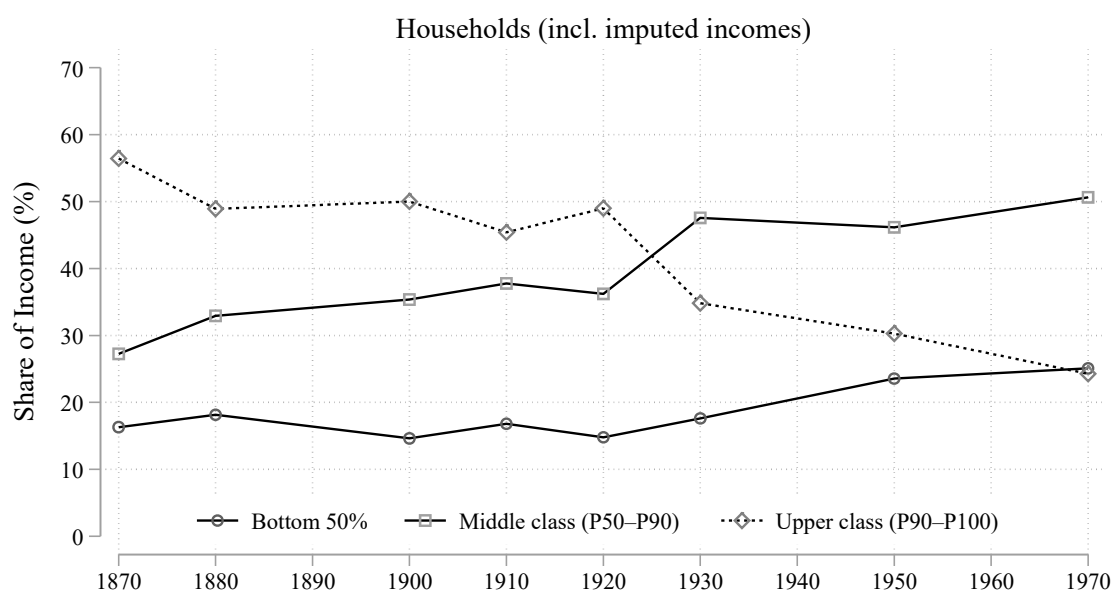
Patterns of relatively stagnant income shares in the bottom half have been shown before: for example, Gustavsson, Husz and Söderberg (2009, p. 99) for Stockholm in the 1913 to 1930 period, Fisher-Post (2020) on the United States, and Garbanti, Goupille-Lebret and Piketty (2020, p. 69) who traced the share of the bottom half in France at 10-15 percent from 1900 to 1930, 20 percent in 1945 and stagnancy from then until the late 1960s. A stagnant bottom half share is a subject of recurring debate in the research on the contemporary period also (Piketty, Saez and Zucman (2018, pp. 581–586; Galbraith 2020), so it becomes interesting to disentangle, with our individual- and household-level data, who composed the bottom half in our context over eighty years. The pro-poor profile of income growth in Stockholm 1920 to 1950 is without correspondence in France (Garbinti, Goupille-Lebret and Piketty 2018) or in the US (Fisher-Post 2020), and both the similarities and differences between Stockholm and the other places make it pertinent to break down how the bottom half of income earners changed in composition over time.

Figure 3 shows the composition of the bottom half of the income distribution from 1870 to 1950. As we see, women dominate the bottom half (with individual data): they comprise about 85 percent until 1920, when the male share grows to about 25 percent. More than 55 percent of the city's population were women, but they are still over-represented in the bottom half. Non-workers such as wives and widows form a very large share, about 60 to 70 percent in every year. Female workers are about 25 percent in the early years, but the share decreases to 15 percent in 1940 and 10 percent in 1950. Taken together, Figures 3 and 4 show

that gender was an important aspect of inequality at the end of the nineteenth century and first half of the twentieth.

Panel B of Figure 3 shows the title of the household head for the bottom half of households. We see that women-headed households dominated the bottom half of households and that households headed by a non-working woman such as a retired widow were also important. Male workers – say metal workers, transport workers or construction workers who might play an important role in our mental image of the "bottom half" in industrial society – were actually only a small fraction of the bottom half, both in terms of individuals and in terms of households. This means that the income share patterns shown in Figure 2 should not necessarily be interpreted as a result of changes in pay differentials, say the difference between the yearly pay for a factory worker and that for a teacher. Instead, the wages for female service sector workers, and the paid labor intensity for women, are crucial for the fate of the income growth of the bottom half. This resonates with Emma Griffin's (2014) recent argument that unemployment and the availability of paid jobs are major aspects of the living standards of the working class during industrialization.[9] The broad trend of a declining share of households headed by low-income female workers points to the role of the upgrading of the occupational structure, in particular the decline in domestic services, for the shift in the income distribution after 1920.

---

[9] It should also be noted that a very small shift in income was required to end up in the bottom 50 percent as compared to, say, the 70th percentile. As an example, the difference in 1920 was less than 30 percent. The shifts in the composition of the bottom 50 percent are therefore very sensitive to assumptions about incomes of non-filers.

Figure 3. The composition of the bottom half of the income distribution, 1870–1950

Panel A. Individuals

Panel B. Households

Note: This figure shows the composition of the bottom 50 percent of income-earning units. Panel A refers to individuals and Panel B to households.

A great strength of the micro data is that they give us abundant detail on the income distribution. All income distributions in the end consist of people, and these people can disappear when we use tabulated data. However, the micro data used here allow us also to calculate the typical traits of the people in the various parts of the distribution. From Figure 3

we know the importance of gender and employment intensity for the bottom half. Table 3 goes into further detail by listing the three most common occupational titles for the bottom half, the "middle class", and the top decile from 1870 to 1950. Table 3 again brings home the importance of gender: job titles with female endings in the original Swedish dominate the bottom half in the early years: maids, cleaners, housekeepers. We see that the clerk appears only in 1920, as the most common job in the "middle class", but that, by 1950, 'clerk' is both most common in the middle group and the third most common title in the bottom half. This indicates an upgrading of the structure of jobs; we come back to this below.

Table 3. Three most common occupational titles by quantile and year

Quantile	No.	1870–1880	1900	1920	1950
Bottom 50	1	Maid (*Piga*)	Maid (*Piga*	Servant (*Tjänare*)	Housekeeper (*Hembiträde*)
	2	Cleaner (*Städerska*)	Seamstress (*Sömmerska*)	Seamstress (*Sömmerska*)	Cleaner (*Städerska*)
	3	Housekeeper (*Hushållerska*)	Housekeeper (*Hushållerska*)	Housekeeper (*Hembiträde*)	Clerk (*Kontorist*)
Middle	1	Workman (*Arbetskarl*)	Worker (*Arbetare*)	Clerk (*Kontorist*)	Clerk (*Kontorist*)
	2	Seamstress (*Sömmerska*	Workman (*Arbetskarl*)	Worker (*Arbetare*)	White collar worker (*Tjänsteman*)
	3	(Farm) hand (*Dräng*)	Shop assistant (*Handelsbiträde*)	Shop assistant (*Biträde*	Chauffeur (*Chaufför*)
Top decile	1	Bookkeeper (*Bokhållare*)	Merchant (*Grosshandlare*)	CEO (*Direktör*)	CEO (*Direktör*)
	2	Trader (*Handlare*)	Trader (*Handlare*)	White collar worker (*Tjänsteman*)	Civil engineer (*Civilingenjör*)
	3	Merchant (*Grosshandlare*)	Bookkeeper (*Bokhållare*)	Engineer (*Ingenjör*)	Businessman (*Köpman*)

Going back to the "big picture", Figure 4 also gives the broad long run perspective, but focusing on top income shares: the share of incomes going to the top decile, to the top percentile, and the top 0.1 percent. The difference for the top 1 percent and the top 10 percent is interesting:

the top 1 percent share grows between 1900 and 1920 while the top 10 percent share does not. This top-skewed increase in inequality also reinforces why it is so difficult to make precise estimates for 1920. In Appendix B we show, comparing our results with results using tabulated data for all tax payers (and adjustments for non-filers), that 1920 is unusual in this regard.

**Figure 4. Top income shares in Stockholm, 1870–1950**
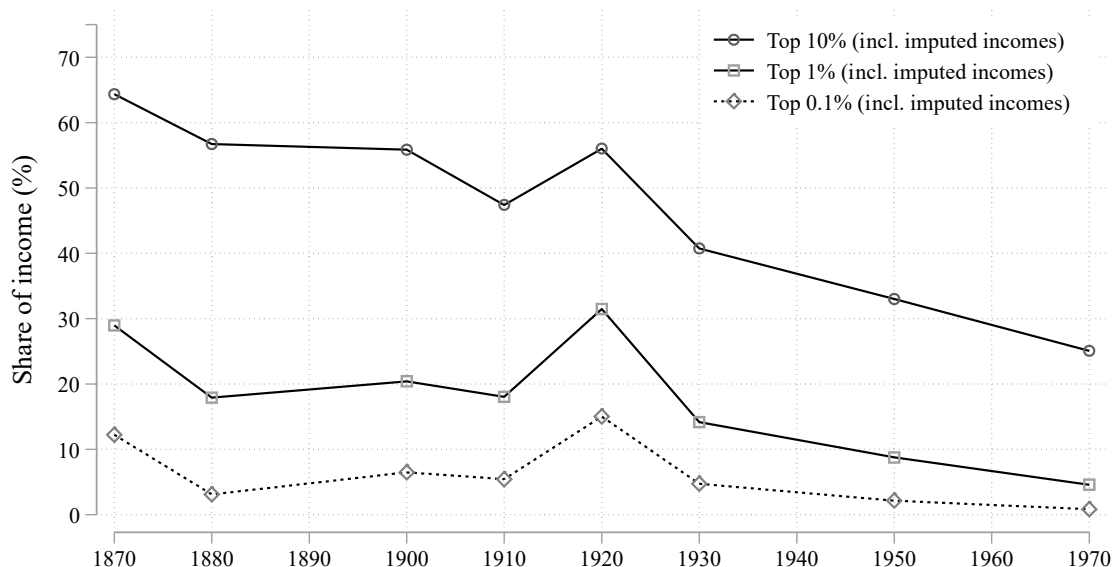
What conclusions can we draw about the timing of the equalization? With the top income share measures as with the Gini coefficients in Figure 1, we find that inequality peaked in 1920, and fell steeply to 1940, and further to 1950. Gustavsson, Husz and Söderberg (2009, p. 96) have argued that the Depression and the Second World War were "more destructive phases to the wealthy of Stockholm than the growth of the welfare state." This is not incompatible with our major fall from 1920 to 1930 and 1940. 1920 was extremely unequal and much happened between 1920 and 1930 (cf. Bengtsson and Molinder 2021), but also from 1930 to 1940 and from 1940 to 1950. Roine and Waldenström (2008, p. 370) found for the country as a whole that about two thirds of the decline in income inequality – measured as top income shares – which occurred from a World War One-era peak to the lowest levels in the early 1980s, took place before 1950. Our findings for Stockholm concur with this; but the explorations in Figure 3 and Table 3 point to the importance of gender and the upgrading of the jobs structure, and not only the destruction of capital in the Depression or the Wars, as important factors of equalization.

Figure 5 sheds more light on this process by showing income growth per income group. Here the income earners are divided into a bottom half, and then into deciles above the median. The top 1 percent is also shown separately. In 1870 to 1900 the middle groups,

deciles 6, 7 and 8, had the best income growth, while the bottom half and the top decile made poor growth. The pattern of strong income growth for the middle groups fits with the findings of healthy wage growth for key worker groups such as construction workers in this period (Berglund 1982; Söderberg 1987; Molinder and Ericsson 2020). In 1900 to 1920, the bottom half lost out in real terms while the middle groups, especially deciles 6 and 7, did reasonably well, the 91st to 99th percentiles did badly with almost no real income growth, and the top percentile evolved further than any other. The First World War saw great profits among merchants and others who made profits from goods in short supply and high prices – Roine and Waldenström (2008) find that 1916 was the year with the highest top percentile share of incomes in Sweden's modern history. In 1920 this elite-driven inequality persisted, something which is also seen in micro data studies of the smaller cities of Malmö in southern Sweden and Esbjerg in Denmark (cf. Svensson and Bengtsson 2021; Abildgren 2019). This is further discussed in Appendix B.

Between 1920 and 1940 the bottom half rebounded in style, with real income growth above 3 per cent per year. The income growth pattern is supremely egalitarian, with the income growth smaller the further up one goes in the distribution, and the top 1 percent are in the red between 1920 and 1950. This was a period of Social Democratic Fordism, creating many unskilled jobs in manufacturing (cf. Schön 2014, pp. 266–298) and with an important restructuring of the female labor force with, as we will see, a widespread shift out of domestic service into better-paid jobs in the service sector and in manufacturing. Gustavsson, Husz and Söderberg (2009, p. 99) in their comparison of income distribution in Stockholm in 1913, 1930, 1950 and 1962 found that the bottom third or so earned about 10 per cent of pre-tax incomes in all years, so that the equalization over the period was driven especially by the increasing shares of the middle groups, from the fourth to the eight deciles. The equalization from 1913 to 1930 in their data was driven mostly by gains for deciles 7-9 and the loss of decile 10, while the equalization from 1930 to 1950 was broader. Our results for 1920 to 1950 also suggest that equalization during this period was broader than the earlier pattern driven by middle groups.

Figure 5. Average Yearly Real Growth of Income by Quantile
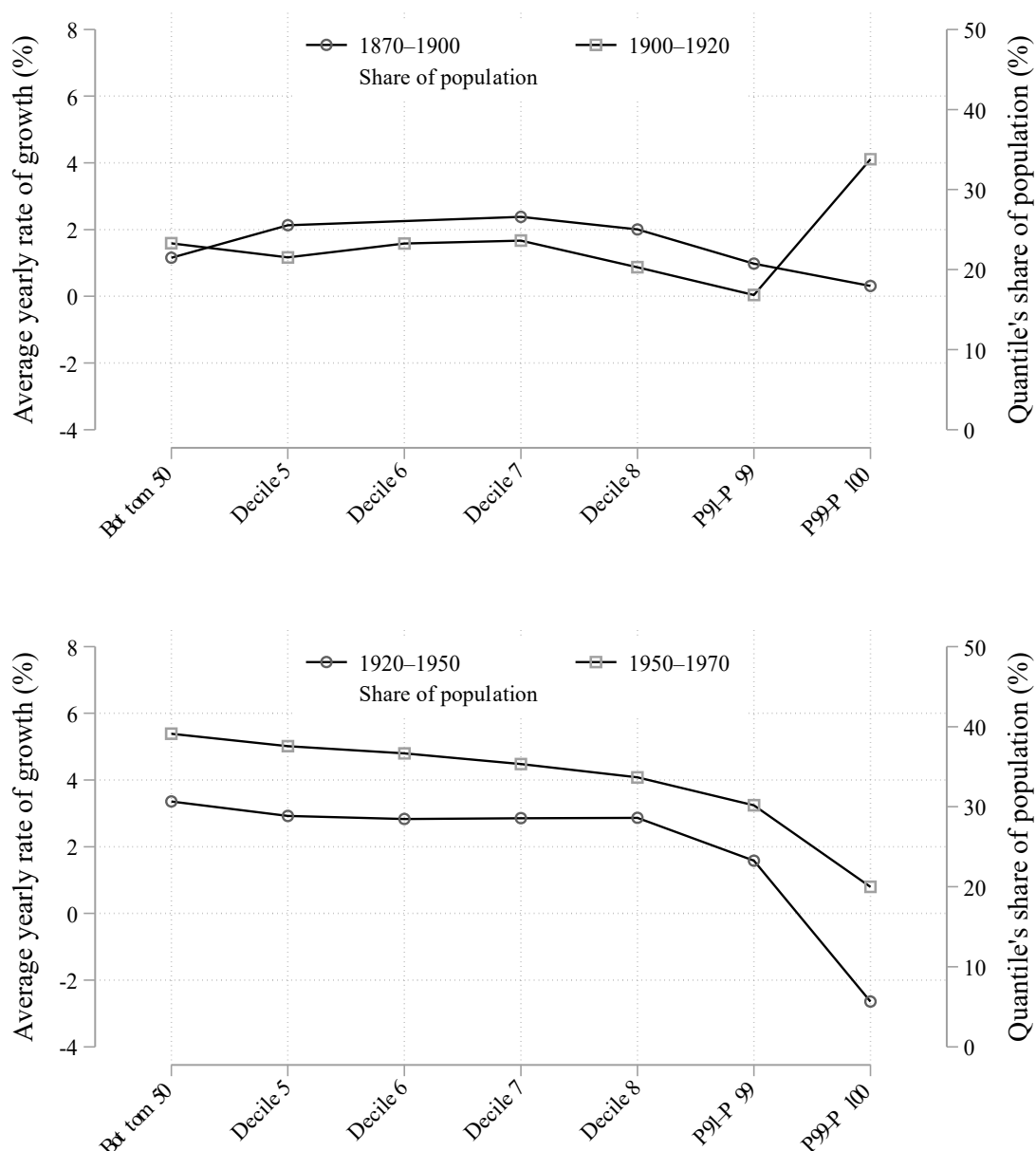
Note: We pooled 1870 and 1880 to get enough observations. The unit of measurement is the individual.

4.3 The upgrading of the jobs structure

We found a significant equalization of the income distribution after 1920. With our individual level data, we were able to examine one important driver: the composition of jobs. Specifically, we decomposed the employment of the working class. This broad group corresponds to social group III in the Swedish three-class taxonomy of the twentieth century

and to the category of "blue-collar workers" in English (*arbetare* in Swedish), as opposed to "white-collar" workers (*tjänstemän* in Swedish).[10]

In Figure 6 we have decomposed the composition of the working class into a much finer classification, building on sector and skill level. The most striking result is that in 1870 the female-dominated sector of domestic work in Stockholm made up about 45 percent of the working class, but in 1950 made up only about 10 percent. Given that this was a low-wage group, the finding has wide implications for income distribution. The decline of inequality occurred over the whole time period, but was especially steep between 1900 and 1920, and between 1920 and 1940. While the main story for women was a shift from low-paying work in domestic service to unskilled jobs in industry and services, for men the principal change was instead from unskilled to skilled work. All of these shifts – decreased domestic service for women, and a shift from unskilled to skilled jobs for men – imply higher wages.

In the contemporary literature on jobs polarization, it is noted that when top incomes grow, these individuals will increase their demand for low-paid, labor-intensive services, such as domestic services (cf. Barany and Siegel 2019). Stockholm in the decades after 1920 experienced the opposite, in a kind of "whipsaw" of inequality: top incomes were curbed by decreasing profits, progressive taxation, and more egalitarian norms (cf. Bengtsson and Molinder 2021), so the top groups could demand fewer domestic servants. At the same time, strong job growth and the rising demands of the young women who were employed as domestic servants pushed up wages in this sector and moved workers to factories and offices. In this way inequality was "whip-sawed", squeezed at both the top and the bottom. This process has been described by social historians: Nordlund Edvinsson and Söderberg (2010) have pointed to a "servant crisis" in Stockholm in the 1930s and 1940s, where it became more and more difficult to recruit young women into jobs with low wages, long working hours and little personal freedom. The upgrading of jobs was of importance also for income inequality.

---

[10] *Arbetare* was a distinct group in Sweden in this time, as they for example had weekly wages as opposed to the monthly salaries of *tjänstemän*, and with fewer privileges such as pensions systems and sickness insurance from their employer. This makes it relevant to aim decomposition specifically at the working class/the group of *arbetare*.

Figure 6. The composition of the working class, 1870–1950
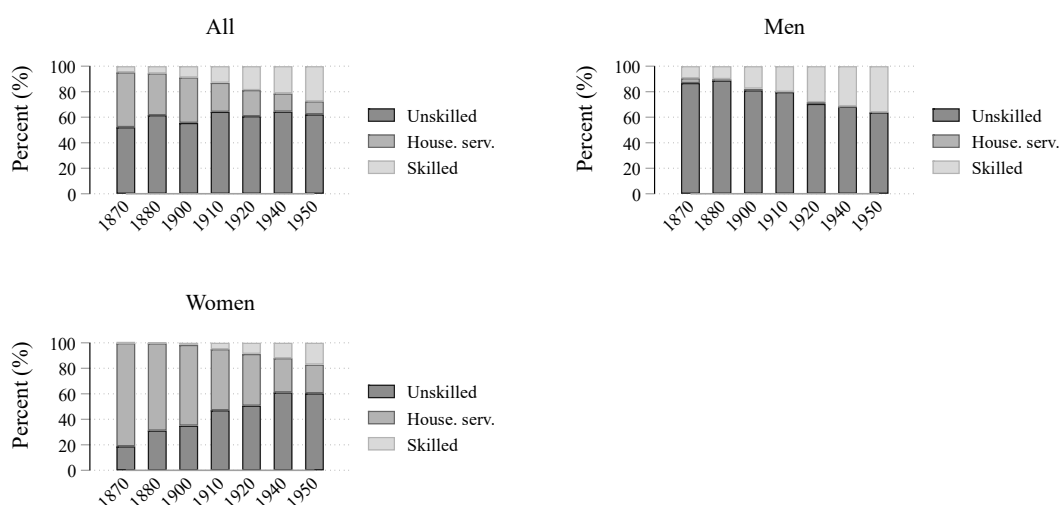
Note. Example of unskilled jobs: "fabriksarbeterska", "städerska" or "affärsbiträde", "byggnadsarbetare", "lagerarbetare". Example of skilled jobs: "montör", "poliskonstapel" or "föreståndarinna", "snickare" or "målare", "spårvagnskusk" or "styrman",

We also want to emphasize that in relation to a previous study which distinguishes wage differentials as an important driver of equalization (Järnek 1971, Olsson 1972, Bengtsson and Prado 2020), the decomposition here points to the fact that, on the aggregate, incomes equalized not only because of reductions in pay differentials but at least as much because of compositional shifts in the work force.

4.4 Decompositions: the role of gender, age, sector, and class in the equalization of incomes

We now consider jointly the roles of gender, age, and class in the equalization of incomes in Sweden. We already know from the jobs decompositions above that gender played an important role in the determinants of one's job and income in this period. Women earned less than men, but the difference diminished over time. Another classic factor in analyses of income distribution is age. Paglin (1979) noted that much of the inequality observed in a cross-section of incomes results from the life cycle pattern of earnings. We earn less when we are young and when we are old than in our prime. This inequality is not necessarily a normative problem. Figure 7 describes in the left-hand graph the life cycle of earnings for people in Stockholm between the ages of 20 and 65 in year 1950, and in the right-hand panel shows the gender income gap by age for years 1900 and 1950.

Figure 7. Life cycle of earnings for women and men ages 20-65

Life-cycle earnings, men and women, 1950

Female income gap by age, 1900 and 1950

Note. In the left-hand panel, the series "sex=0" is men, and "sex=1" is women.

Figure 7 shows that the variation in incomes over the life cycle was remarkably small for women, while men had more of the expected pattern of low income in the early twenties and a peak from 40 to 60 before a decline.

In a final assessment of the role played by gender, age, sector, and class in the equalization of incomes we performed a decomposition of the variance of log income following the methodology developed by Fields (2003) (see also Yun, 2006 and Brewer and Wren-Lewis, 2012. This was done using the Stata command *ineqrbd* (Fiorio and Jenkins, 2010). The share of the variance explained by each factor was estimated taking the covariance of income and the variable capturing the factor and dividing it by the standard deviation of income. The regression equation we used for the decomposition was the following:

$$\log(Income) = \alpha + \beta \, sex + \gamma \, age + \delta \, age^2 + \varepsilon \, sector + \epsilon \, class \qquad (1)$$

where $\log(Income)$ is the natural logarithm of the measure of income including imputations that we used elsewhere in the paper, $\beta \, sex$ is a dummy that take the value 0 if the individual is a man and 1 if it is a woman, $\gamma \, age$ and $\delta \, age^2$ are continuous variables measuring the individual's age while allowing for a non-linear relationship between income and age,

$\varepsilon\ sector$ is a vector of dummy variables measuring the individual's economic sector, and $\epsilon\ class$ is a vector of dummy variables capturing social class.

The decomposition results are shown in Figure 8. We focused on the individual level data and our 1920, 1940, and 1950 benchmarks for which we had a large enough sample and all the information on individuals needed to perform the decomposition. The results show the radical impact of social class on economic inequality; The fall in overall inequality over the 1920 to 1950 period is almost completely explained by the fall in the class component. This goes back to our discussion above of "Social Democratic Fordism": the growing demand for unskilled and skilled workers especially in manufacturing in the 1920s, 1930s and 1940s (Schön 2014), alongside the rapidly growing influence of trade unions in society and over wage setting (Lundh 2002, pp. 116–128, 132–133, 140–194) – probably also, in terms of social norms, over executive pay (cf. Frydman and Molloy 2012). Second in importance is gender, which slightly increased over time despite the decline in the difference in average incomes between men and women. Sector and age played small roles throughout, the impact of the sector – the importance of which for wage differentials has been a very widespread debate in Swedish twentieth century history – remaining roughly constant while age declined in significance.

Figure 8. Decomposition of Inequality
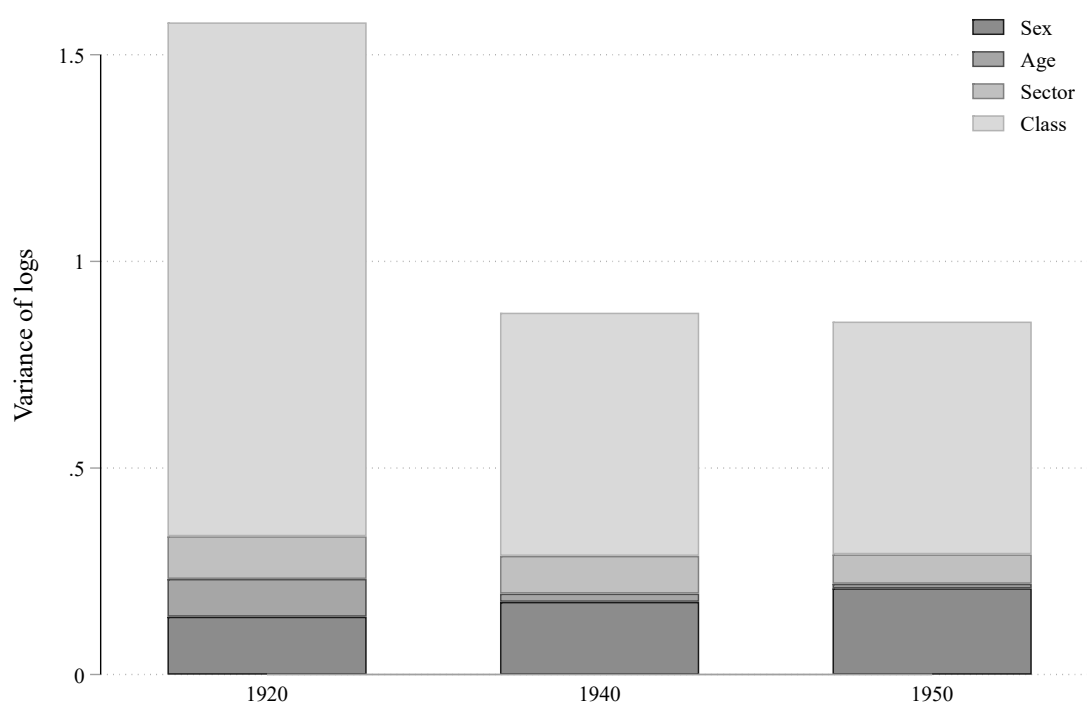
Note: The graph shows the decomposition of the predicted variance of logs for individual income (including imputations). The regression results that were used to perform the decomposition were derived from estimating Equation 1.

5. Conclusions

As argued in the introduction, the contribution of the present paper is four-fold. One, it extends the time frame of Swedish income inequality studies back to 1870 and shows the legacy of high inequality in Sweden before year 1903, where previous long-run income inequality studies started (Roine and Waldenström 2008, 2010). Two, it provides a new in-depth view of the equalization after 1920. We have shown the importance of jobs upgrading for the working class and not least for women, and that while gains of the middle part of the income distribution at the expense of the top 10 percent were important, in Stockholm the bottom half of the distribution also benefited from equalization.

The third contribution is related to the second: the importance of gender differentials for income inequality, and the importance of women's changing patterns of paid employment. In particular, the decline of domestic service and growing importance of more productive blue-collar jobs flattened income differences from the 1920s to the 1960s.

The fourth contribution is the decomposition which has weighted the relative importance of class, sex, sector and age for inequality in 1920, 1940 and 1950. Age and sector were relatively unimportant; gender was important, but class (which, of course, given the

sorting of men and women to jobs of varying quality, also correlated with gender) was the greatest determinant. This also reinforces the emphasis on reducing capital incomes for equalization in previous research by Roine and Waldenström (2008) and Gustafsson and Johansson (2003).

To conclude, then: what does it add up to? What is the overall take-away verdict for the long-run inequality literature from this inspection of income growth and inequality from 1870 to 1970? First, it should be stressed that our study concerns pre-tax, pre-redistribution incomes. A more comprehensive study also considering taxes and transfers, which would have been impossible for this time period with the given data, would most likely find a more comprehensive levelling of incomes, especially from the 1930s on. Previous research has emphasized the growing progressivity of the tax system in the 1930s and 1940s as an important cause of equalization (Roine and Waldenström 2010; Gustavsson, Husz and Söderberg 2009; Gustafsson and Johansson 2003), although transfers and benefits have not been studied. (Here we also differ markedly from the DINA literature with its more encompassing ambition vis-à-vis types of income.)

More substantively, we may note that in our discussion we have emphasized the role of the jobs composition of the working class and the shifting jobs structure over time. This is very much in line with previous studies using micro data (Järnek (1971) on Malmö and Olsson (1972) on Gothenburg), and in contrast to the studies using tabulated data and focusing on the top income earners, which have given more weight to the fall in capital incomes (Roine and Waldenström 2008; Gustavsson, Husz and Söderberg 2009, pp. 100–102 for Stockholm; Gustafsson and Johansson 2003, p. 205 for Gothenburg). Does this simply show that your conclusions are driven by the type of data you look at? Surely to some degree. (But see also Bengtsson and Molinder 2021 which studies top income earners with micro data.) The top income studies have found an important role for capital incomes, since they were indeed an important share of incomes for the top groups. Of course, before 1920 our income data lacked dividends, so it was more difficult to discuss the incomes of the very top. Conversely, we focused especially on the lower income earners, since these are the people who before were missing from the top incomes emphasis, and we drew attention to the factors which affect the incomes of ordinary workers such as maids and industrial workers. Given the originality of the data we have found it worthwhile to spend much effort on a comprehensive exploration of the data, mostly of a descriptive nature. A promising way forward, for a further paper or for a further version of this study, would be to work more on integrating the concerns

of the top incomes studies (taxes and capital incomes) with the matters discussed in more depth here.

In the introduction, we noted the parallelism of our ambitions to those in the Distributional National Accounts literature, regarding the comprehensive view of the income distribution. How do our findings, then, relate to the findings of the work on France by Garbinti, Goupille-Lebret and Piketty (2018), the US by Fisher-Post (2020), and Norway by Aaberge, Atkinson and Modalsli (2020)? The comparison is complicated by the fact that we have studied a city and the other writers have studied countries, but a tentative discussion may be had. Fisher-Post finds impoverishment of the bottom groups during the Great Depression and no secular growth of the bottom half's share of incomes until after the Depression. The level is higher from the late 1930s to the mid-1950s than it was before. We in contrast find lower inequality in 1930 than in 1920, and a quite stable rate of inequality decline for every benchmark after 1920. This fits with the emphasis in Swedish economic history literature on the relatively mild economic crisis of the early 1930s (Schön 2014, pp. 304–313), and with the argument of Bengtsson (2019) of very high inequality before democratization in 1918–1921, a contrast with the new democratic order. As regards the US, Fisher-Post points to a World War Two-period bump in the demand for labor as one explanation for equalization, but for the 1945 to 1975 period points only to Lindert and Williamson's (2016) list of possible explanations: demography, trade policy, financial sector retrenchment, technological change favoring unskilled workers, and growth in human capital.

Aaberge, Atkinson and Modalsli (2020, p. 10) similarly find in Norway that, while the top percentile share of incomes fell secularly from the 1910s to the 1970s, the Gini among the bottom 99 per cent actually grew until the 1940s. They argue that the evolution of inequality in Norway is best characterized "as a series of episodes identified with sub-periods" (p. 11), pointing to a financial crisis of 1899, a profits boom during World War I, and the large-scale introduction of collective bargaining in the 1930s as important inequality-changing events and processes. Garbinti, Goupille-Lebret and Piketty (2018, p. 71) find that income growth in France 1900 to 1950 was remarkably equal, with negative growth for the top 1 percent and the best growth per year (about 2.0 percent per year) for the bottom 25 percentiles and around percentiles 50 to 70. They find that a fall of top capital incomes is crucial in explaining the decline in inequality. Taken together, it seems that the literature is still in an inductive state for explaining the equalization after World War I: capital incomes

mattered, but beyond this, many different factors and events are proposed.[11] We would argue that even though this paper studies only a city, not a country as a whole, the micro data investigation suggests that the change in jobs composition, and especially the shift out of domestic service and into better paid jobs, mattered immensely. The tandem of changing wage differentials between various jobs, and the shift in the share of different jobs as shares of the labor market, would be a promising way forward for further micro data studies of the first half of the twentieth century.

---

[11] In recent studies building on the top incomes approach (Bartels 2019 for Germany) or social tables (Gómez León and De Jong 2019 for Germany and Britain), collective bargaining, unionization, and the social effects of World War I, as mediated by political institutions and power, seem to matter a good deal for inequality in this period.

References

Aaberge, Rolf, Anthony B. Atkinson and Jørgen Modalsli. 2020. "Estimating long-run income inequality from mixed tabular data: Empirical evidence from Norway, 1875–2017". *Journal of Public Economics* vol. 187, Article 104196.

Atkinson, Anthony B. 2007. "Measuring Top Incomes: Methodological Issues", pp. 18-42 in Atkinson and Thomas Piketty (eds.) *Top Incomes over the 20th Century: A Contrast Between European and English-Speaking Countries*. Oxford: Oxford University Press.

Atkinson, Anthony B. and Thomas Piketty. 2007. "Towards a Unified Data Set on Top Incomes", in Atkinson and Piketty (eds.) *Top Incomes over the Twentieth Century: A Contrast Between European and English-Speaking Countries*. Oxford University Press, pp. 531–565.

Auten, Gerald and David Splinter. 2019. "Income Inequality in the United States: Using Tax Data to Measure Long-term Trends". Unpublished paper, U.S. Treasury Department/U.S. Congress. Available from http://davidsplinter.com , dated December 20, 2019.

Bárány Zsófia L. and Christian Siegel. 2018. "Job Polarization and Structural Change". *American Economic Journal: Macroeconomics*, Vol. 10, No. 1 pp. 57-89

Bartels, Charlotte. 2019. "Top Incomes in Germany, 1871–2014". *Journal of Economic History* 79 (3).

Bengtsson, Erik. 2019. "The Swedish Sonderweg in Question: Democratization and Inequality in Comparative Perspective, c. 1750–1920". *Past & Present*, Vol. 244, No. 1, pp. 123–161.

Bengtsson, Erik and Jakob Molinder. 2017. "The Economic Effects of the 1920 Eight-Hour Working Day Reform in Sweden". *Scandinavian Economic History Review*, vol. 65 no. 2, pp. 149–168.

Bengtsson, Erik and Jakob Molinder. 2021. "What Happened to the Incomes of the Rich during the Great Levelling? Evidence from Swedish Individual-level Data, 1909–1950". Lund Papers in Economic History No. 230, October 2021.

Bengtsson, Erik, Anna Missiaia, Mats Olsson and Patrick Svensson. 2018. "Wealth Inequality in Sweden, 1750–1900". *The Economic History Review*, vol. 71 no. 3, pp. 772–794.

Bentzel, Ragnar. 1952. *Inkomstfördelningen i Sverige*. Stockholm: IUI.

Berglund, Bengt. 1982. *Industriarbetarklassens formering: arbete och teknisk förändring vid tre svenska fabriker under 1800-talet*. Dissertation in Economic History, Gothenburg University.

Bernhardt, Eva. 1972. "Fertility and economic status — some recent findings on differentials in Sweden". *Population Studies* 26(2): 175–184.

Björklund, Anders, and Mårten Palme. 2000. "The Evolution of Income Inequality During the Rise of the Swedish Welfare State 1951 to 1973". *Nordic Journal of Political Economy* 26: 115–128.

Bodin, Otto and Herman Palmgren. 1902. *Självdeklaration: Praktisk handbok för skattskyldige och beskattningsmyndigheter*. Stockholm: P.A. Norstedt & söner.

Cassel, Gustav. 1906. "Utvecklingslinjer i svensk skattelagstiftning". *Ekonomisk Tidskrift*, Årg. 8 (1906), pp. 191-216.

Edin, Karl Arvid and Edward P. Hutchinson. 1936. *Studies of Differential Fertility in Sweden*. London: P.S. King and Son.

Edvinsson, Rodney and Therese Nordlund Edvinsson. 2017. "Explaining the Swedish 'housewife era' of 1930–1970: joint utility maximisation or renewed patriarchy?". *Scandinavian Economic History Review* 65 (2): 169–188.

Enflo, Kerstin and Joan Ramon Roses. 2015. "Coping with regional inequality in Sweden: structural change, migrations, and policy, 1860–2000". *Economic History Review* 68: 191-217.

Fisher-Post, Matthew. 2020. "Examining the Great Leveling: New Evidence on Midcentury American Inequality". WID.world Working Paper No 2020/01. Paris: World Inequality Database.

Frydman Carola and Raven Molloy. 2012. "Pay Cuts for the Boss: Executive Compensation in the 1940s". *Journal of Economic History* 72 (1).

Galbraith, James K. 2019. "Sparse, Inconsistent and Unreliable: Tax Records and the World Inequality Report 2018". *Development and Change* 50(2): 329–346.

Garbinti, Bertrand, Jonathan Goupille-Lebret, and Thomas Piketty. 2018. "Income inequality in France, 1900–2014: Evidence from Distributional National Accounts (DINA)". *Journal of Public Economics* 162: 63–77.

Glete, Jan. 1987. *Ägande och industriell omvandling: Ägargrupper, skogsindustri, verkstadsindustri 1850–1950*. Stockholm: SNS.

Gómez León, María and Herman J. De Jong. 2019. "Inequality in turbulent times: income distribution in Germany and Britain, 1900–50". *Economic History Review* 72 (3): 1073–1098.

Griffin, Emma. 2014. *Liberty's Dawn: A People's History of the Industrial Revolution*. Paperback edition. New Haven, CT: Yale University Press.

Gustafsson, Björn and Mats Johansson. 2003. "Steps toward Equality: How and Why Income Inequality in urban Sweden changed during the period 1925-1958". *European Review of Economic History*, 7, 191–211.

Gustavsson, Martin, Orsi Husz, and Johan Söderberg. 2009. "Collapse of a Bourgeoisie? The Wealthy in Stockholm, 1915–1965", *Scandinavian Economic History Review*, 57 (1), 88–105.

Hammarström, Ingrid. 1970. *Stockholm i svensk ekonomi 1850–1914*. Stockholm: Stockholms Stadsarkiv.

Harvey, Laura A., Jochen O. Mierau and James Rockey. 2017. "Inequality in an Equal Society". Luxembourg Income Study Working Paper # 26.

Henrekson, Magnus and Mikael Stenkula. 2015. "Swedish Taxation Since 1862: An Overview". IFN Working Paper No. 1052. Stockholm: Research Institute of Industrial Economics.

Holmberg, Kajsa. 2013. *Behind the counter: Female inroads, obstacles and careers in the Swedish commercial bank sector, 1885-1937*. Dissertation in Economic History, Lund University.

Järnek, Martin. 1971. *Studier i hushållens inkomsterförhållanden 1925–1964*. Lund: Skrifter utgivna av Ekonomisk Historiska föreningen i Lund.

Larrimore, Jeff, Jacob Mortenson and David Splinter. 2021. "Household Incomes in Tax Data: Using Addresses to Move from Tax-Unit to Household Income Distributions". *Journal of Human Resources* 56 (2): 600–631.

Lindert, Peter H. and Jeffrey G. Williamson. 2016. *Unequal Gains: American Growth and Inequality since 1700.* Princeton, NJ: Princeton University Press.

Lodin, Sven-Olof, Gustaf Lindencrona, Peter Melz, Christer Silfverberg and Teresa Simon-Alvendal. 2017. *Inkomstskatt: en läro- och handbok i skatterätt*. Del 1. 16th edition. Lund: Studentlitteratur.

Lundh, Christer. 2002. *Spelets regler: Institutioner och lönebildning på den svenska arbetsmarknaden 1850–2000*. Stockholm: SNS.

Malmer, Håkan. 2003. "Granskningen av inkomstdeklarationer under 100 år", pp. 21–58 in Deklarationen 100 år och andra tillbakablickar. Stockholm: Riksskatteverket.

Matović, Margareta R. 1984. *Stockholmsäktenskap: familjebildning och partnerval i Stockholm 1850–1890*. Doctoral dissertation in History, Stockholm University.

Molinder, Jakob. 2022. "Historical Roots of the Dual-earner Model: Women's Labor Force Participation in Sweden, 1870–1960". Unpublished paper, Uppsala University.

Molinder, Jakob and Johan Ericsson. 2020. "Economic Growth and the Development of Real Wages: Swedish Construction Workers' Wages in Comparative Perspective, 1831–1900", *The Journal of Economic History,* 80(3): 813–852

Molitoris, Joseph. 2017. "Disparities in Death: Inequality in Cause-specific Infant and Child Mortality in Stockholm, 1878–1926." *Demographic Research* 36: 455−500.

Molitoris, Joseph and Martin Dribe. 2016. "Ready to Stop: Socioeconomic Status and the Fertility Transition in Stockholm, 1878–1926." *Economic History Review* 69(2): 679–704.

Nistotskaya, Marina and Michelle D'Arcy. 2018. "Getting to Sweden: The Origins of High Compliance in the Swedish Tax State", pp. 33–55 in Sven Steinmo (ed.) *The Leap of Faith: The Fiscal Foundations of Successful Government in Europe and America*. Oxford: Oxford University Press.

Nordlund Edvinsson, Therese and Johan Söderberg. 2010. "Servants and bourgeois life in urban Sweden in the early 20th century". *Scandinavian Journal of History* 35 (4): 427–450.

Nyberg, Anita. 1989. *Tekniken – kvinnornas befriare? Hushållsteknik, köpevaror, gifta kvinnors hushållsarbetstid och förvärvsdeltagande 1930-talet – 1980-talet*. Doctoral dissertation, Linköping University.

Olsson, Kent. 1972. *Hushållsinkomst, inkomstfördelning och försörjningsbörda – en undersökning av vissa yrkesgrupper i Göteborg 1919–1960.* Doctoral thesis in Economic History, Gothenburg University.

Paglin, Morton. 1975. "The Measurement and Trend of Inequality: A Basic Revision". *American Economic Review*, 65, 598–609.

Piketty, Thomas. 2003. "Income Inequality in France, 1901–1998". *Journal of Political Economy*, vol. 111, no. 5, pp. 1004-1042.

Piketty, Thomas. 2014. *Capital in the Twenty-First Century*. Cambridge, Mass.: Harvard University Press.

Piketty, Thomas, Gilles Postel-Vinay and Jean-Laurent Rosenthal. 2014. "Inherited vs self-made wealth: Theory and evidence from a rentier society (Paris 1872–1927)". *Explorations in Economic History* 51: 21-40.

Piketty, Thomas, Emmanuel Saez and Gabriel Zucman. 2018. "Distributional National Accounts: Methods and Estimates for the United States". *Quarterly Journal of Economics* 133 (2): 553–609.

Roikonen, Petri and Sakari Heikkinen. 2020. "A Kuznets rise and a Piketty fall: income inequality in Finland, 1865–1934". *European Review of Economic History* 24: 46-79.

Roine, Jesper and Daniel Waldenström. 2008. "The Evolution of Top Incomes in an Egalitarian Society: Sweden, 1903–2004". *Journal of Public Economics*, 92 (1-2), 366–387.

Roine, Jesper and Daniel Waldenström. 2010. "Top Incomes in Sweden over the Twentieth Century", in Anthony B. Atkinson and Thomas Piketty (eds.), *Top Incomes in Global Perspective*, Oxford: Oxford University Press, pp. 299–370.

SCB. 1969. *Historisk statistik för Sverige Del 1. Befolkning 1720–1967*. Andra upplagan. Stockholm: SCB.

Schön, Lennart. 2014. *En modern svensk ekonomisk historia*. Fourth edition. Lund: Studentlitteratur.

Söderberg, Johan, Ulf Jonsson, and Christer Persson. 1991. *A Stagnating Metropolis: The Economy and Demography of Stockholm, 1750–1850*. Cambridge: Cambridge University Press.

Splinter, David. 2019. "Who pays no tax? The declining fraction paying income taxes and increasing tax progressivity". *Contemporary Economic Policy* 37 (3), 413–426.

Stockholms Stads Statistiska Kontor. 1922. *Statistisk Årsbok för Stockholm 1922*. Stockholm.

Stockholms Stads Statistiska Kontor. 1951. *Statistisk Årsbok för Stockholms Stad vol. 82, 1951*. Stockholm.

Svensson, Anton and Erik Bengtsson. 2021. "Income Inequality in an Industrial City during the Great Leveling: Individual-Level Evidence from Malmö, 1905–1950". Unpublished paper, Lund University.

Svensson, Lars. 2003. "Explaining Equalization: Political Institutions, Market Forces, and Reduction of the Gender Wage Gap in Sweden, 1920-95". *Social Science History* 27 (3): 371–395.

Thomas, Dorothy Swaine. 1937. Book review of Edin and Hutchinson's *Studies of Differential Fertility in Sweden*, in *American Sociological Review* 2 (1): 126–128.

Widell, Ludvig. 1907. "Några ord om en svensk finans- och inkomststatistik". *Ekonomisk Tidskrift*, Årg. 9, pp. 323-338

Appendix A. Additional graphs and tables, and robustness checks

Figure A1. Gini coefficient with 95 % confidence interval 1870–1970, households
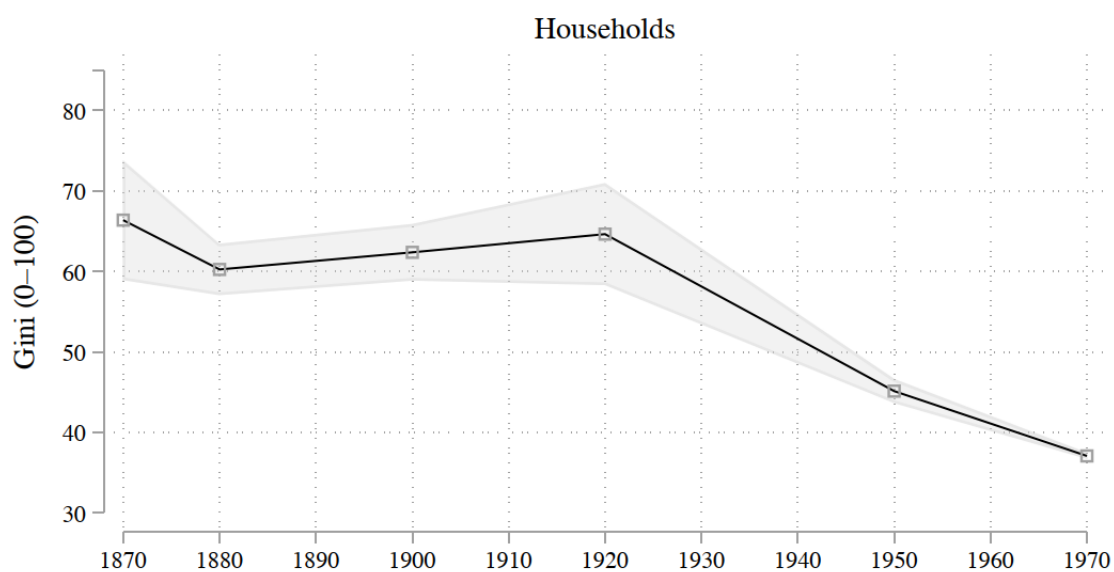
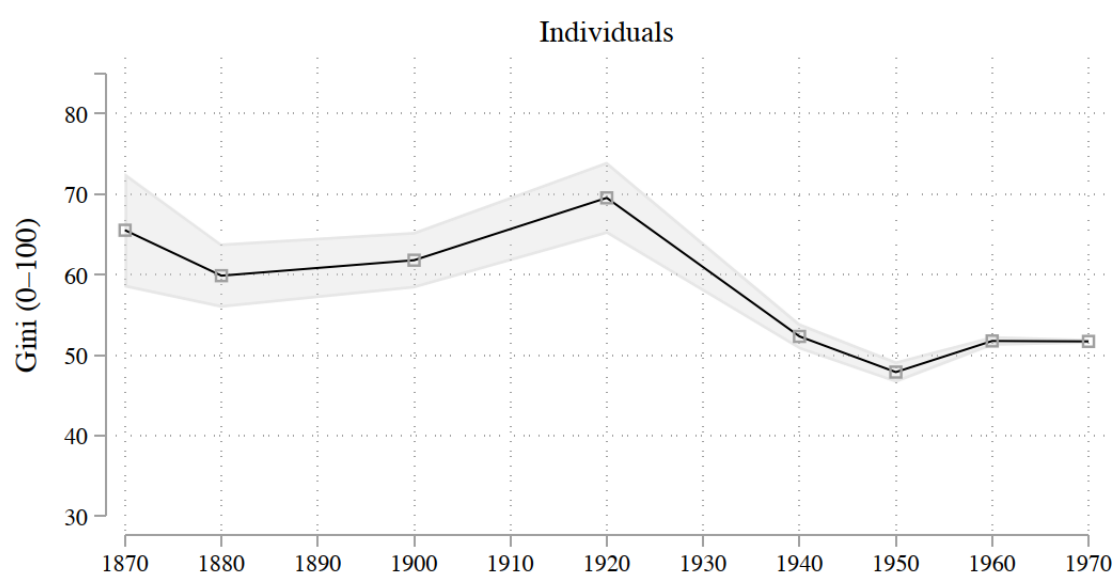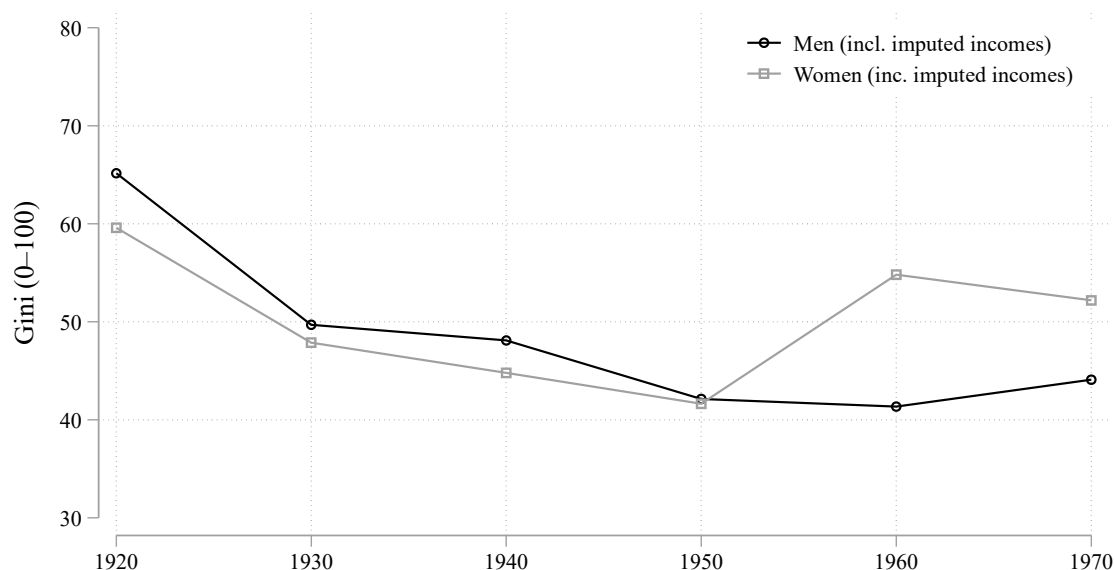*Figure A2. Gini coefficient with 95 % confidence interval 1870–1970, individuals*

Figure A3. Gini coefficient for men and women, 1920–1970, individuals

A robustness check for the inclusion of stock dividends

As discussed in the Data section of the paper, the *bevillning* imposed in 1862 did not cover stock dividends, because these were considered to be taxed at the corporation level. From 1920 on dividends are included. This means that in our calculations for 1870, 1880, 1900 and 1910, we have a downward bias of incomes, especially top incomes (since dividends are unequally distributed in the population). How big a problem is this?

From 1848 to 1897, one needed a special permit from the King to form a new stock corporation (Broberg 2006). This limited the number of corporations in the country, but after deregulation in 1897, the number took off. For 1870 and 1880, then, the bias in our estimates resulting from the exclusion of dividends ought be limited. Broberg (2006) dates the breakthrough of the joint stock companies in the Swedish economy to the first two decades of the twentieth century. According to Flodström (1911, p. 143), in 1881 there were 1,264 corporations in Sweden with a combined capital of 455 million; by 1908 the numbers had increased to 4,919 corporations and 2 034 million.

Waldenström (2014, data appendix, Table II.A6.1 Market data of the Stockholm Stock Exchange, 1863–, and Table II.A6.2 Stock price and return indices, 1870–) presents comprehensive data on stock market capitalization in Sweden in the relevant period. Using his data on the market capitalization (from Table II.A6.1) and dividend yields (from Table II.A6.2), we calculate total dividends in Sweden for 1871 (when the dividend yields data begin), 1880, 1900, and 1910. As expected, the 1871 and 1880 sums are relatively small: 3.8 million SEK in the country as a whole in 1871, and 7.6 million in 1880. Still in 1900 the sum

is small, 6.2 million, while in 1910 dividends have become much more important in the Swedish economy: 82.4 million.[12]

There is no information on the distribution of dividends across the country. Therefore, we use the distribution of taxed capital incomes in the Bengtsson-Molinder-Prado (2022) countrywide dataset to proxy the share of dividends which accrued to Stockholm residents. Stockholm residents had 17 percent of Sweden's taxed capital incomes in 1900, and 29 percent in 1920. Unfortunately, here 1910 is lacking, but we calculate Stockholm's share of capital incomes in 1910 as a linear interpolation between 1900 and 1920, i.e. 23 percent. We thus use these shares to allocate dividends.

In our baseline estimates, the top decile got 56.5 percent of incomes in Stockholm in 1900 (top percentile: 21.9 percent), or 140 million of 247 million SEK and 48.8 percent in 1910 (top percentile: 19.5 percent), or 166 million of 339 million SEK.

We allocate the 17 percent of 6.2 million SEK dividends in 1900 to Stockholm residents, assuming that half of the dividends accrued to the top percentile, and half to the 90th to 98th percentiles. In 1900, the top decile's share of incomes grows from 56.5 to 56.7, and the top percentile's share from 21.9 to 22.0 percent. In 1910, when dividends have become more significant, the underestimation of inequality in our baseline estimates is bigger: the top decile's share grows from 48.8 to 51.5 percent, and the top percentile's share from 19.5 to 21.1 percent. Given the small size of dividends in the 1870s and 1880s, the underestimation for 1870 and 1880 can be assumed to be small.

Thus, the downward bias for our top income shares in 1870, 1880 and 1900 is very small, while it is larger in 1910: around 2-3 percentage points. The key take-away point is that inequality was even higher in 1870 to 1920 than we estimate, and that the reduction in inequality after 1920 was even steeper.

References Appendix A

Bengtsson, Erik, Jakob Molinder and Svante Prado. 2022. "Swedish Income Inequality 1870–1970: A Micro Data Approach". Work in progress, Lund University/Uppsala University/University of Gothenburg.

Broberg, Oskar. 2006. *Konsten att skapa pengar: Aktiebolagens genombrott och finansiell modernisering kring sekelskiftet 1900*. Doctoral dissertation in Economic History, Göteborg University.

---

[12] For previous studies of dividend yields in the early 1900s, see especially Flodström (1911) and Heckscher's (1911) critique of Flodström. Unfortunately, they only study the years 1906–1908, which of course are not included in our study, but they form an interesting background to Waldenström's investigation. See also Gernandt, Palm and Waldenström (2012).

Flodström., Isidor. 1911. "Inländska aktiebolag och solidariska bankbolag. Kapital, vinst och utdelning för åren 1906, 1907 och 1908, under samma år erlagda utskylder till stat och kommun samt reserv- och övriga egna fonder vid 1908 års slut." Published in the series Finansstatistiska Utredningar utgifna genom Kungl. Finansdepartementet, volume IV. Stockholm: K.L. Beckmans Boktryckeri.

Gernandt, Otto, Thomas Palm and Daniel Waldenström. 2012. "Liquidity premia during the industrial breakthrough: evidence from the Stockholm Stock Exchange, 1901–1919". *European Review of Economic History* 16, 247–269.

Heckscher, Eli F. 1911. "Aktiebolagens räntabilitet". *Ekonomisk Tidskrift*, Årg. 13, häfte 3, pp. 97–112

Waldenström, Daniel. 2014. "Swedish stock and bond returns, 1856–2012", in Rodney Edvinsson, Tor Jacobson and Daniel Waldenström (eds.) *Historical Monetary and Financial Statistics for Sweden, Volume II: House Prices, Stock Returns, National Accounts, and the Riksbank Balance Sheet, 1620–2012*. Stockholm: Sveriges Riksbank and Ekerlids. The data is available from https://sites.google.com/view/danielwaldenstrom/data-programs (accessed 23 June, 2022).

Appendix B. Further information about the data and sample

Stockholm 1870–1970

The city of Stockholm is centered around the "City Island" (*Stadsholmen*), known more colloquially as the "Old town". Since the city's expansion in the sixteenth century, the peninsulas north and south of the *Stadsholmen* have also been populated., as well as a few surrounding islands of which *Kungsholmen* is the most populous (see also Map A1.) When our study begins in 1870, a population of roughly 136 thousand people were living in the Inner city. During our study period, the city expanded administratively three times. In 1913, the parish of Brännkyrka, south of the city, was incorporated, and in 1916 the same happened with Bromma, west of the city. In 1930 parts of Nacka, east of the city, were incorporated, and in 1949 Spånga parish farther south. Table A1 shows the population of the city as it was, but also for Brännkyrka, Bromma and Spånga, the city's incorporations, 1870 to 1950.

Table B1. Population in Stockholm, 1870–1950

|	Inner city	Brännkyrka	Bromma	Spånga	Stockholm, following 1950 borders
1870	136 016	2 038	740	1 418	140 212
1880	168 775	4 478	871	1 669	175 793
1890	246 454	5 666	1 201	1 662	254 983
1900	300 624	8 402	1 834	2 352	313 212
1910	342 323	18 379	4 684	8 378	373 764
1920	375 528	34 016	9 896	9 840	429 280
1930	427 647	49 876	24 690	13 013	515 226
1940	458 683	72 169	59 651	14 065	604 568
1950	427 019	217 916	85 071	15 930	745 936

Note. Source is *Statistisk årsbok för Stockholms stad 1952*, Table 10. Brännkyrka in 1940 and 1950 includes Enskede parish; the parish was divided in two in 1931.

The 1970s saw further expansion of the city limits, but this falls outside of our study period.[13]

To further break down the population development of the city, Figure B1 shows the total population growth on a yearly basis, plus a decomposition into (a) a birth/death surplus (the dashed line), and (b) an in-moving/out-moving surplus (the dotted line). As we can see, population growth was overall above all driven by in-migration.

---

[13] Ericson Wolke (2016, p. 30) in his history of Stockholm argues that there have been three major expansion phases in the history of the city since the founding in the 1200s: the 1600s when Stockholm became the capital, the industrialization period of 1870 to 1914, and the post-1945 period.

Figure B1. Population changes in Stockholm, 1870–1970
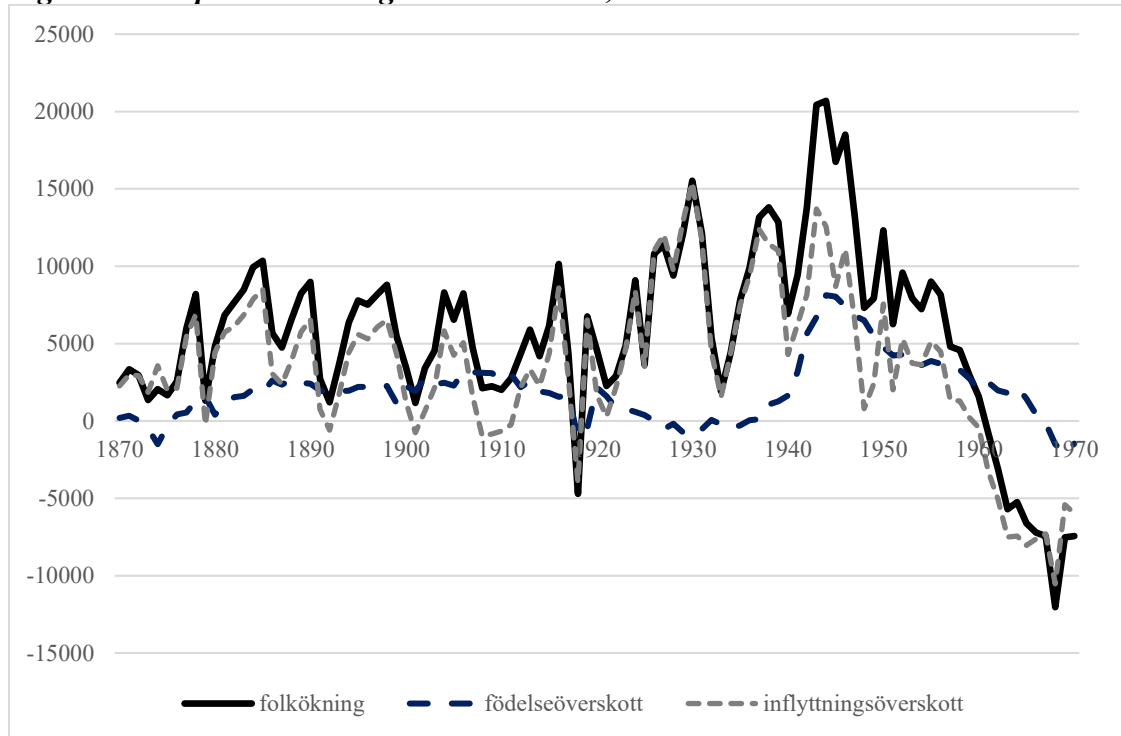
Note. The numbers come from *Statistisk Årsbok för Stockholm*, year 1918, Table 28; year 1941, Table 37; year 1952, Table 35; year 1961, Table 37; and year 1971, Table 35.

The sample

Between 1878 and 1926, Stockholm was administratively organized in *rotar*, each designed to encompass about 10,000 people. The *rotar* followed the borders of the parishes (*församlingar*) and were either identical with a parish, or more often, a part ("lower", "upper", etc.) of a parish.

In 1870 and 1950, our sampling builds on the parishes – before and after the 1878-1926 rote period, this is how the taxation lists were organized. Table B2 gives the contents of our sample.

Table B2. The composition of the sample

year	sample	Sample N	Parishes / *rotar*
1870	1 %	2941	Klara upper (*övre*), Klara lower (*nedre*), Jakob, Ladugårdslandet upper, Ladugårdslandet lower
1880	1 %	2510	rote 1 (Storkyrko), 2 (Clara), 9 (Hedvig Eleonora), 10 (Hedvig Eleonora) och 16 (Maria).
1900	1 %	5385	rote 2 (Clara), 7 (Jacob o Johannes), 9 (Hedvig Eleonora), 10(Hedvig Eleonora), 15 (Maria), 18 (Adolf Fredrik), 19 (Maria), 24 (Hedvig Eleonora).
1910	1 %	5797	rote 6 (Observatorieroten, Gustav Vasa parish), 7 (Jakobroten), 12 (Danviksroten, Katarina parish), 13 (Stadsgårdsroten, Katarina parish), 15 (Maria kyrkorote), 20 (Narvaroten, Oscar parish), 23 (Karlsviksroten, Kungsholm parish), 25 (Barnängsroten, Katarina parish), 26 (Norrtullsroten, Matteus parish), 29 (Kristineroten, Kungsholm parish)
1920	1 %	6581	rote 2 (Clara), 7 (Jacob), 9 (Hedvig Eleonora), 14 (Katarina), 16 (Maria-Högalid), 17 (Kungsholm), 23(Kungsholm-S:t Göran), 24 (Oscar), 30 (Matteus), 31 (Maria), 35 (Bromma), 36 (Katarina)
1940	0.5 %	9759	Gustav Vasa, Johannes, Matteus, Oscar, S:t Göran
1950	0.5 %	9454	Gustav Vasa, Johannes, Oscar, S:t Göran

Note. Sample % = percentage of taxpayers sampled. Some of the parishes – Klara, Hedvig Eleonora, Maria, Jakob, and others – exist throughout the period but there were some reforms in 1906-1907. In 1906 Gustav Vasa and Matteus parishes were created out of the Adolf Fredrik parish. Johannes was created in 1907 as a splinter of Jacob och Johannes. Oscar was created in 1906 as a splinter of Hedvig Eleonora.

Map B1. Stockholm's rotar in 1900
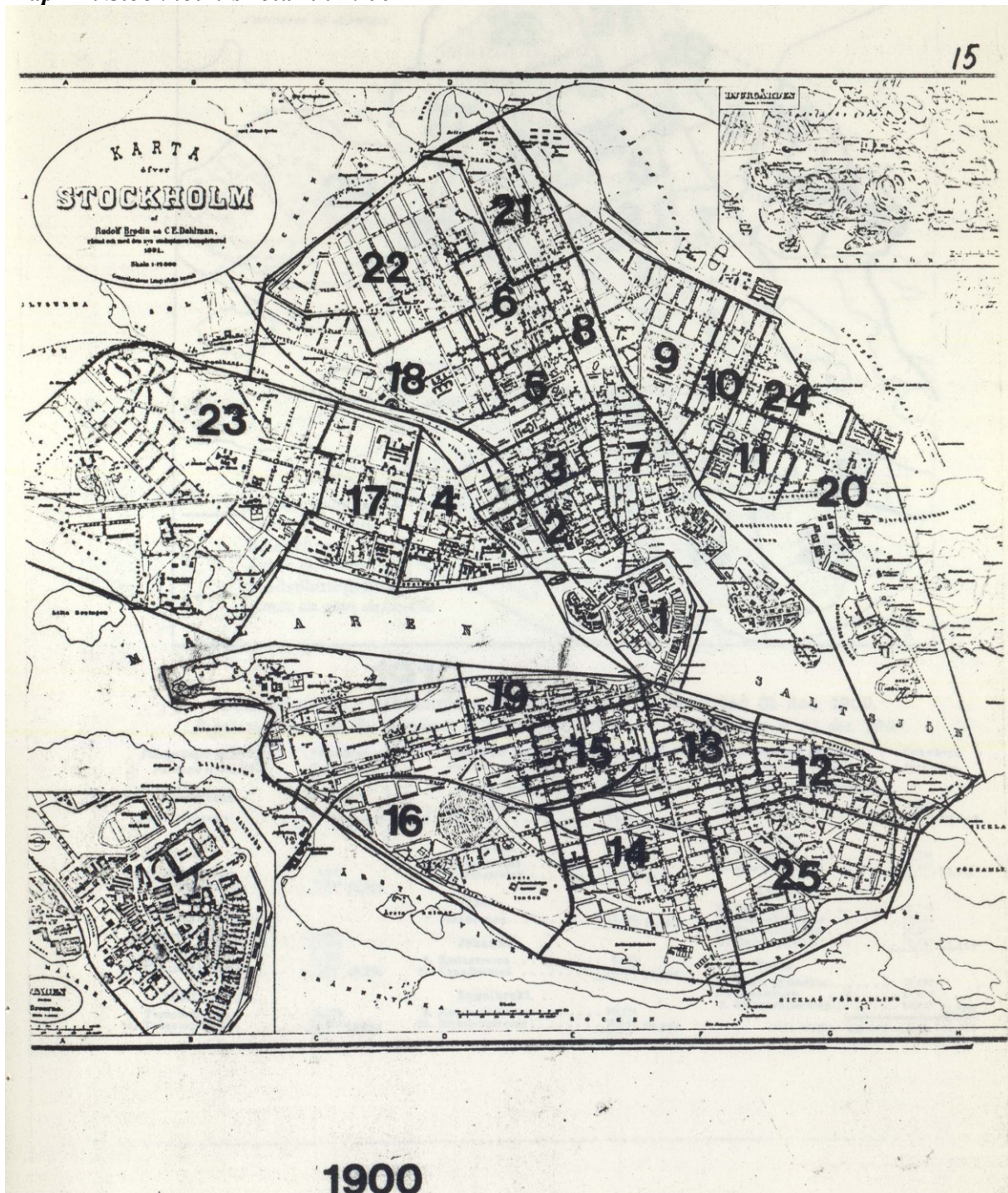
Note. The information on parishes and rotar and the map is taken from the Stockholmskällan website. The map is created by unknown and the copyright is CC-BY. The map is held by the Stockholm City Archive, its object ID is *Stockholms stadsarkiv Stockholms historiska databas/Rotekarta 1900*, and it is available from https://stockholmskallan.stockholm.se/post/18597

The representativeness of the sample: a comparison with data for the city as a whole

To ascertain that our random sample is representative, we compare it with data for the city's population from the Stockholm city statistical bureau. The statistics bureau started to publish yearbooks in 1904, and the first book also included data for the year 1900. This allows the first comparison between our sample and the city as a whole to apply to 1900.

For 1900, the Statistics bureau presented the income distribution over 34 income groups (plus the implicit group which was quite large at the time: those who did not earn enough to pay any income tax at all). For 1920, it is presented with 43 categories, and for 1950, with 20 categories. It is very helpful that the tabulated distributions in 1900 and 1920 are so detailed, since our own data are less comprehensive for these dates (especially 1900). The tabulations allow us to make robustness checks for our inequality estimates.

To ascertain that the sample is reliable, we can compare our sample properties with those of the (tabulated) population on several scores. Figure B2 below compares our sample with the population on the share of the adult population who paid tax. As we see, the measures are very much in agreement for all the years but 1950. In 1900, in our sample 35 percent of adults were taxpayers; according to the tabulated statistics, the share was 33.9 percent. In 1920, the difference was even smaller: the share in our sample was 67 percent, and in the tabulated population 67.6 percent. In 1940, the share in our sample was 71 percent, and in the tabulated population 72.6 percent. It is only in 1950 that our measures differ. The root of the divergence is that children above 18 years old living with their parents were not included in the taxation lists in 1950 (though they were included in previous years) unless they paid tax. This also explains why the 1950 sample is slightly smaller than the 1940 sample. Taken together, the estimates in Figure B2 suggest that our sample is representative of the population of adults in Stockholm.

Figure B2. Share taxpayers (population >15 years old) in Stockholm 1900-1950: sample versus tabulated data
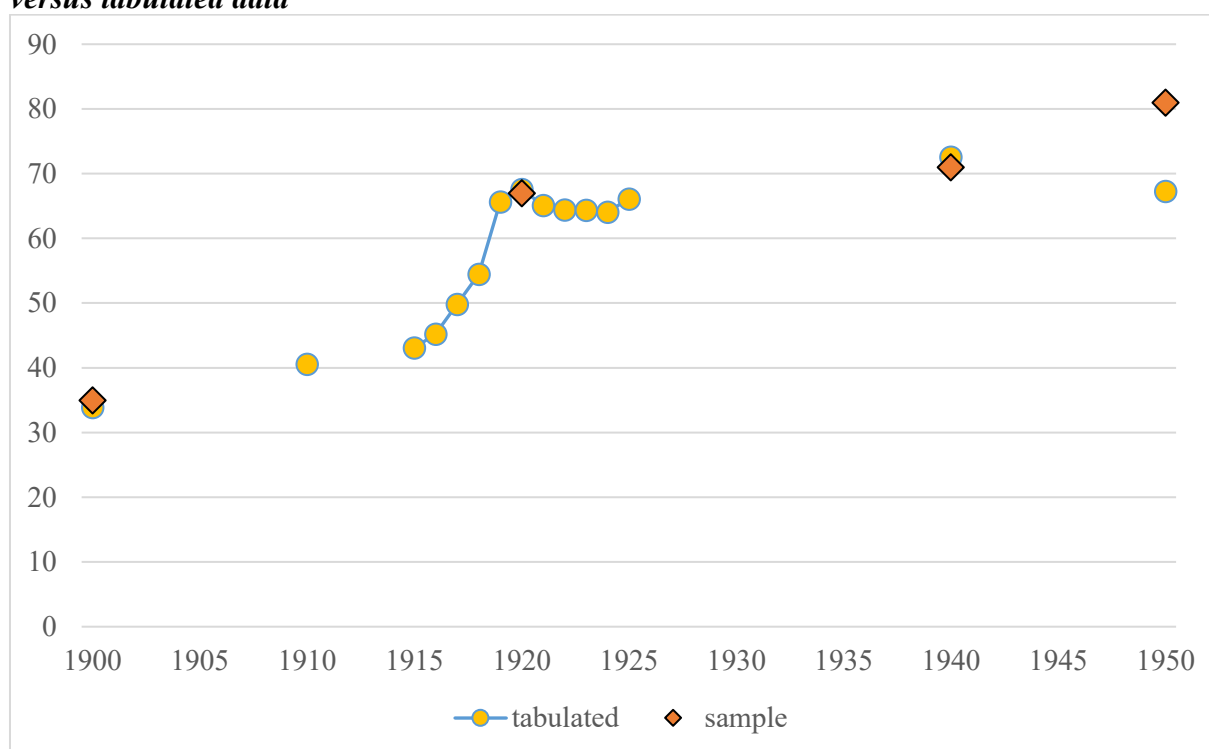
Note. Sources for the series "tabulated" are the Statistical yearbooks for Stockholm City, 1904 to 1952. The series "sample" is the dataset that is used in the paper.

Let us now move to a comparison of inequality estimates with our sample and with the tabulated data. Figure B3 shows three different series, with comparisons for 1900, 1920, 1940 and 1950. The first series replicates our baseline Gini taking individuals as the income-earning unit and with imputed incomes for those below the threshold for paying taxes (described in Appendix C). The second series shows the Gini using tabulated data on the full taxpaying population published by the Stockholm statistical office. Individuals are assumed to earn 50 percent of the range in the tabulated data. Thus, if the range is 900-1000 crowns, income is set to 950. Those in the top bracket are assumed to earn exactly the amount to fall in this bracket, because there is no upper limit for inclusion in this group. Everyone below the threshold for paying taxes is assumed to earn 50 percent of the threshold.[14] The third series

---

[14] Another measure on which we can compare the sample and the Stockholm City yearbook's tabulations is the average income of Stockholm adults. In 1900 we calculate an average income of 789 kr (based on the sample, adjusted for incomes of working non-filers), while our estimate based on the yearbook tabulation gives an average income of 834 kr; the difference is only 5 percent. Similarly, in 1920 we calculate an average of 3 544 kr while the yearbook tabulations give an average of 3693 kr; the difference is only 4 percent. In 1940 the difference is 12 percent and in 1950 the difference is 4 percent; the only year in which we have a significant difference between the calculations is 1910, when our estimated average income is 20 percent lower than that given by the tabulated data. Given that 1910 does not provide an outlier in terms of the inequality estimates in the paper, and that the difference for 1900, 1920, 1940 and 1950 is limited to 4-12 percent, we see our sampling strategy as a sound one.

"tabulates" our own sample using the same brackets and assumptions about incomes as in the second series, allowing for a more direct comparison between the two.

The tabulated data show greater inequality than our estimate in 1900. However, the difference is still within the confidence interval of our original point estimate. The tabulated data then show a lower level in 1920 than we have, which reveals a slight decline from 1900 to 1920 instead of an increase. As we saw in the main body of the paper, the increase from 1900 to 1920 was within the confidence interval. Furthermore, the 1920 estimate is very sensitive to the very top incomes which were booming in this year (more on this below), and the tabulated data do not contain information on the actual incomes. For 1920, they only show that there were three men who earned more than a million crowns (the top bracket), but without exact information about their incomes: it could have been 2 million each, or 3 million each. This lack of detail on top incomes gives a downward bias for the inequality estimates.

After the slight deviation in 1920, the tabulated data also like our main data show a decrease in inequality from 1920 to 1940 and 1950.

Figure B3. Comparison of the Baseline Inequality Estimate to Tabulated Data and Tabulated Sample
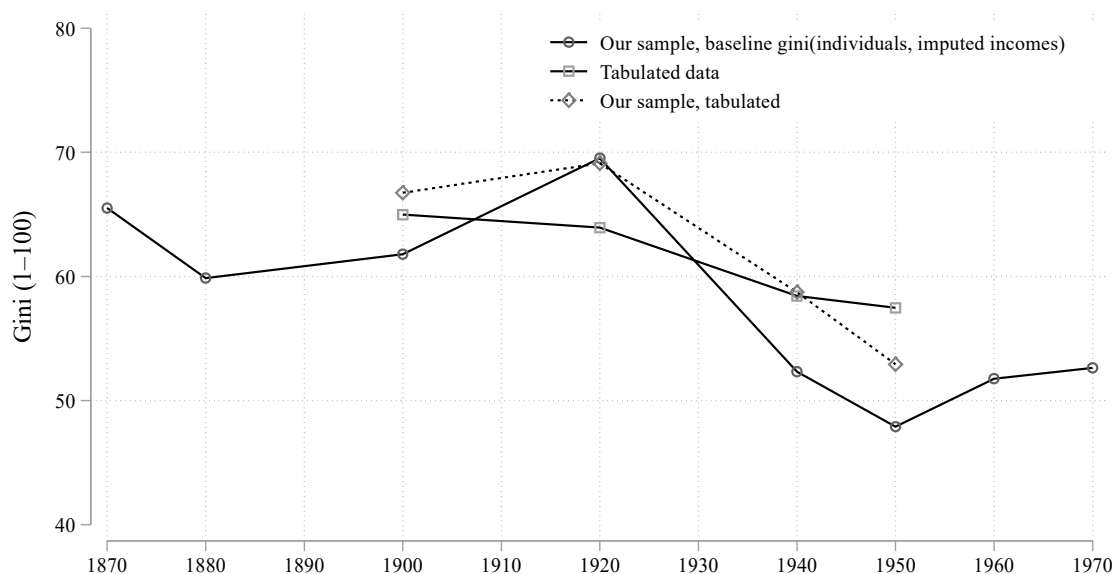
To take a more detailed look at what parts of the income distribution are driving these differences, Figure B4 presents the Lorenz curve for the tabulated full-count data and the tabulated version of our sample for each year. The difference in 1920 seem to be driven by the

---

We can trust that the 1 percent sample (0.5 percent in 1940 and 1950) captures the structure of Stockholm incomes well.

top of the distribution, while the gap for 1950 is concentrated in the middle. The difference in 1950 is probably the result of the higher share of taxpayers in our sample than in the tabulated data.

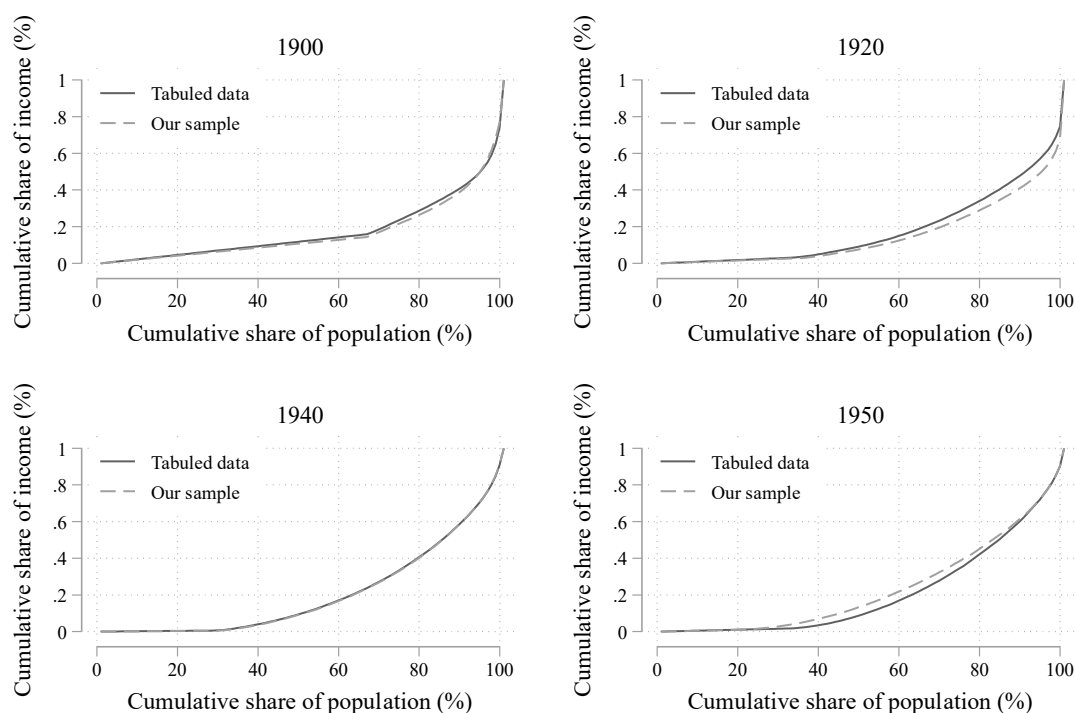*Figure B4. Lorenz curves for the sample and the tabulated data*

Figure B5 highlights the problem with 1920 by presenting the comparison in a slightly different way: it is in the form of a scatterplot with each percentile's share of income in the tabulated full-count data on the x-axis and each percentile's share of income in our sample on the y-axis. If the two datasets corresponded completely, the share of income for a certain percentile in our sample should lie on the 45-degree line. As is clear from the graph, this is also the case for every percentile except the top 1 percent. Their share of income in our data is 32 percent but only 25 percent in the tabulated data.

Figure B5. An in-depth look at the 1920 distributions
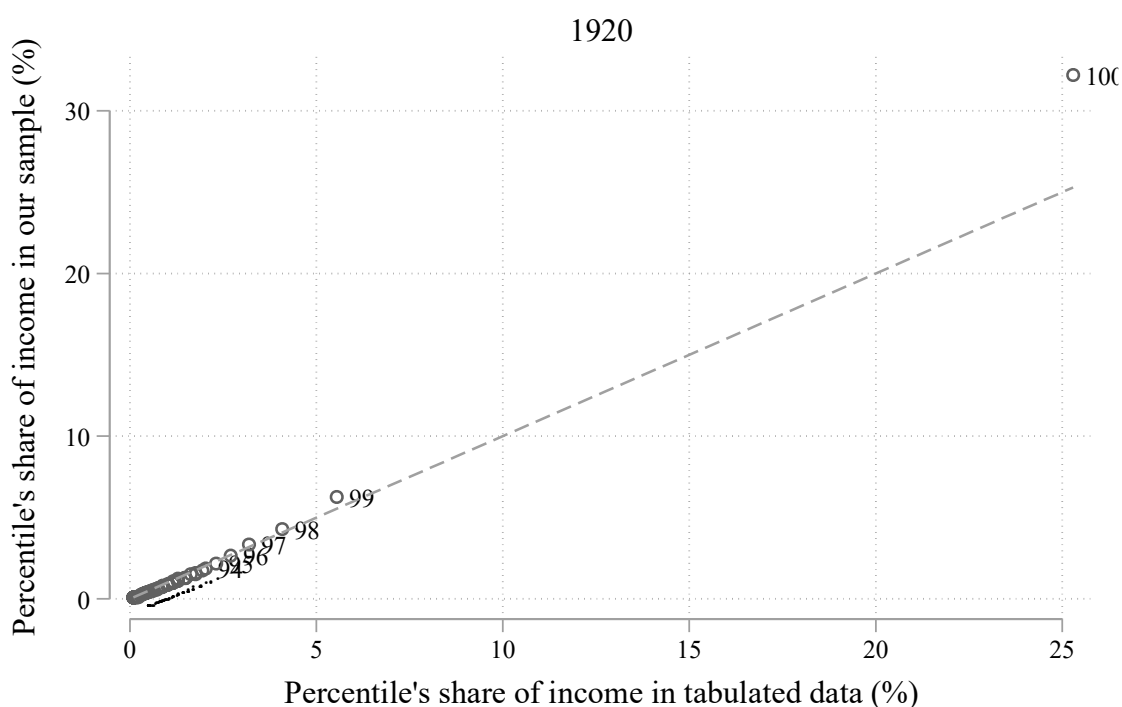
The take-away from Figures B4 and B5 is that 1920 was a very unusual year. Top incomes boomed during the inflation of the First World War – Roine and Waldenström's (2008) series for the country as a whole show a historical peak of inequality in 1916 – and just after, and with this great dispersion and very large incomes at the top, the inequality estimates become especially sensitive to the measurement of top incomes.

Beyond the problem with year 1920, we conclude from the comparison with tabulated data that our sample is a representative one and that the estimates of incomes and income inequality presented in the main body of the paper are reliable and quite precise.

References Appendix B

Ericson Wolke, Lars. 2016. *Stockholms historia*. Lund: Historiska Media.

Roine, Jesper and Daniel Waldenström. 2008. "The Evolution of Top Incomes in an Egalitarian Society: Sweden, 1903–2004". *Journal of Public Economics*, 92 (1-2), 366–387.

Statistical yearbooks for Stockholm City, 1904 to 1952. They are all digitalized and available from the Stockholm City archive website: https://sok.stadsarkivet.stockholm.se/Databas/statistik-arsbok-1904-2006/Sok

Appendix C. Imputed Incomes of Non-filers

In this Appendix we assess the robustness of our results to different assumptions about the incomes of non-filers. This is an important problem for all studies using tax data to study incomes, as acknowledged and discussed especially by Atkinson (2007), but a problem which has not been investigated enough in the existing long-run inequality literature. Therefore, in this Appendix we try out several alternatives to the baseline imputation method used in the body of the paper.

In our baseline imputations, we drew on secondary sources and information about these individuals' social group (textile worker, widow, retired, etc.) to estimate the income of non-filers. A key general reference is Bagge et al (1933) but for the groups discussed here we also used Leffler (1897), Hesselgren (1992), Norlander (2000) and other detailed studies of specific groups and workplaces. See Bengtsson (2021) for a detailed discussion. Table C1 shows imputed incomes for the key groups; for other non-filers we impute a uniform income equal to women textile workers and for men metal workers from Table C1.

Table C1. Imputed incomes for important groups of non-taxpayers

	Gender	1870	1880	1890	1900	1910	1920
Textile	Women	225	225	255	255	305	475
Domestic	Women	205	205	250	250	300	450
Retired	Women	125	125	125	237	350	350
Spouse or widow	Women	150	150	225	225	275	275
Domestic	Men	300	350	400	475	475	475
Metal	Men	350	350	400	425	525	599
Transport	Men	300	350	350	400	400	599
Retired	Men	175	175	175	275	375	375
Spouse or widower	Men	325	325	425	525	525	525

Note. Source is Bengtsson (2021) *under strecket appendix*, Table X2. For non-filers not discussed in the table, incomes are set as follows. For women 1870-1920: 200, 200, 250, 250, 300, 500 kr. For men: 350, 350, 400, 400, 500, 599. There was no public pensions system before 1913. On working-class saving for old age, see Bäcklund and Lilja (2014) who among other things use household surveys from 1908-1909. On old people's household formation and income strategies in a small-town context (Sundsvall) 1830–1930, see Högman (1999). In 1940 and 1950 only a small proportion of adults paid no tax (see Table 2) and we set a uniform income of non-income taxpayers as 250 kr for women and 250 kr for men. The negative selection of this group was much stronger in these years, and so the lower imputed income is motivated: it is less likely that they were gainfully employed, or at least not full time, and more likely that they were out of the labor force with little or no income of their own.

An alternative method is to use regression models to estimate how the income for tax-filers with incomes at the lower end of the income distribution varied with social class, sex, and age, and then use those regression estimates to impute incomes for non-filers. We call this method "regression-based imputations".

A third kind of method is simply to assume that all non-filers earned a certain fraction of the threshold for paying taxes. We try out two such method., one where we assume that all non-filers earned the threshold income, which we call "at threshold", and the other where we assume that all non-filers earned half of the threshold, which we call "at threshold*0.5".

Figure C1 shows the resulting Gini using these different imputation methods. The figure also shows the Gini if non-filers are assumed to have had no income.

Figure C1. Gini Using Different Imputation Methods.
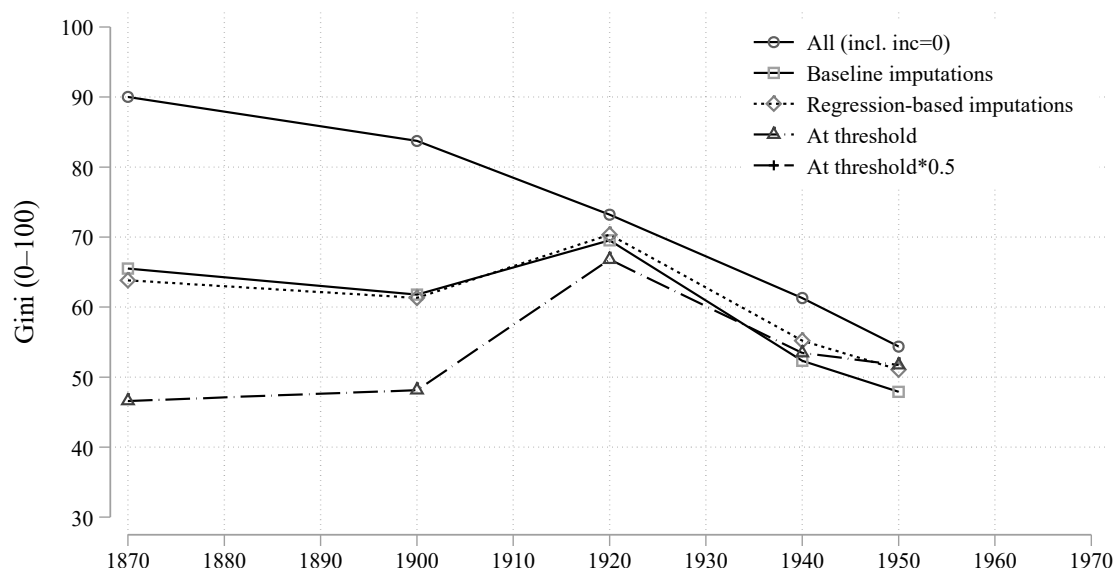
Figure C2 displays this by showing Lorenz curves for 1900 and 1920 using the different imputation methods. In 1900, the data without imputations (the series with circles) show an extremely skewed distribution of income, for it implies that as much as 60 percent of the population had zero income. The imputation which assumes that all non-filers earned at the threshold for paying taxes (the series with triangles) yields that the bottom 60 percent would earn about 30 percent of incomes. These are the estimates relying on the most extreme assumptions. The three intermediate methods: the baseline imputations, the regressions-based imputations, and assuming that all non-filers earned 50 percent of the threshold, generate very similar Lorenz curves.

Figure C2 shows that by 1920 the share of tax-filers had increased so much that there was no longer any major discrepancy between the different imputation methods in terms of inequality estimates, which is also evident from the trend in the Gini shown in Figure C1. The correspondence between our baseline imputations, the regressions-based imputations, and assuming that all non-filers earned 50 percent of the threshold, makes us confident that our main results are not driven by the particular approach to the imputation of incomes for non-filers.

Figure C2. Lorenz Curves Using Different Imputations Methods, 1900 and 1920
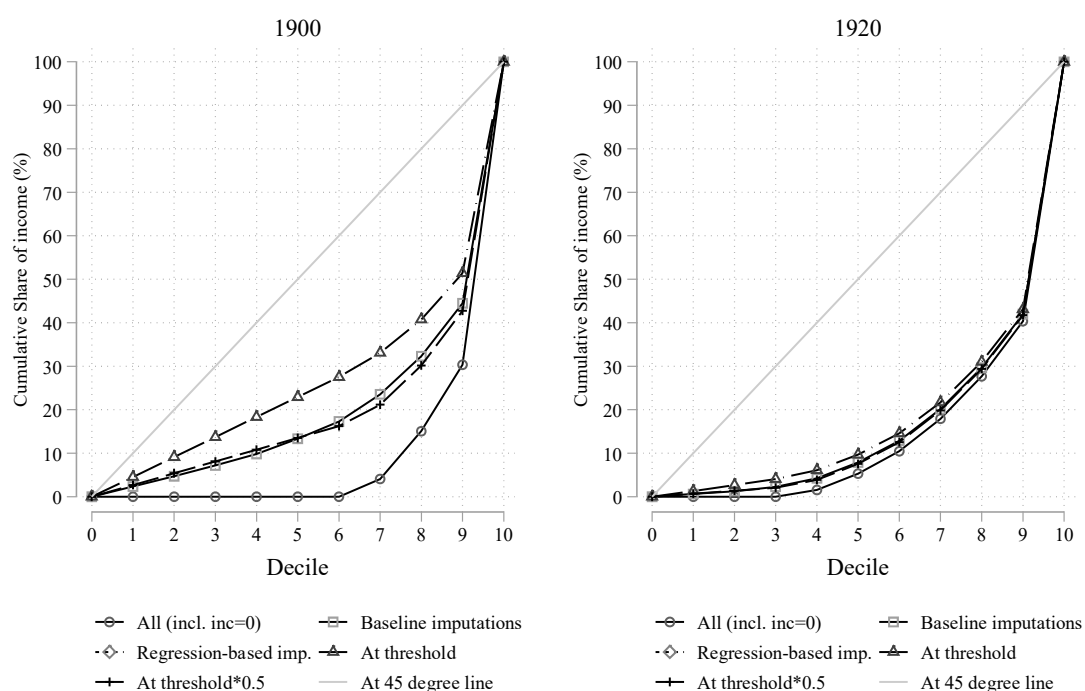
References Appendix C

Atkinson, Anthony B. 2007. "Measuring Top Incomes: Methodological Issues", pp. 18-42 in Atkinson and Thomas Piketty (eds.) *Top Incomes over the 20th Century: A Contrast Between European and English-Speaking Countries*. Oxford: Oxford University Press.

Bäcklund, Dan and Kristina Lilja. 2014. "Variation och förnyelse: Arbetarsparande i Sverige 1870–1914". *Historisk Tidskrift* 134 (4): 615–646.

Bagge, Gösta, Erik Lundberg, and Ingvar Svennilson. 1933. *Wages in Sweden 1860–1930. Part One: Manufacturing and Mining.* London: P.S. King.

--- 1935. *Wages in Sweden 1860-1930. Part Two: Government and Municipal Services, Agriculture and Forestry, the General Movement of Wages in Sweden 1860–1930.* London: P.S. King.

Bengtsson, Erik. 2021. "Estimating the Incomes and Living Standards of Non-Income Tax Payers: Sweden, 1870–1950". Work in progress, Lund University.

Hesselgren, Eva. 1992. *Vi äro tusenden: arbets- och levnadsförhållanden inom svensk textilindustri med särskilt avseende på Gamlestadens fabrikers AB 1890–1935.* Doctoral thesis in economic history, Stockholm University.

Högman, Ann-Kristin. 1999. *Ageing in a changing society. Elderly men and women in urban Sweden 1830–1930.* Dissertation in History, Umeå University.

Leffler, Johan. 1897. *Zur Kenntnis von den Lebens- und Lohnverhältnissen industrieller Arbeiterinnen in Stockholm.* Stockholm: Koersners.

Norlander, Kerstin. 2000. *Människor kring ett företag. Kön, klass och ekonomiska resurser. Liljeholmens Stearinfabriks AB 1872–1939.* Doktorsavhandling i ekonomisk historia, Göteborgs universitet.

LUND
UNIVERSITY