INCOME INEQUALITY AND STRUCTURAL TRANSFORMATION: EVIDENCE FROM SWEDISH MICRO DATA, 1870–1970

ERIK BENGTSSON
JAKOB MOLINDER
SVANTE PRADO

WORKING PAPER N°2026/14

JUNE 2026

WORLD INEQUALITY LAB

Income Inequality and Structural Transformation: Evidence from Swedish Micro Data, 1870–1970¶

Erik Bengtsson*, Jakob Molinder†, and Svante Prado‡

June 8, 2026

Abstract

We study the relationship between the structural transformation of the economy and the income distribution, focusing on the case of Sweden from 1870 to 1970, with extra attention paid to the 1870–1950 period, for which we produce extensive new data. Average income increased fivefold between 1870 and 1950, and the share employed in agriculture declined from 72 to 23 per cent. To study the evolution of the income distribution, we collected new data, including 232,000 individual income tax returns, 13,000 property tax returns, and a rich set of complementary sources. Using these micro data, we calculate Gini coefficients, top income shares, capital shares, skill premia, and occupation- and gender-specific income levels and ratios, providing new evidence on the long-run evolution of income and inequality. Our income data and decomposition analyses demonstrate that the shift of labour out of agriculture, which was a severely unequal sector in Sweden in the late nineteenth and early twentieth centuries, accounts for much of the decline in income inequality, together with the expansion of more productive jobs in manufacturing and offices. This process was aided by labour mobility, as well as by educational and other policies that facilitated structural transformation. Focusing on structural transformation can help explain two paradoxes in the literature on twentieth century income inequality: that much equalisation occurred before the growth of the welfare state, and that non-belligerents in the World Wars, like Sweden, saw similar levels of equalisation as belligerent countries.

Keywords: *incomes, inequality, Sweden, structural transformation, economic history*

JEL codes: *D31, J30, N30, O52*

*Department of Economic History, Lund University; Fellow, World Inequality Lab erik.bengtsson@ekh.lu.se

†Department of Economic History, Uppsala University; Department of Economic History, Lund University; Uppsala History of Inequality and Labor Lab (UHILL) jakob.molinder@ekhist.uu.se

‡Department of Economy and Society, University of Gothenburg svante.prado@econhist.gu.se

¶ Research financed by Swedish Research Council (*Vetenskapsrådet*) project grant 2018-01853 and Jan Wallanders och Tom Hedelius stiftelse samt Tore Browaldhs stiftelse grant P18-0197. We are grateful to Johanne Arnfred, Björn Chebrell, Erik Hellberg, Valeria Lukkari, Pia Lundström, Markus Matovac, Erik Olsson, Jens Olsson and William Skoglund for research assistance. The paper has been presented at workshops, conferences and seminars in Gothenburg, Lund, Stockholm, Oslo, Aalborg, Leiden, and Copenhagen; thanks to participants for comments and suggestions. We are very grateful to Rolf Aaberge, Andreas Bergh, Pierre Brassac, Vincent Delabastita, Casper Worm Hansen, Suvi Heikkuri, Saska Heino, Martin Karlsson, Felix Kersting, Tomas Korpi, Jørgen Modalsli, Therese Nilsson and Thomas Piketty for comments. Any errors and misunderstandings are our own.

1 Introduction

What shapes income distribution in the long run? This paper contributes to the substantial and important research on this question by studying Sweden over one hundred years, 1870 to 1970, with a novel micro data approach. Our research period encompasses most of the middle six decades of the twentieth century, c. 1920 to 1980, which saw the greatest-ever recorded equalisation of incomes and wealth in Europe and America. Inequality has risen since, but not returned to the levels of the early twentieth century. A large body of research literature in Economics and Economic History has mapped the evolution of income distribution, building in particular on tabulated tax data available from public authorities.[1] Detailed survey data or administrative micro data is typically only available from the late 1960s or 1970s.[2] Explanations of the great mid-century equalisation have centred on the rise of progressive taxation and the destructive effects of wars and the Depression on capital (Piketty, 2014; Scheidel, 2017) as well as supply and demand of skills (Goldin and Katz, 2009; Lindert and Williamson, 2016; Heikkuri, 2025).[3] However, the role of structural transformation has been considered to a much lesser degree. This is associated with the need for nuanced data, preferably micro data, if one wants to explore the implications of sectoral shifts for inequality in the economy and changes in job composition.[4]

The contribution of this paper is to bring in the role played by structural transformation in the twentieth-century transformation of income distribution, building on a new micro dataset for Sweden. Our view of structural transformation is a comprehensive one, focusing on the shift from agriculture as the largest sector for value added and employment, to a more diversified, industrialised economy with a large services sector, but also taking in related aspects such as the average size of firms, and the decline in family-run businesses (cf. Gollin and Kaboski, 2023). Sweden in the late 1800s, like all now-industrialised European economies bar Britain, was an agricultural economy with family-run businesses, especially family farms, playing a key role (for a US parallel cf. Ruggles, 2015; Ngai et al., 2024). From 1870 to 1900, the share of the workforce employed in agriculture shrank from 72 per cent to 58 per cent, then to 23 per cent in 1950, and to 8 per cent in 1970 (Edvinsson, 2005, Table O).

[1] The starting point was Piketty (2003) and Piketty and Saez (2003); for an overview of the literature see Roine and Waldenström (2014). Later studies by Piketty et al. (2018); Garbinti et al. (2018) and others in the Distributional National Accounts literature have adopted a much more ambitious and comprehensive approach to income, but the detailed data of these studies have typically been restricted to the post-1960 period, again leaving out the structural transformation analysed in this paper.

[2] See the discussion for example in Auten and Splinter (2024); Piketty et al. (2024).

[3] There is a debate on methodological choices in studying early twentieth century income inequality — see Geloso et al. (2022) – but the pattern of high income inequality in the early 1900s, followed by a drastic fall after 1920, holds also with revisionist assumptions.

[4] We suggest that there are two unexplained puzzles in the literature on twentieth-century equalisation, given the focus in the literature on taxation, war destruction and war-induced taxation (Piketty, 2014; Scheidel, 2017). Firstly, much of the equalisation occurred before the evolution of progressive taxation and redistributive welfare states (Roine and Waldenström, 2008). Secondly, countries not involved in the World Wars, such as Sweden, saw as much or more equalisation as the belligerent countries. This means that Sweden is an especially pertinent case for re-examination.

We document the significant role of structural transformation in equalising Sweden's income structure. We have constructed our data in two steps. Firstly, we have built a new dataset of about 232,000 individual income tax returns collected from the National Archives, supplemented by incomes from the 1930 Census, and with Statistics Sweden micro data for 1960 and 1970. These data include information on income by source, sex, occupation, sector, and place of residence. Secondly, we have added group-specific imputations of incomes for individuals falling below the tax-paying threshold, and for farmers who paid wealth tax instead of income tax (imputations for farmers also built on a new data set of 13,000 property tax returns). This rich set of data facilitates the investigation of how the shares of employment across sectors and location shaped income distribution. Given our historical context, the data are imperfect and the construction of the dataset requires assumptions, but we show throughout and in the Online Appendices that the main results are robust.

The modern conceptualisation of the connection between structural transformation and income inequality began with Simon Kuznets' (1955) influential model. Kuznets assumed that agriculture was a poorer but also more egalitarian sector than industry. Following this assumption, he foresaw that as industrialisation progresses, inequality first increases as the more unequal industrial sector expands. As a result of this expansion, the gap widens between those in agriculture, the lower-income egalitarian sector, and those in industry, the higher-income inegalitarian sector. At some point, inequality begins to decline, as everyone has transitioned to the urban sector, converging on a higher income.[5]

Against the Kuznetsian conceptualisation of the structural transformation–inequality connection, we may counterpose a perspective from recent research in "macro development", which concerns the intersection of macroeconomics and development economics. This literature emphasises sectoral differences in productivity levels and between-sector mobility barriers (Gollin et al., 2014; Gollin and Kaboski, 2023). It argues that low productivity in agriculture keeps incomes low for the vast majority and that productivity may be held back by imperfectly small farm sizes (Foster and Rosenzweig, 2022). Furthermore, market failures such as restrictions on labour mobility or lacking access to capital that facilitate the expansion of more productive firms may prevent workers from moving from agriculture to industry, which delays structural transformation and impedes economic growth. The eventual decline in inequality, as Kuznets envisioned, may then never occur. This more recent literature also gives a more optimistic view of the relationship between industrialisation and inequality, emphasising that the manufacturing industry has a unique capacity to produce highly productive, well-paid jobs for the less educated, which equalises income distribution (Rodrik, 2016). Thus, "premature deindustrialisation" is a problem for the opportunity for many people in today's developing economies to be able to earn a good living.

[5]The Kuznets Curve has attracted plenty of criticism (e.g., Acemoglu and Robinson (2002); Piketty (2014), pp. 11–15, 237; Alfani and di Tullio (2019), pp. 9-14) but is still a workhorse model and has been extended within the longrun inequality research for example with van Zanden's (1995) "Super Kuznets Curve" and Milanovic's (2016) "Kuznets Waves".

In contrast, the developed countries, who began their journey in the nineteenth century, have benefited from long-lasting industrialisation and a share of manufacturing sufficiently large to absorb the inflow of workers from agriculture.[6] To the macro development literature, our study contributes with a comprehensive empirical analysis of how incomes grew historically and how their distribution changed with industrialisation, including the mechanisation of agriculture.

Our results show that income inequality was high in the agrarian economy of Sweden in 1870 to 1920. The share of total incomes captured by the top 10 per cent of the income distribution was 54 per cent in 1870 and 48 per cent in 1900, but fell to 37 per cent in 1950. The broader inequality measure of the Gini coefficient also highlights very high inequality at the beginning of our period, with a Gini of 66.9 per cent in 1870, declining to 61.3 per cent in 1900 and 58.4 per cent in 1950. The decline in top income concentration was greater than the decline in the Gini, which highlights another of our results, that the equalisation of Sweden's (pre-tax) income distribution was driven in particular by growing shares of income going to the middle groups of the income distribution. Another descriptive result is that in the early twentieth century, in terms of its inequality Sweden was at a comparable level with France and the United States (Garbinti et al., 2018; Piketty et al., 2018).[7]

To explain how inequality declined, we employ a set of counterfactual exercises and inequality decompositions. The counterfactuals of the Gini coefficient explore how the Gini would have evolved in various scenarios: without any change in the sectoral composition of employment, or without change in within-sector inequality, or without change in sectoral mean income. Secondly, taking leave of the Gini coefficient for the more flexibly decomposable inequality measure, the Mean Log Deviation (cf. Cowell and Flachaire, 2024), we decompose changes in inequality into within-group inequality, between-group inequality, and population share effects. Both the counterfactuals and the decomposition reveal that Kuznetsian factors (inequality growing because of industrialisation) were much less important than, in particular, the massive decline in inequality *within* sectors. While the urban sector was more unequal than the rural, it was not markedly so. Documenting the occupational composition of the economy with a focus on low-paid occupations, we highlight the decline of large, low-income groups: more than a fifth of the labour force were agricultural workers in 1870 and 1890, but less than 10 per cent in 1940 and 1950, and the share of unpaid family assistants (mainly farmers' wives and unpaid adult children on the farm) declined from also about a fifth in 1870 to 1910 to 5 per cent in 1940 and even less than that in 1950. Here we actually connect back to one of the first studies of the contemporary literature on long-run inequality, Piketty (2003, p. 1026) who in his seminal article on French income distribution from 1901 to 1998 stressed that the share of domestic workers in the French

[6] Ager et al. (2026) exemplify this mechanism for early twentieth-century Norway, in a setting much like the Swedish, where mechanisation of milking served to channel female labour into service and manufacturing, which entailed an upgrading of the workforce at large. Our macro study of incomes and how they changed with structural transformation has many analytical concerns in common with their micro study of the effects of one particular innovation. See also Heikkuri (2026)

[7] See also Ericsson and Molinder (2026) for a comparison of Sweden to other countries.

labour force fell from 5 per cent to 1 per cent from the First World War to the 1950s, and that "The parallelism between this evolution and the evolution of top income shares is striking." In Piketty's analysis, the decline of domestic work was related to income concentration, especially since, post-equalisation, the less rich top groups couldn't afford to employ as many servants as before. In our analysis, the decline in low-paid groups appears to be in itself a driver of shrinking income concentration, as the replacement of low-paid jobs by better-paid jobs mechanically expanded the income of the bottom groups and the working classes.

Our account of the evolution of long-run income inequality is by no means a monocausal one. We show that the decline in capital income and the expansion of education á la Piketty (2014, 2020) and Goldin and Katz (2009) were important factors. Capital income was extremely concentrated, and its shrinking share of national income after 1920 contributed to equalisation, but not in such a way that it can replace structural transformation as an explanation. The expansion of education surely mattered too, but in 1930 still less than one per cent of the labour force had a college degree and less than two per cent an upper secondary school diploma (SOU, 1935), and the expansion of colleges and secondary schools was concentrated to the post-1950 period. Thus, the contribution of education per se in the Swedish case was instead concentrated to the second half of the twentieth century, as was the contribution of progressive taxation.

In a back-of-the-envelope comparison of Sweden with other developed economies, we show that there is no reason to believe that Sweden had an exceptionally unequal agricultural sector at the turn of the twentieth century, or a much more agrarian economy. This means that it is likely that structural transformation mattered for the observed equalisation of incomes that was also observed in other developed economies over the twentieth century, even though this has not been visible in existing long-run inequality studies based on tabulated data.

The paper proceeds as follows. In section 2 we lay out our analytical framework, with a focus on decomposing changes in inequality over time. Section 3 describes the sources used and the data sets constructed for the analysis. Section 4 presents an overview of Swedish income inequality from 1870 to 1970, with Gini coefficients, group income shares and occupational incomes. Section 5 turns to the relationship between income inequality and structural transformation, focusing on the role of the rural-to-urban and agricultural-to-industrial shift and the role of occupational composition. Section 6 extends this analysis by applying counterfactual and decomposition techniques. Section 7 places the Swedish case in comparative perspective. Section 8 relates our case study to two explanatory factors often invoked in the research on long-run inequality: capital income, and education. Section 9 concludes.

2 Analytical framework

This paper explores the interaction between structural transformation and income inequality. It is important to state that the paper does not build on a causal design: it is not a design where we have located an exogenous shock that affects various parts of the income distribution in different ways, thus changing the distribution. In the language of Acemoglu et al. (2014), ours is an exploration of *proximate*, not *fundamental* causes. Acemoglu et al. (2014, p. 876) quote North and Thomas's (1973) classical account of "the rise of the Western world" to the point that factors such as innovation, economies of scale, education, and capital accumulation "are not causes of growth; they *are* growth." Analogous to this argument concerning economic growth, we could say that factors such as the narrowing of occupational income ratios or diverging sectoral average incomes are not causes of changes in inequality, they *are* changes in inequality. Our guiding question is: *how* did the income distribution of Sweden change over a hundred years? Thus, our methods focus on description, decomposition, and counterfactuals: methods with great strengths in mapping the factors that account for changes in inequality, but not for investigating *causes* behind the factors as in, say, an instrumental variable design. (Compare the discussion of decomposition methods in Fortin et al., 2011.)

With a wide perspective on structural transformation as in Gollin and Kaboski (2023, p. 340), the pertinent aspects of structural transformation are the shift from agricultural to industrial and service-sector employment; urbanisation; the shift from self-employment on farms and in skilled trades to employment in larger and more complex firms; and a shift out of home-based production. All these aspects are relevant to the transformation of the Swedish economy from 1870 to 1970 (Morell, 2001; Flygare and Isacson, 2003); the clearest single indicator is probably that the share of employment in agriculture declined from 72 per cent in 1870 to 23 per cent in 1950 (Edvinsson, 2005, Table O).

Our paper focuses on income inequality, but income inequality is not a straightforward concept. Indeed, the literature offers a wide array of measures of inequality: concentration measures that capture the share of income received by top income earners; comprehensive indices like the Gini or Theil, and so on (Cowell, 2011; Piketty, 2014; Jenkins, 2017; Blesch et al., 2022; Cowell and Flachaire, 2024). Let us therefore say a brief word about the measures used here. To describe the distribution, we will initially use group shares – the top decile's share of income, the middle 40 per cent share, the bottom 50 per cent – in the tradition of Piketty (2003, 2014); Piketty et al. (2018). As argued by Piketty (2014), these measures give,a concrete reality to the distribution as we can see – unlike with synthetic measures such as the Gini coefficient – *where* in the distribution incomes converge or diverge. With our micro data we also found it important to identify the social composition of the various income groups, to give a sociological sense of how the income distribution – a quite abstract concept – mapped onto social reality in Sweden in the 1800s and 1900s (cf. Milanovic, 2024). For the same reason, occupational income differences (i.e., landlord versus farm worker) are also important indicators in our analysis.

However, to evaluate the importance of the various aspects of structural transformation for the income distribution, we will also use decomposable measures of inequality: the Gini Coefficient and the Mean Log Deviation (cf. Mookherjee and Shorrocks, 1982; Cowell, 2011; Cowell and

Flachaire, 2024). We use these to quantify the arguments from analysts of structural transformation such as Kuznets and Lewis: for example, how important for overall inequality was within-sector inequality and its changes as opposed to the possible narrowing of between-sector inequality? We will also use decomposition techniques to quantify the importance of the shrinking (and possible equalisation) of capital income as opposed to the equalisation of labour income as drivers of overall inequality.

From this focus on the "outcome variables" (even if our analysis is descriptive and an accounting exercise rather than a causal design), let us move on to our conceptualisation of the explanatory factors. Our first approach to the question of how this transformation shaped the income distribution is the decomposition of inequality by sector (presented in Section 5). Based on contemporary statistics from the US and India and late nineteenth century statistics from Prussia, Kuznets (1955, pp. 7–8) proposed that incomes in the agricultural sector were both lower and more equally distributed than incomes in the industrial sector. Industrialisation thus increases inequality by increasing the difference between individuals in the two sectors, and by re-weighting the sectoral shares of the total so that the within-sector inequality of the more unequal industrial sector dominates.

However, this is not the only possible conceptualisation of the income differences between industry and agriculture. Another canonical model of development economics, that of Lewis (1954), presents an even more pessimistic view of the agricultural sector, with different implications for the income distribution. For Lewis, the agricultural sector in poor economies is marked above all by a surplus of labour and, following this, inefficient use of labour, amassing workers and family members in small, low-productive family enterprises which swallow so much labour not because they make productive use of it but because the labourers have few other options (cf. Fields, 2004, pp. 728–730). This informal sector is matched by a much smaller, capital-intensive, high-productive, formal sector where agrarian capitalists employ labourers with greater productivity and higher wages, but also a high degree of income concentration as the latifundia owners control much of production. With such a conceptualisation, we may imagine a quite unequal agricultural sector: a small group (imagine owners of great estates or latifundia) with high incomes, and a large group with almost subsistence-level incomes. If, in the tradition of Shorrocks (1984), we see sectoral inequality as the composite outcome of the incomes of the various groups (labourers, farmers, estate owners and so on), the relative size of the groups, and within-group inequality, then the difference in inequality between a Kuznetsian agricultural sector and a Lewisian one is clear: a compressed economy of middle-income farmers in the first case, and a spread-out, bimodal distribution of low and high incomes in the other.

A further, even stronger assumption of the Lewis model is that the elastic supply of low-skilled, under-employed labour in the stagnant agricultural sector keeps wages down also in the industrial sector even if that sector experiences productivity growth, as the agricultural sector can release so much labour with low demands on remuneration (cf. Fields, 2004). However, it is equally plausible, and compatible with a neoclassical growth model, that industrial growth instead will absorb labour from agriculture and improve wages in both sectors (Rodrik, 2016) as the demand for labour grows and as agricultural production itself is mechanised (Manuelli and Seshadri, 2014).

In order to formalise our analytical framework and test the Kuznetsian and Lewisian predictions regarding the role of the agricultural/rural and the industrial/urban sectors (Kuznets was not completely clear about the distinction), we begin by decomposing the Gini coefficient[8], expressing the overall Gini coefficient (G) as:

$$G = G_w + G_b + G_r \tag{1}$$

where G_w is the within-group inequality component, G_b is the between-group inequality component (dependent on the difference between average incomes in industry and in agriculture, μ_i and μ_a), and G_r is the residual (overlapping) component. These components are, in turn, derived as follows. To calculate within-group Inequality (G_{w}), for each group k (e.g., urban and rural), we calculate firstly the population share (P_k):

$$P_k = \frac{N_k}{N} \tag{2}$$

where N_k is the population of group k and N is the total population. Then, the income share (S_k):

$$S_k = \frac{Y_k}{Y} = \frac{N_k \mu_k}{N_\mu} \tag{3}$$

where Y_k=$N_k\mu_k$ is the total income of group k, μ_k is the mean income of group k, and μ is the overall mean income. Finally, the weights (w_k) are calculated as $w_k = s_k p_k$. The within-group inequality component is then calculated as:

$$G_w = \sum_k w_k G_k) \tag{4}$$

Between-group inequality is calculated as the Gini coefficient of the group mean incomes μ_k:

$$G_b = Gini(\mu_1, \mu_2, \ldots, \mu_K) \tag{5}$$

[8]The Gini coefficient as a measure of inequality has been criticised for its unbalanced sensitivity to changes in the middle of the distribution and for the lack of transparency (Piketty, 2014, pp. 266-268; cf. Aaberge, 2007, Fortin et al., 2011, p. 57, Cowell and Flachaire, 2024). We use it here for its flexibility to decomposition, but will also show other measures of income concentration and inequality. See also the general discussion of decomposition in Shorrocks (1984).

And finally, the residual component accounts for the income overlap between groups:

$$G_r = G - (G_w + G_b) \tag{6}$$

In the Kuznets model, the two operative parts are the G_b ($\mu_i > \mu_a$) and that $G_{\text{wi}} > G_{\text{wa}}$, where G_{wi} is inequality within industry and G_{wa} is inequality within agriculture. If we instead accept a modified Lewis model of the Swedish economy of the late 1800s, we expect a high G_{wa} and that industrialisation will actually shrink within-agriculture inequality by absorbing low-paid, surplus labour from this less productive sector. In a strong version of the Lewis model, the dynamic sector will not be able to absorb the surplus labour but instead the availability of surplus labour will mean that G_{wi} will grow radically as capitalists benefit from industrial development, but workers do not. (See discussion and overview of later research in Fields, 2004, pp. 725–730 and Gollin, 2014, pp. 75–82.)

To test the Kuznetsian and Lewisian predictions, the decomposition by sector is important – and our data are flexible enough that we can test both the urban/rural dimension and the agricultural/industrial dimension (as in Equation 1). In a country like Sweden with a significant degree of rural manufacturing industry (Schön, 2014), it is important to consider both distinctions. To test the hypothesis that inequality grew as the more unequal and high-income industrial sector grew, we also employ counterfactual analysis and further decompositions.

In our counterfactual analyses, we separate out the various components of the agricultural-to-industrial transition: changes in mean incomes (the Kuznets argument that $\mu_i > \mu_a$), population shares (the shift in labour from one sector to another), and within-group Gini coefficients to the development of overall inequality (the Kuznetsian $G_{\text{wi}} > G_{\text{wa}}$). By fixing each of the three components to the baseline year (1870) in alternative counterfactuals, we isolate what would happen if everything else changed but this aspect remained as it was in 1870. So, to start with, if the share of employment in agriculture (P_a) had remained 72 per cent as in 1870 rather than declining to 23 per cent in 1950, how would inequality have evolved if this share had been fixed? Or, if the relative incomes of agriculture and industry (μ_i, μ_a) had been frozen? We calculate three counterfactual Gini series, G^{CF1}, G^{CF2}, and G^{CF3}, capturing a scenario with fixed mean incomes in the sectors, fixed employment shares, and fixed within-sector inequality, respectively. Then, we compare these Gini series with the actual Gini coefficient (G) to assess how changes in mean incomes, population shares, and within-group inequalities affected overall inequality.

The second set of decomposition analyses follow Cowell and Flachaire (2024) and others by using the very practical Mean Log Deviation measure of inequality. It is less intuitive than the top decile share (and more obscure than the Gini) but also allows us to decompose the evolution of inequality by: (a) within-subgroup changes, (b) shifts to more or less unequal subgroups, (c) population shifts between groups with higher or lower incomes, and (d) changes in subgroup means.

A third type of measure, or rather type of information, used throughout is the occupation-wise income and the occupation-wise size of the population. As we have seen in relation to Equation 1, between-sector income differences matter for overall income inequality as G_b, and it is informative to look at the incomes of some important occupations (farmers, manufacturing workers and others) from the various sectors. For overall inequality it is also crucial to know the share of the population that is in various occupations with various mean incomes, for example, breaking down the P_a into its constituent parts: how large a share were farmers, farm labourers, and estate owners; and how did this change over time? It is also part of the Lewis model that the existence of surplus labour in the informal sector implies that industrial wages will stay flat even with productivity growth, yielding a yawning gap between the incomes of capital owners and workers in the industrial sector. We will test this idea by directly measuring the incomes of capital owners and workers in agriculture and industry separately.

As stressed above, our account, while focusing on structural transformation and its role, is not a mono-causal one. We rather posit that the focus on structural transformation is warranted by the fact that while it is clear that pre-tax income distribution saw important changes over the twentieth century (Bozio et al., 2024), the role played by structural transformation and the occupational composition has not been explored in such depth, with such fine-grained data before. We will also use decomposition techniques (of the Gini) to evaluate the importance of capital income for long-run inequality, a factor well known from the previous literature (Piketty, 2014; Milanovic, 2017; Ranaldi and Milanovic, 2022). With our rich data, we can provide the first historical evidence of the correlation of capital income and labour income in the population, thus providing a fuller account of the role of capital in inequality.

3 Sources and data

For our purposes, it is crucial to work with micro data that allow for the consideration of various factors – in particular occupation, sector, sex and region. However, until now and like most other wealthy countries, Sweden has only had micro data available since the 1960s (Björklund and Palme, 2000), with the exception of a few studies of specific towns (Kuuse, 1970b; Gustafsson and Johansson, 2003). To address this, we have invested considerable effort in collecting data for this paper.

The main source that we used to capture Sweden's income structure over a hundred years is the taxation of income and property. The income tax (*inkomstbevillning*) was introduced in 1862. It took 1 per cent of incomes (a flat rate) and encompassed in principle anyone who did not mainly live off farming; farmers instead paid a property tax (*egendomsbevillning*) corresponding to 1 per cent of an assumed 5 per cent yield off their farmland (Hedlund-Nyström, 1972, p. 65; see Online Appendix A for further discussion of the taxation sources). Wage earners with low incomes did not have to pay the *inkomstbevillning*: the income threshold to pay tax was originally an annual income of 400 kronor, which was raised to 500 kronor in 1884, 600 kronor in 1903 and 1200 kronor in the late 1940s (Gustafsson and Johansson, 2003, p. 194). We used the *bevillning* from 1870 to 1920; in 1928 it was transformed into a municipal income tax and this is our main source for 1940 and 1950. The taxation lists contain the following information which we have extracted: municipality, place (town/rural municipality), occupational title, name (from which we code sex), and income of three types: labour, capital, and self-employment. A great strength of the data, in the historical context, is that the tax was implemented in the same way in each municipality, which ensures that income figures are comparable across the entire country. Furthermore, the basic design of the tax was stable over time, even if the tax rate was raised, as we will discuss below.

The original taxation lists are kept at the Swedish National Archives (*Riksarkivet*) in Stockholm and we have sampled 1 per cent of taxpayers in 1870, 1880, 1890, 1900, 1910 and 1920, and 0.5 per cent of taxpayers in 1940 and 1950, in total sampling about 232,000 original tax returns, with the sample size varying between c. 6,000 in 1870 and 1880 to c. 74,000 in 1950. (See Table 1.) Our sampling strategy was that we first randomly selected one fifth of municipalities (see maps in Online Appendix A). Within the municipalities, we sampled every twentieth page of taxpayers, so that we would capture 1 % = 0.20 * 0.05 of taxpayers. In practice, before 1920, when a large share of the population earned too little to pay the tax, many municipalities had fewer than twenty pages of taxpayers, so before 1920, our sample is slightly larger than 1 % of taxpayers, and is closer to 2 %. (Observations are weighted by rural/urban so representativity should not be a concern.) See discussion in Online Appendix A. By 1940 and 1950, a very large share of the population paid the tax, and we judged it sufficient to sample 0.5 per cent of taxpayers: every twentieth page of tax records for every tenth municipality. (0.10*0.05= 0.5 per cent.) Prior to our project, the taxation returns had not been digitalised, with the exception for Stockholm City from 1862 to 1915 (Stadsarkivet, 2025), and so with that exception, we have collected all the taxation data for the paper by visiting the archives, photographing the taxation lists, and entering the information into a database. The database will be available with the replication package for the paper.

For 1930, we use individual-level information from the Population Census, which is unique among Swedish Censuses in that it includes income information, which is also in the process of being digitised by the National Archives (Riksarkivet, 2025). The 1930 sample has been used in previous research, for example to estimate the incomes of return migrants from America (cf. Ejermo et al., 2025). For 1960 and 1970 we use Statistics Sweden's "Born on the 15th" sample which is a random sample of adult Swedes born on the 15th day of a month.[9] This is less rich in contextual variables than our 1870 to 1950 taxpayer data and in many of our analyses we only employ the 1870 to 1950 data.[10]

[9] See https://riksarkivet.se/born-digital.

[10] A notable methodological aspect of our study is that we measure income inequality on the level of the individual. This follows the structure of the taxation lists, where we cannot always identify who co-resides within a

There are three important challenges with the *bevillning* tax data: the tax material's coverage of the population; the fact that the tax did not include stock dividends before 1920; and possible tax evasion. We discuss these important problems here, one at a time.

Table 1: Population, Sample of Taxpayers, and Tax Coverage, 1870–1950

Year	Population (age 16+)	Sample of taxpayers	Estimated no. of taxpayers	Share covered
1870	2,670,135	5,937	282,250	10.6%
1880	2,983,099	6,144	296,652	9.9%
1890	3,144,733	7,536	307,148	9.8%
1900	3,367,264	16,852	483,935	14.4%
1910	3,716,238	18,477	799,544	21.5%
1920	4,061,119	40,661	1,850,564	45.6%
1930	4,509,718	—	2,134,941	47.3%
1940	4,972,663	55,195	2,914,310	58.6%
1950	5,309,045	74,356	3,825,761	72.1%

Notes: Data for 1870–1920 are based on a 1 per cent sample of *bevillning* income taxpayers in Sweden; 1940 and 1950 use a 0.5 per cent sample. The 1930 figures derive from the 1930 Census, digitised by the Swedish National Archives. Population figures refer to individuals aged 16 and above. See Online Appendix A for further discussion of the sample.

The problem of coverage of the population concerns how we can draw conclusions on the overall income distribution when the taxation data does not include the entire distribution. Less than universal coverage of income taxation systems is an eternal issue in the literature on income distribution; in the contemporary United States, about one third of adults do not pay income tax (Splinter, 2019), and in the early twentieth century, the income tax sources used in important top-income studies in many cases covered only about the top 1–5 per cent of the population (Atkinson and Piketty, 2007, Table 13.0). Of course, in these studies and precisely because of the lack of income information for the lower reaches of the distribution, the researchers focused on top incomes and on top income shares. The coverage rate of the income tax that we use is shown in Table 1. Coverage was about 10 per cent of adults in the late 1800s before a rapidly increasing trend began. This trend had two causes: the shift from farming to paid labour (and hence from the property *bevillning* to the income ditto), and that the growth of nominal incomes

household. The difference between inequality measures on the individual level and the household level can be important, and discussion of issues such as demographic impacts on inequality measures are front and centre in research on long-run inequality (e.g. Auten and Splinter, 2024, pp. 2192-2193). To ascertain that our results are not driven by the individual-level design, in Online Appendix C we experiment with varying the population. Especially, in our long-run historical context, it makes a large difference if one counts farmers' wives or not. However, we show that the trends are very much the same over time, also when varying the population. In a paper on Stockholm, where the taxation lists contain richer demographic information, we also show that in that urban context, shifting between household-level and individual-level measures does not change the results (Bengtsson and Molinder, 2024, Appendix B2).

pushed an ever larger share of the population above the income tax threshold. Thus, by 1920 45.6 per cent of the adult population paid the tax, and by 1950, 72.1 per cent. For a historical setting, this is an impressive level of comprehensiveness, but we must still estimate incomes of the non-filers as well.

As discussed by Atkinson (2007), there are two main approaches to estimating the incomes of non-filers in historical inequality studies building on tax data. The first approach is the simpler one: to impute homogenous incomes below the tax payment threshold for all non-filers, while calibrating the resulting total income (incomes of taxpayers plus imputed incomes of non-filers) to an existing income total from the national accounts. This is the top-down-approach and its merits are clear: it is a simple approach which demands little information, the external check with the national accounts data makes sure that the income total is reasonable, and it is indeed likely that the incomes of non-filers have little variation, as by definition they earn less than the tax threshold. So for example in our case, we know that, bar the farmers, all non-filers in 1870 and 1880 had incomes y bounded as $0 < y < 400$ kronor, and in 1900, the incomes were bounded as $0 < y < 600$ kronor. (For reference, Bagge et al., 1933, Table 1, estimated the average annual wage of a male worker in large firms in mining and manufacturing to be 501 kronor in 1870 and 947 kronor in 1900.) These are small sums that most likely do not make much of a difference for the income distribution; thus it is quite uncontroversial to impute homogenous incomes of non-filers or to estimate a parametric distribution among them, within the bounds set by the taxation system (cf. Jenkins, 2017).

We have, however, chosen the more demanding approach, labelled bottom-up by Atkinson (2007) to tackle the issue of non-filers. We begin by comparing the occupational structure of the population, identified using census information, with the occupational structure of the taxpayers: which occupations are missing in the tax data and what is the occupational structure of the non-filers like? The task is then to estimate incomes for all occupations among the non-filers, incomes for farmers based on the property tax and other sources, and for other (wage-earning) non-filers' incomes bounded between 0 kronor and the tax-filing threshold, based on other sources. We impute in this way specific incomes for a large array of sex- and region-specific occupations (large regions: Norrland, Svealand and Götaland). The advantage of the bottom-up approach is the greater precision of the estimates and the possibility to explore the different dimensions of inequality, such as by sector and occupation. To estimate the incomes of the farmers, we have drawn upon a multitude of sources, especially detailed household budget studies from the 1910s, 1920s, 1930s and 1940s.[11] See Online Appendix B for how we calculated incomes for farmers and workers below the tax threshold. Let us also emphasise that, as with the top-down approach,

[11] A conceptual comment on what "income" is and what we want to measure when we measure income inequality, is in order. The approach here is to include subsistence production, which was not valued in money, not transferred via the market and so did not take the form of an income stream. For our understanding of living standards, the wider measure is more relevant, and it becomes especially important in a society with a large agricultural sector, which Sweden was in our period. See also Lindahl et al. (1937, p. 10). For further discussion of the fit between our imputed incomes, the taxed incomes, and the national accounts, see Online Appendix B.4.

we have also estimated the national accounting post of ”personal income”, as the income of all adult persons, constituting a combination of the incomes of the taxpayers and the (estimated) incomes of the non-filers. We have compared this personal income in detail to the personal income from Historical National Accounts (see Online Appendix B.4) and ascertained that our income estimates do not deviate from the national accounts totals. Thus we are quite certain that our estimated/imputed incomes for non-filers are robust. We also show in Online Appendix C.2 that the aggregate inequality estimates are robust in relation to a switch to a top-down approach.

The second challenge with the *bevillning* tax data is the coverage of capital incomes. Before 1920, stock dividends paid out to individuals were not covered by the income *bevillning* tax; the tax system taxed corporations alongside individuals, and to tax dividends was considered double taxation.[12] Since dividends are always concentrated to the top income earners, this means that before 1920 we have a downward bias in our inequality estimates. Building on financial history sources, we have estimated the extent of this bias (Online Appendix A.4) . In 1870, 1880 and 1890 there were few stock companies, and total dividends in 1870 corresponded only to 1.2 per cent of GDP. By 1900 and 1910, however, dividends were significantly larger, reaching 5 per cent of GDP in 1910. The bias to our inequality estimates is negligble in 1870, 1880 and 1890, but in 1900 and 1910, our baseline estimates underestimate the top decile’s share of income by 2–3 percentage points, and the Gini by about 2 points. We will return to this factor when we discuss our results below.

The third challenge with the sources is possible tax evasion. In 1870–1900, the tax rate was 1 per cent and flat, so there were no strong incentives to evade taxes. Thereafter, the tax rate grew and became more progressive, causing greater worries about tax evasion. It is well known from the literature on the history of taxation that in the early 1900s Sweden pursued less progressive income tax regimes than in the UK and the US, and it is fair to say that it was only in the 1930s and 1940s that the tax system became significantly progressive (Steinmo, 1989; Henrekson and Stenkula, 2015; Torregrosa Hetland and Sabaté, 2026). For the late nineteenth and early twentieth centuries, there were thus no strong material incentive to evade taxes. On the other hand, incomes in 1870 were estimated by a committee; reporting becoming more stringent over time. From 1883 onwards, all companies were required to hand in information on wages and salaries for all their employees to the taxation authorities, and from 1903 there was self-reporting of income, with checks against other sources. Swedish historians have taken an interest in the precision of the income tax records; one reason for this interest is that before 1909, being taxed for a certain income was necessary to be able to vote for the Riksdag (the Swedish parliament), and the number of votes a person was allotted in local elections was in proportion to their taxed income (Öhngren, 1974, pp. 130–131, Bengtsson, 2019). Thus, there was an interesting trade-

[12] This was also the case in Norway (Berger and Vagle, 2017) and Finland (Heino, 2026). The principle met criticism from contemporaries (Flodström, 1902) and in Sweden the system was reformed in 1920 to include dividends.

off between the good of paying less (or no taxes) and the good of gaining votes and political power. Historians have investigated the source value of the taxation lists in light of this and as a source of incomes and living standards. The results of the investigation of Öhngren (1974, pp. 117–133) for the city of Eskilstuna from the 1870s to the early 1900s and Kuuse (1970b, p. 167) and Olsson (1972, pp. 77–81) for specific groups of workers and white-collar workers for the interwar period are encouraging: in many cases there are small discrepancies between the wages and salaries reported in the employers' papers and the incomes reported by the employees to the taxation authorities. But these discrepancies are minor and should not cause bias in our results. The results from their intensive micro investigations are compatible with the country-wide estimate of Bentzel (1952, pp. 61–62), that about 2 to 7 per cent of incomes were withheld from taxation. We conclude that our taxation data underestimate the incomes of the taxed, as in any taxation system[13], and that this under-estimation most likely pertained to low-paid and high-paid workers alike. Incentives for tax evasion definitely increased over time as taxes became progressive, especially in the 1940s, but the checks by the taxation authorities also improved.

4 Swedish income inequality, 1870–1970

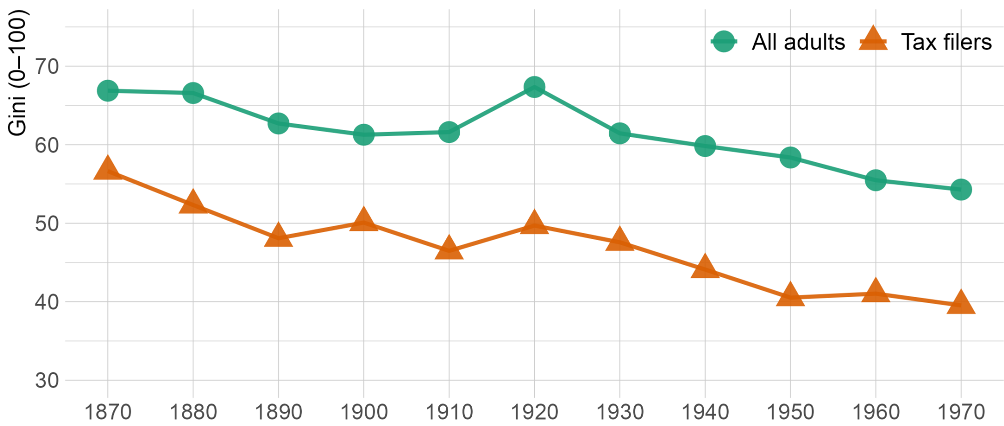
We begin our discussion of the evolution of income inequality in Sweden with Figure 2. The graph plays a dual role. Firstly, it compares our estimates of income concentration to the top decile in comparison with those for the post-1903 period by Roine and Waldenström (2008), which were based on tabulations of a different income tax and with a different method for adjusting for non-filer incomes. The second purpose of the graph is to show the new descriptive insights given by our study. With our richer micro data, we can also calculate the distribution below the top decile.

There are four main takeaways from Figure 2. The first is that our results in the aggregate are quite similar to those of Roine and Waldenström for the top decile's income share, which is the measure that they also provide. Our series displays less volatility than theirs, but the level and trend in the two series concur. This suggests that in the aggregate, our imputed incomes for non-filers are reasonable and do not diverge from actual incomes in a way that would distort our estimates. On this point, see also Figure 1 which shows the difference between our main estimates – in the graph labelled "All adults" – and estimates which only take into account taxpayers. This comparison shows that inequality levels are much lower when we only consider taxpayers, but that the trends are consistent when we take into account non-filers.

The second takeaway from Figure 2 concerns the broad picture of stability and change over time. Broadly, the 1870 to 1970 period falls into two broad sub-periods when it comes to the levels of and trends in inequality. 1870 to 1920 saw high and stable inequality – and recall that adding in dividends (as in Online Appendix A.4), missing in the data used here, raises inequality levels

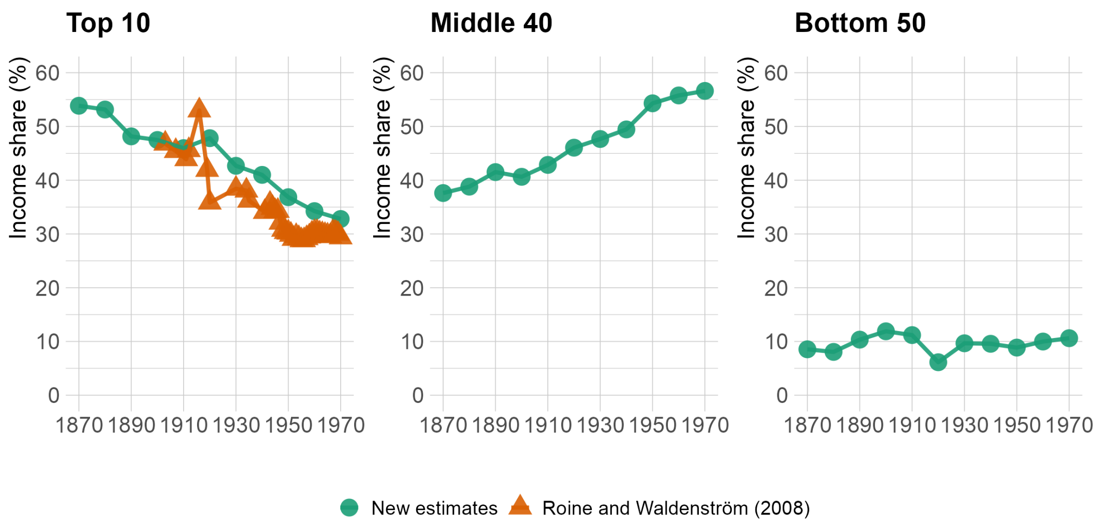
[13]For an engaging debate on the tax evasion in the United States since the 1960s, its frequency across the income distribution and how it influences inequality estimates, see Auten and Splinter (2024) and Piketty et al. (2024).

Figure 1: New Gini estimates, 1870-1970



Note: Data are drawn from the *bevillning* tax dataset for 1870, 1880, 1890, 1900, 1910, 1920, 1940 and 1950, the Census of 1930, and the Statistics Sweden Born-on-the-15th-dataset for 1960 and 1970. See discussion in Section 3 and Online Appendices A and B.

Figure 2: Quantiles' share of income, 1870-1970



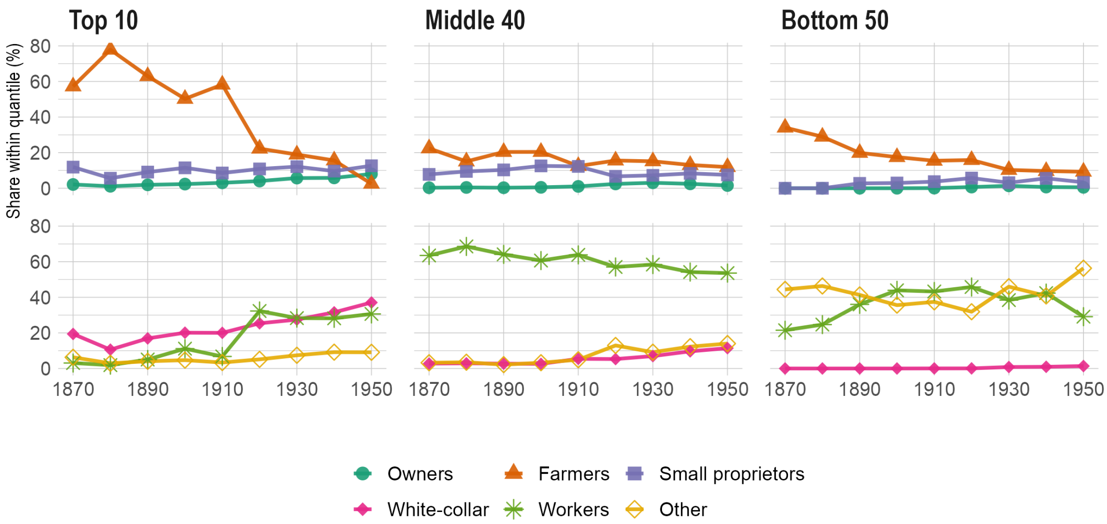
Note: The series labelled "new estimates" are our baseline estimates, based on the micro data constructed for this paper: the *bevillning* tax data adjusted for non-filers; the 1930 Population Census; and Statistics Sweden data for 1960 and 1970. See discussion in Section 3 and Online Appendices A and B. The comparison series is from Roine and Waldenström (2008). The crucial differences between the baseline series and the Roine-Waldenström series are that the R-W series builds on the 1903 *inkomstskatt* income tax, which from 1903 to 1928 existed in parallel to the *bevillning*, and that Roine and Waldenström use a top-down approach, not a bottom-up approach when adjusting for non-filer incomes. See Online Appendix C.2 for a replication of our 1870 and 1900 results, using our micro data but a top-down approach.

in 1900 and 1910, making the lines even more stable. The top decile accounted for about 50 per cent of total income throughout these years. The second sub-period is 1920 to 1970, which saw a quite steady decline in inequality from the high of 1920: to 42.7 per cent in 1930, 41.0 per cent in 1940, and 36.8 per cent in 1950.

The third takeaway is that equalisation was driven in particular by the growing income share of the middle income groups – not by improved shares for the bottom half of the (pre-tax) income distribution. This is also reflected in the difference between the decrease in top income concentration and the decrease in the Gini. The top decile share decreased by 31.6 per cent between 1870 and 1950 (from 53.8 to 36.8 per cent), while the Gini "only" decreased by 12.5 per cent in the same period (from 66.7 to 58.4). This divergence highlights that there was a great deal of redistribution from the top groups to the middle groups – but not any full equalisation across the distribution. As Aaberge (2007, pp. 313–316) points out, the Gini is more sensitive to changes in the bottom than at the top. This means that this measure gives a "conservative" view of the Swedish top-to-middle redistribution process. To further understand this aspect of the changes in the income distribution, Figure 3 reveals the occupational composition of each of the three main income groups: the top, the middle and the bottom. A striking result in this figure is that what in sociological terms were working-class people – industrial workers, construction workers, transport workers and the like – actually made up more than half of the middle-income group. This means that we should be wary of labelling, in a historical context, "middle-income" as "middle class" in a sociological sense. The lack of improvement in the share of the bottom half in Figure 2 should not be interpreted as overall stagnation in working-class incomes, for two reasons. Firstly, the bottom half of the income distribution to a large extent was composed of peripheral groups vis-a-vis the labour market (the category "Other" in Figure 3; see also see Online Appendix Table C.3). Secondly, the stagnating *share* of incomes in Figure 3 means that the average income of the persons in the bottom half of the distribution grew fivefold over the 1870 to 1950 period, i.e. in line with average income growth (as demonstrated in Schön and Krantz, 2015).

The fourth takeaway from Figure 2 concerns the comparative level of income inequality in the early twentieth century. There are few comparable estimates from the this period, and the difference in methodology makes it somewhat difficult to compare with estimates from other countries. However, our estimate for the top decile share of income, which arrives at about 50 per cent in 1890, 1900 and 1910, is rather similar to, and just on the high end of, the estimates for the United States starting in 1917 and for France starting in 1900 (Piketty et al., 2018; Garbinti et al., 2018). (Recall also that our pre-1920 estimates have a downward bias due to the exclusion of dividends.) The most important qualification on this comparison is that we measured inequality at the individual level, while the benchmark series of Piketty et al. (2018) and Garbinti et al. (2018) uses an equal-split assumption for spouses who file taxes jointly. The purely individual-type estimate yields higher inequality: in the US case the top decile's income

Figure 3: Occupational composition of income quantiles, 1870-1950



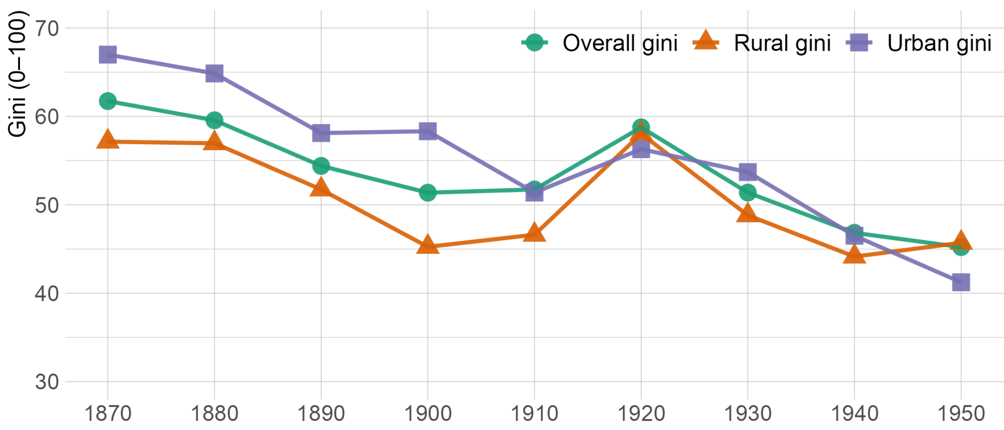
Note: Data are drawn from the *bevillning* tax dataset for 1870–1920, 1940 and 1950, and the 1930 Population Census. See discussion in Section 3 and Online Appendices A and B. On the occupational classifications see Online Appendix A.3.

share is about five percentage points higher using the individual-level measure (Piketty et al., 2018, pp. 590–594). When we take this into account, we judge the inequality level in early twentieth-century Sweden to have been very similar to that in France and in the United States. This result accords with the previous research on income inequality by Roine and Waldenström (2008), with different sources and methods, and also with their research on wealth inequality, which also gives a picture of a quite unequal economy in the late nineteenth and early twentieth centuries (Roine and Waldenström, 2009). In Online Appendix C.1 we present alternate estimates with an equal-split assumption and in a version without farmers' wives; the estimated inequality levels are lower in these versions, but the trends are essentially the same.

5 Structural transformation and Swedish income inequality

The previous analysis has revealed that the concentration of income was very high in Sweden between 1870 and 1920 and began to decline at an increasing rate after 1920, before the ravages of the Depression and World War Two – facts that have often been pointed to as drivers of twentieth-century equalisation (Scheidel, 2017), in particular in the literature on the United States (Geloso et al., 2022). To analyse what was going on in Sweden, we will now move to a more comprehensive inequality measure, the Gini coefficient, and explore how structural transformation shaped its course over time. Figure 4 illustrates the $G_{\mathrm{wi}} > G_{\mathrm{wa}}$ mechanism suggested by Kuznets, focusing

Figure 4: Gini comparison

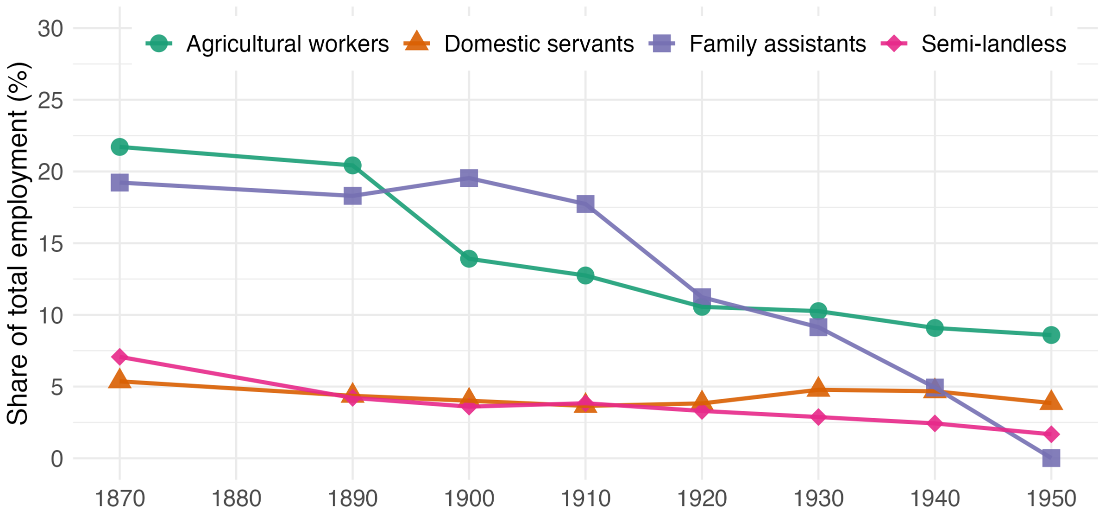


Note: Data are drawn from the *bevillning* tax dataset for 1870–1920, 1940 and 1950, and the 1930 Population Census. The sample includes employed individuals, excluding married women; alternative specifications are reported in Online Appendix C.1. The y-axis is truncated to improve readability.

on the urban-to-rural distinction. The Gini coefficient was indeed higher in the urban sector than in the rural sector, but within-sector inequality fell steeply over time. At the same time, the labour force shifted into industry and into urban sectors generally. In our data, the urban share of employment grows from 14 per cent in 1870 to 24 per cent in 1900, 34 per cent in 1920, and 61 per cent in 1950. The typical income in the urban sector was also higher than that in rural areas, as we would expect: median income in the urban sector was 57 per cent higher in 1870, 40 per cent higher in 1900, and still 40 per cent higher in 1950. The complexity of the process, where within-sector inequality, G_b, and the relative weights of the sectors all change at the same time, motivates the decomposition approach. Another important distinction to consider here is that urban $\neq$ industrial and rural $\neq$ agricultural. For that reason, in Online Appendix C (Figure C.4) we also calculate the Gini coefficient by sector of employment: agriculture, industry, or services. This investigation highlights that income inequality in agriculture throughout the 1870 to 1910 period actually exceeded that of industry; levels around 50 versus levels around 40. This again gives cause to question both the Kuznetsian focus on the $G_{\text{wi}} > G_{\text{wa}}$ dimension of structural transformation and inequality, and to go deeper into the structure of agriculture: how could this sector be so unequal?

Here it is worthwhile to revisit the distinction between proximate and fundamental causes introduced in Section 2. If, as North and Thomas (1973, p. 2) argued, innovation, economies of scale, education, and capital accumulation "are not causes of growth; they *are* growth", then we can equally say that the occupational structure and its pay ratios are not a cause of changes in

Figure 5: The decline in low income groups in agriculture



Note: Measure is % of all employment. Our estimates are based on Census information on occupations. See the methodological discussion in Molinder (2023). The group family assistants in agriculture is missing in 1950 because of a different Census classification for adult children living at home in this year.

inequality, it *is* changes in inequality. (See also Shorrocks, 1984.) Following this, the most important fact to understand regarding the surprising degree of inequality in the rural sector in the late nineteenth century and its ongoing decline is that Swedish agriculture underwent a massive social transformation. In the 1870s and 1880s a very large share of the agricultural population were landless workers, either wage labourers, semi-landless contract labourers, or unpaid family labourers. Over the following 50 years or so this would change. The number of landowning farmers *grew* from 1870 to 1920 (Morell, 2001, pp. 76-81), while the number of labourers in agriculture fell drastically. This process is illustrated in Figure 5, which shows the share of Sweden's total employment occupied by four low-paid groups, mostly in agriculture: agricultural workers, unpaid family assistants, domestic servants, and semi-landless contract labourers (*torpare*). The decline in agricultural labourers and unpaid family assistants in particular is massive: from more than 20 per cent in 1870 to below 10 per cent in 1940 and 1950 for the labourers, and from around 20 per cent to less than 5 per cent for the family assistants. In 1870 a male agricultural worker earned around 300 kronor a year, a female agricultural worker around 150 kronor, and the farmer of a mid-size farm about 1250 kronor (see Online Appendices B.1 and B.2), so this was a process of declining low-paid groups and stable employment of the relatively well-paid groups.

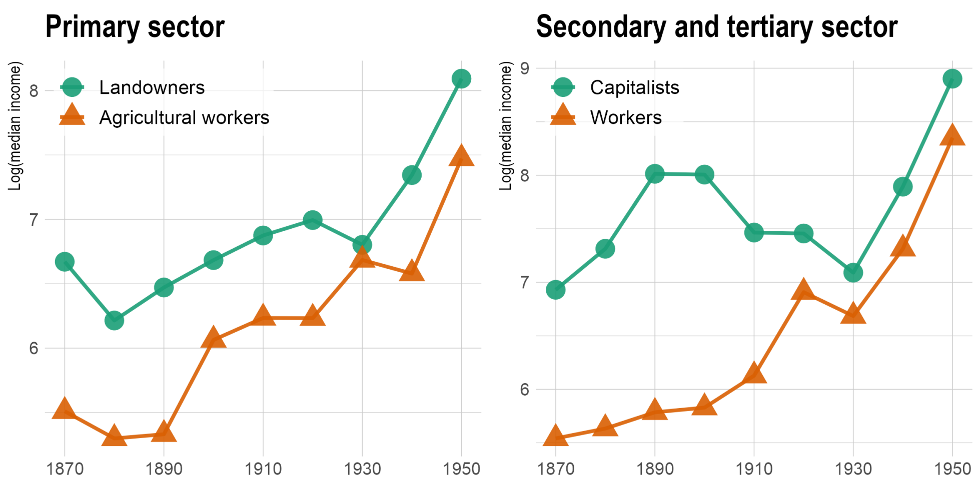
This drastic development calls for an answer to the question: how could this structural transformation happen so smoothly in a largely agrarian country like Sweden in 1870? This question leads us away from the proximate causes of changing inequality to fundamental causes (Acemoglu et al., 2014), and a full answer would require other types of data and methods than the ones used here. But we will just say a few words on this issue, possibly pointing the way for fruitful further research. The macro development literature studying contemporary economies like India offers ample evidence that workers tend to get stuck in less productive occupations because of market imperfections: limits to workers' mobility to better jobs and sectors, as well as poorly functioning capital markets restricting the growth of the more productive companies (Foster and Rosenzweig, 2022; Alvarez-Cuadrado and Poschke, 2011; Manuelli and Seshadri, 2014; Gollin and Kaboski, 2023). In contrast, it appears that the Swedish labour force was highly mobile during the period studied here. This presupposes not only relatively unrestricted movement for workers themselves but also that productive firms can access capital and expand and hire more workers – and that the agricultural firms (i.e., farms) that saw a massive outflux of labourers in the post-1870 period could respond by mechanising production, which also presupposes access to capital markets.

The preliminary evidence for such an interpretation is as follows. Decreasing wage differentials across regions and labour markets indicate the formation of a unified labour market in the late nineteenth and early twentieth century (Enflo et al., 2019; Collin et al., 2019). Massive mobility of workers has been observed in the literature on emigration to America and on internal migration (Bohlin and Eurenius, 2010; Andersson, 2024), and research employing record linkage to follow individuals over time indicates that migrants benefited greatly, from emigration to the United States (Castillo, 2024) and from a more modest internal migration from the countryside to the city (Andersson, 2024). This indicates that for the rural labour force, it was not their own low skills that held them back (cf. discussion in Gollin and Kaboski, 2023, pp. 351–352): given opportunity, they could flourish. Universal primary schooling was most likely a factor here (Jayes et al., 2024).

The productive restructuring of agriculture is evinced in the first instance by the rapidly growing output per working person, clear from the historical national accounts: by 27 per cent from 1870 to 1900, and by 59 per cent from 1900 to 1930 (Schön and Krantz, 2015, Tables III and VIII). A crucial aspect of growing productivity was mechanisation; threshers, originally drawn by horses, saw a breakthrough on larger farms as early as in the 1870s, and from the 1870s to the 1940s, most other sub-operations became more capital-intensive (Kuuse, 1970a, Morell, 2001, pp. 267–288). This is not the place for a rigorous analysis of cause and effect between low-paid workers leaving agriculture and the mechanisation of production (as in Ager et al., 2026); causality was likely two-way: an induced dearth of labour making mechanisation more attractive, and proactive mechanisation replacing labourers.[14] Mechanisation was likely associated with scale effects, and so it would have helped that the market for land was dynamic (Bohlin and Prado, 2011; Dackling,

[14] There is anecdotal evidence for the induced dearth of labour–mechanisation connection (as in Hornbeck and Naidu, 2014). The American consul in Gothenburg in 1903 reported back to the USA about the possibilities to sell agricultural machines to Sweden, argueing that there was now a Swedish market for such products because "wages in agriculture have risen so quickly, that farmers complain that they can impossibly pay the wages demanded. This rise in wages should depend partly on migration to cities, and partly on emigration, which decreases the supply of labour in agriculture..." (cited from Källvemark, 1972, p. 127).

Figure 6: The incomes of capitalists and workers, agricultural and industrial sectors, 1870–1950



Note: The graph reports the log of median nominal income within each group: unskilled workers within agriculture (left) and industry (right), and large owners (landowners left, capitalists right). Data are drawn from the *bevillning* tax dataset for 1870–1920, 1940, and 1950 and the 1930 Population Census. See discussion in Section 3 and Online Appendices A and B.

2013), and farmers who intended to expand could draw upon a dense net of institutional and informal loan providers (Erikson, 2018), unlike many development contexts today (Karlan and Morduch, 2010). At the same time, the growth of industrial firms was facilitated by loans from the modernising banking sector and capital imports from France, Germany and Britain (Schön, 2014).

In the transformation of agriculture, we have so far focused on the decline in the low-paid groups. But what about pay ratios and wage growth? As discussed in Section 2, in a Lewis (1954) economy, a persistent surplus of labour ensures that even when labour productivity grows in the industrial sector, the wages of unskilled workers stay flat, as they can always be replaced by greater influx from the informal or agricultural sector. Such a flattening of workers' wages did not happen in Sweden. Figure 6 shows that, according to our data both agriculture and industry, saw rising wages in the decades following 1870, in line with findings using other data and methods (Lundh and Prado, 2015; Ericsson and Molinder, 2020). As a result, the capitalist-to-worker income ratio did not increase as the Lewis model would predict.

With the data and methods used here, it is beyond the scope of this paper to provide a fundamental explanation of this wage development. What we can say is that Swedish industrialisation advanced from the 1870s to the 1950s (Schön, 2014), increasing capital intensity and energy intensity (especially hydropower), and great demand for labour (Lundh, 2010). The picture in Figure 6 where the growth of workers' wages keeps pace with or outpaces the growth of compensation for capital owners is compatible with an interpretation of Fordist economic growth with voracious demand for low-skilled labour. (On Fordist growth and its connection to income inequality, see Goldin and Katz, 2009, pp. 114-125 on the US in the twentieth century.) However, further research is needed on the determination of wages and wage growth in this period (cf. Skoglund, 2025b,a).

6 Counterfactuals and decompositions

With these building blocks in place – the changing structure of the agricultural sector and the transfer of labour from agriculture to industry – we can now proceed to our counterfactual analysis and a more ambitious decomposition. We know that the 1870 to 1950 period saw a large shift in the population from the rural to the urban setting, at the same time as sectoral incomes grew at various speeds and within-sector inequality declined, as we saw in Figure 4. But which of these (related) processes was quantitatively most important for the changes in overall inequality? In this section, we address this question using two complementary methods: a counterfactual analysis and a decomposition of inequality. While the counterfactuals isolate the contribution of broad factors by holding them fixed, the decomposition provides an exact account of the channels through which these factors operate.

We begin with the counterfactual approach, where we simulate counterfactual Gini coefficients in three scenarios. The first counterfactual (CF1) is that we fix the mean incomes of the urban and the rural sector, μ_k, to their 1870 levels, simulating a scenario where industrialisation and urbanisation proceeds as it actually did in terms of the percentage of the population, but without income growth. Step one in this counterfactual is to recalculate income shares:

$$s_k = \frac{N_k^{\mu_k}}{N^{\mu}} \tag{7}$$

where

$$\mu = \frac{\sum_k N_k \mu_k}{N} \tag{8}$$

Step two is to recalculate weights:

$$w_k = s_k p_k \tag{9}$$

Step three is to recalculate within-group inequality:

$$G_w^{\text{CF1}} = \sum_k^{w_k} G_k \tag{10}$$

And step four is to recalculate between-group inequality:

$$G_b^{CF1} = Gini(\mu_1, \mu_2, \ldots, \mu_K) \tag{11}$$

The next counterfactual (CF2) is a scenario where urbanisation does not progress after 1870 but rather the urban and rural sectors' respective shares of the population, P_k, are fixed to their 1870 levels. Step one in this counterfactual is to recalculate the weights:

$$w_k = S_k P_k \tag{12}$$

Step two is to recalculate within-group inequality:

$$G_w^{CF2} = \sum_k^{w_k} G_k \tag{13}$$

And the third step of the counterfactual is that between-group inequality remains unchanged, $G_b^{CF2} = G_b$.[15] The third counterfactual scenario (CF3) is one where the mean incomes of the two sectors grow as they actually did, and urbanisation progresses as it actually did, but where within-sector inequality, G_k, is fixed at its 1870 levels. Step one of this counterfactual is to recalculate within-group inequality:

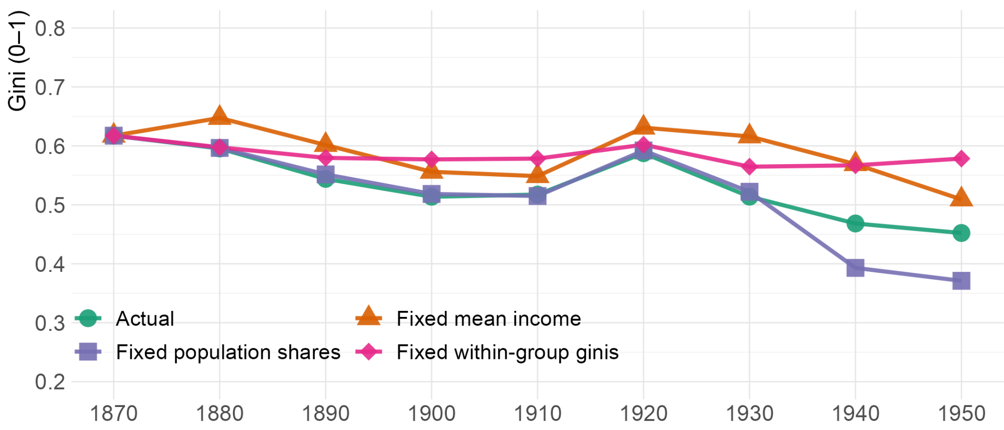
$$G_w^{CF3} = \sum_k^{w_k} G_k \tag{14}$$

At the same time, between-group inequality remains unchanged, $G_b^{CF3} = G_b$.

By comparing the counterfactual Gini coefficients G^{CF1}, G^{CF2}, and G^{CF3} to the actual Gini coefficient G, we assess how changes in mean incomes, population shares, and within-group inequalities affect overall inequality. Figure 7 shows the actual Gini G next to the three counterfactuals. The greater the divergence is between counterfactual and actual Gini, the greater the importance of the factor held constant. The difference between the green line with circles, G, and the pink line with diamonds, G^{CF3}, shows that the actual decrease in *within-sector inequality* made the largest contribution to the overall decline in inequality. Without within-sector changes, but with the actual movement of labour to industry and services and with the actual change in between-sector income difference, pre-tax income inequality would not have decreased much in Sweden from 1870 to 1950. This is rather surprising, and we interpret the lack of an important

[15] Since the Gini coefficient is not additively decomposable, the counterfactual exercises—particularly CF3—should be interpreted as accounting decompositions rather than exact counterfactual income distributions. In CF3, fixing within-group Gini coefficients isolates the contribution of changes in within-group dispersion, but does not uniquely define a corresponding income distribution.

Figure 7: Gini coefficients of income in Sweden, 1870–1950: actual and in counterfactual scenarios



Note: Data are drawn from the *bevillning* tax dataset for 1870, 1880, 1890, 1900, 1910, 1920, 1940 and 1950, and the Census of 1930. See discussion in Section 3 and Online Appendices A and B. Note that the decomposition excludes the undercounted employed women that we include in our Gini calculation for the full population, since we are not able to allocate them between rural and urban areas. For an explanation of how the counterfactuals have been constructed, see the text and equations 7 through 14.

between-sector component of inequality reduction as evidence of a high degree of factor market integration, which we would suggest was due to a combination of competitive factor markets and a productive and fast-growing manufacturing industry. Such an interpretation is compatible with the strong wage growth both in agriculture and industry as shown previously in Figure 6.

While the counterfactual analysis provides a shift-share decomposition, isolating one factor at a time, we will now move to an alternative decomposition, that of the mean log deviation (MLD) of income, following Mookherjee and Shorrocks (1982). The MLD is not as intuitive in its levels (it is bounded to zero at the bottom but has no upper bound) but it is eminently decomposable (Cowell and Flachaire, 2024; see Belfield et al., 2017 for an application to contemporary Britain). With this approach, we decompose the change in income inequality between two points in time with respect to three factors: class, sector, and the urban/rural dimension.

To operationalise class, we divide the population into three groups: capitalist owners and working owners; salary earners; and wage earners. This corresponds clearly to one of the classic dimensions of income inequality: the difference between capitalists and workers (as already discussed in Figure 6), or, often discussed in research on the twentieth century, differences between salaried white-collar workers and blue-collar wage earners (e.g. Piketty, 2014; Lindert and Williamson, 2016). For sector, we use 18 sectors: two within the primary sector (agriculture, fishing), nine within industry (e.g., metal and manufacturing; textiles) and seven within the tertiary sector (e.g., trade; public service; domestic service). (For a detailed discussion of the sectoral classification, see Online Appendix A.3.) This decomposition helps to analyse some of the classical questions about structural transformation raised in Section 2: in a period of rapid economic growth and industrialisation, how did the changing sectoral makeup of the economy shape income distribution? The urban–rural decomposition builds on a binary division by residence and thus differs from the sectoral decomposition, as there were plenty of industrial and service sector jobs also in rural areas. (See the discussion of Figure 4 above.)

Regarding each factor, we decompose changes in inequality into four components. The first component is the change due to trajectories of within-group inequality (Equation 15a). Did inequality within the groups – be it class, employees by sector, or income-earners by urban/rural residence – increase or decline? The second component is the change due to population share shifts between groups with more or less within-group inequality in the base period. Do groups that within themselves are more unequal grow more than the less unequal groups? (Equation 15b) The third component is the change due to population share shifts between groups with higher and lower average incomes in the base period. If groups with mean incomes closer to the population mean grow more (less) than groups with incomes farther away from the population mean, then inequality will decline (increase) (Equation 15c); we can think of this as a process where the distribution becomes more (less) cluttered around a modal point. The fourth component is change due to shifts in the groups' mean incomes (Equation 15d). This captures increasing or decreasing income differences between groups, for example the gender income gap or the income advantage of industry compared to agriculture ($\mu_i > \mu_a$) that Kuznets pointed to.

The four equations are the following:

$$\sum_k \upsilon_k \, \Delta I_0^k = \text{within subgroup inequality} \tag{15a}$$

$$\sum_k \bar{I}_0^k \, \Delta \upsilon_k = \text{population share effect on within component} \tag{15b}$$

$$\sum_k \left(\bar{\lambda}_k - \overline{\log \lambda_k}\right) \Delta \upsilon_k = \text{population share effect on between component} \tag{15c}$$

$$\sum_k \left(\bar{\theta}_k - \bar{\upsilon}_k\right) \Delta \log \mu_k = \text{change in subgroup means} \tag{15d}$$

where $\upsilon_k = \frac{n_k}{n}$ is the population share of subgroup k, $\lambda_k = \frac{\mu_k}{\mu}$ is the relative mean income of subgroup k, and $\theta_k = \upsilon_k \cdot \lambda_k$ is the income share of subgroup k. The sum of the four components aggregates to the total inequality change, in terms of the Mean Log Deviation.

We have performed this analysis for two sub-periods, splitting our detailed sample into two sub-periods of equal length: 1870–1910 and 1910–1950. The first sub-period comes from the period of high and relatively stable inequality evinced by Figures 1 and 2. The second encompasses the post-1920 period of equalisation of incomes, ending with the markedly more egalitarian distribution of 1950. Were the driving forces – contradictory or complementary – of changes in the distribution the same or different in these two epochs of quite different macro circumstances? (See also Online Appendix C for an alternative calculation with the sub-periods 1870–1900 and 1920–1950. The results are essentially similar.)

The results are shown in Figure 8.[16] The Figure maps equation 15a to the green bars, 15b to orange, 15c to purple, and 15d to pink. So, for example, in the urban/rural decomposition, the orange bar (representing equation 15b), and the purple bar (representing equation 15c), represent the two mechanisms of Kuznets' argument on increasing inequality during industrialisation: a shift to the more unequal sector (corresponding to $G_{\mathrm{wi}} > G_{\mathrm{wa}}$ in our discussion of equation 1), and the growing pertinence of the urban-rural average income difference (corresponding to G_b in equation 1).

The 1870 to 1910 period sees a decline along class lines. The period saw growing inequality within classes (green bar) but this was in Figure 8 more than counteracted by a population shift towards the more "average" groups of farmers and salary earners (lilac bar), with small contributions by decreasing income differentials between groups (pink bar) and a population shift to less unequal groups (orange bar). However, we must remember that the data here overstate the equalisation during the period, as we are missing dividend incomes for all pre-1920 years. In Online Appendix A.4 we show that dividends between 1870 and 1910 grew from about 1 per cent of personal income to about 6 per cent. That they are missing before 1920 means that here we are underestimating in particular inequality between classes, as dividends were likely concentrated to capital owners and salary earners. We cannot put exact numbers on the dividends-related bias here, but it seems reasonable that, considering dividends, the 1870–1910 period probably saw less decline in class-related inequality than Panel A of Figure 8 suggests, but still some decline.

The sectoral picture for the 1870 to 1910 period suggests that sector-related inequality declined, because of changes in within-sector inequality as well as an employment shift towards less unequal sectors (recall the unexpected result in Figure C.4, that inequality was markedly high in agriculture) with some contribution from decreasing differences in average income by sector. This speaks against a Kuznetsian (1955) interpretation of this period of high industrialisation

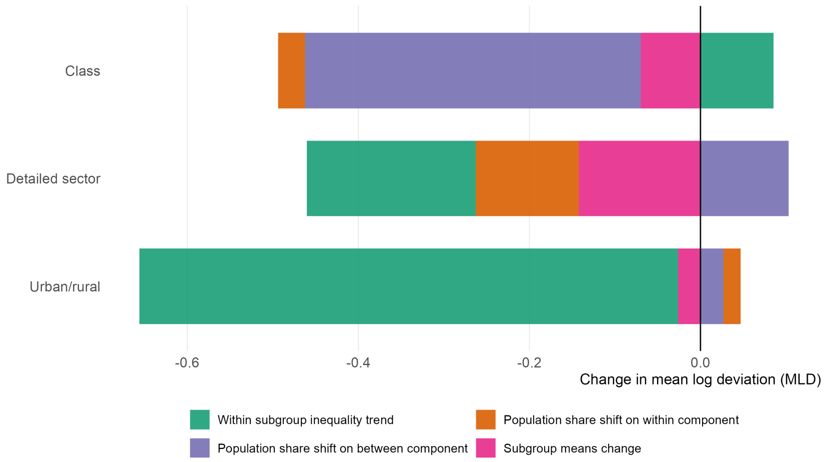
[16] There are some limitations to the analysis that we can do. First, for the rural/urban distinction we can only analyse inequality among the employed, since we do not have information on the rural or urban residency of wives (not employed or estimated to be employed). Second, to make the decomposition, all groups must be present in both years. For class and sector, however, some groups disappear over time. In those cases, we have had to impute one observation for each missing group. For this reason, the totals for the inequality change are not exactly the same across all group classifications. We also exclude widows from the analysis (`st_class` = 13), since this group drives the change in the MLD, which is highly sensitive to the bottom of the distribution. The MLD computed on this restricted sample is reported alongside the full-population MLD in the "MLD (decomp.)" column of Tables C.1 and C.2, and the difference between the two series is discussed in Online Appendix C.6.

in Sweden (Schön, 2014): following Kuznets, sectoral inequality should have *grown* during these years, both along the employment share to more unequal sectors (orange bar) and growing income differences between sectors (pink bar). The Kuznetsian effects largely fail to materialise in Sweden.

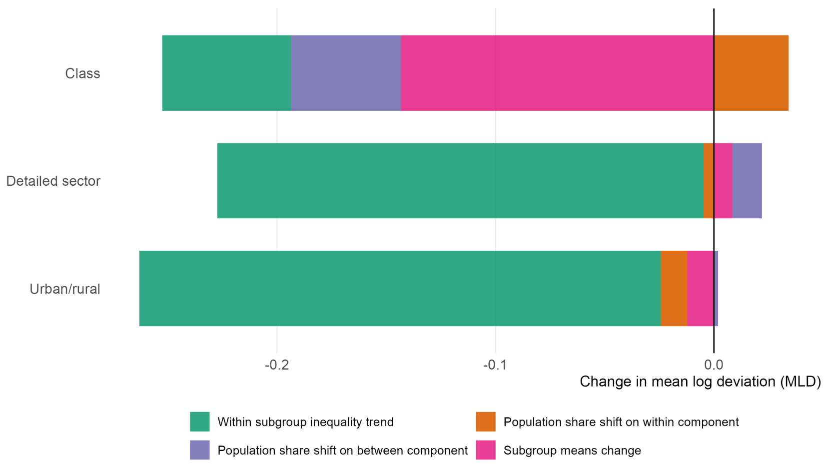
However, the most striking result – note the different x axes for panels A and B of Figure 8 – for the 1870 to 1910 period is the great decline in inequality by the urban/rural dimension, almost entirely driven by declining within-group inequality. This, we would argue, again speaks to the implications of structural transformation for income inequality, and especially the upgrading of the workforce with the drastic decline in the number of agricultural workers and semi-landless workers (Figure 5). Just as the counterfactuals in Figure 7 have suggested, for the urban-rural dimension, the Kuznets effects (workers moving to more unequal sectors with higher income) are drowned out by the within-sector effect (green bar), with a small negative role for inequality played by a decreasing difference between urban and rural incomes (pink bar).

In the equalisation period from 1910 to 1950, changes in the various dimensions overall are smaller than in the 1870–1910 period: the maximum of the x-axis is now about -0.25 rather than -0.6 as in the previous period. Along class lines inequality fell, driven above all by declining income differences between classes (pink bar) but also with contributions from within-class inequality (green bar) and an employment shift to groups close to the population mean (lilac bar). Along the sectoral dimension, inequality declined above all through the within-sector mechanism. The between-classes redistribution and the within-sectoral equalisation observed here are surely related to the strong wage growth of this period, observed in agricultural as well as industrial and service sectors (Figure 6; see also Skoglund, 2025b). This evolution is in evidence also in the decomposition along the urban/rural distinction, where the reduction in inequality is driven almost completely by decreasing within-rural and within-urban inequality. Thus, the Mookherjee-Shorrocks decomposition of the mean log deviation strengthens the argument made from the counterfactual analysis: in 1870 inequality was high in agriculture as well as in industry, and it was decreasing inequality within the sectors, rather than shifts between the sectors, that drove the equalisation of incomes. The Kuznetsian framework thus does not take us far in understanding the evolution of inequality in Sweden.

Figure 8: A Mookherjee-Shorrocks decomposition of inequality, 1870–1910 and 1910–1950



(a) 1870–1910



(b) 1910–1950

Note: Data are drawn from the *bevillning* tax dataset for 1870, 1880, 1890, 1900, 1910, 1920, 1940 and 1950, and the Census of 1930. See discussion in Section 3 and Online Appendices A and B. The bars correspond to the contribution of each factor as identified in Equation 15. The underlying sample drops the widow class `st_class` = 13; the urban/rural panel additionally keeps only the employed (`st_marwom` = 1), because rural/urban residency is not known for non-employed wives. The MLD computed on this restricted sample – to which the decomposition in Figure 8 refers – is reported in the "MLD (decomp.)" column of Tables C.1 and C.2, alongside the full-population MLD.

7 Generalizing from the Swedish trajectory

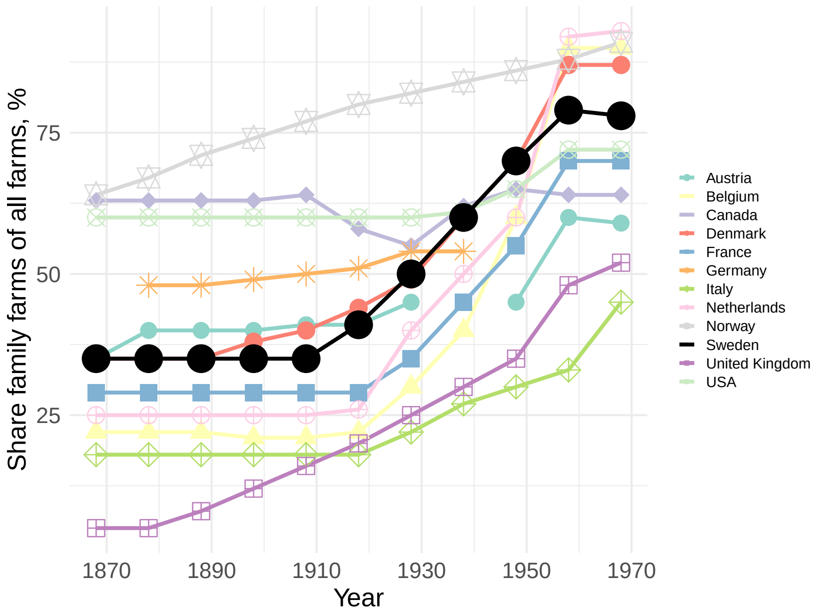
Our account of the evolution of income distribution in Sweden has demonstrated the importance of structural transformation: the removal of low-paid jobs especially in agriculture and the great reduction in income inequality within sectors that ensued. Does this account have generalisability, the ability to travel to other national contexts? We offer here two pieces of evidence of cross-national generalizability: the role of agriculture in GDP, and a crude measure of inequality within the agricultural sector. The logic is as follows: we have demonstrated that inequality, measured properly, was an extremely unequal sector in Sweden – defying Kuznetsian (1955) expectations – and its restructuring, from a large and low-paid (or even unpaid) labour force to a smaller cadre of farmers working in a highly mechanised sector, made an important contribution to the decrease in inequality. (More specifically, it was a proximate cause of the inequality decline.) Thus, it is important to know whether there is reason to believe that Sweden had a uniquely outsized agricultural sector in the late 1800s and early 1900s, and/or a uniquely unequal agricultural sector. If not, the mechanisms or proximate causes found here, with a novel methodological approach, are likely to have some purchase also in other contexts.

To address the first question, we look at the agricultural sector's share of GDP. In 1900, the contribution of agriculture to GDP was 27.6 per cent in Sweden (Schön and Krantz, 2015, Table III) 29.9 per cent in Germany (Hoffmann, 1965), 30 per cent in Denmark (Henriksen, 2009, Table 5.10), 25.7 per cent in France (Vivier, 2009, Table 8.7), and 40.1 per cent in Portugal (Lains, 2009, Table 13.2). Thus, Sweden places itself firmly in the middle of the pack as regards to the role of agriculture in the turn-of-the-twentieth century economy.[17]

The second aspect is whether the Swedish agricultural sector in the late 1800s and early 1900s was uniquely unequal. Here we may begin to note that on the contrary, a common view within historical political economy is that Sweden was a relatively equal agrarian economy in the nineteenth century. In influential political science studies (Ansell and Samuels, 2010; Albertus, 2017), the stable Swedish democratisation is partly explained by precisely the legacy of an egalitarian social structure in agriculture, with pre-dominantly family-owned farms. It is very difficult to find comparable detailed data on historical agrarian inequality across countries, and comparative political scientists tend to use Vanhanen's (1997; 2003) dataset which, for a large set of countries, provides a crude inequality measure. The measure of Vanhanen is the share of units in agriculture held by owner-occupiers ("family farmers"); the logic is that agrarian economies with a large share of family farms are identified as more homogenous and equal, while economies with fewer family farms and a larger share of manors and estates are identified as more unequal.

[17] As labour productivity was lower in agriculture than in other sectors, employment shares of agriculture were higher. Around 1910, agriculture's share of employment was 31.8 per cent in the US, 44.1 per cent in the UK, and 37.9 per cent in Germany (Broadberry, 2009, Table 3.1). Overall, Broadberry (2009) estimates the shares for West Europe as 39.5 per cent and for South Europe 51.6 per cent. In Sweden, the share in 1910 was 47.8 per cent (in 1900: 52.3 per cent) (Schön and Krantz, 2015, Table VIII). Historical national accounting is difficult, and comparisons are complicated, but the employment data broadly support that Sweden was no outlier in Europe in terms of agriculture's weight in the economy.

Figure 9: Share of family farms in agriculture in twelve countries, 1868–1968



Note: Source is Vanhanen (2003). The variable is "family farms (%)", labelled q2 11 for 1868, q3 11 for 1878, and so on.

Figure 9 shows this measure, for Sweden and a large set of comparable countries, throughout our time period. According to this data, in the late nineteenth and early twentieth century Sweden was less dominated by family farmers than the US and Norway, the same as Denmark, and more so than France, Belgium, the UK and Argentina. The middling position of Sweden indicates that the high level of agrarian inequality found here – with much richer and more fine-grained data than the crude measure used by Vanhanen – is likely to be found also in comparable countries, should the data exist. (Indeed, Modalsli, 2018, using a dataset of rare quality for Norway in the late 1860s, found very high inequality there too.)

Taken together, the information in Figure 9 and the share of agriculture in GDP suggests that our focus on agriculture and structural transformation also has a bearing on other countries: it is not the case that Sweden in the late 1800s and early 1900s was an exceptionally agricultural economy, or had an exceptionally unequal agricultural sector.[18]

So far, the discussion in this section has been quite focused on what we have called, following North and Thomas (1973) and Acemoglu et al. (2014), "proximate" causes. One may ask whether there were "fundamental" causes shaping Swedish income inequality over the time period that have been out of view here and that would make the Swedish case unique. As discussed before, we

[18]Furthermore, Chakravorty et al. (2019) document that income inequality in India's agricultural sector today is quite high, a Gini coefficient around 0.6, and that this depends on inequality of ownership. This result indicates that a perspective focusing on agriculture and structural change is relevant also for contemporary societies with significant shares of the population employed in agriculture.

cannot test fundamental explanations with our data, but to place the Swedish case in context, we will just say a few words about our interpretation and how it relates to previous research. While our decompositions, counterfactuals and accounting exercises have been quite focused on the market side of Swedish social transformation after 1870, another body of literature on Sweden puts a much greater emphasis on Social Democracy, a generous welfare state and ambitious redistribution (e.g. Piketty, 2020, pp. 185-189). It has also been shown that Sweden had a greater equalisation of incomes than other industrialised economies over the twentieth century (Roine and Waldenström, 2008; Piketty, 2020). Obviously, such a unique feature cannot be explained by a factor – structural transformation – common to all countries. If structural transformation – industrialisation, the mechanisation of agriculture, the growth of the service sector – was all there was to it, then other European and American economies that industrialised in the decades after 1870 should have had a similar development in terms of inequality. To some degree we think so, but not wholly. To start with, we only analyse pre-tax, pre-benefits incomes, and from the 1940s onwards, Sweden developed a progressive taxation system (Torregrosa Hetland and Sabaté, 2026), and an increasingly ambitious welfare state. Thus, had we focused more on the post-1950 period and on the distribution of post-tax incomes, then political factors and the welfare state would have earned a much more central place in the story, as in Piketty (2020).

However, politics and institutions do not only come into the story through taxes and the welfare state's social insurance systems. The great emigration to America before 1914 is a case in point: slashing the population in the order of about one quarter, it affected inequality in Sweden through its effects on the labour supply, but also through institutional mechanisms, as it strengthened the bargaining power of workers (Karadja and Prawitz, 2019). Swedish trade unions were relatively well-organised already in the 1910s, reaching a position of being able to bargain over wages and benefits with employers (Lundh, 2010), and recent estimates suggest that they had positive effects on blue-collar wages at least as early as in the 1930s (Skoglund, 2025b). It is also the case that Swedish agricultural policy may have been unusually determined to further rationalise, and to pursue structural transformation. This is due to the specific political context, with a very strong Social Democratic labour movement after 1932. Gunnar Sträng, Social Democratic Assistant Minister of Agriculture 1945–1947, Minister of Agriculture 1948–1951, Minister of Social Affairs in 1951–1955, and Minister of Finance in 1955–1976, started his political career as a trade union organiser for agricultural workers, where he drew the conclusions that the sector's small, inefficient farming units necessarily led to penny-pinching farmers who would refuse to pay living wages for their workers. Therefore, Sträng concluded, a massive rationalisation of agriculture – the consolidation of hundreds of thousands of farms, with the transfer of previous farmers to industry – was necessary to create a productive agricultural sector with respectable wages. Similar conclusions were drawn in the 1930s by the leading Social Democratic economist and government minister Gunnar Myrdal (1938, esp. pp. 33–35, 66–70, 100–106). Swedish agricultural policy from the early 1930s focused on structural rationalisation, reducing the number of farms and ensuring that the remaining farms were sufficiently profitable to offer a respectable standard of living – by industrial society's yardstick – for its inhabitants (Nanneson, 1946). Thus, policy aided the movement of labour and capital to more productive lines of business, counteracting the kinds of market frictions that the development economics literature has documented (see the survey in Gollin and Kaboski, 2023).

Thus, there is reason to believe that structural transformation was especially conducive to equality in Sweden, given the relatively strong bargaining position of trade unions from the 1910s on and the pro-labour, pro-rationalisation line of policy and institutions from the early 1930s. To test this hypothesis, however, requires data and methods outside the scope of this paper. For now, we will only say that the results here are compatible with such an interpretation, and that the evidence presented in this section on agriculture's share of GDP and family farms' share of agriculture suggest that in terms of agriculture's role in the economy and social structure, at least at the starting point, c. 1870–1900, Sweden was a country quite comparable to other European and North American economies.

8 The role of capital and top incomes

Previous studies of the twentieth-century income structure in Sweden and elsewhere have emphasised the role of capital incomes (Roine and Waldenström, 2008; Gustafsson and Johansson, 2003; Piketty, 2003, 2014) Given the very stark concentration of capital incomes among the wealthy, a decrease in the capital share of income is associated with a fall in inequality (Bengtsson and Waldenström, 2018). To investigate the importance of this factor, Table 2 decomposes total income into labour income and capital income, and the share of the Gini coefficient explained by the two types of incomes. The concentration of capital income is very strong, with Gini coefficients ranging from 0.96 to 0.99. Given that Gini coefficients of capital income typically are above 0.9 (Milanovic, 2017), this is basically as expected. The capital share of incomes might seem surprisingly low, but we must remember that there is a major discrepancy between capital shares of national income in the national accounts and capital shares of household incomes, since much capital income is held by the corporate sector and therefore not distributed to individuals.[19]

As we can see in Table 2, most of the overall inequality is attributable to labour income. This attribution does not mean that capital income is unimportant, but it does mean that even without capital income, the overall message of Figures 1 and 2 would still hold: Sweden was a very unequal economy in the late 1800s and early 1900s, and then had a continuous decline in inequality for several decades. The decreasing share of capital income did contribute to equalising incomes, but it was not necessary for the overall trend of equalisation. A further interesting result in Table 2 is that the correlation between capital and labour income at the

[19] Piketty et al. (2018), pp. 561–563, 568–569, show that over the twentieth century in the US, about one third of national accounts capital income was distributed among individuals. Milanovic (2017) shows that for the US in the 2000s, taxation data reports that 11 per cent of personal income was capital income, while survey data reports 7 per cent.

individual level (β) changed quite drastically during the period. In 1920, the correlation was strongly negative, -0.505, suggesting a separation between rentiers/capitalists and employees. In 1950, the correlation was much weaker, -0.170, suggesting a degree of combination, an existence of a group who are at once capital-rich and labour-rich, as suggested in contemporary settings by Milanovic (2017) and Smith et al. (2019). Plotting the capital income to labour income correlation at the individual level (plots shown in Online Appendix C) reveals a transition from a linearly negative correlation to a U-shaped correlation, yielding the weaker coefficient shown in Table 2.

Table 2: Decomposing the income Gini into labour and capital incomes, 1920–1950

	Share of total income		**Share of total Gini explained by factor**		**Gini**			β
Year	Labour	Capital	Labour	Capital	Total	Labour	Capital	
1920	91%	9%	88%	12%	0.65	0.63	0.99	-0.505
1940	96%	4%	95%	5%	0.56	0.55	0.98	-0.340
1950	98%	2%	97%	3%	0.53	0.53	0.96	-0.170

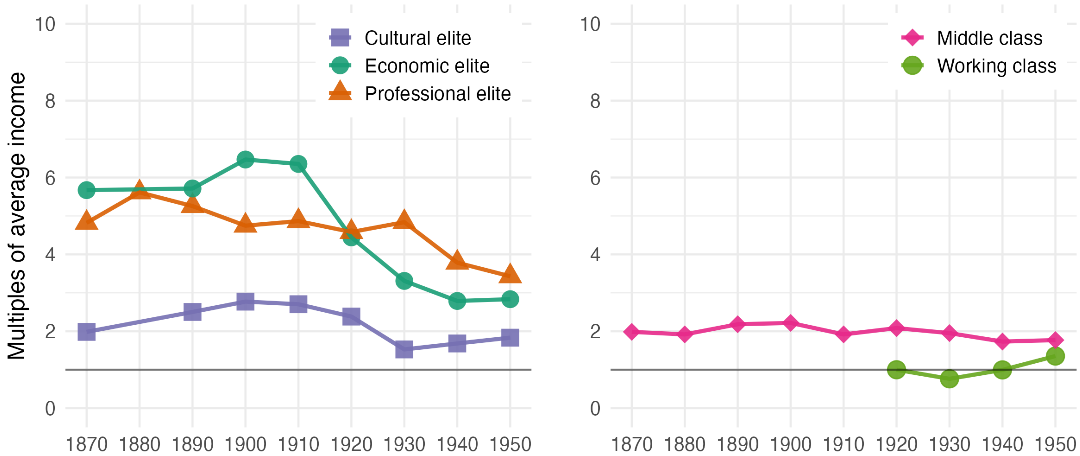
Note. Data are drawn from the *bevillning* tax dataset. See discussion in Section 3. Calculated with the command *sgini* in Stata. Capital incomes are underestimated in the *bevillning* data 1870–1910 since dividends were not covered by the tax in these years (see Online Appendix A.4). For this reason, the analysis in Table 2 only builds on data from 1920 onward. The 1930 Census does not include the information necessary for this calculation. The capital to labour income correlation is estimated for taxpayers using the regression specification: $\log(1 + \text{capitalincome}_i) = \beta \log(1 + \text{laborincome}_i) + \varepsilon_i$. The β shown in the Table is this estimated coefficient.

Another way to examine the importance of capital and other factors is to look at the income ratios for top income groups using the occupational information from our micro data. Here we ask: how important were capital incomes for the income advantages of elite groups compared to the average? Figure 10 shows income ratios, first only considering labour income and business income, and then in panel (b) also considering capital income. The income advantage of elite groups is higher when one includes capital incomes, but the long-run trend is much the same even without capital income.

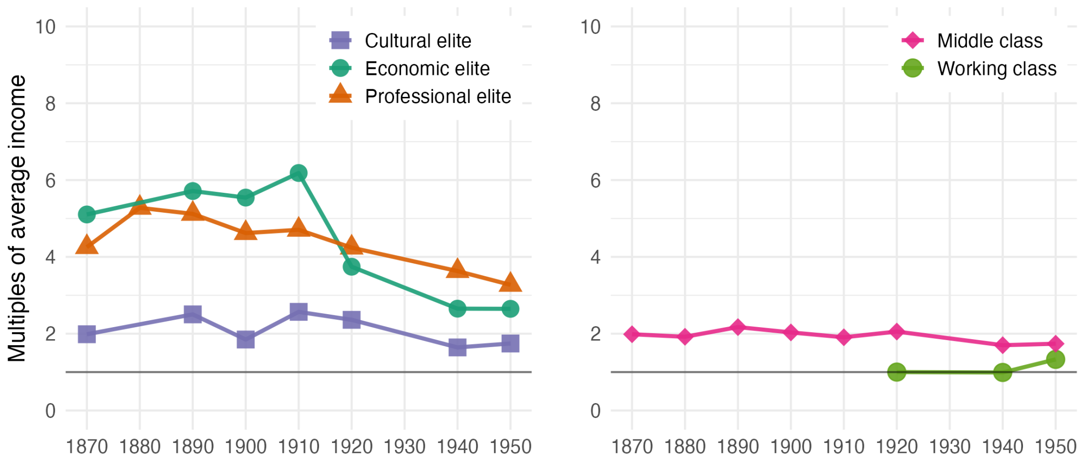
The great elite advantage in early twentieth century Sweden is surely related to the fact that Sweden was a laggard in providing a full secondary school education, and any education beyond that, to broad swathes of its population – a more fundamental factor shaping the inequality patterns observed here. The supply of highly educated professionals was very low until the mid-twentieth century; in 1930, only two per cent of the workforce had an upper secondary school diploma (*studentexamen*, equivalent to thirteen years of school)[20] and only one per cent had a college/university degree (SOU, 1935). The income difference between a professor or a doctor and a worker was much larger in early twentieth century Sweden than in the United States (cf.

[20]For an overview of the rather complex Swedish education system during the period see Larsson and Prytz (2011, pp. 126–129).

Figure 10: Income ratios with and without capital incomes



(a) All income



(b) Earned income (labor + business)

Note: Data are drawn from the *bevillning* tax dataset for 1870, 1880, 1890, 1900, 1910, 1920, 1940 and 1950, and the Census of 1930. See discussion in Section 3 and Online Appendices A and B. Economic elite defined as executives, rentiers and similar occupations. Professional elite defined as engineers, professors and similar highly educated professions. Cultural elite defined as journalists, authors and the like. Middle class defined as teachers, clerks and the like. See also Bengtsson and Molinder (2025) for an application of these categories to two high-income areas in Greater Stockholm, 1909 to 1950.

Bengtsson and Molinder, 2025; Goldin and Katz, 2009, p. 68). Taken together, Table 2 and Figure 10 show that Piketty style effects of the changing role of capital and Goldin and Katz style supply and demand effects are very much present in the Swedish case. To gain a full view of the changes of the income structure over the twentieth century, the structural transformation perspective pursued here needs to be supplemented with an analysis of the role of capital, taxation and transfers, and the education system. This is, however, outside the scope of the present paper.

9 Conclusions

The main contribution of our study is to use new micro data, uniquely detailed for a historical setting like the one studied here, to shed light on the importance of structural transformation for long-run income inequality. The literature on the equalisation of incomes in the developed economies in the twentieth century (Piketty and Saez, 2003; Piketty et al., 2018; Aaberge et al., 2020; Geloso et al., 2022) has not been able to account for the role of structural change, as the tabulated data used has not included information on the sectors and occupations of the income earners. This study demonstrates the importance of structural transformation for economic inequality in the industrializing economies of the late 1800s and early 1900s and, by extension, for contemporary economies (Gollin et al., 2014; Foster and Rosenzweig, 2022).

We have framed our contribution in the language of proximate and fundamental causes of changes in inequality, following North and Thomas (1973) and Acemoglu et al. (2014). The approach used here has its greatest strengths in revealing the proximate causes of inequality, *how* inequality changed. Our investigation, using decompositions of the Gini coefficient into rural and urban components, descriptions of the evolution of the job composition and wages, counterfactuals without structural transformation, and a Mookherjee–Shorrocks decomposition of the income distribution, shows that structural transformation was crucial to the equalisation of the income structure in Sweden in 1870–1950. Incomes were compressed as mechanisation and labour mobility incentivised the large pool of low-paid workers and "family assistants" working informally in agriculture to seek better-paid jobs in industry and services. Looking at crude indicators of the relative degree of egalitarianism in agriculture, and the role of agriculture in GDP in 1900, we have suggested that these mechanisms should have mattered for the equalisation of the income distribution in other developed economies too. In this regard, there was nothing outstanding about the trajectory in Sweden. We have also shown that structural transformation remains influential after controlling for the decline in capital incomes and the growth of human capital (as in Piketty, 2014, 2020; Goldin and Katz, 2009).

However, with our approach we are very limited in what we can say about the *fundamental* causes behind how structural transformation shaped the income distribution. it is likely that policy and institutions in Sweden from the 1930s on were especially friendly to structural transformation, and structural transformation in a labour-friendly fashion. This, we suggest, explains why the equalisation of pre-tax incomes became especially drastic in Sweden; at least such a hypothesis about "pre-distribution" (Bozio et al., 2024) is consistent with the facts documented here. To investigate this issue lies beyond the scope of this paper, but we would suggest that our results on the importance of agriculture, employment composition and education points to several relevant further avenues of research. In further research, the role of politics and the rise of the welfare state should also be enjoined into the analysis of structural transformation.

References

Aaberge, R. (2007). Gini's Nuclear Family. *Journal of Economic Inequality*, 5:305–322.

Aaberge, R., Atkinson, A. B., and Modalsli, J. (2020). Estimating Long-run Income Inequality from Mixed Tabular Data: Empirical Evidence from Norway, 1875–2017. *Journal of Public Economics*, 187:104196.

Acemoglu, D., Gallego, F. A., and Robinson, J. A. (2014). Institutions, Human Capital, and Development. *Annual Review of Economics*, 6:875–912.

Acemoglu, D. and Robinson, J. A. (2002). The Political Economy of the Kuznets Curve. *Review of Development Economics*, 6(2):183–203.

Ager, P., Goñi, M., and Salvanes, K. G. (2026). Gender-biased Technological Change: Milking Machines and the Exodus of Women from Farming. *American Economic Review*, 116(1):246–286.

Albertus, M. (2017). Landowners and Democracy: The Social Origins of Democracy Reconsidered. *World Politics*, 69(2):233–276.

Alfani, G. and di Tullio, M. (2019). *The Lion's Share: Inequality and the Rise of the Fiscal State in Preindustrial Europe*. Cambridge University Press.

Alvarez-Cuadrado, F. and Poschke, M. (2011). Structural Change out of Agriculture: Labor Push versus Labor Pull. *American Economic Journal: Macroeconomics*, 3:127–158.

Andersson, J. (2024). *Migrants and Towns: Self-Selection and Occupational Attainment of Rural-Urban Migrants in Swedish History*. Doctoral dissertation, Uppsala University.

Ansell, B. and Samuels, D. (2010). Inequality and Democratization: A Contractarian Approach. *Comparative Political Studies*, 43(12):1543–1574.

Atkinson, A. B. (2007). Measuring Top Incomes: Methodological Issues. In Atkinson, A. B. and Piketty, T., editors, *Top Incomes over the Twentieth Century: A Contrast Between European and English-Speaking Countries*, pages 18–42. Oxford University Press.

Atkinson, A. B. and Piketty, T. (2007). Towards a Unified Data Set on Top Incomes. In Atkinson, A. B. and Piketty, T., editors, *Top Incomes over the Twentieth Century: A Contrast Between European and English-Speaking Countries*, pages 531–565. Oxford University Press.

Auten, G. and Splinter, D. (2024). Income Inequality in the United States: Using Tax Data to Measure Long-Term Trends. *Journal of Political Economy*, 132(7):2179–2227.

Bagge, G., Lundberg, E., and Svennilson, I. (1933). *Wages in Sweden 1860–1930. Part One: Manufacturing and Mining*. P.S. King, London.

Belfield, C., Blundell, R., Cribb, J., Hood, A., and Joyce, R. (2017). Two Decades of Income Inequality in Britain: The Role of Wages, Household Earnings and Redistribution. *Economica*, 84(234):157–179.

Bengtsson, E. (2019). The Swedish Sonderweg in Question: Democratization and Inequality in Comparative Perspective, c.1750–1920. *Past and Present*, 244(1):123–161.

Bengtsson, E. and Molinder, J. (2024). Incomes and Income Inequality in Stockholm, 1870–1970: Evidence from Micro Data. *Explorations in Economic History*, 92:101568.

Bengtsson, E. and Molinder, J. (2025). What Happened to the Incomes of the Rich during the Great Levelling? Evidence from Swedish Individual-level Data, 1909–1950. *Journal of Economic History*, 85(2):411–441.

Bengtsson, E. and Waldenström, D. (2018). Capital Shares and Income Inequality: Evidence from the Long Run. *Journal of Economic History*, 78(3):712–743.

Bentzel, R. (1952). *Inkomstfördelningen i Sverige.* IUI, Stockholm.

Berger, E. and Vagle, H. B. (2017). Equality Before the Welfare State: The Norwegian Income Distribution 1892–1929. Master's thesis, Norwegian School of Economics.

Björklund, A. and Palme, M. (2000). The Evolution of Income Inequality During the Rise ofthe Swedish Welfare State 1951 to 1973. *Nordic Journal of Political Economy*, 26:115–128.

Blesch, K., Hauser, O. P., and Jachimowicz, J. M. (2022). Measuring Inequality Beyond the Gini Coefficient May Clarify Conflicting Findings. *Nature Human Behavior*, 6:1525–1536.

Bohlin, J. and Eurenius, A.-M. (2010). Why They Moved: Emigration from the Swedish Countryside to the United States, 1881–1910. *Explorations in Economic History*, 47:533–551.

Bohlin, J. and Prado, S. (2011). A Market with Strings Attached: Agricultural Property Price Movements. In Antonsson, H. and Jansson, U., editors, *Agriculture and Forestry in Sweden since 1900: Geographical and Historical Studies*, page 123–138. The Royal Swedish Academy of Agriculture and Forestry, Stockholm.

Bozio, A., Garbinti, B., Goupille-Lebret, J., Guillot, M., and Piketty, T. (2024). Predistribution versus Redistribution: Evidence from France and the United States. *American Economic Journal: Applied Economics*, 16(2):31–65.

Broadberry, S. (2009). Agriculture and Structural Change: Lessons from the UK Experience in an International Context. In Lains, P. and Pinilla, V., editors, *Agriculture and Economic Development in Europe Since 1870*, pages 76–94. Routledge, London.

Castillo, M. (2024). *Roots Across the Atlantic : Social Networks and Economic Outcomes from the Swedish Mass Migration to the United States, 1880–1920.* Doctoral dissertation, Lund University.

Chakravorty, S., Chandrasekhar, S., and Naraparaju, K. (2019). Land Distribution, Income Generation and Inequality in India's Agricultural Sector. *Review of Income and Wealth*, 65(S1):S182–S203.

Collin, K., Lundh, C., and Prado, S. (2019). Exploring Regional Wage Dispersion in Swedish Manufacturing, 1860–2009. *Scandinavian Economic History Review*, 67(3):249–268.

Cowell, F. A. (2011). *Measuring Inequality.* Oxford University Press, Oxford, third edition.

Cowell, F. A. and Flachaire, E. (2024). Inequality Measurement and the Rich: Why Inequality Increased More than We Thought. *Review of Income and Wealth*, 70(2):254–277.

Dackling, M. (2013). *Släktgårdens uppkomst: jord och marknad i Skaraborg, 1845–1945.* Doctoral dissertation, University of Gothenburg.

Edvinsson, R. (2005). *Growth, Accumulation, Crisis: With New Macroeconomic Data for Sweden 1800–2000.* Doctoral dissertation, Stockholm University.

Ejermo, O., Enflo, K., Eriksson, B., and Prawitz, E. (2025). Home, Sweet Home: Returns to Returning in the Age of Mass Migration. *American Economic Journal: Applied Economics*, 17(4):29–59.

Enflo, K., Lundh, C., and Prado, S. (2019). The Role of Migration in Regional Wage Convergence: Evidence from Sweden 1860–1940. *Explorations in Economic History*, 52(1):93–110.

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

Ericsson, J. and Molinder, J. (2026). How Deep are the Roots of Swedish Egalitarianism? A Multidimensional Approach. *European Review of Economic History*, 30(1):1–27.

Erikson, M. (2018). *Krediter i lust och nöd. Skattebönder i Torstuna härad, Västmanlands län, 1770–1870.* Doctoral dissertation, Uppsala University.

Fields, G. S. (2004). Dualism in the Labor Market: A Perspective on the Lewis Model After Half a Century. *The Manchester School*, 72(6).

Flodström, I. (1902). *Öfversikt af 1900 års bevillningstaxering.* P.A. Norstedt och söner, Stockholm.

Flygare, I. and Isacson, M. (2003). *Det svenska jordbrukets historia, V: Jordbruket i välfärdssamhället, 1945-2000.* LTs förlag, Stockholm.

Fortin, N., Lemieux, T., and Firpo, S. (2011). Decomposition Methods in Economics. In Ashenfelter, O. and Card, D., editors, *Handbook of Labor Economics, Volume 4A*, pages 1–102. Elsevier, Amsterdam.

Foster, A. D. and Rosenzweig, M. R. (2022). Are There Too Many Farms in the World? Labor Market Transaction Costs, Machine Capacities, and Optimal Farm Size. *Journal of Political Economy*, 130(3):636–680.

Garbinti, B., Goupille-Lebret, J., and Piketty, T. (2018). Income Inequality in France, 1900–2014: Evidence from Distributional National Accounts (DINA). *Journal of Public Economics*, 162:63–77.

Geloso, V. J., Magness, P., Moore, J., and Schlosser, P. (2022). How Pronounced Is the U-curve? Revisiting Income Inequality in the United States, 1917–60. *Economic Journal*, 132(647):2366–2391.

Goldin, C. and Katz, L. (2009). *The Race between Education and Technology*. Harvard University Press, Cambridge, MA, paperback edition.

Gollin, D. (2014). The Lewis Model: A 60-Year Retrospective. *Journal of Economic Perspectives*, 28(3):71–88.

Gollin, D. and Kaboski, J. P. (2023). New Views of Structural Transformation: Insights from Recent Literature. *Oxford Development Studies*, 51(4):339–361.

Gollin, D., Lagakos, D., and Waugh, M. E. (2014). The Agricultural Productivity Gap. *Quarterly Journal of Economics*, 129(2):939–993.

Gustafsson, B. and Johansson, M. (2003). Steps Toward Equality: How and Why Income Inequality in Urban Sweden Changed during the Period 1925–1958. *European Review of Economic History*, 7(2):191–221.

Hedlund-Nyström, T. (1972). The Finance Act of 1862 and Its Effect on the Economy. *Economy and History*, 15(1):58–82.

Heikkuri, S. (2025). Unravelling Income Differentials: Exploring the Decline in Skill Premiums in Twentieth-century Sweden. *Scandinavian Economic History Review*, 74(1):62–83.

Heikkuri, S. (2026). Technological Change and Structural Shifts in Sweden's Labor Market, 1870–1930: Skill Upgrading or Deskilling? *Scandinavian Economic History Review*, pre-published online, 17 February 2026:1–25.

Heino, S. (2026). Was It Really So High? New Estimates on Income Inequality in Finland, 1865–1959. *Scandinavian Economic History Review*, pre-published online, 19 May 2026:1–23.

Henrekson, M. and Stenkula, M. (2015). Swedish Taxation Since 1862: An Overview. IFN Working Paper 1052, IFN, Stockholm.

Henriksen, I. (2009). The Contribution of Agriculture to Economic Growth in Denmark, 1870–1939. In Lains, P. and Pinilla, V., editors, *Agriculture and Economic Development in Europe Since 1870*, pages 117–147. Routledge, London.

Hoffmann, W. G. (1965). *Das Wachstum der deutschen Wirtschaft seit der Mitte des 19 Jahrunderts*. Springer, Berlin. data freely available from GESIS: https://doi.org/10.4232/1.8260.

Hornbeck, R. and Naidu, S. (2014). When the Levee Breaks: Black Migration and Economic Development in the American South. *American Economic Review*, 104(3):963–990.

Jayes, J., Molinder, J., and Enflo, K. (2024). Power for Progress: The Impact of Electricity on Individual Labor Market Outcomes. CEPR Discussion Paper 18973, CEPR Press, Paris and London.

Jenkins, S. P. (2017). Pareto Models, Top Incomes and Recent Trends in UK Income Inequality. *Economica*, 84(334):261–289.

Karadja, M. and Prawitz, E. (2019). Exit, Voice, and Political Change: Evidence from Swedish Mass Migration to the United States. *Journal of Political Economy*, 127(4):1864–1925.

Karlan, D. and Morduch, J. (2010). Access to Finance. In Rodrik, D. and Rosenzweig, M., editors, *Handbook of Development Economics, Volume 5*, pages 4704–4784. Elsevier, Amsterdam.

Kuuse, J. (1970a). *Från redskap till maskiner: Mekaniseringsspridning och kommersialisering inom svenskt jordbruk 1860–1910*. Department of Economic History, Gothenburg.

Kuuse, J. (1970b). *Inkomstutveckling och förmögenhetsbildning: En undersökning av vissa yrkesgrupper 1924–1959*. Doctoral dissertation, Gothenburg University.

Kuznets, S. (1955). Economic Growth and Income Inequality. *American Economic Review*, 45(1):1–28.

Kälvemark, A.-S. (1972). *Reaktionen mot utvandringen: Emigrationsfrågan i svensk debatt och politik 1901–1904*. Scandinavian University Books, Lund.

Lains, P. (2009). Agriculture and Economic Development in Portugal, 1870–1973. In Lains, P. and Pinilla, V., editors, *Agriculture and Economic Development in Europe Since 1870*, pages 333–353. Routledge, London.

Larsson, E. and Prytz, J. (2011). Läroverk och gymnasieskola. In Larsson, E. and Westberg, J., editors, *Utbildningshistoria: en introduktion*, pages 121–142. Studentlitteratur, Lund.

Lewis, W. A. (1954). Economic Development with Unlimited Supplies of Labour. *The Manchester School*, 22(2):139–191.

Lindahl, E., Dahlberg, E., and Kock, K. (1937). *The National Income in Sweden 1861–1930. Part One*. WS King & Son, London.

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

Lundh, C. (2010). *Spelets regler: Institutioner och lönebildning på den svenska arbetsmarknaden 1850–2000*. SNS, Stockholm, 2 edition.

Lundh, C. and Prado, S. (2015). Markets and Politics: The Swedish Urban-Rural Wage Gap, 1865–1985. *European Review of Economic History*, 19(1):67–87.

Manuelli, R. E. and Seshadri, A. (2014). Frictionless Technology Diffusion: The Case of Tractors. *American Economic Review*, 104(4):1368–1391.

Milanovic, B. (2016). *Global Inequality*. Harvard University Press, Cambridge, MA.

Milanovic, B. (2017). Increasing Capital Income Share and Its Effect on Personal Income Inequality. In Boushey, H., DeLong, J. B., and Steinbaum, M., editors, *After Piketty: The Agenda for Economics and Inequality*, pages 235–258. Harvard University Press, Cambridge, MA.

Milanovic, B. (2024). How Rich Were the Rich? *Explorations in Economic History*, 93:101592.

Modalsli, J. (2018). The Regional Dispersion of Income Inequality in Nineteenth-Century Norway. *Explorations in Economic History*, 67:62–79.

Molinder, J. (2023). Historical Roots of the Dual-Earner Model: Women's Labour Force Participation in Sweden, 1870–1960. Lund Papers in Economic History 244, Lund University.

Mookherjee, D. and Shorrocks, A. (1982). A Decomposition Analysis of the Trend in UK Income Inequality. *The Economic Journal*, 92(368):886–902.

Morell, M. (2001). *Jordbruket i industrisamhället 1870–1945*. Natur och Kultur / LTs förlag, Stockholm.

Myrdal, G. (1938). *Jordbrukspolitiken under omläggning*. Kooperativa förbundets bokförlag, Stockholm.

Nanneson, L. (1946). Rationalitetsvariationerna inom det svenska jordbruket. Technical report, Statistiska centralbyrån (SCB), Stockholm. Utredning på uppdrag av 1942 års jordbrukskommitté.

Ngai, R. L., Olivetti, C., and Petrongolo, B. (2024). Gendered Change: 150 Years of Transformation in US Hours. NBER Working Paper 32475, National Bureau of Economic Research, Cambridge, MA.

North, D. C. and Thomas, R. P. (1973). *The Rise of the Western World: A New Economic History*. Cambridge University Press, Cambridge.

Olsson, K. (1972). *Hushållsinkomst, inkomstfördelning och försörjningsbörda: En undersökning av vissa yrkesgrupper i Göteborg 1919–1960*. Doctoral dissertation, University of Gothenburg.

Piketty, T. (2003). Income Inequality in France, 1901—1998. *Journal of Political Economy*, 111(5):1004–1042.

Piketty, T. (2014). *Capital in the Twenty-First Century*. Harvard University Press.

Piketty, T. (2020). *Capital and Ideology*. Harvard University Press.

Piketty, T. and Saez, E. (2003). Income Inequality in the United States, 1913–1998. *Quarterly Journal of Economics*, 118(1):1–39.

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

Piketty, T., Saez, E., and Zucman, G. (2024). Income Inequality in the United States: A Comment. Unpublished paper, Paris School of Economics/UC Berkeley, 16 September 2024.

Ranaldi, M. and Milanovic, B. (2022). Capitalist Systems and Income Inequality. *Journal of Comparative Economics*, 50(1):20–32.

Riksarkivet (2025). Swedish 1930 Census. Available from: https://sok.riksarkivet.se/folkrakningar/. The 1930 Census is digitalized by the Swedish National Archives, Riksarkivet.

Rodrik, D. (2016). Premature Deindustrialization. *Journal of Economic Growth*, 21:1–33.

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

Roine, J. and Waldenström, D. (2009). Wealth Concentration over the Path of Development: Sweden, 1873–2006. *Scandinavian Journal of Economics*, 111(1):151–187.

Roine, J. and Waldenström, D. (2014). Long Run Trends in the Distribution of Income and Wealth. In Atkinson, A. B. and Bourguignon, F., editors, *Handbook of Income Distribution*, volume 2, pages 469–592. North-Holland, Amsterdam.

Ruggles, S. (2015). Patriarchy, Power, and Pay: The Transformation of American Families, 1800-–2015. *Demography*, 52(6):1797–1823.

Scheidel, W. (2017). *The Great Leveler: Violence and the History of Inequality from the Stone Age to the Twenty-First Century*. Princeton University Press.

Schön, L. (2014). *En modern svensk ekonomisk historia*. Studentlitteratur, Lund, 4th edition.

Schön, L. and Krantz, O. (2015). Swedish Historical National Accounts 1560–2010. Available from: https://www.lusem.lu.se/organisation/department-economic-history/research-department-economic-history/databases-department-economic-history/economic-history-data/swedish-historical-national-accounts-1300-2022. Dataset created by Lennart Schön and Olle Krantz, maintained and updated by the Department of Economic History, Lund University.

Shorrocks, A. (1984). Inequality Decomposition by Population Subgroups. *Econometrica*, 52(6):1369–1385.

Skoglund, W. (2025a). *Power Matters: Essays on Labor Market Power*. PhD thesis, Acta Universitatis Upsaliensis.

Skoglund, W. (2025b). Union Wage Effects in Sweden: Evidence from the Interwar Period. *Explorations in Economic History*, 96:101655.

Smith, M., Yagan, D., Zidar, O., and Zwick, E. (2019). Capitalists in the Twenty-First Century. *Quarterly Journal of Economics*, 134(4):1675–1745.

SOU (1935). *Betänkande med undersökningar och förslag i anledning av tillströmningen till de intellektuella yrkena*. Number 1935:52 in Statens Offentliga Utredningar. Department of Education, Stockholm.

Splinter, D. (2019). Who Pays No Tax? The Declining Fraction Paying Income Taxes and Increasing Tax Progressivity. *Contemporary Economic Policy*, 37(3):413–426.

Stadsarkivet (2025). Mantals- och kronotaxeringslängder Stockholms stad. Available from: https://sok.riksarkivet.se/mantalslangder-stockholms-stad. In Stockholm City, the Register of population, the Mantalslängder, were also the taxation lists. The Stockholm City Archive (Stadsarkivet) have digitalized all lists up to 1915.

Steinmo, S. (1989). Political Institutions and Tax Policy in the United States, Sweden, and Britain. *World Politics*, 41(4):500–535.

Torregrosa Hetland, S. and Sabaté, O. (2026). Income Taxes and Redistribution in the Early Twentieth Century. *Economic History Review*, 79(2):634–660.

Van Zanden, J. L. (1995). Tracing the Beginning of the Kuznets Curve: Western Europe during the Early Modern Period. *The Economic History Review*, 48(4):643–664.

Vanhanen, T. (1997). *Prospects of Democracy: A Study of 172 Countries*. Routledge, London.

Vanhanen, T. (2003). Democratization and power resources 1850–2000. http://urn.fi/urn:nbn:fi:fsd:T-FSD1216. Version 1.0 (2003-03-10). Finnish Social Science Data Archive [distributor].

Vivier, N. (2009). Agriculture and Economic Development in France, 1870–1939. In Lains, P. and Pinilla, V., editors, *Agriculture and Economic Development in Europe Since 1870*, pages 210–233. Routledge, London.

Öhngren, B. (1974). *Folk i rörelse: Samhällsutveckling, flyttningsmönster och folkrörelser i Eskilstuna 1870–1900*. Doctoral dissertation, Uppsala University.

A Online Appendix A: The taxation sources, sampling strategy, and adjustments

A.1 The taxation system

During the period for which we create an original dataset, 1870 to 1950, the Swedish income taxation system saw two parallel taxes for a period (1902–1928), and the transformation in 1928 of the older income tax into a municipal tax.

The first income tax, *bevillningen*, was presented as a proposal at the 1859–1860 Riksdag, decided upon in October 1860 (Riksdagen, 1860), and for the first time implemented in 1862.[21] The income *bevillning* removed all special rules which were rife for income taxes before and imposed a simple system: all net incomes should be taxed at 1 %, with the exception that those earning less than 400 kronor were exempt, and that those earning 400-1800 kronor could deduct 300 kr from tax. (Hedlund-Nyström, 1972; Lodin et al., 2017, p. 9). In the simplicity and the removal of special rules and privileges, the prominent economist Gustav Cassel (1906, pp. 191–192) saw the reform as one modernizing step away from the old society of privileges. It is sometimes said that the "modern" income tax system was created in Sweden in 1902 with the imposition of another, parallel income tax, *inkomstskatten* (Andersson, 2021; Torregrosa Hetland and Sabaté, 2026), but the 1862 *bevillning* should, given its great simplification of the tax system and its universal application, share some of this recognition.[22]

The *bevillning* had two parts: that for income, and that for property. Farmers were due to pay tax on property, not on income: the politicians argued, as summarised by one historian, that "most property owners did not use a system of book-keeping that made possible a definite assessment, even by themselves, of their net income." (Hedlund-Nyström, 1972, p. 59) The assumption was that farmers earned a return of 5 per cent on their property, and the tax rate was 1 per cent of this assumed 5 per cent. Since the farmers were not covered by the income *bevillning* we have imputed their incomes based on other sources, including a sample of roughly 13 000 property *bevillning* returns. Our estimates do not build on the 5 per cent returns assumption, as this was, as already recognised by contemporaries (Fahlbeck, 1893), artificially low. (See Online Appendix B.1.)

The practical taxation procedure for the income *bevillning* was from 1862 conducted in three steps. In step 1, a bevillning committee (*bevillningsberedning*) consisting of local public officials, and a chairman appointed by the County, suggested the levels of taxation. In step 2, another committee (*taxeringsnämnd*) of 2-4 public officials reviewed the set taxation levels and established them. In step 3, if a taxpayer had a complaint, he or she could appeal to the board of appeals (*prövningsnämnd*). (The procedure is described in Andersson, 1995, p. 301.) Note that in every case, the taxation committee of one year organised the taxes to be paid on the income of the *previous* year. Thus, the *bevillning* of 1911 covered the incomes of 1910, the bevillning of 1912 the incomes of 1911, and so on. So when we analyse the incomes of 1870, 1880 and 1890, it

[21] The name of the tax, *bevillningen*, is an old Swedish word for "grant" and in the taxation context it refers to the fact that the King needed the acceptance of the Estates (until 1866, the parliament was an Estates Diet) to impose new taxes. The word was old-fashioned by the late nineteenth century and fell out of use around 1930.

[22] The *inkomstskatt* is the tax used by previous tax-based estimates of long-run income inequality in Sweden by Roine and Waldenström (2008, 2010). For contemporary discussions of the relation and difference between the two income taxes and their source value, see Flodström (1906) and Widell (1907). For a historical discussion of the relation between the two taxes and the politics around them see Andersson, 1995, pp. 89–143. The legislation of the income tax is available in Riksdagen (1902) and Riksdagen (1910).

means that we use the taxation sources of 1871, 1881, and 1891.[23]

In 1902, the second income tax, *inkomstskatten*, was created. This instated the necessity for individual tax filing with more stringent checks on one's precise incomes. The tax filing (*deklarationen*) was from then on used as the basis for both the *bevillning* and the *inkomstskatt*. The *bevillning* from then on became the municipal income tax, while the *inkomstskatt* became the state tax. Since one's voting rights on the municipal level was linked to one's income and wealth, the *bevillning* estimate of income and wealth played a crucial role in municipal politics.

The *inkomstbevillning* system was reformed in 1920 and 1928 (Riksdagen, 1919, 1928). The 1920 reform (decided in 1919, implemented in 1920) meant, as discussed in the main body of the paper (Section 3), that corporate dividends earned by individuals started to be taxed. We discuss the implications of the exclusion of dividends before 1920 in Online Appendix A.4. The 1928 reform is not of relevance for our purposes, to estimate incomes. But from then on there was not any longer any *bevillning*, and the municipal income tax was only called that (*kommunal inkomstskatt*). (The municipal tax reforms are described in SOU, 1933, pp. 236-239.)

A great strength of the *bevillning* as a source to study income distribution is that the system, even before the 1920 and 1928 reforms, was implemented homogenously across all municipalities within the country. This can not be taken for granted in early twentieth-century settings.

A.2 Sampling strategy and the dataset

The purpose of our archival work was to create the first nationally representative dataset of incomes in Sweden in the pre-1960s period.[24] To create a representative dataset, we have put emphasis on the broad sampling of all kinds of regions and settings.

The taxation lists are archived at the National Archives (*Riksarkivet*) in Stockholm. For the purposes of sampling, the crucial fact is that they are organised by municipality (*kommun*).[25] They were two types of municipalities during this time period, rural (*landskommun*) and urban (*stadskommun*), and any feasible sampling strategy must work from this organisation of the taxation lists. The lists are bound in large books, collected by county (*län*). Within each county, rural municipalities were sorted by bailiwick (*fögderi*), while urban municipalities had their own volumes.[26]

In principle, every page of every municipality has been preserved for every year from 1862 onwards. Each year of *bevillning* records encompasses hundreds of volumes, over time thousands, each volume with hundreds of pages, each page with roughly 20 taxpayers. The extreme wealth of the sources necessitated a sampling strategy. To begin with, we chose to only study every

[23] Details regarding how Swedes (especially wealthier Swedes) went about filing their taxes in 1910s can also be gathered from Palmgren's (1915) contemporary guide to filing your taxes.

[24] Cf. discussion in Björklund and Palme (2000) and the contemporary discussion in Esberger and Malmquist (1972) who also discuss the early history of the "Born the 15th" sample which we use for 1960 and 1970.

[25] The municipalities were created in 1862 and were given some of the responsibilities previously held by the Churchs's parishes (*socken*), notably schooling and poor relief.

[26] To take an example, purely for illustration. Our 1900 sample from Malmöhus County consists of five towns/urban municipalities (Malmö, Lund, Landskrona, Helsingborg and Trelleborg) and 45 rural municipalities. The rural municipalities are nested within eight bailiwicks: for example, Luggude bailiwick contains seven sampled municipalities, Allerum, Brunnby, Ekeby, Höganäs, Kattarp, Norra Vram and Väsby. All these are included in the two volumes of taxation lists (volume nos. 20 and 21 in the Malmöhus series) from Luggude bailiwick in 1900. Conversely, our sampled Malmöhus municipalities Bara, Burlöv, Genarp and Lyngby are located in Torna and Bara bailiwick, volumes 26 and 27 of the Malmöhus series. Note that the bailiwick level hasn't been a factor in our sampling, but only matters on the practical level since this is the way that the taxation lists are sorted.

tenth year from 1870 to 1950, as it is the purpose of our study to capture the long-run changes in the income distribution and how these relate to structural transformation. This means, of course, that our dataset is less useful for studying short-run fluctuations.

The second necessary choice was: which taxpapers to sample each year? Our aim was to sample 1 per cent of taxpayers every tenth year, from 1870 to 1920, and (for reasons of work economy) 0.5 per cent in 1940 and 1950.[27] We randomly drew rural and urban municipalities into our sample, using a population weighting strategy so that more populated municipalities (within each type) were more likely to be sampled. There were about 2,400 rural municipalities in Sweden during the period, and, in 1870, 90 urban municipalities. For 1870–1920 we sampled one fifth of municipalities, and for 1940 and 1950 one tenth. Within each sampled municipality, we sampled every twentieth page of taxpayers. Thus, the sample in principle adds up to 5 per cent of the taxed population of the sampled municipalities, and to 1 per cent (1870-1920: 20 % * 5 %) or 0.5 per cent (1940 and 1950: 10 % * 5 %) of the taxed population in the country as a whole.

The samples of 1900, 1920 and 1950 are visualised in the maps in Figure A.2. As they well illustrate, our sampling strategy is not a stable panel of municipalities, but a new random sample for each year. The maps also illustrate the very broad sampling of Sweden, geographically. Given the great variation in municipality population (and taxpaying population), our sample varies significantly in the density of observations per municipality. In 1900, for example, the sample includes 6 taxpayers from Öglunda municipality in Skaraborg County, but about 5 400 taxpayers from Stockholm City.

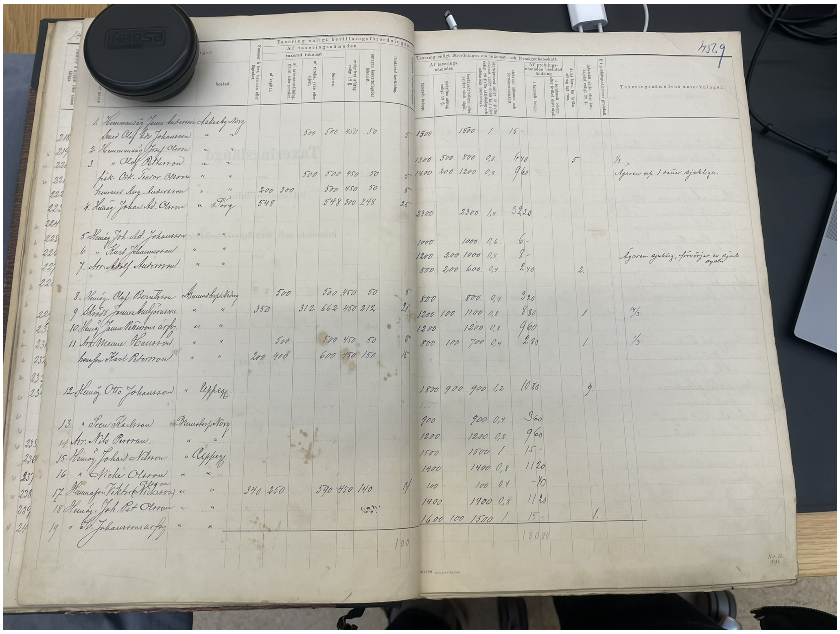
Building the dataset from our sampling strategy, we ran into a problem of oversampling of rural taxpayers. This problem – in relation to the 0.5 / 1 per cent goal – arose from the unforeseen fact that it was common among rural municipalities in the pre-1920 years, when taxation coverage still was relatively low (Table 1) to have only one page or a few pages of taxpayers. In the extreme case of there being only one page of taxpayers, and we have sampled that page (1/n), then we have sampled 100 per cent of the taxpayers in this particular municipality, while we only intended to sample 5 per cent of taxpayers there. This is not an issue for urban municipalities, as they were more populous (and had a higher proportion of the population covered by the income *bevillning*), but for rural municipalities, to make our calculations of income levels and the income distribution, we have applied a re-weighting of observations due to pre-1920 over-representation of taxpayers from small rural municipalities.

Our reweighting procedure builds on the ratio of taxpayers to municipalities. From the public statistics (SCB, 1906), we gathered for each of the 24 counties (a) the rural population and (b) the number of rural municipalities, and from this we calculated the ratio of population/municipality (a/b).[28] The median population for a rural municipality in 1900 varied from the low of 411 in Gotland, 647 in Skaraborg and 843 in Uppland to 3056 in Norrbotten, 3840 in Gävleborg and 3844 in Västerbotten. Thus, in high municipality density areas such as Gotland and Skaraborg, our sampling strategy leads to overrepresentation.

[27] The threshold to pay the tax was raised to 500 kr in 1884, 600 kr in 1903 and 1200 kr in the late 1940s (Gustafsson and Johansson, 2003, p. 194). However, since nominal incomes grew throughout the period, an increasingly large share of the population ended up above the threshold, as discussed in the Data section (Section 3) of the paper.

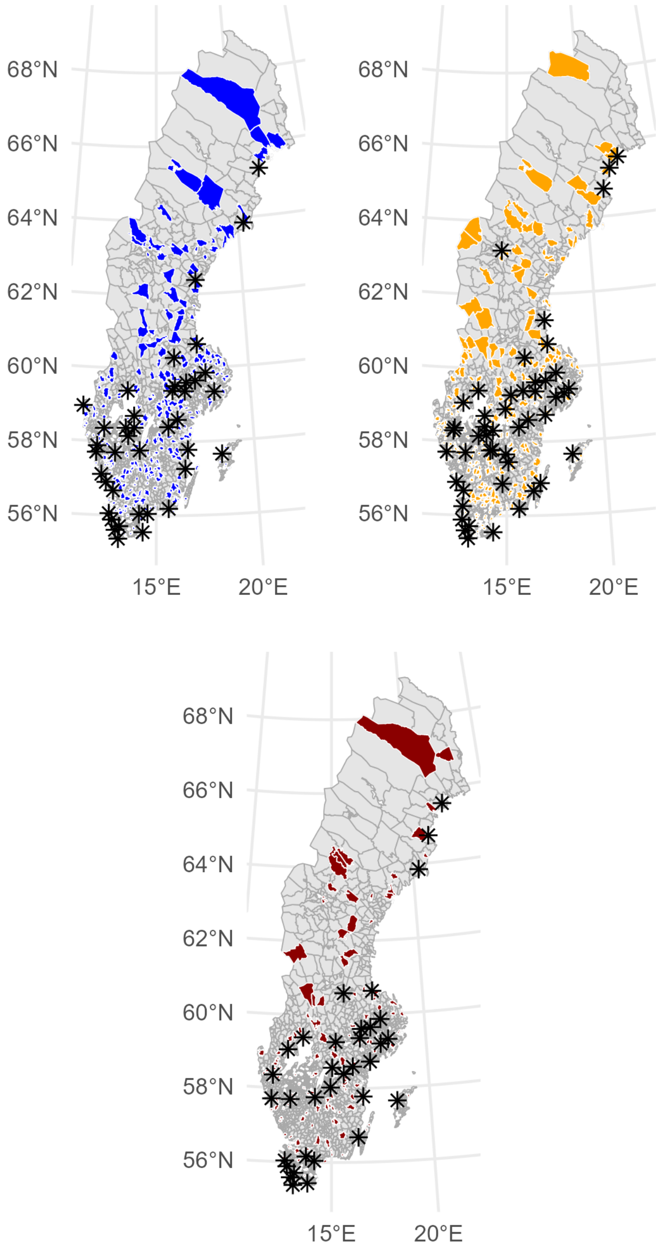
[28] The variation goes back to the long history, as the municipalities in 1862 were formed from the existing parishes. Thus, areas more populated in the Middle Ages and with slower population growth from then until 1862 had higher municipality/population ratios in our period of investigation, until the large mergers of municipalities in the 1940s. The historical persistence is tempered by any transformations of rural parishes into towns (with city privileges) in the meanwhile, as urban municipalities are not included in the calculations here.

Figure A.1: Example of a taxation list: Säve municipality, 1910



Note. This is a tax list from Säve Municipality in Göteborg and Bohus County from taxation year 1911, covering the incomes of 1910. The taxpayers are sorted by place of residence (sorted alphabetically) within the municipality. The list begins with *hemmansägare* (farmer) Janne Andersson in Askerby with an income from self-employment of 500 kr. The columns on the left hand side, the ones that we use, are in order: occupation and name, place of residence, income of capital, income of labour, income of self-employment, total income ("Summa"), allowed deductions, taxable income. The page number on the top right side of the right page of the fold, 4569, is a reminder of the immense extent of the taxation records: this is page 4569 of rural taxpayers for just one of the 24 counties of this year. Note that the left page presents the income taxed under the *bevillning* and the right page the income taxed under the *inkomstskatt*, the two tax systems as discussed in Appendix A.1.

Figure A.2: Maps of the 1900, 1920 and 1950 samples



Note: The areas filled with colour are sampled rural municipalities. The black stars represent sampled urban municipalities. Note that the 1900 and 1920 samples encompass 1 % of taxpayers and the 1950 sample 0.5 %; hence the sparser coverage in the map for 1950.

To calculate the extent of rural oversampling, we counted the number of pages in the tax lists for each municipality in three counties: Malmöhus, Skaraborg, and Norrbotten. The three counties were chosen to represent various densities of (rural) municipalities per (rural) population: Skaraborg on the high end, Norrbotten on the low end, and Malmöhus in the intermediate group (median population for a rural municipality in 1900 was 1262).[29] The results are shown in Table A.1.

Table A.1: Number of pages in taxation registers per rural municipality, year 1900

	Skaraborg		Malmöhus		Norrbotten	
Pages	n	%	n	%	n	%
1	187	70.6%	60	23.4%	1	4.2%
2	35	13.2%	63	24.6%	4	16.7%
3	13	4.9%	40	15.6%	0	0
4	13	4.9%	21	8.2%	1	4.2%
5	2	0.8%	10	3.9%	0	0
6	5	1.9%	11	4.3%	0	0
7	2	0.8%	6	2.3%	1	4.2%
8	0	0	6	2.3%	1	4.2%
9	0	0	3	1.2%	1	4.2%
10	1	0.4%	3	1.2%	1	4.2%
11–20	2	0.8%	12	4.7%	5	20.8%
>20	5	1.9%	21	8.2%	9	37.5%

Note. The table presents the number of pages in the income *bevillning* taxation registers per rural municipality for the taxation year 1900. So for example, in Skaraborg county there were 187 rural municipalities which only had one page of income taxpayers, 35 rural municipalities with two pages, and so on. The calculation builds on the original taxation registers archived at the National Archives (*Riksarkivet*) in Stockholm.

Table A.1 shows as we would expect that the upward bias in our sample size is significant in Skaraborg, moderate in Malmöhus, and lesser in Norrbotten which had fewer and larger municipalities. To avoid over-representation of sparsely populated areas with a high ratio of municipalities per capita, we in our calculations weight, for the pre-1920 years, observations from rural municipalities by our calculated ratio of sampled taxpayers to actual taxpayers.[30] In the abstract, this weight will always be 20, as we have sampled every 20th page of taxpayers. But because of the discrepancy shown here, the number of rural municipalities with < 20 pages of taxpayers, using the weight of 20 will overstate the number of total taxpayers in the countryside. For 1870, 1880, 1890, 1900 and 1910, we thus use the weight 4 for rural municipalities (thus assuming that in practice, we have sampled 25 % of taxpayers within the sampled rural municipalities). This makes sense in the light of Table A.1 and we control the validity of the assumption by comparing the resulting estimate of the number of total taxpayers, using this adjustment for rural municipalities and no adjustment for urban ones (i.e. the weight is 20), with the estimate of the number of total taxpayers by county in 1900 made by Flodström (1902). His estimate for the country as a whole is that 13.9 % of the population paid the tax, and our estimate with the

[29] We here bracket the fact that the relationship between population and taxpayer population is not fixed, since the share over the income threshold to pay the tax varied across regions.

[30] By 1920 the number of taxpayers had increased drastically as a consequence of real income growth and inflation, so the bias that tax registers in rural areas had less than 20 pages is less of a concern. See discussion of Table 1 in the main body of the paper.

assumptions discussed here is 14.2 %, a satisfying degree of correspondence.

A.3 Coding of professions and gender

The taxation lists contain a multitude of professional titles: both because there genuinely were a multitude of jobs in Sweden, and because the conventions regarding spelling, abbreviations and so on vary across the regions, towns and years. We have used Stata to homogenise the professional titles.

We have coded each standardised occupation along four dimensions: (i) economic sector, (ii) subsector, (iii) occupational class, and (iv) labour force status. This classification draws on the social-table tradition in economic history, which provides a detailed yet tractable way to describe the occupational composition of a population (cf. Milanovic, 2024, Gómez León and Gabbuti, 2025).

The first dimension, *sector*, distinguishes four broad categories: (1) primary (agriculture, forestry, and fishing), (2) secondary (mining, manufacturing, and construction), (3) tertiary (trade, transport, and services), and (4) other or undefined. Within each sector, a *subsector* variable captures finer economic divisions. The secondary sector is divided into nine subsectors: mining; metal and engineering; quarrying; wood; pulp, paper, and printing; food, beverage, and tobacco; textile, clothing, leather, and hair; chemical and rubber; building and construction (including gas and waterworks); and an undefined residual category. Similarly, the tertiary sector distinguishes between seven subsectors and an unknown category: trade; transport and communication; insurance and banking; personal and social services; education and professions; administration and military; domestic services; and an undefined category. The primary sector is divided into agriculture and forestry on the one hand and fishing on the other.

The third dimension, *occupational class*, captures the individual's position within the hierarchy of production. In the primary sector, the class variable differentiates between large landowners; large-scale farmers; middle-scale farmers; small-scale farmers; semi-landless; salary-earners; wage-earners; assistants (family workers); and former or retired workers. In the secondary and tertiary sectors, the class categories are factory owners or executives; working owners or artisans; salary-earners; wage-earners; and former or retired workers. The "other or undefined" sector additionally includes categories for the poor and paupers, students, and individuals listed only with familial titles (such as wife, widow, or daughter) whose own occupation is unrecorded.

The fourth dimension, *labor-force status*, classifies each individual as either in the labor force, outside the labor force, or of unknown labor-force status. Individuals listed with familial titles, students, and those in poor relief institutions are coded as outside the labor force or of unknown status.

Each standardised occupation is manually coded along all four of these dimensions using Stata `replace` commands. For example, a *skomakare* (shoemaker) is coded as belonging to the secondary sector, the textile, clothing, leather, and hair subsector, and the working owners or artisans class. A *gruvarbetare* (mining worker) is assigned to the secondary sector, the mining subsector, and the wage-earner class. A *skollärare* (school teacher) is placed in the tertiary sector, the education and professions subsector, and the salary-earner class. Several entries in the taxation lists carry no substantive occupational information—for example, individuals listed only as *fru* (Mrs.), *herr* (Mr.), or *son* (son)—and are therefore left without a class coding in this step. The coding also handles individuals registered with the prefix *fd.* ("former"), indicating that they are retired from a given profession; these are assigned to the former or retired class within the relevant sector.

In total, approximately 5,900 unique standardised occupations were coded along these four dimensions, producing roughly 25,400 lines of Stata code.

The taxation lists contain no gender variable, so we have coded the taxpayers as men or women depending on their (1) first name, and (2) their occupational title – Swedish job titles in our period of investigation often carried gendered suffixes for women, such as *affärsidkerska* for a woman but *affärsidkare* for a man ("shop owner"), or *arbeterska* for a woman but *arbetare* for a man ("worker").

A.4 The issue of dividends before 1920

As discussed in the Data section of the paper, dividends from shares were not taxed under the *bevillning* until 1920. This means that we have a downward bias in our estimates of incomes as well as income inequality before 1920, given that dividends were more unequally distributed than the average incomes. How grave is this problem? To answer this question we need two pieces of information: firstly, the importance of dividends in the Swedish economy; secondly, their distribution in the population.

To map the importance of dividends, we have collated information on gross operating surplus and dividends for all incorporated firms between 1870 and 1950. In doing so, we have drawn upon previous research, contemporary investigations before 1910, and official statistics thereafter (Flodström, 1902, 1911; Möller, 1962; Broberg, 2006). To contextualise, from 1848 to 1897, a special permit from the King was required to form a new stock corporation (Broberg, 2006). This requirement limited the number of corporations in the country, but after deregulation in 1897, the number took off; Broberg (2006) dates the breakthrough of the joint stock companies in the Swedish economy to the first two decades of the twentieth century.[31] As one would expect given this background, the total sum of dividends for 1870 and 1880 are relatively small: 11 million SEK in 1870, and 28 million in 1880. In 1900 the dividends had risen to 96 million SEK, and in 1910 they had reached 146 million SEK. To understand the relative magnitude of these figures, and the possible bias resulting from their exclusion from our *bevillning* data, we must compare them to the personal income sum, that is the share of national income that is distributed among individuals (Schön and Krantz, 2015; see also discussion in Appendix B.4). This is the basis for our calculations of inequality and thus the relevant comparison for dividends. Dividends related to total personal income (without dividends) was 1.2 per cent in 1870 and 6 per cent in 1910.[32]

That is the size of the dividends. What about their distribution? We do not have actual data on their distribution, so we need to proxy the pre-1920 distribution (that is, before they are included in our sources) by the post-1920 distribution of capital income. We build these assumptions on Table 2 in the main body of the paper, which shows that the Gini coefficient of capital income exceeded 0.9 in 1920, 1940 and 1950. It is highly unlikely that capital income was more equally distributed before 1920. Since we do not have precise information, we will make our calculations of bias using several assumptions on the distribution.

We begin by calculating the downward bias for our pre-1920 estimates of the top decile's share of incomes. Firstly, we add the amount of dividends in kronor, D_t to the personal income total, Y_t. Then, to calculate the adjusted top decile share, T_t^* we make three different assumptions on how the dividends were distributed, with the distribution of dividends denoted α. In the

[31] 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.

[32] There was a very high peak in 1920, when dividends were already included in the *bevillning* system, at 12.5 per cent. This was a turbulent time of inflation and great profits in Swedish corporations. The share of dividends in personal income then fell continuously, to 5.9 per cent in 1930 and 3.8 per cent in 1950.

first and most likely scenario, 100 % of dividends went to the top decile (of the distribution without dividends), so $\alpha = 1$. In the second scenario, 90 %, and in the third scenario, 80 %. The equation shows the difference between the adjusted top decile share T_t^*, which takes into account the dividends, and the original estimate, T_t:

$$T_t^* = \frac{T_t Y_t + \alpha D_t}{Y_t + D_t}$$

Then, to the estimation of bias for our pre-1920 Gini coefficients. This is a trickier adjustment as we cannot simply add in dividends to total income and to top group income as in the previous exercise. We do not have full distributional data for the dividends, but just the estimated Gini of somewhere between 0.95 and 0.99, based on Table 2. We define G_t as observed Gini without dividends, C_t^D as the concentration coefficient of dividends, and w_t as the dividends' share of adjusted personal income. We get then a new adjusted Gini G_t^* which comes out as a function of the original Gini, weighted by the original personal income's (Y_t as above) share of adjusted personal income ($Y_t + D_t$), plus the concentration of dividends:

$$G_t^* \approx (1 - w_t)G_t + w_t C_t^D$$

Again following the data in Table 2 on the distribution of capital income in 1920, 1940 and 1950, we set three alternative values for the concentration of dividends (C_t^D): 0.95, 0.97, and 0.99.

Table A.2: Robustness check for the exclusion of dividends, 1870–1910

	Top decile share				Gini			
Year	Baseline	$\alpha = 0.8$	$\alpha = 0.9$	$\alpha = 1$	Baseline	$C_t^D = 0.95$	$C_t^D = 0.97$	$C_t^D = 0.99$
1870	53.8	54.1	54.2	54.3	66.9	67.2	67.3	67.3
1880	53.1	53.8	54.0	54.2	66.6	67.3	67.3	67.4
1890	48.2	49.6	50.1	50.5	62.7	64.1	64.2	64.3
1900	47.5	49.2	49.7	50.3	61.3	63.1	63.2	63.3
1910	46.0	47.9	48.4	49.0	61.6	63.4	63.5	63.7

Note: Baseline estimates for top decile's share as presented in Figure 2 and baseline estimates for Gini as presented in Figure 1 in the main body of the paper. Adjustments as discussed in the text. The size of dividends estimated from Flodström (1902, 1911) and Möller (1962). The assumptions on the distribution of dividends made on the basis of the distribution of capital income in 1920, 1940 and 1950 as presented in Table 2.

In Table A.2 we show our baseline estimate of the top decile's income share and of the Gini coefficient, in each case alongside three alternative calculations of what inequality would have been had we had data encompassing dividends. The alternative scenarios are presented from left to right in ascending degree of inequality: for the top income share, from $\alpha = 0.8$ (80 per cent of dividends accrue to the top decile) to $\alpha = 1.0$ (100 per cent), and for the Gini, from a concentration coefficient of dividends (C_t^D) of 0.95, to 0.99. Note that the assumptions on concentration of dividend income are quite strong and especially the $\alpha = 1.0$ and $C_t^D = 0.99$ estimates are estimates of the maximum level of bias.

For the size of the bias under varying assumptions, compare the baseline estimates to the estimates under the various assumptions. The results in Table A.2 shows that the dividends-related bias is very minor in 1870 and 1880. By 1900 the bias has become larger as the dividends have become a larger share of personal income in the country. With the strongest assumption on the concentration of dividends, that only people in the top decile of the income distribution received dividends ($\alpha = 1$), by 1900 we underestimate the top decile's share of incomes by 2.8 percentage points (50.3 compared to our 47.5 pp baseline) and by 1910, the underestimatation is 3.0 percentage points (49.0 compared to 46.0 pp). The Gini coefficient also sees little difference (< 1 point) in 1870 and 1880 even with the most inegalitarian assumption on the distribution of dividends ($C_t^D = 0.99$), while in 1900, the bias is 2.0 points and in 1910, 2.1 points. The overall takeway from Table A.2 is that a proper consideration of dividend incomes reinforces our interpretation of inequality in Sweden between 1870 and 1920 as high and trendless. In the baseline estimates in the main body of the paper, 1920 is a temporary peak in inequality which contrasts to the small decline from 1900 to 1910, but the calculations here rather supports the idea of not much change in 1900, 1910 and 1920, before the onset of inequality decline after 1920.[33] The results of this exercise, therefore, suggest that our main conclusions and interpretations are robust to the exclusion of dividends from the *bevillning* data before 1920.

[33] Such an interpretation accords well with the social and political history of Sweden during these years. 1909 famously saw a General Strike with bitter enmity between labour and employers, and the 1910s saw severe political polarisation over issues such as the power of parliament and the power of the King. See on labour conflict Lundh (2010), ch. 3, and on political conflict Lewin (2017, chs. 3–4).

B Online Appendix B: Construction of group-specific incomes for non-filers

As discussed in the Data section (Section 3) of the main body of the paper, not all adults paid the income tax (the income *bevillning*) in the period of study. In 1870–1910 about 85 % of adults did not pay the tax, in 1920 around half, and in 1950 about one third; see Table 1. Not paying the tax came from two things: either being a farmer, or not earning enough to pay the tax. Thus, in this Appendix we focus on the construction of income measures for these two groups: farmers, and the low-income earners. Appendix B.1 discusses farmers, B.2 workers in agriculture, and B.3 workers in other sectors. In each case, the imputed incomes build on extensive work with historical sources, published statistics, and previous research. In Appendix B.4 we compare the implied total personal income from our data with personal income from historical national accounts, and show that our total income estimates closely align with the national accounts. Another robustness check is presented in Appendix C.2 where we compare our inequality estimates with estimates where we instead have imputed homogenous incomes for non-taxpayers, the top-down approach discussed by Atkinson (2007).

B.1 Incomes of farmers

Unlike the wage earners, who were primarily covered by the income *bevillning* and for whom the non-filer incomes are therefore bounded by the income threshold to pay the tax (so for example in 1870 any wage-earning non-filer's income y is bounded $0 \leq y < 400$ kronor), the farmers were covered by the property *bevillning* and their incomes are therefore *not* bounded by the threshold. For this reason, imputations of farmers can be wider, will make more a difference for the resulting distribution, and must be treated separately.

The farmer group was internally heterogenous and we estimate incomes for farmers divided by two factors. Firstly, the size of the farm, from small to large, and secondly, by broad region, from Götaland in the south to Norrland in the north.

We define small farms as <10 hectares of arable, medium-size farms as 10-50 hectares, and large farms as >50 hectares. This typology has been created based on the agrarian history literature of the period, which clearly delineates the different conditions under which small farmers and the wealthier ones operated (Kuuse, 1970; Bäcklund, 1988; Morell, 2001), and based on the public agricultural statistics, which throughout the period presents data on the number of farms by farm size.[34] The regional division is important because of the quite different climate in northern and southern Sweden and the differing conditions especially for growing grain.

Our estimates of farmers's incomes begin with the best source available: the agronomist Ludvig Nanneson's farm surveys from 1914–15, 1920–21, 1944, and other interwar years (Nanneson, 1916, 1923, 1946).[35] Nanneson's farm surveys (the 1914-15 survey encompassed 100 farms, the

[34] The official agricultural statistics of Statistics Sweden during the 1885–1910 period used four group farm size groups, <2 ha, 2-20 ha, 20-100 ha and >100 ha, and for 1919–1951 they used <2 ha, 2-10 ha, 10-20 ha, 20-30 ha, 30-50 ha, 50-100 ha and >100 ha. In our typology the <2 ha group is defined as semi-landless as it is clear from the historiography and historical farm surveys that a family could not get by on so little land without labouring for others.

[35] For a history of Nanneson's pioneering work see Hjelm, 1995. Nanneson, who was born in 1883, was a teacher at an agricultural school in the early 1900s, wrote a textbook in "Rational Farm Accounting" in 1913, founded a string of farm accounting socities in 1915 and onwards, and became a professor of agricultural economics in 1933. From 1914 to 1950 he published 36 studies under the headline *Accounting Results from Swedish Farms* (*Räkenskapsresultat från svenska jordbruk*).

1920-21 survey 300-400 farms) were meticulous mappings of all kinds of labour performed at the farm (paid or unpaid) and all kind of income sources (farm income, subsistence production, incomes from by-employments), and they are incredibly helpful to map the incomes of farmers under various conditions (farm sizes, geography).

Nanneson and his team were highly interested in the incomes and living standards of farmers specifically as they were concerned with the sustainability of Swedish agriculture. For this reason, they spent a lot of effort mapping the balance between labour input and income on the farms, and throughout compared the income of farmers with the incomes of workers, as they were concerned with an exodus from agriculture if being a farmer yielded too little (relative to one's labour input) compared to being a wage earner. In the 1920-21 study, for example, Nanneson demonstrated that a farmer owning around 10 hectares could bring home an income equivalent to that of an industrial worker, but that the farm's overall income less that income would provide no surplus income, analogous to profits on invested capital. The study of the 1940s, on the other hand, found that a 10-20 hectare farm could deliver the income of an industrial worker, plus a benchmark 5 per cent return on invested capital (SOU, 1946, pp. 55-56). We use these detailed studies of labour input, capital input and incomes of Swedish farms as the benchmarks for how large incomes farmers (of various sizes and in various regions) had.

For the 1930s and 1940s, the sources for incomes and living standards on farms of various types are even better than the sources for the 1910s and 1920s. Nanneson-style studies got more support as agricultural policy in the 1930s increasingly consciously was oriented towards a structural transformation of the sector: decreasing the number of farms through consolidation to more efficient farm sizes, enhancing productivity, and improving the efficiency of farm labour. In the 1930s, agricultural policy explicitly set the goal that a "standard-size" farmer should achieve an income equivalent to a male industrial worker (Nanneson, 1946; Morell, 2001, pp. 171–180). The 1937 Agricultural Census identified that only a quarter of Swedish farms were big enough to live up to the new stated policy aim that a typical farm should yield an income level for the family farmer equivalent to the income of a male worker in manufacturing. Half of the farms were classified as "incomplete family farms" (*ofullständiga familjejordbruk*), in need of growth through consolidation, and about a quarter as larger farms which were mostly worked by labourers rather than the owners themselves (Grönlund, 1942, p. 166). Again we see how important the income comparison with the worker outside of agriculture (in practice, male workers in manufacturing) was for the contemporary policy discourse around agriculture. This is helpful for our purposes as it reinforced policymakers' and statisticians' interest in calculating farmers' incomes.[36]

In line with these policy aims, Nanneson (1946) presented calculations of returns for farms of various sizes (from 2–5 ha of arable land to 100+ ha) and in various regions for the years 1939–1943. The largest return per hectare was on the small farms, with a declining tendency according to size. Net returns on the other hand had the opposite tendency, increasing with farm size. For the very small farms the net rate of return was 3 per cent, the mid-size with 10-20 hectares had 3.9 per cent, the group with 30-50 hectares had 6.2 per cent, and those with 50-100 hectares had 6.3 per cent (Nanneson, 1946, p. 19). Nanneson also shows that regional differences were very large. Gross returns for small farms were almost twice as large in the very South as in the very North (where the climatic conditions for agriculture are less amenable); for mid-size farms, it was 67 per cent higher. The unequal yields for differently sized farms were a constant concern

[36] In the agricultural policy of 1947, the definition of "basic farm" (*basjordbruk*) which for the farmer should deliver the industrial worker-norm living standard was set to 10–20 hectares of crop acreage, and in the 1950s the norm was revised to 20 to 30 hectares. The 1952 regulation of agrarian prices was set to, again, guarantee a living standard corresponding to that of an industrial worker for these farmers. See discussion in Odhner, 1953, pp. 175–202.

for policy in the postwar period (e.g., SCB, 1955).[37]

These sources, however, only begin in the 1910s. To calculate pre-1910 farmer incomes, we have used the Nanneson estimates as benchmarks from the 1910s, and projected incomes back in time especially built on three pieces of information: (a) the farmer/worker income ratio of the 1910s (with the farmer group divided by size and region), (b) farm prices from 1870 to the 1910s, and (c) wages from 1870 to the 1910s.

Nanneson's 1914-1915 study indicates that farmers earned significantly more than farm workers. Depending on the size of the farm and the geographical area, the farmer to worker ratio ranged from 3 to 9 in 1915. A backward projection of these ratios should take into account that the earnings of these two groups have changed over time. We proxy the changing ratio of farmer/worker income by the land price/wage ratio, available from Åmark (1923) and Bohlin and Larsson (2007).[38] Åmark (1923) constructed his price series based on 4,854 sales transactions, presenting the data in three-year averages stretching from 1876/78 to 1918/20 for counties in central and southern Sweden.[39]

The backward projection of farmers' incomes using land prices shows that farmers' incomes relative to farm workers' peaked in 1880, when the ratio ranged from 3.7 to 13.6 depending on farm size and region. We have hence corroborated the literature on the wage rental ratio, which showed that wages have outgrown land prices in the longer term, also across the twentieth century (Bohlin and Prado, 2011). The declining income gap between the two groups occurred against the backdrop of a discussion of an "agrarian crisis" in the 1880s and 1890s, a turn to tariffs, and a later recognition that the 1870s and 1880s were the heydays of the Swedish farmers, when their political and social (and economic, as shown here) advantage over the workers was the most marked (Morell, 2001, pp. 84–120).

The farm surveys of Nanneson and the public statistics using the same methodology are our main source for farmers' incomes from the 1910s to 1950, and the land price/wage ratio has been our guiding principle in extrapolating back to 1870. However, these are not the only pieces of evidence available. We have also considered the wealth of farmers and the semi-landless as indicated by probate inventories (Bengtsson and Svensson, 2019, 2022) and above all, we have made a new sample of about 13 000 property *bevillning* taxpayers, to complement our much larger sample of the income *bevillning*. We now turn to this dataset.

As mentioned in Section 3 in the main body of the paper, farmers were taxed in the property *bevillning*, instated alongside the income *bevillning* which is the main source for our study. We have for this paper gathered a 1 per cent random sample of property tax returns in ten counties (out of 24) in 1900, and for four counties also the years 1870 and 1920 (Bengtsson, 2022). In total, the analysis builds on 12 709 property tax returns. The four counties with more extensive

[37] At least two investigations from the relatively poor north of Sweden show that the 1940s and 1950s saw an exodus of small farmers from their agrarian upkeep and into becoming manufacturing workers (Johansson, 1996, pp. 28-29; also Isacson, 1994 for Gagnef municipality in the northern-central region of Dalarna). This indicates that in fact small farms by the 1940s and 1950s were in economic terms less attractive compared to wage labour in manufacturing, at least farms on the less fertile soils in the northern parts of the country.

[38] The motivation for using land price as a proxy for income from land, i.e. farmer income, is the following. The price of a piece of land mirrors the future stream of income from owning it, discounted to present value by applying a suitable rate of interest. If income from owning land increases, whether attributable to increased productivity or increasing terms of trade for agricultural products, land prices will follow suit.

[39] The investigation further excluded farmlands smaller than 5 hectares, and farms with forestland and pasture that were more than three times the size of the crop land. The number of purchases varied greatly by the size of the arable land. Purchases of farm properties with more than 100 hectares of crop land do not figure prominently in the sample; in some locations and years, there were only a few purchases of properties of that size.

property tax samples are Malmöhus, known as a markedly unequal rural society with large estates and many landless workers, Älvsborg, an intermediate county with some rich plains but also poorer forested areas, Jönköping, a county with few estates and many small and mid-sized farms, and Västmanland, a county from the central Swedish industrial belt. (On the background see Gadd, 2000, pp. 42–47; Morell, 2001, pp. 36–37.) The estimates of mean and median wealth for rural property taxpayers by county in Table B.1 shows that there were marked differences between the counties, as one would expect. The Malmöhus property taxpayers, living on the best soil in the country (cf. Bengtsson and Kersting, 2026) were on average much wealthier than the ones in Älvsborg and Jönköping, but the mean–median discrepancy in Malmöhus also reveals that the distribution there was highly skewed, more so than in Älvsborg or Jönköping with their more farmer-dominated social structures.

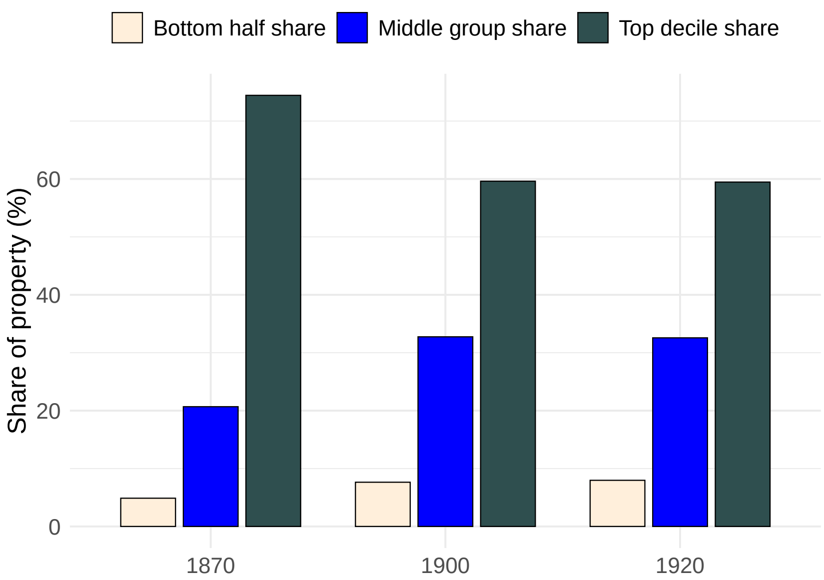
Table B.1: Evolution of taxed property in the countryside, four Swedish counties 1870–1920

County	Measure	1870	1900	1920
Malmöhus	Average	19 360	8 168	15 994
	Median	2 000	1 500	3 800
	P90	21 400	15 600	27 880
Älvsborg	Average	4 342	3 776	8 758
	Median	2 000	2 200	3 800
	P90	7 330	8 520	16 000
Västmanland	Average	11 537	11 188	9 487
	Median	4 500	3 050	5 000
	P90	29 340	16 600	24 000
Jönköping	Average	NA	5 062	7 988
	Median	NA	3 000	5 000
	P90	NA	9 290	16 800

Note. Built on a 1 per cent random sample of property taxpayers in four Swedish counties: Malmöhus, Älvsborg, Jönköping and Västmanland. The Jönköping sample in 1870 is too small to make calculations. Only property owners are included in the calculations. See Bengtsson (2022, Table 6) for details.

The inequality among farmers comes out even stronger in Figure B.1 which highlights the distribution of taxed property among the property taxpayers in 1870, 1900 and 1920. In 1870, the ten per cent of taxpayers with the most real estate owned 82.2 per cent of taxed real estate in Malmöhus, 53.2 per cent in Älvsborg, and 49.7 per cent in Västmanland. Inequality declined in Malmöhus and Älvsborg to 1900, when the top decile shares were 62.7 and 43.7 respectively but rose in Västmanland. In Jönköping, where the sample is too small in 1870 so our calculations only begin in 1900, the top decile share in 1900 was 44.0 per cent. Inequality was relatively stable from 1900 to 1920: it grew in Älvsborg, but decreased in Västmanland, and was stable in Malmöhus and Jönköping. When all four counties are taken together, it reinforces the picture of especially high inequality in 1870: the top decile's share measured on the level of all four counties fell from 74.4 per cent in 1870 to 59-60 per cent in 1900 and 1920. The Gini coefficient for rural wealth in the four counties is estimated as 81.1 in 1870, 70.7 in 1900 and 70.4 in 1920. A decomposition shows that within-county and between-county inequality are both important parts of inequality as a whole (Bengtsson, 2022). The high inequality of the agrarian sector shown in B.1 reinforces our conviction that it is important to differentiate our estimates of farmer incomes by farm size, and also supports the results in Figures 4 and C.4, that rural and agricultural

Figure B.1: Distribution of property in the countryside, four Swedish counties 1870–1920



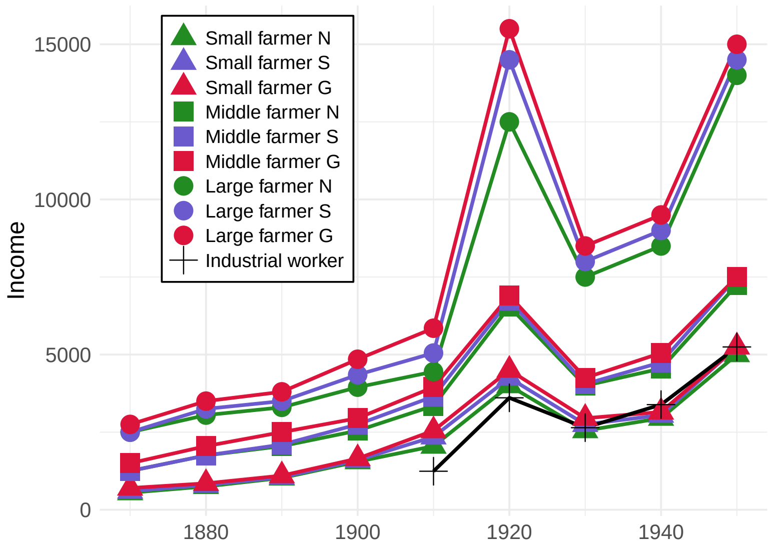
Note. Built on a 1 per cent random sample of property taxpayers in four Swedish counties: Malmöhus, Älvsborg, Jönköping and Västmanland. Only property owners are included in the calculations. "Middle group" is percentiles 50 to 89, that is, all property owners between the bottom half and the top decile. The Gini coefficient for the four-county sample is 81.1 in 1870, 70.7 in 1900 and 70.4 in 1920. For 1900 we also have a 1 % sample of six other counties: Uppland, Södermanland, Kalmar, Halland, Göteborg och Bohus län, and Skaraborg. Calculating inequality for the 10 counties sample for 1900 yields a Gini of 70.5, so no essential change compared to the four county sample. See Bengtsson (2022) for details.

inequality was in fact steep in Sweden in the late nineteenth and early twentieth century.

Building on the sources discussed so far, especially the farm surveys from the 1910s to the 1940s, the land prices and the wage series, Figure B.2 shows our estimated incomes for farmers: divided into the three size groups of farms, and into the three main regions of the country. For comparison, we also include the average wage of a male worker in engineering industry as a benchmark; as we have discussed above, for much of the period this was an explicit frame of reference for agricultural policy. The estimates highlight the wide variation between farmers of different wealth levels, as implied already by Figure B.1, but also how manufacturing workers in the 1930s and 1940s caught up with the small farmers, income-wise. For the interpretation of the graph it is important to consider the relative weights of the various farm categories: by 1950, there were 282 000 farms with up to 10 hectares of arable land (defined as "small" here), while there were only 89 000 farms in the middle category of 10–50 hectares, and only about 8 000 farms in the large category. (Calculations built on SCB, 1959, Table D.14.) Thus, the huge income advantage of the category of large farmers reflect their relative exclusivity.

Given the rich set of sources that we have worked with to estimate farmer incomes – farm household surveys, land prices, property tax returns, probate inventories, agricultural statistics, contemporary policy reports, and wage series – we are confident that the estimates are sensible. See also Online Appendix B.4 for a comparison of our implied total income (for all groups) with

Figure B.2: Estimated incomes of farmers, 1870–1950



Note. Group-specific estimations as discussed in the text. Farmers divided by region: N = Norrland, S = Svealand, G = Götaland. The series "Industrial worker" is an adult male industrial worker from the HILD database, Lundh and Prado (2020).

the total income from historical national accounts.

B.2 Landless and semi-landless workers in agriculture

The group of landless and semi-landless in agriculture was substantial in size in the late nineteenth and early twentieth centuries, and typically did not earn enough to pay income tax. Our estimates, comparing the taxation sample as described in Section 3 of the main body of the paper with census data, suggest that still in 1940 and 1950, only about one third of agricultural wage-earners paid the income tax. Alongside domestic servants, agricultural workers were the most low-paid group of workers, and their plight awoke much concern in Sweden in the late 1800s (von Feilitzen, 1890; Hultgren and Landergren, 1891) as well as in the 1930s and 1940s (see discussion in Lundh, 2008).

It is of help that in principle, unlike for the farmers who were covered by the property tax system, the incomes of non-filing *wage-earners* are bounded by the threshold to pay the income tax: 500 kronor in 1890 and 1900, 600 kronor in 1910, 1920 and 1940, and so on. However, for the agricultural wage-earners this principle is complicated by the fact that much of their pay was not in cash but in-kind[40], and that even though this in theory was valued and covered by the income tax, it appears that the subjective nature of the value of lodging and food made the income tax coverage of the agricultural workers to lag behind other workers, even at comparable income levels. We draw this conclusion based on a comparison of the wage data available for agricultural workers in comprehensive public statistics (collected in Lundh and Prado (2020))

[40] Lundh (2012) shows that agrarian contract labourers in the 1880s received more than 70 per cent of their remuneration in kind, in housing and food. In 1920, the share had fallen to a still significant 57 per cent.

as well as in extensive historical research (Lundh and Olsson, 2011; Lundh, 2012; Lundh and Prado, 2015), with the tax coverage that we have calculated.

We divide the labouring classes within agriculture in two main groups: the semi-landless, and the "pure" wage-earners. The semi-landless group (*torpare*) had access to some land on a tenancy basis, and combined small-scale subsistence agriculture on this land with labour for the landlord to pay the landrent, and wage labour more generally. (See Bengtsson and Svensson, 2022; Uppenberg and Nilsson, 2025.) Their living standards were slightly better than those of the "pure" wage-earners who had no land at all.[41]

To estimate the incomes of the semi-landless and the landless agricultural labourers from 1870 to 1950, we have used both contemporary investigations of living standards (von Feilitzen, 1890; Hultgren and Landergren, 1891), public investigations (Kommerskollegium, 1911; Socialstyrelsen, 1923) and the extensive historical research into agricultural wages and living standards (Lundh and Olsson, 2011; Lundh, 2012; Lundh and Prado, 2015; Bengtsson and Svensson, 2022).

Figure B.3 shows our estimates for the lower classes in agriculture: agricultural workers, and crofters. These are equivalent to the estimates for farmers in Figure B.2 and again, these are compared to a male engineering wage, the black line. Unlike the middle and large farmers, the agrarian labourers were in the deficit compared to manufacturing workers, and more markedly so over time (cf. discussion in Lundh, 2008, 2012; Lundh and Prado, 2015.)

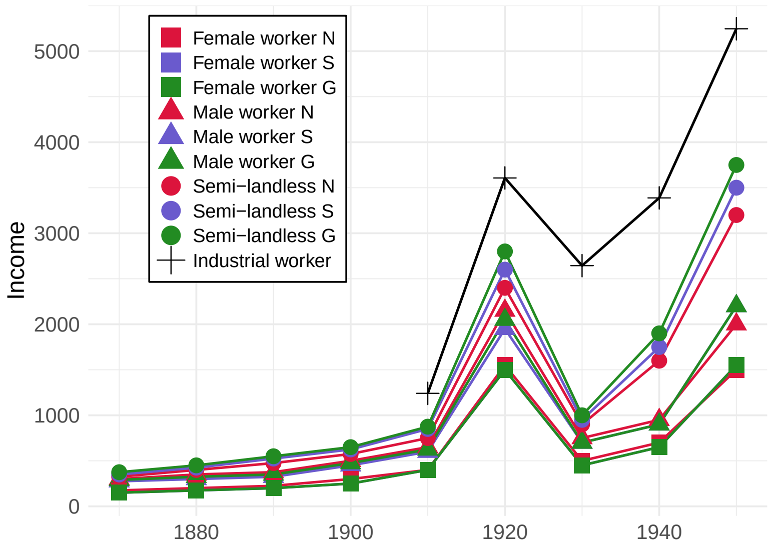
The regional pay differences estimated in Figure B.3 are informed by the rich availability of county level wages from 1913 onwards, produced by Statistics Sweden and collected in the HILD database (Lundh and Prado, 2020). Overall these suggest that for an important group such as the farm labourer, differences between the three regions were relatively small but that a nominal pay difference for Norrland in 1870 was turned into an advantage for Svealand and Götaland in 1940 and 1950.

B.3 Incomes of non-agrarian workers below the taxpaying threshold

For estimating the incomes of workers below the tax-payment threshold, a great strength is that there is a large literature in Swedish economics and economic history describing workers' wages, building especially on corporate archives. The main previous study of manufacturing wages in our period is *Wages and Prices in Sweden*, volume 2:1 (Bagge et al., 1933). This massive work presents much firm-level wage data, and we have used their wage data as a set of useful reference points, but only as such. There are, as Hamark et al. (2025) clearly explain, several reasons why the *Wages in Sweden* data in fact overestimate average wages. To begin with, the authors of WiS themselves, Bagge, Lundberg and Svennilson, had the intention to capture wage trends

[41] There is actually a third group: unpaid family labour. At family farms during our period of study, this was an important source of labour inputs: from unpaid wives and children. Public investigations at the time (SOU, 1938, pp. 96-99) as well as research in agrarian history (Flygare and Isacson, 2003) demonstrates the centrality of unpaid wives' labour. It is also very clear that teenage and young adult children provided much labour input; the social historian Sten Carlsson (1968, p. 38) has curtly remarked that sons and daughters staying at home were for the farmer "a very cheap solution of the labour problem, as the wage typically was constricted to food and housing". The children would, argued Carlsson, stay and work for free for two reasons: waiting to take over the farm, and because of the lack of alternative employments. The role of this unpaid group of wives and children is of course a central component of our discussion of inequality in agriculture and the role of structural transformation in the main body of the paper. Of course, the importance of unpaid family labour was not a unique Swedish feature: cf. US studies by Craig (1991), Folbre and Wagman (1993) and Chiswick and Robinson (2021). Precisely because this was such a universal feature of family agriculture – and is still today (cf. Fields, 2004, pp. 728-730; Gollin and Kaboski, 2023, p. 1349) – we belive that the disapperance of this no-pay group is of great importance for the study of long-run income inequality.

Figure B.3: Estimated incomes for semi-landless and agrarian workers, 1870–1950



Note. Group-specific estimations as discussed in the text. Agrarian groups divided by region: N = Norrland, S = Svealand, G = Götaland. The series "Industrial worker" comes from the HILD database, Lundh and Prado (2020). The HILD data begins in 1913.

rather than precise wage levels; they called this a "kinetic" method. Therefore, they were more concerned about consistency over time than about precision at any given point. As Hamark et al. highlight, the WiS data overrepresent skilled workers, full-time rather than casual workers, and employees of larger firms. All these factors mean that all *Wages in Sweden* wage series should be considered ceiling estimates of manufacturing wages: very relevant as reference points, but not useful as actual estimates of average wages.

Prado (2010) offers a more recent comprehensive account of wages, focusing on average wages within each manufacturing sector, for men and for women. Several more specialised studies are also available: Larsson (1986) on an iron works from the 1820s to 1905, Leffler's (1897) survey of women workers in Stockholm in the late 1890s, Norlander (2000) about the Liljeholmen candle factory in Stockholm from 1872 to 1920, and Stanfors et al. (2014) on the turn-of-the-century tobacco industry. In the 2010s, economic historians at the University of Gothenburg presented the most comprehensive resource on historical wages: the HILD database (Lundh and Prado, 2020). We make extensive use of this database to infer the movements of wages below the tax threshold: even though of course the HILD database captures workers overall – taxpayers as well as non-filers – it is useful for the movement of wages.

Domestic servants constituted a large low-wage group within the service sector and only rarely earned enough to pay the *bevillning* tax. We estimate – comparing our tax payer sample with Census data – that in 1900, only 1 per cent of the domestic servants earned enough to reach the *bevillning* threshold. For this group we have existing wage estimates from the 1911 from a public investigation (Kommerskollegium, 1911), from the 1920s and 1930s from public Stockholm statistics and historical research (SSSK, 1924; Nordlund Edvinsson and Söderberg, 2010), and extensive statistics from the mid-1930s from a public investigation (SOU, 1938, pp. 100-101).

B.4 A robustness check on the imputations: comparing our national income to the Historical National Accounts

A potential concern is that the incomes we assign to farmers and workers below the *bevillning* threshold are imputed rather than based on individual-level income observations. We would reply that, firstly, incomes below the tax threshold are, by definition, bounded above by the threshold itself (400 kronor, 500 kronor, and so on). Thus, the imputations for groups below the threshold can have only a limited impact on aggregate inequality measures.

This is a general insight of the long-run income inequality literature building on tabulated data (e.g. Atkinson, 2007), and we demonstrate its relevance in our case in Online Appendix Section C.2 where we compare our inequality estimates for 1870 and 1900 to estimates employing a "top-down" approach to non-filers' incomes. Moreover, as highlighted in Appendices B.1, B.2 and B.3, we have taken care to anchor all estimates of farmers and workers below the tax threshold in contemporary sources. Our final validation test is that in this section, we compare our implied national income to that produced by Historical National Accounting (HNA), which has estimated GDP and national income *from the production side*. The comparison is particularly informative because our comprehensive reconstruction of personal incomes is, analytically, close to an *income side* national account. Apart from retained corporate earnings and certain tax items, aggregating incomes for taxpayers and non-taxpayers alike yields a measure comparable to national income. At the same time, the underlying sources and estimation procedures differ fundamentally from those used in production-side historical national accounting.[42] This means that the comparison with the production-side historical national accounts estimates provide a strong test of the validity of our estimated incomes for non-filers.

Since Lindahl et al. (1937), Swedish HNA research has primarily relied on the production-side approach, aggregating value added across sectors to form GDP; this is also true for the two most recent vintages of Swedish HNAs (Historical National Accounts), those of Edvinsson (2014) and Schön and Krantz (2015). The production-side approach differs fundamentally from our approach, which aggregates incomes across population groups to form GDP.[43] However, in principle, national accounting implies that value added can be decomposed into payments to the factors of production. In theory, hence, our method should equal GDP obtained from the production side. To make our income estimates comparable we make a few adjustments, adding in the distributed dividends before 1920 (as in Online Appendix A.4) and also adding in the implied flows of income from owner-occupied housing, a post which does not exist in our *bevillning*-based data.

Table B.2: Implied national income estimates, compared to the Historical National Accounts

Year	1870	1880	1890	1900	1910	1920	1930	1940	1950
Edvinsson	1.07	1.03	1.09	1.03	1.05	0.85	0.89	1.03	0.96
Krantz and Schön	1.03	1.07	1.14	1.02	1.02	0.82	0.86	0.99	0.93

Note. The table presents the ratio between the national income implied by our bottom-up approach, and the national income estimated by two research efforts within Historical National Accounts, using the output method of estimating GDP: that of Edvinsson (2014), and that of Schön and Krantz (2015).

[42]For discussion of the discrepancy between personal income and national income from an inequality studies perspective, see Atkinson (2007, pp. 29–32) and Piketty et al. (2018, pp. 560–573).

[43]For discussion of how income-based and output-based estimates of national income or GDP can diverge in the historical national accounts for this period, see for example Fremdling's (1988) study of Germany 1850–1913.

Table B.2 compares the national income estimates implied by our bottom-up approach with those reported in the output-based HNA series. With the exception of the highly volatile year 1920, our estimates differ from the HNA estimates by no more than 11 per cent. Given the complexity of historical national accounting and the fact that the two approaches rely on fundamentally different sources and estimation procedures, this degree of correspondence provides strong support for the validity of our income estimates, including the imputations for individuals not observed in the *bevillning* records.

C Online Appendix C: Additional estimates

C.1 Alternative population concepts

One of the most difficult issues in measuring income inequality over long periods of time is how the results are affected by demographic change. In the studies of the United States that have been highly influential in the inequality literature, the measurement unit is typically the tax unit, which is a married couple or a single person. Auten and Splinter (2024) contend that some of the growth in inequality observed since the 1970s by Piketty et al. (2018) is simply due to demographic change: top income earners are as likely as before to marry (and thus form a single tax unit out of two income-earners) while lower- and middle-income individuals are less likely to marry today, and more likely to live in single-earner households.[44] This "inflates" the number of tax units (per capita) compared to the 1970s and makes concentration measures like the top percentile share of income look steeper. Analyzing instead the income distribution among individuals, yields lower inequality estimates, argue Auten and Splinter (2024, pp.2192-2193); see also Larrimore et al. (2021).

We study a very long period and so demographic patterns and demographic change are therefore highly relevant to our analysis. We may note that the income tax data that we use report individual income, but that the non-reporting for farmers' wives implies that, in the agrarian sector, in practice our measure approximates a tax unit level estimate, since we analyse the head of household in farm couples, alongside single men and women. This applies, for example, to the Gini estimates in Figure 1 of the main body of the paper, and the Gini and counterfactual Ginis in Figure 7.

Measuring incomes of women in historical settings is difficult given the state of the sources. In current research in economics, the distribution of resources between spouses in agrarian and/or informal sector households is an active area of research (Chiappori, 1993; Buchmann et al., 2025). Figure C.1 shows our baseline Gini estimate and then the Gini without farmer wives, or when imposing an assumption of an equal split of incomes between farmers and farmers' wives. The trend is the same over time, but the level of inequality becomes substantially lower if we exclude farmers' wives. Figure C.1 clearly highlights what a considerable difference the treatment of farmers' wives makes. This mirrors Piketty et al's (2018, Figure VI) demonstration that the top decile share of income is 5–10 percentage units lower in the United States from the 1910s to the 1970s when using the equal split assumption rather than the individual measure.

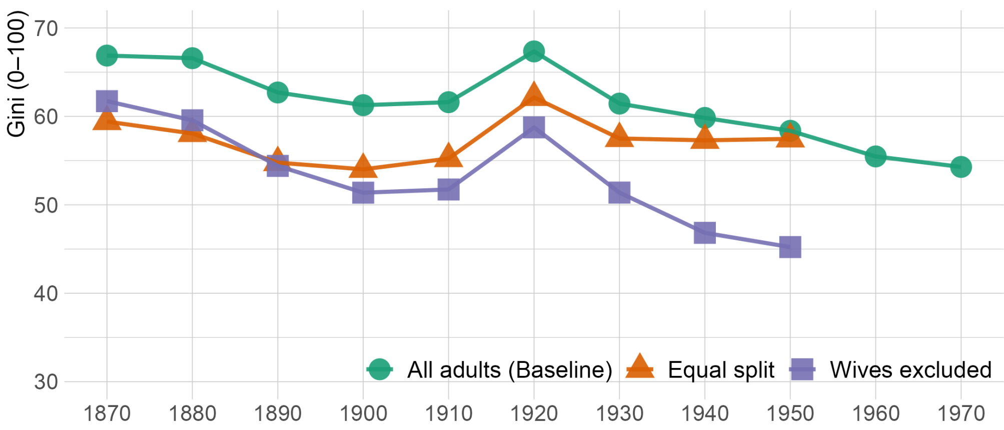
From Figure C.1, we conclude that our baseline individual-level estimates as presented in the main body of the paper indeed represents a ceiling estimate of income inequality. We cannot estimate household level inequality with our data, but the difference shown by the estimates when farmers' wives are excluded, indicates that household level inequality was lower than individual level inequality.

C.2 A comparison of bottom-up and top-down approaches to non-filers

As we have discussed in Section 3 in the main body of the paper, and in Online Appendix B, we have devoted considerable effort to estimating incomes for non-filers. This stands in contrast to the "top-down" approach where a fixed income is assigned to all non-filers, for example 50

[44] They reference Larrimore (2014) who estimates that the class-specific changes in marriage rates explains 23 % of the growth in the household-level Gini coefficient from 1979 to 2007. Our individual-level estimates are not sensitive to the changes in joint tax filing, but of course, one's labour behaviour is influenced by one's family situation.

Figure C.1: Gini with various population concepts



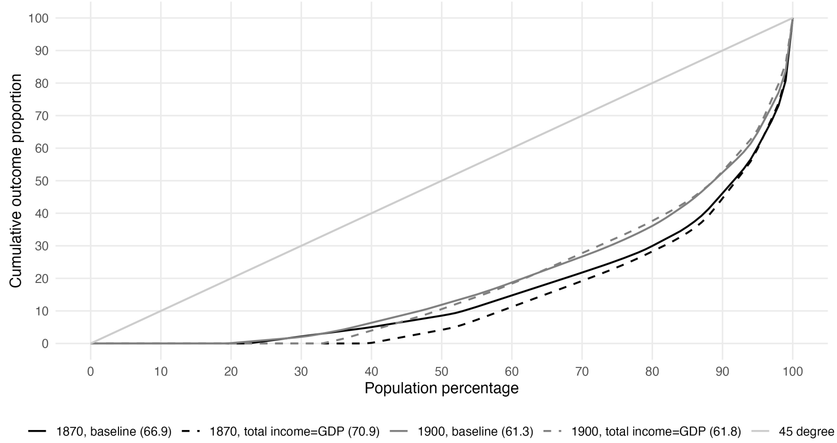
Note: Illustrates the importance of the population concept. The series "Equal split" splits the income of a married farmer equally between him and his wife. The series "Wives excluded" drops all farmer wives from the calculations.

per cent of the tax threshold (cf. Atkinson, 2007). Our "bottom-up" approach is important as it makes possible our sector specific and occupation specific estimates. But it is less common in the historical income inequality literature; most other studies either assign a uniform income below the threshold, or use a parametric distribution to assign incomes to non-filers. A prominent example of the first approach is the canonical study by Piketty and Saez (2003) on the United States, where they assign an income of 20-50 per cent of the threshold[45]; another example is the study of income inequality in Sweden since 1903 by Roine and Waldenström (2008).

Against this background, it is crucial to demonstrate that our main inequality results are not sensitive to the choice between bottom-up and top-down imputations. Figure C.2 shows that in the aggregate, had we used the top-down approach, the results would have been similar. We implement this alternative assumption for two particularly challenging years, characterized by relatively low taxpayer coverage: 1870 and 1900. In 1870, the estimated Gini coefficient with our preferred approach is 66.9, while with a top-down approach, it would have been 70.9, a difference of about four Gini points. In 1900, our baseline estimate is 61.3, while with the top-down approach the estimate is 61.8, a difference of less than one Gini point. Both baseline figures coincide with the Gini for all adults reported in Figure 1, as they are computed on the same population and with the same estimator. The top-down approach yields marginally higher inequality here because anchoring total income to output-side GDP, which exceeds the income total implied by our bottom-up reconstruction, scales the imputed non-filer incomes upward and thereby widens the gap to the large group of adults with little or no recorded income. The resulting differences are nonetheless small and do not alter our interpretation of long-run inequality trends or the role of structural transformation emphasized in the main text. This is particularly noteworthy given that 1870 and 1900, the years studied in Figure C.2, are among

[45]This assumption is still underlying the pre-1962 estimates in Piketty et al. (2018); cf. methodological discussion in Geloso et al. (2022), pp. 2379–2380.

Figure C.2: Comparing the Lorenz curve and the Gini with our estimates and an alternate top-down approach to non-filers



Note: The figure compares our baseline bottom-up imputations for non-filers with an alternative top-down approach where total income is forced to equal GDP as estimated from the output side (Schön and Krantz, 2015). Gini coefficients shown in parentheses in the legend.

the years with the largest non-filer imputations. If the bottom-up versus top-down choice makes a difference anywhere, it is here.

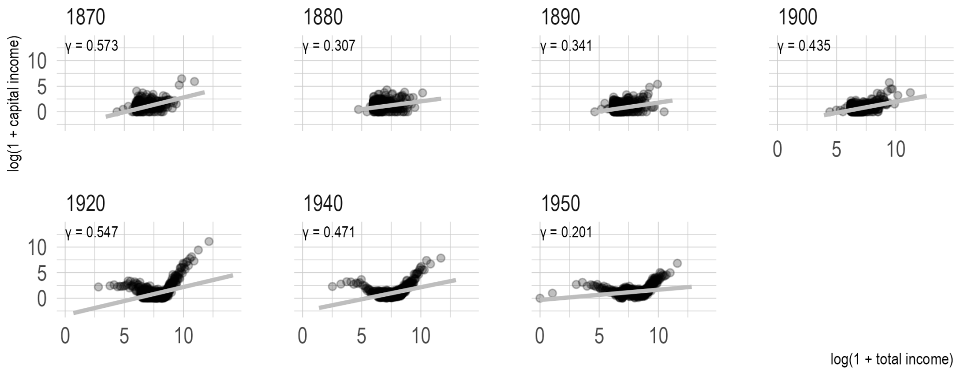
C.3 The role of capital incomes

In the literature on long-run income inequality, capital incomes play a crucial role, not least through the work of Piketty Piketty (2003, 2014). We show in Table 2 in the main body of the paper that capital incomes were indeed highly concentrated (Gini coefficients > 0.9) and so, all else equal, a larger share of capital income was associated with more inequality. However, as stressed by Milanovic (2016), the implication of the capital share for income inequality depends not only on the degree of concentration of capital income, but also on the position in the income distribution of those who receive capital income. It could be that capital incomes are concentrated to high-income-earners (as we would expect), but it could also be that they are concentrated to some group (petty rentiers) with capital income but nothing else. Figure C.3 highlights the varying correlation between a taxpayer's capital income, shown on the Y axis, and his or her total income, shown on the X axis, over time. It is important to note that the coefficient γ reported here is a different one to the coefficient β reported in Table 2. In Table 2 we correlate capital income and labour income; in Figure C.3 we correlate capital income and total income.

We see in Figure C.3 that the correlation, γ, fluctuated over time. In 1870 there appears to have been a linear and rather strong positive association – people with high total income also had high capital income. By 1950, capital income was no longer concentrated solely among top-income earners. Instead, it was disproportionately found both among low-income individuals (primarily older rentiers) *and* among high-income earners (*homoploutia* as described by Ranaldi and Milanovic, 2022). This means that the association illustrated in the 1950 panel of C.3 rather is U-shaped.

How do the results here complement those from Table 2? The declining negative correlation between capital and labour income found in Table 2 indicates a move away from a strict separation

Figure C.3: The correlation between an individual's capital income and his or her total income, 1870-1950



Note: Builds on the 1/0.5 per cent sample of *bevillning* income taxpayers in Sweden 1870, 1880, 1890, 1900, 1920, 1940, and 1950, as discussed in section 3 of the main body of the paper, and in Online Appendix A.

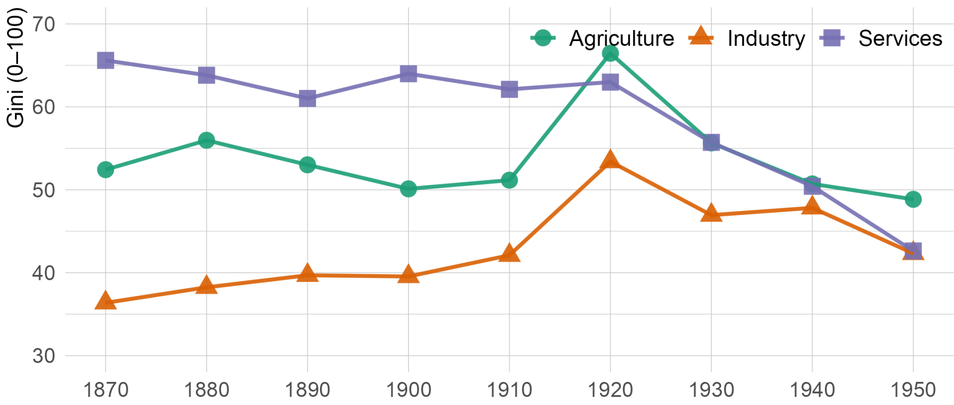
between capitalists and workers. At the same time, the changing relationship between capital income and total income – from strongly positive to more U-shaped – in Figure C.3 shows that capital income became less exclusively concentrated at the top of the distribution. Taken together, these results suggest that capital income shifted from being both highly concentrated and socially segmented to being more broadly distributed and combined with labour income, thereby making capital income less inequality-enhancing than before.

C.4 Income inequality by sector

In Section 5 of the main body of the paper, we analyse inequality by urban and rural residence, examine changes in the agricultural employment structure, and compare the incomes of workers and capital owners in agriculture and industry. However Figure 4 presents the Gini coefficients for *urban and rural* income earners separately, and as we emphasise in the text, "urban" is not synonymous with "industrial", nor is "rural" synonymous with "agricultural". It is an essential fact of Swedish industrialisation that much industry in was located in rural areas (Schön, 2014). For this reason, Figure C.4 reports inequality by economic sector rather than by urban–rural status.

The results in Figure C.4 again challenge the Kuznetsian assumptions of egalitarian rural sectors and unequal urban sectors. Granted, the service sector is the most unequal one in the years 1870 to 1910, but inequality within agriculture consistently exceeded inequality within the industrial sector, lending further support to the view that industrialization could promote greater income equality, through the creation of homogenous, high-productive jobs for low-skill workers (cf. Rodrik, 2016).

Figure C.4: Gini coefficient by sector



Note: Builds on the 1/0.5 per cent sample of *bevillning* income taxpayers in Sweden 1870, 1880, 1890, 1900, 1920, 1940, and 1950, as discussed in section 3 of the main body of the paper, and in Appendix A.

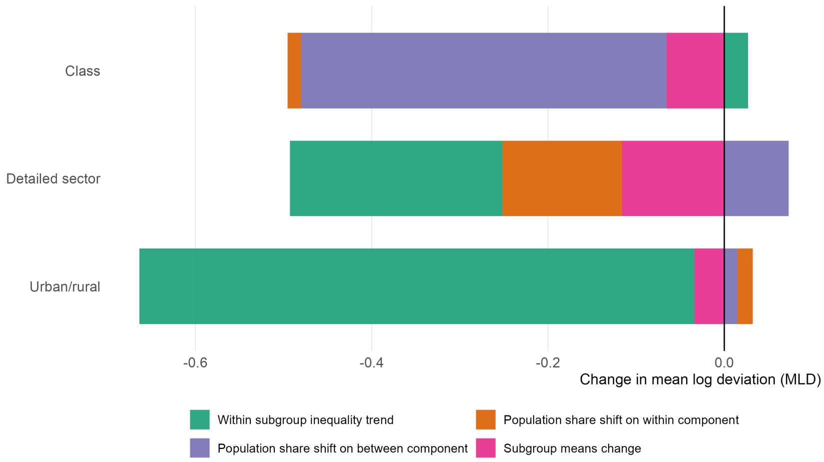
C.5 Decomposition analysis using an alternative periodisation

In Section 6 we have presented a Mokherjee-Shorrocks decomposition of income inequality, measured by the Mean Log Deviation (MLD) for two subperiods: 1870 to 1910, and 1910 to 1950. The 1870–1910 and 1910–1950 division is our preferred subdivision of the study period because of its symmetry (subperiods of equal length) and its overlap with two periods of different patterns of change: the 1870–1920 period of high and mostly stable inequality, and the 1920–1950 period of declining inequality. In this section we present the equivalent analysis using a different periodisation: 1870–1900 and 1920–1950. There are two reasons for this alternative periodisation. The first is that the 1870 to 1910 data exclude dividend income (Online Appendix A.4). As a result, comparisons between 1910 and 1950 understate the decline in inequality, since dividends had become quantitatively important by 1910 and therefore our baseline estimate for 1910 understates the level of inequality then. For that reason, it is important to here also show results considering only the 1920–1950 period, when the data are strictly comparable. The second motivation is to assess the robustness of the results from Figure 8 to modest changes in periodisation.

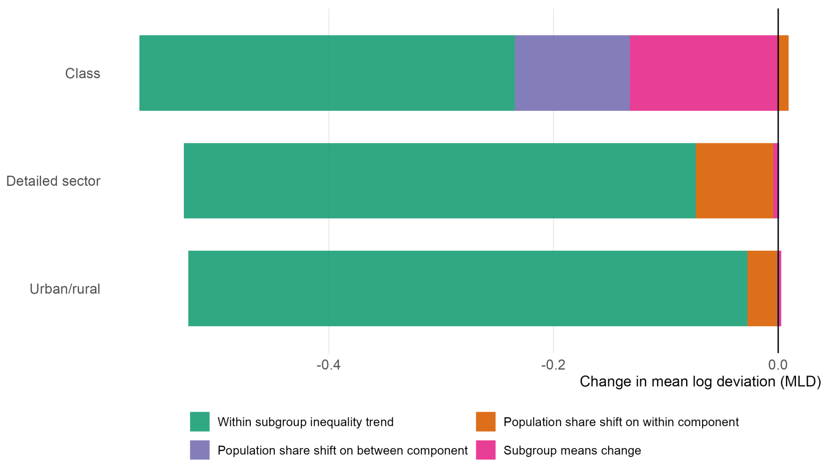
A comparison between the 1870–1900 decomposition (Panel A in Figure C.5) and the 1870–1910 decomposition (Panel A in Figure 8) shows that the modest change in periodisation has virtually no effect on the results. This is unsurprising given the very similar levels of inequality in 1900 and 1910 shown previously (Figures 1 and 2).

The comparison between the 1920–1950 and 1910–1950 decompositions (Panels B of Figures C.5 and 8) shows that the selection of 1910 or 1920 as base year has a substantially larger impact on the decomposition results. This makes sense given the inclusion of dividends in the 1920 data, as well as the turbulent economic changes of the years around 1920. The 1920 to 1950 period saw more substantial equalisation than the 1910 to 1950 period – the x-axis of Figure 8 goes to -0.5 while the x-axis of Figure C.5 only goes to about -0.3 – which is as expected given the higher levels of inequality in 1920 and the inclusion of dividends in the 1920 data but exclusion in the 1910 data. The class decomposition shows that that the decrease in class income differentials

Figure C.5: A Mookherjee-Shorrocks decomposition of inequality, 1870–1900 and 1920–1950



(a) 1870-1900.



(b) 1920-1950.

Note: Data are drawn from the *bevillning* tax dataset for 1870, 1880, 1890, 1900, 1910, 1920, 1940 and 1950, and the Census of 1930. See discussion in Section 3 and Online Appendices A and B. The bars correspond to the contribution of each factor as identified in Equation 15. The underlying sample drops the widow class `st_class` = 13; the urban/rural panel additionally keeps only the employed (`st_marwom` = 1), because rural/urban residency is not known for non-employed wives. The MLD computed on this restricted sample – to which the Mookherjee–Shorrocks decomposition refers – is reported in the "MLD (decomp.)" column of Tables C.1 and C.2, alongside the full-population MLD.

(pink bar) was larger between 1920 and 1950 (> -0.1) than between 1910 and 1950, plausibly related to the extraordinary profits of the late 1910s and early 1920s discussed previously in Roine and Waldenström (2008). The sectoral decomposition shows that within-sector equalisation was more substantial 1920–1950 than 1910–1950, again likely reflecting the great incomes of capital owners and executives around 1920 (as observed in Bengtsson and Molinder, 2025). There was also more decrease 1920–1950 in inequality due to an employment shift to less unequal sectors (orange bar). Along the urban/rural dimension the decrease in the mean log deviation was above -0.4 for the 1920–1950 period but "only" around -0.2 for the 1910–1950 period. The direction is the same but the very high inequality levels of the early 1920s, when profits in the industrial sector led to a stock market boom at the same time as high food prices benefited estate owners who produced for the market, lead to a stronger compression when considering the 1920–1950 period specifically.

C.6 Alternative inequality measures and Lorenz curves

The main body of the paper presents our findings primarily through the Gini coefficient and the top decile's share of income. As Cowell and Flachaire (2024) emphasise, however, every inequality index rests on a particular axiomatic foundation – the Gini coefficient satisfies the transfer principle but not the monotonicity-in-distance principle; the Mean Log Deviation (MLD) satisfies both – and different indices place different weights on different parts of the distribution. (See also Aaberge, 2007 and Cowell, 2011.) Two indices applied to the same data can therefore tell somewhat different stories about the pace and magnitude of distributional change. For example, Jenkins (2024) highlights that the Gini is very sensitive to changes in the middle of the distribution while the Theil is more sensitive to changes at the top, which means that if we ask the question "how did income inequality change in the UK since the 1990s?", our answer will be different depending on which measure we look at. If we only look at the Gini, the answer is: a slight increase, but if we instead look at the Theil, the answer is: a large increase. (Conversely, the MLD is especially sensitive to changes in the bottom of the distribution, so could give yet another picture of the evolution.) That different measures pick up different aspects of inequality and the distribution means that it is important to use several measures for a fuller evaluation. In this section we make our full set of baseline estimates available so that readers can judge whether our conclusions are robust to the choice of measure.

Table C.1 reports all baseline inequality measures for the *All adults* population (i.e., using the broader demographic coverage discussed in Appendix C.1), and Table C.2 reports the same measures for the *Tax filers* population. For each year from 1870 to 1970 we report the Gini coefficient, the MLD, the Theil index, the decile shares D_1–D_{10} (where D_{10} is the top 10 per cent share), the top 1 per cent share, the P90/P50 and P50/P10 percentile ratios, and the median income in nominal SEK. In addition, both tables report a second MLD series in the "MLD (decomp.)" column, computed on the same restricted sample that underlies the Mookherjee–Shorrocks decompositions in Figure 8 where widows (class 13) are excluded. Gini, deciles and other measures that tolerate zero-income observations are computed on the raw income data; MLD and Theil, which involve the logarithm of income, are computed after adding one nominal SEK to every income to avoid log(0) (the "Gini (+1)" column reports the Gini on the same shifted vector for transparency; the shift compresses Gini by at most 0.2 points in the earliest years). P50/P10 returns "–" for the *All adults* samples in which the tenth percentile is zero.

Table C.1 shows that inequality decreased by different intensities depending on which measure we look at. The concentration measures give the most drastic impression: the top 1 percentile's share of incomes shrank by 49.5 per cent from 1870 to 1950 (from 19.2 to 9.7 per cent) while the top decile's share of incomes shrank by 31.6 per cent (from 53.8 to 36.8 perc ent). Thus, the top 1 per cent – the elite of the elite – relatively speaking lost out more than the top 10 per cent. The change in the Gini was smaller, as noted in the main body of the paper: it decreased by 12.7 per cent. With the more top-sensitive Theil measure, the decline was again larger, −38.2 per cent. But with the Mean Log Deviation (the "MLD" column), computed on the full *All adults* population and with the measure's sensitivity to the bottom of the distribution, there was actually no decrease at all in pre-tax inequality from 1870 to 1950 – the level in 1950 (1.870) is in fact slightly above the level in 1870 (1.558). This stands in stark contrast to the decompositions in Section 6, which are computed on the restricted sample reported in the "MLD (decomp.)" column. On that sample, MLD falls from 1.176 in 1870 to 0.552 in 1950 – a decline of about 0.6 log points, matching the magnitude of change displayed on the x-axes of Figure 8. The gap between the two MLD series highlights the measure's great sensitivity to the bottom of the distribution: the groups excluded from the decomposition sample (most importantly non-employed wives, who cannot be allocated to an urban/rural area or to a specific sector) contain many zero or near-zero incomes, and their growing share of the recorded population over time pushes the full-population MLD upward even as inequality among the rest of the distribution compresses. The "MLD (decomp.)" column in Table C.2 shows that the same restriction has a much smaller effect once we already condition on tax filers, since most of the affected groups are already absent from that sample.

Table C.1: Inequality measures, All adults, 1870–1970.

Year	Gini	Gini (+1)	MLD	MLD (decomp.)	Theil	D1	D2	D3	D4	D5	D6	D7	D8	D9	D10	Top 1	P90/P50	P50/P10	Median (SEK)
1870	66.9	66.7	1.558	1.176	1.063	0.0	0.0	2.2	2.9	3.5	6.2	7.0	8.2	16.2	53.8	19.2	6.98	–	124
1880	66.6	66.4	1.631	1.079	0.925	0.0	0.0	1.6	2.7	3.8	5.9	8.2	8.7	16.0	53.1	14.4	7.64	–	158
1890	62.7	62.6	1.491	0.874	0.829	0.0	0.0	1.9	3.6	4.8	7.0	7.9	9.2	17.4	48.2	14.1	3.76	–	288
1900	61.3	61.2	1.430	0.706	0.886	0.0	0.1	1.9	4.3	5.5	6.9	7.9	9.4	16.4	47.5	16.3	3.22	–	340
1910	61.6	61.5	1.476	0.755	0.817	0.0	0.2	1.7	4.2	5.1	6.6	7.2	10.9	18.2	46.0	15.0	4.14	–	431
1920	67.3	67.3	1.690	1.080	0.995	0.0	0.3	1.0	2.1	2.9	4.4	8.0	13.5	20.2	47.8	17.1	7.40	–	550
1930	61.4	61.4	1.647	0.773	0.792	0.0	0.0	1.4	3.4	4.9	6.2	8.7	13.6	19.2	42.7	13.8	4.33	–	600
1940	59.8	59.8	1.597	0.536	0.751	0.0	0.1	1.5	3.2	4.6	7.0	10.4	13.9	18.2	41.0	13.6	3.99	–	1000
1950	58.4	58.4	1.870	0.552	0.657	0.0	0.0	0.8	2.7	5.3	7.7	11.5	15.4	19.6	36.8	9.7	3.31	–	2330
1960	55.5	–	–	–	–	–	–	–	–	–	–	–	–	–	34.2	–	–	–	–
1970	54.3	–	–	–	–	–	–	–	–	–	–	–	–	–	32.8	–	–	–	–

Table C.2: Inequality measures, Tax filers, 1870–1970.

Year	Gini	Gini (+1)	MLD	MLD (decomp.)	Theil	D1	D2	D3	D4	D5	D6	D7	D8	D9	D10	Top 1	P90/P50	P50/P10	Median (SEK)
1870	56.6	56.6	0.551	0.573	0.871	3.0	3.1	3.1	3.3	3.9	4.5	5.7	8.2	13.9	51.1	20.9	5.00	1.25	500
1880	52.3	52.3	0.458	0.468	0.612	3.3	3.4	3.4	4.0	4.2	5.0	6.3	8.6	15.4	46.5	12.4	4.91	1.32	530
1890	48.1	48.0	0.389	0.396	0.564	3.6	3.8	4.1	4.5	5.1	5.8	6.7	8.4	13.7	44.1	14.0	3.57	1.40	700
1900	50.1	50.0	0.430	0.416	0.767	3.3	3.6	4.1	4.6	5.2	5.7	6.6	8.0	12.1	46.9	19.0	3.12	1.60	800
1910	46.5	46.5	0.373	0.365	0.560	2.9	3.8	4.5	5.1	5.8	6.6	7.6	9.1	12.5	42.1	14.8	2.64	1.75	1000
1920	49.7	49.7	0.445	0.419	0.639	1.8	2.8	3.6	4.5	5.7	7.1	9.1	11.2	14.2	40.0	16.2	2.71	2.50	2000
1930	47.6	47.5	0.402	0.385	0.530	1.8	3.4	3.9	4.7	5.7	7.0	9.0	11.6	15.3	37.6	13.5	2.82	2.20	1100
1940	44.1	44.1	0.342	0.307	0.456	2.2	3.3	4.2	5.3	6.5	7.8	9.6	11.6	14.4	35.1	12.3	2.34	2.54	2030
1950	40.5	40.5	0.329	0.249	0.321	1.5	2.9	4.4	6.0	7.7	9.4	11.0	12.6	15.1	29.4	7.9	2.01	3.77	4530
1960	41.0	–	–	–	–	–	–	–	–	–	–	–	–	–	28.3	6.5	–	–	–
1970	39.5	–	–	–	–	–	–	–	–	–	–	–	–	–	27.1	5.7	–	–	–

It is a clear result from our research that in terms of the pre-tax, pre-transfers income distribution, there was much more equalisation in Sweden between 1870 and 1950 when comparing top groups and middle groups, than if we compare middle groups and bottom groups. This makes it especially pertinent to analyse the composition of the bottom groups: who were these groups who had an income growth at pace with the general population, but not more? Table C.3 examines the composition of the bottom decile (D_1) more closely. Since more than 10 per cent of the *All adults* population reports zero income in every year, D_1 consists entirely of zero-income individuals. The composition of this group shifts markedly over time. In 1870 and 1880, D_1 is dominated by the poor, paupers, and vagrants (72 and 67 per cent, respectively), with a substantial share of former and retired individuals (28 and 23 per cent). The gender split is roughly even (56 and 55 per cent women). From 1890 onwards, however, D_1 becomes increasingly dominated by individuals recorded under familial titles – wives, widows, sons, and daughters – who rise from 67 per cent in 1890 to 100 per cent in 1940 and 1950. By 1940, the bottom decile consists entirely of married women (housewives) classified as outside the labour force (see also discussion in Appendix C.1). This shift reflects both improved classification of married women's labour-force status in the sources and the gradual disappearance of extreme poverty categories from the tax records.

Table C.3: Composition of the bottom decile (D_1), All adults, 1870–1950

Year	Poor, paupers, vagrants (%)	Former, retired (%)	Familial titles (wives etc.) (%)	Women (%)	Housewives (%)
1870	72	28	0	56	0
1880	67	23	10	55	0
1890	25	8	67	84	60
1900	7	13	80	89	71
1910	7	15	78	87	73
1920	0	15	76	86	75
1930	0	17	83	89	79
1940	0	0	100	100	100
1950	0	0	100	100	100

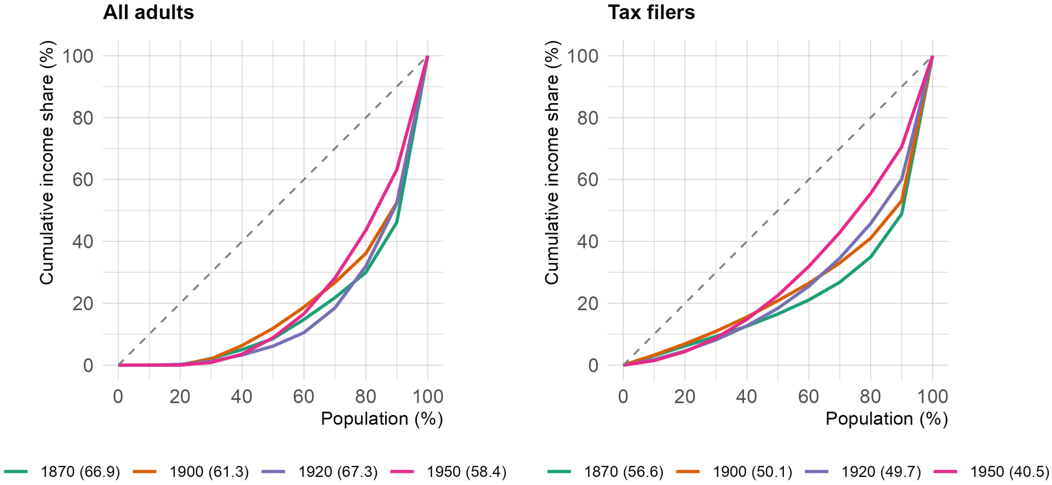
Note: The table reports the weighted share of the bottom decile (D_1) of the *All adults* population belonging to selected categories. "Poor, paupers, vagrants" corresponds to occupational class 11; "Former, retired" to class 10; "Familial titles (wives etc.)" to class 13. "Women" is the share with $sex = 1$. "Housewives" is the share with $st\_marwom = 3$, i.e. married women classified as outside the labour force. Since more than 10 per cent of the population reports zero income in every year, D_1 consists entirely of zero-income individuals.

That the bottom groups did not improve their share of the income over the 1930 to 1950 period, as the welfare state developed and Sweden became an industrial welfare society, is somewhat surprising. However, in the 1950s and 1960s the first detailed surveys of incomes in Sweden were produced (and the large datasets for 1960 and 1970 that we use here), and the contemporary commentary on what these surveys revealed, as well as later research within economics and sociology, does give perspective on this result. Björklund and Palme (2000, Table 4) show that in surveys from 1951, 1956, 1967 and 1973, a quite large share reported zero incomes: more than half of women in 1951 and 1956, and still 39 per cent in 1967. Even on the family level 8.4 per cent in 1951 and 8.8 per cent in 1967 reported zero income. Similarly, Jansson (1990, pp. 63–64) shows that in 1967 there was still a comparatively large share of households on markedly low living standards in relative terms, but that to 1974 and 1981, this group shrank drastically. This was related to the fact that low-income earners became a prioritised group on the political agenda in the late 1960s and 1970s, motivated precisely by the surprise and disappointment of welfare state reformers when they found lingering poverty in the social studies and surveys of the 1960s (Smedberg, 2018).

Thus, our relatively pessimistic view of postwar equalisation in Sweden should be seen in the light of three facts. Firstly, that our unit of measurement is the individual. Table C.3 brings out very clearly that the bottom decile of the income distribution over time to a larger and larger proportion (three quarters in 1920, 100 per cent in 1940 and 1950) was made up of housewives. Had the population of our inequality measures been tax units or households, these housewives would have been counted with their spouses who had incomes, and inequality would have seen a steeper decline. (See also Figure C.1.) Secondly, that postwar equalisation reached a new phase just as our period of investigation came to an end: in the late 1960s and 1970s, the incomes of the very bottom groups began growing much faster (Jansson, 1990). Thirdly, that redistribution via taxes and transfers grew in importance in the 1950s, 1960s and 1970s and that had we measured incomes as post-tax, post-transfer – which is at most available from 1968 onwards – equalisation would also look steeper. In the most recent study (Hammar et al., 2022, Figure 6), the Gini coefficient of incomes on the household level in the 1970s looks 6–8 per cent lower after taxes and transfers than before, and redistribution was sparse before the 1940s (Torregrosa Hetland and Sabaté, 2026). This means that the 1940 to 1970 decrease in inequality would have been steeper with a post-tax-and-transfer measure, and that the bottom groups would see a more beneficial development in this period.

Figure C.6 complements the tables with Lorenz curves for four benchmark years – 1870, 1900, 1920, and 1950 – for both populations. The curves make visually apparent the gradual compression of the distribution in the All-adults panel between 1870 and 1950, and the substantially larger compression in the Tax-filers panel over the same period.

Figure C.6: Lorenz curves for selected years, 1870–1950



Note: Weighted Lorenz curves for the *All adults* population (left panel: $st\_marwom \in \{1, 2, 3\}$) and the *Tax filers* population (right panel: $inddata = 1$). Gini coefficients are shown in parentheses next to each year in the legends. Based on the same micro data as underlying Figures 1 and 2.

References

Aaberge, R. (2007). Gini's Nuclear Family. *Journal of Economic Inequality*, 5:305–322.

Andersson, L. I. (1995). *Kommunalskattefrågan under mellankrigstiden: Skatterättvisa i och mellan kommuner*. Doctoral dissertation, Lund University.

Andersson, P. F. (2021). Fiscal Capacity in Non-democratic states: The Origins and Expansion of Income Tax. WIDER Working Paper 2021/176, United Nations University World Institute for Development Economics Research.

Atkinson, A. B. (2007). Measuring Top Incomes: Methodological Issues. In Atkinson, A. B. and Piketty, T., editors, *Top Incomes over the Twentieth Century: A Contrast Between European and English-Speaking Countries*, pages 18–42. Oxford University Press.

Auten, G. and Splinter, D. (2024). Income Inequality in the United States: Using Tax Data to Measure Long-Term Trends. *Journal of Political Economy*, 132(7):2179–2227.

Bagge, G., Lundberg, E., and Svennilson, I. (1933). *Wages in Sweden 1860–1930. Part One: Manufacturing and Mining*. P.S. King, London.

Bengtsson, E. (2022). Agrar ojämlikhet i Sverige studerad med bevillningstaxeringen som källa, 1870–1920. Lund Papers in Economic History No. 242.

Bengtsson, E. and Kersting, F. (2026). The Social Origins of Democracy and Authoritarianism Reconsidered: Prussia and Sweden in Comparison. *Comparative Political Studies*, 59(5):979–1009.

Bengtsson, E. and Molinder, J. (2025). What Happened to the Incomes of the Rich during the Great Levelling? Evidence from Swedish Individual-level Data, 1909–1950. *Journal of Economic History*, 85(2):411–441.

Bengtsson, E. and Svensson, P. (2019). The Wealth of the Swedish Peasant Farmer Class, 1750–1900: Composition and Distribution. *Rural History*, 30(2):129–145.

Bengtsson, E. and Svensson, P. (2022). The Living Standards of the Labouring Classes in Sweden, 1750–1900: Evidence from Rural Probate Inventories. *Agricultural History Review*, 70(1):49–69.

Björklund, A. and Palme, M. (2000). The Evolution of Income Inequality During the Rise ofthe Swedish Welfare State 1951 to 1973. *Nordic Journal of Political Economy*, 26:115–128.

Bohlin, J. and Larsson, S. (2007). The Swedish Wage-rental Ratio and its Determinants, 1877–1926. *Australian Economic History Review*, 47.

Bohlin, J. and Prado, S. (2011). A Market with Strings Attached: Agricultural Property Price Movements. In Antonsson, H. and Jansson, U., editors, *Agriculture and Forestry in Sweden since 1900: Geographical and Historical Studies*, page 123–138. The Royal Swedish Academy of Agriculture and Forestry, Stockholm.

Broberg, O. (2006). *Konsten att skapa pengar: Aktiebolagens genombrott och finansiell modernisering kring sekelskiftet 1900*. Doctoral dissertation, University of Gothenburg.

Buchmann, N., Dupas, P., and Ziparo, R. (2025). The Good Wife? Reputation Dynamics and Financial Decision-Making Inside the Household. *American Economic Review*, 115(2):525–570.

Bäcklund, D. (1988). *I industrisamhällets utkant: Småbrukets omvandling i Lappmarken 1870-1970.* Doctoral dissertation, Umeå University.

Carlsson, S. (1968). *Yrken och samhällsgrupper: den sociala omgrupperingen i Sverige efter 1866.* Almqvist och Wiksell, Stockholm.

Cassel, G. (1906). Utvecklingslinjer i svensk skattelagstiftning. *Ekonomisk Tidskrift*, 8:191–216.

Chiappori, P.-A. (1993). Collective Labour Supply and Welfare. *Journal of Political Economy*, 100(3):437–467.

Chiswick, B. R. and Robinson, R. H. (2021). Women at Work in the United States since 1860: An Analysis of Unreported Family Workers. *Explorations in Economic History*, 82:101406.

Cowell, F. A. (2011). *Measuring Inequality.* Oxford University Press, Oxford, third edition.

Cowell, F. A. and Flachaire, E. (2024). Inequality Measurement and the Rich: Why Inequality Increased More than We Thought. *Review of Income and Wealth*, 70(2):254–277.

Craig, L. A. (1991). The Value of Household Labor in Antebellum Northern Agriculture. *Journal of Economic History*, 51(1):67–81.

Edvinsson, R. (2014). The Gross Domestic Product of Sweden within Present Borders, 1620–2012. In Edvinsson, R., Jacobson, T., and Waldenström, D., editors, *Historical Monetary and Financial Statistics for Sweden, Volume II: House Prices, Stock Returns, National Accounts, and the Riksbank Balance Sheet, 1620–2012*, pages 101–182. Sveriges Riksbank and Ekerlids, Stockholm.

Esberger, S. E. and Malmquist, S. (1972). *En statistisk studie av inkomstutvecklingen.* Monographs in connection with the 1960 census of population and housing in Sweden. Allmänna förlaget, Stockholm.

Fahlbeck, P. (1893). *Det svenska jordbrukets afkastning.* Gleerupska Universitets-Bokhandeln, Lund.

Fields, G. S. (2004). Dualism in the Labor Market: A Perspective on the Lewis Model After Half a Century. *The Manchester School*, 72(6).

Flodström, I. (1902). *Öfversikt af 1900 års bevillningstaxering.* P.A. Norstedt och söner, Stockholm.

Flodström, I. (1906). Skatteväsendets omreglering. *Ekonomisk Tidskrift*, 8:169–171.

Flodström, I. (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.* Number series F, volume IV in Finansstatistiska Utredningar utgifna genom Kungl. Finansdepartementet. Beckmans Boktryckeri, Stockholm.

Flygare, I. and Isacson, M. (2003). *Det svenska jordbrukets historia, V: Jordbruket i välfärdssamhället, 1945-2000.* LTs förlag, Stockholm.

Folbre, N. and Wagman, B. (1993). Counting Housework: New Estimates of Real Product in the United States, 1800-1860. *Journal of Economic History*, 53(2):275–288.

Fremdling, R. (1988). German National Accounts for the 19th and Early 20th Century: A Critical Assessment. *VSWG: Vierteljahrschrift für Sozial- und Wirtschaftsgeschichte*, 75(3):339–357.

Gadd, C.-J. (2000). *Det svenska jordbrukets historia, band III: Den agrara revolutionen 1700–1870*. Natur och Kultur/LTs förlag, Stockholm.

Geloso, V. J., Magness, P., Moore, J., and Schlosser, P. (2022). How Pronounced Is the U-curve? Revisiting Income Inequality in the United States, 1917–60. *Economic Journal*, 132(647):2366–2391.

Gollin, D. and Kaboski, J. P. (2023). New Views of Structural Transformation: Insights from Recent Literature. *Oxford Development Studies*, 51(4):339–361.

Grönlund, O. (1942). Den senaste allmänna jordbruksräkningen i Sverige och därmed sammanhängande undersökningar. *Statsvetenskaplig Tidskrift*, 45(2-3):159–176.

Gustafsson, B. and Johansson, M. (2003). Steps Toward Equality: How and Why Income Inequality in Urban Sweden Changed during the Period 1925–1958. *European Review of Economic History*, 7(2):191–221.

Gómez León, M. and Gabbuti, G. (2025). Wars, Depression, and Fascism: Income Inequality in Italy, 1901–1950. *Explorations in Economic History*, 98:101715.

Hamark, J., Collin, K., and Prado, S. (2025). A New Benchmark for Industrial Wages in Mid-1880s Sweden. *Essays in Economic and Business History*, 42(1):116–151.

Hammar, O., Roth, P., and Waldenström, D. (2022). The Swedish Income Distribution, 1968–2016. Unpublished paper, dated 7 February 2022. Available from https://sites.google.com/site/ollehammar85/research.

Hedlund-Nyström, T. (1972). The Finance Act of 1862 and Its Effect on the Economy. *Economy and History*, 15(1):58–82.

Hjelm, L. (1995). Ludvig Nanneson – föregångsman inom svensk lantbruksekonomi. In Engwall, L., editor, *Föregångare inom företagsekonomin*, pages 125–138. SNS, Stockholm.

Hultgren, E. and Landergren, E. (1891). *Undersuchung über die Ernährung Schwedischer Arbeiter bei frei gewählter Kost*. Samson och Wallin, Stockholm.

Isacson, M. (1994). *Vardagens ekonomi: arbete och försörjning i en mellansvensk kommun under 1900-talet*. Gidlunds, Hedemora.

Jansson, K. (1990). *Inkomst- och förmögenhetsfördelningen 1967–1987. Bilaga 19 till LU90*. Allmänna förlaget, Stockholm.

Jenkins, S. P. (2024). Getting the Measure of Inequality. *Oxford Open Economics*, 3(S1):i156–i166.

Johansson, L.-O. (1996). *Levebrödet: den informella ekonomin i 1930-talets Dalarna*. Doctoral dissertation, Uppsala University.

Kommerskollegium (1911). *Till belysning af landtarbetarnas arbets- och löneförhållanden i Sverige år 1910*. Kommerskollegium, Stockholm.

Kuuse, J. (1970). *Från redskap till maskiner: Mekaniseringsspridning och kommersialisering inom svenskt jordbruk 1860–1910*. Department of Economic History, Gothenburg.

Larrimore, J. (2014). Accounting for United States Household Income Inequality Trends: The Changing Importance of Household Structure and Male and Female Labor Earnings Inequality. *Review of Income and Wealth*, 60(4):683–701.

Larrimore, J., Mortenson, J., and Splinter, D. (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.

Larsson, M. (1986). *Arbete och lön vid Bredsjö bruk: En studie av löneprinciper och lönenivåer för olika yrkeskategorier vid Bredsjö bruk 1828–1905*. Doctoral dissertation, Uppsala University.

Leffler, J. (1897). *Zur Kenntnis von den Lebens- und Lohnverhältnissen industrieller Arbeiterinnen in Stockholm*. Koersners, Stockholm.

Lewin, L. (2017). *Ideologi och strategi: Svensk politik under 130 år*. Carlssons, Stockholm, sixth edition edition.

Lindahl, E., Dahlberg, E., and Kock, K. (1937). *The National Income in Sweden 1861–1930. Part One*. WS King & Son, London.

Lodin, S.-O., Lindencrona, G., Melz, P., Silfverberg, C., and Simon-Alvendal, T. (2017). *Inkomstskatt: en läro- och handbok i skatterätt*. Studentlitteratur, Lund, 16th edition.

Lundh, C. (2008). Statarnas löner och levnadsstandard. In Lundh, C. and Olsson, M., editors, *Statarliv - i myt och verklighet*, page 113–165. Gidlunds, Hedemora.

Lundh, C. (2010). *Spelets regler: Institutioner och lönebildning på den svenska arbetsmarknaden 1850–2000*. SNS, Stockholm, 2 edition.

Lundh, C. (2012). Wage Forms, Costs of Living and the Urban–rural Wage Gap: Southern Sweden, 1881–1930. *Scandinavian Economic History Review*, 60(2):123–145.

Lundh, C. and Olsson, M. (2011). Contract-Workers in Swedish Agriculture, c. 1890s–1930s: A Comparative Study of Standard of Living and Social Status. *Scandinavian Journal of History*, 36(3):298–323.

Lundh, C. and Prado, S. (2015). Markets and Politics: The Swedish Urban-Rural Wage Gap, 1865–1985. *European Review of Economic History*, 19(1):67–87.

Lundh, C. and Prado, S. (2020). The Historical Labour Database (HILD). University of Gothenburg, Department of Economy and Society. https://www.gu.se/en/school-business-economics-law/economy-society/the-historical-labour-database-hild.

Milanovic, B. (2016). *Global Inequality*. Harvard University Press, Cambridge, MA.

Milanovic, B. (2024). How Rich Were the Rich? *Explorations in Economic History*, 93:101592.

Morell, M. (2001). *Jordbruket i industrisamhället 1870–1945*. Natur och Kultur / LTs förlag, Stockholm.

Möller, B. (1962). *Svensk aktiemarknad*. Gothenburg university, Gothenburg.

Nanneson, L. (1916). *Räkenskapsresultat från svenska jordbruk i bokföringsåret 1914–1915*. Meddelanden från Kungl. Lantbruksstyrelsen N:r 201. AB Skånska Centraltryckeriet, Lund.

Nanneson, L. (1923). *Räkenskapsresultat från svenska jordbruk bokföringsåret 1920–1921*. Meddelanden från Kungl. Lantbruksstyrelsen N:r 246. Röhrs boktryckeri, Malmö.

Nanneson, L. (1946). Rationalitetsvariationerna inom det svenska jordbruket. Technical report, Statistiska centralbyrån (SCB), Stockholm. Utredning på uppdrag av 1942 års jordbrukskommitté.

Nordlund Edvinsson, T. and Söderberg, J. (2010). Servants and Bourgeois Life in Urban Sweden in the Early 20th Century. *Scandinavian Journal of History*, 34(4):427–450.

Norlander, K. (2000). *Människor kring ett företag. Kön, klass och ekonomiska resurser. Liljeholmens Stearinfabriks AB 1872–1939.* Doctoral dissertation, University of Gothenburg.

Odhner, C.-E. (1953). *Jordbruket vid full sysselsättning.* KF:s bokförlag, Stockholm.

Palmgren, H. (1915). *Självdeklaration: Praktisk handbok för skattskyldiga och beskattningsmyndigheter.* P.A. Norstedt och söner, Stockholm, 9th edition.

Piketty, T. (2003). Income Inequality in France, 1901—1998. *Journal of Political Economy*, 111(5):1004–1042.

Piketty, T. (2014). *Capital in the Twenty-First Century.* Harvard University Press.

Piketty, T. and Saez, E. (2003). Income Inequality in the United States, 1913–1998. *Quarterly Journal of Economics*, 118(1):1–39.

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

Prado, S. (2010). Nominal and Real Wages of Manufacturing Workers, 1860–2007. In Edvinsson, R., Jacobson, T., and Waldenström, D., editors, *Historical Monetary and Financial Statistics for Sweden: Exchange rates, Prices and Wages, 1277–2008*, pages 479–527. Ekerlids, Stockholm.

Ranaldi, M. and Milanovic, B. (2022). Capitalist Systems and Income Inequality. *Journal of Comparative Economics*, 50(1):20–32.

Riksdagen (1860). Bihang till samtlige riks-ståndens protokoll vid lagtima riksdagen i Stockholm åren 1859 och 1860. Tionde samlingen. 1sta afdelningen: Rikets ständers bevillnings-stadga och riks-stat samt riksdags-beslutet. Available from the Royal Library website of historical Riksdag materials: https://weburn.kb.se/riks/metadata/63/21798563.html. Decision by the Swedish Riksdag 28 October 1860.

Riksdagen (1902). Riksdagens skrifvelse till Konungen, i anledning af Kongl. Maj:ts proposition angående en inkomstskatt för år 1903. Available from the Royal Library website of historical Riksdag materials: https://weburn.kb.se/riks/metadata/65/19566765.html. Decision by the Swedish Riksdag 14 May 1902, Riksdagens Skrifvelse N:o 197.

Riksdagen (1910). Riksdagens skrifvelse till Konungen, i anledning af dels Kungl. Maj:ts proposition med förslag till förordning om inkomst- och förmögenhetsskatt, dels ock i ämnet väckta motioner. Available from the Royal Library website of historical Riksdag materials: https://weburn.kb.se/riks/metadata/37/19567137.html. Decision by the Swedish Riksdag 9 June 1910, Riksdagens Skrifvelse N:o 106.

Riksdagen (1919). Riksdagens skrivelse till Konungen, i anledning av dels Kungl. Maj:ts proposition med förslag om vissa, för ernående av jämkningar i fördelningen av den kommunala skattebördan erforderliga ändringar i förordningen den 28 oktober 1910 angående bevillning av fast egendom samt av inkomst, dels ock en i ämnet väckt motion. Available from the Royal Library website of historical Riksdag materials: https://weburn.kb.se/riks/metadata/84/19567784.html. Decision by the Swedish Riksdag 18 June 1919, Riksdagens Skrifvelse N:o 373.

Riksdagen (1928). Riksdagens skrivelse till Konungen i anledning av Kungl. Maj.ts proposition med förslag till kommunalskattelag m. m. jämte i ärendet väckta motioner. Available from the Royal Library website of historical Riksdag materials: https://weburn.kb.se/riks/metadata/81/19568281.html. Decision by the Swedish Riksdag 1 June 1928, Riksdagens Skrifvelse N:o 344.

Rodrik, D. (2016). Premature Deindustrialization. *Journal of Economic Growth*, 21:1–33.

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

Roine, J. and Waldenström, D. (2010). Top Incomes in Sweden over the Twentieth Century. In Atkinson, A. B. and Piketty, T., editors, *Top Incomes: A Global Perspective*, pages 299–370. Oxford University Press, Oxford.

SCB (1906). *Befolkningsstatistik*. Bidrag till Sveriges officiella statistik (BiSOS) A. Statistiska Centralbyrån, Stockholm. Digitalized by SCB and available at https://www.scb.se/hitta-statistik/aldre-statistik/innehall/bidrag-till-sveriges-officiella-statistik/Befolkning-1851-1910/.

SCB (1955). *Det mindre jordbrukets möjligheter att uppnå bättre lönsamhet*. Betänkande avgivet av småbruksutredningen. Statistiska Centralbyrån, Stockholm.

SCB (1959). *Historisk statistik för Sverige II: Väderlek, lantmäteri, jordbruk, skogsbruk, fiske t.o.m. år 1955*. Statistiska Centralbyrån, Stockholm. available online from the SCB website, https://www.scb.se/hitta-statistik/aldre-statistik/innehall/serien-historisk-statistik-for-sverige/.

Schön, L. (2014). *En modern svensk ekonomisk historia*. Studentlitteratur, Lund, 4th edition.

Schön, L. and Krantz, O. (2015). Swedish Historical National Accounts 1560–2010. Available from: https://www.lusem.lu.se/organisation/department-economic-history/research-department-economic-history/databases-department-economic-history/economic-history-data/swedish-historical-national-accounts-1300-2022. Dataset created by Lennart Schön and Olle Krantz, maintained and updated by the Department of Economic History, Lund University.

Smedberg, C.-F. (2018). Låginkomsttagarna: Expertis, politik och mediering i formandet av en ny kategori omkring 1968. *Scandia*, 84(1):61–85.

Socialstyrelsen (1923). *Levnadskostnaderna på landsbygden i Sverige vid år 1920*. Sveriges officiella statistik: Socialstatistik. Isaac Marcus Boktryckeri-Aktiebolag, Stockholm. Digitalized by the Gothenburg University Library and available at https://gupea.ub.gu.se/handle/2077/43429.

SOU (1933). *Skatteutjämningsberedningen III: Principbetänkande*. Number 1933:4 in Statens Offentliga Utredningar. Department of Finance, Stockholm.

SOU (1938). *Betänkande angående gift kvinnas förvärvsarbete m.m. angivet av Kvinnoarbetskommittén*. Number 1938:47 in Statens Offentliga Utredningar. Department of Finance, Stockholm.

SOU (1946). *Riktlinjer för den framtida jordbrukspolitiken del II. Betänkande avgivet av 1942 års jordbrukskommitté*. Number 1946:46 in Statens Offentliga Utredningar. Ministry of Agriculture, Stockholm.

SSSK (1924). *Statistisk årsbok för Stockholms stad 1924*. Stockholms stads statistiska kontor, Stockholm. available online from the Stockholm City Archive's website, https://sok.stadsarkivet.stockholm.se/?template=view_post&id=507.

Stanfors, M., Leunig, T., Eriksson, B., and Karlsson, T. (2014). Gender, Productivity, and the Nature of Work and Pay: Evidence from the Late Nineteenth-century Tobacco Industry. *Economic History Review*, 67(1):48–65.

Torregrosa Hetland, S. and Sabaté, O. (2026). Income Taxes and Redistribution in the Early Twentieth Century. *Economic History Review*, 79(2):634–660.

Uppenberg, C. and Nilsson, M. (2025). The Crofter Is a Woman: Gender Division of Labour in Rural Semi-landless Households, Sweden 1850–1900. *Agricultural History Review*, 73(1):72–95.

von Feilitzen, U. (1890). *Tjenare, statare och torpare: En del statistiskt material till jordarbetarfrågans lösning*. Samson och Wallin, Stockholm.

Widell, L. (1907). Några ord om en svensk finans- och inkomststatistik. *Ekonomisk Tidskrift*, 9:323–338.

Åmark, K. (1923). *Undersökning angående jordegendomsvärdenas utveckling i Sverige och vissa främmande länder*. Number 1923:45 in Statens Offentliga Utredningar. Ministry of Finance, Stockholm.