Top Incomes in Sweden over the Twentieth Century

Jesper Roine* Daniel Waldenström**

SSE/EFI Working Paper Series in Economics and Finance No. 602

2005-08-15

Preliminary version, please do not quote without checking with the authors.

Comments are most welcome.

Abstract

This paper presents homogenous series of top income shares in Sweden from 1903 to 2003 using individual tax returns data. We find that Swedish top incomes have developed more similarly to the US, Canada and the UK than to other continental European countries when capital gains are included. The top income shares are U-shaped over time, falling steadily until around 1980 when they start increasing again. Around 2000 they reach levels similar to those found around 1950, before the expansion of the Swedish welfare state. However, unlike the Anglo-Saxon countries, where the recent increases were mainly driven by increased wage earnings inequality, Swedish top income shares have risen almost exclusively due to capital gains, a finding consistent with relatively high marginal wage taxes and internationally high price increases in financial and real estate markets since 1980. When excluding capital gains the increase in top income shares since 1980 almost disappears and the Swedish experience looks more like that of continental Europe. Furthermore, we also find that the largest decrease of top income shares happens between 1935 and the beginning of the 1950s, but not (as in the US and in France) during the war years, but before 1939 and after 1945 suggesting that the Swedish development was more driven by policy than by exogenous shocks.

* Department of Economics, Stockholm School of Economics, P.O. Box 6501, SE-11383 Stockholm, Ph: +46-8-7369000, e-mail: Jesper.Roine@hhs.se

** Department of Economics, Stockholm School of Economics, P.O. Box 6501, SE-11383 Stockholm, Ph: +46-8-7369000, e-mail: Daniel.Waldenstrom@hhs.se

Table of Contents:

1 Introduction

Over the past couple of years there has been an increased interest in income inequality at the top of the distribution, both academically and in the policy debate. The well documented long-run trends, with clear increases in income inequality in the US (Piketty and Saez, 2003) and the UK (Atkinson 2002) over the past decades, has made some conclude that policy has shifted so as to tolerate widening gaps in society. Others see signs of "technological change" and a new "Industrial revolution" which, in line with Kuznets' hypothesis, initially leads to increased inequality.

In the Swedish context recent corporate scandals and increased income gaps between "the elite" and "the average worker" have also caused debates about top income earners and to what extent the Swedish welfare state is being dismantled. These discussions of inequality should be viewed in light of the fact that Sweden, at least until the early 1990s, had one of the most equal distributions of income among industrialized countries (see e.g. Atkinson, Rainwater and Smeeding (1995) and Gottshalk and Smeeding (1997)). However, as pointed out by Björklund and Palme (2000), "when and how Sweden achieved its equal distribution of income is more of an open question. Is it a rather recent outcome of the growth of the welfare state during the 1960s and 1970s? Or is it a historical inheritance from rather a long time ago?"(p. 115). In this paper we try to contribute to answering this type of question, focusing on the top of the income distribution. What has been the development of top income shares in Sweden over the twentieth century? What were the effects of the World Wars and the depression in the 1930s? What was the effect of the expansion of the welfare state in the 1960s and 1970s? Has there been an increase in top income shares over the past decades as in the Anglo-Saxon countries, or is the development in Sweden closer to that found in France and the Netherlands, where top income shares have been rather stable in the recent past? And what does this recent development look like in relation to the rest of the Twentieth Century?

To be able to shed light on questions like these we need *long-run, homogenous*, series of income inequality, which ideally should be *comparable across countries*. In this paper we construct such series of pre tax income for Sweden over the one hundred years between 1903 and 2003. Using data from tax returns combined with National Accounts statistics and population statistics we construct series for the share of "total income" going to the top decile, the top five percent, the top percent, etc., down to the top 0,05 percent of the adult (tax) population.

As far back as possible we also calculate shares of different sources of income (wages, capital income, business income and capital gains income) for each fractile.

There are three main contributions in the paper. First, we construct, for the first time a *long-run series of income inequality in Sweden* spanning the whole of the Twentieth century. In the absence of micro data few have ventured further back than to 1967. A recent exception is a study by Björklund and Palme (2000) who focus on the period 1951–1973. Before them, e.g., Spånt (1976, 1979) reports on the period 1920-1976, Spånt and Selander (1969) 1951–1966, Lydall (1968) 1920-1960, Bentzel (1953) 1930–1948, Lindstrand (1949) 1935–1947, Quensel (1944), 1930–1941 and Hagstroem (1944, 1949) covers the period 1912-1947. Bentzel (1953) stands out as a particularly interesting reference since his method is very similar to the one we use.[1] In addition, Gustafsson and Johansson (2003) have carried out a study of the income distribution between 1925 and 1958. Their study is, however, limited to the City of Göteborg. To the best of our knowledge no one has previously constructed longer series than the ones mentioned above and no one has studied income inequality in for the very first years of the Twentieth Century. Furthermore, none of the above focus on the very top of the income distribution as we do in this paper.

Second, our series are constructed so as to be *homogenous* over the whole period. Previous studies all use slightly different methods and concepts, and are not immediately comparable. Throughout this paper, we focus on building homogenous series, and to the extent that this means abstracting from some detail which may be of importance at one point in time but for which we have no information at other points in time, we chose to have comparability over time. Our series are hence, best suited for studying the long-run trends in the data (even though, as we shall see, our series are also in line with previous studies based on more detailed information).

Third, we focus on the development of the *top income earners* including data for fractions of the top percent. We know of only one previous study on Swedish data with this focus by Magnusson (2004) which is confined to the period 1985-2002. Most previous studies on Swedish data have focused on welfare and especially on disposable incomes amongst those in

[1] To the extent that Piketty and Saez (2001, 2003) extend and generalize the work by Kuznets on US data in the early 1950s, we extend and generalize the work by Bentzel (1953) on Swedish data for top incomes, both by going further back in time, as well as constructing series for the 50 years that has past since Bentzel's publication.

the lower parts of the distribution. Some would even say that it does not matter what happens at the top, and certainly not the very top of the distribution, since this (at least in a Swedish context) concerns such a small number of individuals. We do not think this is the case. There are a number of reasons for why it is interesting to know more about the development of top incomes. First of all, the development of the top incomes gives important indications of the structure of long-term economic development, as suggested by Kuznets (1953). Second, high income individuals typically control large assets, and make decisions which potentially affect the economy at large. What happens in this group is therefore of interest for the whole economy. Third, those with the highest incomes also constitute a potentially important source of tax revenue. Last, but not least, not going beyond the income shares of the top decile or at best the top five percent, one runs the risk of completely missing what is really driving changes in the top of the distribution. As we shall see, much of the development of the top income shares in Sweden over the past one hundred years actually stem from what has happened in the top percent of the distribution. As pointed out by Atkinson (2004) changes at the very top can also have a substantial impact on the Gini coefficient and, hence, not carefully studying the development of the top may lead to incorrect inferences about what is driving changes in inequality.

Internationally the use of income tax data for the construction of long time series of income distribution is far from new. Besides Kuznets' classic work on the *"Shares of Upper Income Groups in Income and Savings"* (1953), Champernowne (1973) has studied data based on "super-tax" records which was introduced in 1908 in the UK, Hartog and Veenbergen (1978) construct income distribution estimates for the period 1914-1972 using income statistics data for the Netherlands, Sörensen (1993) has made use of Danish income statistics data for the period 1870-1986. Morrison (2000) provides an overview of historical data for Europe. Notwithstanding the many previous studies, it is fair to say that the approach of Kuznets has only recently been fully revived, starting with the work of Piketty (2000, 2003) on France, Piketty and Saez (2001, 2003) on the US, and Atkinson (2004) on the UK. In the past couple of years a number of studies have used very similar methods to construct series for a number of countries, e.g. Atkinson and Salverda (2005) who compare top income shares in the Netherlands and the UK, Atkinson and Leigh (2004) who compare top income shares in a number of Anglo-Saxon countries, Banerjee and Piketty (2004) who study top income shares in India over the Twentieth Century, Dell (2003, 2005) who constructs series for Germany and Switzerland, and Saez and Veall (2003) who construct series for Canada. In this paper we do the same for

Sweden, using similar methods as those used in the recent papers mentioned above, so as to allow comparisons between our series and those for other countries.

In the following section we briefly describe our methodology and the data used. All details concerning adjustments, alternative specifications and sources have been placed in a number of appendices. In Section 3, we present our main results, which we then compare to the findings from other countries in Section 4. Section 5 concludes.

2 Methodology and data

This section briefly describes our method and the data used. Details concerning the data, the sources, and the adjustments made are placed in a number of appendices. For some of the most important concepts, however, Table 1 gives a brief overview.

Table 1: Overview of income data used, adjustments and reference totals[a].

<table>
<tr><th>Income years</th><th>Main income concept in tax statistics [Swedish term]</th><th>Adjustments to get comparable income concepts</th><th>Reference total income</th><th>Reference total population</th></tr>
<tr><td>1903–1909</td><td>Total income [Taxerad inkomst]</td><td>No adjustments</td><td rowspan="2">Share of personal sector total income (in National accounts) based on estimates of items not included in preferred def. (1903–1942)</td><td rowspan="2">Adult population (16 yrs and above) minus married women (–1951)</td></tr>
<tr><td>1910–1942</td><td>Total income plus a share (–1937: 1/60; 1938–1942: 1/100) of taxable wealth [Taxerat belopp]</td><td>Removal of wealth shares using data on wealth distribution per income class</td></tr>
<tr><td>1943–1950</td><td rowspan="2">Total net income = Total income minus deficits at source [Sammanräknad nettoinkomst]</td><td>No adjustments</td><td rowspan="3">Tax statistics income plus estimates of non-taxed items included in preferred def. plus estimated incomes of “non-filers” (1943–2003)</td><td rowspan="2">Adult population (16 yrs and above) less housewives (1951–1971)</td></tr>
<tr><td>1951–1970</td><td rowspan="2">Excluding tax returns from those under the age of 16 in each income group.</td></tr>
<tr><td>1971–2003</td><td>Total income [Sammanräknad inkomst (1971–1990), Sammanräknad förvärvs- och kapitalinkomst (1991–)]</td><td>Adult population (16 yrs and above) (1971–)</td></tr>
</table>

[a] All concepts referred to in the table are elaborated upon in the Appendix. The below refers to our preferred versions of the reference totals. To check the robustness of our results we have calculated a number of different series showing that our main conclusions and the trends in the series do not depend on the precise choice of reference total.

Our basic concern is to analyze the share of total income before tax (from all sources) earned by the top decile (P90-100), the top 5 percent (P95-100), the top 1 percent (P99-100) and so on. Furthermore, we distinguish between series including and excluding capital gains respectively and we will also decompose the series according to the source of income. The shares are constructed in relation to a *reference total* for the (tax) population as well as for income. It is important to note that this is different from measuring the share of “total *taxed* income”

received by some share of the *taxpaying* population. Such a measure would fluctuate with changes in tax policy regarding who should pay taxes and what incomes are taxable, without there being any underlying changes in the actual income distribution. Our measures are instead top shares in relation to estimates of 1) the total income as if all income (our preferred definition of it) was subject to taxation and 2) the total tax population as if everyone filed a tax return. The reason for this being the appropriate way to construct homogenous income shares over time is of course that we do not want changes in tax policy regarding, for example, exemption levels and what sources of income are subject to taxation, to affect our income shares.

Given our focus on the top income earners, there are two basic requirements for our series to be correct. First, income according to the tax return should be equal to the actual income of everyone in the top decile. If the *taxable amount in the income tax returns* contains something which we do not want to include, or if it does not contain something we would like to include, we must have information on the *distribution* of this item over income groups so as to be able to deduct or add the right amounts to the right group. Similarly, it must be the case that everyone who actually belongs to the top decile also files a return. The main problem regarding this first requirement is of course potential underreporting of income. Second, we must be able to estimate *aggregate* numbers for all incomes and all potential tax units in the economy. Importantly, however, we do not need information on the distribution of excluded items as long as they do not belong in the top decile.

Our calculations are based on data from income tax returns starting when the modern income tax was introduced in Sweden in 1902. The series are based on pre-tax total income, that is market income from all sources (labor, capital, business and capital gains, the last also being treated separately). The relevant concept in the Swedish income statistics is, hence, *Sammanräknad inkomst* (or the equivalent, see Table 1 above). To make the income data consistent with our definition of reference total population we have excluded all income earners below the age of 16 in each income class.[2]

[2] To be precise we have excluded the number of income earners below the age of 16 in each income class, but we have not been able to deduct their incomes. The income shares of those below 16 are, however, insignificant and much smaller than their share of the total number of income earners. See Appendix C3 for details.

Needless to say, there are a number of problems with tax statistics data; it is collected as part of an administrative routine in which individuals have incentives to underreport income, it tells us nothing per se about the welfare of individuals, tax units do not necessarily correspond to the units we are interested in, initially it only covers a small part of the population, etc. Nevertheless, as long as we think that tax statistics, at least for the top income earners, approximate actual incomes, and as long as the problems with the statistics have not changed systematically over time, they are a useful source. And, importantly, it is the only available source for the whole of the Twentieth Century.

In terms of data availability it is useful to distinguish between two periods; one between 1903 and 1942, the other from 1943 and onwards. In the first period, taxation data divided into income groups are only available for the years 1903 (only the very top), 1907, 1911, 1912, 1916, 1919 and 1920, but the tax returns also formed the bases for income distribution estimates calculated in the Census (*Folkräkningen*) in 1920, 1930, 1935 and 1945, as well as for some studies of income distribution made at the time (e.g. SOU 1936:18 for the year 1934, and Quensel, (1944) for 1937).[3] Combining these sources and making adjustments so as to get homogenous series we have sufficient data to be able to study the development over the whole period, with the very unfortunate exception of the 1920s for which there is only income distribution data for the income year 1920. Appendices A-C provide more details on the data, the adjustments made and sensitivity of our final results with respect to these adjustments. Starting in 1943 taxation data for income classes have been compiled yearly and, hence, data is available for every year after that. The statistics are not however completely comparable for the whole period after 1943 since a number of changes have occurred in definitions and methods of compiling data. Again details of these changes have been confined to the appendices.

The form of the basic data is typically tabulations for a given year of the total number of tax units in a given income class and the sum of their incomes. For example, according to the tax statistics for the income year 1912, there were 245,792 tax units (i.e. singles or married couples, who at the time were treated as one tax unit) with a taxable income between 800 and 900 Swedish kronors (SEK). Their taxable income sums up to SEK 185,247,200. Similarly, there were 148,464 tax units with income between 900 and 1,100 SEK earning together SEK 154,744,400, etc. all the way up to 698 tax units having above SEK 80,100 and a total of SEK

[3] In the Swedish statistics one makes a distinction between the *Taxation year* and the *Income year,* the latter being one year before. Throughout this study all years refer to the *Income year*.

130,830,500 in taxable income.[4] In 1912 taxable income, however, included a wealth share of 1/60 of taxable wealth. Fortunately we have data on the wealth shares per income group, which allows us to subtract the wealth share for each income group so as to get the actual incomes. Taxable income also differed from actual income in a number of other ways, but this difference mainly concerned incomes in the lower part of the distribution, and to the extent that they affected top income earners their effects were negligible. Based on this data we can construct cumulative numbers of incomes and tax.

The main income concept in the official Swedish income statistics is *total income*, which is the gross income of labor, capital, business and capital gains before taxes and most deductions (yet with slight variations over time as shown in Table 1).

To calculate the top income shares we now want to relate these numbers to *total income* in the economy as well as to the *total number of tax units*. For the year 1912 the *total taxed income* (adjusted for the wealth shares) was 1,353,046 thousand SEK but according to our preferred estimate the *total* (individual) *income* was 2,126,596 thousand SEK. In that same year the total number of *tax filers* was 714,919 tax units but the total number of *tax units* in the economy was according to our preferred estimate 2,888,302. Relating these numbers to the cumulative number of tax filers and their total income divided into income classes, we get that 0.099 percent of all tax units (2,851/2,888,302) had 8.968 percent (190,704,625/2,126,596,000) of all income, 0.047 percent (1,353/2,888,302) had 6.704 percent (142,566,868/2,126,596,000) of all income, etc. Based on these shares and the assumption that the distribution is of Pareto form, we can then use standard interpolation techniques to get the income share of the different desired shares of the population (see Appendix D for details on the inter- and extrapolation techniques used).

2.1 Reference Total for Income

In principal there are two ways of constructing the reference total income. Either, one starts with data on "Total Personal Sector Income" reported in the National Accounts and deduct items not included in the preferred definition, or one can start from the "Tax Statistics Income" and add items not included in the tax base and income estimates for individuals not included in the tax statistics. As outlined in Atkinson and Salverda (2003);

[4] See also Table A1 for another example (1945) of how the income tables appear in the published sources.

Personal sector total income
− Non-household income (e.g. charities)
− Items not included in preferred definition of income (like employers' social security contributions).
= Preferred Reference Total Income Definition
− Items not included in the tax base (like certain social security benefits)
− Taxable income of those not included in tax statistics ("non-filers")
= Tax statistics income

We have calculated the reference total income from "both sides" and our final preferred series is a combination of the two for reasons described below (see Appendix B for details and comparisons of different reference totals). The problem, in the Swedish case, with starting with the National Accounts data for Personal sector income is that there are no homogenous series which are detailed enough for the whole period. The most recent and up-to-date National Accounts series have been published by Edvinsson (2005) from which we use data for "Wages and salaries (including social benefits)" and "Imputed labor income of self-employed (including social benefits)" from 1903-2000. Edvinsson (2005) does not, however, publish separate figures for Property income and Capital income. What we do have is aggregate tax statistics for these incomes starting in 1922. We therefore add these aggregate numbers to wages and salaries and the labor income of the self-employed (implicitly assuming that all incomes from these sources were included in the tax statistics). In the period before 1922, for which we have no data on property- and capital income, we add a fixed share of labor income based on the ratio of property- and capital income to labor income in 1922. The resulting series is on average 0.7 times the GDP (calculated from the expenditure side) reported in Edvinsson (2005) with a standard deviation of 0.03 and forms our estimate for total personal income (which still includes some items, such as employers' social security contributions, which we would like to exclude in our preferred definition of reference total income).

Starting from the other end, with the total income according to the tax statistics, gives a relatively precise picture in recent years, as a majority of the adult population file tax returns (actually close to 100 percent (!) since 1978, when government institutions, employers, banks, etc. started to send income statements, so called "*kontrolluppgifter",* to the tax authorities) and as most social security benefits are taxable. Hence, the tax statistics income is close to our preferred total income definition for the past couple of decades. However, going back in time, both the number of "non-filers" and the number of items not included in the tax base increase. To get a homogenous total reference income series we must add the incomes of non-filers and

items not included in the tax base but which form part of total income. The major reform in terms of including social security benefits in the tax base was the 1974 tax reform. This year sick leave benefits, unemployment benefits and parental leave compensation all became part of taxable income. We therefore add the aggregate government outlays for all of these items to the tax statistics income in the period until 1973. We have not included a number of smaller items which were tax free before, and remain tax free after 1974 (such as study grants, child allowances, etc.). When it comes to estimating the incomes of "non-filers" there are two basic approaches which have been used. Piketty and Saez (2001, 2003) and Atkinson and Salverda (2003) assume that non-filers have a fixed share of the average income of those who pay tax (like 0.3 times the average income). Bentzel (1953) and Spånt and Selander (1969) instead assume that non-filers have 0.5 times the threshold income. That is, if everyone with income above SEK 600 should pay taxes (as was the case in Sweden between 1919 and 1951) they assume that everyone not filing a tax return had an income of 300 SEK.[5] We have calculated series in both ways and believe that using a share of the threshold income is the best way to proceed in the Swedish case. Even though this at times gives quite low estimates (and one could argue that it would not be possible to survive on such a small income) we think these same individuals probably get parts of the (at the time tax free) social welfare benefits. Therefore, imputing larger incomes to them we would risk double counting items which we have already added when adding tax free social benefits to the reference total. We think however, that the average income for non-filers is closer to the threshold than 0.5. In our preferred series we impute 0.8 times the threshold income.

Having calculated an estimate for total personal income (which we know to be an upper limit to the preferred income total since it contains items which are not included in our preferred total income definition) and an estimate for the "Preferred Reference Total Income Definition" starting with the tax statistics income, adding income for non-filers and items not included in the tax base we can compare the two series. This leads us to the following preferred specification for the reference total income: 1) For the period 1903-1942 it is 0.89 times the estimated total personal income and 2) for the period 1943-2003 it is the income according to tax statistics, plus an average income of 0.8 times the tax threshold for non-filers, plus social benefits that were not part of the tax base before 1974 (see Table 1). Appendix B contains all calculations and compares the different methods, which in the end does not change the results

[5] See Figure AX showing the lowest taxable income and its share of average income 1903–2003.

very much, nor does it change the trends at all (as is shown in Appendix C). Figure 1 shows the different series as shares of GDP.

Figure 1 (from the xls. File "Finalreftotals1903-2003")

2.2 Reference Total for Population

Just as we do not want to construct top income shares in relation only to taxed income, but to an estimate of "total income", we do not want to define the top shares of the population as a share of those who pay taxes, but as a share of total number of potential tax units, i.e. as if everyone would have filed a tax return. The major problem in constructing such a total in Sweden is that the concept of tax unit has changed from a "family based" system, where married couples filed a joint tax return, to an individually based system where individuals (whether married or not) file separate returns. In the tax legislation this shift took place in 1971, with an option to file separate returns from 1966. In terms of *tax statistics*, however, this change occurred (at least to some extent) already in 1951. Before this, tax statistics were based on the entire tax population and figures referred to "tax units" i.e. married couples counted as one income earner.[6] Before 1951 the obvious reference population is therefore the adult population (which we take to be everybody aged 16 or above) less married women (since a married women formed one tax unit together with her husband). After 1951, however, statistics changed to being based on a representative sample (ten percent) of the population, with married couples where both had an income, now treated as two income earners in the statistics even though they were still *taxed* as one. The problem is that in cases where the women did not work, or had low income, she was not necessarily counted (as is pointed out in many of the sources). This means that income statistics between 1951 and 1971 is a mix between a family based system and an individually based system including some women but not all. Starting 1971, the reference total is again relatively unambiguous, now obviously being the adult population.

In terms of choosing the appropriate reference population in Sweden, the period 1903-2003 can, hence, be divided into the following three periods: 1) 1903-1950, the total population aged 16 or above minus married women, 2) 1951-1970, the total population aged 16 or above minus women likely to be excluded in the statistics, 3) 1971-2003, the total population aged 16 or above. In our preferred specification we define the reference total for the population to

[6] Note that this is the case for *tax statistics* before 1951 but not income figures in the Census (*Folkräkningen*).

be the total population aged 16 or above minus married women in until 1951, the population aged 16 or above minus the number of housewives until 1967, the population aged 16 or above minus a declining share of housewives 1967-1971, and the population aged 16 or above from 1971 and onwards. (We also make some minor adjustments for individuals who have died during the year depending on how they are treated in the statistics). The reason for choosing this particular way of defining the reference total for the population is that it makes for the smoothest transitions at years when statistics (and the number of tax returns) change, without being an ad-hoc adjustment. Figure 2 shows the ratio of the total number of tax returns for a number of alternative specifications of the reference total population. As with the reference total for income we have calculated top income shares using a number of different reference totals for the population and the results can be found in Appendix C. Appendix B contains all details of how the different reference totals for population have been constructed.

Figure 2 (from the xls. file "FinalIncShares1903-2003")

3 Results

We have organized the presentation of our results as follows: We first present the main trends in Swedish top incomes – with and without capital gains - over the Twentieth Century, then we study the composition of income for the different fractiles. All results presented here have been calculated based on our preferred series for adjusted taxed income and our preferred estimates of the reference totals, but in Appendix C we present the main results using different reference totals, which show that while point estimates and levels may change somewhat, the main trends hold true for all our different specifications. We have also checked whether the results are sensitive to changing the age limit for the relevant population from 16 to 20. This does not change the results in any significant way.

3.1 Main trends in Swedish top income shares over the twentieth century

In short, the Swedish top income shares over the twentieth century can very broadly be characterized as 80 years, between 1903 and around 1980, of relatively steady decline, followed by 20 years, between around 1980 and 2003, of increasing top income shares when including capital gains (after 1980 the series peak in the year 2000 after which the top income shares have fallen quite substantially). However, this increase over the past twenty years almost disappears in the data when capital gains income is excluded.

The most rapid period of decline seems to be between the mid 1930s and the beginning of the 1950s. There are not, however, any clear signs that top shares would have declined especially dramatically during the war years as seems to have been the case in France and the US (see Section 4 in this paper). If anything the most rapid decline seems to have been the period just after 1945, suggesting that the decrease in top income shares in Sweden had more to do with policy than exogenous chocks. Seen over the whole period, most of the decline happened before the early 1950s. Figure 3 shows the top income shares of the top ten percent (P90-100), the top five percent (P95-100) and the top one percent (P99-100) including capital gains for the whole period, and also excluding capital gains for the period 1967-2003. The latter series show that when excluding capital gains there is no clear upward trend in the top income shares, suggesting that most of this development is a result of differences in capital gains rather than other types of income.

Figure 3 (P90-100, P95-100, P99-100) (fig3 in the excl-file)

Most of the changes in the top income shares come from changes in the very top of the distribution. The income share of those between the 90th and 95th percentile (P90-95) is almost constant over the past one hundred years(!) The income share of those between the 95th and 99th percentile is also relatively constant for the period after 1950 (before that there is a relatively steady decline in P95-99). This suggests that studying the very top of the distribution is of great importance to understand what is driving changes in measures of the top decile. Again most of the action over the past decades seems to come from capital gains.

Figure 4 (P90-95, P95-99, P99-100, from the fig4 xls. file "FinalIncShares1903-2003")

Looking at the top percent and the top 0.5 percent of income, we see a similar pattern of relatively steady decline, again with most of the decline before the early 1950s and with the sharpest drops in the years 1946-1952 and before the Second World War.

Figure 5 (P99-100, P99.5-100, fig5 from the xls. file "FinalIncShares1903-2003")

For the very top of the distribution, the top 0.1 percent and the top 0.05 percent, the pattern is again similar, but the magnitude of the changes are even larger than for the other groups. Roughly, the share for the top 0.1 and top 0.05 percent of income earners decreased by a fac-

tor of 10 between the beginning of the 1900s and 1980. It then increased by a factor of about 5 between 1980 and 2000. (For the top 10 percent of income earners the drop between 1903 and 1980 was less than a factor 2 and the subsequent increase between 1980 and 2000 about 1.3).

Figure 6 (P99.9-100, P99.95-100, fig6 from the xls. file "FinalIncShares1903-2003")

Another way of looking at the data is to calculate the income thresholds (in constant prices) for the various fractiles and to express these as multiples of average income. Figure 7 illustrates the development of the income needed to be in the respective fractiles of the distribution expressed in 2003 prices.

Figure 7 (income thresholds Fig7 from the excel file).

Figure 8 illustrates these numbers but expressed as multiples of average income. This again suggests that the difference between an individual just at the top ten (or five) percent threshold and the average income earner has not been very large and has not changed much over the past one hundred years, while the income of a top 0.1 percent compared to the average has first dropped dramatically but then also increased substantially in the recent decades (when including capital gains).

Figure 8 (income thresholds as multiples of average income)

3.2 Composition of top incomes

Examining the composition of top incomes is important for the understanding of the development of top income shares. For example, shocks to capital income during World Wars I and II explain much of the secular decline in French top incomes (Piketty, 2003) while drastic increased wage and salary earnings for the already rich has been the prime factor behind the tremendously increased income inequality in the U.S. during the 1980's and 1990's (Piketty and Saez, 2004). The composition of Swedish top incomes also changes significantly during the twentieth century, but not always in tandem with other Western economies.

The sources of income which are distinguished in Swedish tax laws are labor (wages and salaries), capital (mainly interest earnings and dividends), business income and capital gains.[7] Apart from capital gains, these are the standard income sources used by most countries' authorities reporting tax and income statistics, and hence the ones analyzed in previous studies of top incomes (see, e.g., Piketty 2003; Piketty and Saez 2004; Saez and Veall 2003, 2005; Dell 2003, 2005). Capital gains are somewhat different from the other income sources, primarily because they are lumpy outcomes dependent on the timing of their realization. Capital gains can be seen as the result of an accumulated stream of wealth increases (or decreases) over some period of time, and their appearance on the tax return may not coincide with the period when they actually emerged. Therefore, unlike labor income, income from capital gains at time *t* may already have been capitalized by the income earner through past consumption or other economic activities. In practice, however, the other income sources can be well as lumpy as capital gains (consider the income chock of losing or getting a job!) and sometimes capital gains are left out of the compositional analyses simply because they are not included in the country's income statistics (as in France where they were not subject to taxation before the 1980's). As for Sweden, capital gains have for most of the twentieth century been an integral part of income taxation and the official income statistics. A more detailed discussion of the compositional data and the construction of the series can be found in section A2.1.2 in the Appendix. Important to note, however, is that all composition series from 1991 onwards are substantially less reliable than earlier years for two reasons. First, the main income concept used in the tables shifted to earned income (excluding capital income and capital gains) along which the incomes in some composition tables are sorted. Second, the post–1991 concept of capital income also includes capital gains, and though we have estimated the shares of the two sources for the last years they are still indirect observations.

Figures 7, 8 and 9 show compositional cross-sections for Swedish top income fractiles in three years, 1945, 1978 and 1997, both when capital gains are included (solid lines) and excluded (broken lines). Prior to 1945 there are no comprehensive data on income composition (except for capital income, see below) and after 1997 data on all income sources are not freely available from Statistics Sweden (see further Appendix A2). As can be seen in Figure 7 the top incomes in 1945 were dominated by wages in all fractiles except for the top 0.05 percent,

[7] As discussed in Appendix A, the Swedish income statistics reported six different sources of incomes until 1990 and only three thereafter. Using available data and mainly the latter definitions, we are able to construct consistent and continuous series of the four above-mentioned sources for the entire postwar period.

although its share was falling in the level of income. While having much lower shares, both business and capital income increase in income. Capital gains, finally, are completely insignificant at all income levels. In 1978, the dominance of labor income is augmented at all income levels, representing around 90 percent or higher for all income earners below the absolute top (top 0.1 percent and above). Compared with 1945, the increasing shares of labor income and capital gains are the most significant changes. Looking at the composition in 1997, the increase in the share of capital gains is greatly augmented. When including capital gains, the share of wages decrease rapidly as one increases the level of income. For the top 0.05 percent, capital gains represent three quarters of total income while wages only one seventh. Excluding capital gains, however, the picture looks more similar to previous years. Also interesting is the decline in business income, which only represents 1–2 percent of total income at all income levels in 1997, i.e., the same share for above-median and the top 0.05 percent income earners.

Figure 7: Income composition across top fractiles: 1945
Figure 8: Income composition across top fractiles: 1978
Figure 9: Income composition across top fractiles: 1997

The long-run evolution of income composition for fractiles P90–100, P99–100 and P99.95–100 for the period 1945–2003 is shown in Figures 10a-b, 11a-b, 12a-b. In all cases, we present series with and without capital gains (although they are always part of the underlying total income concept). While wages and salaries are still relatively important for the top 10%, they are only about two thirds of the income when capital gains are excluded (about half when they are included). Specifically, after 1980 the share of capital income and capital gains increase significantly for the highest fractiles.[8] For example, when including capital gains, their income share for P99–100 increase from an average of 3.7 percent in the 1970's to 12.5 percent in the 1980's and 30.4 percent in the 1990's. This represents a percentage increase of 243 percent and 143 percent, respectively. For P99.95–100, the average shares for the same periods were 12.1 percent, 39.9 percent and 62.3 percent, reflecting increases of 226 percent and 58 percent, respectively. Interestingly enough, P99–100 increased their capital gains share from 1.1 percent to 3.8 percent and 13.2 percent, which represent percentage changes of 238 per-

[8] It should be noted that capital gains were actually included in labor income before 1967 why their exact shares are unknown in the early years. Census data in 1945 and 1951 suggests fairly modest shares, however, with point estimates of 1.23 percent and 0.42 percent of total net income, respectively.

cent and 250 percent. In other words, in terms of wealth payoffs there was a clear convergence during the 1990's within Swedish top income groups. The exact types of capital gains earned are not disclosed in the data sources we use, but special investigations by Statistics Sweden (Statistics Sweden 2001, 2002, 2003) suggest that the main source is sales of financial assets, and especially so among top income earners. Swedish financial and real estate markets experienced a well-documented boom during the 1980's and 1990's, coming in two waves separated by the financial crisis of 1991–1993. Over the whole twenty-year period, however, the stock prices in Stockholm increased 30 times compared to only ten times in New York, London and Paris.[9]

Another finding is that labor income represented a substantial part of Swedish top incomes during most of the postwar period, since 1970 about 80–90 percent for the top 1 percent and about half for the top 0.05 percent when capital gains are excluded. Over time, however, its share exhibits an inverted U-shaped tendency, peaking around 1980 and then decreasing faster in the 1990's, especially when capital gains are counted. This development could be closely related to the variance in top marginal income tax rates, which when they reached their highest postwar levels in the 1970's and 1980's may have induced top income earners to work less.[10] The tax reform in 1991 also seems to have mattered by its more comprehensive definition of taxable capital income.

The share of business income displays perhaps the most uniform time pattern of all income sources. It declines constantly throughout the postwar period. From initial levels of about 30 percent of total income down to a mere couple of percentage points in the 1990's (after 1997 we have no data on business income specifically) for both the top 1 percent and top 0.05 percent groups. This marked decay is partially explained tax rule changes after 1991, when self-employed small-firm owners became increasingly employed by their own firms and thereby able to choose whether to get their income in the form of wages, dividends or capital gains. A more economic-structural explanation, however, could also be that the Swedish economic postwar institutions have systematically generated incentives for people to be wage earners, and thereby enjoy the full set of public insurances, instead of being self-employed entrepreneurs, with less security in case something goes wrong but still only reaping a fraction of eventual profits due to high marginal taxes (e.g., Henrekson 1996).

[9] These numbers are based on stock price indexes in the *International Financial Statistics* published by IMF.
[10] Klein (1995) finds the tax revenue elasticities with respect to tax rates to be the highest during ...

Figure 11a: Income composition (excluding capital gains) for P90–100, 1945–2003
Figure 11b: Income composition (including capital gains) for P90–100, 1945–2003
Figure 12a: Income composition (excluding capital gains) for P99–100, 1945–2003
Figure 12b: Income composition (including capital gains) for P99–100, 1945–2003
Figure 13a: Income composition (excluding capital gains) for P99.95–100, 1945–2003
Figure 13b: Income composition (including capital gains) for P99.95–100, 1945–2003

Over an even longer time period, 1912–2003, Figure 14 shows the share of capital income of total income in the top.[11] Just before World War I, capital income represented 30–55 percent of top 10 percent–0.05 percent incomes, and this share increased in 1916 to 50–80. While the high prewar levels correspond well with the kind of wealth accumulation witnessed in many European countries during the era of intense industrialization and peace (see, e.g., Piketty et al. 2005 on the French wealth distribution during this time), the remarkable Swedish war boom 1914–1917 with greatly increased corporate profits and stock prices is less commonly observed.[12] After the war, Sweden experienced a severe financial and monetary crisis which depressed markets, and capital incomes, significantly. The build-up during the 1920's resulted again in relatively higher capital incomes in 1930 but then world depression and, especially for the Swedish top wealth, the so called Kreuger Crash in 1932 represented a major blow to Swedish top capital incomes. The late 1930's saw another economic upswing and, interestingly, which for Sweden lasted until after World War II. World War II does not seem to have had any particular impact on capital income shares for Swedish top income earners. During the early postwar years, however, the shares declined substantially (if one is to believe these isolated point estimates) and the decrease was even larger than during the turbulent early 1930's. Between 1950 and about 1980, top capital income shares were relatively stable at their lowest level during the entire century. Thereafter, however, they started to increase and by the late 1990's they had reached approximately the same levels as during the pre-World War I period, indicated a long U-shaped pattern of Swedish top capital income shares.

Figure 14: Capital income share of total income (excl- capital gains), 1912–2003.

[11] Before 1945, we compute this share as a flat rate of return of the personal wealth estimated for different income classes (see further appendix A2).
[12] See, e.g., Östlind (1945, pp. 258ff) or Schön (2000, pp. 275ff).

3.3 Other possible explanations for the trends in Swedish top income shares

3.3.1 Changes in the tax legislation

There are many instances through which changes in tax laws could significantly affect the income of top income earners. For example, drastic changes in the highest marginal income taxes have significant incentive effects on the work efforts of people with high incomes, but also on their propensity to spend resources on avoiding or evading taxation. In the first case, a registered change in top income shares would reflect a true income share change in the specific top fractile whereas in the latter case our observed changes would not correspond to changes in actual income shares but merely the ones reported to tax authorities.[13] Clearly the latter variant is much more problematic for a study of before-tax incomes like ours.

In several of the appendices below, we describe some of the most significant changes in the Swedish tax laws during the twentieth century in order to detect possible correlations with the trends in top income shares. While there are specific times when influences of the tax system on top income shares are apparent, we find no systematic, and especially not one-directional, relation between the tax system and top income shares.

4 International comparisons

4.1 Comparing top income shares

Comparing our results to the evolution of top income shares in other countries reveals a number of interesting differences and unexpected similarities between the Swedish experience and that of Anglo-Saxon countries on the one hand, and "Continental Europe", on the other.[14] These differences and similarities may not only reflect what has happened to top income shares, but also point to more fundamental economic variation across countries.

Figure 15 shows the income share for the top five percent of the population, P95-100, for the Netherlands, the UK (both from Atkinson and Salverda, 2003), the US (from Piketty and Saez, 2003), Canada (from Saez and Veall, 2003) and France (from Piketty, 2003) together with our figures for Sweden. The broad patterns seen here go through for the top ten percent as well as for the top one percent. For Sweden, Canada and the US we distinguish series in-

[13] For a discussion on the endogeneity of efforts and incomes with respect to the tax system, see Bergh (2005).
[14] These are admittedly broad terms given that we only make the comparison between Sweden and the US, UK, the Netherlands and France. However the points we will refer to hold true when including Canada, New Zealand and Australia to the Anglo-Saxon group and Switzerland to "Continental Europe".

cluding and excluding capital gains. In the case of France this distinction is, according to Piketty (2001, p. 20n), not very important as the capital gains share is very small even for the top income earners. For the UK and Netherlands there is no separate data including capital gains.

Figure 15 from excel file Internatcomp2, FigTop5

This figure reveals the most striking aspect about Swedish top incomes namely the fact that Sweden seems to show an Anglo-Saxon pattern over the past decades when including capital incomes but a continental European pattern when capital gains are excluded. Even if the Swedish income share for the top five percent around 1980 starts at a much lower level in Sweden than the corresponding shares in the US, the UK and in Canada, the pattern 1980-2003 is unmistakably Anglo-Saxon when including capital gains. However, when excluding capital gains the Swedish series are as flat as the French and the Dutch ones. This is not the case for any of the Anglo-Saxon countries where the difference with and without capital gains are not very dramatic. (This pattern becomes even more pronounced looking at the top one percent and the top 0.1 percent, see below). Schematically the different experiences can be grouped as shown in table below.

Table 2: Development of top income shares in the period 1980-2003

	Increase	**Flat**
Including capital gains	US, Canada, UK, **Sweden**	France, Netherlands (?)
Excluding capital gains	US, Canada, UK (?)	France, Netherlands, **Sweden**

Development of top income shares in the period 1980-2003 for different countries when including and excluding capital gains."?" denotes that we do not have the exact information on the treatment of capital gains in the cases of the UK and the Netherlands.

Apart from the development over the past decades there are a number of other interesting comparisons to be made. In chronologic order we can first note that the spike during the First World War is there in the Dutch data as well, suggesting that the First World War economy increased top income shares (there is also a sharp increase in the top one percent share for the

US and France during these years as can be seen below). As pointed out above we do not, unfortunately, have sufficient data for Sweden to see any detailed effects of the 1920s, nor the effects of the 1930s depression, which in the US, France and the Netherlands was a time of increasing top shares in the 1920s and then a decrease in the early 1930s. A first major interesting difference between Sweden and other countries comes during World War II. The period 1939-1945 is one of very sharp decline in top income shares in the US as well as in France. This does not seem to be the case in Sweden. If anything the war years in Sweden seem to constitute a slowdown in the decline of top income shares which started in the mid 1930s and then accelerated again after the war. (Holland and the UK are difficult to judge since there is not enough data for the war years). After 1945, however, when the fall in the US and France stops and top shares even recover slightly, they drop sharply in Sweden until the early 1950s. After the early 1950s there is a slow gradual decrease in top shares in Sweden, but not very different from other countries. Our series suggest that, to the extent that Swedish top income shares were different from other countries in the late 1970s, this was a difference established already established in the beginning of the 1950s.

As noted above, top income shares in the US, Canada and in the UK start increasing around 1980. Looking first at the figures for the top five percent of income earners including capital gains, there is not much of an upturn in Sweden until about 1985, (perhaps not until 1990). However, if we look at the top one percent, and the top 0.1 percent of the population, we see that something seems to be happening, starting at the very top of the distribution, around 1980.

Figure 16 from excel file Internatcomp2, FigTop1
Figure 17 from excel file Internatcomp2, FigTop0.1

4.2 Comparing the composition of top incomes

The composition of Swedish top incomes does not fully match the top income composition of any of the other Western countries analyzed previously. For example, while French (in Piketty 2001, 2003), U.S (Piketty and Saez 2001, 2003) and Swedish capital income shares all display marked declines since World War I, they are seemingly more related to the wars (and warfare?) in the two former cases whereas the Swedish decline, especially in the late 1940's and early 1950's, seems more policy-related. Also while U.S. and French top capital income

shares continue decreasing throughout the postwar period, they start increasing again in Sweden around 1980 and by the mid-1990's they reached pre-WWI levels.

As for the distribution of sources of income in the different top fractiles, Sweden is rather similar to most other countries. For example, the share of wages and salaries of total income is strictly decreasing in income in Canada (Saez and Veall, 2003, 2005), France (from Piketty, 2003), and the United States. Similarly, the share of capital income is increasing in the level of total income. Business or entrepreneurial (or self-employment) income is less homogenous across countries, and also less monotonically related to income. In Canada and France it increases up to a certain point (about the top 1 percent–top 0.1 percent) and then starts to decline for the highest fractiles. In the U.S. and Sweden, however, the relation is varying over time between being either quite flat or weakly increasing in income.

Linking top income composition to the overall development of top income shares of the total income in the economy, i.e., changes in income inequality over the twentieth century, has received considerable attention in the past studies of top incomes. While authors explain most of the increased inequality in Anglo-Saxon countries (Great Britain, USA, Canada, Australia and New Zealand) with wage increases for the "working rich" (for a discussion see, e.g., Atkinson and Leigh, 2004) the secular decline many continental European countries (France, Netherlands, Germany) has been explained by both war-related shocks to capital income and relatively compressed distribution of wage increases (see, e.g., Piketty, 2003; Atkinson and Salverda, 2005 and Dell, 2005).

As noted in the previous section, Sweden is similar to continental European countries when capital gains are not counted as part of income. This means that despite observing an increased share of capital income in Swedish top incomes since 1980, this was of sufficiently limited magnitude not to affect overall income shares. When including capital *gains* in total incomes, however, Sweden jumps over to the Anglo-Saxon group of countries experiencing sharp increases in income inequality since 1980 which for Sweden is predominantly driven by a boost in capital gains. As was argued in section 3.2, this Swedish "anglo-saxon pattern" may be related to the formidable upswing on Swedish financial and housing markets during the two last decades of the century. Between 1980 and 2000, Swedish stock prices increased by a

factor of fifty, compared to about a factor ten for the U.S. and U.K.[15] Furthermore, since French stock prices also increased substantially (though not as much as in Sweden - approximately a factor of fifteen) it seems as if Swedish top income earners also were more market-oriented than the French ones who only experienced marginal increases in capital gains over this period (Piketty, 2001, p. 20n).

5 Concluding remarks

So what conclusions about the development of top incomes in Sweden over the twentieth century can be drawn at this stage? Three broad findings strike us as especially important. First, Swedish top income shares have decreased relatively steadily between 1903 and around 1980. Even though the most rapid decline in Sweden was between 1930 and 1950, as in the US and in France, it does not seem to be as clearly related to the Second World War as in many other countries. If anything, the war years seem to be a period of slowdown in the fall in top income shares, which in Sweden was most pronounced just after 1945. This suggests that policy was driving this trend rather than exogenous shocks. Second, Swedish top income shares have – when including all sources of income – increased since the beginning of the 1980s. The pattern of development in Sweden is in this respect close to that in Anglo-Saxon countries, but different from continental Europe. The main reason for this increase in top income shares in Sweden is increases in capital gains rather than wage earnings as in the US and in the UK. Third, the composition of income in the top of the distribution has changed dramatically in Sweden over the past decades. The share of wages and business income has decreased while capital income and especially capital gains have become steadily more important.

References

Atkinson, Anthony B. (2003) "Measuring Top Incomes: Methodological Issues", Working Paper, Nuffield College: Oxford.

Atkinson, Anthony B. (2004), "Top Incomes in the UK over the twentieth Century", *Journal of the Royal Statistical Society A* 168(2), 325–343.

Atkinson, Anthony B., Lee Rainwater and Timothy M. Smeeding (1995), "Income Distribution in OECD Countries: Evidence from the Luxemburg Income Study (LIS)", *Social Policy Studies* 18. Paris: OECD.

Atkinson, Anthony B. and Andrew Leigh (2004), "Understanding the Distribution of Top Incomes in Anglo-Saxon Countries over the Twentieth Century", Working Paper, Nuffield College, Oxford.

[15] Data on stock prices were take from IMF's database *International Financial Statistics*.

Atkinson, Anthony B. and Wiemer Salverda (2003), "Top Incomes in the Netherlands and the United Kingdom over the Twentieth Century", Working Paper, Nuffield College.

Atkinson, Anthony B. and Salverda, Wiemer (2005), "Top Incomes in the Netherlands and the United Kingdom over the Twentieth Century", *Journal of the European Economic Association* 3(4): 883–913.

Banerjee, Abhijit and Thomas Piketty (2005), "Top Indian Incomes, 1922–2000", *World Bank Economic Review* 19(1): 1–20.

Bentzel, Ragnar (1953), *Inkomstfördelningen i Sverige*. Uppsala: Almqvist & Wiksell.

Bergh, Andreas (2005), "On the counterfactual problem of welfare state research: How can we measure redistribution?", *European Sociological Review* 21(4): .

Björklund, Anders and Mårten Palme (2000), "The Evolution of Income Inequality during the Rise of the Swedish Welfare State 1951 to 1973", *Nordic Journal of Political Economy* 26(2), 115–128.

Champernowne, D. G. (1953), "A Model of Income Distribution." *Economic Journal* 63(250): 318–351.

Dell, Fabien (2003), "Top Incomes in Germany throughout the Twentieth Century", forthcoming in A.B. Atkinson and T. Piketty (eds.) Oxford: Oxford University Press.

Dell, Fabien (2005), "Top Incomes in Germany and Switzerland over the Twentieth Century", *Journal of the European Economic Association* 3(2–3): 412–421.

Edvinsson, Rodney (2005), *Growth, Accumulation, Crisis*. *With New Macroeconomic Data for Sweden, 1800–2000*. Stockholm: Almqvist & Wiksell International.

Feenberg, Daniel and James Poterba (1993), "Income Inequality and the Incomes of Very High Income Taxpayers: Evidence from Tax Returns", in J. Poterba (ed.) *Tax Policy and the Economy*, Cambridge: MIT Press.

Feenberg, Daniel R. and James M. Poterba (2000). "The Income and Tax Share of Very High-Income Households, 1960–1995." *American Economic Review Paper & Proceedings* 90(2), 264–270.

Flodström, Isidor (1906) *Redogörelse för statistisk utredning i och för utarbetande af förslag till förändrade bestämmelser dels angående bevillning af fast egendom dels ock om inkomstskatt*, Stockholm.

Flodström, Isidor (1909), *Taxeringen till inkomstskatt år 1907 samt taxeringen till bevillning s. å. för inkomst under 1,000 kronor af kapital och arbete*, Ministry of Finance, Stockholm.

Flodström, Isidor (1914), *Taxeringen till inkomst- och förmögenhetsskatt år 1912*, Ministry of Finance, Stockholm.

Flodström, Isidor (1915), *Taxeringen till inkomst- och förmögenhetsskatt år 1913 jämte uppgifter om däri ingående förmögenhetsdelar*, Ministry of Finance, Stockholm.

Gottschalk, Peter and Timothy M. Smeeding (1997), "Empirical Evidence on Income Inequality in Industrialized Countries" in A.B. Atkinson and F. Bourgignon (eds.), *The Handbook of Income Distribution*." London: North Holland Press.

Gustafsson, Björn and Mats Johansson (2003), "Steps toward equality: How and why income inequality in urban Sweden changed during the period 1925–1958" *European Review of Economic History* 7(2), 191–211.

Hagstroem, K. G. (1944), "Inkomstutjämningen i Sverige", *Skandinaviska Bankens Kvartalstidsskrift*, April, pp. 35–40.

Hagstroem, K. G. (1949), "Inkomstinflation och inkomstnivellering i Sverige", *Skandinaviska Bankens Kvartalstidsskrift*, April, pp. 39–41.

Hartog, Joost and J. G. Veenbergen (1978), "Long-Run Changes in Personal Income Distribution", *De Economist* 126: 521–549.

Henrekson, Magnus (1996) *Företagandets villkor. Spelregler för sysselsättning och tillväxt*. Stockholm: SNS.

Jakobsson, Ulf and Göran Normann (1974), *Inkomstbeskattningen i den ekonomiska politiken. En kvantitativ analys av systemet för personlig inkomstbeskattning 1952–1971*, Stockholm och Uppsala: Industriens Utredningsinstitut och Almqvist & Wiksell.

Klein, Paul (1996), "Swedish Income Taxes 1951–90", mimeo, Stockholm University.

Krantz, Olle (1987), *Historiska Nationalräkenskaper för Sverige. 6, Husligt arbete 1800-1980*. Ekonomisk-historiska föreningen; Lund.

Kuznets, Simon (1955), "Economic Growth and Economic Inequality", *American Economic Review* 45(1), 1–28.

Lindstrand, M. (1949). *Inkomstfördelningen före och efter kriget*. Meddelande från Konjunkturinstitutet, Ser B:10, 89–100. Stockholm.

Lydall, Harold (1968), *The Structure of Earnings*, Oxford: Clarendon Press.

Löwnertz, Susanne (1983), *De svenska skatternas historia* [en artikelserie tidigare publicerad i RSV Info åren 1982-83], Stockholm: Riksskatteverket och Liber.

Magnusson, Katarina (2004), "Utvecklingen av de högsta inkomsterna i Sverige, 1985–2002", *TCO granskar* No. 13/04.

Morrison, Christian (2000), "Historical Perspectives on Income Distribution: The Case of Europe", in *Handbook of Income Distribution*, *Vol. 1*, Elsevier.

Piketty, Thomas (2001), "Income Inequality in France, 1900–98", CEPR Discussion Paper No. 2876.

Piketty, Thomas (2003), "Income Inequality in France, 1900–1998", *Journal of Political Economy* 111(5): 1004–1042.

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

Quensel, Carl-Erik (1944) *Inkomstfördelning och skattetryck*, Lund: Sveriges industriförbund.

Saez, Emmanuel and Michael L. Veall (2003), "The Evolution of High Incomes in Canada, 1920–2000", NBER WP #9607.

Saez, Emmanuel and Michael L. Veall (2005), "The Evolution of High Incomes in Canada, 1920–2000", *American Economic Review* 95(3): 831–849.

Schön, Lennart (2000), *En modern svensk ekonomisk historia. Tillväxt och omvandling under två sekel*. Stockholm: SNS.

SOU 1936:18 (1936), *Undersökningar rörande det samlade skattetrycket i Sverige och utlandet*, Ministry of Finance, Stockholm.

SOU 1949:9 (1949), *Betänkande angåendebeskattning av realisationsvinster, m.m. samt ackumulerade inkomster*, Stockholm: Isaac Marcus.

SOU 1951:51 (1951), *Betänkande angående den statliga direkta beskattningen*. Stockholm: K. L. Beckmans.

SOU 1975:53 (1975), *Beskattning av realisationsvinster. Betänkande avgivet av Realisationsskattekommittén*, Stockholm: Liber.

SOU 1986:37 (1986), *Reavinst, aktier och obligationer. Slutbetänkande av kapitalvinstkommittén.* Stockholm: Gotab.

Spånt, Roland (1976), *Den svenska inkomstfördelningens utveckling*, Ph.D. dissertation in economics No. 4, University of Uppsala, Uppsala: Almqvist & Wiksell International

Spånt, Roland (1979), *Den svenska inkomstfördelningens utveckling 1920–1976*, background material to Långtidsutredningen 1979:4, Ds E 1979:4.

Spånt, Roland and Pebbe Selander (1969), "Inkomstfördelningens utveckling i Sverige 1951–66", mimeo, Uppsala University.

Statistics Sweden, *Statistical Yearbook*, various years.

Statistics Sweden (1923) *Taxeringen till inkomst- och förmögenhetsskatt år 1920.* Stockholm: Ivar Häggströms AB.

Statistics Sweden (1927) *Folkräkningen den 31 december 1920. V. Yrkesräkningen, 2: yrke, inkomst och förmögenhet kombinerade inbördes samt med kön, civilstånd och ålder*, Stockholm: P. A. Norstedt & Söner.

Statistics Sweden (1937) *Folkräkningen den 31 december 1930. VII. Folkmängden efter yrke, inkomst och förmögenhet: 2 avd.*, Stockholm: P. A. Norstedt & Söner.

Statistics Sweden (1938) *Folkräkningen den 31 december 1930. VIII. Folkmängden efter yrke, inkomst och förmögenhet: 3 avd.*, Stockholm: P. A. Norstedt & Söner.

Statistics Sweden (1940) *Särskilda folkräkningen 1935/36. VIII. Partiella folkräkningen i mars 1936: Yrke. Yrkesväxling. Skol- och yrkesutbildning. Inkomst och förmögenhet.* Stockholm: P. A. Norstedt & Söner.

Statistics Sweden (1951) *Folkräkningen den 31 december 1945. VIII:1. Partiella undersökningar (bottensamplingen). Behandlar delar av statistiken över inkomst.* Sveriges Officiella Statistik: Folkmängden och dess förändringar. Stockholm: K.L. Beckmans Boktryckeri.

Statistics Sweden (1956) *Folkräkningen den 31 december 1950. VII. Urvalsundersökningar. Statistiken över inkomst.* Sveriges Officiella Statistik: Folkmängden och dess förändringar. Stockholm: Statistiska Centalbyrån.

Statistics Sweden (2001), *Reavinster och reaförluster 1999*, Rapport 2001. Statistics Sweden: Programmet för ekonomisk välfärdsstatistik.

Statistics Sweden (2002), *Rekordåret 2000. Reavinster och reaförluster. Förmögenhet*, Rapport 2002. Statistics Sweden: Programmet för ekonomisk välfärdsstatistik.

Statistics Sweden (2003), *Minskade vinster 2001. Kapitalvinster och kapitalförluster. Förmögenhet*, Rapport 2003. Statistics Sweden: Programmet för ekonomisk välfärdsstatistik.

Swedish National Tax Board (2000), *Skattestatistisk årsbok 2000*, Stockholm: Fritze.

Swedish National Tax Board (2004), *Skattestatistisk årsbok 2004*, Stockholm: Fritze.

Söderberg, Johan (1991). ”Wage differentials in Sweden, 1725-1950, in Y. S. Brenner, H. Kaelble and M. Thomas *Income distribution in historical perspective*, Cambridge, Cambridge University Press: 76-95.

Sörensen, R. S. (1993), “Changes in the Personal Income Distribution 1870-1986” in: Atkinson, A. B. (ed.) *The Economic Development of Denmark and Norway since 1870*, Cheltenham: Edward Elgar.

Tengblad, Åke (1994) “Beräkning av svart ekonomi och skatteundandragandet i Sverige 1980–1991” in Malmer, Håkan, Annika Persson and Åke Tengblad, *Århundradets skattereform. Del II.* Stockholm: Fritze.

Östlind, Anders (1945), *Svensk samhällsekonomi 1914–1922: Med särskild hänsyn till industri, banker och penningväsen*, Stockholm: Svenska bankföreningen.

Appendix A: Income inequality data

A1: List of sources

A1.1: Sources for the income data

The income data used in this study come from official sources published by tax and statistics authorities as well as a few specific investigations led by the Ministry of Finance. Before 1943, there were no regularly published annual statistics of the Swedish income distribution. Therefore, we are confined to investigations of the Minstry of Finance and and Census material. In particular, for 1903–1912, we use tabulated income distributions made by Isidor Flodström and the Minsitry of Finance. For years 1916, 1919 and 1920, we use income tables published in the *Statistical Yearbook* by Statistics Sweden. For 1930, 1935 and 1940, we use numbers from Bentzel (1952) who in turn uses Census material (i.e., individual incomes) which he has adjusted to make comparable with later tax returns-based material in which tax units are analyzed. For 1934 and 1937, we use special investigations by the Ministry of Finance and adjusted by Quensel (1944). From 1943 onwards, we take the main tables on income distribution for the whole tax population published annually by Statistics Sweden in *Skattetaxeringarna samt inkomstfördelningen inom yrkesgrupper* (later renamed).[16] All income sources for all years are listed in Table A1.

Specific data on the sources of income, e.g., labor or capital income, are not as good as for total income. Prior to 1967, there exist no systematic annual publications on the composition of incomes divided in income levels. On the aggregate level, there are annual figures in the *Statistical Yearbook* starting in 1921. In the Censuses from 1945 and 1950 (Statistics Sweden, 1951, 1956), however, we have nevertheless located two point observations covering all sources of income. As for capital income, we have been able to construct estimates all the way back to 1912 using the fact that a share of personal wealth was included in the taxable income and that Censuses from that time tabulated amounts of wealth for various levels of total income. We simply assume a fixed rate of return flowing from this wealth to the income earners and take this as capital income (not including capital gains). Similar estimates were made by Flodström (1915, p. 46–47) and Statistics Sweden in the *Statistical Yearbook of 1929*. We use the same rates of return used to compute the capital income as Flodström and Statistics Sweden, and as we argue in our main text in this paper these rates are in line with market rates at each point in time. Unfortunately, after 1991 the compositional data becomes less detailed in the income statistics, mainly because of changes in the publication strategy of Statistics Sweden (see further in section A2 about income sources).

Table A1: Sources on income distribution used in the study, 1903–2003.

Table A2: Example of tabulated income data from 1945.

A1.2: Sources for other data

Beside the income data we also need population data to calculate our reference totals for population and incomes respectively. Here the main sources are Statistics Sweden's popula-

[16] In Sweden, the *taxation year*, i.e., the year when tax returns are filed at the tax authorities, is always the year after the *income year*, i.e., the year when the income was made. Hence, 1943 was the first income year for which the new income statistics was used but 1944 was the first taxation year.

tion statistics, *Programmet för befolkningsstatistik*, and the Statistical Yearbook of Sweden, *Statistisk Årsbok,* for various years starting 1914. Our main source for historical national accounts data is Edvinsson (2005).

A2: Details of Swedish income data

A2.1: Definitions of income concept and income sources

A2.1.1: Concept of income

The concept of income used in our main income series is *total income*, as defined in the Swedish income statistics throughout the period as listed in Table 1. Total income is a gross income concept, recorded before deductions and taxes. In terms of specific sources of income (discussed in detail below), total income is the sum of wages, capital and business income and capital gains, although we present calculations of top income share both with and without capital gains included in the total income to make the series internationally comparable.

Above from these main features, the Swedish tax-based assessed total income has also undergone some significant, but yet noteworthy, adjustments over the twentieth century. For example, during 1910–1942 the assessed income from the standard sources were also including a fixed share of the personal net wealth. This "wealth share" entered the taxable income by 1/60 in 1910–1937 and 1/100 in 1938–1948. In the income statistics starting in 1943, however, this share was never counted and for earlier years we have subtracted it from the rest of the income for all income levels based on estimation from Census material on income and wealth.

Since the introduction of the annual income statistics of Statistics Sweden in 1943, the main concept of income has undergone some adjustments. During the period 1943–1970, the main concept used in all tables was *total net income* (in Swedish, *sammanräknad nettoinkomst*), defined as the total income less deductions of deficit in any income source (*underskott i förvärvskälla*), mainly interest payments. After a smaller reform of the tax system in 1971, a new main income concept was introduced by Statistics Sweden: *total income* (*sammanräknad inkomst*). This concept equaled the former except that deductions of deficits in sources were no longer subtracted. In terms of the time consistency of our series, however, this adjustment had no significant impact since the deductions were in general quite small. A somewhat more important change occurred in 1974 when the Swedish government ruled that social benefits, e.g., unemployment insurance, social security transfers and state pensions, were made taxable. Hence, the incomes filed on the tax return changed in structure from one year to another, but this also has little effect on our study of top incomes since the social benefits had a small share of total income in the higher income classes. Moreover, our reference total income is defined so as to include social benefits ever since the 1940's to ensure comparability over time.

Lastly, in 1990–1991 there was the most recent major Swedish tax reform, and along this event Statistics Sweden changed its main concept of income in the income statistics to *total earned income* (*sammanräknad förvärvsinkomst*), which is the sum of labor and business income, hence excluding capital income and capital gains. Fortunately, however, Statistics Sweden continued to publish some summary tables in which they use total income (now called *sammanräknad förvärvs- och kapitalinkomst*) as concept of income and hence we have been able to use the same concept since 1971.

A2.1.2: Definitions of sources of income

As already mentioned in previous appendices, the Swedish tax laws and income statistics define different sources of income which are supposed to be specified on the tax returns. During the twentieth century, these definitions have been remarkably stable although there have been one major change, in 1991. In the official income statistics, however, compositional data have not been published for different classes of incomes until 1967, with exception from two early Census observations in 1945 (Statistics Sweden, 1951) and 1950 (Statistics Sweden, 1956).

The definition of the sources of income before 1991 rested on a division of six different kinds:[17] labor income (*inkomst av tjänst*), mainly wages and salaries; capital income (*inkomst av kapital*), mainly interest earnings and dividends; entrepreneurial income (*inkomst av rörelse*), mainly firm profits and royalties; farm income (*inkomst av jordbruksfastighet*), mainly of sales of agricultural and forestry products and leases; real estate income (*inkomst av annan fastighet*), mainly rents and in-kind payments and capital gains (*inkomst av tillfällig förvärvsverksamhet*) from sales of real estate and securities.[18]

After the tax reform in 1991, the number of official income sources was reduced to three: labor (*inkomst av tjänst*), business (*inkomst av näringsverksamhet*) and capital (*inkomst av kapital* (*överskott*)). Compared with the earlier period, labor income was defined in basically the same way. Business income, however, included not only the previous entrepreneurial income, but also all of farm incomes and a small part of real estate income emanating from rental apartments. In the new concept of capital income, the previous capital income was included but also most of former real estate income coming from private rental and, notably, all forms of capital gains.

To ensure that our compositional analysis is consistent for the whole period, we have defined four main income sources that are used in our analyses. These primarily follow the definitions of the post-1991 for computational reasons, and are constructed as follows:

- *Labor:* Includes wages and salaries and is basically defined in the same way both before and after 1991.
- *Capital:* Includes interest earnings, dividends and real estate income. In the period before 1991, we add "capital income" (interests and dividends) and "real estate income" together.[19] After 1991, estimate capital income from the "new capital income", which includes both the old concept and capital gains. Hence, we break out interest earnings and dividends (called *inkomst av ränta* in the income statistics), private rental income (*inkomst av uthyrning av privatbostad*) and special rental income (*inkomst av positiv räntefördelning*) as is described in detail in section A2.1.3.

[17] In the late 1960's, there was also a specific entry for income from partnerships (*inkomst av delägarskap i vanligt handelsbolag etc*), but this was included in entrepreneurial income from the 1970's onwards and we do this also for these years when it was reported separately.

[18] Detailed descriptions of the income sources are found in, e.g., Statistics Sweden (1945, pp. 50–67) and Statistics Sweden (1975), *Inkomst och förmögenhet 1973*, pp. 25–26.

[19] Formally, one part of the real estate income was also included in business income after 1991, namely income from public rental buildings. However, this only concerned so-called "physical persons" (private individuals) and not "judicial persons" (public and private companies) which instead had to report all of their income (including that from real estate) as entrepreneurial income and which was the largest part of the two incomes. Leif Johansson at Statistics Sweden (from a discussion on June 15, 2005) also would believe that the absolute majority of the real estate income before 1991 should refer to what would after 1991 have been included in capital income. For these reasons, we place all of real estate income in the capital income in our long-run series.

- *Business:* Includes mainly income from privately held firms. Before 1991, we add together "entrepreneurial income" and "farm income". After 1991, we use "business income".
- *Capital gains:* Includes net gains from sales of real estate and other assets. Before 1991, capital gains were reported as a separate source of income. After 1991, we have to estimate it from the new concept income from capital as described in section A2.1.3.

A2.1.3: Estimating the income composition using post-1990 data

Unfortunately, the compositional data after 1990 are much less detailed than and also not fully comparable with earlier years. This is both because of the tax reform of 1990–1991 and the redefinition of the sources of income. A more tangible cause, however, is the fact that Statistics Sweden started listing all compositional data according to different classes of *earned* income, instead of *total* income which was the case for the period up to 1990. Remarkably, this income numeraire is also used when listing various kinds of income of capital despite the fact that capital income was not even included in earned income.[20] Hence, there is a large portion of income earners with positive capital income but zero earned income (e.g., children with significant wealth). To get comparable series with earlier years we therefore have to two adjustments: First, we must map these compositional tables into classes of total income. Second, we must separate out the capital gains part of the new income of capital-concept. This cannot be done with perfect precision since we do not have the primary individual taxation material. To make sure our series are roughly consistent approximations, we carry through a number of checks against data from the years just before 1991.

We deal with these problems as follows. We start by transforming the three new income sources (labor, business and capital, including capital gains), listed in classes of earned income, into classes of total income. This is done by summing wages and business income together for each class of earned income. Recognizing that earned and total incomes are classified in the exact same nominal amounts, we then subtract the earned income classwise from the total income. The resulting difference should then be roughly the total capital income in each class listed in classes of total income. A slight correction is made to ensure that the sum of these differences is the same as the reported aggregate capital income.[21] This gives us income sums for all three sources in classes of total income, adding up to their correct income totals.

The next step is to separate out capital gains from the new capital income-concept and achieve them listed in total income classses. Unfortunately, it is not straightforward to do this since capital gains and capital income are listed in gross terms (incomes and losses) whereas the new capital income-concept (*inkomst av kapital* (*överskott*)) shows the sum of all individual surpluses, i.e., what is left after incomes and losses are netted out against each other whenever possible.[22] Since 1991, the tax laws do not differentiate between different kinds of capital losses and gains why there is no way to tell from the publications which kind has been used for what. Instead, we calculate the share of capital gains from their share of the total gross

[20] Table 6.1 and 6.2. in SM, called "Inkomsttagarna med inkomst av capital fördelade på inkomster och avdrag efter kön, ålder, sammanräknad förvärvsinkomst".
[21] We use the reported capital income (*inkomst av capital (överskott)*) as true total. The correction is very small, never larger than roughly one percentage point of the total and often practically zero.
[22] The existence of a surplus (total capital income) is due to the fact that all individuals with capital gains did not have corresponding capital losses which they could use to net out the taxable gains. The gross income is bigger than total capital income since it contains incomes which were later netted out by losses.

capital income (*summa kapitalinkomster*) and assume the net gains to be the same of corresponding net capital income. We then multiply these shares with the capital income (including capital gains) for each income class of total income using the approximated series derived in step one.

The third, and final, step is to compute the capital income and capital gains. This is simply done by multiplying the share of capital gains of total capital income (and one minus that share for capital income) with the reported total capital income for each income class.

An obvious problem with this approach is the unobserved individual variation that potentially could bunch together income earners with different total income in the same total income class. For example, our estimate gives the same capital income to all persons with some earned income, say almost zero, despite the fact that they can be either practically poor and unemployed or wealthy *rentiers*. These two groups would end up in different classes of total income, but this is nothing our method can take into account. For our approach to be roughly adequate, we need individuals with similar income composition to end up in approximately their correct income class. To "test" the plausibility of the approach, we compare the shares of capital income across income classes between two years before the tax reform, when this problem did not exist, and the first two years after the reform. We do not wish to use more than two years since the income composition varied quite drastically during these years. Inspecting the two-year average ratios of earned to total income 1989-90 and 1991-92, they seem to match remarkably well, especially from the median income and above. Using a Kolmogorov-Smirnov test to test formally for similarity between these two distributions of capital income over total income, we get a *D*-statistic of 0.3012 and a *p-value* of 0.265, i.e., we cannot reject the null hypothesis that they are drawn from the same distribution.

A2.1.4: Estimating the top shares of capital income, 1912–2003

Thanks to early wealth data in the tax statistics for income earners in different classes of total income, we are able to construct shares of capital income of total income as far back as 1912 and for some more years until the postwar period when we use the compositional sources described previously.

Specifically, the shares before 1945 are computed by assuming that capital income is a fixed rate of return flowing from the individuals' net wealth. Information about net wealth in different classes of income is available from the tax-based income statistics due to the fact that 1/60 of that wealth was to be added as taxable income until 1938 when the share was reduced to 1/100 and 1943 when it was removed altogether (recall Table 1). The approach was previously used by, e.g., Flodström (1915, pp. 46–47) and Statistics Sweden (1927). Capital income is then computed as the annual rate of return from this wealth. We assume that the yield is flat and the same for all income earners, being 5 percent in the years 1912 and 1919, 5.5 percent in 1920, 4.5 percent in 1930 and 3 percent in 1935. These are the same rates that Flodström and Statistics Sweden use (except for 1920 when they use 5 percent).[23] Unlike them, however, we can also motivate our choice of these rates by referring to three other reference interest rates from the same particular years. Specifically, the yearly averages of the minimum lending rate (*diskontot*) set by the Swedish central bank, the average deposit rate at Swedish savings banks and the effective Swedish Government bond yield were in 1912: 4.81, 4.35 and 4.80; in 1916: 5.23, 4.76 and 5.09; in 1919: 6.38, 5.08 and 5.71; in 1920: 6.92, 5.16 and 7.00;

[23] Unfortunately, no income data were collected in the Census of 1940, why we have no information about wealth shares in different classes of income.

in 1930: 3.71, 5.22 and 4.18; and in 1935: 2.50, 3.59 (in 1933) and 3.30 (*Svensk Sparbankstidskrift* 1934, p. 825).

Finally, the share of capital income of total income across the various top fractiles is computed using Pareto interpolation in the same way as in the rest of the compositional analysis.

A2.2: Concept of income earners

The concept of income earners is problematic in most countries' tax-based income statistics, often associated with redefinitions and changes in the underlying demographic structure of the country. Sweden is no exception although the situation seems not more complicated than elsewhere. First of all, income earners in the Swedish income statistics refer to physical persons who lived in Sweden during the income year (January 1–December 31 the year before hading in the tax return) and who also filed a personal tax return.[24]

We use the definition of income earners according to published income statistics, which is mostly, but not always, identical with what is written in the tax legislation. Until 1971 Sweden had a *family-based* tax system, which means that families were the tax units and all incomes of family members added together and filed under a single tax return. Under this system, married couples hence counted as one income earner. While before 1954 the wife's income was automatically assessed as a part of her husband's income, however, between 1954 and 1965 spouses filed separate tax returns after which their incomes were lumped together and taxed as one tax unit according to a specific rate of "joint taxation" (*sambeskattning*). In the period of 1966 to 1970, the system was further adjusted so that married couples could choose whether to have their income being taxed separately or as one couple according to the specific scale. Finally, in 1971 the Swedish tax system changed to being *individual-based* altogether. Married couples were from that on hence treated as two income earners.

In the period 1943–1950 the income statistics followed the tax system by being family-based, using the total number of filed tax returns as primary material. Due to processing constraints, however, only a few variables could be collected for each tax unit and therefore it was decided to switch to a sample-based system that allowed more background information to be collected and analyzed. During 1951–1966, the income statistics was hence based on a representative 10 percent sample of the tax population. In practice this meant that the income statistics became individual-based while having a family-based tax system since all persons with positive income had to file an individual tax return regardless of whether they were eventually taxed jointly with their spouses or parents.[25] The 10-percent sample was drawn from the population of all adults aged 16 years or above and born on either the 5th, 15th or 25th in each month.[26] To avoid undersampling of income earners in the highest income classes, a full sampling of these groups was made.[27] The sample-based income statistics lasted until 1967 when Statistics Sweden returned to basing the income statistics on the complete tax population with the help of new data processing techniques.

---

[24] Formally, unfinished death estates and family foundations are also counted as income earners, but they only represent about 1 percent of the total number of income earners.

[25] The switch to using a population sample followed the instructions of a governmental statute (*kungörelse den 21 december 1951*, No. 832).

[26] Having in fact 365.25 days calendar year, the chosen sample was actually smaller than 10 percent of the population and instead of multiplying each income earner with 10 (for those jointly assessed 5) it should have been 10.146 (and 5.340). As noted by Statistics Sweden in *Inkomst och förmögenhet 1968*, p. 26 (see Appendix sources), this could have some minor effects on the comparability of the data before and after 1967.

[27] The definition of high income was SEK 30,000 or above during 1951–59 and with income above and SEK 50,000 or above in 1960–66.

Above from these major changes in the income earner definitions, there have been several smaller adjustments and related changes that have affected the income earner concept in particular. For example, in income years 1972 and 1973 all retirees receiving public pension only (*folkpensionärer*) were granted extra large tax deductions to avoid paying taxes.[28] Furthermore, from 1978 both employers and employees were required to report all incomes paid and received, which in itself increased the tax liable population by a couple of hundred thousand income earners who were most likely previously avoiding taxes altogether.

A3: The Swedish income tax system, 1903–2003

A3.1: General overview

The modern Swedish income tax system was introduced in 1902, with the first income year being 1903. During the twentieth century there have been some major reforms and numerous minor adjustments of the income tax system. This section lists some of the most important ones.

The Partial Income Tax Act (*lagen om partiell inkomstskatt*) of 1902 was the first comprehensive progressive tax on income in Sweden.[29] The word partial referred to the fact that Sweden already had an "income tax" since 1812, called *bevillning*. This was a proportional tax (generally 5 percent of all incomes) and fulfilled other functions than merely being a source of revenue to the state. For example, the income registers based on the *bevillning* were used to decide income weights for the adult men who were allowed to vote for government in income-weighted electoral system. The new tax in 1903, did not only introduce progressivity but also a requirement that everyone with an annual income of SEK 1,000 or above had to file a personal tax return (*deklaration*) with detailed description of amount and type of incomes. The progressivity of the tax was though fairly moderate, with marginal tax rates from 0.2 percent to 5 percent. In the first tax year 191,515 families and individuals filed taxes, representing about seven percent of the entire tax population (see Table A1). Soon after the introduction of the partial income tax, inconsistencies arouse in the tax system by having two parallel income taxes. Decision-makers therefore decided to keep the new income tax (which generated the most revenues) while abrogating parts of the *bevillning*.[30]

In 1910, a new state income tax came, the Income and Wealth Tax Act (*lagen om inkomst- och förmögenhetsskatt*).[31] This law increased progressivity, with marginal tax rates between 0.4 and 6 percent. Furthermore, the concept of taxable income was expanded, adding to the taxable income (from labor, business, farming, capital, and capital gains) also a share (one sixtieth) of the tax filer's net personal wealth. This law was kept intact until 1947 with the exception that the wealth share was lowered to 1/100 in 1938 when another wealth tax was introduced.

[28] See, e.g., Statistics Sweden, *Inkomst och förmögenhet 1973*, p. 15.
[29] However, almost a hundred years earlier, in 1810–1811, an early progressive income tax was introduced but removed after only two years of practice (Löwnertz, 1983).
[30] Bevillning was also launched on people's estate. A brief description of these matters (in Swedish) can be found in Löwnertz (1983).
[31] Law (*förordning*) of October 28, 1910.

In 1928, the Municipal Tax Act of 1928 (*kommunalskattelagen*) extended the taxation rights of the regions and in fact shifted the major part of the overall income tax burden from the state level to the municipalities.[32]

The next reform of the state income taxes came in 1947, with the State Income Tax Act (*lagen om statlig inkomsskatt*).[33] This reform separated the taxation of income and wealth by the state and removed the addition of a wealth share to the taxable income. Other changes in the reform was that it introduced tax collection at source which made employers responsible to deduct taxes before paying wages and salaries.

The most recent change of the Swedish tax system was the tax reform of 1990–1991, which changed the structure of the entire tax system. In particular, tax bases were broadened to increase the uniformity of treatment of different sources of income. At the same time, tax rates were lowered to prevent a general increase in the overall taxation.

A3.2: Marginal and average income tax rates in Sweden, 1903–2003

Marginal taxes were gradually increased between 1910 and 1940, from levels of 0.4–6 in 1910, to 3–15 percent in 1919 and 2–31 percent in 1938.[34] With the new law of 1947, marginal taxes were further raised to between 10 and 70 percent. This level was kept until 1971 when the highest marginal taxes were lowered to 65 percent.[35] In the 1980's, marginal tax rates reached their highest levels of the entire century, peaking on 85 percent during several years [More to come...].

A3.3: Changes in capital gains taxation and its effect on our results[36]

A3.3.1: Rules for capital gains taxation, 1903–2003

Capital gains have been taxed as income in Sweden since the 19th century, although the forms and rates of taxation in different periods have varied substantially. Only to define what capital gains really are has been a delicate matter for tax authorities and legislators, e.g., concerning the differences between capital gains and various kinds of business or capital income. During our period of analysis, one can discern two main periods of capital gains taxation: 1903–1990, during which capital gains were added to and taxed together with labor income, and 1991– , when capital gains were taxed separately at a flat rate different from other income tax rates. Within these two periods, however, the taxation of capital gains have undergone several smaller adjustments that are described in the following.

Before 1910, Swedish tax laws did not explicitly mention capital gains but only that the "income from non-professional sales" should be added to labor income and tax accordingly. Capital losses were generally deductable from labor income in the same way. From 1910, new laws defined capital gains (*inkomst av tillfällig förvärvsverksamhet*) as the gains from non-professional sales of property (agricultural or urban real estate) or other assets (e.g., securities) acquired through purchase or barter and owned for no more than 10 years (for property) or 5 years (for other assets). This definition implied that gains from sales of previous gifts or be-

[32] SFS 1928:370.
[33] SFS 1947:576.
[34] SOU 1964:25 (1964), p. 251.
[35] Jakobsson and Normann (1974), Table 59, pp. 191–194.
[36] This section relies on discussions in SOU 1949:9 (1949), pp. 20–52; SOU 1975:54 (1975), pp. 47–70; SOU 1986:37 (1986), pp. 53–66, Statistics Sweden, *Inkomst, skatter och bidrag 1997*, If 20 SM 9901, pp. 10–11, and Swedish National Tax Board (2000), pp. 82–83.

quests were not taxable. Capital losses could be deducted against gains only if they emanated from the same type of assets (i.e., losses from stock sales could not be deducted against gains from real estate or lottery sales). The tax law of 1928 did not change the capital gains taxation except through further restricting the deduction possibilities of losses through, this time in terms of where they were made (i.e., losses from real estate sales in Stockholm were not deductible against gains from real estate sales in Gothenburg).

In 1951, the laws were significantly changed in that the holding period of the property or assets that were sold with a gain was shortened to qualify for taxable capital gains. In a falling scale, possessions of less than two years rendered taxability of 100 percent of the gains, between two and three years 75 percent of the gains, between three and four years 50 percent of the gains, between three and four years 25 percent of the gains and, finally, longer than five years gains were not taxed at all. Moreover, the tax base was extended to include also gifts and bequests. Between 1967 and 1976, a further tax base extension was to make taxable 10 percent of all gains from selling assets held longer than five years and rendering at least a five percent profit. Between 1977 and 1990, the falling scale was replaced by a division into "younger" (owned less than two years) "older" (otherwise) assets, with full taxability for the former and 40 percent for the latter.

The tax reform in 1990–1991 separated the taxation of capital gains from the rest of the income taxation. The tax rate was flat and for real estate only half the gains were due to tax until 2001 when two thirds were made taxable. The rates have varied since 1991, from 30 percent in 1991–1992, 25 percent in 1994, and finally 30 percent since 1995.[37]

A3.3.2: The effect of changes in capital gains taxation on our study

[To be written]

A3.4: Lowest taxable income threshold

Since the introduction of the Swedish income taxation in 1903, there have been rules defining the smallest annual income that requires the income earner to file a personal tax return (in Swedish *deklarationspliktgräns* or *"skattestreck"*). As shown in Figure A1, the lowest taxable income was SEK 1,000 in 1903 and then lowered to SEK 800 in 1910 and SEK 600 in 1919, which remained the case until 1951 when it was doubled. In the postwar period, this nominal threshold has been raised at a few occasions until it was linked to an inflation-index in the 1990's. Figure A1 also shows how the threshold relates to the average annual income, defined as ratio between the reference total income and the reference total population. Over time, the share of the lowest taxable income threshold of the average income has declined sharply, meaning that lower incomes were made taxable income. As for our analysis, this affects the size of our reference total income concept, which includes both the total assessed income (i.e., the income reported by all tax filers) and a fixed fraction of the lowest taxable income threshold for all those outside the tax population (see below). The reference total population, however, is not sensitive for these changes since it is based on total population counts rather than number of tax filers. For that reason, not even the discrete threshold increases in 1952, 1962

[37] More correctly, the 15 percent tax on real estate-related capital gains is a 30 percent tax on half the gains. The variation in tax rates across income years corresponds to switching political colors of the government: being Social Democratic before 1991 and after 1995 and right-wing in between.

and 1971 which reduced the number of tax filers by at most up to three percent, has any significant bearing on the top income shares analyzed by us.[38]

Figure A1: Lowest taxable income and its share of average total income, 1903–2003.

A3.5: Tax evasion and tax avoidance

Among the graver problems with using tax returns when analyzing income distribution is the presence of *tax avoidance* and/or *tax evasion*. While tax avoidance is the minimization of tax liability by legally exploiting the tax regime, tax evasion denotes efforts by individuals who evade tax payments by breaking the law. Arguably both these activities are endogenous to the tax system as a whole, and countries with relatively high marginal and average tax rates, such as Sweden, would be expected to have relatively more of them. Empirically, however, both are difficult to observe and hence measure, why we know little about their extent both today and in the past.

The presence of tax avoidance and tax evasion primarily affect our results when they differ systematically in extent across different time periods as well as levels of income. While both reduce the overall assessed amount of income, and hence also our reference total income which is used as denominator for measuring income shares, they do not affect the top income shares if all income earners in all income classes practice them to the same extent (e.g., as percentage of their incomes). Only if there is a systematic difference across income classes we may have problems. If, say, only the below-median income earners evade taxes by working substantially in the "black market", the reference total income is reduced by a large amount while keeping top incomes intact and hence leading to an overestimation of their income shares. Similarly, if only the top 10% practiced significant tax avoidance and evasion, they would get smaller shares of total income than would be if all their "true" income was assessed. As the anecdotal evidence from the history of Sweden shows, it is difficult to conclude anything about significant differences across income earners with different income both due to lacking data and the use of various measurement techniques. On the whole, however, it seems like both low- and high-income earners evade and avoid taxes but in different ways. Whereas the former more often work on the "black market" where they get paid in cash or in kind, the latter more often hide their sales of foreign financial assets and use advanced legal tax planning to avoid taxes.

The earliest statement found on tax evasion is in 1919 when a special inquiry into the extent of past evasion was carried through (Statistics Sweden, 1923, p. 13*). It was found that 331 income earners had evaded taxes partly or fully and that their (eventually) taxable incomes amounted to a sum of were SEK 3,640,720, i.e., an average taxable income (including wealth shares) of SEK 11,700. Compared with the overall average income (excluding wealth shares) in 1919, SEK 2054.5, the tax evaders were hence people earning about five times the average household which actually corresponds to somewhere around the 97–98th percentiles of all income earners (see Table AX – "Income thresholds, as multiples of average income). While being distributionally skewed towards the top, the tax evaders were not so significant in relation to the total number of top income earners. In 1919, there were about 3,000 income earners in the top 1% group and hence...

[38] The doubling of the threshold in 1962 was estimated to decrease the number of income earners by about 125,000, representing about 3 percent (Statistics Sweden, *Skattetaxeringarna samt fördelningen av inkomst och förmögenhet taxeringsåret 1963*, p. 21).

For the late 1940's and early 1950's, estimates of the aggregate evaded, or "under-assessed", income point at levels of 8–11 percent of all personal income, of which half was labor income and the rest business (entrepreneurial) income (Bentzel 1952: pp. 64–65, 156–157). In a public investigation of direct taxation from the same period, further anecdotal evidence comes from surveys among tax collection officials. They report extensive evasion among practically all sources of income, but especially business income, where wages are replaced by in-kind payments, and the amounts written on invoices reduced while the rest is paid in cash (SOU 1951:51, pp. 352–361). If one would assume entrepreneurs to be more often top income earners than the rest of the population, this evidence hence suggests that evasion of top incomes around 1950 was significant.

For most of the postwar period, wage earners have a harder time avoiding, or evading, taxes than have self-employed since Sweden after 1947 practiced income taxation at source (see section A3.1). Estimations from the late 1960's suggest that tax evasion in Sweden was moderate (see, e.g., Spånt, 1976: pp. 27–29). For example, local tax collection bureaus (*taxeringsnämnderna*) made complaints on deductions in over 600,000 tax returns corresponding to on average taxes of SEK 1,100, this would if true only result in about 0.8 percent higher total income of the population. For top income earners, however, the addition would be somewhat larger as about every forth such tax return was filed by a top 5% income earner. When considering all tax returns actually checked by the tax authorities, the increased assessed income was only about SEK 20 million, which corresponds to or 0.02 percent of total taxed income).

After the tax reform of 1990–1991, a series of estimations of the size of the aggregate tax evasion has been made. In Tengblad (1994), the estimated amount of income evaded from taxation was SEK 60 billion, i.e., approximately five percent of the total assessed personal income. Later this figure was lowered, and for 1997 the amount was taken to be about SEK 20–40 billion, or about 1.5–3 percent of total assessed income (Swedish National Tax Board, 2004: p. 218). In interview surveys during 1997–1998, it was found that between 11 and 14 percent of the respondents, representing between 650,000 and 800,000 individuals, had themselves worked on the "black market". Their average hourly wage of was SEK 112, which is only slightly higher than the average private sector wage and hence indicates that very few, if any, of these respondents belonged to the top 10% income earners.[39]

Income from capital and capital gains in the 1980's and 1990's has also been hidden from taxation, but at quite varying degrees. Domestic interest and dividend payments were hardly hidden at all due to extensive control mechanisms and double reporting rules.[40] As for foreign holdings, however, one estimate suggests that 40 percent of all individuals receiving foreign capital income did not assess them on their tax returns. Realizations of sales of stocks were also improperly assessed, leading in 2001 to reduced tax income of approximately SEK 900 billion (Swedish National Tax Board, 2004: p. 222).

Appendix B: Construction of reference totals

In this Appendix we explain in greater detail exactly how our reference totals have been constructed. The different reference totals are used in Appendix C to test the robustness of our series to the choice of reference total.

[39] The average hourly wage for private sector blue-collar workers was SEK 106.8 in 2000 (data from Statistics Sweden: *Lönestrukturstatistik, privat sektor, genomsnittlig timlön (totallön), arbetare*).
[40] Estimates suggest that the amounts hidden represented about one percent of total capital income during the 1980's and a tenth of the for the 1990's (Swedish National Tax Board, 1994: p. 221).

B1: Reference total population

As described above, there has been one major change in Swedish tax legislation in the Twentieth Century which has fundamentally changed the concept of tax unit, namely the 1970 tax reform shift from a family based tax unit to an individually based concept. In terms of *tax statistics*, however, this change occurred (at least to some extent) already in 1951. Before this tax statistics were based on the entire tax population and figures referred to "tax units" i.e. individuals as well as married couples counted as one income earner.[41] Before 1951 the obvious reference population is therefore the adult population (which we take to be everybody aged 16 or above) less married women (since a married women formed one tax unit together with her husband). After 1951, however, statistics changed to being based on a representative sample (ten percent) of the population with married couples, where both had income, now treated as two income earners in the statistics even though they were still *taxed* as one unit. The problem is that in cases where the women did not work, or had low income, she was not necessarily counted. This means that income statistics between 1951 and 1971 when the individually based system was fully introduced (for labor income, tax on capital income remained family based) is a mix between a family based system and an individually based system including some women (those with substantial income) but not all. Starting 1971, the reference total is again relatively unambiguous, now obviously being the adult population.

Apart from the quantitatively more substantial decisions discussed above there are a number of smaller adjustments which can be considered. Over the course of a year individuals move in and out of the country, some die, some turn 16 after the population count but before taxes are filed, etc. Based on recent years when we believe that the coverage in the tax statistics is close to complete we have concluded that correcting for deaths is most important. The tax statistics before 1951 contain tax returns for those who died during the previous year (the income year), in the period 1951-1973 these are not present in our data, but from 1974 and onwards they are again part of the statistics. We have therefore added deaths to our reference total for the population before 1951 and after 1973.[42] For these periods we therefore add the number of deaths during the year when calculating the reference total population.

In terms of choosing the appropriate reference population the period 1903–2003 can, hence, be divided into the following three periods: 1) 1903–1950, the total population aged 16 or above minus married women, 2) 1951–1970, the total population aged 16 or above minus women likely to be excluded in the statistics, 3) 1971–2003, the total population aged 16 or above.

For the period 1903-1950 the reference total population is:

The population aged 16-	(from Statistics Sweden, Population statistics, *SCB Programmet för befolkningsstatistik*)
– married women	(from Statistics Sweden, Statistical Yearbook of Sweden, *Statistisk Årsbok,* various years)

[41] Note that this is the case for *tax statistics* before 1951 but not income figures in the Census (*Folkräkningen*).
[42] To be precise, deaths are not in the statistics 1951-1966 (though they are taxed) while they are separetly accounted for in the period 1967-1973 and hence we can exclude them from our tables. References for the treatment of deaths are e.g.: for the period before 1951, *Statistics Sweden, Inkomst och förmögenhet 1969, p. 11,* for the period 1951-1966, *Statistics Sweden, Skattetaxeringarna…1966*, p. 32,for the period 1967-1973 *Statistics Sweden, Inkomst och förmögenhet 1969*, p. 13-15, 20-21, and after 1974 *Statistics Sweden,*SCB SM N 1976:4 (p.2) and SCB OE 21 SM 0501.

+ deaths during the year	(from Statistics Sweden, Statistical Yearbook of Sweden, *Statistisk Årsbok,* various years)

For the period 1951-1971 the preferred reference total population is:

The population aged 16-	(from Statistics Sweden, Population statistics, *SCB Programmet för befolkningsstatistik*)
– "housewives"	(before 1950 data is from Krantz (1987) and after 1950 the difference between men and women in paid work (see Edvininsson (2005), p. 140)). All "housewives" are subtracted before 1967 (when individual taxation became voluntary) there after a linearly declining share between 1967-1971.

For the period 1972-2003 the preferred reference total population is:

The population aged 16-	(from Statistics Sweden, Population statistics, *SCB Programmet för befolkningsstatistik*)
+ deaths during the year	(added after 1973 since they reappear in the statistics in 1974, from Statistics Sweden, Statistical Yearbook of Sweden, *Statistisk Årsbok,* various years)

To check the robustness of our results we have calculated a number of alternatives which differ mainly in the period 1951-1971. These are sometimes not "alternatives" in the sense that we may know that they are clear over-, or underestimations, but rather they serve the purpose of giving bounds to our estimates. Figure XYZ shows the population aged 16 and above, the number of tax returns and the different alternative specifications. The alternative specifications are the following:

- *Preferred series* = (Pop 16-)-Married W + deaths for 1903-1950, (Pop 16-)-(Housewives for 1951-1966), and from 1967- Pop 16-, subtracting declining share of housewives 1967-1971 and adding deaths after 73 (1974-).
- *Alt tax units 1* = (Pop 16-)-Married W for 1903-1950, and from 1951- Pop16-
- *Alt tax units 2* = (Pop 16-)-Married W for 1903-1950, (Pop 16-)-Housewives for 1951-1966, and from 1967-, Pop 16-
- *Tax units 3* = (Pop 16-)-Married W + deaths for 1903-1950, (Pop 16-)-(0,5 x Housewives for 1951-1966), and from 1967- Pop 16-, subtracting declining share of housewives 1967-1971 and adding deaths after 73 (1974-)
- *Tax units 4* = (Pop 20-)-Married W for 1860-1950, and from 1951- Pop20-

Figure Diagreftotpop1 from FinalIncShares1903-2003June2805

Looking at the behavior of the ratio between the number of tax returns and our reference series, especially around the critical years when there are changes in the definition of tax unit, i.e. 1951, 1967 and 1971, indicates which series seem best. Put simply, we do not want there to be any sudden jumps in the ratio unless there are underlying real changes in the tax base. To exemplify, in 1919 the tax threshold was dropped from SEK 800 to SEK 600 leading to a real major expansion of the tax base. Here we expect the ratio to go up sharply. In 1951, however, the change was only in the type of statistics, not in the actual underlying number of tax eligible individuals (units), so here we should not expect a break in the ratio. To the extent

that the number of returns increase, this should be compensated by an increase in the reference total. At the same time, we do not, of course, wish to make *ad hoc* adjustments to keep the ratio fixed, since there are real changes in the number of tax filers. Figure ABC shows the ratio between the number of tax returns and our preferred series with indications of critical breaks (the ratios between the number of filed tax returns and the alternative series above are shown in Figure 2, in Section 2.2 above).

Figure Diagreftotpop3 from FinalIncShares1903-2003June2805

B2: Reference total income

In constructing our reference total income we have used three basic approaches. The first two are based on that we can (as explained in Section 2 above) arrive at the "Preferred Total Income Definition" either by 1) starting with "Total Personal Sector Income" and deducting items not included in our preferred definition, or 2) by starting from the "Tax Statistics Income" and add items not included in the tax base and income estimates for individuals not included in the tax statistics. The third - which is mainly included as a point of reference - is based on the assumption that our preferred income total can be approximated as a fixed share of GDP.

Starting with the first approach, we need homogenous estimates of "Total Personal Sector Income" from which we want to deduct items not included in our preferred definition of total income. The best homogenous National Accounts series which span the whole period which we study are those by Edvinsson (2005). These, however, contain only aggregate series for *Wages and salaries of employees (including social benefits)* and *Imputed labor income of self-employed (including social benefits)*. To these we have added aggregate *capital income* and *property income* reported in the tax statistics giving us an estimate of "Personal sector total income".[43] This, hence, becomes:

Wages and salaries of employees (including social benefits)	(from Edvinsson, 2005)
+ Imputed labor income of self-employed (incl. social benefits)	(from Edvinsson, 2005)
+ individual capital income	(from *Taxeringarna…,* 1922-1988, and corresponding sources thereafter, and estimated before 1922).
+ individual property income	(same as for capital income above)
= Estimated "Personal sector total income"	

This estimate fluctuates around 0.7 times GDP (calculated from the expenditure side, reported in Edvinsson 2005) with a standard deviation of 0.03.

Starting from the tax statistics income we use the following method to get at our preferred "Reference total income":

[43] These are available from the aggregate taxation statistics *Taxering till inkomst- och förmögenhet* 1922-1988, for the years before we add shares based on the observations 1922, and after 1988 we add the corresponding figures in the new tax statistics.

Tax statistics income	(the aggregates from the same sources as the income statistics described above in A2 (sometimes corrected for wealth shares))
+ items not included in the tax base	(we make the assumption that all important sources of income including certain social security benefits are included in the tax base after 1974 (hence abstracting from child allowances, *allmänt barnbidrag,* and study grants, *studiebidrag,* which are tax free) and add aggregate government expenditures for unemployment benefits (*arbetslöshetsersättning*), payments for sickleave (*sjukpenning*) and payments for mothers (*moderskapsförsäkring*, which in 1974 was replaced by "parenthood insurance", *föräldrarförsäkring,* which was taxed) based on figures in the Statistical Yearbook of Sweden from 1948- (before that they are not listed but can be assumed to be a small share).
+ estimated income for "non-filers"	(in our preferred specification we take (reference population - tax filers) x (0,8 times the tax threshold). As an alternative specifications we use 0.25 times the average income of tax filers).

= "Preferred reference total" (starting from the tax statistics income)

Figure DEF shows the alternative specifications over the whole period as shares of GDP, as well as 0.63 times GDP. What we can say with some certainty is that the estimate of "Personal sector total income" is an over estimate of our preferred reference total. We can also say with some certainty that at least since 1974 the tax statistics income is relatively close to our preferred reference total since most people file taxes and everything we wish to include as income is included in the tax base. We can also note that in the period 1930-1990 our "Preferred reference total" calculated starting with the tax statistics income follows the estimated "personal sector total income very closely. In fact, taking 0.89 times the latter, yields numbers which follow the former with very small deviations.[44] We also note that for the early years (1903-1920) imputing 0.8 times the threshold (or 0.25 times average income) clearly yields over estimates of reference income. This is to be expected since when most individuals are below the threshold small changes in assumptions about their average income make a big difference and at this point in time the average income amongst tax payers was certainly much higher then later implying that imputing similar shares to non-filers as later means overestimating their income a lot.

Figure 1 in Finalreftotals

Given the behavior of these series we have chosen to use 0.89 times our estimated "personal sector total income" as our reference total for the period 1903-1942 and then (as tax statistics become yearly) our calculated reference total income starting with tax statistics income. As with the reference total population we have calculated top income shares using a number of alternatives as well. The resulting series can be found in Appendix C.

[44] The standard deviation is 0.02 and the maximum deviation is 0.05.

Appendix C: Sensitivity analysis: various adjustments, changing reference totals etc

Adjustments and comparisons to other studies of the period before 1950
Sensitivity of the results with respect to changing reference totals.

C1: Sensitivity analysis of choosing using different reference totals

Figure RST shows the top income shares when using all the alternative specifications of reference total population discussed in Appendix B1, while Figure UVX shows the top shares using the different reference totals for income.

Figure RST
Figure UVX

Again, we would like to stress that these are not all "alternatives" some series (like the reference population 1 is known to be incorrect but it creates a bound for the maximum deviation if the estimate was maximally incorrect.

C2: Sensitivity analysis of choosing either individuals or families as tax units

Our income series are computed from the tax returns-based income statistics for most years, and as we describe in Appendix A2 this implies that we use two different concepts of income earners over the twentieth century. Before 1951, the income earner in our data is the *family* (or household), i.e., married couples with, or without, children, single men 16 years and older, and single women 16 years or older. From 1951 onwards, our income earner is the *individual*, meaning all men and women 16 years or older. Hence, while we in the first period count married couples as one income earner, they are counted as two income earners in the latter period.

This section offers some partial explorations of how this switch of income earner concept may influence the overall results of our study. As our historical data were chosen largely due to availability constraints, we cannot make a fully-fledged comparison as there are simply no parallel datasets based on tax data available. What we can do, however, is to compare our family-based series with the series in which individuals are the basis. This can be done from the years from which we use the *Census material* (the years 1920, 1930, 1935 (partial census), 1945 (partial census), and 1950) when the primary material is individual-based but adjusted by us and others (especially Bentzel, 1952) to be consistent with the family-based series from the years before 1920 and in between the other years (1934 and 1937).

Figure C1 shows the income shares of the top fractiles (from top 10 percent to the top 0.05 percent). Solid lines represent our main family-based income series used in our analysis (called "Family") whereas the broken lines are the unadjusted, individual-based census series (called "Individual"). Note that since we use different concepts of income earners in the two cases, we must also use two different reference total populations to calculate the correct population shares. In our family-based series, we use the adult population 16 years and above minus married women and in the individual-based series the adult population 16 years and above is used. For this reason, the level of the shares may not be fully corresponding to each other although as Figure C1 shows they as a matter of fact are to quite some extent. As for the changes in shares over the period, they are pretty much coinciding in all cases for all fractiles, and importantly there is no systematic tendency in some direction of either series. For example, whereas the individual-based series produce slightly larger declines between 1935 and 1950 for the top 10 percent to top 0.5 percent income earners, the family-based series do it for

the top 0.1 to top 0.05 percent fractiles. Altogether, we feel confident with our choice of income earner concepts and have not found any systematic biases when contrasting them with alternative definitions.

C3: Age adjustments and effects of censoring the youngest income earners

Similar to the top income earners analyzed in previous studies for other countries, we use an age cutoff to remove young people from our data and thereby ensure that the series are conceptually consistent over the years. Specifically, we use an age cutoff of 16 years meaning that all income earners aged 16 and older are included in the series. We choose 16 years for several reasons. First, this was the lowest allowed age by Statistics Sweden in the sampling of income earners when constructing the income statistics 1951–1966 (see section A.3). Second, ever since the late 1970's this has been the lowest reported age level followed by 20 years. Third, 16 years has for long also been the first year after the Swedish compulsory secondary education ended. For robustness purposes we have also run the analysis excluding all income earners under 20 years of age and present some comparative results in other sections of this paper. As Atkinson (2003) and Atkinson and Leigh (2004) show, however, the effects on the results from varying the age cutoff a few years are rather modest.

In practice, our age cutoff means that we subtract the number of income earners age 15 or less from our reference total population and from the main top income series. It should be noted that the age data only includes the number of income earners in each age-income class without disclosing any income sums for each such group. This means that our age adjustments in practice is a subtraction of the number of income earners below 16 years in each income class while leaving the total income as it stands. As shown in the following discussion, however, this has no effect whatsoever on the overall results of our study since these very young income earners practically had no significant incomes at all.

Another data constraint is that before 1951 there exist no age data in the tax material (Statistics Sweden only collected data on income and occupation). This also should have no effect our study since before that year the Swedish income statistics was family-based and the tax units counted were mainly family fathers and single-living households (men and women). The ones below 16 years of age in these two groups should be quite few.

A more intriguing data problem is that the income statistics in some years report a higher lowest age group, 0–16 years during 1957–1966 and 0–17 years during 1971–1977. To get numbers for the (hence unobserved) 0–15 group we use interpolation based on the continuously observed 0–19 group and the relation between this group and the 0–15 group during the years just before and after the missing data periods. This minor bridging of the series seems to have no unwarranted effects.

The effect of removing the youngest income earners from the population of income earners seems to be insignificant on our analysis of top income earners. Figure C1 shows how the number and incomes of income earners under 16 and 20 years as share of the total population developed since 1951 (the first year for which age data were collected). The youngest group, 0–15 years, was clearly insignificant as income earners as their share of 0.1 percent of all incomes shows. Their share of the number of income earners, however, increased disproportionately in 1978 and 1992. In 1978, new tax collection routines required employers to submit income statements for all employees and in the same year the number of income earners below 16 years increased fivefold. In 1992, the drastic increase in the share of young income earners was related to new rules in the tax reform. More specifically, all capital income above

SEK 100 was made taxable and hence almost one million children, roughly one ninth of the entire Swedish population, became tax units overnight.[45] Notably, the number of income earners between 16 and 20 years old hardly increased at all (about 3 percent, compared to 571 percent for those 15 years old and younger). In other words, by excluding the youngest income earners we avoid some unwarranted heterogeneity in the income earner shares caused by the tax reform of 1990–1991.

Figure C1: Shares of tax population and total income of young income earners (below 16 and 20 years), 1951–2003.

(add section showing results with different ref totals (figure B1 and others)

Appendix D: Estimation theory

D1: Pareto interpolation: description and examples

All our income data come from tables showing the number of income earners and their summed incomes in classes of total income (see Table X for an example of these tables, showing the year 1945). Since the groups of income earners most often represent uneven shares of the population whereas we wish to analyze income shares for a given share of the population, e.g., the top 1%, we need to use interpolation techniques. This is a standard methodological approach used in all previous studies of top incomes, although the specifics of the methods may differ somewhat. One of the most commonly methods is to assume that the top incomes (when lined up for each individual in a row) are Pareto distributed (see, e.g., Feenberg and Poterba 1993, 2000; Piketty and Saez, 2003; Dell, 2005), and this is also the basic assumption used in this study. As described in, e.g., Atkinson (2003), if $H(y)$ is the cumulative proportion of income earners with income y or higher, then for some arbitrary constant A and model parameter α Pareto's formula states that

$$H(y) = \mathrm{A}\, y^{-\alpha}. \tag{1}$$

The cumulative share of total income for persons with income y or higher is

$$G(y) = [\alpha/(\alpha-1)]\ \mathrm{A}\, y^{-(\alpha-1)}. \tag{2}$$

Dividing equations (1) and (2) to get the average income for those income earners who have an income equal to y or higher, we get the following expression,

$$G(y)/H(y) = [\alpha/(\alpha-1)]\, y. \tag{3}$$

Now, in order to calculate income shares for specifically defined top shares of the population, we define the specific share i (say 0.01, or the top 1 percent) of the population $H_i(y)$ and that their corresponding share of total income is $G_i(y)$. We eliminate y in (1) and (2) by substitution and get $G_j(y)^{\alpha}$ as a function $H_j(y)^{\alpha-1}$ times a constant. The Pareto distribution is then used for interpolation between two population shares surrounding the specific share (say, top 1% or

[45] Formally, the new rules were in practice already in 1991 but in that year's income statistics Statistics Sweden made an adjustment to exclude the new bulk of very young income earners. They excluded all income earners below 18 years of age with labor income less than SEK 12,000 (Statistics Sweden, *Inkomst- och skattestatistik 1991,* Be 20 SM 9301, p. 9).

top 0.1%) for which we wish to compute the share of total income. Using these two points to achieve relative income shares $G_i(y)/G_j(y)$ and relative population shares $H_i(y)/H_j(y)$ we get our main expression for computing top income shares

$$G_i(y)/G_j(y) = [H_i(y)/H_j(y)]^{(\alpha-1)/\alpha}. \quad (4)$$

From this expression, we compute the value of α and then the income share of specific population shares.

As an example, we take the incomes in 1945 and compute shares of the top 1% of all income earners. According to that year's tabulated incomes, the classes representing shares of population surrounding the 1%-level are those earning SEK 15,000 or above, having a share of population of 1.53 percent and of total income of 12.38 percent, and those earning SEK 20,000 or above, being 0.83 percent of all income earners earning 8.86 percent of all incomes. Using equation (4) in logarithmic form, we get that $\alpha/(\alpha-1) = \ln(0.0153/0.0083)/\ln(0.1238/0.0886) = 1.8439$, and $\alpha = 2.185$. Then to compute the income share of the top 1%, again use equation (4) and set $H_i(y) = 0.01$, $H_j(y) = 0.0083$, $G_j(y) = 0.0886$, but leave $G_i(y)$ (the income share of the top 1%) undecided. Then using the value of α, we get that $G_i(y) = 0.0886 \cdot [0.01/0.0083]^{1/1.8439} = 0.0983$, or 9.83 percent, which is hence the income share of the top 1% income earners.

D2: Robustness issues of Pareto interpolation

The Pareto interpolation methodology does not come without problems, as many authors have shown (see, e.g., Atkinson 2003, 2005). In particular, the income distributions may in fact not be Pareto distributed, not even in the top levels where they are generally conceived to be so. However, robustness checks against micro-level data in Feenberg and Poterba (2000) and Piketty (2003) suggest that in at least for U.S. and French top incomes the Pareto approximation seems generally adequate. Moreover, Atkinson (2003) shows that the Pareto interpolations between two reference points ($G_i(y)$ and $G_j(y)$) may give different final shares depending on which of the two points one inserts in the formula (4) for computing the wanted income share. The standard practice is to use the information from the share being closest to the wanted share, and in our example above we used the numerical information associated with the top 0.86% to interpolate the income share for the top 1% since this was closer located than the top 1.53%. If we instead would have used the latter share, however, we would have gotten the exact same share for the top 1% (9.82 percent).

A more serious problem with using this method arises if the tabulated income intervals in the top are too low to admit interpolation of the top income shares. For example, whereas the share of income earners located in the highest (open-ended) income interval in 1945 was 0.0048 percent, which is sufficient to interpolate the income shares of the P99.95–100 fractile, the same group in 1976 represented 1.09 percent of the total population, which instead means that we need to *extrapolate* the income shares of most top fractiles. These problems arise if tax or statistics authorities do not recurrently update their tabulated nominal income brackets and if there is significant inflation in the economy. Although there have been no systematic quantifications of how severe the effects of extrapolations some comparisons made by Atkinson (2003) indicates that they may be considerable. When we make some simple checks on the Swedish data, however, the effects seem not so radical at all. We start by censoring the 1945 income data (same as above) at the income level SEK 20,000 (instead of the true highest level SEK 200,000), representing the top 0.83% instead of the 0.005%, and extrapolate the shares of the higher fractiles we get small deviations. For the top 0.5%, the share increases

from 6.73 percent to 6.75 percent, and for the top 0.05% (our highest fractile), the share increases from 1.84 percent to 1.93 percent of total income. Similarly, when comparing the years 1976 and 1977, when the highest income level was raised from SEK 120,000 (containing the top 1.09%) to SEK 1,000,000 (containing the top 0.002%) the income shares for the top fractiles change only marginally and in the direction as the overall decreasing trend would suggest.

To at least know the extent of extrapolations in our analysis of Swedish top incomes, Figure D1 depicts both the highest tabulated income boundary in the income statistics during 1903–2003 as well as the share of those income earners having this income or higher of all income earners. As the Figure shows, the need to extrapolate the share of top incomes has not been overly common throughout the period. For the top 1%, we only did it in 1976 when the top group represented 1.087 percent of the whole population. For the top 0.5%, we extrapolated the share six times, (1971–1975, 2000), for the top 0.1% 32 times and for the top 0.05% 44 times, out of a total of 74 observations. From inspecting the final top income series, we do not trace many unwarranted variations of shares that seem attributable to the extrapolations. Quite the contrary, in the 1970's when the potential extrapolation problems are the worst we have remarkably stable shares of top 0.1% and top 0.05% and in particular consistent with later years when the highest reported income level was raised and no extrapolation needed.

Figure D1: Highest reported income boundary (right scale) and the share of income earners in the highest income class of reference total population (left scale), 1903-2003

Table A.2: List of sources, 1903–2003.

Income year	Main source [a), b)]	Tables in source	Pages	Series [c)]
1903	Flodström (1906)		1, 3	
1907	Flodström (1909)		XI–XII	FU
1911	Flodström (1914)		11	FU
1912	Flodström (1915)		13º	FU
1916	Statistics Sweden, *Statistical Yearbook 1929*.		286–287	SOS
1919	Statistics Sweden, *Statistical Yearbook 1929*.		286–287	SOS
1920	Statistics Sweden, *Statistical Yearbook 1929*.		286–287	SOS
1930	Bentzel (1952)			
1934	SOU 1936:18	10	47	SOS
1935	Bentzel (1952)			
1937	Quensel (1944)			
1940	Bentzel (1952)			
1943	Skattetaxeringarna (1) ... taxeringsåret 1944	L	31*	SOS
1944	Skattetaxeringarna (1) ... taxeringsåret 1945	Q	43*	SOS
1945	Skattetaxeringarna (1) ... taxeringsåret 1946	P	42*	SOS
	Statistics Sweden (1951), *Census of 1945*	4	2–3	SOS
1946	Skattetaxeringarna (1) ... taxeringsåret 1947	R	47*	SOS
1947	Skattetaxeringarna (1) ... taxeringsåret 1948	V	51*	SOS
1948	Skattetaxeringarna (1) ... taxeringsåret 1949	Q	48*	SOS
1949	Skattetaxeringarna (2) ... taxeringsåret 1950	R	48*	SOS
1950	Skattetaxeringarna (2) ... taxeringsåret 1951	S	51*	SOS
1951	Skattetaxeringarna (2) ... taxeringsåret 1952	Å, 8	63*, 26–27	SOS
	Statistics Sweden (1956), *Census of 1950*	7	20–21	SOS
1952	Skattetaxeringarna (2) ... taxeringsåret 1953	Z, 8	53º, 26–27	SOS
1953	Skattetaxeringarna (2) ... taxeringsåret 1954	Z, 8	49º, 26–27	SOS
1954	Skattetaxeringarna (2) ... taxeringsåret 1955	Z, 8	47º, 26–27	SOS
1955	Skattetaxeringarna (2) ... taxeringsåret 1956	Z, 8	46º, 28–29	SOS
1956	Skattetaxeringarna (2) ... taxeringsåret 1957	Z, 8	47º, 28–29	SOS
1957	Skattetaxeringarna (2) ... taxeringsåret 1958	Y, 8	47º, 28–29	SOS
1958	Skattetaxeringarna (2) ... taxeringsåret 1959	Å, 8	50º, 34–35	SOS
1959	Skattetaxeringarna (2) ... taxeringsåret 1960	J, 8	28º, 32–33	SOS
1960	Skattetaxeringarna (2) ... taxeringsåret 1961	I, 10	28º, 32–33	SOS
1961	Skattetaxeringarna (3) ... taxeringsåret 1962	I, 10	28º, 34–35	SOS
1962	Skattetaxeringarna (3) ... taxeringsåret 1963	J, 10	29º, 34–35	SOS
1963	Skattetaxeringarna (3) ... taxeringsåret 1964	J, 10	43º, 36–37	SOS
1964	Skattetaxeringarna (3) ... taxeringsåret 1965	K, 10	44º, 36–37	SOS
1965	Skattetaxeringarna (3) ... taxeringsåret 1966	J, 10	43º, 116–117	SOS
1966	Skattetaxeringarna (3) ... taxeringsåret 1967	L, 9	43º, 118–119	SOS
1967	Inkomst och förmögenhet 1967	2, 7	44–45, 58–61	SOS
1968	Inkomst och förmögenhet 1968	2, 7	50–51, 64–67	SOS
1969	Inkomst och förmögenhet 1969	2, 7	50–51, 64–67	SOS
1970	Inkomst och förmögenhet 1970	2, 7	48–49, 62–65	SOS
1971	Inkomst och förmögenhet 1971	3, 12	68–69, 90–93	SOS
1972	Inkomst och förmögenhet 1972	1, 3, 14	54–55, 70–71, 102–105	SOS
	Inkomst- och förmögenhetsfördelningen år 1972	7	19	SM N 1973:94
1973	Inkomst och förmögenhet 1973	3, 14	68–69, 100–103	SOS
1974	Inkomst- och förmögenhetsfördelningen år 1974	1, 7	11, 33	SM N 1976:4
1975	Inkomst- och förmögenhetsfördelningen år 1975	1, 7	13, 35	SM N 1976:23

1976	Inkomst- och förmögenhetsfördelningen år 1976	1, 7	18, 41, 43	SM N 1977:24
1977	Inkomst- och förmögenhetsfördelningen år 1977	1, 7	22, 46–47	SM N 1978:22
1978	Inkomst- och förmögenhetsfördelningen år 1978	1, 4.1, 4.2	29, 38, 41	SM N 1980:9
1979	Inkomst- och förmögenhetsfördelningen år 1979	1, 4.1, 4.2	20, 27, 30	SM N 1981:9.1
1980	Inkomst- och förmögenhetsfördelningen år 1980	1, 4.1, 4.2	7, 14, 17	SM N 1976:4
1981	Inkomst- och förmögenhetsfördelningen år 1981	1, 4.1, 4.2	7, 14, 17	SM N 1976:4
1982	Inkomst- och förmögenhetsfördelningen 1982	1, 4.1, 4.2	14, 21, 24	SM Be 1984:6.1
1983	Inkomst- och förmögenhetsfördelningen 1983	1, 4.1, 4.2	14, 21, 24	Be 20 SM 8501
1984	Inkomst- och förmögenhetsfördelningen 1984	1, 3.1, 3.2	15, 19, 22	Be 20 SM 8601
1985	Inkomst- och förmögenhetsfördelningen 1985	1, 2.1, 2.2	15, 18, 21	Be 20 SM 8701
1986	Inkomst- och förmögenhetsfördelningen 1986	1, 2.1, 2.2	17, 20, 23	Be 20 SM 8801
1987	Inkomst- och förmögenhetsfördelningen 1987	1, 2.1, 2.2	17, 20, 23	Be 20 SM 8901
1988	Inkomst- och skattestatistik 1988	1, 2.1, 2.2	16, 19, 22	Be 20 SM 9001
1989	Inkomst- och skattestatistik 1989	1, 2.1, 2.2	16, 20, 23	Be 20 SM 9101
1990	Inkomst- och skattestatistik 1990	1, 2.1, 2.2	15, 20, 23	Be 20 SM 9201
1991	Inkomst- och skattestatistik 1991	1, 4.1, 4.2, 6.1, 6.2	15, 22–23, 28–29	Be 20 SM 9301
1992	Inkomst- och skattestatistik 1992	1, 4.1, 4.2, 6.1, 6.2	14, 21–22, 27–28	Be 20 SM 9401
1993	Inkomst- och skattestatistik 1993	1, 4.1, 4.2, 6.1, 6.2	14, 21–22, 27–28	Be 20 SM 9501
1994	Inkomst- och skattestatistik 1994	1, 4.1, 4.2, 6.1, 6.2	15, 22–23, 28–29	Be 20 SM 9601
1995	Inkomster, skatter och bidrag 1995	1, 4.1, 4.2, 6.1, 6.2	19, 26, 28, 36, 38	Be 20 SM 9701
1996	Inkomster, skatter och bidrag 1996	1, 4.1, 4.2, 6.1, 6.2	20, 27, 29, 37, 39	If 20 SM 9801
1997	Inkomster, skatter och bidrag 1997	1, 4.1, 4.2, 6.1, 6.2	20, 29, 33, 45, 49	If 20 SM 9901
1998–2003	–			[d]

a) Some publications titles are abbreviated. Skattetaxeringarna (1) = *Skattetaxeringarna samt inkomstfördelningen inom yrkesgrupper*; Skattetaxeringarna (2) = *Skattetaxeringarna samt fördelningen av inkomst och förmögenhet inom yrkesgrupper*; Skattetaxeringarna (3) = *Skattetaxeringarna samt fördelningen av inkomst och förmögenhet taxeringsåret*.

b) The publications since 1982 also have the subtitle *Totalräknad statistik*.

c) "FU" denotes *Finansstatistiska utredningar* (Fiscal Surveys) and "SOS" *Sveriges officiella statistik* (Swedish Official Statistics).

d) During 1998–2003, Statistics Sweden stopped publishing income data in the form of earlier years. Instead we had to purchase data for these years directly from Statistics Sweden.

Table A.1: Example of how tax data appears in Swedish income statistics: Income year 1945.

Income earners earning at least SEK...	*Number of income earners*	*Sum of incomes (SEK)*
600	2,955,890	10,758,167,000
1,000	2,666,978	10,524,967,000
1,500	2,264,436	10,021,845,000
2,000	1,917,702	9,417,016,000
2,500	1,609,106	8,724,037,000
3,000	1,327,728	7,951,446,000
3,500	1,076,653	7,136,837,000
4,000	861,737	6,332,420,000
4,500	683,508	5,576,476,000
5,000	541,999	4,905,747,000
6,000	351,915	3,870,010,000
7,000	244,620	3,177,222,000
8,000	178,803	2,686,052,000
10,000	108,483	2,061,673,000
12,000	74,035	1,686,006,000
15,000	47,127	1,327,306,000
20,000	25,614	958,934,000
30,000	10,901	606,652,000
50,000	3,771	339,106,000
100,000	682	89,564,000
200,000	174	55,913,000
Sum	2,955,890	10,758,167,000

Source: See Table XX.

Figure 1

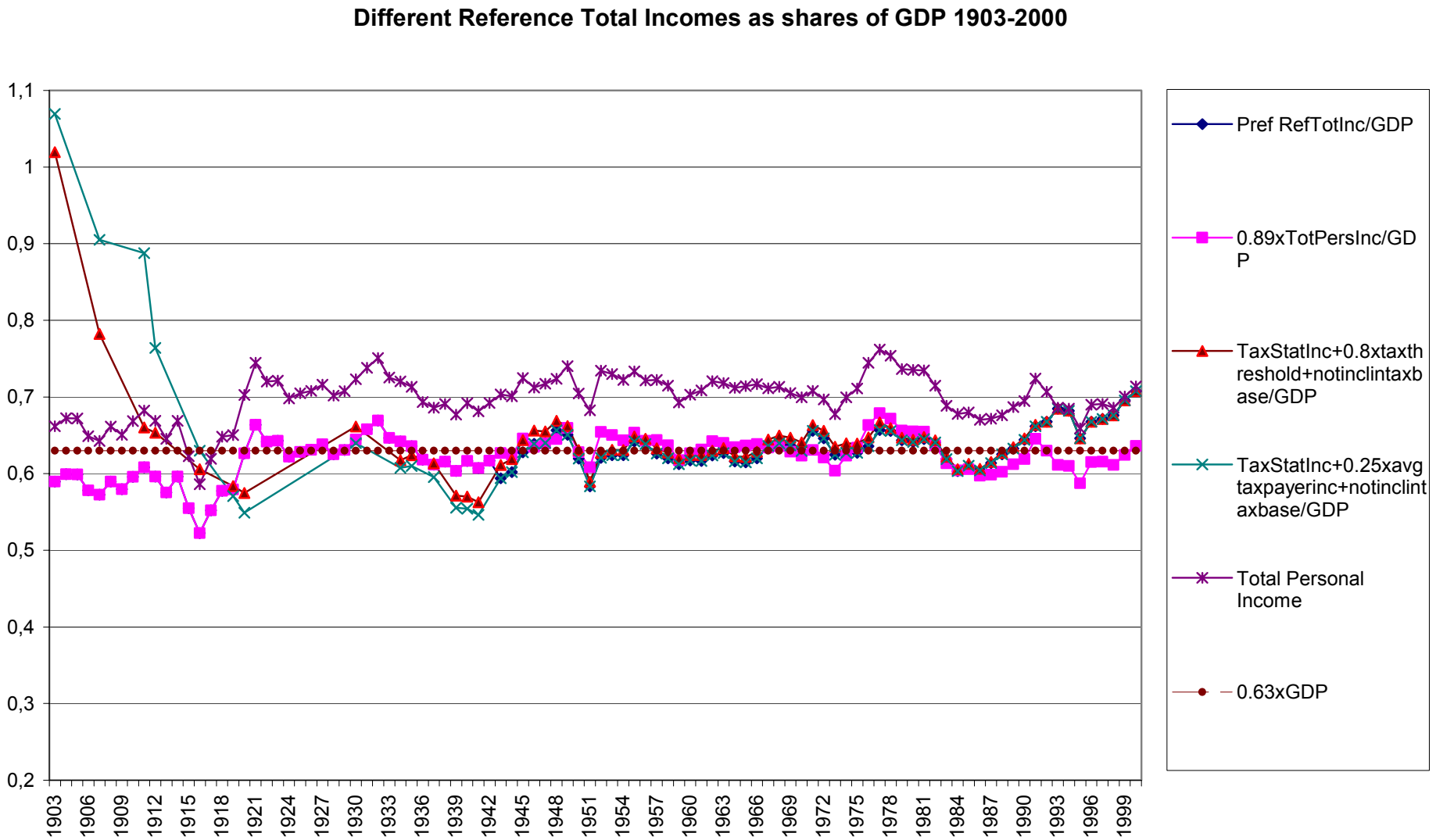

Different Reference Total Incomes as shares of GDP 1903-2000
1,1
1
0,9
0,8
0,7
0,6
0,5
0,4
0,3
0,2
1903
1906
1909
1912
1915
1918
1921
1924
1927
1930
1933
1936
1939
1942
1945
1948
1951
1954
1957
1960
1963
1966
1969
1972
1975
1978
1981
1984
1987
1990
1993
1996
1999
Pref RefTotInc/GDP
0.89xTotPersInc/GDP
TaxStatInc+0.8xtaxthreshold+notinclintaxbase/GDP
TaxStatInc+0.25xavgtaxpayerinc+notinclintaxbase/GDP
Total Personal Income
0.63xGDP

Figure 2

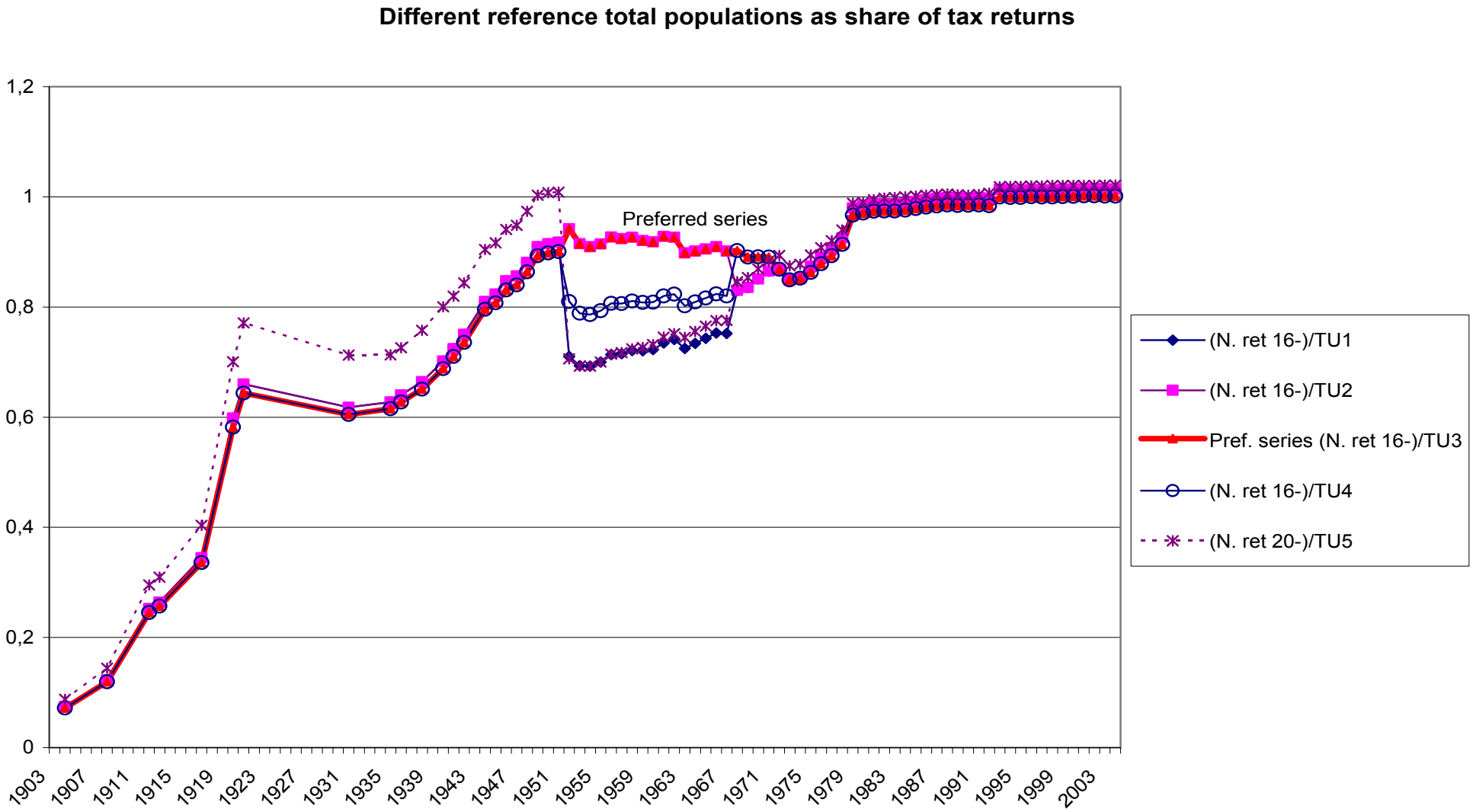

Different reference total populations as share of tax returns
1,2
1
0,8
0,6
0,4
0,2
0
Preferred series
1903
1907
1911
1915
1919
1923
1927
1931
1935
1939
1943
1947
1951
1955
1959
1963
1967
1971
1975
1979
1983
1987
1991
1995
1999
2003
(N. ret 16-)/TU1
(N. ret 16-)/TU2
Pref. series (N. ret 16-)/TU3
(N. ret 16-)/TU4
(N. ret 20-)/TU5

Figure 3

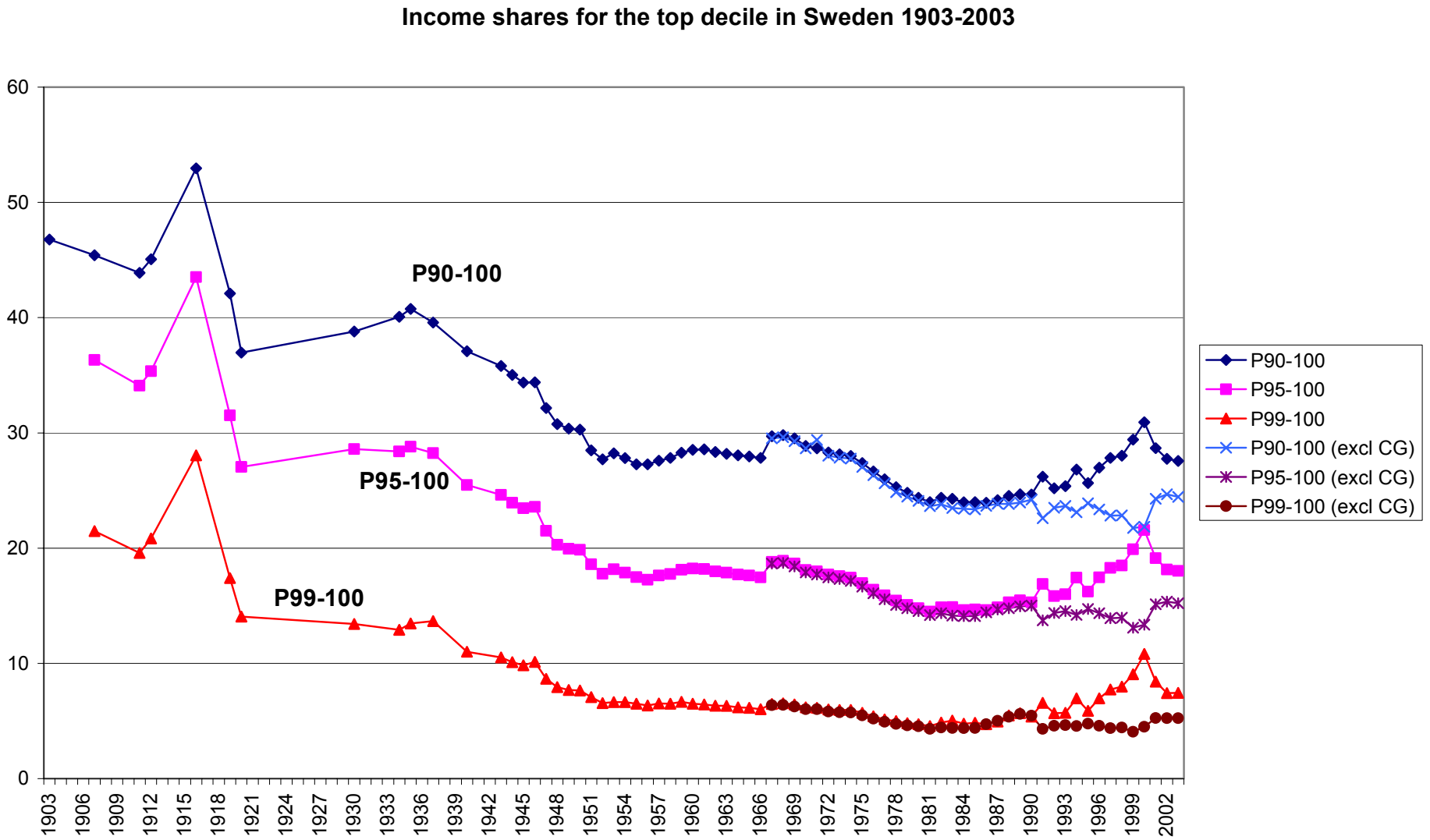

Income shares for the top decile in Sweden 1903-2003
60
50
40
30
20
10
0
1903
1906
1909
1912
1915
1918
1921
1924
1927
1930
1933
1936
1939
1942
1945
1948
1951
1954
1957
1960
1963
1966
1969
1972
1975
1978
1981
1984
1987
1990
1993
1996
1999
2002
P90-100
P95-100
P99-100
P90-100
P95-100
P99-100
P90-100 (excl CG)
P95-100 (excl CG)
P99-100 (excl CG)

Figure 4

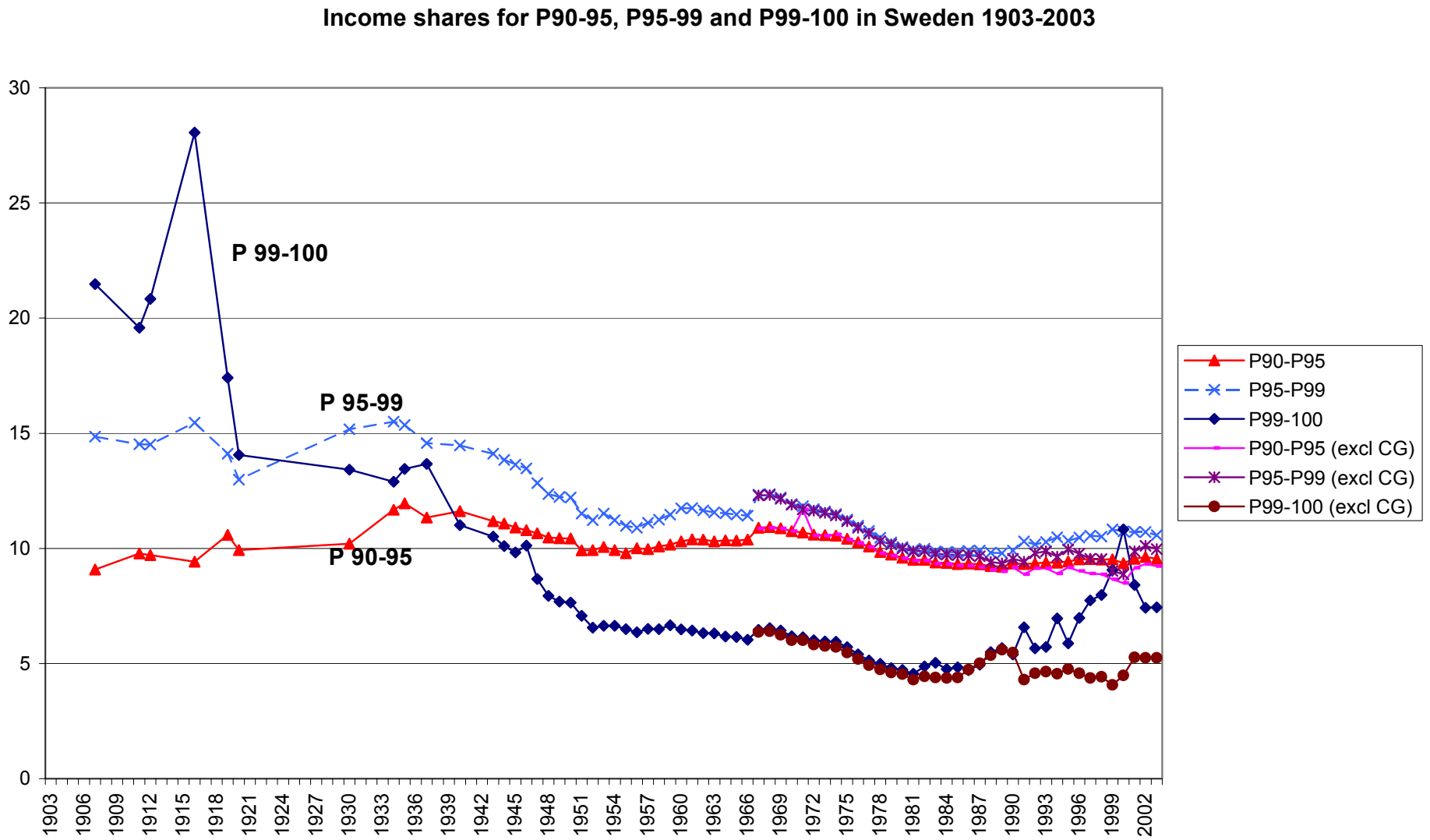

Income shares for P90-95, P95-99 and P99-100 in Sweden 1903-2003
30
25
20
15
10
5
0
P 99-100
P 95-99
P 90-95
1903
1906
1909
1912
1915
1918
1921
1924
1927
1930
1933
1936
1939
1942
1945
1948
1951
1954
1957
1960
1963
1966
1969
1972
1975
1978
1981
1984
1987
1990
1993
1996
1999
2002
P90-P95
P95-P99
P99-100
P90-P95 (excl CG)
P95-P99 (excl CG)
P99-100 (excl CG)

Figure 5

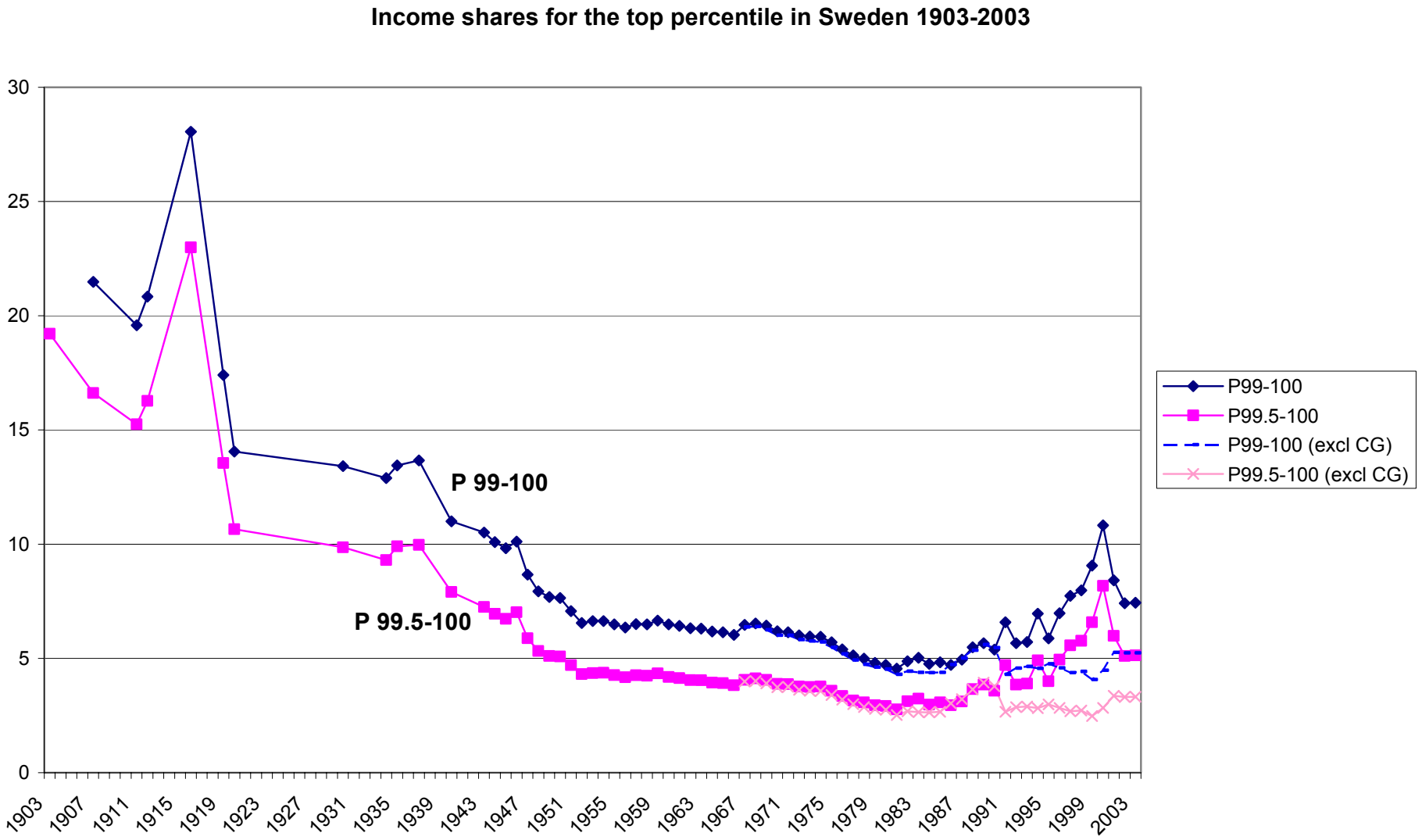

Income shares for the top percentile in Sweden 1903-2003
30
25
20
15
10
5
0
1903
1907
1911
1915
1919
1923
1927
1931
1935
1939
1943
1947
1951
1955
1959
1963
1967
1971
1975
1979
1983
1987
1991
1995
1999
2003
P 99-100
P 99.5-100
P99-100
P99.5-100
P99-100 (excl CG)
P99.5-100 (excl CG)

Figure 6

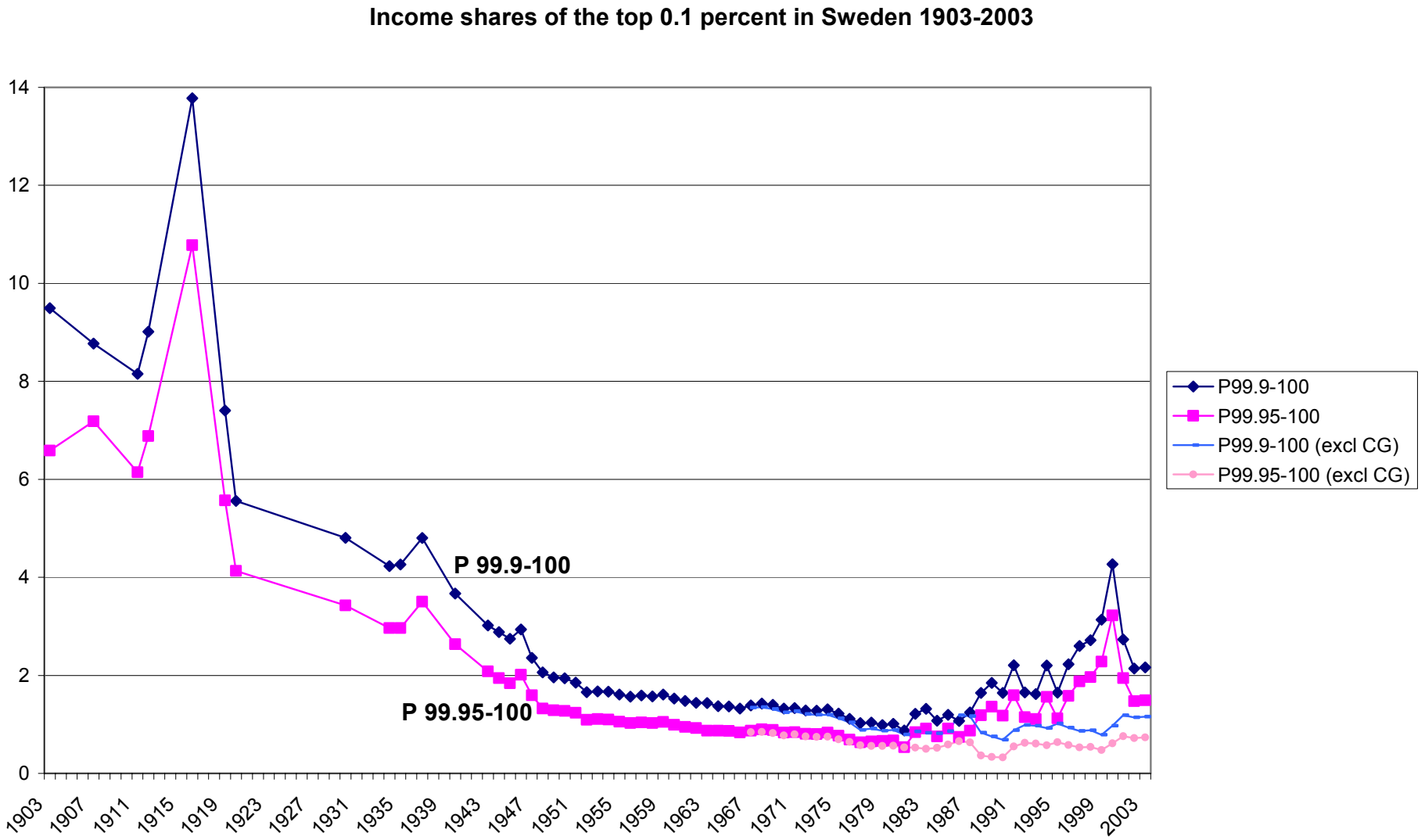

Income shares of the top 0.1 percent in Sweden 1903-2003
14
12
10
8
6
4
2
0
1903
1907
1911
1915
1919
1923
1927
1931
1935
1939
1943
1947
1951
1955
1959
1963
1967
1971
1975
1979
1983
1987
1991
1995
1999
2003
P 99.9-100
P 99.95-100
P99.9-100
P99.95-100
P99.9-100 (excl CG)
P99.95-100 (excl CG)

Figure 7

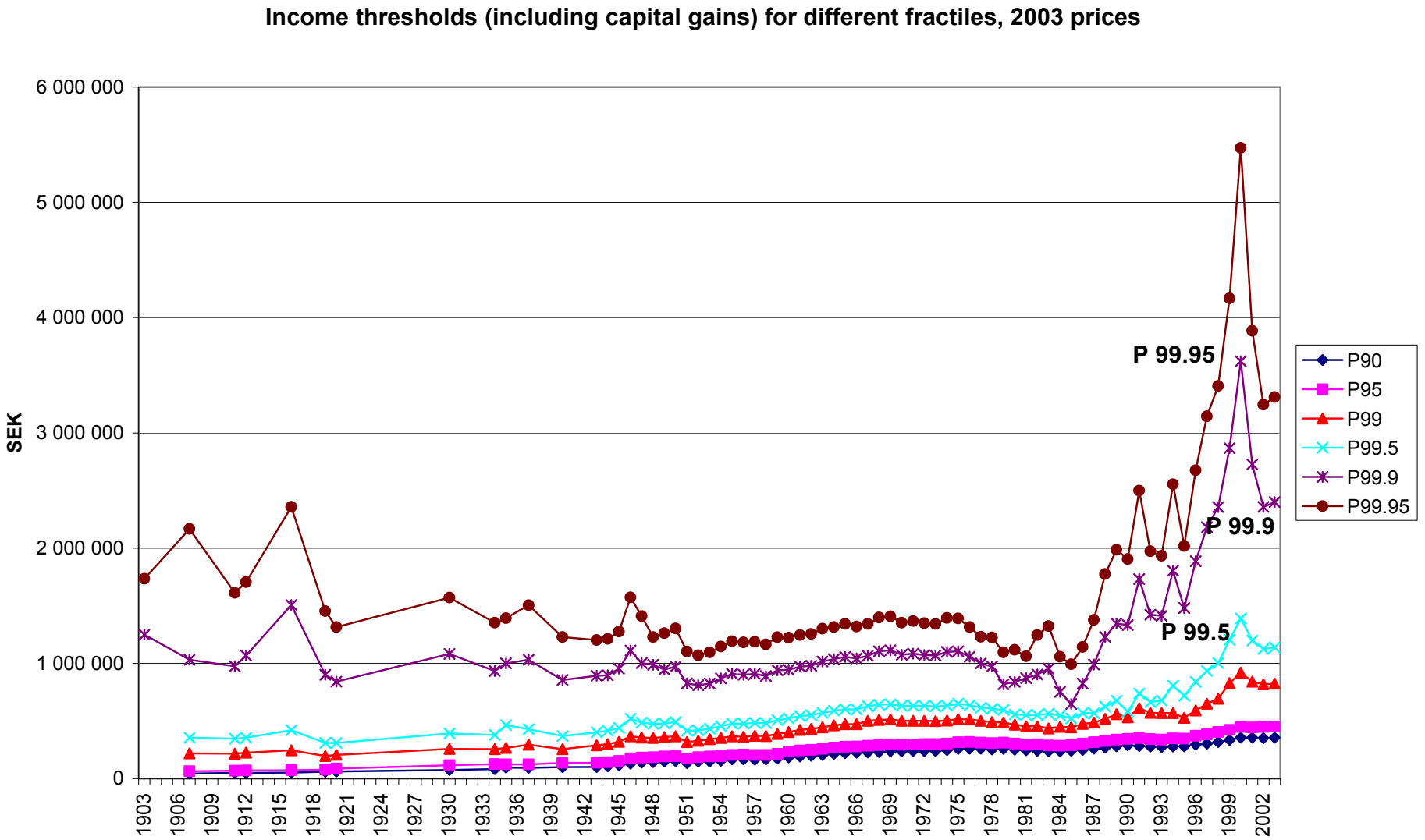

Income thresholds (including capital gains) for different fractiles, 2003 prices
SEK
6 000 000
5 000 000
4 000 000
3 000 000
2 000 000
1 000 000
0
1903
1906
1909
1912
1915
1918
1921
1924
1927
1930
1933
1936
1939
1942
1945
1948
1951
1954
1957
1960
1963
1966
1969
1972
1975
1978
1981
1984
1987
1990
1993
1996
1999
2002
P 99.95
P 99.9
P 99.5
P90
P95
P99
P99.5
P99.9
P99.95

Figure 8

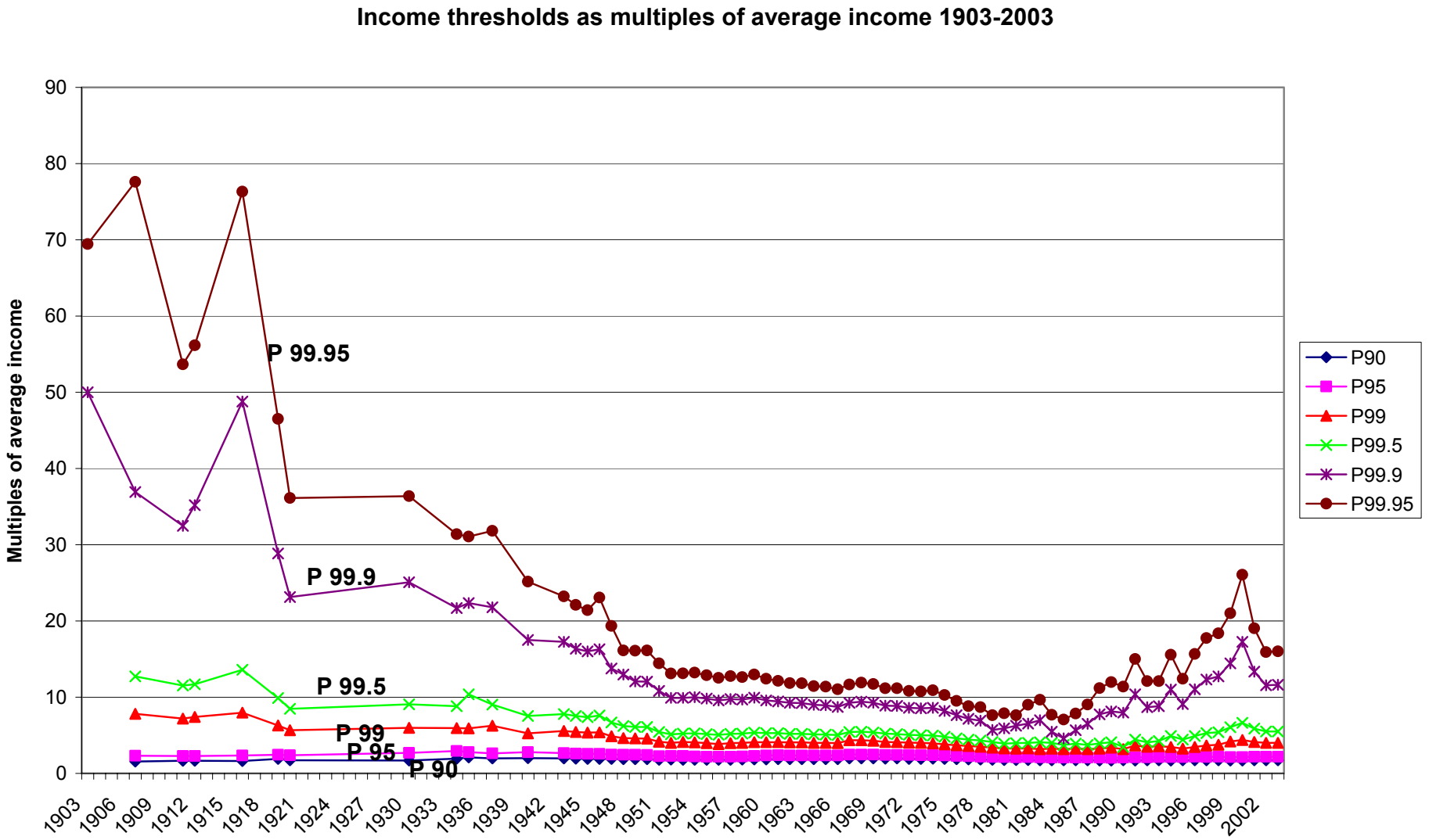

Income thresholds as multiples of average income 1903-2003
Multiples of average income
90
80
70
60
50
40
30
20
10
0
P 99.95
P 99.9
P 99.5
P 99
P 95
P 90
1903
1906
1909
1912
1915
1918
1921
1924
1927
1930
1933
1936
1939
1942
1945
1948
1951
1954
1957
1960
1963
1966
1969
1972
1975
1978
1981
1984
1987
1990
1993
1996
1999
2002
P90
P95
P99
P99.5
P99.9
P99.95

Figure 9: Income composition 1945 (solid: including capital gains, broken: excluding capital gains)

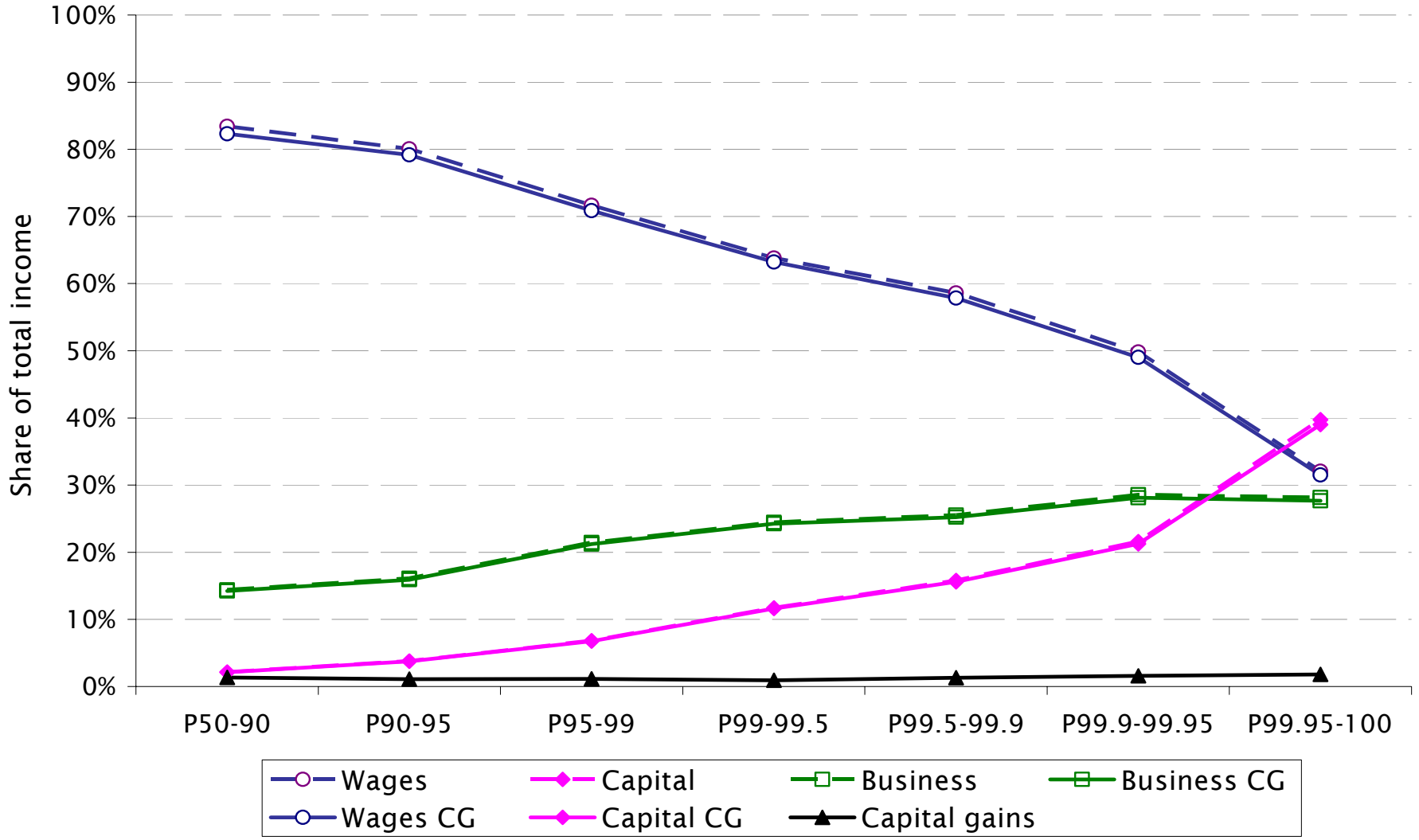


Figure 10: Income composition 1978 (solid: including capital gains, broken: excluding capital gains)

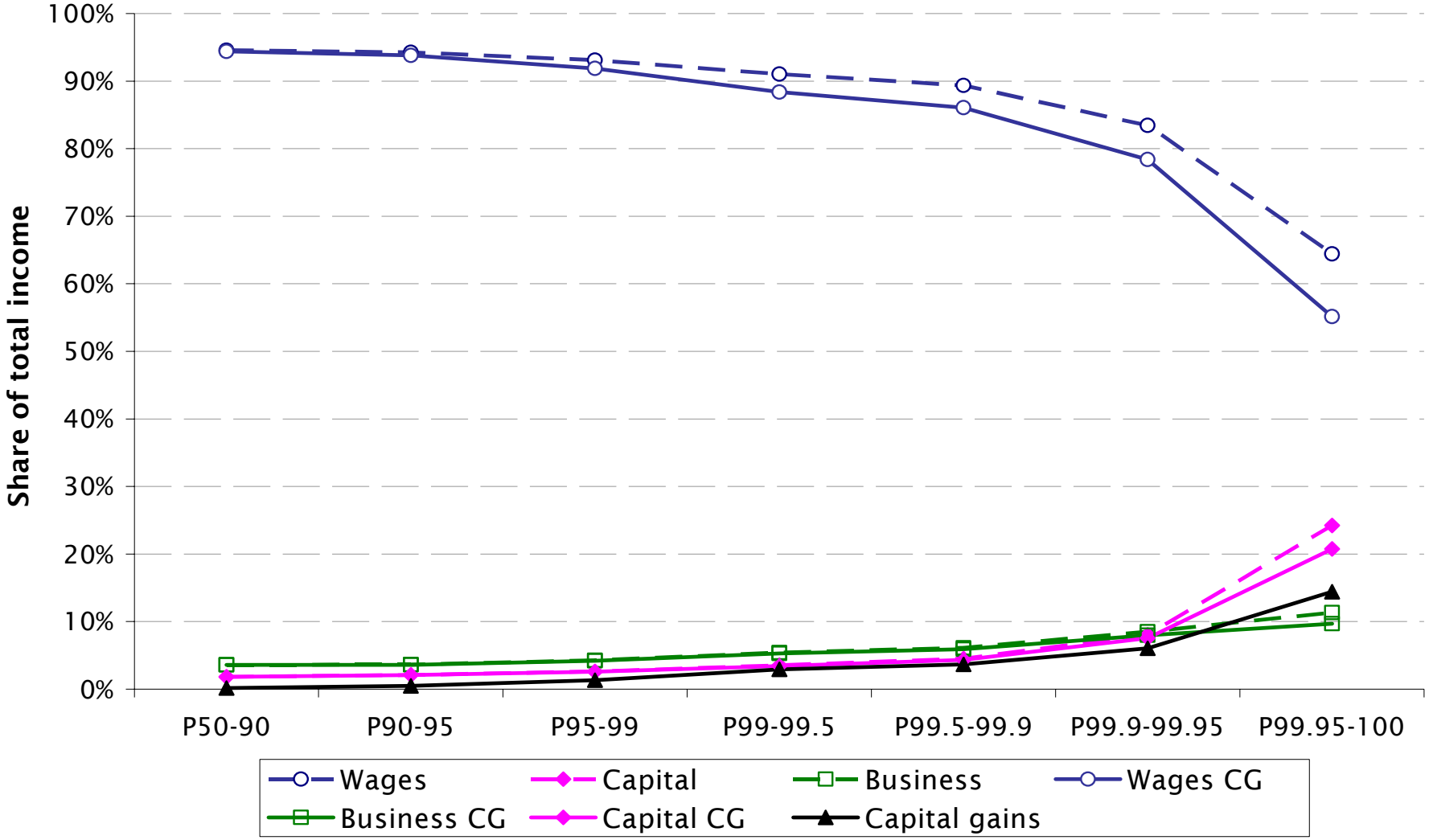


Figure 11: Income composition 1997 (solid: including capital gains, broken: excluding capital gains)

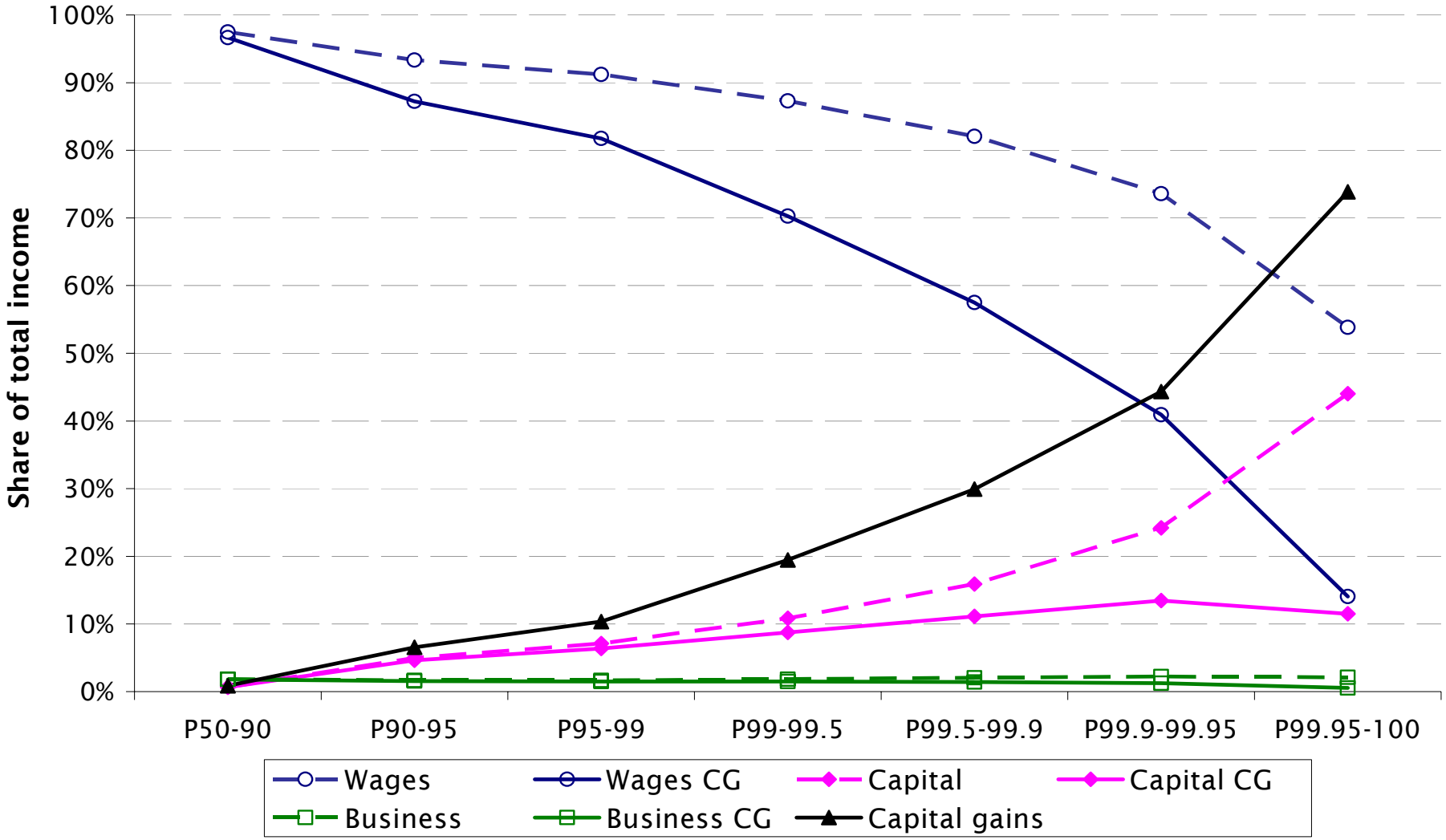
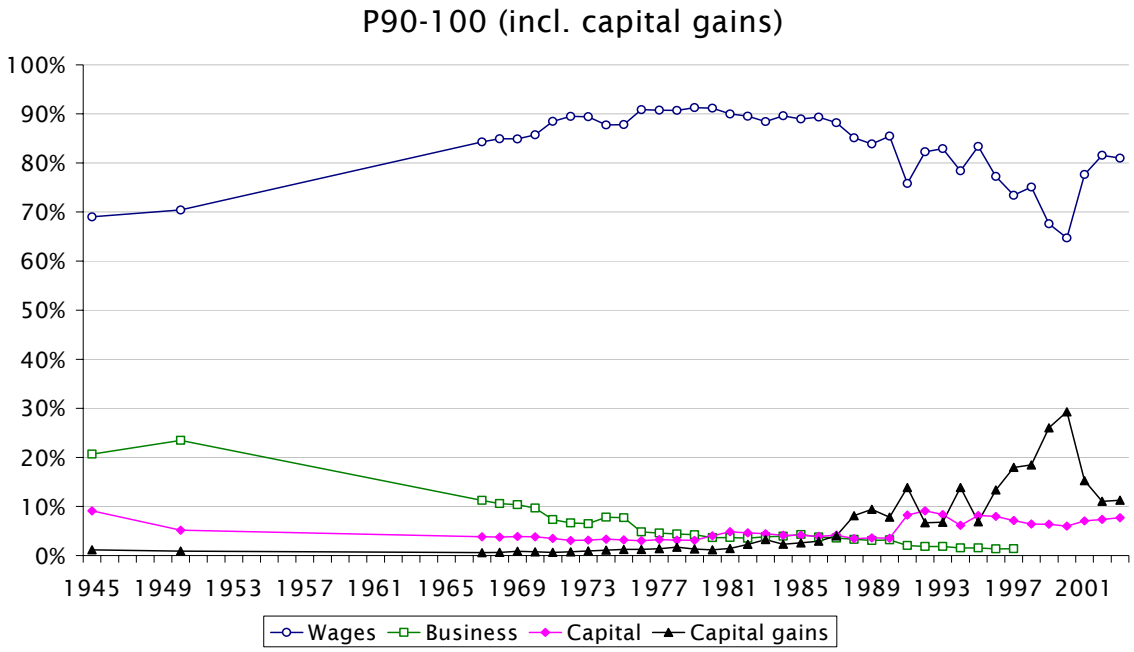


Figure 12a-b: Long-run income composition of the top 10 percent, 1945–2003



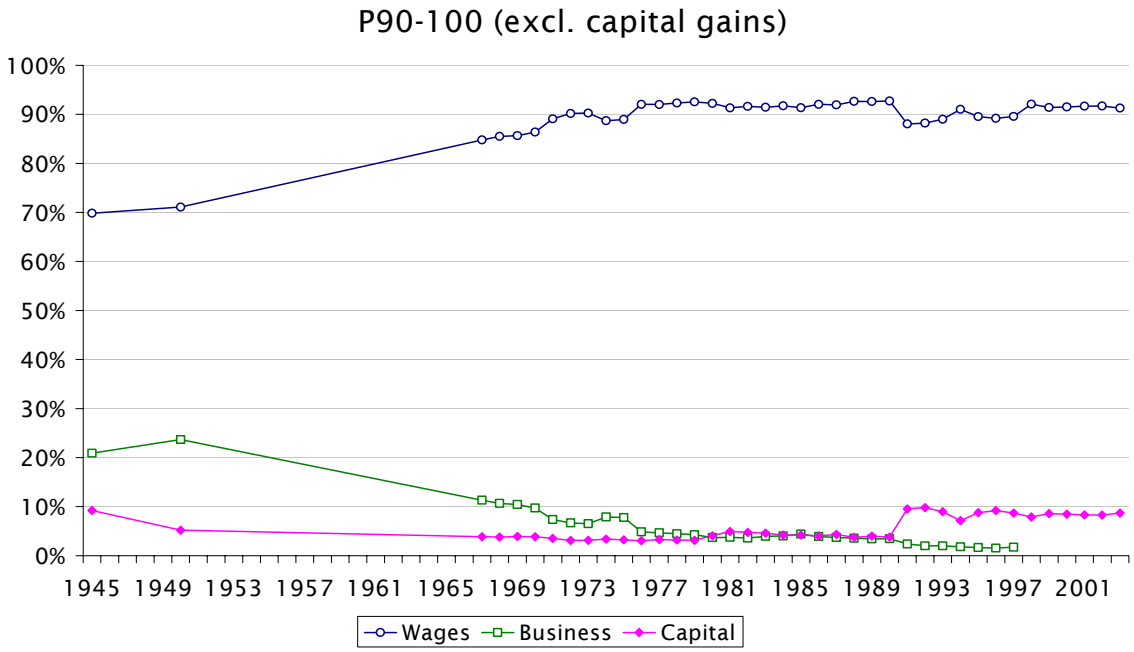
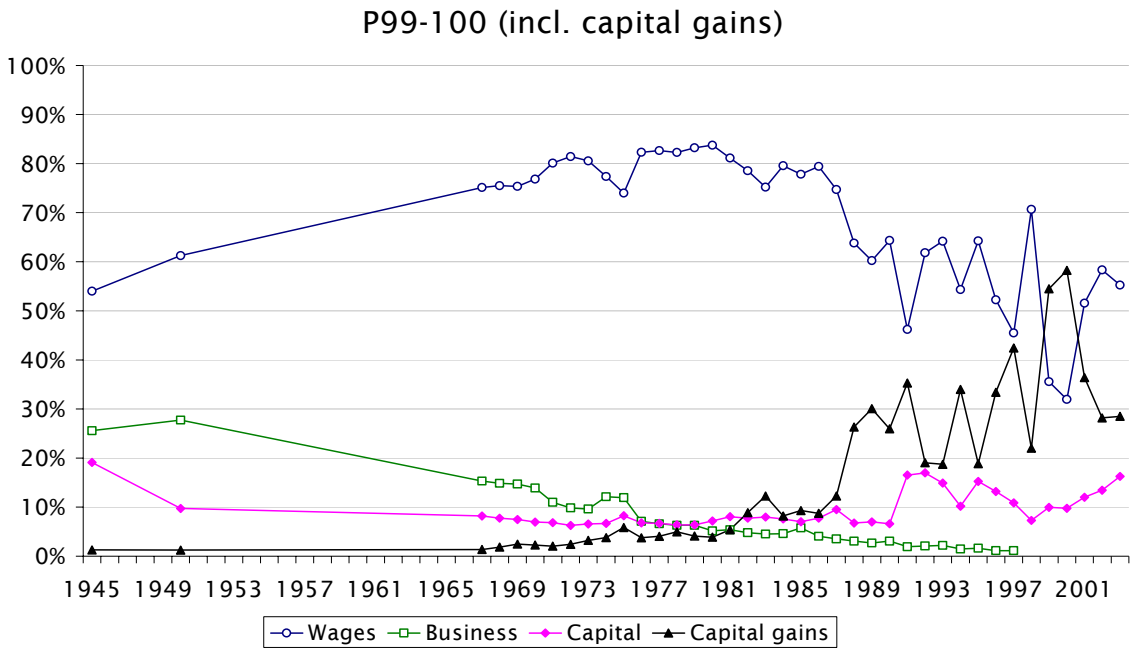


Figure 13a-b: Long-run income composition of the top 1 percent, 1945–2003



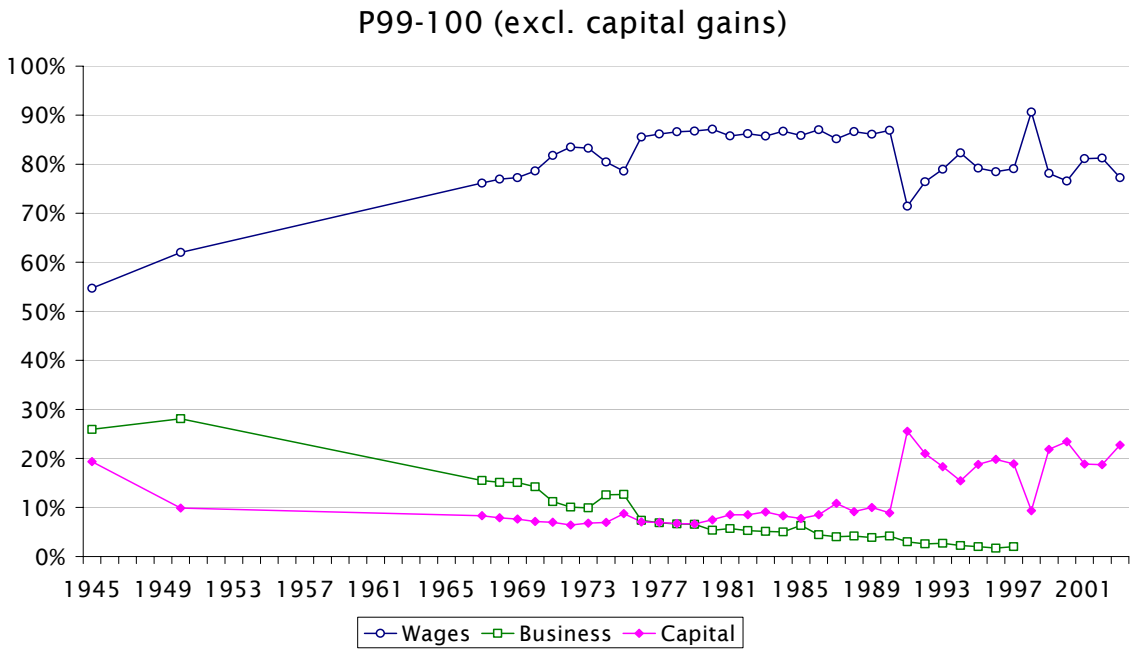
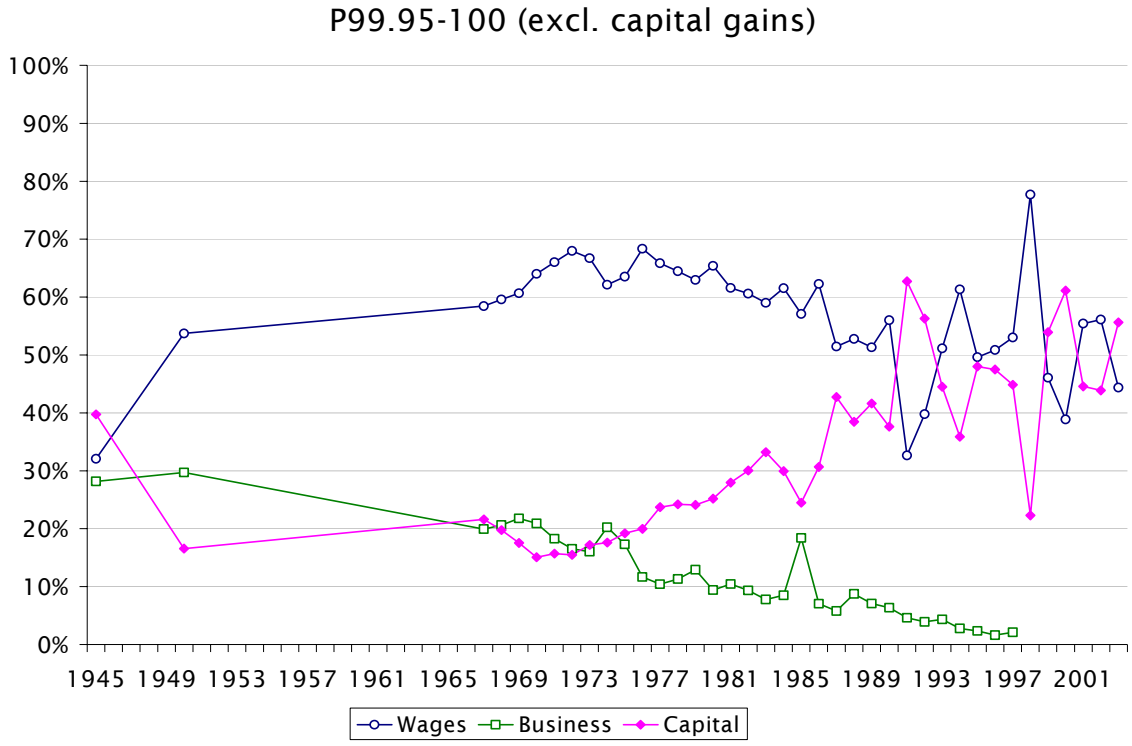


Figure 14a-b: Long-run income composition of the top 0.05 percent, 1945–2003



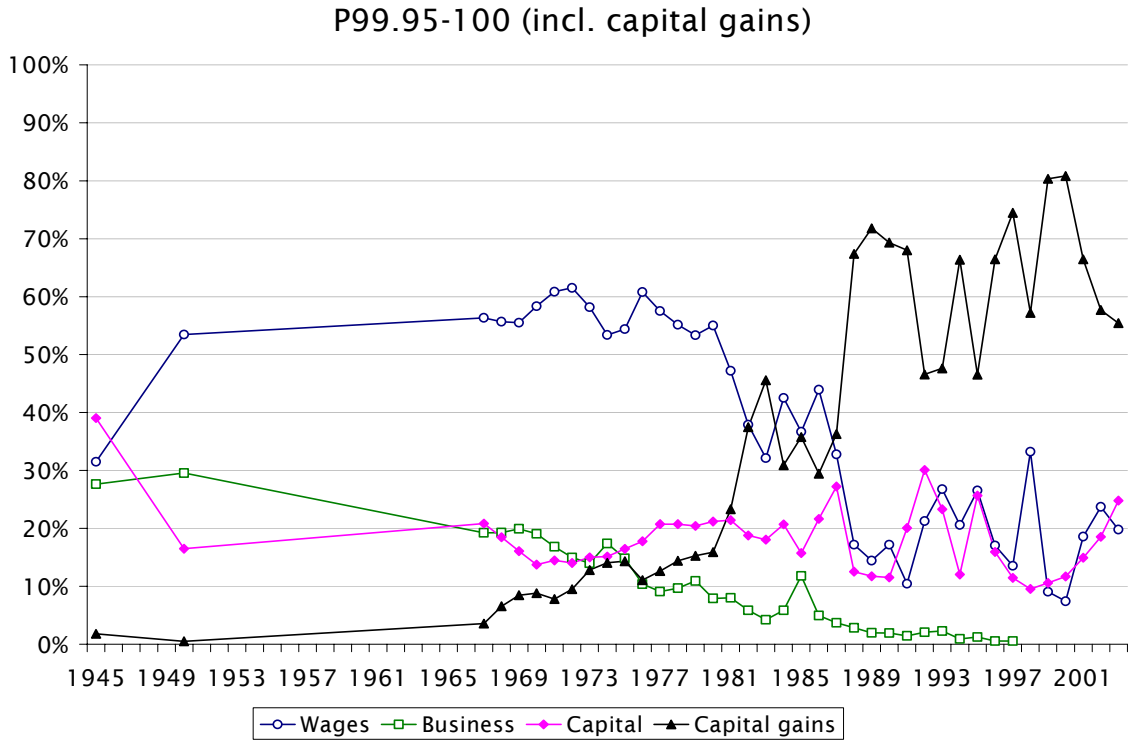


Figure 15:

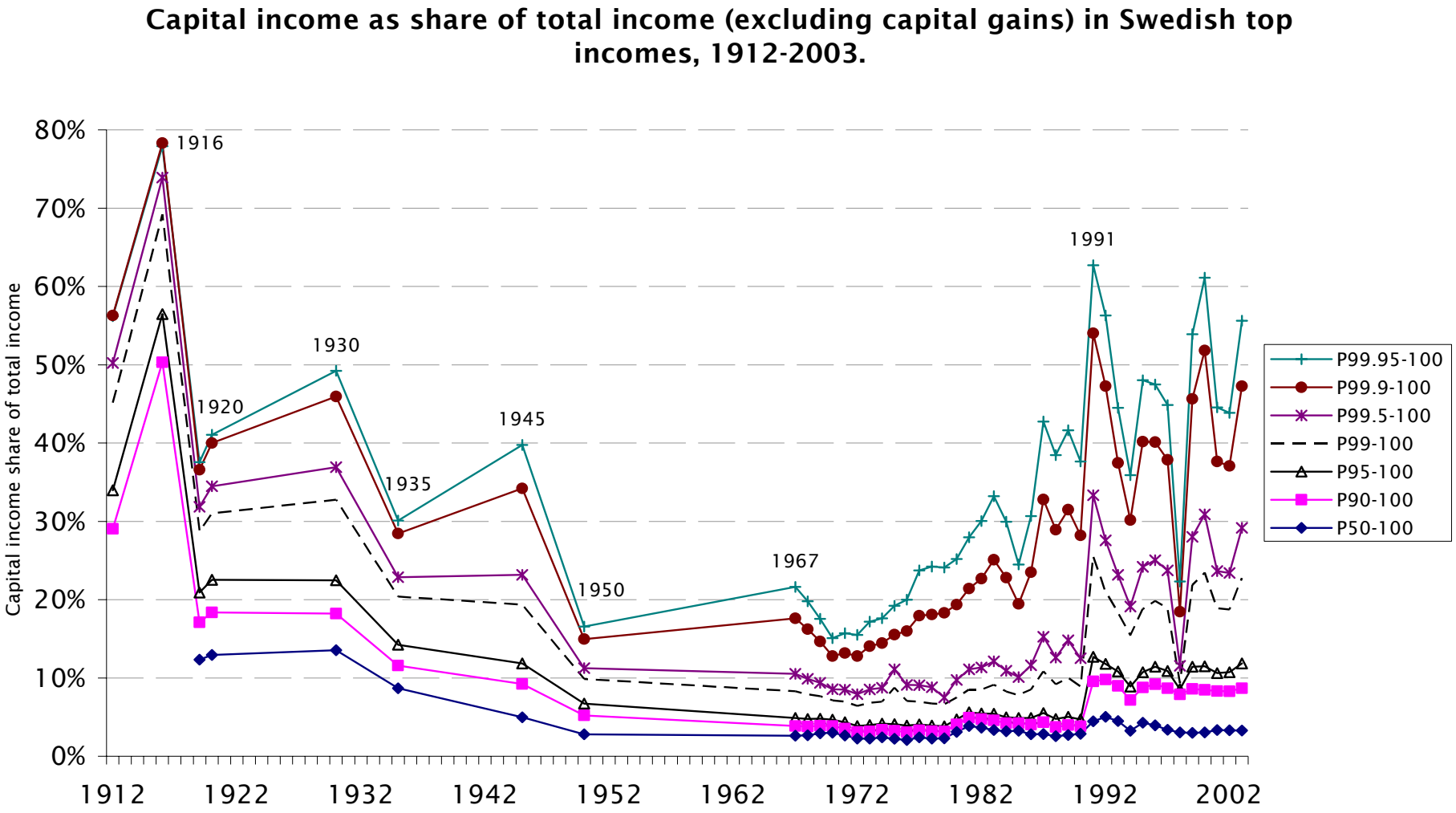

Capital income as share of total income (excluding capital gains) in Swedish top incomes, 1912-2003.
Capital income share of total income
80%
70%
60%
50%
40%
30%
20%
10%
0%
1912
1922
1932
1942
1952
1962
1972
1982
1992
2002
1916
1920
1930
1935
1945
1950
1967
1991
P99.95-100
P99.9-100
P99.5-100
P99-100
P95-100
P90-100
P50-100

Figure 16:

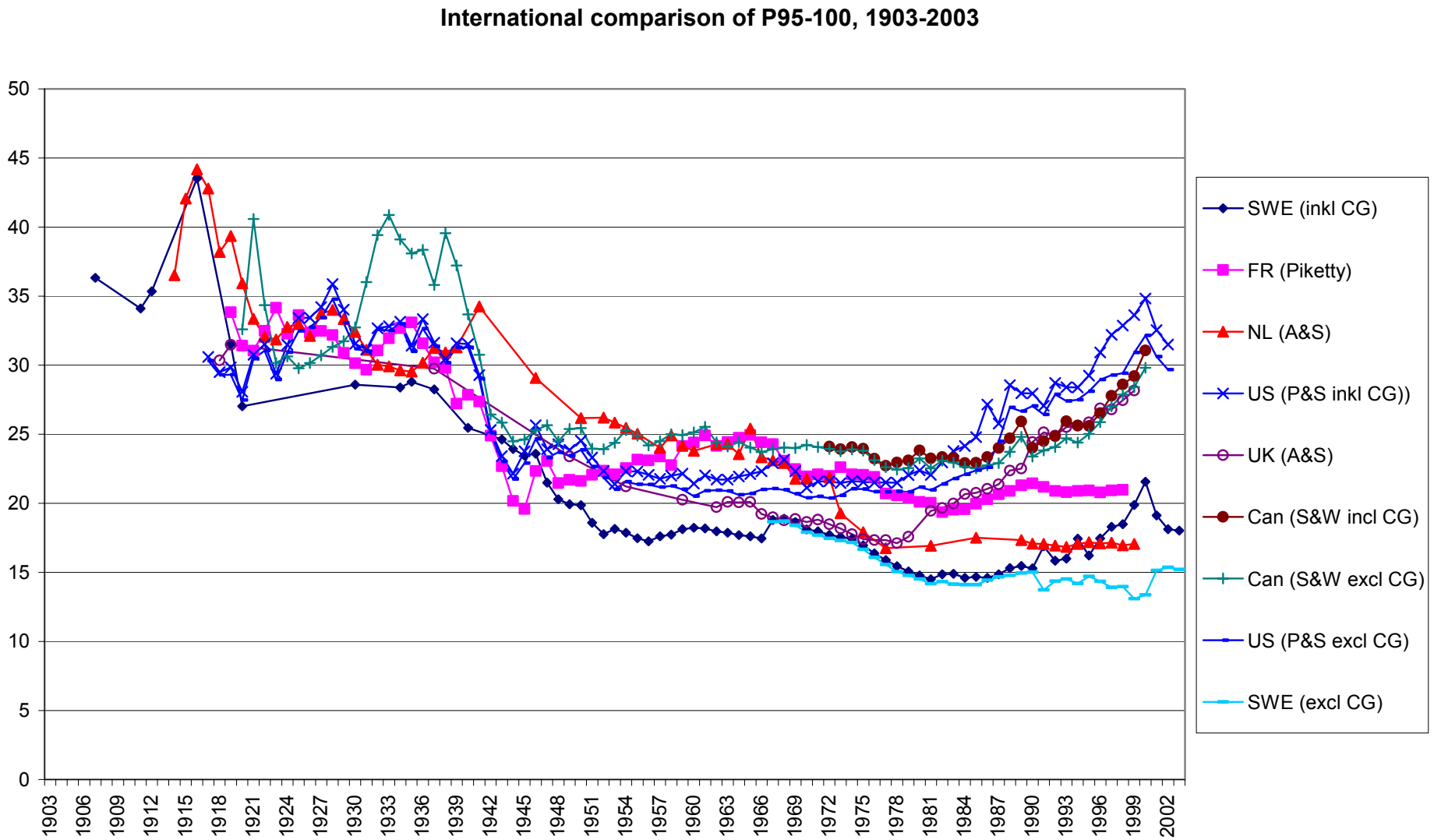

International comparison of P95-100, 1903-2003
50
45
40
35
30
25
20
15
10
5
0
1903
1906
1909
1912
1915
1918
1921
1924
1927
1930
1933
1936
1939
1942
1945
1948
1951
1954
1957
1960
1963
1966
1969
1972
1975
1978
1981
1984
1987
1990
1993
1996
1999
2002
SWE (inkl CG)
FR (Piketty)
NL (A&S)
US (P&S inkl CG))
UK (A&S)
Can (S&W incl CG)
Can (S&W excl CG)
US (P&S excl CG)
SWE (excl CG)

Figure 17:

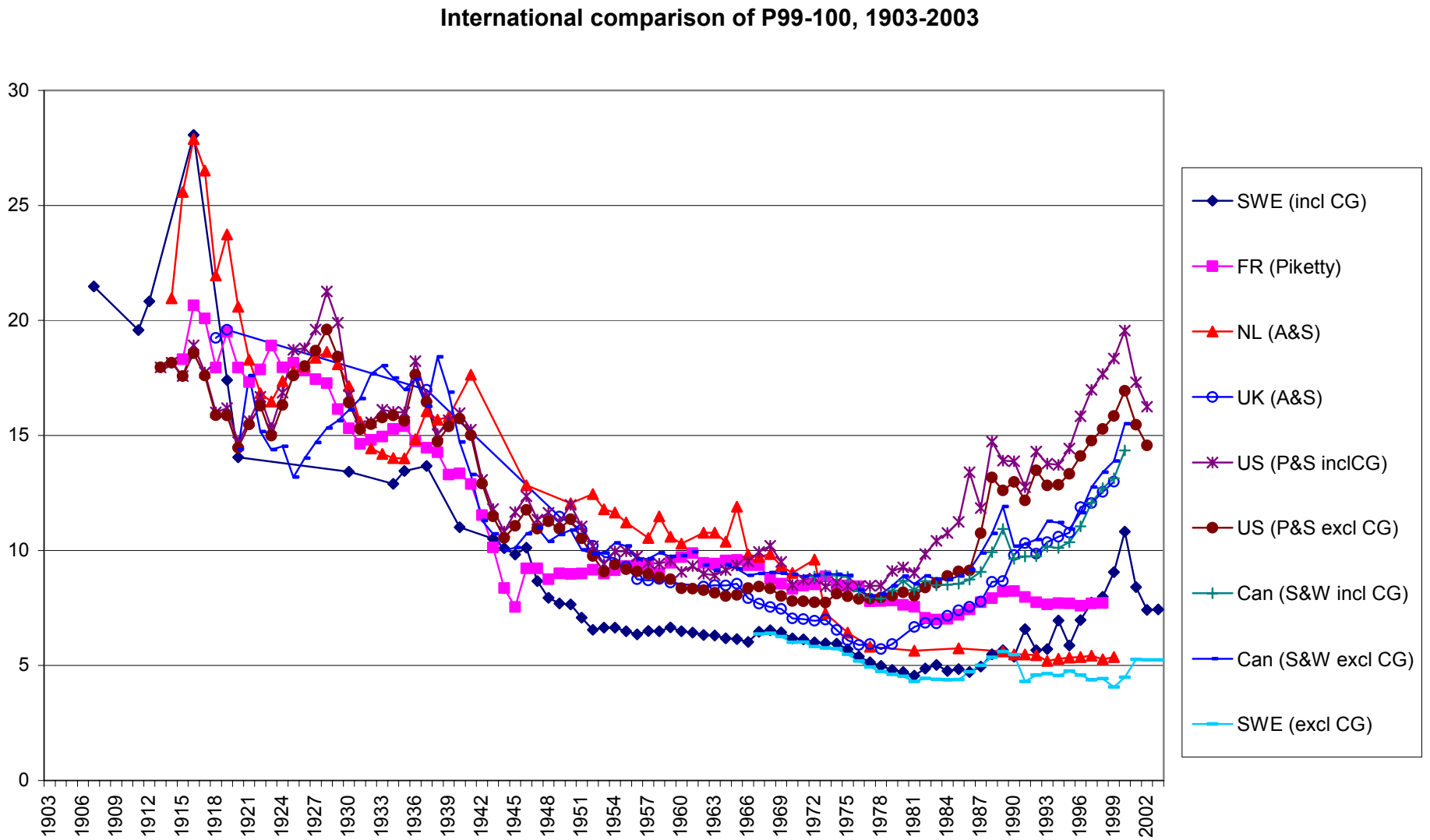

International comparison of P99-100, 1903-2003
30
25
20
15
10
5
0
1903
1906
1909
1912
1915
1918
1921
1924
1927
1930
1933
1936
1939
1942
1945
1948
1951
1954
1957
1960
1963
1966
1969
1972
1975
1978
1981
1984
1987
1990
1993
1996
1999
2002
SWE (incl CG)
FR (Piketty)
NL (A&S)
UK (A&S)
US (P&S inclCG)
US (P&S excl CG)
Can (S&W incl CG)
Can (S&W excl CG)
SWE (excl CG)

Figure 18:

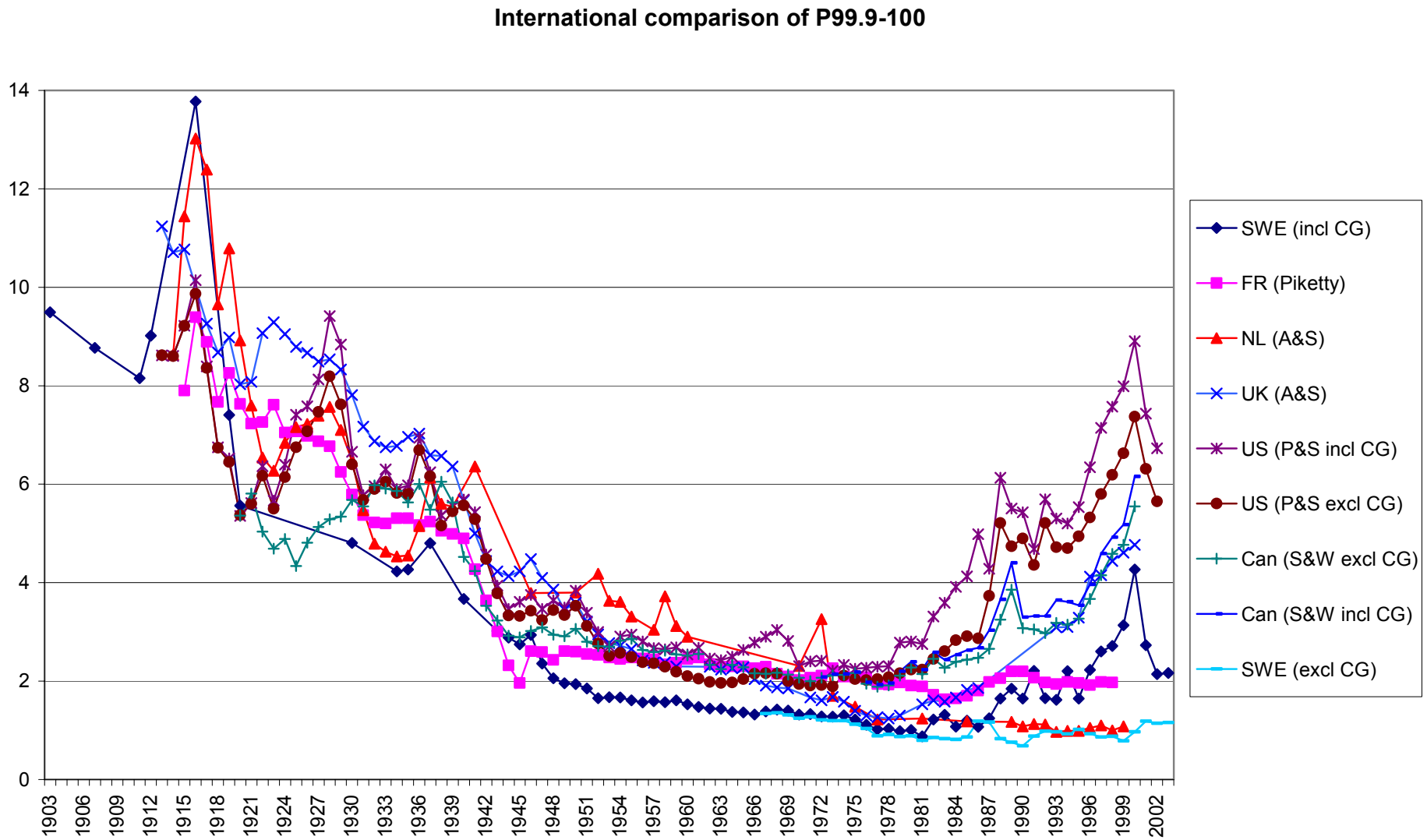

International comparison of P99.9-100
14
12
10
8
6
4
2
0
1903
1906
1909
1912
1915
1918
1921
1924
1927
1930
1933
1936
1939
1942
1945
1948
1951
1954
1957
1960
1963
1966
1969
1972
1975
1978
1981
1984
1987
1990
1993
1996
1999
2002
SWE (incl CG)
FR (Piketty)
NL (A&S)
UK (A&S)
US (P&S incl CG)
US (P&S excl CG)
Can (S&W excl CG)
Can (S&W incl CG)
SWE (excl CG)

Figure A1: The lowest taxable income and its share of average income (Reference total income/Reference total population).

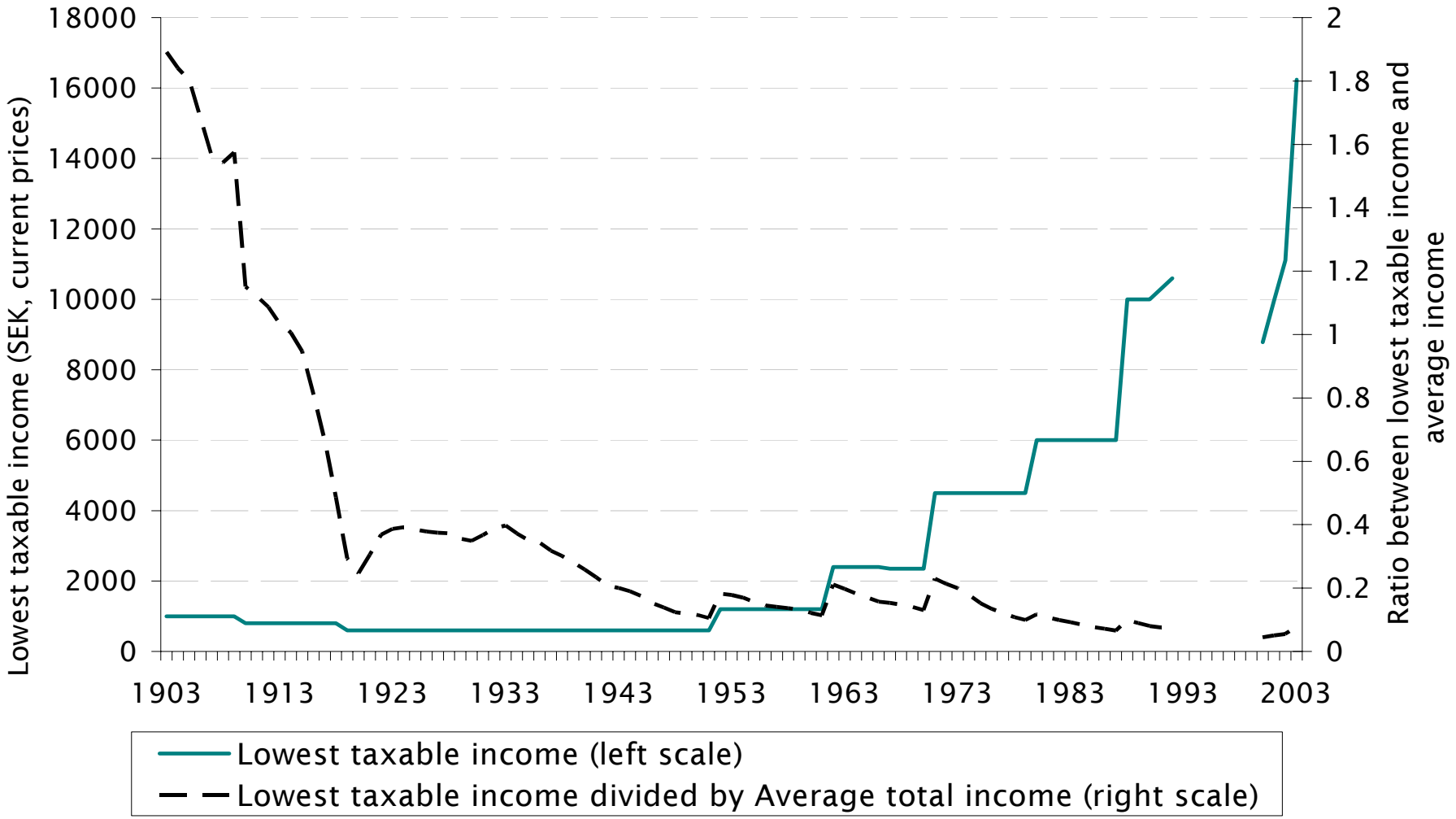


Figure B:

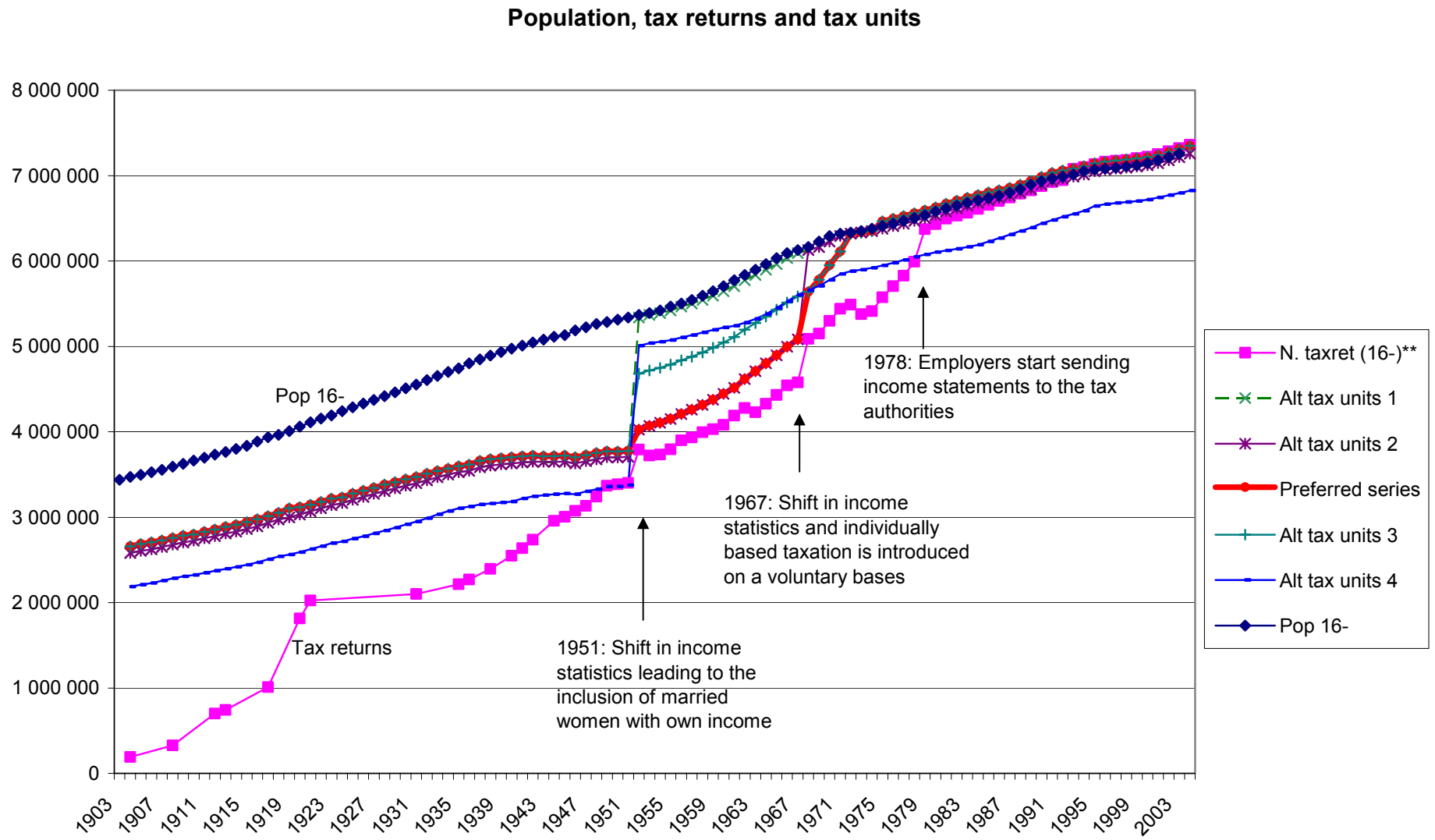

Population, tax returns and tax units
8 000 000
7 000 000
6 000 000
5 000 000
4 000 000
3 000 000
2 000 000
1 000 000
0
1903
1907
1911
1915
1919
1923
1927
1931
1935
1939
1943
1947
1951
1955
1959
1963
1967
1971
1975
1979
1983
1987
1991
1995
1999
2003
Pop 16-
Tax returns
1951: Shift in income statistics leading to the inclusion of married women with own income
1967: Shift in income statistics and individually based taxation is introduced on a voluntary bases
1978: Employers start sending income statements to the tax authorities
N. taxret (16-)**
Alt tax units 1
Alt tax units 2
Preferred series
Alt tax units 3
Alt tax units 4
Pop 16-

Figure B:

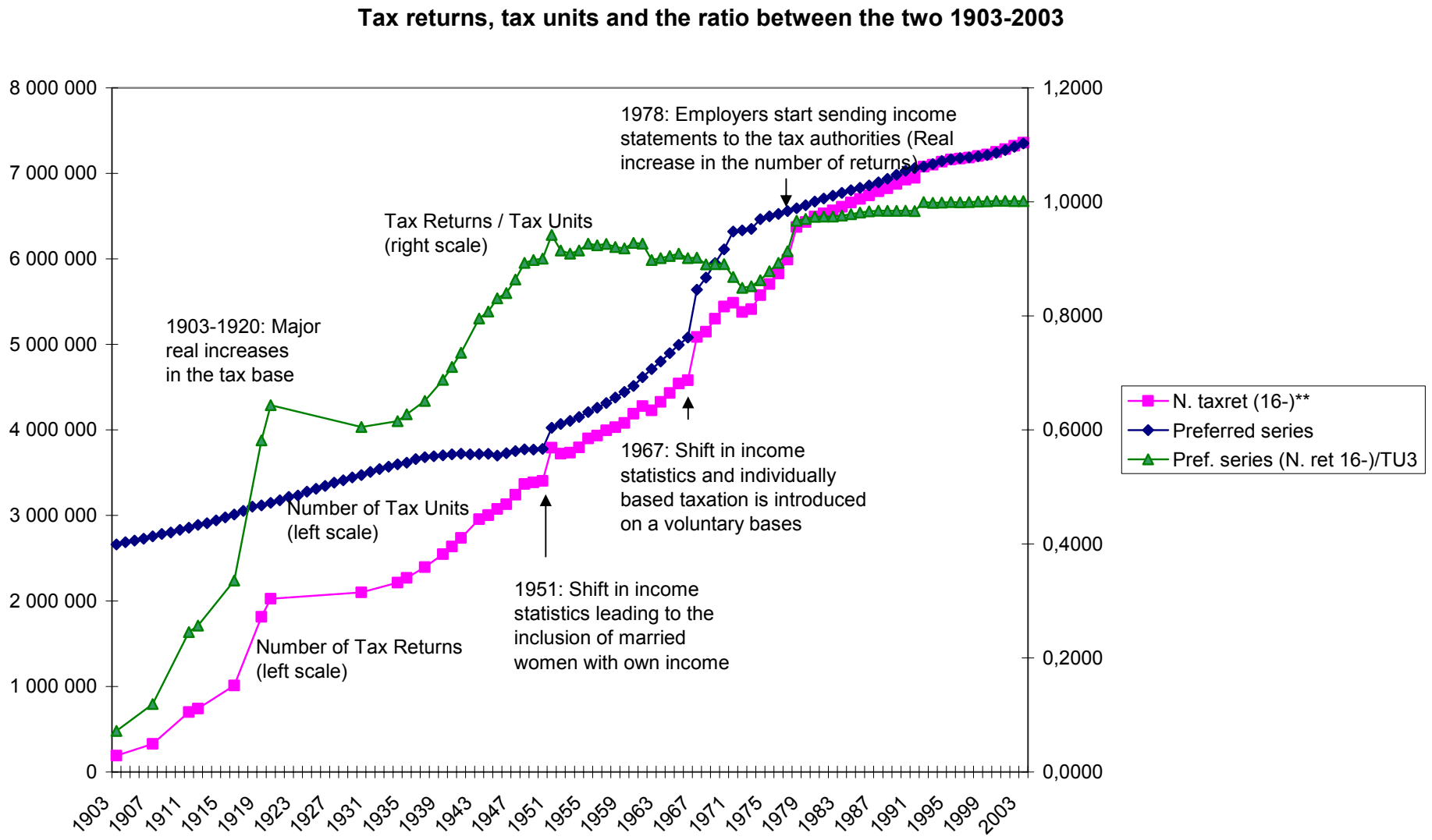

Tax returns, tax units and the ratio between the two 1903-2003
8 000 000
7 000 000
6 000 000
5 000 000
4 000 000
3 000 000
2 000 000
1 000 000
0
1,2000
1,0000
0,8000
0,6000
0,4000
0,2000
0,0000
1903
1907
1911
1915
1919
1923
1927
1931
1935
1939
1943
1947
1951
1955
1959
1963
1967
1971
1975
1979
1983
1987
1991
1995
1999
2003
1978: Employers start sending income statements to the tax authorities (Real increase in the number of returns)
Tax Returns / Tax Units (right scale)
1903-1920: Major real increases in the tax base
Number of Tax Units (left scale)
1967: Shift in income statistics and individually based taxation is introduced on a voluntary bases
1951: Shift in income statistics leading to the inclusion of married women with own income
Number of Tax Returns (left scale)
N. taxret (16-)**
Preferred series
Pref. series (N. ret 16-)/TU3

Figure B:

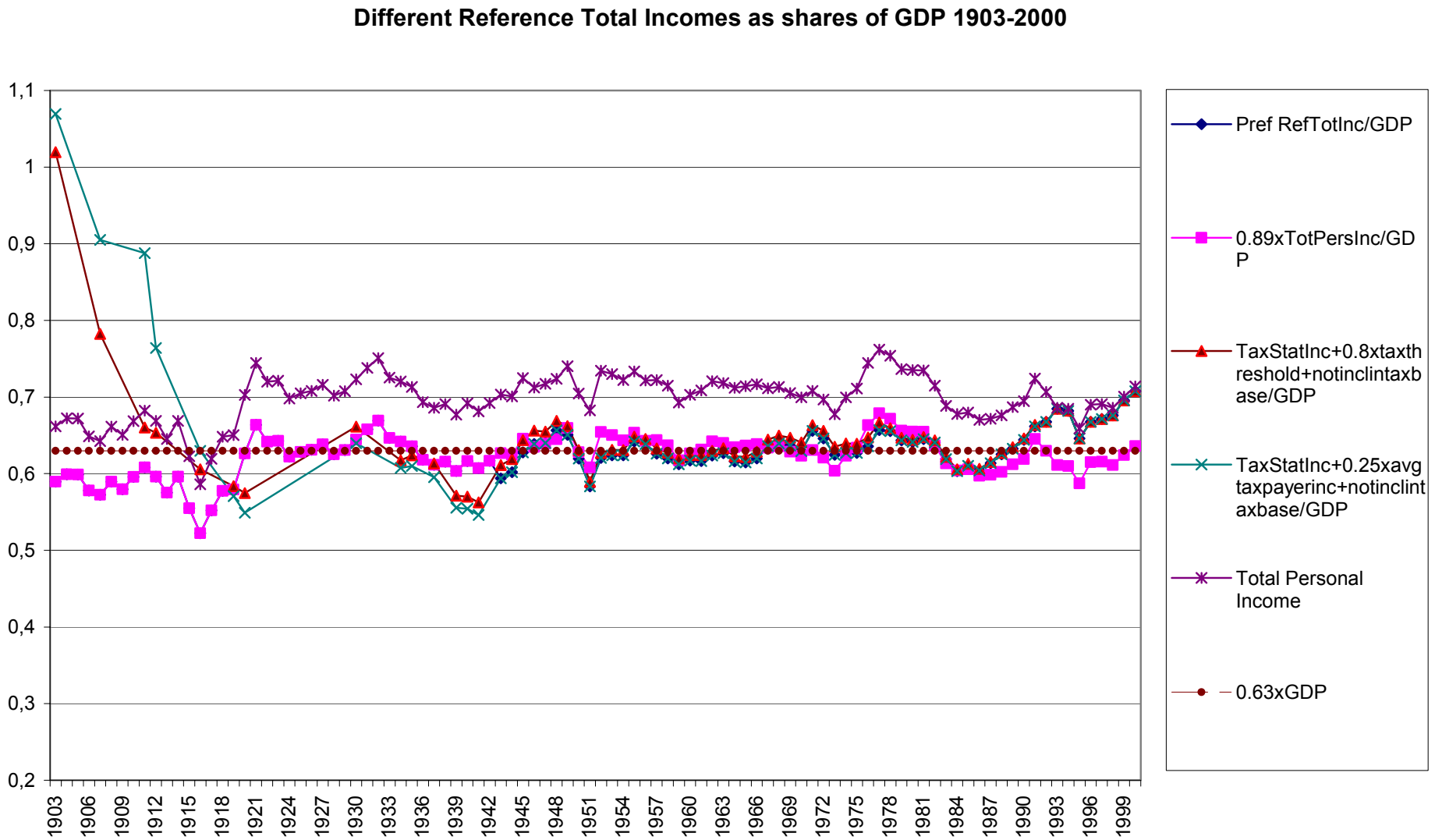

Different Reference Total Incomes as shares of GDP 1903-2000
1,1
1
0,9
0,8
0,7
0,6
0,5
0,4
0,3
0,2
1903
1906
1909
1912
1915
1918
1921
1924
1927
1930
1933
1936
1939
1942
1945
1948
1951
1954
1957
1960
1963
1966
1969
1972
1975
1978
1981
1984
1987
1990
1993
1996
1999
Pref RefTotInc/GDP
0.89xTotPersInc/GDP
TaxStatInc+0.8xtaxthreshold+notinclintaxbase/GDP
TaxStatInc+0.25xavgtaxpayerinc+notinclintaxbase/GDP
Total Personal Income
0.63xGDP

Figure B:

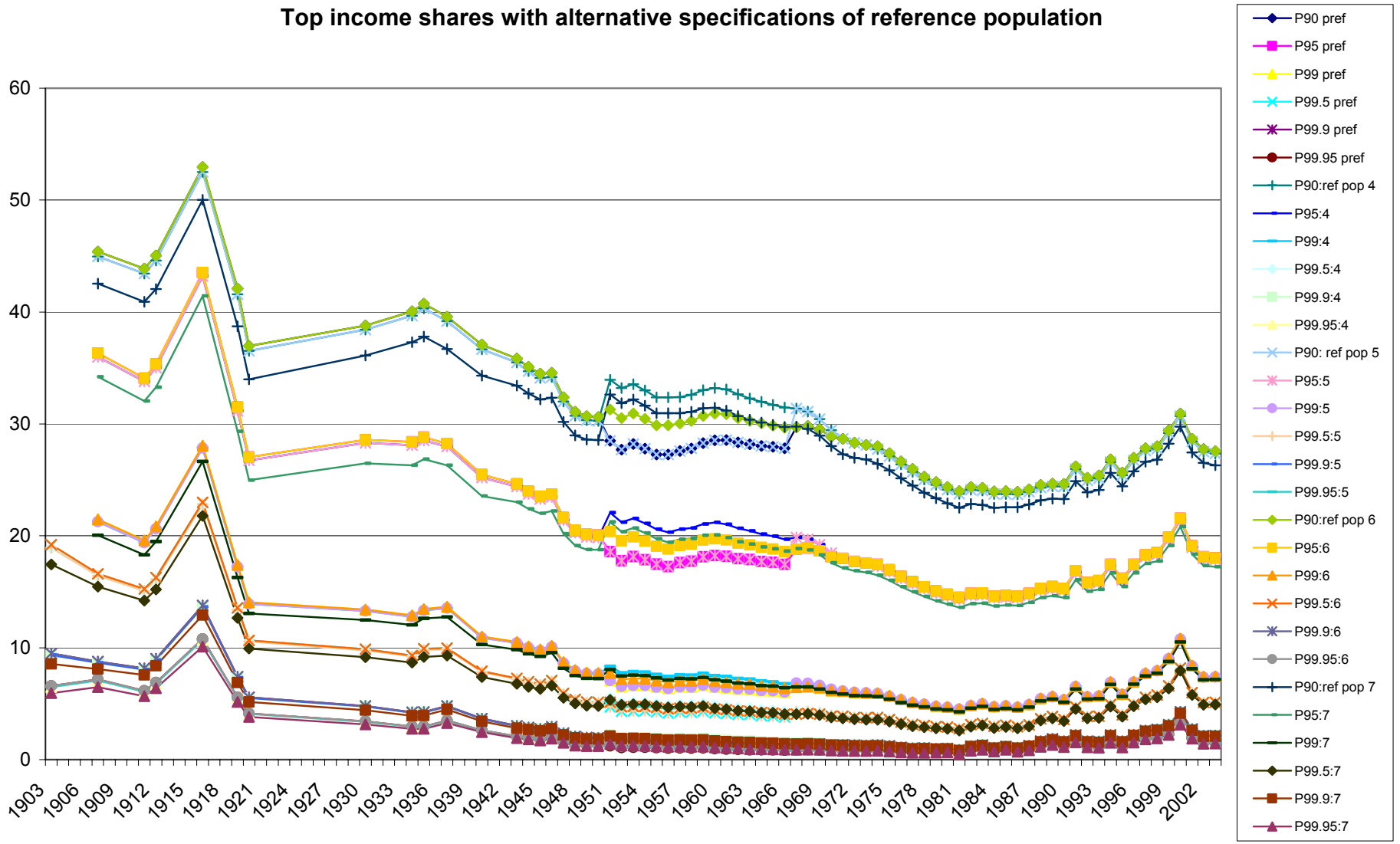

Top income shares with alternative specifications of reference population
60
50
40
30
20
10
0
1903
1906
1909
1912
1915
1918
1921
1924
1927
1930
1933
1936
1939
1942
1945
1948
1951
1954
1957
1960
1963
1966
1969
1972
1975
1978
1981
1984
1987
1990
1993
1996
1999
2002
P90 pref
P95 pref
P99 pref
P99.5 pref
P99.9 pref
P99.95 pref
P90:ref pop 4
P95:4
P99:4
P99.5:4
P99.9:4
P99.95:4
P90: ref pop 5
P95:5
P99:5
P99.5:5
P99.9:5
P99.95:5
P90:ref pop 6
P95:6
P99:6
P99.5:6
P99.9:6
P99.95:6
P90:ref pop 7
P95:7
P99:7
P99.5:7
P99.9:7
P99.95:7

Figure B:

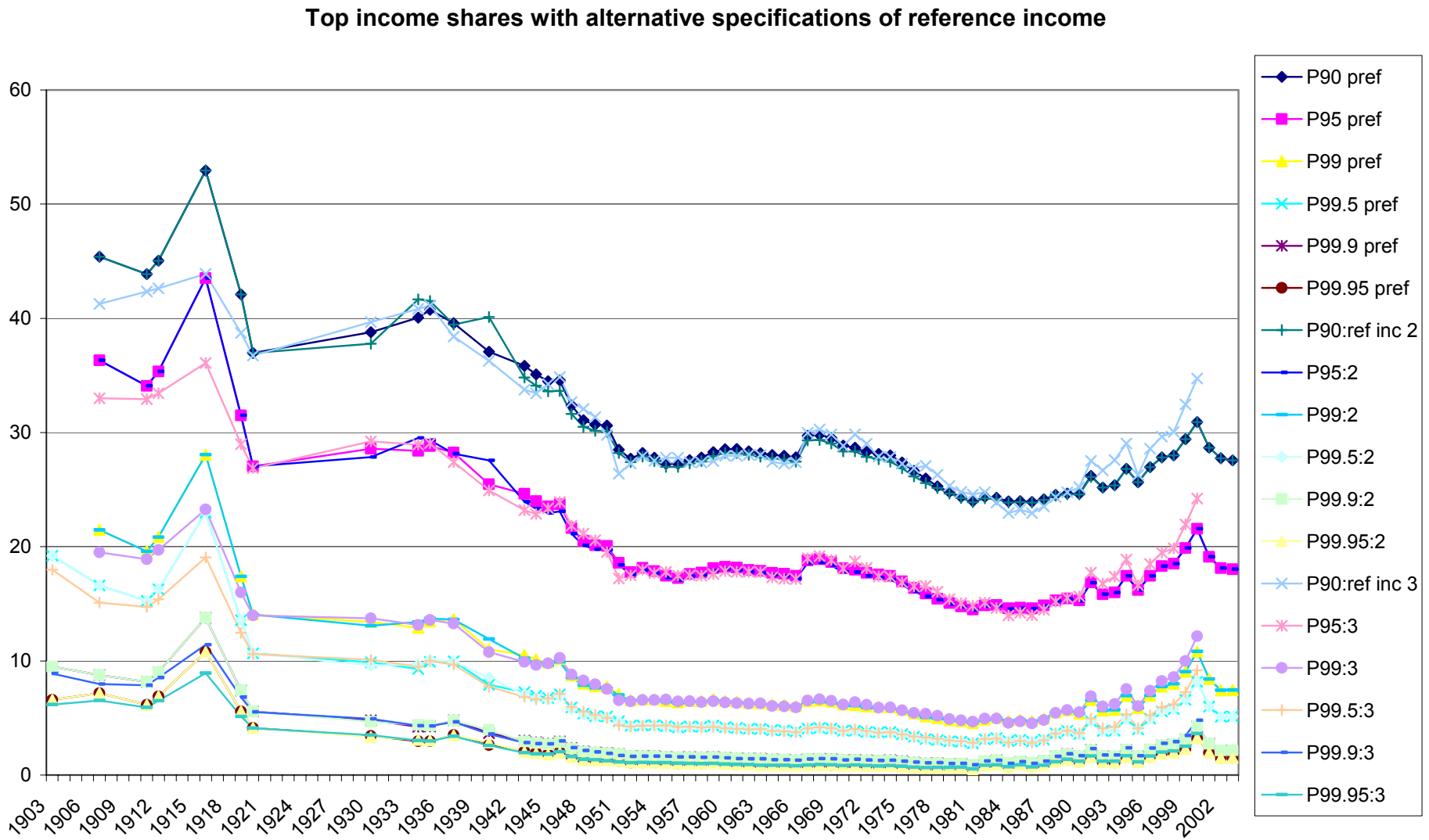

Top income shares with alternative specifications of reference income
60
50
40
30
20
10
0
1903
1906
1909
1912
1915
1918
1921
1924
1927
1930
1933
1936
1939
1942
1945
1948
1951
1954
1957
1960
1963
1966
1969
1972
1975
1978
1981
1984
1987
1990
1993
1996
1999
2002
P90 pref
P95 pref
P99 pref
P99.5 pref
P99.9 pref
P99.95 pref
P90:ref inc 2
P95:2
P99:2
P99.5:2
P99.9:2
P99.95:2
P90:ref inc 3
P95:3
P99:3
P99.5:3
P99.9:3
P99.95:3

Figure C1: Sensitivity analysis: Comparison of shares when income earner concept is either *individuals* or *households*

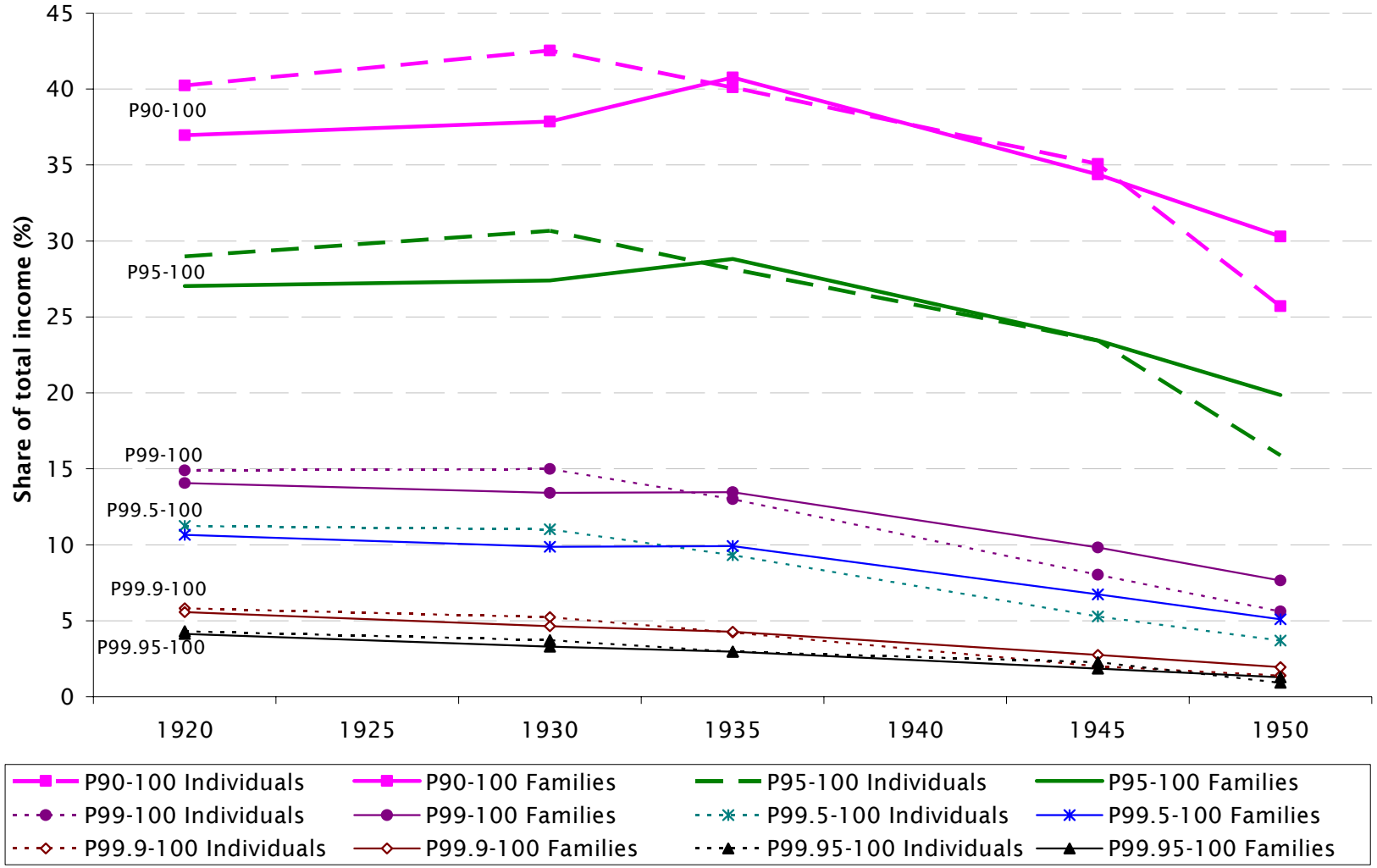


Figure C2: Shares of total income and income earners of the youngest age groups, 1951–2003

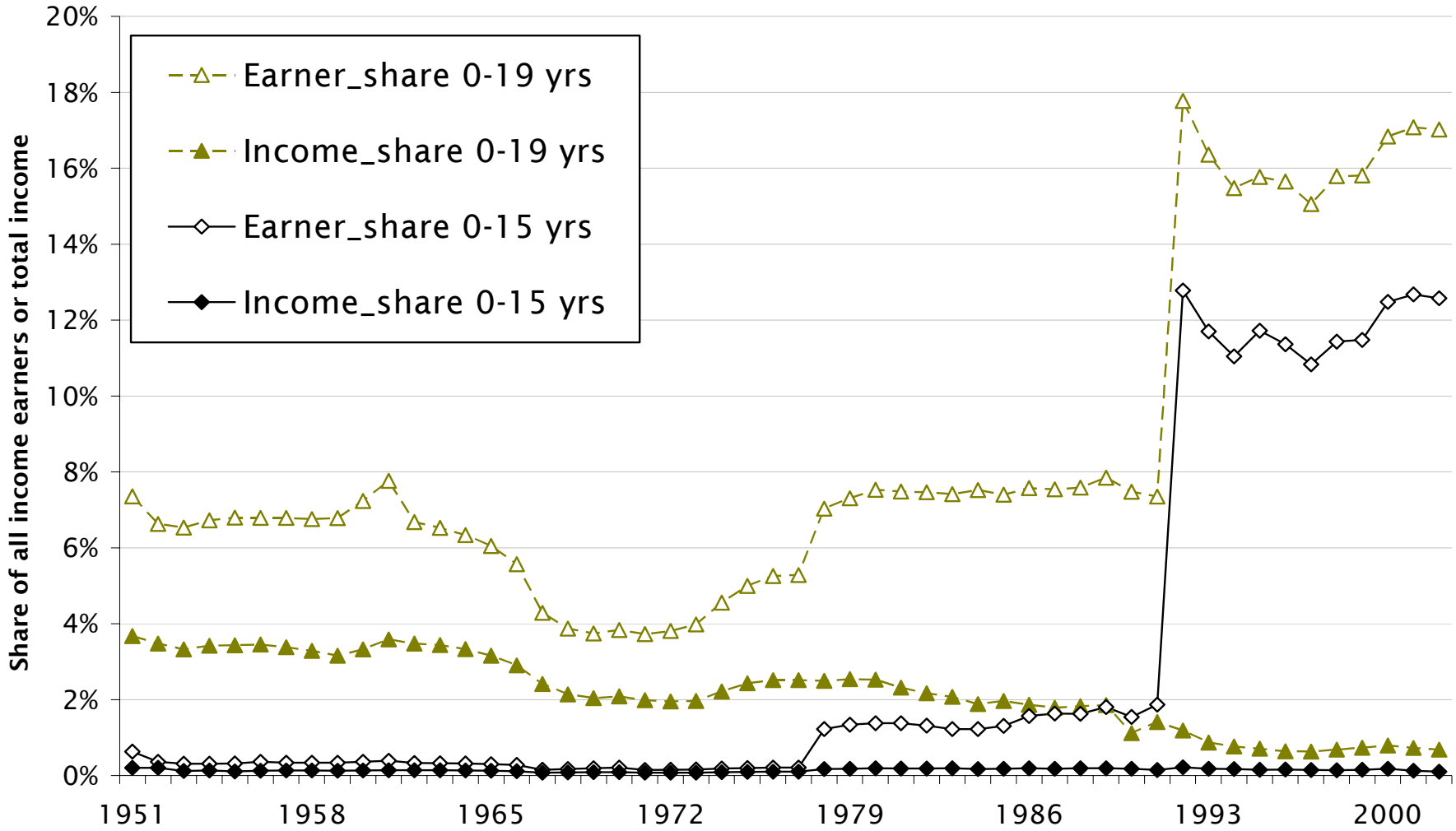


Figure D1: Highest report income level and share of income earners in highest income class.