The Suitability of Tax Data to Study Trends in Inequality

A theoretical and empirical review with tax data from Switzerland

Oliver Hümbelin[a]
Rudolf Farys[b]

[a] Bern University of Applied Sciences, Switzerland
[b] University of Bern, Switzerland

Abstract

In many countries results of inequality trends are ambiguous, because different methodological approaches blur the picture or because reliable data are not available. In this paper we assess whether tax data are suitable for the analysis of inequality trends. We do so by comparing tax data measurement concepts concerning income definition, statistical units and population coverage to theoretical-ideal concepts. We use Swiss tax data as an example to obtain a sense of the general direction and magnitude of potential biases and advantages. We therefore estimate the impact of the methodological options for measuring inequality based on tax data by comparing aggregated tax statistics and micro tax data results to corresponding results taken from surveys. While there are clear advantages to using tax data, such as long-term availability and reliable population coverage in more recent years, there are also drawbacks that lead to an overestimation of inequality based on aggregated tax statistics and hinder comparability over time. In sum, tax data are a source that should be used with care, but nonetheless seem to be indispensable for the analysis of inequality. Finally our estimations raise doubts about whether surveys are able to adequately track changes in income distribution tails, due to the undercoverage of very poor and very rich households.

Keywords: Tax data, Inequality trend, Income distribution, Switzerland

1 Introduction

Economic resources might be seen as key indicators for life chances. Therefore, the distribution of resources matters not only with respect to inequality of consumption, but also with respect to health status and even life expectancy (Wilkinson and Pikett 2009). Considering the rising economic inequality in the majority of western countries over the last decades (OECD 2008; OECD 2011; Gornick and Jäntti 2013; Salverda et al. 2014), it is not surprising that concerns about the widening gap between rich and poor are increasingly expressed by global leaders (World Economic Forum 2013). Although inequality did not rise uniformly, a common pattern seems to be identifiable; this is generally described as the "hollowing of the middle class," meaning that middle class households are moving towards the top and the bottom of the distribution (Alderson and Doran 2013). This is especially problematic as the middle class stands at the core of western democracies or, as stated by Stiglitz (2012, 117): by hollowing the middle class, "our democracy is being put at peril."

Given the importance of the subject, a constant reflection on reliability of empirical data seems appropriate. While thinking about the future needs Atkinson (2013:7) notices advances in technology and methodology regarding household surveys, the core sources of inequality research. Despite these improvements, household surveys are labor-intensive, expensive and they suffer from low response rates, which undisputedly affect the assessment of inequality. Korinek et al. (2006) showed, for example, that the probability of responding to a survey is highly driven by the position in the income distribution, leading to an overrepresentation of middle-income households and imperfect estimations of inequality. These concerns have led to the search for alternative data sources that can supplement survey data. The technological progress and the modernization of public administration improved access to several inequality relevant administrative registers like personal income or social benefit records. Especially interesting are tax data, because records reach relatively far back in time. While the use of tax data received significant attention recently with the bestseller of Piketty (2014), this approach had already been applied before. Kuznets (1955) started working with tax data to examine the relationship between economic growth and the distribution of personal incomes. More recently, Piketty (2001; 2003) and Piketty and Saez (2003) popularized the use of tax data. Following Piketty's approach, many top income studies have been conducted in several countries (Atkinson and Piketty 2007; Atkinson and Piketty 2010). Today, all time series that are based on top income tax statistics are collected and accessible through the World Top Incomes Database (Alvaredo et al. 2015).

While there is already an extensive body of literature using tax data to focus on top incomes (showing a sharp increase in English speaking countries in the last decades (Atkinson, Piketty and Saez 2011)) the utility of tax data for studies of overall inequality has not been discussed thoroughly and its potential is not yet clarified, although many researchers are interested in changes in every part of the distribution, not only the top. In this paper we therefore provide a theoretical and an empirical review of tax data for overall inequality studies. In Section 2 we describe the current standards for measuring economic inequality and analyze the theoretical advantages and disadvantages of tax data by comparing tax and survey data. In Section 3 we empirically test the extent to which tax data deviate from theoretically ideal data. We do this using federal and cantonal tax data from Switzerland, which we compare to results from surveys. We provide a summary of key findings that distinguish major from minor methodological issues with respect to the magnitude of related biases in Section 4.

2 Standards in assessing economic inequality

2.1 Income concepts

Although the OECD (2013) recommends looking at income, consumption and wealth simultaneously to adequately measure economic well-being, inequality in the distribution of income still receives most scholarly attention. While this implies a common simplification inequality studies have to declare clearly which kind of incomes they use, because the degree of inequality is connected to the chosen income concept.[1] In Figure 1 we present a stylized framework, which includes an overview of income definitions that are commonly used for inequality studies.[2] Most people earn labor income while some also have capital income. These incomes are a direct product of the market outcome and the sum of them is called the primary income. But households do not only rely on their primary income. Every western society maintains, to some degree, a system of redistribution. This includes transfers paid (taxes and direct inter-household transfers) and transfers received (pensions, social security insurances and transfers from other households). Incomes adjusted for these transfers are called disposable incomes. It is the income that is finally at disposal for consume. For international comparison of income inequality it is most common to include the effect of both government transfer and tax policies (see Atkinson & Brandolini (2001). In addition, for research purposes incomes are often equalized with an equivalence scale (see OECD 2013, 173; Buhmann et al. 1988) to make individual economic well-being among individuals comparable even if they are living in households of different size (see also the subsection on statistical units below).

[1] E.g. pensioners, unemployed or welfare recipients appear poorer, when looking at primary incomes compared to disposable incomes, because received transfers payments are neglected.

[2] For detailed discussion see: OECD (2013, 44) and United Nations (2011, 24).

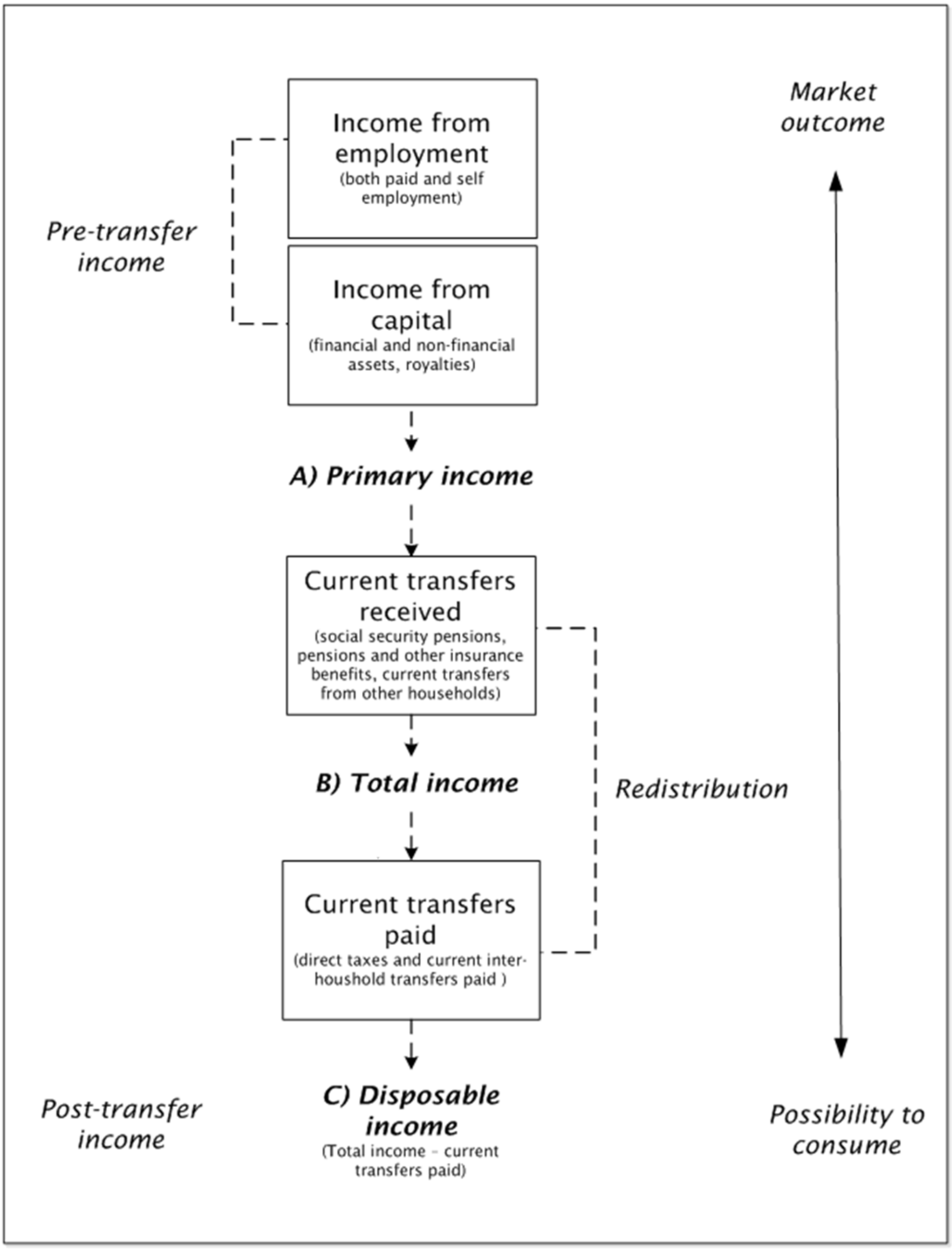

Figure 1: Income definitions from primary to disposable income

Source: OECD (2013, 44), own diagram

With tax data, *concepts of economic resources and definitions of key measures* are strongly data-driven, because tax data are collected for administrative and not for scientific purposes. Tax statistics are often easily available in an aggregated form, showing tax units per taxable income/wealth brackets, but without any information on individuals. The missing link on the micro level implies therefore that there is no possibility of doing a conjoint analysis of income and wealth. Researchers therefore are only able to analyze the distribution of either income or wealth, but not both simultaneously. In addition, information on consumption is missing entirely. The definition of key measures is often restricted too, because only tax-relevant measures are reported. Taxable incomes in Switzerland for example include direct social transfers (e.g. rents), but no mean-tested benefits (e.g.

social assistance) and taxes are not subtracted. Thus, a researcher using taxable income can look at neither a pre- nor a post-transfer measure (see Figure 1). Taxable income is rather something in between. Furthermore, deductions impose changes to income measures, which can bias the result, when deductions change over time. Aside of transfers and deductions Atkinson et al (2011) identify changes in taxation of capital income and capital gains that potentially hinder comparability over time especially for top income analysis. The situation is far better with micro tax data. If income and wealth are taxed, a complete conjoint distributional analysis is possible. Key measures can also be constructed quite flexibly, because individual tax data contain information on pre-tax income (before deductions) as well as most important expenditures like taxes. However, detailed information on consumption is still missing. Nonetheless, with respect to concepts of economic resources and definitions of key measures survey data are clearly superior, because concepts and measures can be tailored carefully to the needs of scientists.

2.2 Inequality measures

Today there are a plethora of inequality measures with different properties (Hao and Naiman, 2010; Cowell, 2011). Widely used in social sciences are *quantile function*-based measures like *top income shares, the quantile ratio or the Gini coefficient*, which is undoubtedly the most prominent inequality measure in the academic literature as well as for government statistics. As it is derived from the Lorenz curve, the quantified amount of inequality can be described simply in a formal and visual way. Therefore the Gini coefficient is easy to understand. However, several drawbacks are reported in the literature. The Gini coefficient is more sensitive to changes in the middle of the distribution, which is not necessarily a desired feature. Most importantly, being a single aggregate measure, the Gini coefficient cannot tell if it is driven by a few rich or many poor individuals. This can be problematic for comparison between countries or over time. In extreme cases two totally different distributions share the same Gini coefficient (Cowell 2011, 69). Another widely used measure is the Atkinson index. It is derived from a *social welfare function*. Atkinson (1975, 47) noted that inequality "cannot, in general, be measured without introducing social judgments." Measures such as the Gini coefficient are not purely 'statistical' and they embody implicit judgments about the weight to be attached to inequality at different points on the income scale (i.e. sensitivity in the middle of the distribution). Therefore, the Atkinson index incorporates a sensitivity parameter (ε), which can range from 0 (meaning that the researcher is indifferent about the nature of the income distribution) to infinity (where the researcher is concerned only with the income position of the very lowest-income group). One obstacle to using this measure is that the researchers must actively choose, and thus justify, their choice of sensitivity parameter. Similar to the Atkinson index, measures derived from *information theory* (e.g. Theil index) incorporate a sensitivity parameter that varies in the weight given to different parts of the income spectrum. A beneficial property of information theory-based measures is that they are decomposable; that is, they can be broken down into component parts (i.e. population subgroups). This enables analyses of between- and within-group effects.

The *estimation of inequality measures* is flexible when data are available on the micro level – as is commonly the case with survey data and also with micro tax data. As surveys relate on samples the estimation includes inherently a statistical uncertainty, which means tax data relying on full population are more precise. The problem gets more accentuated with measures focusing on the extremes of the distribution. As income distributions are usually left skewed with long tails this is especially true for measures focusing on the upper part of the distribution. If a researcher has to deal with aggregated tax data, however, calculation of inequality measures is restricted. First, the precision of the measures suffers because of the aggregation. Different methods have been used for interpolation, such as the Pareto interpolation and the split histogram (see Atkinson 2005). Second, it is not possible to decompose the measure by household characteristics. Nonetheless, all common measures (like the Gini coefficient or Theil index) can still be calculated.

2.3 Statistical units

Commonly, households, not individuals, are the statistical units for inequality analysis (OECD 2013, 60f). Indeed, although individuals receive an income, own assets and consume goods and services, their possibility of doing so is strongly tied to the concept of the household. Following the OECD a household is defined as all persons living in one housing unit and combining incomes to provide themselves with food and other essentials of living. Data are collected on the household level instead of the individual level because it is assumed that people in the same household share resources and therefore pool their incomes (when two or more earners live together) and use the household income to provide the essentials for every household member (also non-earning members, like children). Correspondingly, there are economies of scale for people sharing living space and commodities. For the comparison of the individual economic well-being among individuals living in different households, usually equivalence scales are used, as mentioned above.

The adequate *statistical units* are easier to identify with survey data, because the household situation can be identified directly as part of the survey process. The statistical units of tax data, however, are tax units or fiscal households, but these do not necessarily correspond to real households. In fact in some countries, such as Canada, New Zealand from 1963, or the United Kingdom from 1990, the tax unit is the individual (Atkinson et al. 2011). In other countries, like France, the United States or Switzerland tax units represent families (i.e. singles or married couples). At least they should. Indeed, there are situations where members of the same household submit several tax forms. A common case is an unmarried couple living together. With changing household structures, this issue becomes increasingly important. Another problem are adult children living with their parents but who are taxed independently (Burkhauser et al. 2012)

2.4 Population coverage

Generally, inequality studies try to make a statement about the whole population of interest (e.g. nation). But resources and/or options strongly determine whether such a venture has success, as these restrictions shape the way data are collected. When total population data are not at hand, researchers usually work with samples and try to infer from samples to the population. This is a thorny task for surveys because nonresponse is a major source of bias (Bethlehem et al. 2011). As Korinek et al. (2006) show, the position in the income distribution influences the probability to participate in a survey. Low-income and high-income households are more likely to refuse survey response, which leads to an overrepresentation of middle-income households. Missing data in household surveys are therefore not missing at random, which has an impact on the measures of inequality. Additionally surveys lose validity through incomplete response because of unintended (and intended) reporting errors. Therefore most surveys impose top coding to limit the effects of measurement error on aggregates, which particularly affects top income analysis (see Brewer et al. (2008) for an example in the United Kingdom). Alternatively, researches can use income data from registers, when suitable administrative data and a legal basis to use them for statistical purposes exist. In fact, nearly a third of all countries that participate in the European Union's Statistics on Income and Living Conditions (EU-SILC) collect at least some of their income data from registers (OECD 2013, 93). Tax statistics are attractive because they technically provide total population coverage. Compared to surveys they are not subject to sampling bias. They may, however, suffer from undercoverage or missing data as well. A critical issue is tax evasion, which can definitely bias the assessment of inequality. Evasion occurs when individuals do not fill out tax returns or misreport incomes. Alvaredo and Saez (2009) for example consider estimates of Spanish top incomes prior to 1981 to be unreliable due to widespread tax evasion. With respect to coverage, tax laws that define the taxed population are crucial. This is especially problematic for older tax statistics, because many countries started taxation with very progressive tax schemes and high exemption levels that narrowed the tax to only a little group. Atkinson et al. (2011: 20) gathered information on key features of tax information for several countries. The taxed population in many countries initially covered only around 10% of total population or even less.

2.5 Comparison of tax data and survey data – overview of advantages and disadvantages

To define a standard of measuring economic resources and related inequality, we introduced four key areas researchers need to address. Ideally, researchers want to (1) look at income, wealth and consumption together, (2) have data suitable to calculate all types of inequality measures in a precise way, (3) do that for disposable resources on a household level and (4) calculate an unbiased estimate of a chosen inequality measure. Table 1 summarizes the comparison of tax and survey data on these four dimensions.

Table 1: Comparison of tax data and survey data

	Aggregated tax statistics	Micro tax data	Surveys
Concepts of economic resources and definition of key measures	strongly data-driven	data-driven	theory-driven
Estimating inequality measures	restricted, rather precise	flexible, precise	flexible, imprecise
Statistical unit	tax units	tax units	households
Coverage problems	tax evasion, tax exemption	tax evasion, tax exemption	nonresponse, undercoverage

The main advantage of aggregated tax statistics not mentioned so far is *availability*. First, tax statistics are often publicly available. Second, tax statistics exist in many countries for very long time periods. This makes them an interesting data source, aside from the mentioned restrictions. For several countries the availability of tax records reaches back in time over 100 years, allowing assessment of time trends that cover substantially longer periods than is possible with survey data. Nonetheless, concerning comparison over time, scientists have to test comparability, because measures and population might be affected by changes in the tax systems or the way tax statistics are reported. The availability of micro tax data, however, can be restricted, because of privacy reasons and also because of limited archiving resources. While a document with aggregated tax statistics is a neat and parsimonious way of historicizing information, the requirements for complete micro data archiving are far greater. In the United States micro tax data is at least available since 1960 (for example (Piketty & Saez, 2003). In Switzerland federal micro tax data is available since 1973/1974 but complete micro tax data is only available after the millennium and it is accessible only for some cantons. Finally, household surveys are easy to access for scientific purposes, if they exist at all. In the European Union, for example, many countries did not implement household surveys for distributional analysis before 2003 or even later (Eurostat 2015). The potential to assess inequality trends with survey data therefore is restricted to relatively short periods in many countries. Burkhauser et al. (2012) compared income inequality trends in the US using Current Population Survey (CPS) and Internal Revenue Service (IRS) tax return data. They investigate on the importance of income inequality measure, statistical unit and income definition and conclude that tax data income and statistical unit definition increase observed levels of income inequality but do not greatly impact trends. Differences in inequality trends observed by researchers using these two data sources are not primarily due to deficiencies in either data source but rather to the traditions of income inequality measures used in the two literatures (top-income shares vs Gini). The authors, however, neither did investigate on coverage issue, nor did they disentangle the role of income definition and statistical units in a systematic way.

3 Empirical case study with tax data from Switzerland

As we will show, results of studies on income inequality in Switzerland are inconsistent, which makes Switzerland an interesting case to have a closer look at methodological aspects. Looking at official data for Switzerland, there are three main data sources: the Statistics on Income and Living Conditions (EU-SILC), the Household Budget Survey (HBS) and the Luxembourg Income Study (LIS). Figure 2 shows Gini coefficients of equalized disposable income calculated from these three sources plus a time series we calculated on the basis of aggregated tax statistics published by the Swiss Federal Tax Administration (FTA). To date, EU-SILC is the main source used for policy monitoring at EU-level. The main focus of EU-SILC is to collect data on a common framework to ensure comparability among EU and European Free Trade Association countries. As a non-EU member, Switzerland did not join SILC in the first year of data collection (2004), but rather waited until 2007. Therefore, this time series does not cover the period before 2007. According to the results from EU-SILC, income inequality decreased from 2007 to 2012. The second important source concerning the distribution of income is the HBS. The main focus of this survey lies in providing detailed information on household budgets. Since 2000 the survey has been conducted on a continuous basis, which allows looking at a consistent time series from 2000 to 2011. As is evident from Figure 2, the trend using HBS is rather stable. Both time series (EU-SILC and HBS) cover a relatively short time period. A longer period is covered in the LIS dataset (1982-2004). The LIS data are harmonized using three surveys: the Swiss Income and Wealth Survey (1982), the Swiss Poverty Survey (1992) and the Income and Consumption survey (2000, 2002 and 2004). The harmonization done in the LIS dataset provides the longest time series on inequality for Switzerland. Analyzing these data, Gornick and Jäntti (2013) found a quite substantial decrease in income inequality for Switzerland, the opposite trend as in most other western countries. The time series we constructed from federal tax data however cover a longer time period; they suggest overall higher inequality and a slight increase in recent years. This result is in line with Foelmi and Martinez (2014), who calculated top income shares for this period. The question arises: Why do the series differ and which one is most accurate?

Differences might be explained with factors introduced in Sections 2. First, coverage of low and top incomes is assumed to be better within tax data than within survey data due to nonresponse bias. If this is true, inequality assessed with surveys is underestimated. The FTA series, however, is based only on taxed subjects (tax units below the taxation threshold do not show up in the statistics). Second, different income concepts were used. The tax data time series is based on taxable incomes, while the surveys rely on disposable income and use an equivalence scale. As Modetta and Müller (2012) have shown, the income distribution is strongly affected by governmental redistribution through social transfers and taxes, reducing inequality substantially. Third, the statistical units within tax data are fiscal households and not real households, which again are the base of analyses for the survey studies. With a trend of unmarried cohabitation, this could lead to a bias within tax data. To sum up: using different data sources and different concepts can lead to substantially different results. Because misspecifications overlap, it is hard to disentangle the single sources that potentially lead to a bias and therefore it is hard to say where truth is hidden.

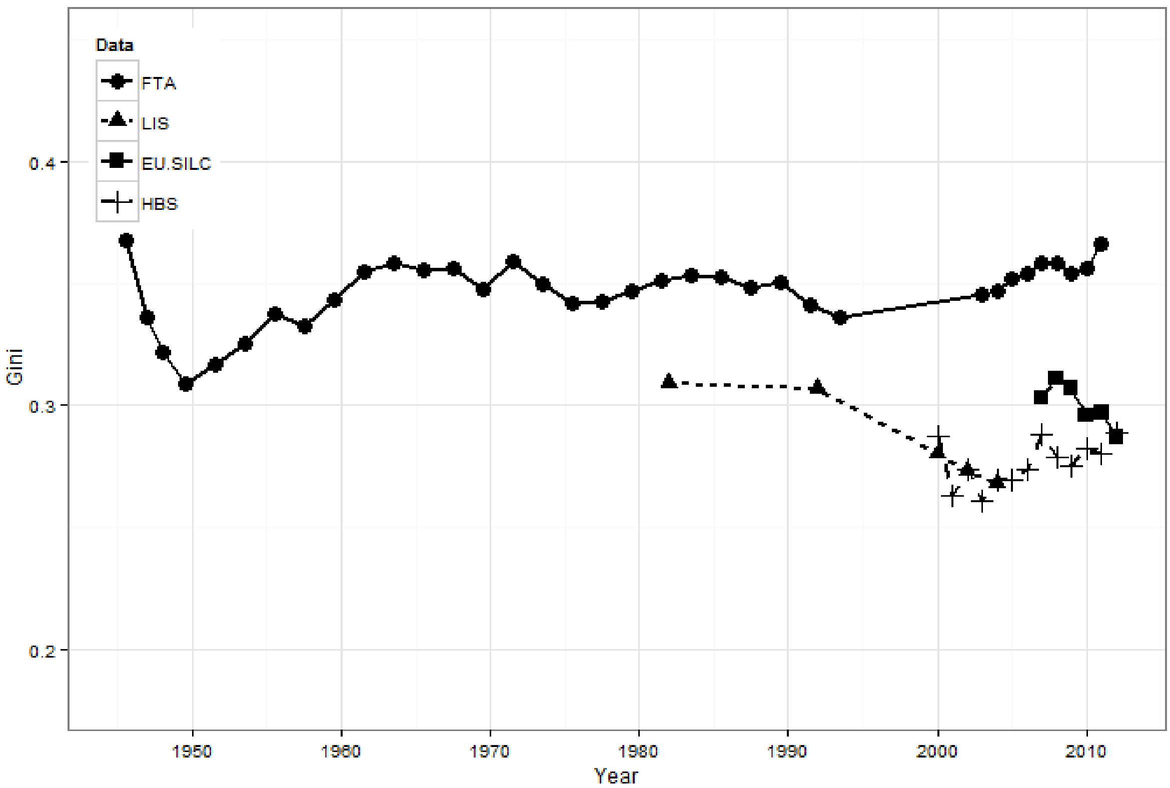

Figure 2: Income inequality trends in Switzerland

Source: Luxembourg Income Study Inequality Key Figures (LIS)[3], Household Budget Survey (HBS)[4],

[3] http://www.lisdatacenter.org/lis-ikf-webapp/app/search-ikf-figures

European Union Statistics on Income and Living Conditions (EU.SILC)[5]*, Swiss Federal Tax Administration (FTA)*[6]

In this section we isolate all potential sources of error following the four areas (income concepts, inequality measures, statistical units and population coverage). We discuss in what way theoretical concepts introduced in Section 2 can be addressed with tax data in Switzerland, and we quantify empirically the direction and magnitude of several theoretical data-specific misspecifications. Exceptions are tests (3) and (4), which we provide in addition to the data-specific tests to show how different statistical techniques can be applied to aggregated tax statistics and how top and bottom sensitivity measures change interpretations. The results of the formulated questions below are intended to serve as guidelines to identify issues that are relevant when working with tax data in general, while at the same time shedding light on the contradiction presented in Figure 2.

Income concepts

1. How do tax data-based income definitions alter inequality measurement?
2. What is the impact of using an equivalence scale derived from tax data?

Inequality measures

3. Do different measures (Gini, Theil, Atkinson) report different trends?
4. On top of these common measures, what can we learn from comparing full income distributions?

Statistical units

5. How important is observing real households instead of tax units?

Population coverage

6. How do survey and tax data differ with regard to population coverage?
7. Do we have to worry about so-called "special tax cases"?
8. How large is the bias due to not observing non-taxed units?

3.1 Data and methods

Our main data source is income tax data published by the Swiss FTA.[7] Federal taxes are collected and documented by the FTA since 1915. For this paper we use data from 1945 to 2011, covering 35 tax periods.[8] While the FTA provides data electronically readable since 1973, we collected earlier data by scanning hard copies. In general, data are provided by the FTA in an aggregate form for privacy reasons, i.e. they are classified into numerous income brackets. Because these data do not always contain all desired information, we use additional data sources (see the column *Data* in Table 3 in the appendix). This includes FTA-published key figures based on the federal micro tax data.[9] These figures include Gini coefficients and percentiles ranging from 1973/1974 to 2011 for individuals who had to pay federal taxes and from 1995/1996 for all taxable individuals. Additionally, we use cantonal

---

[4] Calculated and kindly provided by Modetta and Müller (2012).

[5] http://ec.europa.eu/eurostat/tgm/table.do?tab=table&init=1&language=en&pcode=tessi190&plugin=1

[6] Own calcuations.

[7] http://www.estv.admin.ch/dokumentation/00075/00076/00701/index.htm

[8] We did not use tax data before 1945 although they are accessible from 1915 because data before 1945 comprise only a minority of potential tax units. According to estimations of Dell et al. (2007) the share of tax filers before 1945 was below 50% and sometimes even below 15%. In addition, we have a gap in our data between 1993 and 2003, where the annual presence taxation (Praenumerando System) was implemented. Before 1993, tax periods comprise two years, because taxes were levied with the Postnumerando System (taxation based on income generated two years in the past). Cantons implemented the change in different years, which is why there are no exact data available for Switzerland in the transition period.

[9] These calculations were done on commission of the FTA within the SNF project Sinergia Nr. 130648, "The Swiss Confederation: A Natural Laboratory for Research on Fiscal and Political Decentralization", by Raphael Parchet and Stefanie Brilon in coordination with Prof. Dr. Marius Brülhart.

micro tax data for tests that are not possible with FTA tax statistics, but nonetheless provide information in regard to tax data in general. We are able to use micro tax data from the canton Bern, one of the largest canton in Switzerland, which has a fairly representative mix of rural and urban areas. Using the micro tax data from Bern we can construct more flexible income concepts, which is necessary to answer question (1). Additionally we were able to link these data to personal register information to generate a unique register-based household ID, which allows us to address questions (5) and (6). For question (6) we furthermore use the Household and Consumption Survey (HBS). This survey is commonly used for distributional analysis by the federal statistical office in Switzerland (ESTV 2014), and incomes are provided on a very detailed base, which enables us to make comparisons to incomes derived from tax data.

In general we base the analyses on the longest available time series. Because the availability of data or certain information can change over time, we are forced to restrict certain analyses to specific time periods. Table 3 in the appendix gives more detailed and standardized information about the data source, population, time frame, income concept and method used to conduct the analyses. For the analyses, we use several statistical techniques (see the column *Method* in Table 3). To assess the development of inequality over time, we calculate Gini coefficients for all possible time points. For test (3) we additionally calculate the Atkinson and Theil indices. Then we apply relative distribution methods where we think an in-depth distributional analysis provides a more insightful understanding of distributional differences than comparing measures conflating information on one distribution into a single statistic.

3.2 Income concepts

As described in Section 2.1, an analysis of income inequality should simultaneously look at income, wealth and consumption. But the OECD (2013, 13) also states:" [...] integrated analysis at the household level has significant data requirements that go beyond the measurement efforts currently undertaken in most countries." This last statement holds for Switzerland too, although the HBS study is strongly influenced by the recommendations of the Canberra group handbook (United Nations, 2011), which in turn is part of the ICW framework of the OECD. Although the FTA publishes statistics on income, wealth and federal taxes, it is not possible to analyze the joint distribution on the micro level. In addition, measures of consumption and taxes are missing in aggregate tax data. These problems can be better addressed with cantonal micro tax data. These data contain information on income, wealth and all direct taxes. It is therefore possible to analyze how the assessment of income inequality is affected by using different income definitions that are present within the tax data (3.2.1). Then we evaluate the impact of using an equivalence scale tailored to tax data (3.2.2).

3.2.1 Income definitions within tax data

When focusing on income, the key measures reported in tax statistics are tax measures. To assess the effect of income definitions within aggregated tax data we get three income measures:

- *Net income:* total income (earnings, income from property and current transfers received) minus some deductions (excluding social deductions).[10]
- *Taxable income:* net income minus social deductions.[11]
- *Taxable income after federal taxes:* By taking account of the reported federal taxes per taxable income bracket, we can construct an income measure, which is a kind of pseudo disposable income.[12]

[10] These deductions include: professional expenses, travel expenses, interest on debt, alimonies, training costs, party contributions, private pension provision "Säule 3a", buying into the pension plan, medical expenses over 5% of income and charitable donations.

[11] Social deductions include deductions for: married couples, single-parent households, second earner deductions, insurance premiums, interests earned by savings, and deductions for children and supported persons.

[12] We call it a pseudo disposable income, because important expenses like cantonal and municipal taxes, which represent the bulk of taxes in Switzerland and also the cost of health insurance, are not covered at all.

These tax measures do not correspond directly with theoretically defined measures like primary income (before redistribution) or disposable income (after redistribution). Rather, they are situated between the poles of market outcome (primary income) and income left for consumption (disposable income) (see also Figure 1 on page 4). Using these three income definitions we calculate Gini coefficients. As Figure 3 shows, the series cover different time periods, depending on the reported information by the FTA. The longest time period is covered using taxable income and taxable income after federal taxes (1945 to 2011). Information on net income only reaches back until 1981/1982. The three measures develop in parallel with the exceptions of the 1980s and 2011. In these periods the Gini coefficient for net income deviates from the other series. This is due to changes in the tax exemption threshold (e.g. inflationary adjustments or extended deductions; see section 4.5.3) and shows that longitudinal data need to be interpreted considering changes in taxation or regulation systems. In general, inequality assessed with taxable income is higher than inequality assessed with net income or taxable income after federal taxes. This is not surprising: Federal taxes reduce inequality slightly because of the tax progressivity. In addition, inequality is higher for taxable income than for net income, because of social deductions (see footnote 11), which are fixed-rate deductions related to household characteristics. Hence, subtracting social deductions from net income results in over proportional reduction of lower incomes.

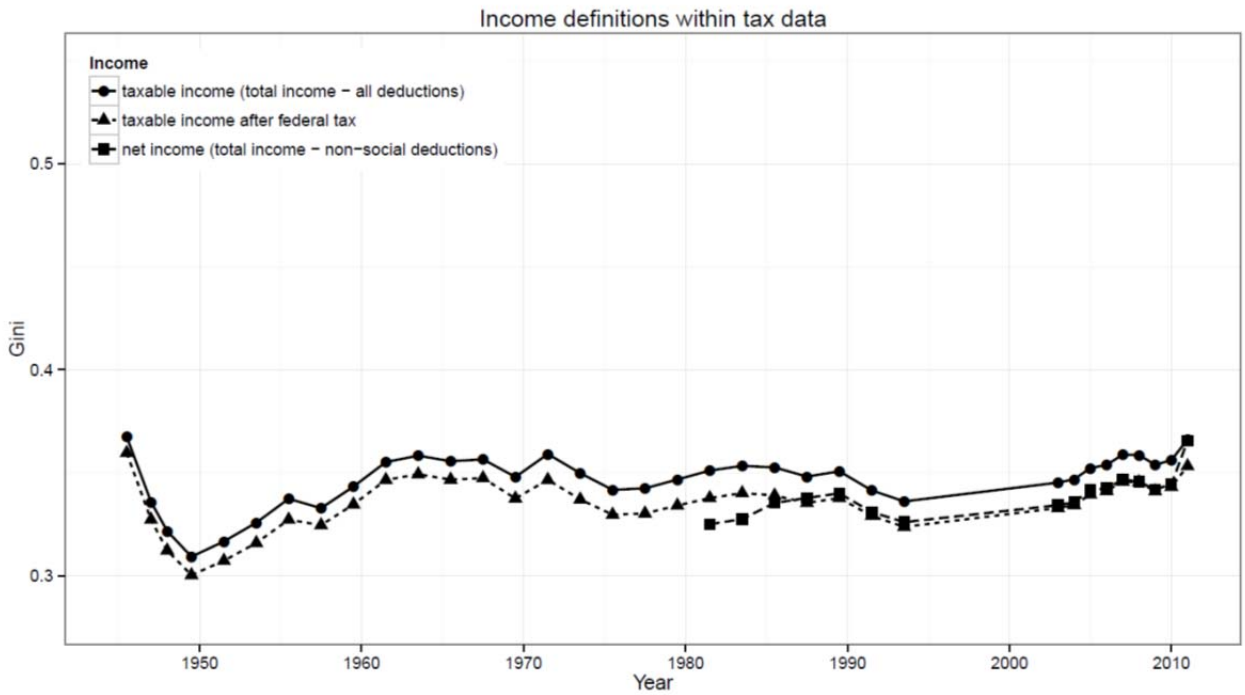

Figure 3: Inequality trends using different within tax data income definitions
Source: Aggregated Tax Statistics (FTA)

Using micro tax data from Bern, we are able to quantify how much Gini coefficients calculated with taxable income deviate from a coefficient based on disposable income. We additionally provide a time series based on total income, to be able to relate differences either to deductions or to taxes. Figure 4 shows that the Gini coefficient based on taxable income is highest and that the difference between the theoretically more sound disposable income (total income minus taxes and private transfers paid) and the often available taxable income is huge (roughly Δ 0.1 each year). Surprisingly, a bigger part of the difference is explained by deductions, while an inequality reduction through progressive taxation is present, but with lower impact.

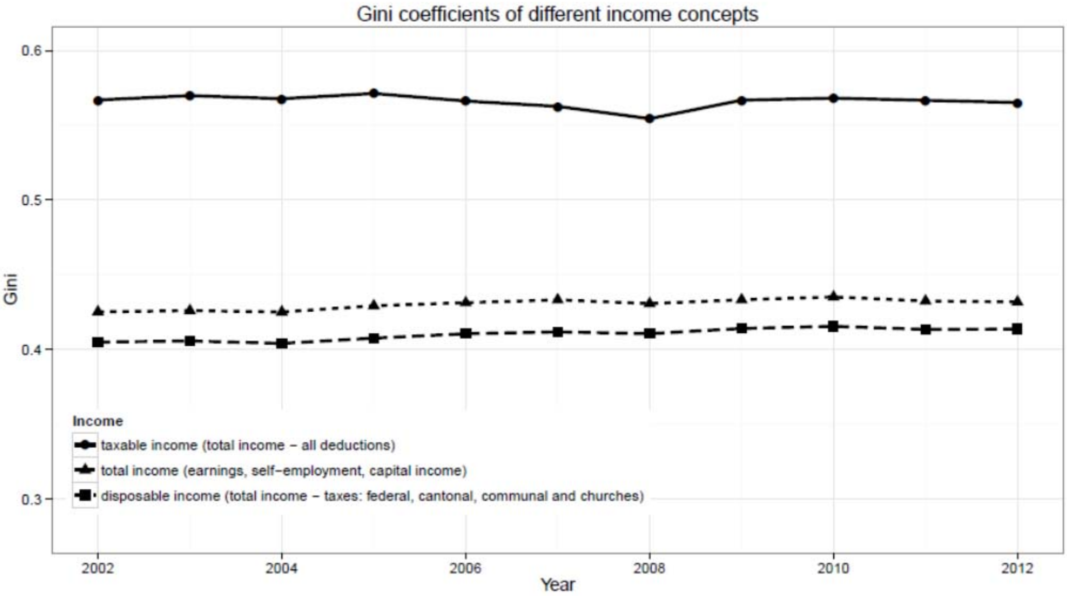

Figure 4: Inequality trends comparing taxable income to disposable income
Source: Micro tax data (Bern)

3.2.2 Using income corrected with an equivalence scale based on tax information

Income inequality studies often work with an *equivalence scale* to account for the number of household members that potentially share income and resources. Because tax data refer to fiscal households and not to real households, it is only possible to use an approximation of the equivalence concept, which uses a scale that is based on information from tax data and applied to tax units. The incomes of single households are divided by 1 (no change), while for married tax units the equivalence factor is 1.5. For every child and person supported by the tax unit, a value of 0.3 is added to the denominator. These calculation steps follow the logic of the modified OECD scale (OECD 2013, 173).[13] We compare the Gini coefficient with and without equivalence scale to find out, how strong the assessment of inequality is affected by the scale. As excluding the group of non-taxed individuals (on the influence of non-taxed individuals see Section 3.5.3) leads to a longer time series, we provide four time series in total (two possibilities to compare the effect of the equivalence scale).

[13] The implementation of this equivalence scale is not done by us. It is part of the key figures provided by the FTA.

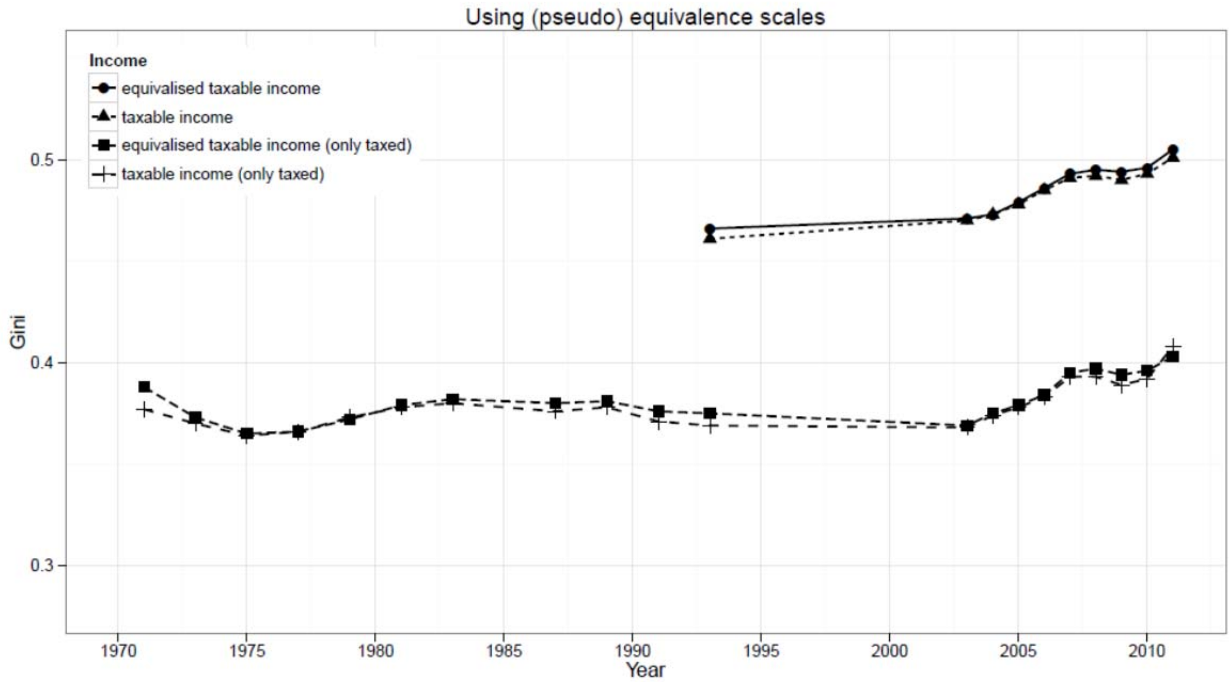

Figure 5: Inequality trends using a tax-based equivalence scale

Source: Tax data-based key figures (FTA)

The implementation of an equivalence scale does not have a major impact on the assessment of inequality (see Figure 5). Over the observed time period, the two lines, which can be compared, move more or less parallel and differ only slightly. Because tax units depict households only approximately, the implemented equivalence scale has conceptual drawbacks.

3.3 Inequality measures

So far we have shown Gini coefficients, the most common measurement of inequality. However, the coefficient has certain restrictions. It is generally acknowledged that the Gini coefficient is more sensitive to the middle part of the distribution and accordingly less sensitive to changes at the extremes. Hence, it is possible to identify periods where inequality increased or decreased, but it is not possible to understand which part of the distribution was affected. To overcome these restrictions, we calculate additional measures (3.3.1) and expand the analysis with relative distribution methods (3.3.2).

3.3.1 Change over time using Gini, Atkinson and Theil

To overcome the restricted focus on the middle part of the income spectrum, we compare the Gini coefficient time series to inequality measures that are more sensitive to other parts of the distribution. For that purpose we calculate the Atkinson index and the Theil index. We choose $\varepsilon = 1$ for the Atkinson and the Theil (GE(α=1)) indices to compare how the development of inequality changes over time, when comparing the middle part-sensitive Gini coefficient to the bottom-sensitive Atkinson index and the top-sensitive Theil index (De Maio 2007). We choose rather moderate variants of the Atkinson/generalized entropy families, because we do not want to focus on the extremes. Cowell and Flachair (2007) show that these measures are very sensitive to high/low incomes when high values for $\varepsilon > 1$ and $\alpha > 1$ respectively are chosen.

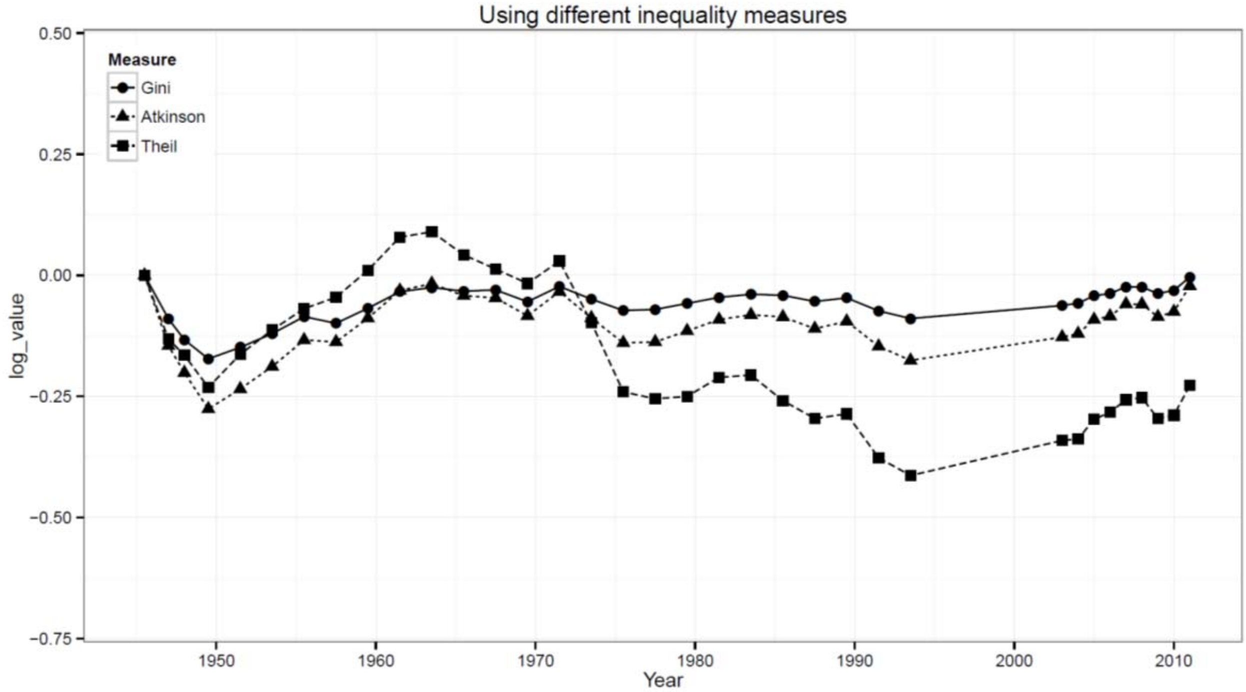

Figure 6: Inequality trends using different inequality measures

Source: Aggregated tax statistics (FTA)

Figure 6 shows the three time series based on taxable income for taxed units published in the aggregated tax statistics. We used the log of the indices and indexed each series to its value in 1945/1946. Anchoring the index makes it impossible to interpret the level of each series, but makes changes over time comparable across series. The trends follow quite a similar pattern, but they differ in volatility. This suggests that the borders of the distribution are much more prone to changes. A pattern, which is probably better revealed with full population tax data that cover the extreme parts of the distribution more precise. Following the strong changes of the Theil index, this is especially true for the upper part of the distribution. During the 1950s and the early 1960s higher incomes grew faster, which resulted in an inflated Theil index. Then in the 1970s and 1990s, the Theil index drops below the other measures, suggesting a relative decline of higher incomes in these periods.

3.3.2 Change over time using relative distribution

The comparison of bottom-, mid- and top-sensitive measures can give a clue to the nature of changing inequality. Even more light is shed on the changing patterns when we expand the analysis by using relative distribution methods (Handcock and Morris 1999). This approach compares probability densities of two populations comprehensively. To review the change of the income distribution over time, we use the published percentiles of the distribution of taxable income from the FTA key figures dataset.[14] By comparing the income distribution of 2011 to that of 2003, we shed light on the area after the post-dotcom bubble crisis, which in Switzerland was followed by a period of steady economic growth and recurring debates on rising salary for top earners. In terms of the Gini

[14] We prefer these measures over the calculated measures out of the published income bracket statistics, because they represent the distribution at both tails more accurately since they are based directly on the information about every single tax unit. When calculating percentiles out of the income bracket statistic we lose relevant information at the edges. First, we do not have information about taxable income of tax units falling below the income threshold for federal taxation (see also Section 3.5.3). We only know how many persons fall in this category. However, the percentiles reported on the FTA webpage are based on the true taxable income (also for units below the threshold), which allows a more precise estimation of the lower percentiles. Secondly, it is especially hard to estimate the highest top income percentiles out of the aggregated tax statistics, leaving us with information only until the 95th percentile, while the reported percentiles reach the 99.99th percentile.

coefficient, inequality rose from 0.47 to 0.50—a moderate increase. The in-depth distributional analysis allows us to see where in the distribution this change occurred.

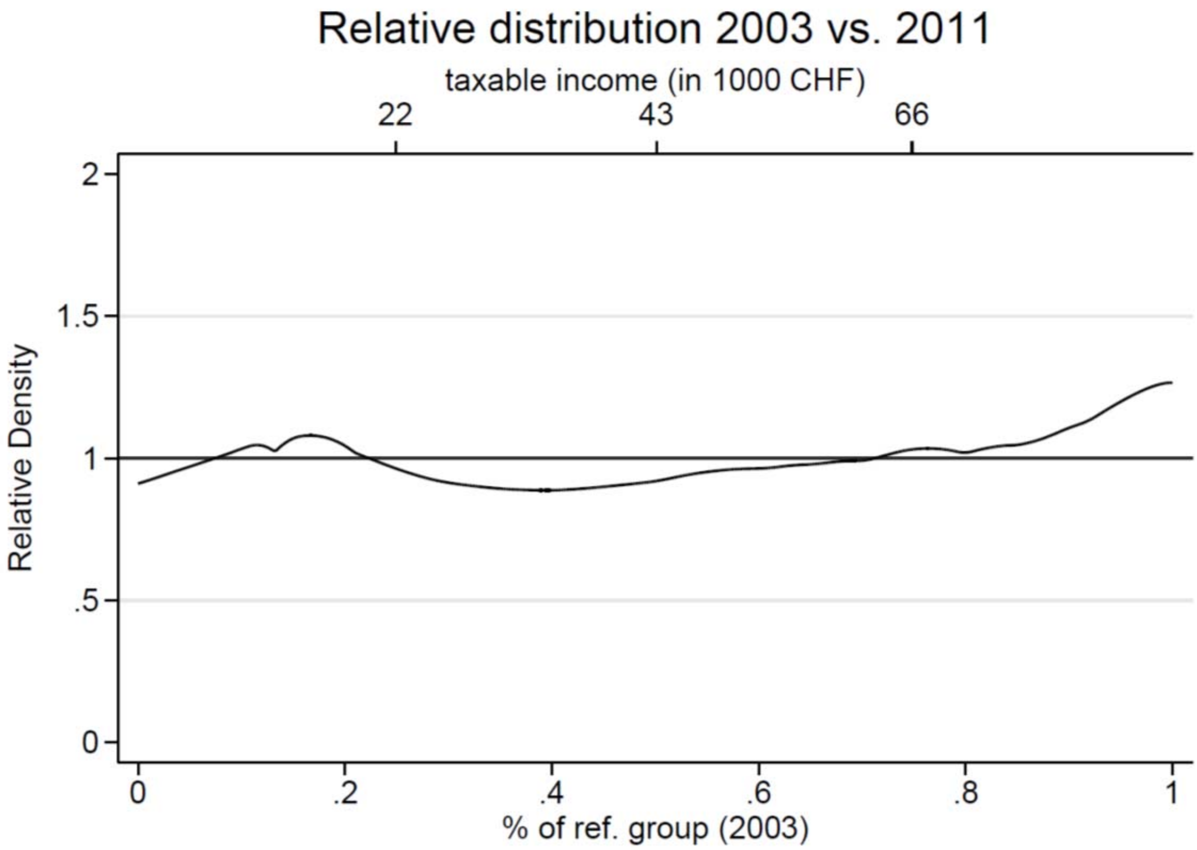

Figure 7: Relative distribution over time

Source: Tax data-based key figures (FTA)

By performing a complete distributional comparison it is clearly visible where changes occurred. When looking at the relative density of the 2011 versus 2003 tax data (Figure 7), a moderate polarization is visible, which is represented in a lower relative density in the middle deciles ($D_{0.2}$ to $D_{0.7}$), while the density ratio is notably higher in the top two deciles but also in the area below $D_{0.2.}$[15] On a substantive level, this analysis shows that the rise of inequality in the post-dotcom bubble area can be attributed not only to an increase of top earners, but also to an increase of units with low incomes. Additionally the full distributional trend analysis shows the importance of complete coverage inequality estimation, as the distributional changes occur at the tails of the distribution. It can be hypothesized that the stable/declining trend reported by surveys is related to estimation with surveys that cover the extreme parts of the distribution inadequately.

3.4 Statistical units

The usual units to assess inequality are households because the possibility of experiencing economic well-being is strongly connected to households (see Section 2.3). In tax data, however, the units are represented according to administrative rules and fiscal households do not necessarily represent true households. It is not straightforward to derive households and household income from tax data. This might influence the assessment of inequality development, taking into account the change from traditional households and family structures over the last century.

[15] We compare full distributions although we work with percentiles. To achieve this, we created data that represent the distribution described by these percentiles, by imputing cases between adjacent percentiles in a linear fashion.

To examine the sensitivity of measuring inequality to the statistical unit, we use micro tax data from the canton Bern. This data includes housing information added from personal registers that allow construction of a household identifier for tax units. Because this register harmonization is fairly new, we can only use data for one time point (2012). Nonetheless, we are able to look at the distribution of taxable income with tax units and then compare it to the distribution when pooling income according to the household identifier. By comparing these two distributions, we can test the sensitiveness of inequality regarding different concepts of statistical units.

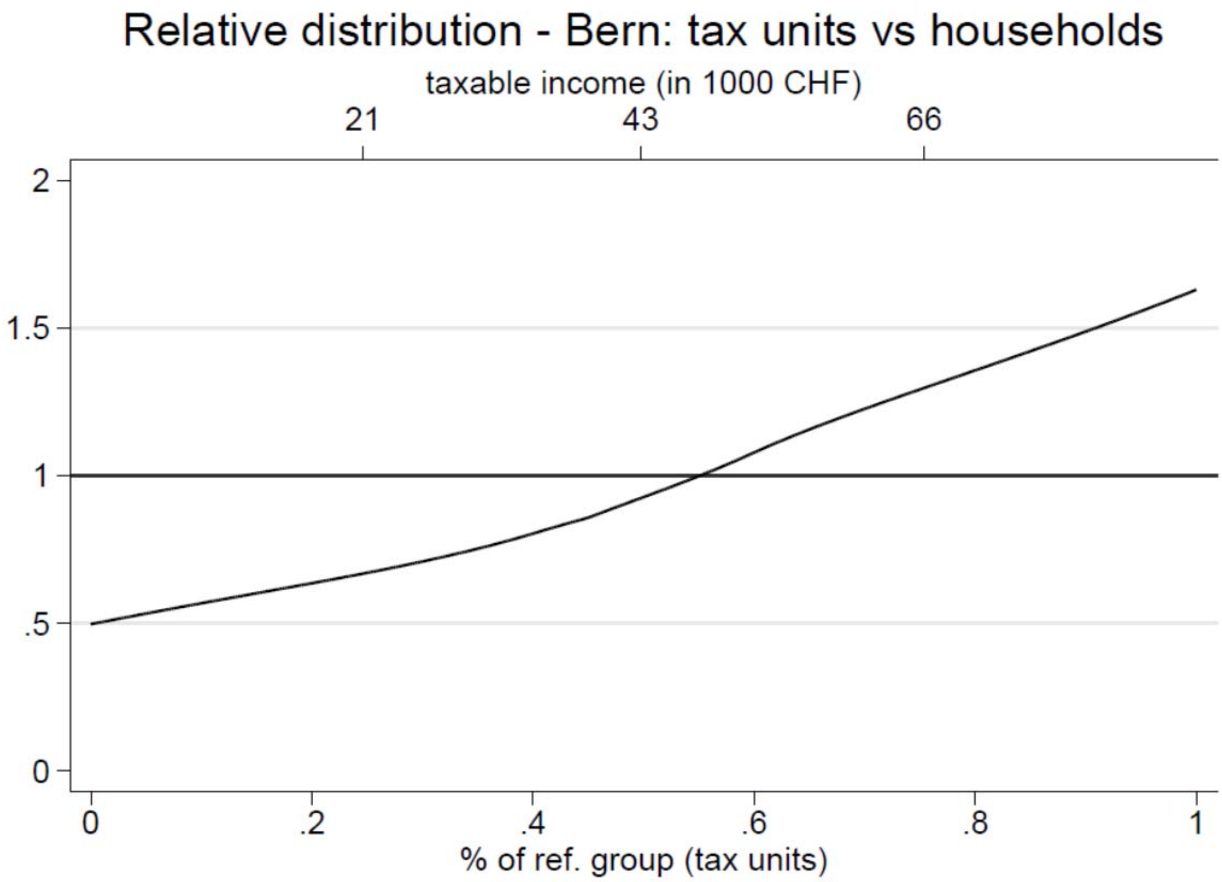

Figure 8: Relative distribution, real over fiscal households

Source: Micro tax data (Bern)

Our test shows substantial higher inequality among tax units (Gini=0.45) than among households (Gini=0.39). This is because the share of persons effectively living alone decreases drastically, when we switch from tax units to real households. Many single-person tax units are not living alone, 66.1% are taxed as single-person tax units although we identify only a share of 36.9% of actual single-person households. This results in pooling of income and an upward shift of former “poor” units. In other words, many units with low income are replaced with fewer units with higher incomes. The related relative distribution illustrates the differences (see Figure 8). In the distribution based on households, lower-income units are underrepresented compared to the distribution based on tax units while there is more mass in the upper part of the distribution.

This mechanism is likely similar for the income distribution of Switzerland derived from the aggregated tax statistics. Looking at the published tax statistics for the year 2011, the proportion of single (62.1%) to married tax units (37.9%) are similar to Bern, meaning that inequality would be lower if assessed on the household level and not among tax units. In addition, we assume that the bias increased in recent decades, and it thus had less influence in times when cohabiting without marriage was less common and the share of tax units corresponding to actual households was bigger.

3.5 Population coverage

While survey samples are suspected to be biased because of nonresponse, the concerns about incomplete coverage are different with tax data. Essentially every permanent resident in Switzerland over 18 years of age (20 years of age prior to 1996) is taxed on a yearly base (or every two years before the change of the tax system). Theoretically this leads to a full representation of the adult population of Switzerland and a complete coverage of the income distribution. Practically, however, tax data distinguishes several subgroups and for some time periods information on certain groups is missing. This can lead to an incomplete representation of the population. First, tax data distinguish normal and special cases. The majority of taxpayers are normal cases; these are tax units residing in Switzerland without foreign-sourced income, liable to taxation for the full year. Special cases include foreign nationals living in Switzerland or individuals who moved to or departed from Switzerland and are therefore not liable to taxation for the whole year. Second, tax statistics separate those who actually pay taxes from those with an income below a threshold that leads to an exemption of direct federal taxes. While information on taxed normal cases is available for longer time periods, information on special cases and non-taxed units are not always reported.

Another source of incomplete coverage within tax data are missing incomes; this includes incomes at the bottom and at the top alike. Incomes at the bottom are not reported properly, because social welfare is not taxed in Switzerland. Income at the top is suspected to be incomplete because of tax evasion. Non-filers are a minor problem, because in Switzerland non-filers are also in the tax statistics as long as they are registered at the local residents' registration office. Incomes are imputed based on older tax returns and employer-reported information. Only non-registered non-filers, like undocumented migrants, are not in the records. An important bias, however, is caused by individuals who misreport incomes. Feld and Frey (2006) examine the role of tax evasion in Switzerland by calculating the difference between the national accounts measures of primary income and the income reported to the tax authorities. They show that the average level of income tax evasion from 1965 to 1995 varies between 13% and 35% and suggest that evasion is heavily driven by capital income tax evasion.

With available tax statistics, we can distinguish three coverage issues with an empirical possibility of testing their relevance for inequality analysis. First, we compare the tax income distribution to survey data, to see if tax data cover extreme incomes more reliably than survey data (3.5.1); then we test if the inclusion or exclusion of special cases has a substantial impact on the assessment of income inequality (3.5.2). Third, (3.5.3) we quantify the extent to which inequality is affected by neglecting those subjects who are not taxed, because their incomes are below the exemption threshold.

3.5.1 Superior coverage with tax data than with survey data

The prevalent scholarly opinion is that tax data cover the extreme parts (lower and upper incomes) of an income distribution more reliably than survey data because the latter suffer sampling error. To test this hypothesis, we perform two tax data comparisons with the Household and Consumption Survey (HBS). This allows us to construct measures that are more comparable to income measures derived from tax data. A successful comparison requires the control of all other relevant differences between tax data and survey data, like differences in income definitions and the fact that HBS represents households and tax data represent tax units. Because it is not possible to construct a perfect comparison, we follow the two best alternative strategies and report results for both:

1. We construct a comparison for the Swiss population for the year 2011, where we use the FTA key figures. To control for the difference of statistical units, we restrict our analysis to married couples. Additionally, we construct a pseudo net income with the HBS that is comparable to the net income from tax statistics. We do this by subtracting social security contributions and transfers to other households from total income (earnings, wealth and direct social transfers). Some differences stemming from fiscal deductions remain, which cannot be reflected within the HBS. Peters (2005) showed that deductions reduced taxable income by almost 30 percent on average. Therefore, it is not surprising that net incomes within tax statistics are

substantially lower on average. We assume that these deductions[16] are proportionally equal across the whole income distribution and hence do not interfere with the comparison. To calculate the relative density we correct this difference with a multiplicative (log of mean) location shift. By adjusting for location differences we are able to analyze potential differences in shape, which is the crucial aspect with respect to distributional inequality. To get a fair benchmark for the tax data distribution, we apply sampling weights.

2. We construct a restricted comparison for the canton of Berne, where we are able to observe both tax units and households, and address the conceptually different statistical units directly. We improve our comparison further by excluding households with more than seven members, which is the highest number within HBS for the canton Berne. We do this to exclude collective households from the comparison, which are by definition not represented within the HBS. We base the comparison on primary income, (a) to get rid of the deductions and (b) to avoid a potential bias from missing information on social welfare, which is not represented in tax data but is in the survey data. As a drawback of this strategy we cannot compare the same years. Tax data represent the year 2012, while the most recent HBS data refer to 2011. We therefore test if the distribution based on tax units in Bern differs between 2011 and 2012. No substantial difference could be identified.

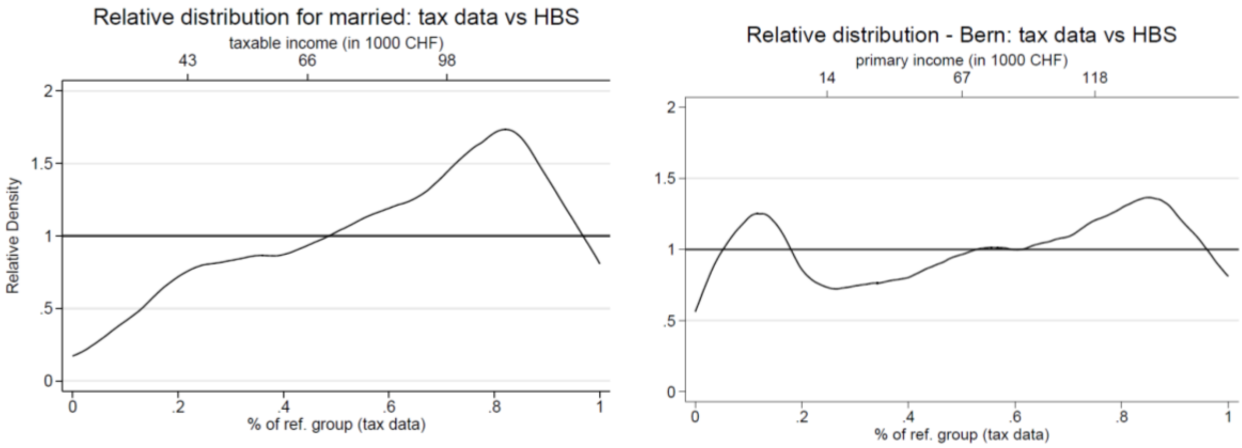

Figure 9: Relative distribution, survey over tax data

Source: Aggregated tax statistics (FTA), micro tax data (Bern) and Household Budget Survey (HBS)

Figure 9 plots the relative density of the HBS distribution (FTA: left, micro tax data: right) with tax data as a reference distribution. The results show a poor overlap of the distributions, which mainly stems from an "upper middle-class bias" within the survey data. This bias seems more pronounced in the plot for married couples than in the plot for Bern. The extreme parts (very rich and poor) are better represented in both plots within tax data. This upper middle-class bias results in an underestimation of inequality. The Gini coefficient for Bern is +0.05 higher in tax data than in the HBS. A comparison of the Gini coefficients for the tax data and HBS for the married couples results in an even higher (by +0.18) coefficient.

3.5.2 Influence of special tax subjects

The question of adequate population coverage for tax data also has to be answered regarding different – rather technical – definitions of tax units. Aggregated tax statistics in Switzerland differentiate between normal and special cases (see Section 2.4). To test the influence of the inclusion of special cases on the income distribution, we compare the distributions of taxable income for normal cases to the pooled distribution (normal and special cases). Unfortunately, the FTA stopped publicly reporting data for special cases after the tax period 1993/94. Therefore we compare two

[16] Not to confuse with the social deductions, which we assume to be fix.

distributions based on aggregated tax statistics for a rather old dataset. However, the FTA key figures do report distributional figures (e.g. percentiles) based on a pool of all cases (normal and special) for more recent periods, which allows us to do a corresponding analysis for 2011 as well.[17]

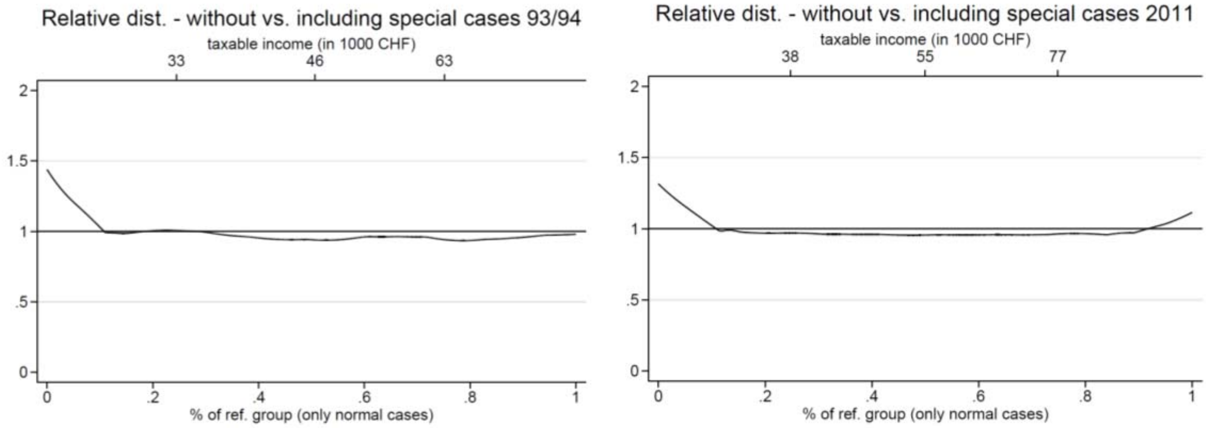

Figure 10: Relative distribution with and without special tax cases

Source: Aggregated tax statistics and tax data-based key figures (FTA)

The pooled dataset of normal and special cases for 1993/94 has a slightly higher density at the lower end compared to data based exclusively on normal cases (see Figure 10 left). Put simply: the population of special cases in 1993/94 holds considerably more tax units with low incomes than does the population of normal cases. For 2011, the picture is similar: Special cases appear more frequent around the lower percentiles of the pooled distribution. However, for 2011 there is an even more remarkable distinction in the upper part of the distribution (see Figure 10 right).

To get a better understanding of the observed patterns, we take a closer look at the special cases subgroups (for detailed definitions see EFD, 2008). First, special cases include individuals who are taxed according to expenditures. More precisely, these are wealthy foreigners who are not employed in Switzerland. These individuals are taxed with special conditions and get an imputed income according to their expenditures. These imputed incomes probably underestimate real incomes, but because they are still higher than average incomes they appear in the upper part of the income distribution. As Table 2 shows, this is a minor group but in the last 20 years their number more than doubled, which supports the hypothesis that rich immigrants led to an increase of inequality in recent years. Inequality also increases with migration at the lower end of the income distribution. There is a larger group of other special cases with diverse circumstances. The most common case is individuals who either moved to or departed from Switzerland and are therefore not liable to taxation for a whole year. Their income in Switzerland is extrapolated to a 12-month income so that their income does not appear artificially low. Other special cases are natives with foreign incomes or foreigners with income in Switzerland. Their incomes represent their true economic situation as taxes are calculated on the base of the incomes they generated in and outside of Switzerland. Lastly, foreigners are also liable to taxes if they own business establishments or property in Switzerland. Because these persons only have to pay taxes for income earned in Switzerland, they appear in tax statistics with lower incomes for technical reasons.

[17] Again it is possible to perform a fully distributional comparison, with a little technical effort. When using aggregated tax statistics we first estimate percentiles via Pareto interpolation (Cowel 2011). Then we create an artificial dataset that represents the distribution described by these percentiles (see also footnote 15).

Table 2: Numbers of taxed normal and special cases 1993/1994 and 2011

	1993/1994		2011	
	abs.	%	abs.	%
normal cases	2'762'419	84.4%	3'152'002	92.0%
taxed according to expenditures	2'730	0.1%	5'530	0.2%
other special cases	506'129	15.5%	267'819	7.8%
Total	*3'271'278*	*100%*	*3'425'351*	*100%*

Source: Aggregated tax statistics from Swiss FTA

All in all, special cases are natives and foreigners who are associated with a foreign country but are nonetheless part of Swiss society and should theoretically be included in the analysis. Their inclusion leads to an increase of income inequality because special cases are strongly polarized, including very low and very high incomes (while even still underestimating high incomes of wealthy foreigners taxed according to their expenditures). In terms of the Gini coefficient, the inclusion of the special cases leads to a moderate increase of +0.02 in 2011.

It has to be mentioned that individuals who are taxed at source are not covered in the tax statistics. These are mainly migrants who live and work in Switzerland but have not yet received a permanent residence permit. These individuals get taxes directly subtracted from their income without filling a tax form. As this is a common case and as these individuals often stay for several years and probably have very diverse incomes, it would be interesting to see how their inclusion would affect the income distribution. Also taxed at source and therefore not included in the tax statistics are individuals who do not have a permanent residence in Switzerland. This includes for example cross-border commuters, consultants, athletes or artists, who earn income in Switzerland while living abroad.

3.5.3 Influence of non-taxed units

From 1995/1996 to 2011 the number of non-taxed units is reported by the FTA, but not for the years before. This means that we are able to quantify the influence of excluding the non-taxed units based on the period from 1995/1996 to 2011.

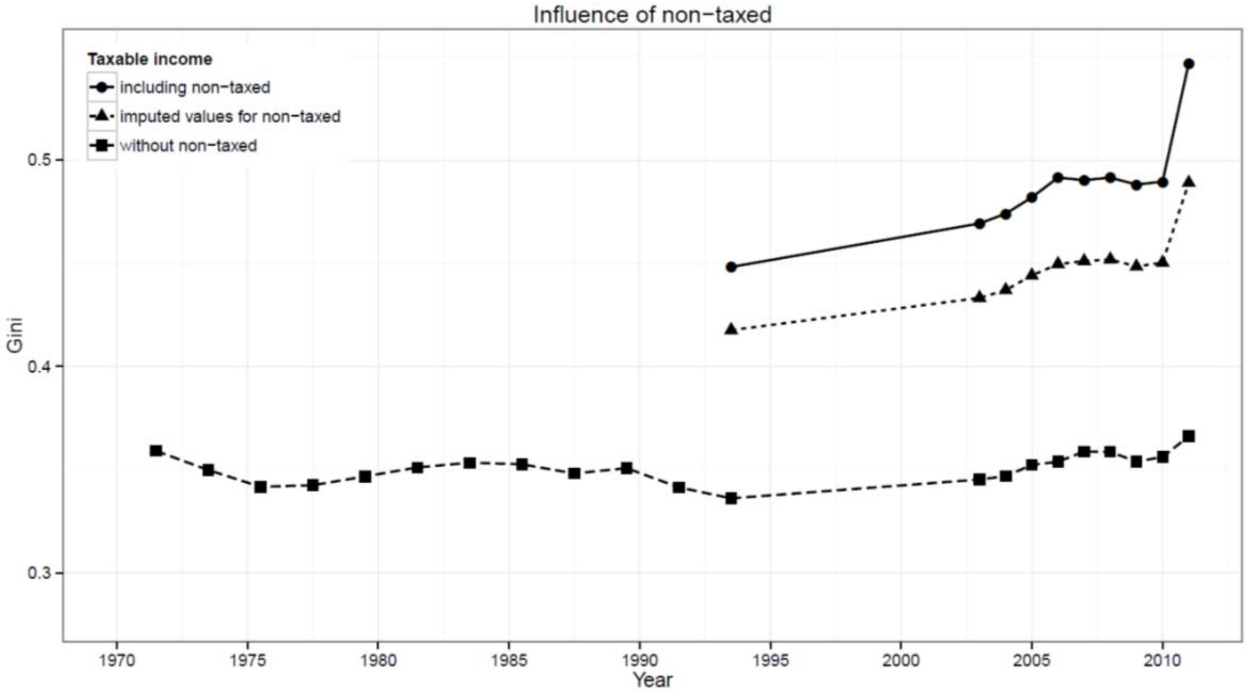

Figure 11: Influence on non-taxed units

Source: Aggregated tax statistics (FTA)

We calculate three Gini time series (see Figure 11). Unsurprisingly, excluding the non-taxed leads to a dramatic drop of the Gini coefficient. At the same time, however, we overestimate inequality by assuming non-taxed tax units have zero taxable income. Rather, we must assume a taxable income between zero and the taxation threshold. We address this by presenting a third time series, where we assume non-taxed units to have a taxable income equal to half the threshold for single tax units.[18] This results in slightly lower, more realistic Gini coefficients.
A second problem related to the exemption threshold is identified through the sharp increase of the Gini coefficient in 2011. Although this rise could be attributed to a more unequal distribution of incomes, fiscal adjustments are another cause of the higher Gini coefficient. This becomes clear when counting the number of non-taxed subjects. In 2010, 906,500 normal tax subjects fell below the exemption threshold, which means that 20.7% of all potential normal tax subjects were not taxed for direct federal taxes. In 2011, however, the number of non-taxed units increased by over 350,000 units to 1,257,075 (28.5% of all tax subjects). This major increase can be explained by the rise of the exemption threshold and the rise of claimed deductions for married couples with children. All in all these fiscal adjustments result in a substantial bigger share of non-taxed units and an artificial increase of the Gini coefficient.

The problem of non-taxed units is worse in earlier tax periods. Although the FTA does not report the share of non-taxed units before 1995/96, Dell et al. (2007) estimated this share from the difference between the Swiss population over 20 (census report) and the number of taxed people. They find the covered part of the population to be lower, the earlier the period. According to their estimates, the share of tax subjects represented in FTA tax statistics varies from 94% in 1993/1994 to 13.7% in 1933. It is highly questionable if analysis based only on a small fraction of the population is appropriate.

[18] We consider only the threshold for single tax units, because married tax units are very seldom exempted from direct federal taxes although the threshold is set at a higher level. We accounted for the variation of the exemption threshold over time. The threshold was raised in 2003 (from CHF 14,900 to CHF 16,100 for single people) and in 2011 (to CHF 17,700).

4 Discussion and conclusion

In this paper we assessed the suitability of tax data to carry out inequality trend research and discussed data related advantages and disadvantages. For the conclusion, it is important to distinguish between aggregated tax statistics and micro tax data. Aggregated tax statistics tabulate tax units according to income brackets. They are available for a variety of countries over long periods of time and they are used to calculate the already established top income shares. Micro tax data, by contrast, here refers to the full fiscal information based on the tax returns of every entity liable to pay taxes. This equates to full population coverage and adequate representation of incomes and taxes that can be used to calculate theoretically sound post-transfer/post-tax incomes. Because the demands of archiving such micro tax data are high, they cover shorter time periods than aggregated tax statistics (micro tax data are available in the US from 1960 onwards, in Switzerland only from the turn of the millennium). While aggregated tax statistics show some crucial conceptual imperfections, micro tax data satisfy the most important requirements of state of the art inequality concepts. In fact, the problem of tax units not necessarily representing households can be resolved by combining tax data with personal register data that allow for generating a household identifier. By doing so it is possible to construct a near to perfect dataset, which we then compared to the major survey used to track income inequality in Switzerland. We were able to show that, indeed, this survey does not cover the whole population adequately, thereby leading to an underestimation of inequality. This shows that establishing a full population representation with samples still poses a serious challenge, even if it is done with great care and the use of adjusting weights. At the same time, in Switzerland micro tax data is not fully available because local government authorities levy taxes, and cantonal privacy law sometimes forbids its provision even for scientific purposes. Finally, it can be supposed that the data linkage required to allow household identification is becoming more and more accessible, which makes – in our opinion – micro tax data the state of the art means to study income inequality. It should not pose any problems, at least in countries where official unique personal identifiers exist, as is the case in the Nordic countries.

Aggregate tax statistics, however, remain the only option to study the long-term evolution of income distribution. The question arises how grave the potential data-driven errors are when using aggregate tax statistics for overall inequality estimations, despite the known imperfections. To answer this question we conducted several analyses on tax data from Switzerland. By estimating the magnitude and direction of assumed biases and advantages, we are able to provide a ranking that helps researchers to differentiate major from minor issues with respect to the assessment of income inequality trends. We build this ranking based on the maximum observed range of Gini coefficients for each section of our analysis:

1. Influence of non-taxed units (max. range of Gini coefficient: + 0.12)
2. Income concepts (+ 0.10)
3. Tax units vs. households (+ 0.06)
4. Avoidance of bias through nonresponse (- 0.05)
5. Influence of special tax subjects (+0.02)
6. Use of income corrected with an equivalence scale based on tax information (+ 0.01)

According to the ranking, the greatest source of bias is related to incomplete information on tax units that fall below the taxation threshold. The level of the threshold influences the share of non-taxed entities, and this has a strong impact on the assessment of income inequality. Fiscal adjustments therefore have a strong influence on inequality measures based on tax statistics. In Switzerland, the number of non-taxed units has at least been reported since 1995/96. For the period from 1933 to 1995/96, Dell et al. (2007) estimated how well tax statistics cover the whole Swiss Population. They show that a minimum of three quarters of the Swiss population has been covered since the 1970s. The taxed population of earlier periods, however, only represents a fraction of the population of interest. As shown, for example, by Atkinson et al. 2011, taxation schemes evolved similarly in most western societies: starting with strongly progressive taxation and high exemption thresholds that lead to the initial coverage of only a minor group. This is a hurdle that top income studies are able to overcome by using external estimates of total population and total income. The second biggest source of bias refers to the importance of the income concept used for distributional analysis. Taxable income, the key income concept within tax data, is neither a pre- nor a post-transfer income measure, but something in between (usually direct social transfers, with the exception of mean-tested benefits, are accounted for, while redistribution through the tax system is not). Our analysis showed that the bias induced by deductions outweighs even the bias of ignoring paid taxes. Of course, the amount of the introduced bias is tied to the specific legislation of particular countries. Our analysis showed, however, that the estimation of income inequality is strongly influenced by the taxation scheme in two ways (deductions and missing information on taxes). The third important bias that results from using tax data is that statistical units are fiscal and not real households, so that in the case of cohabitation, for instance, considering the individuals as two separate tax units also leads to an overestimation of inequality and a bias in the inequality trend, as the "cohabiting-to-married-ratio" increases over time in most western countries. This is a general problem of tax data regardless of whether the taxation system targets individuals or families. The fourth point relates to problems stemming from nonresponse. Leaving aside the issue of non-taxed subjects in this regard, tax data are superior. Our analysis showed that the distributions of tax and survey data differ substantially – even if key methodological differences are controlled for. We claim that this difference stems from the under-representation of very poor and very rich households in survey data, which leads to an underestimation of inequality when working with survey data and to "blind spots" in crucial parts of the distribution. In comparison to the other issues, the influence of special tax subjects and the implementation of the equivalence concept tailored to tax data are rather minor issues. We showed, however, that the inclusion of special cases is necessary to catch the effect of special socio-political developments, such as the recent immigration of rich individuals to Switzerland, who get tax privileges by getting taxed according to expenses. This shows that researchers should be careful about tax laws that divide the population into several subgroups. Researchers have to clarify whether or not specific groups theoretically belong to the group of interest and check if these groups are represented adequately.

The estimated differences give a direct overview of biases related to the Gini coefficient time series based on aggregate tax statistics from Switzerland, but they should not be used to adjust results from other data sources because the reported differences are related to the used data-sets[19]. We nonetheless believe that the ranking can be generalized to other cases by giving an overview of what factors are and are not potentially influential.

A special section of the discussion is dedicated to inequality measures, as the tests performed cannot be included in the above ranking since measures other than the Gini coefficients were used and comparability in the sense of the ranking is therefore not suitable. Nonetheless, the analysis showed that all relevant statistical techniques can be applied to aggregated tax statistics. Furthermore, we showed that trend analysis is indeed influenced by the measurements chosen. The top-sensitive Theil index suggests more volatility in the upper part of the income distribution over the observed period than the series based on the middle-sensitive Gini coefficient and the bottom-sensitive Atkinson index. Single indices that conflate information to a single measure drastically reduce information, while distributional analysis with relative distribution methods allows the precise area of change to be located, but only refers to two single time points. Trend analysis is therefore best done by combining several one-population measures that are sensitive to different parts of the distribution for a first analysis of time patterns. In a second step, the analysis is enriched through relative distribution methods for specific time points to unravel complete distributional differences.

[19] E.g., the difference of the Gini coefficient between the distribution based on tax units compared the distribution of households is affected by the degree to which tax units actually mirror households. It is expected that the bias gets even stronger with the increasing trend towards cohabitation without marriage. Therefore, reported biases vary over time and probably also between countries.

The time series displayed in Figure 2 on page 8 showed inconsistent findings on income inequality trends in Switzerland. Given the results of the methodological tests performed, is it possible to solve this contradiction? Keeping the mentioned imperfections in mind, we know that none of the Gini coefficients displayed are perfectly valid. Most of the factors outlined above imply an overestimation of income inequality based on aggregated tax statistics. At the same time, income inequality measures calculated using survey data underestimate inequality due to nonresponse. Both effects explain why the overall level of income inequality is higher with tax data. The truth probably lies between the presented series from tax data and survey data. With respect to the length of the time series, tax data clearly outperforms survey data. While most imperfections of aggregate tax statistics are rather constant over time, the missing information on non-taxed units varies and therefore introduces a bias to the trend. Following the estimates of Dell et al. (2007), it is not recommended to start interpreting the tax data bases time series before 1970s. The evolution of income inequality directly after World War II is at least plausible. This period was characterized by strong economic growth and an increase in income inequality. It can be assumed that high income percentiles disproportionally profited from the economic upturn. After the oil crisis in 1972, there were alternating periods of economic up- and downturns and the expansion of social welfare began – a period during which income inequality remained quite stable. An interesting period began around the millennium, for which the figures based on tax data can be compared to the results from the major surveys, and the trends clearly diverge. Figures based on survey data suggest a decline in income inequality, while the time series based on tax data indicate an increase. By analyzing the relative distribution of 2011 in comparison to that of 2003 (see Figure 7 on page 15), we have been able to show that a polarization occurred that was driven by the downgrading of low incomes as well as by an increase of top incomes. Since these parts of the income distribution are better covered within tax data than within survey data, whether the recent trend is really a decreasing one, as the analysis of the LIS-data performed by Gornick and Jäntti (2013) has suggested, is open to doubt.

5 Appendix

Table 3: Overview of empirical tests within inequality-related methodological areas

Methodological area	Empirical test	Method	Data
Income concepts	(1) Income definitions within tax data	Time series of Gini coefficients (own calculation)	Aggregated FTA tax statistic – normal cases without non-taxed units – different income measures. Micro tax data from Berne canton
	(2) Using income corrected with an equivalence scale based on tax information	Time series of Gini coefficient (provided)	FTA key figures – all tax units and without non-taxed units – taxable income
Inequality measures	(3) Change over time: difference between one-population measures	Time series of Gini coefficients, Theil and Atkinson indices (own calculation)	Aggregated FTA tax statistic – normal cases without non-taxed units – taxable income
	(4) Change over time: one-population measure vs. relative distribution	Gini differences (provided), relative distribution (own calculation based on provided percentiles)	FTA key figures – all tax units – taxable income
Statistical units	(5) Tax units vs. households	Gini differences, relative distribution (own calculation)	Micro tax data from Berne canton – all tax units – taxable income
Population coverage	(6) Population coverage with tax data compared to survey data	Gini differences, relative distribution (own calculation)	Micro tax data from Berne canton and subsample for Berne from HBS – primary income
	(7) Influence of special tax subjects	Gini differences, relative distribution (own calculation partly based on provided percentiles)	Aggregated FTA tax statistics and FTA Key figures – all tax units – taxable income
	(8) Influence of non-taxed units	Time series of Gini coefficients (own calculation)	Aggregated FTA tax statistics – normal cases with and without non-taxed units – taxable income

6 Bibliography

Alderson, A. S., & Doran, K. (2013). How Has Income Inequality Grown? In J. C. Gornick & M. Jäntti (Hrsg.), *Income inequality: economic disparities and the middle class in affluent countries*. Stanford: Stanford University Press.

Alvaredo, F., & Saez, E. (2009). Income and Wealth Concentration in Spain from a Historical and Fiscal Perspective. *Journal of the European Economic Association*, *7*(5), 1140–1167.

Alvaredo, F., Atkinson, A. B., Piketty, T., & Saez, E. (2015). The World Top Incomes Databas. http://topincomes.g-mond.parisschoolofeconomics.eu/(13/04/2015)

Atkinson, A. B. (1975). *The economics of inequality*. Oxford: Clarendon Press.

Atkinson, A. B. (2013). Foreword. In J. C. Gornick & M. Jäntti (Eds.), *Income inequality: economic disparities and the middle class in affluent countries*. Stanford: Stanford University Press.

Atkinson, A. B., & Piketty, T. (2007). *Top incomes over the twentieth century a contrast between continental European and English-speaking countries*. Oxford; New York: Oxford University Press.

Atkinson, A. B., & Piketty, T. (2010). *Top incomes: a global perspective*. Oxford: Oxford University Press.

Atkinson, A. B., Piketty, T., & Saez, E. (2011). Top Incomes in the Long Run of History. *Journal of Economic Literature*, *49*(1), 3–71.

Atkinson, A. B., & Salverda, W. (2005). Top Incomes in the Netherlands and the United Kingdom over the 20th Century. *Journal of the European Economic Association*, *3*(4), 883–913.

Bethlehem, J., Cobben, F., & Schouten, B. (2011). *Handbook of Nonresponse in Household Surveys*. Wiley.

Brewer, M., Muriel, A., Phillips, D., & Sibieta, L. (2008). Poverty and Inequality in the UK: 2008. London: Institut for Fiscal Studies.

Buhmann, B., Rainwater, L., Schmaus, G., & Smeeding, T. M. (1988). Equivalence Scales, Well-Being, Inequality, and Poverty: Sensitivity Estimates Across Ten Countries Using the Luxembourg Income Study (lis) Database. *Review of Income and Wealth*, *34*(2), 115–142.

Burkhauser, Richard V., Shuaizhang, Feng, Stephen P., Jenkins, & Jeff, Larrimore. (2012). Recent Trends in Top Income Shares in the USA: Reconsiling Estimates from March CPS and IRS Tax Return Data. *The Review of Economics and Statistics*, 94(2), 371–388.

Cowell, F. (2011). *Measuring inequality*. Oxford: Oxford University Press.

Cowell, F. A., & Flachaire, E. (2007). Income distribution and inequality measurement: the problem of extreme values. *Journal of Econometrics*, *141*(2), 1044–1072.

Dell, F., Piketty, T., & Saez, E. (2007). Income and wealth concentration in Switzerland over the 20th century. In *Top Incomes Over the Twentieth Century: A Contrast Between Continental European and English-Speaking Countries* (pp. 472–500). Oxford: Oxford University Press.

De Maio, F. G. (2007). Income inequality measures. *Journal of Epidemiology and Community Health*, *61*(10), 849–852.

EFD. (2008). WEISUNG des EFD über die Erhebung und Lieferung der erforderlichen Daten durch die Kantone gestützt auf Artikel 22 der Verordnung vom 7. November 2007 über den Finanz- und Lastenausgleich.

EUROSTAT. (2015). European Union statistics on income and living conditions (EU-SILC). Retrieved from http://ec.europa.eu/eurostat/web/microdata/european_union_statistics_on_income_and_living_conditions

ESTV (2014). *Verteilung des Wohlstands in der Schweiz*. Bericht des Bundesrates in Erfüllung des Postulats 10.4046 von Jacqueline Fehr vom 07.12.2010. Bern.

Feld, L. P., & Frey, B. S. (2006). Tax Evasion in Switzerland: The Roles of Deterrence and Tax Morale. In *Working paper series: Institute for Empirical Research in Economics* (Vol. 284, pp. 123–153). Zürich.

Foellmi, R., & Martinez, I. (2014). *Volatile Top Income Shares in Switzerland? Reassessing the Evolution between 1981 and 2009* (SSRN Scholarly Paper No. ID 2501458). Rochester, NY: Social Science Research Network.

Gornick, J. C., & Jäntti. (2013). *Income inequality: economic disparities and the middle class in affluent countries*. Stanford: Stanford University Press.

Handcock, M., & Morris, M. (1999). *Relative Distribution Methods in the Social Sciences (Statistics for Social Science and Behavorial Sciences)*. Springer.

Hao, L., & Naiman, D. Q. (2010). *Assessing Inequality*. California, New Delhi, London, Singapore: SAGE.

Korinek, A., Mistiaen, J. A., & Ravallion, M. (2006). Survey nonresponse and the distribution of income. *The Journal of Economic Inequality*, *4*(1), 33–55.

Kuznets, S. (1955). Economic growth and income inequality. *The American Economic Review*, *45*(1), 1–28.

Modetta, C., & Müller, B. (2012). Einkommensungleichheit und staatliche Umverteilung. *Bundesamt für Statistik, Neuchâtel*.

OECD. (2008). *Growing Unequal? Income Distribution and Poverty in OECD Countries*. OECD Publishing.

OECD. (2011). *Divided We Stand. Why Inequality Keeps Rising*. OECD Publishing.

OECD. (2013). *OECD Framework for Statistics on the Distribution of Household Income, Consumption and Wealth*. OECD Publishing.

Peters, R. (2005). *Effet des déductions du l'impöt fédéral direct des personnes physiques*. Bern: Administration fédéral des contributions AFC.

Piketty, T. (2001). *Les hauts revenus en France au XXe siècle: inégalités et redistributions : 1901-1998*. Paris: Grasset.

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

Piketty, T., & Saez, E. (2003). Income inequality in the United States, 1913–1998. *The Quarterly Journal of Economics*, *118*(1), 1–41.

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

Salverda, W., Nolan, B., Checchi, D., Marx, I., McKnight, A., & Toth, I. G. (2014). *Changing Inequalities in Rich Countries: Analytical and Comparative Perspectives*. Oxford: Oxford University Press.

Stiglitz, J. E. (2012). *The price of inequality*. New York; London: W.W. Norton & Company.

United Nations. (2011). *Canberra Group. Handbook on Household Income Statistics*. Geneva: United Nations.

Wilkinson, R. G., & Pickett, K. E. (2009). Income inequality and social dysfunction. *Annual Review of Sociology*, *35*, 493–511.

World Economic Forum. (2013). *Global Risks 2013. Eighth Edition.* Geneva: World Economic Forum.