|  |
| --- |
| Are tax data suitable to assess inequality trends?  Income Inequality in Switzerland from 1945 to 2011  Oliver Hümbelin  Bern University of Applied Sciences  Rudolf Farys  University of Bern  January 2015  *Abstract*  In many countries inequality trends are ambiguous, because different methodological approaches blur the picture or reliable data is not at hand. In this paper we assess whether tax data is suitable for inequality trend analysis. Because concepts used with tax data are strongly data-driven we first compare tax data specific concepts to the current state of the art of inequality analysis. Second, we estimate the impact of tax data related methodological options on inequality with aggregated tax statistics from the Swiss Federal Tax Administration (FTA). While there are clear advantages of tax data like long-term availability and reliable population coverage, there are also drawbacks that lead to an overestimation of inequality or hinder comparability over time. All in all, tax data is a data source that should only be used with care but nonetheless is an indispensable source for inequality analysis. For Switzerland our tax data analysis puts strongly into question the declining inequality trend reported by survey data for the last decades. |
| **Berner Fachhochschule**  Soziale Arbeit |

# Introduction

Economic resources might be seen as central indicators for life chances. Therefore, the distribution of resources does not only matter regarding consumption, but also regarding health and even life expectancy. 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 more and more expressed by global leaders (World Economic Forum, 2013). Although inequality rose not uniform, a common pattern seems to be identifiable, which is generally described as the “hollowing of the middle class” (Alderson and Doran, 2013). Households are moving towards the top and the bottom of the distribution. This is especially problematic as the middle class stands at the core of western democracies or, as it is stated by Stiglitz (2012, 117) that 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. On the one hand, Atkinson (2013) observes improved methods regarding household surveys, the core sources of inequality research. On the other hand, the labor intensive and expensive surveys are subject to budget cuts around the world and they suffer from low response rates, which affect the assessment of inequality undisputedly. These concerns have led to the search of alternative data sources that can supplement the established survey data studies. The technological progress and the modernization of public administration have led to several inequality relevant administrative registers like personal income or social benefit records. Especially interesting is tax data, because records reach relatively far back in time. Already Kuznets (1955) used tax data to examine the relationship between economic growth and the distribution of personal incomes. Then, it took several decades until Piketty (2001, 2003) and Piketty and Saez (2003) made the use of tax data fashionable again. Following Piketty’s approach, studies have been conducted in several countries (Atkinson and Piketty, 2007, 2010). Today, all existing top income tax statistics based time series are collected and accessible through the world top incomes database, some of which date back to the beginning of the 20th century (Alvaredo et al., 2014).

While there is already an extensive body of literature with top income based taxed data studies, the utility of tax data for overall inequality studies is not yet discussed thoroughly. In this paper we therefore provide a theoretical and an empirical review of tax data. To assess the use of tax data theoretically we describe the current standards for measuring economic inequality in section 2. This

constitutes a common ground for the evaluation of the suitability of data sources in section 3, where we analyze the theoretical advantages and shortcoming of tax data with a comparison of tax data and survey data. In section 4 we introduce Switzerland as an interesting example for the empirical section 5. In this section we calculate the impact of tax data specific options on inequality with federal tax statistics from Switzerland. This allows us to distinguish major from minor methodological issues, which are summarized in section 6 to show which methodological aspects inequality researchers working with tax data should treat with care.

# Standards on Assessing Economic Inequality

## Concepts on measuring economic resources

Most studies on inequality focus exclusively on income inequality. However, recent activities emphasize that a broader conceptualization is necessary. A recent publication from the OECD (2013) condensed these ideas and created the ICW framework (income, consumption and wealth), which is an internationally agreed framework on micro-level statistics[[1]](#footnote-1). According to the framework it is best to look at income, consumption and wealth as three separate but interrelated dimensions of people’s economic well-being. To gain policy relevant insight, the framework recommends looking at all three distributions simultaneously. Some households with low income, for example, may report adequate levels of consumption expenditure or wealth holdings, or households with no wealth may have adequate income and consumptions.

Because inequality in the distribution of income gets most scholarly attention, we have a closer look at the definition of income. Although terminology slightly differs, common concepts are identified (for detailed discussion see: OECD (2013, 44) and United Nations (2011, 24)]. In Figure 1 we present a stylized framework, which includes a distinction of common income sources. Most people get an income from labor and some get an income from property. These incomes are a direct product of the market outcome and the sum of them is called the primary income. But households not only rely on their primary income. Redistribution through social transfers has also to be accounted for. This includes transfers paid (taxes and direct inter household transfers) and transfers received (pensions, social security insurances and transfers from other households). Redistribution eventually leads to disposable incomes, which finally determines the possibility to consume. The assessment of income inequality is strongly influenced by the definition of income itself. Primary income or disposable income for example differ by substantial meaning and by the expected degree of inequality. Because not every member of a society participates in market activities - may it be with labor or capital – the distribution of primary incomes is usually more unequal than the distribution of disposable incomes, where redistribution through taxes and direct social transfers took place.

In addition, for research purpose 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 subsection statistical unit below).



Figure 1: Income definitions from primary to disposable income  
*Source: OECD (2013:44), own diagram*

## Measuring inequality

Nowadays a plethora of inequality measures exists. Hao and Naiman (2010) provide a good overview on inequality measures and their properties. Based on their overview, we distinguish five major families: Measures relating to probability distribution, measures based on quantile functions and Lorenz curves, measures derived from social welfare functions, measures from information theory and measures based on relative distribution. Many scholars discuss properties of measures usually with regard to five principles.[[2]](#footnote-2) Here we don’t discuss the usefulness of inequality measures in regard to these five principles exhaustively. We rather want to highlight the main advantages and shortcomings of the most used and prominent inequality-measures.

Measures related to *probability distribution* consist of variance or variance based measures. These measures are straightforward to calculate, but the disadvantage is that coefficients theoretically can take any value between zero and infinity. Hence, it is difficult to identify reasonable levels. Additionally, some measures (Range, log variance and variance of logs) the weak principle of transfer, first introduced by Dalton (1920). The principle states that a transfer from a richer person to a poorer person, other things being equal, should lead to a reduction of inequality. As this is a central property of inequality measures, these variance based measures should be avoided.

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. The Gini coefficient is also often used for international comparison. As it is derived from the Lorenz curve, the quantified amount of inequality can unpretentiously be described in a formal and visual way. Therefore the Gini coefficient is easy to understand. Furthermore it has several desired statistical properties (Hao and Naiman, 2010:64). (1) “Principle of population”: the assessment of inequality is independent of the population size; (2) “scale invariance”: the measure is sensitive for changes of income shares, but not for absolute changes (e.g. doubling of all income) and (3) the already mentioned “weak principle of transfers” or “requirement of Dalton”. 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 (see following discussion of Atkinson index). 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 also be problematic for comparison between countries or over time. In extreme cases two totally different distributions share the same Gini coefficient.

Another widely used measure is the Atkinson index. It derives from the *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 (sensitivity in the middle of the distribution). Therefore, his 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). However, the flexibility of the sensitivity parameter comes with the need to justify the chosen value thoroughly.

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 very beneficial property of the Information theory based measure is that they are decomposable; that is, they can be broken down to component parts (i.e. population subgroups). This enables analyses of between‐ and within‐group effects.

While the aforementioned measures describe inequality referring to one population, polarization indices follow a different approach by comparing two distributions via the concept of *the relative distribution* (Handcock and Morris, 1999). This allows the researcher to analyze distributional differences or changes over time in a more meaningful way than it is possible with single distribution measure. Even if these measures register increasing inequality over time, it is not possible to distinguish a polarization of the distribution (increases in both tails and thinning of the middle) from upgrading (relative increases in the upper tail) or downgrading (relative decreases in lower tail). The polarization index developed by Handock and Morris (1999) addresses this issue.

Given this plethora of inequality measure, how should the right one be chosen? In general, the decision should consider the research question, but broadly speaking, to rely on just one inequality measure is not recommendable. Already classical works from Kolm (1969) to Atkinson (1970) to Sen (1973) warn us against relying on a single inequality measure. Using just a few different inequality measures as suggested by Shorrocks and Slottje (2002) can effectively and accurately compare inequality across populations or over time. More generally, Hao and Naiman (2010:62) suggest paying attention to the middle, to the lower region and the two extremes of a distribution.

## Statistical Units

Commonly, households and not individuals are the statistical units for inequality analysis (OECD, 2013, 60). Indeed, although individuals have an income, own assets and experience economic well-being, their possibility to do so is strongly tied to the concept of household. A household is defined by all persons under the same housing arrangement. The basic underlying assumption for collecting data on the household level instead of the individual level, is 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 of living for every household member (also non-earning members, like children). Additionally, there are economies of scale for people sharing living space and commodities. When comparing the individual economic well-being among individuals living in different households, usually equivalence scales are used as already mentioned above.

## Coverage Issues

Generally, inequality studies try to make a statement about the whole population of interest (e.g. nation). It is mostly resources and/or options of researchers that determine whether such a venture has success, as these restrictions shape the way data is collected. When total population data is 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). 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). However, in some countries register data on income may be incomplete, which may exclude significant proportions of the population. Compilers of income data from administrative data therefore should be aware of the shortcomings of their data.

# Comparison of tax data and survey data – overview of advantages and shortcomings

To define a standard of measuring economic resources and related inequality we introduced four central areas researchers need to address. Ideally we want to (1) look at income, wealth and consumption together, (2) do that for disposable resources on household level, (3) have data suitable to calculate all types of inequality measures and (4) calculate an unbiased estimate of a chosen inequality measure. Table 1 compares tax and survey data on these four dimensions and adds a fifth dimension *possibility to assess inequality trend,* which is not a general need but an implication, when trends are of interest.

Table 1: Comparison of tax-data and survey data

|  |  |  |
| --- | --- | --- |
|  | **Tax-Data** | **Survey-Data** |
| Concepts of economic resources and definition of central measures | data-driven | theory-driven |
| Calculation of inequality measures | restricted | flexible |
| Statistical unit | tax units | households |
| Coverage problems | tax evasion, non-taxed | sample bias |
| Possibility to assess trends of inequality | long | short |

With tax data *concepts of economic resources and respective definitions of central measures* are data-driven, because tax data is collected for administrative purpose. Furthermore, in many countries tax statistics are only available in aggregated form showing tax units per taxable income/wealth brackets without any information on individuals. The missing link on the micro level implies therefore that there is no possibility to do a conjoint analysis of income and wealth. In addition, information on consumption is missing at all. This leaves the researcher with the option to look at income or wealth, but not at both simultaneously. The definition of central measures is also often restricted, because only tax relevant measures are reported. Taxable incomes for example include direct social transfers (e.g. rents) but no taxes are subtracted. Thus, a researcher using tax data can neither look at a pre- nor a post-transfer measure (see Figure 1 on page 5). Taxable income is rather something in between. Furthermore, deductions impose changes to income measures, which can bias the result, when deductions change over time. Concerning this dimension, survey data is clearly superior, because concepts and measures can be tailored carefully to the need of scientists.

The *calculation of inequality measures* is flexible when data is available on a micro level (like it is mostly the case with survey data). If a researcher has to deal with aggregated tax data, however, calculation of inequality measures is restricted. First, the precision of the measures suffers (depending on the degree of aggregation). Second, it is not possible to decompose the measure by features on the micro level (e.g. income source or characteristics of the household). But all common measures (like the Gini coefficient or Theil Index) are still possible to calculate, even though calculation can be tedious.

When looking at *statistical units,* a second drawback of tax data occurs. The statistical units of tax data are tax units (i.e. singles or married) or fiscal households, but these do not necessarily correspond to real households. Indeed, there are constellations where members of the same household submit several tax forms. A common case is an unmarried couple living together. With the change of household-structures and increasing alternative lifestyles this issue becomes increasingly important. Here again surveys are usually able to address the ideal statistical unit in a more appropriate way.

A closer look is necessary regarding *coverage issues*. As mentioned, nonresponse is a general problem of samples and major issue, when working with income data. As Korinek et al., 2006 show, the position in 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. This mechanism can be referred to as the ”middle- class bias” (Diekmann, 2009). Missing data in household surveys is therefore not missing at random, which has an impact on the measures of inequality. Several authors (Särndal et al., 2003) discuss weighting strategies to handle this kind of bias , but they all require a register with information on the complete income distribution, which is rarely available. In contrast, tax based statistic provide total or near-total population coverage. Compared to surveys they are not subject to sampling bias. They may, however, suffer from under-coverage or missing data as well. In many countries tax data is only available for people who file their taxes. Therefore a significant proportion of the population is missing. Another critical issue is tax evasion, which can definitely bias the assessment of inequality. Alvaredo and Saez (2009) for example consider estimates of Spanish top incomes prior to 1981 as unreliable due to widespread tax evasion. Evasion occurs, when individuals try not to fill tax returns or by misreporting of incomes.

The main advantage of tax data is the *possibility to assess trends of inequality*. This makes it an interesting data source albeit the mentioned restrictions. For several countries the availability of tax records reaches back in time for over 100 years allowing to asses time trends that cover substantially longer periods than it is possible with survey data. Nonetheless, when it comes to comparison over time, scientists have to be aware to test the comparability, because measures and population might be affected by changes in the tax systems or the way tax statistics are reported.

# Different trends for income inequality in Switzerland due to methodological differences?

As we will show, results on income inequality for Switzerland are particularly contradictory, which makes it interesting case to have a closer look at methodological aspects. Looking at official data, 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 1 shows Gini coefficients of equalized disposable income calculated from these three sources. 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 EFTA countries. As a Non-EU member, Switzerland implemented the instrument not from the beginning (2004) but from 2007 on. Therefore, this times-series does not cover time periods before 2007. Following 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 lays 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 1, the trend is rather stable. Both time-series (EU-SILC and HBS) cover a relatively short time period. A longer period is covered in the LIS-Data-set (1982-2004). In contrast to the aforementioned surveys, the LIS-data is harmonized out of three surveys: Swiss Income and Wealth Survey (1982), Swiss Poverty Survey (1992) and the Income and Consumption survey (2000, 2002 and 2004). All in all the LIS dataset contains 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, contradictory to the development in most other western countries. This result is supported by Grabka and Kuhn (2012) who analyzed the Swiss Household Panel (2000-2009).

Figure 2: Income inequality trends in Switzerland  
*Source: Luxembourg Income Study (LIS), Household Budget Survey (HBS), European Union Statistics on Income and Living Conditions (EU.SILC), The World Top Incomes Database (top income shares)*

Whereas the aforementioned publications focused on disposable household income from survey data, the revival of tax-data-inequality studies lead to fruitful insights for Switzerland as well. Dell et al. (2007) used tax data from the Federal Tax Administration to assess how the concentration of the highest incomes (top-shares) developed between 1913 and 1996. In contrast to most other examined countries, Switzerland did not experience a reduction in income and wealth concentration from the pre-First World War period to the decades following the Second World War (up to 1996). Using the same approach, Foellmi and Martínez (2013) expand the Dell et al. time series up to 2008 and found out that the share of top incomes did rise. The top 0.01% share even doubled in the last observed 20 years. These results from the top income studies seem to oppose the figures of official data.

To sum it up: survey studies suggest a declining trend in income inequality while top-share studies argue that the concentration of income at the top of the distribution is rising, suggesting that inequality indeed did rise. Differences can be explained with factors introduced in section 2 and 3. First, coverage of top incomes is assumed to be better within tax data than within survey data (non-response bias), which is a crucial issue concerning inequality not only regarding top income shares. Second, different measures of inequality hamper comparability. While some argue that top income shares are an appropriate proxy for overall inequality, others disagree. Following Leigh (2007:600) “top income shares are far from perfect as a measure of distribution of income across society”, although he finds a strong positive correlation with other inequality measures. Third, different income concepts were used. The top income studies work with taxable incomes while the surveys rely on disposable income. As Modetta and Müller (2012) have shown, the income distribution is strongly affected by governmental redistribution through social transfers and taxes, reducing inequality substantially. With the focus on taxable income the change in taxing policy (like tax-rates and tax deductions on sub federal level) is not directly represented. Fourth, the statistical units within tax data are fiscal households and not real households, which again are the base of analysis for the survey studies. With a trend to unmarried cohabitation this could lead to a bias within tax data.

# Testing the stability of income inequality measures with tax data from Switzerland

As discussed above, using different data sources and different concepts leads to substantially different conclusions. In this section we have a closer look at methodical options regarding the four areas introduced in section 2 (defining economic resources, measuring inequality, statistical units and coverage Issues) when working with tax data. We discuss in what way theoretical concepts can be addressed with tax data in Switzerland and we provide empirical stability tests to sort out the crucial topics within the four introduced methodical relevant areas. By doing so, we show which issues are relevant when working with tax data within a more general perspective while at the same time we try to shed light on the contradiction presented in section 4. Table 1 gives an overview on tests that we present in the rest of this paper. For each test we try to calculate time series as long as possible. Because the availability of data or certain information can change over time, we are forced to restrict certain analysis on specific time periods and/or to use different datasets.

Our main data source is income tax data from personal incomes published by the Swiss Federal Tax Administration (FTA).[[3]](#footnote-3) Federal taxes are collected and documented by the FTA since 1915. The time frame we look at in this paper reaches from 1945 to 2011 including 35 tax periods.[[4]](#footnote-4) While the FTA provides data electronically readable since 1973, we collected earlier data by scanning hard copies. In general, data is provided by the FTA in an aggregate form for privacy reasons, i.e. they are classified into numerous income brackets. Because these data not always contain all desired information, we use additional data sources (see column *Data* in Table 2). This includes FTA published key figures based on the federal tax statistics.[[5]](#footnote-5) 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. Furthermore, we use micro tax data from the canton Bern, because this data contains a register based household-ID, which allows us to address test (5) and (6) in a way, that is not possible with FTA tax statistic, but nonetheless shall provide us information in regard to tax statistic in general. For test (6) we finally use the Household and Consumption Survey (HBS).

For the empirical tests, we use several statistical techniques (see column *Method* in Table 2). 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 index. Then we apply relative distribution methods where we think an in-depth distributional analysis provides a more insightful understanding of distributional differences than one population measures. Section 5.2 provides a more thorough discussion on the applied inequality-measures.

Table 2: Overview on empirical tests within inequality related methodological areas

|  |  |  |  |
| --- | --- | --- | --- |
| **Methodological Area** | **Empirical test** | **Method** | **Data** |
| Defining economic resources | 1. Income definitions within tax data | Time series of Gini coefficients (own calculation) | Aggregated FTA tax statistic – normal cases without non-taxed – different income measures |
|  | 1. 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 – taxable income |
| Measuring inequality | 1. Change over time: difference between one population measures | Time series of Gini coefficients, Theil and Atkinson index (own calculation) | Aggregated FTA tax statistic - normal cases without non-taxed – taxable income |
|  | 1. Change over time: one population measure vs relative distribution | Gini differences (provided), relative distribution and polarization index (own calculation based on provided percentiles), polarization index (own calculation) | FTA Key figures – all tax units – taxable income |
| Statistical units | 1. Tax units vs households | Gini differences, relative distribution (own calculation) | Micro tax data from Canton Berne –all tax units - taxable income |
| Coverage issues | 1. Superior coverage with tax data compared to survey data | Gini differences, relative distribution (own calculation) | Micro tax data from Canton Berne and subsample for Berne from Household Budget Survey – primary income |
|  | 1. 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 |
|  | 1. Influence of non-taxed | Time series of Gini coefficients (own calculation) | Aggregated FTA tax statistics – normal cases with and without non-taxed – taxable income |

## Defining Economic resources

As described in section 2.1, an analysis of inequalities in incomes should look at the three factors income, wealth and consumption simultaneously. 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 on consumption are missing in tax data. But we can analyze how the assessment of income inequality is affected by using different income definitions that are present within the FTA tax data (5.1.1). Furthermore we evaluate the impact of using an equivalence scale tailored to tax data (5.1.2).

### *Income definitions within tax data*

When focusing on income the central measures reported in tax statistics are tax measures. To assess the effect of income definition we get three income measures:

* *Net income:* total income (earnings, income from property and current transfers received) minus some deductions.[[6]](#footnote-6)
* *Taxable income:* net income minus social deductions.[[7]](#footnote-7)
* *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.[[8]](#footnote-8)

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 consume (disposable income) (see also Figure 1 on page 5). The measure closest to primary income is net income. The measure closest to disposable income is taxable income after federal taxes.

Using these three income definitions we calculate Gini coefficients. As Figure 3 shows, these three time series cover different time periods, depending on the reported information by the FTA. A long time period is covered with taxable income and taxable income after federal taxes (from 1945 to 2011). Information on net income only reaches until 1981/1982 resulting in a shorter time series. The development for the three defined measures of income is quite parallel except for the 1980s. In this time period the Gini coefficient for net income veers. This is because of a change in regulations of deductions and shows that data over time need to be interpreted carefully, because changes in taxation or regulation systems can affect the outcome. 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 the difference are social deductions and these are basically fixed rate deductions, that relate to household properties. Hence, subtracting social deductions from net income results in over proportional reduction of lower incomes.

### *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 refers 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 out of tax data and applied to tax units. The incomes of single households are divided by 1 (no change), 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)[[9]](#footnote-9). By comparing Gini-time series for taxable income with and without implementation of this pseudo equivalence scale, we examine how strong the assessment of inequality is affected by this scale. Because excluding the group of not-taxed (on the influence of non-taxed see section 5.4.3) leads to a longer time-series, we provide four time-series in total (two possibilities to compare the effect of the equivalence scale).

The implementation of an equivalence scale does not have a major impact on the assessment of inequality (see Figure 3). 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.

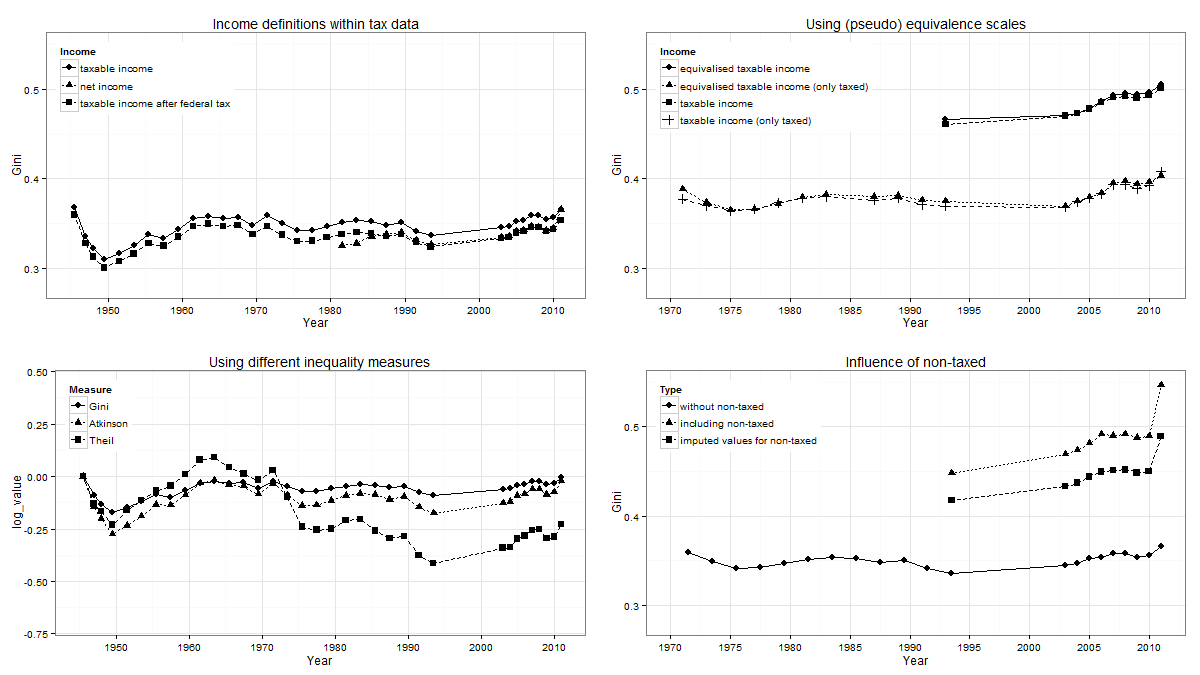


Figure 3: Inequality trends (a) using different income definitions, (b) with/without equivalence scale, (c) using different inequality measures and d) including/excluding non-taxed  
*Source: Aggregated Tax Statistics from FTA (a, c and d) and FTA Key Figures (b) from Swiss Federal Tax Administration (FTA), own calculations*

## Measuring inequality

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, its informative value is limited. Nonetheless, it is possible to identify periods where inequality increased or decreased, but it is not feasible at all to understand which part of the distribution actually was affected. To overcome these restrictions, we calculate additional measures (5.2.1) and expand the analysis with relative distribution methods (5.2.2).

### *Change over time: difference between one population measures*

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 and the Theil index.

The Atkinson index is defined as:

The core of the formula is the term where each individual income () is compared to the mean income ( This term gets 1 if all incomes () are the same and the Atkinson index gets zero (regardless of . The theoretical upper bound of the index is one, while it is driven by the ratio of for incomes below or above the mean. More unequal distributions result in higher ratios leading to an increase of the index. The second central parameter is the inequality aversion parameter . The Atkinson index is defined for each possible value of 0. For values close to zero the Atkinson index gets close to zero as well, regardless of the empirical distribution. This would describe a society indifferent to inequality. Furthermore the index reports higher inequality the higher is set. This would express higher aversion to inequality. Additionally the sensitiveness to the lower part of the distribution increases with increasing (De Maio 2007).

The Theil Index (T) is defined as:

Similar to the Atkinson index the measure is driven by the ratio of , but it’s value lies in the interval [0, log(n)], where 0 equals a complete even distribution and log(n) equals maximum inequality. More generally the Theil index can be assigned to the family of generalized entropy (GE) measures derived from information theory. Similar to the inequality aversion parameter the GE-Measures incorporate a sensitivity parameter. This parameter can be any real number. The higher a positive is, the GE measure gets more sensitive on the upper part of the distribution. Similarly, with values for getting more and more negative the GE measures give more weight to the bottom of the distribution. The Theil index equals the GE measure with =1 making it a measure that gives more weight to the upper part of the distribution.

We choose for the Atkinson and the Theil (GE(=1)) 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. 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 get very sensitive to high/low incomes when high values for respectively are chosen.

The bottom left graph in Figure 3 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 of 1945/1946. By doing so it is not possible anymore to interpret the level of each series, but changes over time are comparable between each series. The series follow quite a similar pattern, but they differ in volatility. This suggests that the borders of the distribution are much more prone to changes. 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 a Theil index above the other inequality measures. Further, in the 1970s and the 1990s, the Theil index drops below the other measures, suggesting a relative decline of higher incomes in this periods.

### *Change over time: One population measures vs relative distribution*

The comparison of bottom, mid and top sensitive measures can give a hunch on the nature of changing inequality. Even more light is shed on the changing patterns when we expand the analysis by using relative distribution methods. We therefore use the published percentiles of the distribution of taxable income from the FTA key figures dataset.[[10]](#footnote-10) We use the reported measures at the cost of time. The longest time-period we can compare out of these data reaches from 2003 to 2011. This includes all tax subjects (non-taxed and special cases). [[11]](#footnote-11) The Gini coefficient changed from 0.47 to 0.50, which equals a moderate increase of inequality. The in-depth distributional analysis allows us to see how this change translated into different shapes of the distributions. Instead of calculating two inequality measures for two separate time points, we apply the concept of relative distribution (Handcock and Morris, 1999) to perform an enriched trend analysis.

To construct the relative distribution we define 2003 as the reference distribution and the distribution of 2011 as the comparison distribution . represents taxable income. First, we calculate the two probability density functions (PDF). The PDF is a function which describes the distribution of probability over the outcome set and is defined for all possible values of .The PDF is characterized by its cumulative distribution function (CDF). The CDF is formulated as , which represents the probability that a randomly chosen value is less than or equal to .The relative distribution of to is then defined as:

(3)

is obtained from by transforming it by the CDF for ,. therefore measures the relative rank of compared to .

(4)

We can calculate the Probability Density Function of , where represents the proportion of values and is the inverse cumulative distribution function, also called the quantile function. might be interpreted as a density ratio, which is defined as the ratio of these two quantities evaluated at every percentile of the reference distribution [0,1]. With a complete overlap of both distributions the probability density function of the is 1 at every point of the PDF. Values higher than 1 represent higher probabilities in the comparison distribution than in the references distribution at this specific point and values lower than 1 respectively represent lower probabilities.

When looking at the relative density of the 2011 versus 2003 tax data (Figure 4, top left) a moderate polarization is visible, which is represented in a lower relative density in the middle deciles (d.20 to d.70), while the density ratio is notably higher in the top two deciles. While graphical displays are an interesting feature of the relative distribution framework, we want to compare summary measures based on the relative distribution to Gini coefficients and show how relative distribution measures complete the assessment of inequality trends. For this purpose we calculate the median relative polarization index (MRP), the upper polarization index (URP) and the lower polarization index (LRP), introduced by Handcock and Morris (1999). The median relative polarization index (MRP) is defined as the mean absolute deviation around the median of the location-adjusted relative distribution, scaled to produce an index that varies between -1 and 1. Given the scaling, a value of zero represents no differences in distributional shape; positive values represent more polarization (increases in the tails of the distribution) and negative values represent less polarization (convergence towards the center of the distribution). The measure catches only differences in distributional shape (not location). Further, the MRP is decomposable along the scale of . This makes it possible to compare the contribution of each section of the distribution to the overall polarization. A natural decomposition is the contributions made by components above (upper polarization index, URP) and below (lower polarization index, LRP) the median (of ). We quantify the visualized pattern with the described polarization indices and get following figures: Median Index=0.06, Lower Index=0.07, Upper Index=0.05. When comparing the lower and the upper index, we see that the polarization is slightly more driven by the downgrading of the low median percentiles.



Figure 4: Relative distribution over time, population and data source  
*Source: Aggregated Tax Statistics and Key Figures from Swiss Federal Tax Administration (FTA), Micro cantonal Tax data (Bern) and Household Budget Survey (HBS), own calculations*

## Statistical units

The usual units to assess inequality are households because the possibility to experience economic wellbeing 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. Therefore, it is not straightforward to elicit households and household income from tax data. This might influence the assessment of inequality development, taking into account the change from traditional household and family structures over the last century.

To examine the sensitivity of inequality regarding to the statistical unit, we use micro tax data from the canton Bern. This data includes housing information added from personal registers and allows constructing 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 we get when we look at pooled incomes of households instead of tax units. By comparing these two distributions, we can test the sensitiveness of inequality regarding different concepts of statistical units.

Our test shows substantial higher inequality among tax units (Gini=0.45) than among households (Gini=0.39). This can mainly be explained by an upward shift. Many single person tax units are indeed 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. When we switch from tax units to households, the share of persons effectively living alone decreases drastically. This results in pooling of income and an upward shift. In other words; a lot of units with low income are replaced with less units with higher incomes, The related relative distribution illustrates the differences (see Figure 4). 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 quite similar to 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 household level and not among tax units. In addition, we suggest that the bias got stronger in recent decades, and it had thus less influence in times when the overlap of fiscal and real households was bigger.

## Coverage issues

While survey samples are suspected to be biased because of non-response, the concerns about incomplete coverage are different with tax data. Essentially every permanent resident in Switzerland over 18 years of age (respectively 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 are missing. This can lead to an incomplete representation of the population. First, tax data distinguishes normal and special cases. Normal cases embrace the majority of taxpayers and are tax units residing in Switzerland without foreign source income and being liable to taxation all year long. Special cases cover mainly foreign nationals living in Switzerland or individuals who moved to or depart from Switzerland and are therefore not liable to taxation for the whole year. Second, tax statistics separate between those who actually pay taxes and 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. In 2011 this comprised 236’133 Individuals.[[12]](#footnote-12) Income at the top are suspected to be incomplete because of tax evasion. Non-fillers are a minor problem, because in Switzerland non-fillers equally show up in the tax-statistics as long as they are registered. These persons get an imputed income based on an older tax return and information given by employers. Only non-registered non-fillers are not in the records. An important bias, however,are individuals that misreport incomes. Feld and Frey (2006) examine the role of tax evasion in Switzerland by calculating the difference of 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.

When focusing on the available tax statistics, we can distinguish three coverage issues with an empirical possibility to test their relevance for inequality analysis. First, we compare the tax income distribution to survey data, to see if tax data covers extreme incomes more reliable than survey data (5.4.1), then we test if the inclusion or exclusion of special cases has a substantial impact on the assessment of income inequality (5.4.2). Third, (5.4.3) we quantify how strong inequality is affected by neglecting those subjects, who are not taxed.

### *Superior coverage with tax data than with survey data*

According to the going scholarly opinion, tax data covers the extreme part (lower and upper incomes) of an income distribution in a more reliable way than survey data because the latter is suspicious of sampling error. To test this hypothesis we perform two tax data comparisons with the Household and consumption Survey (HBS). We use the HBS because it is one of the surveys used for governmental publications (ESTV 2014) and incomes are provided on a very detailed base. This allows us to construct measures that are better 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 represents 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 the difference of statistical units we restrict our analysis to married couples. Afterwards, 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 smaller on average. We assume that these deductions are proportionally equal across the whole income distribution and hence do not interfere with the comparison and correct this difference with a multiplicative (log of mean) location shift. 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 conceptual 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 in the survey data. The drawback of this strategy is that we cannot compare the same years. Tax data represents the year 2012, while the most actual HBS data refers 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.

The lower left and the lower mid graph in Figure 4 plot the relative Density of the HBS distribution with tax data as 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 seams more pronounced in the plot for married than in the plot for Bern. In addition, this difference can be explained with missing welfare incomes in tax data, which leads to an underrepresentation in the lower part of the income distribution within tax data. The extreme parts (very rich and poor) are better represented in both plots within tax data. All in all, the upper middle class bias results in an underestimation of inequality. The Gini coefficient for Bern is +0.06 higher in tax data than in the HBS. A comparison of the Gini coefficients for the tax data and HBS for the married results in even a +0.18 higher coefficient, which is explained by the upper middle class bias and the missing social welfare incomes.

### *Influence of special tax subjects*

Tax statistics distinguishes between normal and special cases. To test what influence has 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 to publicly report data for special cases after tax period 1993/94. Therefore we compare two distributions for a rather old dataset. However, the FTA key figures do report aggregate statistics (e.g. percentiles) based on a pool of all cases (normal and special) for more recent periods which allow us to do a corresponding analysis for 2011 as well.

The pooled data set 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 4 top middle). Put simply: the population of special cases in 1993/94 hold 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 4 top right). According to Figure 4 we can attribute this effect to the top percentiles. To get a better understanding of the observed patterns we have 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 very rich foreigners who are not in gainful employment in Switzerland. These individuals are taxed with special conditions and get an imputed income according to their expenditures. Therefore, they appear in the upper part of the income distribution. As Table 3 shows, this is a minor group but in the last twenty years their number more than doubled, which supports the thesis that rich immigrants lead to an increase of inequality in recent years. Inequality also increases with migration at the lower end of the income distribution. There is a bigger group of other special cases with very diverse constellations. Common are individuals who either moved to or departed from Switzerland and are therefore not liable to taxation for a whole year. Their income in Switzerland gets extrapolated to a 12 month income. Therefore, their incomes should not be artificially low. Other special cases are natives with foreign incomes or foreigners with income in Switzerland. Their incomes too 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 liable to taxes if they own business establishments or property in Switzerland. Because these persons only have to pay taxes for incomes gained in Switzerland, they have lower incomes because of technical reasons.

Table 3: 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 Federal Tax Administration (FTA)*

All in all, special cases are natives and foreigners who are related to a foreign country but are nonetheless part of Swiss society. Their inclusion leads to an increase of income inequality because special cases are strongly polarized including very low and very high incomes. In terms of 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 include migrants who are liable to taxes in Switzerland but who did not yet receive

a permanent residence permit. As this is a common case and these individuals often stay several years in Switzerland and probably have very diverse incomes, it would be interesting to see how their inclusion would affect the income distribution. Also taxed by 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.

### *Influence of non-taxed*

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 based on the period from 1995/1996 to 2011.

We calculate three Gini-time-series (see Figure 3, bottom right). Though this is not surprising, excluding the non-taxed leads to a dramatic drop of the Gini coefficient. At the same time, however, we overestimate inequality when we assume 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 halve the threshold for single tax units.[[13]](#footnote-13) This results in slightly lower, more realistic Gini coefficients.

Regarding the exemption threshold, a second issue occurs when having a closer look at 2011 where the Gini coefficient rose drastically. Although this raise could be attributed to a more unequal distribution of incomes, fiscal adjustments are another cause of the higher Gini coefficient. This becomes clear when we have a look at 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 increased by over 350’000 cases to 1’257’075 (28.5% of all tax subjects). This major increase can be explained with a raise of the exemption threshold and with a raise to allege deductions for married with children. All in all these fiscal adjustments result in a substantial bigger share of non-taxed and an artificial increase of the Gini coefficient.

The problem of non-taxed is even worse for earlier tax periods. Although the FTA does not report the share of non-taxed 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 in question. According to their estimations the share of tax subjects represented in FTA tax statistic varies from 94% in 1993/1994 to 13.7% back in 1933. It is highly questionable if analysis based only on a small fraction of the population is appropriate.

# Summary and Conclusion

In this paper we checked the suitability of tax data to carry out inequality trend research and showed that tax data has strong advantages like long time availability and solid population coverage, but in the same time not all theoretically relevant concepts are addressed properly. After pointing out ideal inequality related theoretical concepts concerning the definition of economic resources, statistical units, measurement of inequality and coverage issues, we evaluated the advantages and disadvantages of real tax data by using Swiss income tax data provided by the Federal Tax Administration (FTA) as an example. In the following, we want to sort out major and minor methodological issues by projecting the conclusions from the case of Switzerland to a more general level of how methodological options affect the assessment of income distributions.

Regarding aggregated tax statistics, some central conceptual imperfections need to be mentioned. While the state of the art is to use data sources that include measures for income, wealth and consumptions alike to allow a holistic view on economic well-being, aggregated tax statistics report only on income or wealth and cover expenses only partially. The fact that the data is presented in an aggregated form also means that analyses with the goal of decomposing inequality components are not applicable. The potential for extended inequality analysis is therefore restricted. Additionally, central definitions of economic resources are data-driven. Regarding income conventional measures like pre- or post-redistribution income measures cannot be addressed. Rather, taxable income is at hand, which has to be situated between the pole of market outcome and income disposable for consumption. Furthermore, the statistical units of tax data are fiscal households and not real households and up to date it was not clear, how the assessment of inequality is affected by this conceptual difference.

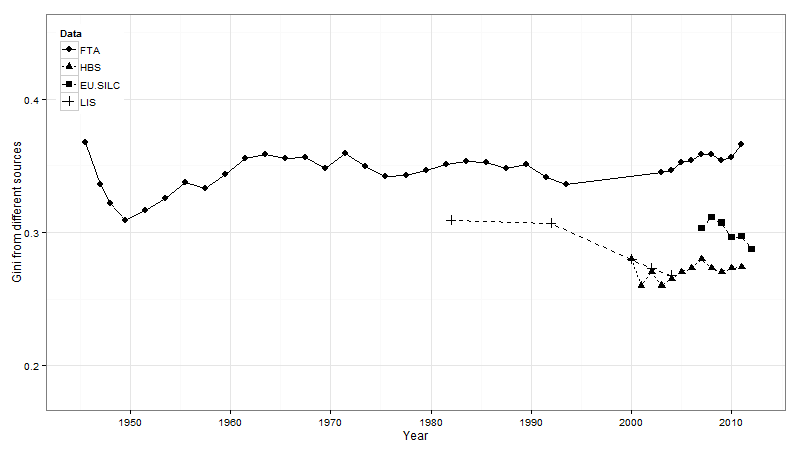
In the empirical section we conducted several stability tests with aggregated tax statistics to sort out major from minor issues regarding the assessment of income inequality trends. Based on the maximum observed range of Gini coefficient for each performed test, we build a ranking:

1. Influence of non-taxed (Max Range of Gini coefficient= 0.12)[[14]](#footnote-14)
2. Superior coverage with tax data compared to survey data (0.07)
3. Tax units vs households (0.06)
4. Income definitions within tax data (0.03)
5. Influence of special tax subjects (0.02)
6. Using income corrected with an equivalence scale based on tax information (0.01)

We do not mean these figures as a “hard” estimate of true differences, in a sense that they could be used to adjust given estimations. The reported differences are strongly related to a given population. E.g. the difference between the tax unit and the household distribution is strongly affected by the actual overlap of the two concepts. Nonetheless, the ranking gives us an overview on what potentially is influential and what not. Following the ranking, the issue of non-taxed is the most central problem when working with tax data. The results in the empirical section show that fiscal adjustments influence the share of non-taxed and this again has an impact on the assessment of inequality. Furthermore, information on non-taxed is only available after 1995/1996 leaving the researcher with only information on taxed. Then again there are periods before 1943/1944 where the subpopulation of taxed represent sometimes only a small fraction of Swiss population as the estimations of Dell et al (2007) suggests. The second point in our list refers to coverage issues. Our analysis showed that the distributions of tax and survey data differ substantially, albeit both cover Swiss population in theory. We argue that this difference stems from an upper middle class bias in the survey data which results in underestimation of inequality. Another central difference of tax data is that statistical units are fiscal and not real households. This is crucial in the case of cohabitation without marriage, which is present in tax data as two single tax units and leads to an overestimation of inequality and certainly to a bias in the inequality trend as the „single-to-married-ratio“ varies in most western countries over time (trend towards more singles). The 4th place in the ranking goes to income definitions within tax data. Compared to other issues this seems to be a minor point, but we have to keep in mind that aggregated tax data do not allow to construct disposable incomes, especially the missing possibility to account for federal and communal taxes carries weight. Hence, the redistributional effect of taxes cannot be depicted, which leads summa summarum to an overestimation of inequality based on taxable income. Compared 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, albeit we showed that the inclusion of special cases is necessary to catch the effect of a special socio-political developments like the recent immigration of rich individuals in Switzerland, who get tax privileges by getting taxed according to expenses.

Measurement issues are especially crucial, as the performed tests cannot be included in the ranking because other measures than the Gini coefficients were used and comparability in the sense of the ranking therefore is not suitable. Nonetheless, the performed test showed that the judgment of inequality is somehow tied to the chosen measurement coefficient and analyses are most fruitful when different measures and approaches are used. We think trend analysis is best done by combining several one population measures for a first analysis of time patterns, which then is enriched with a second analysis using relative distribution methods for specific time points to unravel complete distributional differences.

Given the presented findings from our methodological journey, is it possible to dissolve the contradiction regarding the current state of research concerning the income inequality trend in Switzerland? Figure 5 displays the longest meaningful time series of Gini coefficients that can be calculated for Switzerland out of FTA tax statistics.[[15]](#footnote-15) This time series is based only on normal taxed units (information for special cases and non-taxed are not available for such a long time period) and calculated with taxable income. This series can be compared to the already shown results from the main Swiss Surveys. Having all the imperfections summarized above in mind we know that all these Gini coefficients are not perfectly valid. While most factors discussed above lead to an overestimation of inequality when using tax data and on the same time to an underestimation of inequality when using survey data (non-response) it is not surprising that the level of inequality in general is higher with tax data. The truth probably lies between the presented series from tax data and survey data. But what do we learn if we focus on the possibility to assess the trends?

Figure 5: Income inequality trends in Switzerland: Tax Data vs Surveys  
*Source: Aggregated Tax Statistics Federal Tax Administration (FTA), Luxembourg Income Study (LIS), Household Budget Survey (HBS), European Union Statistics on Income and Living Conditions (EU.SILC)*

It is clearly visible that tax data outperform survey data in regard to the length of the covered time period. While most imperfection of tax data are rather constant over time, the missing information on non-taxed vary and therefore interfere with the pattern. Construe the estimations of the share of taxed (Dell et al.; 2007) more conservatively it is reasonable to start interpreting the series not before 1973. Since then, the share of taxed subjects is rather stable and more than 75% of Swiss population is covered. Before 1973, the share of taxed was noticeable lower and interpretation for those time periods should only be made very carefully. The development of income inequality in the period directly after World War II is at least plausible. This period was characterized by strong economic growth as well as an increase in inequality. One possible interpretation is that high income percentiles over proportionally profited from the economic upturn. After the oil crises in 1972 there were alternating phases of social welfare expansion and economic up- and downturns. Especially interesting is the period past the millennium, where it is possible to compare tax data to the results from survey data and trends diverge clearly. While survey data suggest a decline in income inequality the time series based on tax data rather promote an increase and the question arises which series does represent reality more adequate. By analyzing the relative distribution of 2011 compared to 2003 (see Figure 4 on page 16) during this period a polarization occurred that is slightly stronger driven by downgrading of low incomes but also by an increase of top incomes. Because these parts of the income distribution are better covered within tax data than within survey data it seems feasible that the recent trend is an increasing one and Switzerland therefore is not a special case like the recent analysis of the Luxembourg Income Study performed by Gornick and Jäntti (2013) suggested.

To sum up, tax data deviates substantially from given state of the art concepts. Because of its historic availability it is anyhow an essential data source for trend analysis. Finally, it has to be mentioned that the aggregated tax statistics reported from the FTA are the result in the end of the line in the process of levying taxes. This means there is also micro tax data around that allows addressing most of the mentioned conceptual drawbacks. We recommend researchers to check the availability of such data in their country.

# Literaturverzeichnis

Alvaredo, F., Atkinson, A. B., Piketty, T., & Saez, E. (2013). The Top 1 Percent in International and Historical Perspective. *Journal of Economic Perspectives*, *27*(3), 3–20.

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. doi:10.1162/JEEA.2009.7.5.1140

Atkinson, A. B. (1970). On the measurement of inequality. *Journal of Economic Theory*, *2*(3), 244–263.

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.

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

Brandolini, A., & Smeeding, T. M. (2009). Income inequality in richer and OECD countries. *The Oxford Handbook of Economic Inequality*, 71–100.

Brülhart, M. (n.d.). Proposal Sinergia - The Swiss Confederation: A Natural Laboratory for Research on Fiscal and Political Decentralization.

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. doi:10.1111/j.1475-4991.1988.tb00564.x

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

Dalton, H. (1920). The measurement of the inequality of incomes. *The Economic Journal*, 348–361.

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. doi:10.1136/jech.2006.052969

Diekmann, A. (2009). *Empirische Sozialforschung: Grundlagen, Methoden, Anwendungen*. Reinbek bei Hamburg: Rowohlt-Taschenbuch-Verl.

ESTV. (2014). *Verteilung des Wohlstands in der Schweiz*.

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., & Martínez, I. (2013). Volatile Top Income Shares in Switzerland? St. Gallen. Retrieved from http://www2.unine.ch/files/content/sites/irene/files/shared/documents/SSES/Martinez.pdf

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

Grabka, M., & Kuhn, U. (2012). The evolution of income inequality in Germany and Switzerland since the turn of the millennium. Retrieved from http://papers.ssrn.com/sol3/papers.cfm?abstract\_id=2131220

Handcock, M., & Morris, M. (1999). *Relative Distribution Methods in the Social Sciences (Statistics for Social Science and Behavorial Sciences)*. Springer. Retrieved from http://www.amazon.ca/exec/obidos/redirect?tag=citeulike09-20&path=ASIN/0387987789

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

International Labour Organisation (ILO). (n.d.). *Final report of the 17th International Conference of Labour Statisticians*. 2004.

Kolm, S.-C., & others. (1969). The optimal production of social justice. *Public Economics*, 145–200.

Korinek, A., Mistiaen, J. A., & Ravallion, M. (2006). Survey nonresponse and the distribution of income. *The Journal of Economic Inequality*, *4*(1), 33–55. doi:10.1007/s10888-005-1089-4

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

Leigh, A. (2007). How closely do top income shares track other measures of inequality? *The Economic Journal*, (117), 589–603.

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. (2013a). Basic statistics of Switzerland, 2012: (Numbers in parentheses refer to the OECD average). In *OECD Economic Surveys: Switzerland 2013* (pp. 9–129). OECD Publishing.

OECD. (2013b). *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.

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.

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.

Särndal, C.-E., Swensson, B., & Wretman, J. (2003). *Model Assisted Survey Sampling*. Springer.

Sen, A. (1973). On economic inequality. *OUP Catalogue*. Retrieved from http://ideas.repec.org/b/oxp/obooks/9780198281931.html

Shorrocks, A., & Slottje, D. (2002). Approximating unanimity orderings: An application to Lorenz dominance. *Journal of Economics*, *9*(1), 91–117.

SNA. (2008). *System of national accounts 2008*. New York: UN.

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

UNECE, & EUROSTAT. (2006). *Conference of European Statisticians Recommendations for the 2010 Censuses of Population and Housing*. New York and Geneva: United Nations. Retrieved from

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.

1. Harmonization with other international standards was an important objective that guided the work of the expert group in developing the ICW Framework presented in this publication. Considered main standards were the System of National Accounts (SNA, 2008), the Canberra Group Handbook on Household Income Statistics (United Nations, 2011), the final report of the 17th International Conference of Labour Statisticians (International Labour Organisation (ILO), 2004) and the UNECE/CES recommendations for the 2010 Censuses of Population and Housing (UNECE and EUROSTAT, 2006). [↑](#footnote-ref-1)
2. (1) Weak principle of transfers,

   (2) strong principle of transfers,

   (3) scale invariance,

   (4) the principle of population and

   (5) Decomposability. [↑](#footnote-ref-2)
3. http://www.estv.admin.ch/dokumentation/00075/00076/00701/index.htm [↑](#footnote-ref-3)
4. We did not use tax data before 1945 albeit it is accessible until 1915(?) because data before 1945 comprises only a minority of potential tax units. According to estimations of Dell et al. (2007) the share of tax fillers 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, that’s why there is no exact data available for Switzerland in the transition period. [↑](#footnote-ref-4)
5. 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. [↑](#footnote-ref-5)
6. 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 and charitable donations [↑](#footnote-ref-6)
7. Social deductions include: married, single-parent households, insurance premiums, interests earned by savings, deductions for children and supported persons, second earner deductions. [↑](#footnote-ref-7)
8. We call it a pseudo disposable income, because important expenses are not covered at all like cantonal and municipal

   taxes, which represent the bulk of taxes in Switzerland. Also the cost of health insurance are not covered. [↑](#footnote-ref-8)
9. The implementation of this pseudo equivalence scale is not done by use. It is part of the key figures provided by the FTA. [↑](#footnote-ref-9)
10. We prefer these measures over the calculated measures out of the published income brackets statistics, because they represent the distribution at both tails more accurate since they are based directly on the information about every single tax units. When calculating percentiles out of the income bracket statistic we lose relevant information at the edges. First, we don't have information about taxable income of tax-units falling below the income threshold for federal taxation (see also section 5.4.3). 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 95%-percentiles, while the reported percentiles reach the 99.99%-percentiles. [↑](#footnote-ref-10)
11. See section 5.4 for additional information on special cases. [↑](#footnote-ref-11)
12. <http://www.bfs.admin.ch/bfs/portal/de/index/infothek/lexikon/lex/0.topic.1.html> (13.4.3 Sozialhilfe und Asylwesen) [↑](#footnote-ref-12)
13. 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 14’900 CHF to 16’100 CHF for singles) and in 2011 (to 17’700 CHF) [↑](#footnote-ref-13)
14. Difference to the imputed series. [↑](#footnote-ref-14)
15. We do not show time points before 1943/44, although theoretically information is available, because in this period only small fraction of potential tax units are covered in the tax statistics. [↑](#footnote-ref-15)