|  |
| --- |
| Assessing inequality with tax data  Income Inequality in Switzerland from 1917 to 2011  Oliver Hümbelin  Bern University of Applied Sciences  **[oliver.huembelin@bfh.ch](mailto:oliver.huembelin@bfh.ch)**  Rudolf Farys  University of Bern  **[rudolf.farys@soz.unibe.ch](mailto:rudolf.farys@soz.unibe.ch)**  November 2014  *Abstract*  In many countries inequality trends are inconclusive, because different methodological approaches blur the picture or reliable data is scarce. In this paper we assess whether tax data is suitable to satisfy ideal theoretical concepts (income measurement, population coverage, statistical units) to track inequality development. Using public tax data from Switzerland we show a) how strong inequality is affected by using different income concepts, by b) comparing tax data to survey data, we furthermore quantify the bias from different household concepts (tax units, households) and from sample bias and by c) using relative distribution method, we show how interpretation changes, when using different inequality measures. Finally we present a new income inequality time series for Switzerland showing rising inequality trends in the 1950s and 2000s as well as a decreasing trend in the 1970s and 1990s. |
| **Berner Fachhochschule**  Soziale Arbeit |

Inhaltsverzeichnis

1 Introduction 3

2 Standards on Assessing Economic Inequality 3

2.1 Concepts on measuring economic resources 4

2.2 Statistical Units 5

2.3 Measuring inequality 5

2.4 Coverage Issues 7

3 Comparison of tax data and survey data – overview of advantages and shortcomings 8

4 Conflicting results due to methodological differences in Switzerland? 9

5 Assessing income inequality trend with tax data for Switzerland 11

5.1 Defining Economic resources 13

5.2 Statistical units 14

5.3 Measuring inequality 15

5.4 Coverage issues 18

6 Conclusion 22

7 Literaturverzeichnis 25

8 Anhang 26

# Introduction

Economic resources can be seen as central indicator for life chances. Therefore the distribution of resources does not only matter in regard to the possibility to consume, but also to physical and mental health or even life expectancy. Furthermore the distribution of resources cannot be reduced to the opposition of haves and have not’s on the individual level, it is rather an issue concerning the whole society. Inequality matters for societies, because it is related to negative outcomes like for example criminality, violence, imprisonment, teenage births or social trust (Wilkinson and Pickett, 2009), which is a core dimensions of functionality of societies. In the light of empirical evidence that suggests a rise of 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 the concern about the widening gap is addressed more and more by global leaders (World Economic Forum, 2013). Although the rise was not uniform, a common pattern seems to be identifiable, which can be referred to as the “hollowing of the middle class” (Alderson and Doran, 2013). Households are moving towards the top and the bottom of the distribution relative to the past, which is especially problematic as the middle class can be seen as the core of western democracies or, as it is stated by Stiglitz (2012, 117): “our democracy is being put at peril.”

Given the importance of the subject a constant reflection about reliability of empirical data seems appropriate. Atkinson (2013) observes improved methods regarding household surveys, the core sources of inequality research. On the other hand, the labor intensive and expensive surveys around the world are subject to budget cuts and the instrument itself suffers of low response rates, which affects 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), Piketty and Saez (2003) made the use of tax data fashionable again. Following Pikettys approach, studies on several countries were conducted (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).

Tax and survey data can be identified as the two major data sources concerning the assessment of inequality trends. Both sources predefine the way inequality can be analyzed in fundamental ways. The question arises, to what extent the assessment of inequality is affected by the choice of the data source and consequent possibilities and restrictions. To answer this question, we describe the current theoretical standards for measuring economic resources and inequality in section 2, which gives a common ground to evaluate the suitability of data sources. In section 3 we show in what ways either tax or survey data are superior and where special attention concerning the assessment of inequality should be paid. In section 4 we introduce Switzerland as an interesting example for a closer methodological inspection. In the light of the above, we assess inequality of incomes with federal tax statistics for Switzerland in section 5. We show, how the assessment of inequality is affected by the choices researcher have to make, when working with tax data. In section 6 we summarize which methodological issues are relatively important and which are negligible.

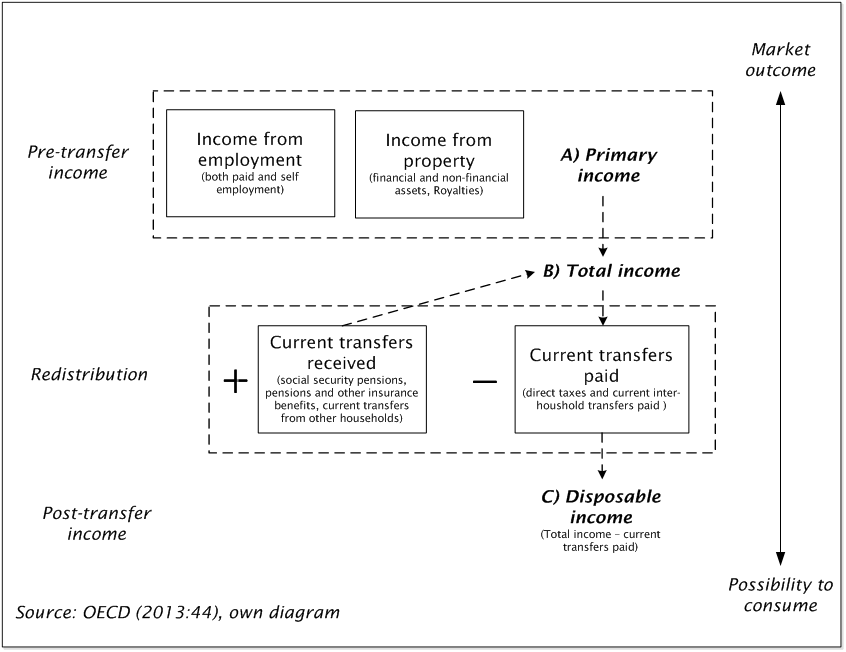
# Standards on Assessing Economic Inequality

We identify four areas, which are crucial concerning the assessment of economic inequality. First of all one has to be clear about the type of economic resources (section 2.1). What kind of resource do we look at when we investigate its distribution? Then one has to define the statistical unit to announce among whom inequality occurs (see section 2.2. Section 2.2 gives an overview on inequality measures and discusses their central advantages and shortcomings. Section 2.4 addresses the importance of coverage issues.

## Concepts on measuring economic resources

Most studies on inequality focus on income inequality solely. However, recent activities emphasize the need of a broader conceptualization. A recent publication from the OECD (2013) condenses these ideas into the ICW framework (income, consumption and wealth), which is meant to be 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, it is recommended to look at the distribution of all three distributions simultaneously. Some households with low income, for example, may report adequate levels of consumption expenditure or wealth holdings, or vice-versa.

Because inequality in income is by far the concept, that gets most scholarly attention, we have a closer look at the definition of income. Terminology can slightly differ, while common concepts can be identified (for detailed discussion see: OECD (2013, 44), United Nations (2011, 24)). Figure 1 shows 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. This direct product of the market outcome is called the primary income. Redistribution then takes place through social transfers. 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 income: the income measure, which finally shapes 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 the latter considers redistribution and the former does not. Additionally, incomes are often equivalised 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 different households (see subsection statistical unit).



## Figure 1 : Income definitions from primary income to disposable income Source: OECD (2013:44), own diagramMeasuring 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 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. Properties of measures are usually discussed in regard to five principles[[2]](#footnote-2). Here we don’t want to 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 on variance or variance based measures. This measures are straightforward to calculate, but have the disadvantage, that coefficients theoreticaly can take any value between zero and infinity. Hence, it is difficult to identify reasonable niveaus. Additionally some measures (Range, log variance and variance of logs) even fail 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 seems to be a central property an inequality measure should have, 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 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 can’t 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 is derived 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 in-different 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). But 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, the polarization indices follows 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, one cannot 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, because this measure is decomposable to distinguish differences in the upper and lower tails.

Given this plethora of inequality measure, how do we choose the right one? In general this has to be decided in regard to the research question, but broadly speaking, it is not recommended to just rely on one inequality measure. Already classical works from Kolm (1969) to Atkinson (1970) to Sen (1973) warn us against relying on a single inequality measure. A use of 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

The agreed standard on the statistical units, which should be the base of inequality analysis, are households not individuals (OECD, 2013, 60). Indeed it are individuals, who receive income, own assets and experience economic well-being, but their possibility to do so, is strongly tied to the concept of household. This comprises all persons under the same housing arrangement. The basic underlying assumption for collecting data on household level instead of individual level is, that people in the same household share resources and therefore pool their incomes (when two or more earners live together) and/or 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 when people share living space and commodities and they therefore benefit from sharing. To compare the individual economic well-being among individuals living in different households usually equivalence scales are used as mentioned above.

## Coverage Issues

Studies in general try to make a statement about the whole population of interest (e.g. nation). The success of such a venture is closely related to the way data is collected. This again is determined by the resources researchers have. When total population data is not at hand, researchers usually work with samples and try to infer from sample results on 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 exists and there is a legal basis to use them for statistical purposes. In fact nearly a third of all countries participating 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 discussed four central areas researchers need to address. To sum up: 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, (4) calculate an unbiased estimate of an inequality measure. Table 1 compares tax-data and survey 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 with an administrative purpose. Furthermore, in a lot of countries tax statistics are only available in aggregated form showing tax units per taxable income/wealth brackets and no information on individuals. The missing of the link on the individual level implies therefore no possibility of a conjoint analysis of income and wealth. In addition, information on consumption is missing at all. This leaves the researcher with the possibility to look at income or wealth. The definition of central measures is also often restricted, because only tax-relevant measures are at hand. Taxable incomes for example include direct social transfers (e.g. rents) but no taxes are subtracted. Ergo a researcher using tax data can neither look at a pre- nor a post-transfer measure but something in between (see Figure 1). 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, if data is available on an individual level (like it is 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 individual 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 has to be mentioned. The statistical units of tax data are so called tax units (i.e. singles or married), but these do not necessarily correspond to households. Indeed there are constellations where members of the same household hand in several tax forms. A common case would be an unmarried couple living together. With the change of household-structures over time this issue is becoming increasingly important. Here again surveys are usually able to address the ideal statistical unit in a more appropriate way.

A closer look should be taken at *coverage issues*. As mentioned, nonresponse is a general problem of samples and major issue, when working with income data. (Korinek et al., 2006) show, that 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. This mechanism can be referred to as the ”middleclass bias” (Diekmann, 2009). Missing data in household surveys is therefore not missing at random, which has an impact on the measures of inequality. Weighting strategies to handle this kind of bias are discussed in the literature (Särndal et al., 2003), but require a register for every unit, that is proportional to income, which is rarely available. On the other hand, 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, when not accounting for this. Another critical issue with tax data is the problem of tax evasion, which definitely can bias the assessment of inequality. [Alvaredo and Saez](#Xalvaredo_income_2009) ([2009](#Xalvaredo_income_2009)) for example consider estimates of Spanish top incomes prior to 1981 as unreliable due to widespread tax evasion. Evasion can occur, 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 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 if the data is truly comparable over time, 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?

Results on income inequality for Switzerland are particularly contradictory, making it an interesting case to have a closer look at methodological aspects. What is known about Switzerland so far? Looking for official data, three main sources have to be mentioned, which can be considered as de facto official data sources: Statistics on Income and Living Conditions (EU-SILC), Household Budget Survey (HBS) and Luxembourg Income Study (LIS). Figure 2 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 doesn’t cover time periods before 2007. As figure 2 shows, following the results from EU-SILC, income inequality decreased from 2007 to 2013. The second important source concerning the distribution of income is the HBS. The main focus of this survey lays in providing detailed data 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 it can be seen from figure 2 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, 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 substantially decreases in income inequality for Switzerland, contradictory to the development in most other western countries. This result is supported by Grabka and Kuhn (2012) analyzing the Swiss Household Panel (2000-2009).



Figure 2: Trends of income inequality in Switzerland.

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 the development of concentration of the highest incomes and wealth (top-shares). 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 to 2008 finding that the share of top incomes has risen, the top 0.01% share even doubled in the last observed 20 years. A result that opposes the outcome 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. Differences can be explained with several factors. As described in section 2 and 3 the choice of data source has pros and contras regarding the ideal standard. First, coverage of top incomes is assumed to be better within tax data than it is within survey data (non-response bias), which is a crucial issue concerning inequality. Second, different measures of inequality hamper the comparability. Following Leigh (2007:600) “top income shares are far from perfect as a measure of distribution of income across soceity”, although he finds a strong positive correlation with other inequality measures. Third, different income concepts were used. As it is shown by Modetta and Müller (2012) 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 taxe-rates and tax deductions on sub federal level) is not directly represented. Fourth, tax data also neglect the household structure, because tax units don’t necessarily correspond to households.

# Assessing income inequality trends with tax data for Switzerland

As shown, the use of different data sources and different concepts can lead to different conclusions. In this section we therefore have a closer look at methodical choices that have to be made concerning the four areas introduced in section 2 (defining economic resources, measuring inequality, statistical units, coverage Issues) when working with tax data. We discuss how good theoretical relevant concepts can be addressed with tax data and we provide empirical evidence for crucial topics within the four introduced methodical relevant areas. Table 1 gives an overview on topics covered in the rest of this paper. The table also shows the data we used for each specific calculation. In general our main strategy is to apply different possible concepts within one of the defined areas while holding other conceptual differences constant. With this strategy we show, where the assessment of inequality is sensitive to conceptual choices and where not. Because the availability of data or certain information can change over time, we are forced to restrict certain analysis on specific time periods and to use different datasets.

The core of our analysis is always tax data. Our main data source are income tax data for individuals published by the Swiss Federal Tax Administration (FTA).[[3]](#footnote-3) Federal taxes are collected and documented by the FTA since 1915. Being called a war-tax in the beginning, the federal tax was renamed to crisis levy in 1934, defense-tax in 1939 and is finally known as direct federal tax since 1983. The time frame we were able to collect ranges from 1945 to 2011 including 44 tax periods[[4]](#footnote-4). While the FTA provides data in machine readable form since 1973 we collected earlier data by scanning hard copies[[5]](#footnote-5).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 additionally use FTA published key figures based on the federal tax statistics[[6]](#footnote-6). This figures include Gini coefficients and percentiles ranging from 1973-1974 to 2010 for individuals, who had to pay federal taxes and from 1995-1996 for all taxable individuals. Furthermore, we use individual cantonal tax data from the canton Berne, because these data contain a register based household-ID[[7]](#footnote-7), which allows us to address test (5) and (6) in a way, that is not possible with FTA tax statistic For test (6) we finally use the Household and Consumption Survey (HBS).

Table 1 : Overview on empirical test on inequality related methodological decisions.

|  |  |  |  |
| --- | --- | --- | --- |
| **Methodological Area** | **Empirical test** | **Method** | **Data** |
| Defining economic resources | 1. Income definitions within tax data | Time series of Ginicoefficients (own calculation) | Aggregated FTA tax statistic - 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 – net income |
| Measuring inequality | 1. Change over time: difference between one population measures | Ginicoefficient, Theil, Atkinson (own calculation) | Aggregated FTA tax statistic - without non-taxed – taxable income |
|  | 1. Change over time: one population measure vs relative distribution | Ginicoefficient (provided) relative distribution (own calculation based on provided percentiles), polarization index | FTA Key figures – all tax units – taxable income |
| statistical units | 1. Tax units vs households | Gini differences, relative distribution (own calculation), | Individual 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) | Individual 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 and 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 Ginicoefficients (own calculation) | Aggregated FTA tax statistics – all tax units – taxable income |

For the empirical test, we use several statistical techniques. 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. While the Gini coefficient is silent concerning the relevant areas of the distribution subject to a change, the Theil and Atkinson indexes can give a hunch about trends within the upper and lower end of the income distribution. The latter aspect is extended by relative distribution methods which we apply for selected periods 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 aspect of different inequality-measures.

## Defining Economic resources

As described in section 2.1on page 3, it is recommended to look at income, wealth and consumption simultaneously, when the interest lays in the distribution of economic well-being. 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 concepts are part of the ICW framework of the OECD. The Federal Tax Administration (FTA) publishes statistics on income and wealth but it is not possible to analyze the joint distribution on the individual or household level. Also measures of consumption are missing in tax data, but we can analyse 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 equivalence scales that are based on tax information(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 (Reineinkommen):* total income (earnings, income from property and current transfers received) minus some deductions[[8]](#footnote-8)
* *Taxable income:* net income minus social deductions[[9]](#footnote-9)
* *Taxable income after federal taxes:* Through accounting the reported federal taxes per taxable income bracket, we can construct the taxable income after federal taxes, which can be understood as a sort of pseudo disposable income[[10]](#footnote-10). [[11]](#footnote-11).

These tax measures don’t correspond directly to theoretically defined measures like primary income (before redistribution) or disposable income (after redistribution). They rather have to be situated between the poles of market outcome (primary income) and income left for consume (disposable income) (see Figure 1). 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 out of the FTA-tax statistics. As it is visible in Figure 3, these three time series cover different time periods, depending on what was reported by the FTA. The longest time period is reached with taxable income (from 1918 to 2011).

As Figure 3 shows, 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 has to do with a change in regulations of deductions and shows that interpretation over time has to be very careful, 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 reduces inequality slightly because of the progressivity of the taxes and inequality is higher for taxable income than for net income, because the difference are social deductions, which are basically fixed rate deductions, that relate to household properties. Hence, subtracting social deductions from net income to get the taxable 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* by accounting for the number of household members that potentially share income and resources. Because tax data refers to fiscal households and not real households only an approximation of the equivalence concept is possible by using a scale which 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. This is principally follows the logic of the modified OECD-Scale implements (OECD, 2013:173). By comparing Gini-time series for net 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 leads to a longer time-series we provide four time-series in total (two possibilities to compare the effect of an equivalence scale). These measures are part of the key figures provided by the FTA and are not calculated by us

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 only approximately depict households, it has to be said, that the implemented equivalence scale automatically has its drawbacks.

## Measuring inequality

In this section we show, how interpretation on the development of inequality is related to the measurement coefficient. So far we showed Gini coefficients, the most common measurement of inequality. But the coefficient has certain restrictions. It is know, that the Gini coefficient is more sensitive to the middle part of the distribution and accordingly less sensitive the 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 other one population 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 for taxable income to inequality measures, that are more sensitive to other parts of the distribution. Namely we calculate the Atkinson and the Theil index.

The Atkinson index is defined as:

The index theoretical range is 0 to 1, while it equals to zero only if all incomes () are the same. The core of the formula is the term, where each individual income () is compared to the mean the higher

Then again the inequality aversion parameter is central, as it allows the researcher to adjust the aversion of inequality. The bigger it get’s the more sensitive is the Atkinson index to the lower part of the distribution

The Theil Index (T) is defined as

We choose for the Atkinson and GE to compare, how the development of inequality changes over time, when comparing the middle part sensitive Gini coefficient to the lower part sensitive Atkinson index and the upper part sensitive Theil index

Ergebnisse

(1) Die Muster sind grundsätzlich sehr ähnlich

(2) Die Krisen 70er und 90er haben die oberen Einkommensschichten stärker getroffen (Theil)

(3) Die oberen Einkommen sind volatiler (Theil schwankt stärker) macht eigentlich Sinn.

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

Here we examine how interpretation can change, when we expand the analysis by using relative distribution methods in comparison to time series of Gini coefficients. We therefore use the published percentiles of the distribution of taxable income from the FTA key figures dataset. 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.[[12]](#footnote-12)

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 2010. This includes all tax subjects (non-taxed and special cases). 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 population and the distribution of 2010 as the comparison population . 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 can be characterized by its cumulative distribution function (CDF). The CDF can be 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

(1)

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

(2)

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. can 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. On the other hand, 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. It is a proper PDF in the sense that it integrates to 1 over the unit interval.

When looking at the relative density of the 2010 versus 2003 tax data (Figure 4, top left) it gets visible that from 2003 to 2010 a moderate polarization occurred, 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[[13]](#footnote-13). 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). Additionally it has several interesting features: 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: Median Index=0.06, Lower Index=0.07, Upper Index=0.05 Comparing the lower and the upper index shows, that the polarization is slightly more driven by the downgrading of the below median percentiles.



## Statistical units

The usual units to assess inequality are households, because the possibility to experience economic wellbeing is strongly connected to households (see section 2.2 on page 4). In tax data, however, the units are represented according to administrative rules and fiscal households don’t represent true households. Tax units rather represent individuals and couples, but only couples, who are married or officially registered. This doesn’t imply that those couples live together, as it is needed to satisfy the definition of a household. On the other hand, is it quite likely that more than one tax unit live in the same household (unmarried/unregistered couples, see Müller and Schoch (2014, 99)). It is therefore not directly possible 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 in regard to the statistical unit, we us special individual tax data from the canton Berne. These data have housing information’s added from personalregisters that allows the construction of a household-identificator 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 households (and pooled income) instead of tax units. By comparing these two distributions, we can test the sensitiveness of inequality in regard to different concepts of statistical units.

Our test shows, that inequality is substantially higher among tax units (Gini=0.48) than among households (Gini=0.44). This can mainly be explained by an upward shift. A lot of single person tax units are indeed not living alone. When we switch from tax units to households, the share of person effectively living alone decreases drastically. 62.3% are taxed as single person tax units, while we identify only a share of 37.4% of actual single person households. This results in pooling of income and the mentioned upward shift. The related relative distribution illustrates the differences (see Figure XY). 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 probably quite similar for the income distribution of Switzerland derived from the aggregated tax statistics. Looking at the published tax statistics for the year 2010 the proportion of single tax units (63.3%) to married tax units (36.6%) are similar than in Bern, meaning that inequality would be lower if assessed when assessed on household level and not among tax units. Additionally the hypothesis can be made, that the bias got stronger in recent decades, and accordingly it was less influencing in times where the overlap of fiscal and real households was bigger.

## Coverage issues

In section 3 we mentioned that survey data are suspected to be biased. The magnitude of this bias in Switzerland, however, is unknown. Strategies to handle this kind of bias are discussed in the literature (Särndal et al., 2003), but require a register for every unit, that is proportional to income. Up to date no such register exists for Switzerland (Müller and Schoch 2014, 43). Currently used micro datasets, which are used for official publications concerning inequality in Switzerland (SILC and HABE) are furthermore confronted with a constructed coverage problem, because these surveys rely on the phone register, which excludes households not having a registered connection.

The concerns of 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 and having income in Switzerland but with a yearly or any other temporary resident permit only. Second, tax statistic separate between those who actually pay taxes and those with an income below a certain threshold that leads to an exemption of direct federal taxes. While information on taxed normal cases is easier available, information on special cases and non-taxed are not always complete.

Another source of incomplete coverage within tax data is tax evasion. Non-fillers are a minor problem, because in Switzerland non-fillers show up in the tax-statistics either way, as long as they are registered. This person gets an imputed income based on an older tax return and information given by employers. Only non-registered non-fillers are not in the records. Not negligible is the circumstance, that individuals 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 can show, that the average level of income tax evasion from 1965 to 1995 varies between 13% and 35%. They suggest that evasion is heavily driven by capital income tax evasion.

When focusing on the available tax statistics, we can distinguish three coverage issues where it is empirically possible, to see 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 show that the inclusion or exclusion of special cases can have a substantial impact on the assessment of income inequality (5.4.2). Lastly we show (5.4.3) how strong inequality is affected by neglecting those subjects, who aren’t taxed.

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

As mentioned it is assumed that tax data covers the extreme part (lower and upper incomes) of an income distribution in a more reliable way than survey data, which is suspected to be confronted with sampling error. To test this hypothesis we perform two tax data comparisons with the Household and consumption Survey (HBS)[[14]](#footnote-14). A successful comparison demands to control 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 we are not able to construct a perfect comparison, we follow two different strategies:

1. We construct a comparison for the Swiss Population for the year 2010, where we use the FTA key figures. To control the difference of statistical units we restrict our analysis to married tax units. Then we construct a pseudo net income, that we believe is comparable to the net income from tax statistics. This includes all relevant income sources (income from labor, wealth and direct social transfer), and equals the gross income, Then we subtract social security contributions and transfers to other households, because these are also subtracted for the calculation of the net income from tax data. Some differences stemming from fiscal deductions remain, which can’t be mirrored within the HBS. Peters (2005) showed that deductions reduced taxable income by almost 30 percent on average, we assume therefore that net income within tax statistics are smaller on average. We assume, that these deductions are proportionally equal across the whole income distribution and hence don’t interfere for the comparison. To get a fair benchmark for the tax data distribution, we apply sampling weights, which are supposed to correct the sampling bias.
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 conceptional difference directly. We improve our comparison further by excluding households with more than seven members, which is the highest number within HBS for 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. Drawbacks of this strategy are that we cannot compare the same years. Tax data represents the year 2012, while the most actual HBS data refers to 2011. We therefore tested, if the distribution based on tax units in Bern differs between 2011 and 2012. No crucial difference could be identified.

Ergebnise

* Gini’s vergleichen (für summary Tabelle). Damit man quantifizieren kann, wie gross der Bias ist
* Relative Distribution beschreiben

### *Influence of special tax subjects*

The FTA distinguishes normal from special cases as described in the data section. To test whether it matters which cases the researcher looks at we want to compare the distributions of normal and special cases. Unfortunately, the FTA stopped to publicly report data for special cases after tax period 1993/94. Therefore we will compare the two distributions for a rather old dataset. However the FTA does 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 2010 as well.

1993/94 a pooled data set of normal and special cases has a slightly higher density at the lower end compared to data based in normal cases only (see figure Figure 4 top middle). Special cases appear to have a slightly lower median income and their distribution is more skewed. Therefore special cases are more polarized than normal cases (see table Table 2) i.e. striving away from the median (positive Median Index of 0.02). This tendency is more pronounced in the lower than upper part of the distribution (Lower Index of 0.029 compared to Upper Index if 0.01). Put simply: the population of special cases 1993/94 hold considerably more tax units with low incomes than does the population of normal cases.

As the special cases consist of a broad mix of individuals it remains unclear which factor explains the differences of both distributions. Possible explanations can be immigrants partly concentrating in lower income percentiles, low income artists who belong to the special cases or a more technical selection effect: tax units not liable for taxation throughout the whole year are special cases; those cases might have lower incomes, e.g. if they moved and stopped working. To get a more complete picture we can look how the two distributions relate to each other in 2010 (see Figure 4 top right).

2010 the picture is similar but more apparent: Special cases appear more frequent around the lower percentiles of the pooled distribution (Lower Index of 0.039), however 2010 there is a more noteworthy effect in the upper part of the distribution (Upper Index of 0.022). According to figure Figure 4 we can attribute this effect to the top percentiles. This gives credibility to the thesis that rich immigrants whose number increased between 1994 and 2010 drive the effect.

### *Influence of non-taxed*

From 1995/1996 to 2010 the number of non-taxed units is reported by the FTA, but not for the years before. This means, we are able to quantify the influence of excluding the non-taxed based on the period from 1995/1996 to 2010.

We therefore calculate three Gini-time-series (see Figure 2 on the right). Excluding zeros leads to a dramatic drop of the Gini coefficient, which is not really surprising. On the other hand inequality is overestimated when assuming non-taxed tax units have zero taxable income. Rather we must assume the taxable income for zeros to lie between zero and the taxation threshold. We address this by presenting a third time-series, where we assume non-taxed to have a taxable income equal to halve the threshold for single tax units (around CHF 8000). This results in slightly lower, more realistic Gini coefficients.

Before tax period 1995/1996 the FTA does not report on non-taxed, hence from then on it is only possible to assess inequality with taxed subjects. To get a feeling how well this group represents the population of Switzerland it is informative to consult estimations on taxed subjects as it is provided by Dell et al. (2007), who used census reports. According to their estimations the share of tax subjects represented in FTA Tax statistic drops from 94% in 1993/1994 to 13.7% back in 1933. It is questionable if analysis based on only a small fraction of the population is appropriate.

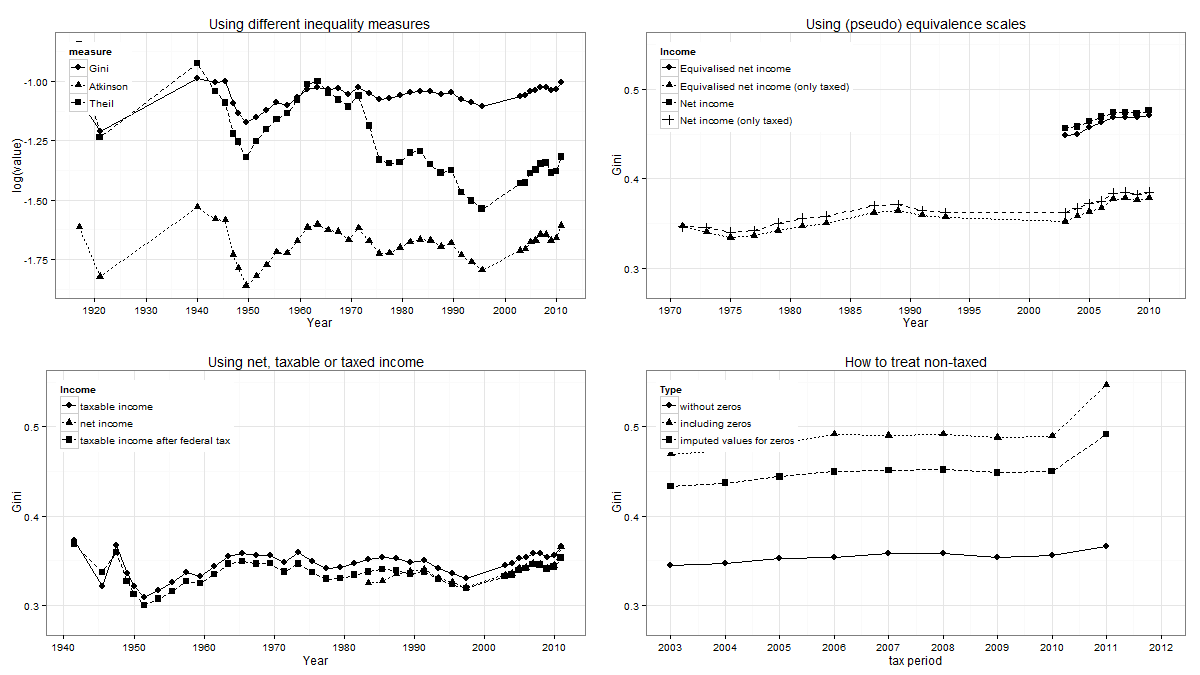


Figure 3: Inequality trends (a) using different inequality measures, b) using different income definitions, c) with/without equivalizing scale and d) including/excluding non-taxed

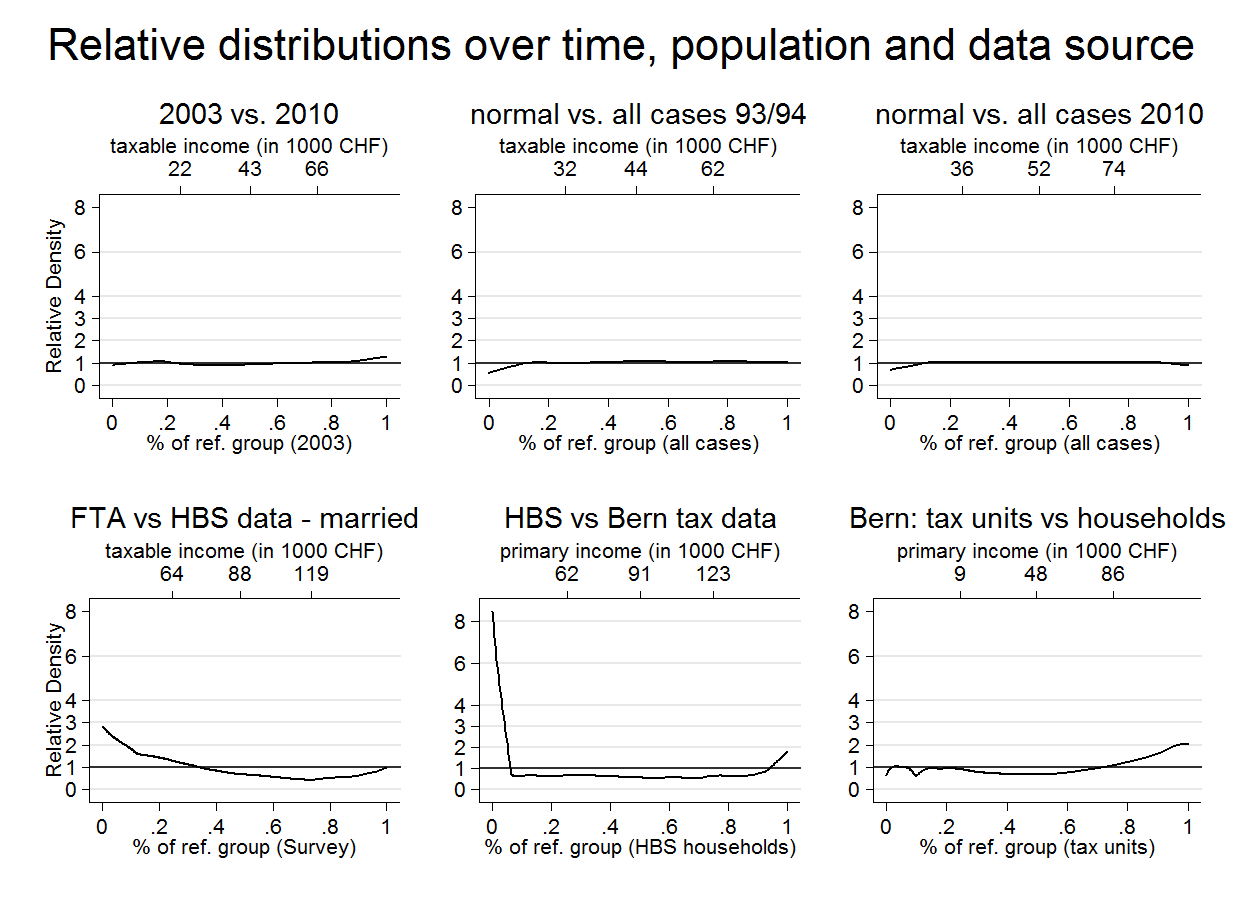


Figure 4: Relative distribution over time, population and data source

# Conclusion

In the course of the paper we checked the suitability of tax data to carry out inequality trend research. After defining ideal theoretical concepts (population coverage, statistical units, income measurement) to depict inequality we evaluate the benefits and downsides of real tax data using swiss tax data provided by the Federal Tax Administration as an example. We finally want to draw both a conclusion of methods and of results, i.e. the actual inequality trend in Switzerland.

Methods conclusion

Concepts within tax data do not equal their ideal theoretical counterpart and neither do other sources of data (like survey data). However we unveiled, which research decisions have a minor, major or crucial impact on calculated outcomes (inequality measures).

|  |  |  |
| --- | --- | --- |
| Decision | Importance | Range (Gini coefficient) |
| Using (pseudo) equivalence scales | Minor | 0.01 |
| Including special cases | Minor | 0.02 |
| Using net, taxable or taxed income | Major | 0.03 |
| How to treat non-taxed | Crucial | 0.14 |

Comparing tax and survey data we revealed another weakness that needs to be declared crucial: Tax units do not properly resemble households. Instead, the common case of cohabitating without marriage is treated as two single incomes. This most certainly leads to biases in the inequality trend as the „single-to-married-ratio“ in Switzerland also varies over time (trend towards more singles).

However, the comparison also reveals one crucial strenght of tax data over survey data, that is the freedom of a sampling bias (middle class bias). When dealing with tax data we finally recommend:

- Impute plausible income values for non-taxed

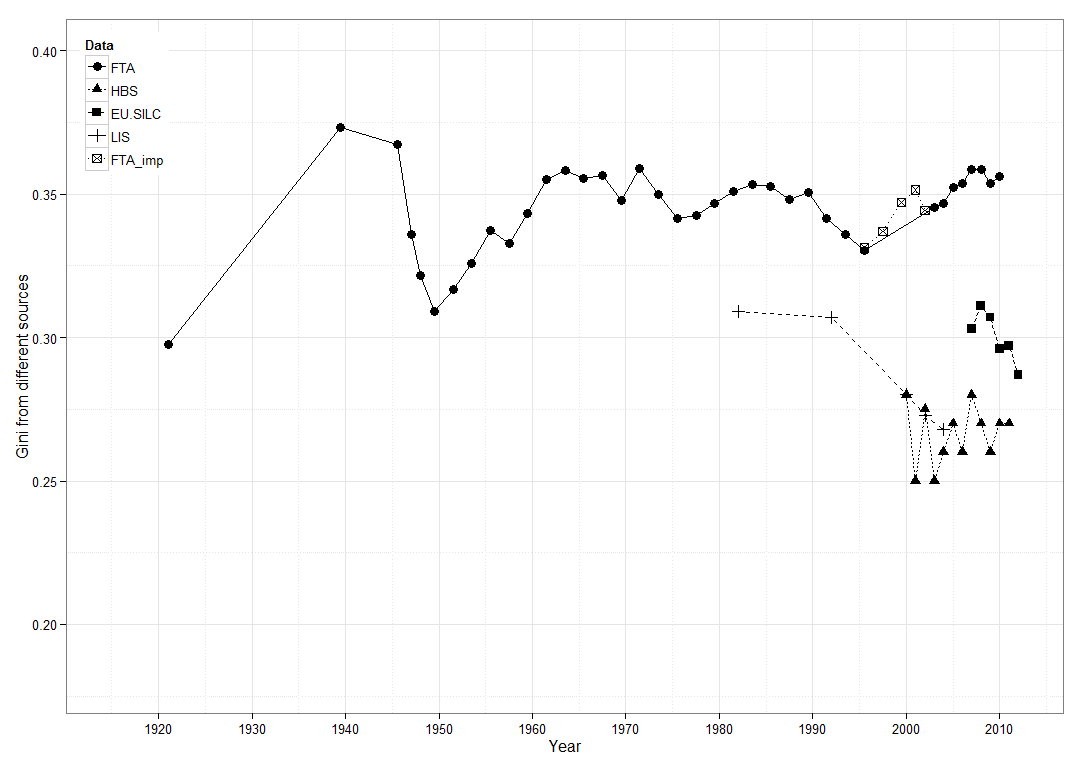
- Include all cases (e.g. the swiss „special cases“)

- Use taxable incomes as basis (after tax deductions if possible)

- Consider separate analyses for singles and married

Inequality trend in Switzerland

As a by-product of our methodological journey we can illustrate some developments for Switzerland. Figure 5 displays the most relevant Gini coefficients that can be calculated for Switzerland. Although we cannot adjust the Gini-coefficients to be perfectly valid, we can discuss the picture against the background of our analyses. For the periods before the second world war it is difficult to draw secure conclusions because these data points are based on unreliable data. 1933 only 13.7% of the population filled in a tax form (Dell et al 2007). During the war we see a tendency towards lower income inequality. This might be attributable to a changed data base as the amount of non-fillers decreased during the war and shortly after. The period after the world war is characterized by strong economic growth as well as an increase in inequality. Our interpretation is that high income percentiles overproportionally profited from the economic upturn. After the oil crises there were alternating phases of social welfare expansion and economic upturns.

 Figure 5: The overall picture of inequality trends in Switzerland

An interesting part of the picture are the years around and past the millennium. Between 1997/98 and 2003 we see a gap in the FTA series caused by the change in the swiss tax system. We imputed the gap by canton-wise interpolating income brackets (cantons switched the tax system in different years so we gain at least some information about the trend within the gap). The spike 2001 might be explained by tax tricks: within the period the tax system changed, individuals were able to save taxes by shifting parts of their incomes into this period. The latest periods show some impact of the 2008 financial crisis. As the incomes include incomes from capital we see a little set-back of the rich percentiles (2008 and 2009), temporarily leading to reversion of the otherwise rising inequality trend that started its rally in the mid nineties.

# Literaturverzeichnis

# Anhang



Figure 6: Bias variation by time and cantons

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.html> [↑](#footnote-ref-3)
4. 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). Between 1993 and 2003 the annual presence taxation (Praenumerando-System) was implemented. Because cantons implemented this change in different years, there is no exact data available for Switzerland in this time period. [↑](#footnote-ref-4)
5. The FTA provides scans starting from the tax period 1947/48. But these scans are not always machine-readable, which made it necessaire to rescan the hard documents. [↑](#footnote-ref-5)
6. 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-6)
7. The ID is constructed out of a register harmonization and is based on a building and an apartment identificationnumber.

   <http://www.bfs.admin.ch/bfs/portal/de/index/news/00/00/06.html> [↑](#footnote-ref-7)
8. This deductions include: professional expenses, travel expenses, interest on debt, alimonies, training costs, two-earner deduction, party contributions, private pension provision “Säule 3a'', buying into the pension plan and sideline deductions, medical expenses, charitable donations, tax-free amounts [↑](#footnote-ref-8)
9. Social deductions include: (married, single-parent households, insurance premiums, interests, deductions for children and supported persons, second earner deductions). [↑](#footnote-ref-9)
10. [↑](#footnote-ref-10)
11. Deductions correspond somehow to obligatory expenses, but itt is definitely not a true disposable income, because important expenses are not covered at all like cantonal and municipal taxes, which represent the bulk of taxes in Switzerland and also the cost of health insurance. [↑](#footnote-ref-11)
12. 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-12)
13. We used reldist for Stata by Ben Jann (2008). [↑](#footnote-ref-13)
14. [↑](#footnote-ref-14)