|  |
| --- |
| Assessing inequality trends with tax statistic  Income Inequality in Switzerland from 1945 to 2010  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)  Oktober 2014  *Abstract*  In many countries inequality trends are inconclusive, because different methodological approaches blur the picture. In this paper we therefore discuss the benefits and shortcomings of tax-statistics concerning the assessment of economic inequality in general and inequality in income in detail, because tax data is becoming more and more popular for inequality analysis. The focus of the theoretical sections lies herein to compare the possibilities of tax data to “state of the art” measurement concepts and show in which way the assessment of inequality is affected by decision researcher have to make. To point out our argument, we have a closer look at Switzerland, where current studies are especially conflicting. Hence we use public tax statistic to assess the trend of income inequality and show, that taxable income is a rather stable income definition, |
| **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 6

3 Comparison of tax data and survey data – overview of advantages and shortcomings 7

4 Conflicting results due to methodological differences in Switzerland? 8

5 Assessing income inequality trend with tax data for Switzerland 10

5.1 Defining Economic resources 10

5.2 Statistical units 11

5.3 Measuring inequality 12

5.4 Coverage issues 14

6 Conclusion 18

7 Literaturverzeichnis 19

8 Anhang 20

# 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 in regard to physical and mental health or even life expectancy. Furthermore the distribution of resources cannot be reduced to the opposition of have 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 (Wilkinson and Pickett, 2009) or social cohesion, 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 advances in technology and methodology, which improve the core sources of inequality research, the household surveys. On the other hand, the labor intensive and expensive surveys around the world are subject to budget cuts and the instrument itself faces problems in form 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 wide back in time. Already Kuznets (1955) used tax data to examine the relationship between economic growth and personal distribution of income. Then it took several decades until Piketty (2001, 2003), Piketty and Saez (2003) made the use of tax data fashionable again. Following his 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 reaching sometimes 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 trend. Both sources predefine the way inequality can be analyzed in fundamental ways. The question arises, to what extend the assessment of inequality is affected by the choice of the data source and consequent possibilities and restrictions. To be able 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 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 in section 5 inequality in income with federal tax statistics for Switzerland and show, how the assessment of inequality is affected by the choices researcher have to make, when working with tax data. Finally in section 6 we summarize by showing which methodological issues are relatively important and which are negligible.

# Standards on Assessing Economic Inequality

We identify four areas, which we think 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 resources 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 2.3 gives an overview on current inequality measures and discus their central advantages and shortcomings. Finally in section 2.4 we discuss the importance of coverage issues, which complete the four areas of inequality relevant methodological 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) condense these ideas into the ICW framework (income, consumption and wealth), which is meant to be an internationally agreed framework on micro-level statistics . and the UNECE/CES recommendations for the 2010 Censuses of Population and Housing (UNECE and EUROSTAT, 2006). 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 and shows the central steps of redistribution, which eventually lead to disposable income: the income measure, which finally shapes the possibility to consume. Within this framework common other income definitions are situated. The assessment of income inequality is strongly influenced by the definition of the income itself. Primary income or disposable incomes for example differ by substantial meaning and by the expected degree of inequality, because the later considers redistribution and the former not.



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

## 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 (see OECD 2013, 173, Buhmann et al. 1988).

## Measuring inequality

Nowadays a plethora of inequality measures exists. Hao and Naiman (2010) provide a good overview on inequality measures and their properties by identifying five main families of measures. 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[[1]](#footnote-1). Here we don’t want to discuss the usefulness of inequality measures in regard to this five principles exhaustively. Rather we want to highlight the main advantages and shortcomings of the most prominent used inequality-measures.

Measures related to *probability distribution* consist on variance or variance based measures, but they fail sometimes 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, variance based measures are commonly avoided.

Widely used in social sciences are *quantile function* based measures like *the quantile ratio or the gini-coefficient*, which is undoubtably the most prominent inequality measure in the academic literature but also 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 easily interpretable. 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. Gini-Coefficent is more sensitive to changes in the middle of the distribution, which is not necessarily a desired feature (see following discussion of Atkinson index). The most important the underlying distributional form of the measured inequality is unknown and it is therefore not possible to see if the measure 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.

Other prominent *quantile function* based measures are *income shares* or more specifically: top income shares. Top income shares are calculated not only with tax data, but corrected using marginal distributions of income and population using external sources (census data and national accounts). This procedure ensures, that the inequality measure is not biased because of non-fillers, who do not appear in the tax statistics (see Dell et. all, 2007). While showing the concentration of economic resources at the top of the distribution, top income shares are silent about other parts of the distribution. Leigh (2007) compares top income shares with other inequality measures and asks, whether they are a useful measure of inequality in a society. He tries to answer this question empirically by comparing measures of inequality based on top income shares with measures of household or family inequality. He finds a strong positive relationship, but concludes (P.600): “top income shares are far from perfect as a measure of distribution of income across society.” Furthermore, top income shares only weakly satisfy the Pigou-Dalton transfer principle (in contrast to the Gini-Coefficent as mentioned above). A transfer from rich to poor will indeed never increase the top income shares, but if the transfer is between individuals, who are either both within the top group or both outside the top group, then the share measure will remain unchanged.

Furthermore the Atkinson index is also a widely used income measure. It is an index that is derived from *social welfare function*. Atkinson (1975:47) noted (p. 47) 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 then again, the flexibility of the sensitivity parameter imposes the need to justify the chosen value thoroughly.

Similar to the Atkinson index, measures derived from *information theory* (e.g. Theil-Index) incorporates a sensitivity parameter that varies in the weight given to inequalities in differing 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 analysis of between‐ and within‐group effects.

Distributional polarization is of particular interest in the study of inequality. However, common inequality indicators (for example Gini or Theil’s index) are not designed to distinguish between differences in the upper and lower tails. Even if the measures register increasing inequality over time, one cannot distinguish a polarization of the distribution (increases in both tails) from upgrading (increases in the upper tail) or downgrading (increases 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. Furthermore, it is based on *the relative distribution* and therefore provides a simple link between what is observed in a graphical display and what is measured by the numerical summary.

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.

## Coverage Issues

Because polices are usually implemented on a national or subnational level, inequality is most often assessed on the respective levels. This implies that inequality-studies usually try to cover the whole population of interest. The success of such a venture is closely related to the way of data collection. We therefore will discuss general issues in regard to coverage issues in the next section, where we will summarize central benefits and shortcomings of tax data opposed to survey 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 and data needs to address. To sum up; if we are interested in the distribution of economic well-being over time, we ideally would look at income, wealth and consumption together; we would look at the disposable resources on household level, ideally the data source is flexible to apply different inequality measures and we would like to have an unbiased estimate of the population measure. Table 1 compares tax-data and survey on this four dimensions and ads 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 |
| Statistical unit | tax units | households |
| Application of inequality measure | restricted | flexible |
| 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. 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 also is often restricted, because only taxable measures are at hand. In regard to income that means, that because often part of direct social transfers (like rents) are included, but – on the other hand – no taxes are subtracted, researchers have to look at neither a pre- nor a postransfer measure (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.

When looking at *statistical units* a second drawback of tax data has to be mentioned. The statistic unit of tax data are tax units, but they do not necessarily correspond to households. Indeed there are constellations where members of the same household hand in several tax forms, like for example an unmarried couple living together. With the change of household-structures over time this issue becomes more and more importance. Here again surveys are able to address the ideal statistical unit in a more appropriate way.

With tax data the *application of inequality measures* can be restricted, if one thinks of tax data as information in aggregated form. Out of that it is possible to estimate the form of the distribution or the quantile function, therefore it is possible to calculate common measures like the gini-coefficent or the income/wealth shares, but measures based on welfare function or information theory need individual data and are therefore not applicable. Survey data usually is collected on household level and therefore leaves every option open a statistician can think of.

A closer look should be taken at *coverage issues*. This is a special thorny task for surveys working with samples, because nonresponse is a major source of bias (Bethlehem et al., 2011). (Korinek et al., 2006) show, that the position in the income distribution influence 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 regard 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. The availability of tax records reaching for several countries back in time for 100 years allowing assessing time trends that cover substantial 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 overt time, because measures and population might be affected by changes in the tax systems or the way tax statistics are reported.

# Conflicting results due to methodological differences in Switzerland?

As we will show in this section, the state of research for Switzerland concerning the trend of income inequality is especially 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: EU-SILC, HBS and LIS-data. Figure XY shows the results stemming from these three sources while looking at Gini of equalized disposable income. Up to the day, EU-SILC (Statistics on Income and Living Conditions) 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. Therefore this times-series doesn’t cover time periods before 2007. As graph XY shows, following the results from SILC, income inequality de-creased from 2007 to 2013. The second important source concerning the distribution of income is the Household Budget Survey (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 graph XY the trend is rather stable. Both time-series (SILC and HBS) cover a relatively short time period. A longer period is covered in the LIS-Data-set (1982-2004). Data-provider for the LIS Data is the Swiss Federal Statistical Office too. In contrast 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 this data Gornick and Jäntti (2013) found for Switzerland a quite substantially decreases in income inequality, 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 in-come 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 income has risen, the top 0.01% share even doubled in the last observed 20 years. A result which 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 comes along with several implications, sometimes these implications get closer the ideal standard, sometimes they are away. First of all, it is assumed that the coverage of top incomes is better in tax data than it is in survey data (non-respondent bias), which is a crucial issue concerning inequality. On the other hand the focus on top income neglects other changes in the distribution of income as it is not possible to see, whether newer concerns like the “hollowing of the middle class” occurred in Switzerland or not, which leads to the second point. Different measure of inequality hampers the comparability. 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 tax data the change in institutional settings is not covered. Fourth, tax data also neglect the household structure, because tax units don’t necessarily correspond to households. Whereas it is unclear how inequality is affected whether one looks at household income or at income of tax units. It can be assumed, that inequality corresponding to different concepts reacts differently on demographic change (change in household structure).

In the end it is unclear, if the different interpretation concerning the trend in income inequality only arises because different concepts were used or because of coverage issues.

# Assessing income inequality trend with tax data for Switzerland

As just showed, the use of different data sources and different concepts can lead to different interpretations, albeit looking at the same time period. In this section we therefore have a closer look at methodical choices that have to be made concerning the four dimensions introduced in section 2 (measurement concepts, statistical units, measuring inequality, coverage Issues) when working with tax data. For this purpose we have a closer look at income tax data for individuals (not legal persons) published by the Swiss Federal Tax Administration (FTA).[[2]](#footnote-2) 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 2010 including 44 tax periods[[3]](#footnote-3). While the FTA provides data in electronic form since 1973 we collected earlier data by scanning hard copies. Data is available for Switzerland plus all cantons and basically covers every individual in Switzerland liable to pay federal taxes. In general data is provided by the FTA in an aggregate form for privacy reasons, i.e. they are classified into numerous income brackets. Additionally the FTA publishes statistical key figures based on the federal tax statistics[[4]](#footnote-4). 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.

In general our main strategy is to apply different possible concepts within one of the defined dimensions while holding other conceptual differences constant. Where possible, we compare results from tax data to results from survey data[[5]](#footnote-5). With this strategy we want to show, where the assessment of inequality is sensitive to conceptual choices and where not. To fulfill the described tasks, we use two statistical techniques. To assess the development of inequality over time, we calculate Gini-coefficients for all possible time points, allowing us to make time trends visible. Because the Gini-coefficient is silent concerning the relevant areas of the distribution subject to a change, we expand the analysis for selected periods with relative distribution methods, which allow an in-depth analysis of distributional differences and therefore compensates the shortcomings of Gini-coefficients. Section 5.4 provides a more thorough discussion on the aspect of different inequality-measures.

## Defining Economic resources

As described in section 2.1 on page 3, the recommendation is to look at income, wealth and consumption simultaneously, when interested in 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. 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.

When focusing on income the central measure reported through 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[[6]](#footnote-6)
* *Taxable income:* net income minus social deductions[[7]](#footnote-7)
* *Taxable income after federal taxes:* This measure is theoretically the closest measure of disposable income[[8]](#footnote-8).

These tax measures don’t correspond directly to theoretical defined measures like primary income (before redistribution) or disposable income (after redistribution). But they can be situated between the pole of market outcome (primary income) and income left for consume (disposable income) (see Figure 1). The measure closest to primary income is net income (less deductions than taxable income). The measure closest to disposable income is taxable income after federal taxes.

Following this three definitions we calculate Gini-coefficients out of the FTA-tax statistics. Because information about net income per income bracket is only reported from 1983/1984 until 2010, we can show differences over this time period.

Was lernen wir aus den Ergebnisen?

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 probably 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. While it is not surprising that federal taxes reduces inequality slightly because of the progressivity of the taxes it is somehow astonishing that adding deductions (from net income to taxable income) increases inequality. This implies deductions to be rather flat not progressive.

## 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. Tax units therefore neither represent individuals in every case nor 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 look (a) at Gini-coefficent with and without equivalence scale and we compare the distribution out of tax data with the weigthed distribution of survey data [besser beschreiben was wir gemacht haben und was aus dem Vergleich raus kommt]

Was zeigen die Ergebnisse

The main issue concerning the unit of analysis is not easy to solve with tax-data, because the concepts of tax-units and households are not perfectly congruent (we will take up this issue, when comparing survey data and the FTA tax data later on.) However we can examine how the assessment of inequality is affected by the implementation of equivalence scale. We do this by looking at Gini-time series for net income with and without implementation of an equivalence scale. The scale is constructed by using information out of tax data. 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 measures are provided directly by the FTA and are not calculated by us. 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).

The implementation of an equivalence scale does not have a major impact on the assessment of inequality (see Figure 3). Over the whole observed time period the two lines, which can be compared, move more or less parallel. Because tax units only approximately depict households, it has to be said, that the implemented equivalence scale automatically has it's drawbacks. This hinders a pure assessment of the effect of a equivalence scale.

The bad household approximation within tax data needs to be addressed in further research. One possible way would be to link tax data to household ID's from the residential register.

## Measuring inequality

Here we examine how interpretation can change, when we expand the analysis by using relative distribution methods and not only look at Gini-Coefficients. We therefore use the published percentiles about the distribution of taxable income on the FTA webpage. We prefer these measures over the calculates 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[[9]](#footnote-9)

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. Gini changed during this time 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.

Formal beschreiben wie aus income brackets Gini-Koeffizienten berechnet werden und Probleme die dabei entstehen (Interpolation) .

To describe how the two distributions are going to be transformed into a relative distribution, we start with defining the two distributions, focusing on a comparison over time (2003 to 2010). This will show us in the end how relative distribution enrich interpretation and is, therefore in our opinion a suitable way to complete trend analysis. We define 2003 as the reference population and distribution of 2010 as the comparison population . represents our measure of interest (taxable income). At 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.

What we got through the above transformation of two distributions is the overall relative probability density. But differences between distributions can be divided into two basic components: changes in location and changes in shape. If the comparative distribution is a simple location-shifted version of the reference distribution, then the difference between the two distributions can be parsimoniously summarized by this shift. Differences that remain after a location adjustment are differences in “shape” (scale, skewness and other distributional characteristics). When both types of shifts are operating, or when factors other than scale are changing in the shape component, we need a way to separate out the various effects. If we want to identify the effect of a location shift and separate it from other changes in the distribution, it is necessary to specify what scale this shift operates on. It is possible to adjust distributions by any measure of central tendency. Here we choose a median location adjustment because the median is a natural, robust and scale invariant unit of measurement. Because our interest lies in analyzing distributional differences concerning the degree of inequality, we will focus in the results section on shape differences and look therefore at the relative distribution after the distributions are adjusted for location differences.

While looking ate relative density of shape differences (see right part of Figure 4) it gets clearly visible, how the two compared distributions differ. The pattern suggest that from 2003 to 2010 a moderate polarization occurred, which is represented in a lower relative density in the middle seven deciles (d.30 to d.90), while the density ratio is higher in the top decile and the two lowest deciles. Comparing 2003 to 2010 tax unites moved to the tails. This pattern is quantified with the inequality indices reported in table 2. Comparing the lower and the upper index shows, that the polarization is slightly stronger driven by the downgrading process.

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-Coefficents and show, how relative distribution measures complete the assessment of inequality trends. For this purpose we use calculate the median relative polarization index (MRP), the upper polarization index (URP) and the lower polarization index (LRP), introduced by Handock 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). And it has several interesting features. MRP can be interpreted in terms of a proportional shift of mass in the distribution from more central to less central values. A value of 0.1, for example, is equivalent to a 10\% population shift from the center of the distribution to the upper and lower quartiles and 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 )

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Table 2: Inequality Indices | | | | |
|  | *Median*  *Index* | *Lower*  *Index* | *Upper*  *Index* | *∆ Gini* |
| 2003 vs. 2010 | 0,058 | 0,072 | 0,045 | 0,025 |
| all vs. special 93/94 | 0,020 | 0,029 | 0,010 | 0,013 |
| all vs. special 2010 | 0,031 | 0,039 | 0,022 | 0,020 |

Ergebnisse beschreiben

## 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. Currently 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 issue of incomplete coverage is less dramatical 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). Essentially this leads to a full representation of the adult population of Switzerland and a complete coverage of the income distribution. This includes a separation of normal cases, which embrace the majority of taxpayers, and the special cases, which cover (not only) foreign nationals living in Switzerland but with a yearly or any other temporary resident permit only. Most important this includes high net wealth individuals taxed according to their expenditures. Special attention has to be paid to tax units with none or very low incomes. Even though they have to hand in a tax return, their income does not show up in the statistics if their income after deductions falls below 15’000 CHF and they are therefore not taxed with direct federal taxes. This is possible for normal and special cases alike. From 1995/1996 until 2010 the number of non-taxed units is reported, but not for the years before. Dell et al. (2007) try to estimate the fraction of non-taxed by comparing the reported numbers of tax units to census reports about the number of adult population. According to their estimations this fraction drops from 94% in 1993/1994 to 63% back in 1945/46.

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. Therefore non-fillers are a minor problem. 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 availabe tax statistics, we can distinguish three further coverage issues where it is empirical possible, to see their relevance for inequality analysis. First we will examine a special differentiation within the population of tax units (5.5.1). There we will show that the inclusion or exclusion of so-called special cases can have a substantial impact on the assessment of income inequality. Then, in section 5.5.2, we show how strong inequality is affected by neglecting those subjects, who aren’t taxed. Finally, we compare tax income distribution to survey data, to see if survey data has a sample bias and if yes, how strong this bias is (section 5.5.3).

### *Special cases*

Furthermore the FTA differentiates between two groups of tax units, so called normal cases and special cases. A normal case is a tax unit residing in a swiss canton without foreign source income and being liable to taxation all year long. All other tax units and very few that are taxed based on the style of living because they don’t work (Pauschalbesteuerte) are special cases.

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 both ends compared to data based in normal cases only (see figure Figure 4). 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).

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).

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.

### *What about non-taxed?*

FTA-Tax statistics sometimes exlcude tax units, which reach below the treshold to be taxed (let's call them zeros). We compare here three Gini-time-series Figure 2. Excluding zeros leads to a dramatical drop of the gini-coefficent, which is not realy surprising. On the other hand inequality is overestimated when assuming non-taxed tax units have zero taxable income. Rather we must assume taxable income for zeros lay between zero and the taxation treshold. We adress this by presenting a third time-series, where we assume non-taxed have a taxable income equal halve the threshold for single tax units (around CHF 8000). This results slightly lower Gini-coefficents.

This exempts all tax units with taxable income below a certain threshold (e.g. 29’200 for a married couple in 2010).

Ergebnis aus kantonaler Analyse: share of zeros over time and cantons

### *Tax data vs Survey Data*

From the discussion in the data section we would expect differences between the income distributions from survey and tax data. Within the FTA data we observe zeros for incomes below the threshold to be liable for federal tax while survey data might cover this range. On the other hand we expect underreporting from both lowest percentiles and highest percentiles (middle class bias) within the survey data. Figure 4 (left) however reveals a more critical issue related to tax data, that is the median location of income compared to survey data. Although we try to measure similar concepts of income, survey data shows a median (93.000 CHF) more than twice as big as tax data (44.600 CHF). The issue here is clearly the assumptions of household composition. While survey data is likely to capture the correct household composition, tax data can only approximate households by using marital status. Splitting the data in married and unmarried tax units underpins this argument. From figure 4 we can see that married tax units (FTA data) and household with married couples (survey data) are better (but still not perfectly) comparable. Figure 4 (right) shows the expected shape difference between the two distributions: survey data has a bias towards the (median-adjusted) 80\% to 90\% percentile (of the FTA data distribution). Top percentiles are badly covered by survey data, suggesting that inequality measures based on survey data might underestimate inequality development that arises from changes in the incomes of the rich. Though, survey data can be of interest if one is interested in the lower 20\% of the income distribution.



Figure 3: Gini over time a) with/without equivalizing scale, b) different income definitions c)

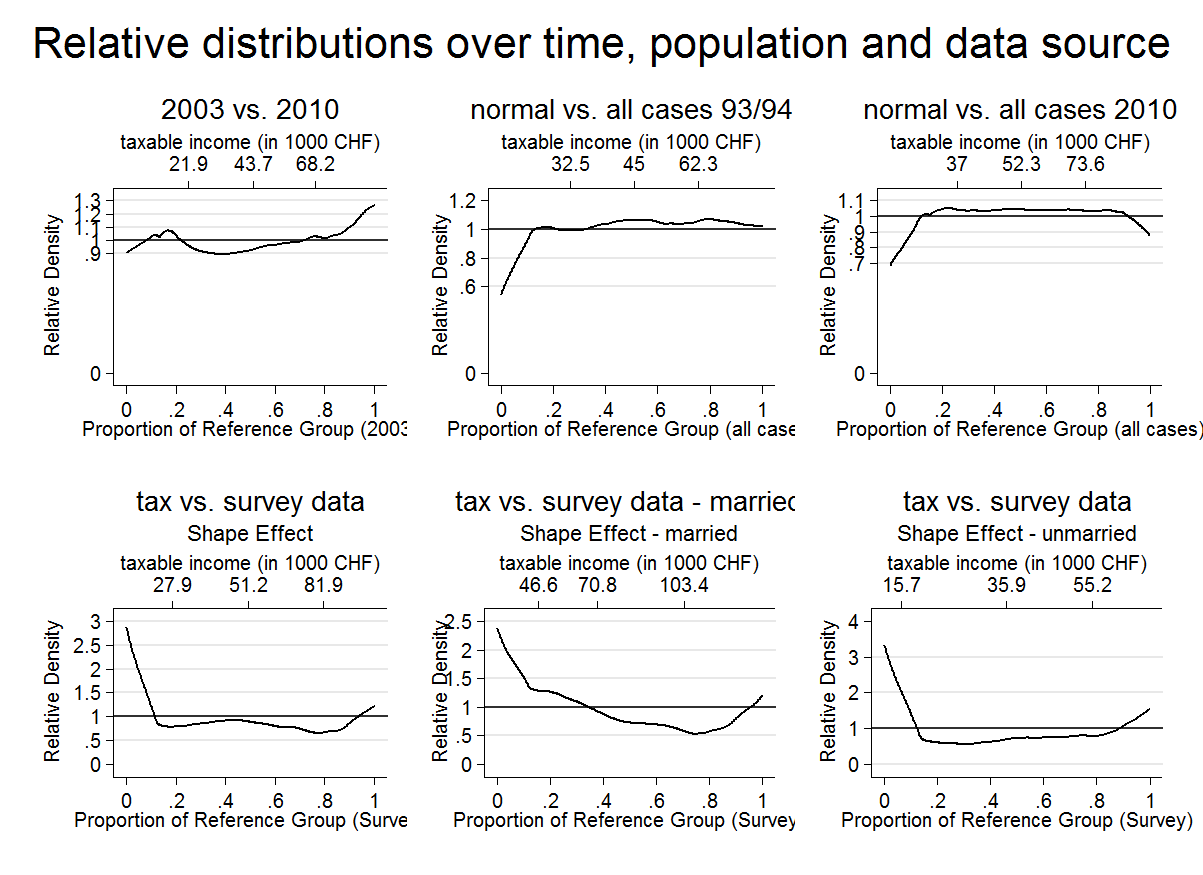


Figure 4: Relative distribution over time, population and data source

# Conclusion

What is the problem. What have we done to solve it

*Methodisches Fazit*

Was ist wichtig, was nicht.

*Inhaltliches Fazit*

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 all 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 disproportionately profited from the economic upturn. After the oil crises there were alternating phases of social welfare expansion and economic upturns.

The more 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 tax system. Our approach to impute the gap is to interpolate the income sums of all FTA-reported income brackets. This is fruitful because cantons switched the tax system in different years so we gain at least some information about the trend within the gap. The spike 2001 can 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.

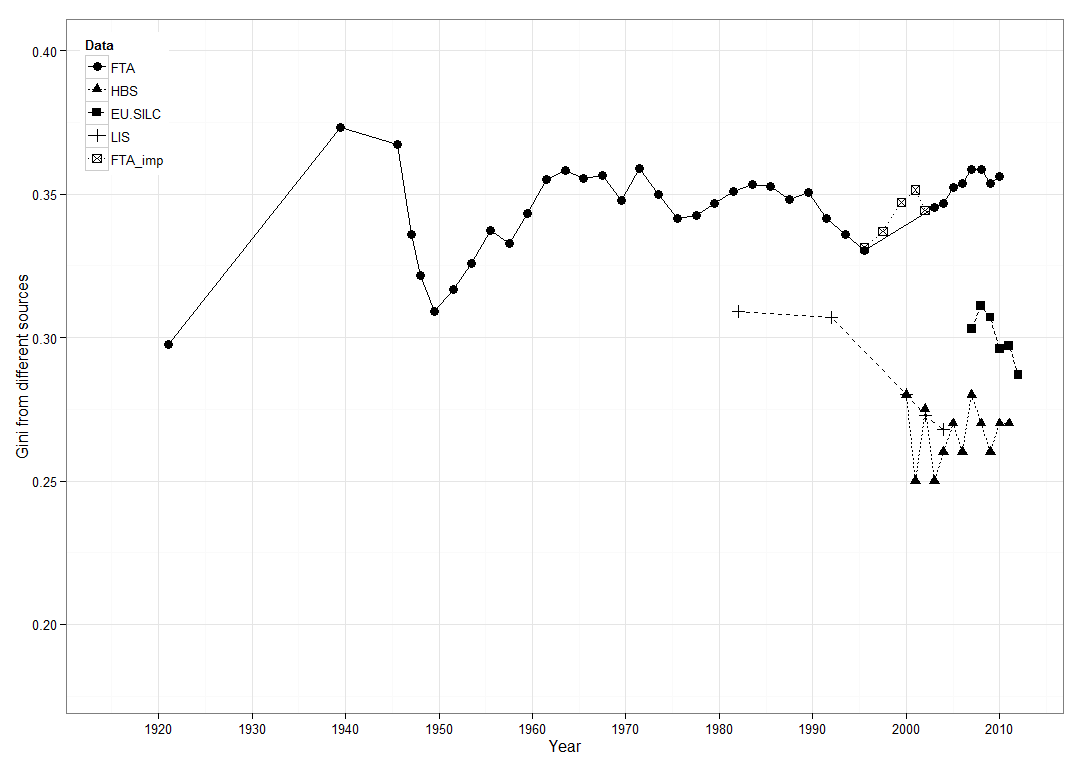


Figure 5: Inequality trend with tax and survey data

After the financial crisis most of the data sources (except EU SILC) state decreasing inequality. Our interpretation is that top percentiles lost relevant parts of their incomes from capital investments.

# Literaturverzeichnis

# Anhang



Figure 6: Bias variation by time and cantons

1. (1) Weak principle of transfers, (2) strong principle of transfers, (3) scale invariance, (4) the principle of population and (5) decomposability. [↑](#footnote-ref-1)
2. <http://www.estv.admin.ch/dokumentation/00075/00076/00701/index.html> [↑](#footnote-ref-2)
3. 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-3)
4. 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-4)
5. We use data from the Household and consumption Survey (HBS) because income is provided on a very detailed base. This allows us to construct measures that are better comparable to income measures derived from tax data. [Wie ist der Measure contructed? Irgendwo muss das genau beschrieben werden] [↑](#footnote-ref-5)
6. 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-6)
7. Social deductions include: (married, single-parent households, insurance premiums, interests, deductions for children and supported persons, second earner deductions). [↑](#footnote-ref-7)
8. Through accounting the reported federal taxes per taxable income bracket, we can construct the taxable income after federal taxes, which is a sort of pseudo disposable income. It is not a true disposable income, because important expenses are not covered like cantonal, municipal taxes, which represent the bulk of taxes in Switzerland and also the cost of health insurance are missing. [↑](#footnote-ref-8)
9. 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. 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-9)