Assessing economic inequality with tax data - Switzerland from 1945 to 2010

Oliver Hümbelin

Bern University of Applied Sciences
oliver.huembelin@bfh.ch

Rudolf Farys

University of Bern
rudolf.farys@soz.unibe.ch

Abstract

There is empirical evidence that economic inequality increased in the majority of western countries over the last decades (Cooperation and Development 2011, Gornick and Jäntti 2013). In Switzerland, however, the development is unclear, as there is evidence for trends in both directions. Part of the inconclusive picture is due to different methodological approaches. In this paper we discuss the role of tax-data concerning the assessment of inequality in income. The focus of the discussion lays herein to show the benefits and shortcomings of tax data compared to current "state of the art" measurement concepts of economic inequality. We present common and new strategies to handle tax data specific methodological difficulties and compare results out of aggregated federal tax statistics to results from the Household Budget Survey (HBS). We can show to which extend survey data underestimate inequality in income. Following the results out of the tax-data Switzerland experienced in slight rise in inequality in recent years, similar to other western countries, but only because of rise in upper percentiles of the income distribution.

I. Introduction

Economic resources can be seen as central indicator for life chances in general and a multitude of outcomes like physical and mental health, life expectancy and crime in particular (Wilkinson and Pickett, 2009). While the study of social inequality can be considered as one of the core subjects of sociology in more recent years the concern about the widening gap was addressed by global leaders (World Economic Forum, 2013) and scholars alike. Empirical evidence acknowledge the supposed trend that economic inequality increased in the majority of western countries over the last decades (OECD 2008, OECD 2011, Gornick and Jäntti 2013, Salverda et al. 2014). Although the rise was not uniform, a common pattern seems to be identifiable, which can be referred to as the "hollowing of the 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."

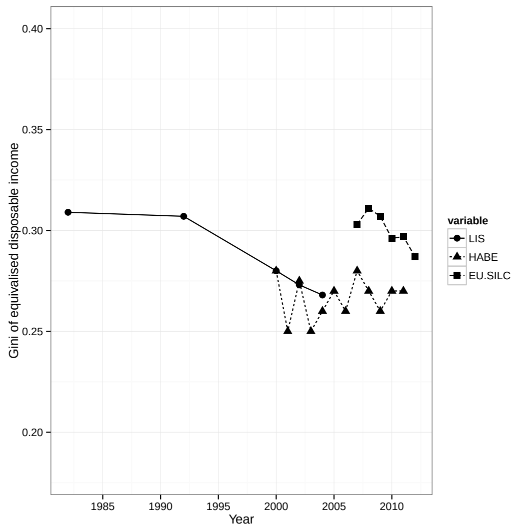
Given the importance of the subject a constant reflection about reliability of empirical data seems appropriate. Atkinson (2013) observes advances in technology and methodology which improves the core sources of inequality research, the household surveys. On the other hand the labor intensive and expensive 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, which can supplement the established survey data studies. 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 (Alvaredo et al., 2014).

As we focus our paper on the case of Switzerland, it is important to embed our work in the context of given publications concerning inequality in income. What is known about Switzerland so far? Looking for official data, three main sources has 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 this three sources while looking at Gini of equivalised disposable income. Up to the day, EU-SILC or 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-countries and countries living around or within the EU. As a Non-EU member Switzerland implemented the instrument not from the beginning (2004) but as from 2007. Therefore this times-series doesn't cover time points before 2007. As graph XY shows, following the results from SILC income inequality decreased from 2007 to 2013.[1] 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. This allows researcher to look at different income concepts like income before and after public transfers. Since 2000 the survey has been conducted on a continuous basis, which allows to look at a consistent time series from 2000 to 2011. As it can be seen from graph XY the trend is rather stable.[2] 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 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).[3]

[1] Data shown in the graph was downloaded from the Eurostat Metadata-portal `http://epp.eurostat.ec.europa.eu/cache/ITY_SDDS/EN/ilc_esms.htm` last accessed 21.Mai 2014.

[2] figures shown in the graph were calculated out of the original datasets, which were kindly provided by Swiss Federal Statistical Office.

[3] A further official database for income distribution is the OECD-Database. It includes measures from Income and Consumption survey as well. Additional data for 2008 is available from EU-Survey of Income and Living Conditions (EU-SILC). But this change in survey is considered as a strict break. Comparison before and since 2008 is not recommended (OECD 2012:315). For the sake of completeness the database construtcted by Deininger and Squire (1996) and the World Income Inequality Database (WIID) have to be mentioned. Both datasets do not contain figures for recent years for Switzerland. A further important database on inequality is the GINI database which has been derived from the GINI Country Reports (Nolan et al., 2014). But this dataset doesn't cover 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 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 published by the Swiss Federal Statistical Office.[4]

Divergence can be explained with several factors. First of all, different data sources were used. The official data providers trust on survey data, whereas the later mentioned publications use tax data. 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 and different units of analysis were used. As it is shown by Modetta and Müller (2012) income distribution is strongly affected by governmental redistribution, reducing inequality substantially. With the focus on tax data the change in institutional settings is not covered. Also neglected is the household structure, 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 react differently on demographic change (change in household structure).

Up to the day, Switzerland can be situated according to the actual level of income inequality in western societies as there is a huge effort to collect data which can be harmonized to comparable measures (see Luxembourg income study, EU-SILC). However, it is unclear how the bias through non-response affects the overall measure of inequality. For the US Atkinson et al. (2009) estimate that CPS survey data fail to capture about half of the overall increase in inequality measured by the Gini coefficient, a result confirmed by Alvaredo (2011). Likewise a long and consistent time-series allowing to identify and explore development patterns on every point of the distribution (not only the top-shares) is missing. Building on recent developments in the field of inequality research, we assess the suitability of the publicly accessible tax data to report inequality and its changes over time. First of all this includes a discussion of the accessible measures in context of a reflection about the state of the art conceptualization of economic resources as an indicator for eco-

[4] There are other studies on Switzerland covering different periods but not the recent years. Flückiger et al. (2002) and also Jeitziner and Peters (2007, 2009) report constant inequality from 1960-1996 respectively from 1995 to 2003. Covering a similar time period Bauer and T. (1994) and Bolzani and Naga (2001) found decreasing inequality. On the other hand Buchmann and Sacchi (1995) and Ernst et al. (2000) found an increase in the 1980s).

nomic well-being. Second, we summarize and apply tax data specific techniques to construct suitable measures of inequality. We expand the given set by an in depth discussion of a newly applied step to handle the incomplete coverage in tax data statistics. Third, we compare our results to results from a relevant Inequality-Survey in Switzerland (HBS) to assess the bias through data source.

II. Data, Measurement concepts and Methods

Studies on inequality have to address several thorny challenges. It starts with answering three crucial questions: First of all one has to define, which concepts should to be looked at. This refers to answering the question about inequality of what. Secondly one has to be clear about the unit of analysis. This refers to answering the question about inequality among whom. Thirdly, one has to choose an appropriate measure of inequality. All these questions are ideally answered considering theory and a given research question. Often it has to be answered in context of a given dataset. Therefore we start this section with a description of the FTA-Tax Data. Based on a review on the literature about the measurement concepts in an ideal world, we discuss the advantages and shortcomings of tax data compared to other data sources - namely survey data. We describe and explain methods and techniques needed to construct time-series of income inequality measures for Switzerland. We will present these time-series in the result section and evaluate the use of the FTA Tax Data to assess the development of inequality in Switzerland

I. Tax statistics in Switzerland

Our data comes from the Swiss Federal Tax Administration (FTA). 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 for cantons and Switzerland. [5] 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 tax unit (individual or household) in Switzerland liable to pay federal taxes. This exempts all tax units with taxable income below a certain threshold (e.g. 15.000 CHF in 2010). 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.

Data is provided by the FTA in an aggregate form for privacy reasons, i.e. they are classified into numerous income brackets. While we don't know individual incomes we still have sufficient information (number of tax units per income bracket plus sum of incomes within each bracket) to calculate percentiles, Gini coefficients and other desired measures. However, equivalization (weighting income by household members) is not possible from these aggregate data. In any parts of the article where we point to equivalized income, the data stem from ready-made calculations (percentiles, Gini coefficients) provided by the FTA[6]. Their measures are calculated from the same source (FTA data) but used the number of children and marriage status as an approximation of household size to calculate inequality

[5] Between 1993 and 2003 there is no exact data available for aggregate Switzerland (but for individual cantons) because of a system change from taxation assessed in arrears (Praenumerando-System) to taxation assessed on current year income (Postnumerando-System) which was implemented by cantons in different years.

[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.

measures that are pseudo equivalized.

The FTA provides two income measures: taxable income and net income. Net income here is an administrative term and means taxable income plus social deductions (children and supported persons) but not including other deductions like donations or health-care costs. Both measures are designed for taxation purposes which might limit the suitability to measure inequality as we will discuss later.

II. Standards for measuring economic resources and inequality

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.[7]. 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. But it is also stated (OECD, 2013, 18):"[...] integrated analysis at the household level has significant data requirements that go beyond the measurement efforts currently undertaken in most countries."[8]

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. Albeit the awareness of an assessment of income, consumption and wealth simultaneously is rising, we focus our analysis on income, which is undoubtedly a crucial indicator of economic well-being. It should be noted, that 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 largely missing in tax data, albeit deductions can in some sense be understood as mandatory consumptions.

Defining income

The assessment of income inequality is influenced by the definition of the income itself. Market income or disposable income for example differ by substantial meaning and by the expected degree of inequality. Therefore the awareness of the analyzed concept is crucial. Terminology can slightly differ, while common concepts can be identified (for detailed discussion see: OECD (2013, 44), United Nations (2011, 24)). Figure **??** shows a stylized framework, which includes a distinction of common income sources [9] 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 central income reported through tax statistics is the taxable income. It includes all

[7]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)

[8]The Luxembourg Wealth Study Database is currently facing this shortcomings by collecting and providing a database following this broader concept of economic well-being. `http://www.lisdatacenter.org/our-data/lws-database/`

[9]Income from production of household services for own consumption is excluded because this income is hard to measure and not covered in the FTA tax data

[10]Means-tested benefits are not taxed and therefore not included in tax data. Income for low income groups are therefore underestimated. However, (Piketty, 2003) note that non-taxable social security benefits grew as a share of personal income

reported incomes (income from employment, income from property and received transfers[10]) minus several deductions. It is therefore neither a pre-transfer income nor a post-transfer income measure. It's rather something in between. As the FTA tax statistics include some but not all deductions[11] it is possible to calculate a sort of total income, which is called "net inomce" (Reineinkommen). As some deductions can be interpreted as compulsory expenses similar to taxes the step towards total income is a step away from the income, which can be used for consumption. Similar when calculating the disposable income out of the taxable income through accounting the reported federal taxes, this is a step towards the income, which is left in the basket for consumption (disposable income). Again it is not a "pure" disposable income, because cantonal, municipal taxes and taxes from churches, which represent the bulk of taxes in Switzerland, are missing.

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 is the 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-earing members, like children). Additionally, there are economies of scale when people share living space and commodities and they therefore benefit from the 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).

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 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.

Measuring inequality or concentration

To be able to make qualifying statements about a distribution or to compare different distributions, the concept of inequality turned out to be the most appropriate and thus the most commonly used dimension. The Gini coefficient is the most known measure and mainly 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 Engelhardt (2000). (1) "**principle of population**": the assessment of inequality is independent of the population size (2) "**Requirement of Bresciani-Turroni**": the measure is sensitive for changes of income shares, but not for absolute changes (e.g. doubling of all income) (3) "*weak principle of transfers*" or "*requirement*

in the US but find that these changes had only a trivial impact on top income shares.

[11]The difference between the real total income and the taxable income are deductions. These 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

of Pigou-Dalton": transfers from richer households lead to a reduction of inequality. However, several drawbacks are reported in the literature. The most important point is, that 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.

The recent wave of tax-data studies do not report Gini-coefficents. Rather top income shares are informed on, which are calculated not only with tax data, but together with external sources to produce the population and personal income control totals (RF 16.06.2014: hä, was? gehts da um die Randverteilung der Bevölkerung?). This procedure ensures, that the inequality measure is not biased because of non-fillers, who do not appear in tax statistics. 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." Top income shares hence inform not completely on how inequality evolves elsewhere in the distribution. 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.

Newer branches of inequality studies emphasize the need for broader measures of inequality, which allow better analyses about the change of inequality and namely statements about the area of change (downgrading/upgrading). The polarizationindex is developed by (**?**) allows for this. Recently this index was applied in the work of (Alderson et al., 2005) and (Alderson and Doran, 2013). This approach is rooted in relative distribution methods. It includes a precise comparison of the shape of two distributions (groups, over time). As the main advantage this approach allows to characterize the change in detail. It is possible to see which parts of the distribution changed, e.g. whether a polarization occurred (reduction of middle class) – which equals an increase of inequality – or this change is driven by either a change in the upper or lower part of the distribution.

The literature mentions several other metrics with desired properties we will not discuss here (see for example Cowell (2000) and Hao and Naiman (2010)).

Population Coverage

Often inequality is assessed on national level, which implies, that studies try to cover the whole population of the country of interest. 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 process can be referred to as the "middle-class bias" (Diekmann, 2009). Missing data in household surveys is therefore not missing at random, which has an impact on the measures of inequality. 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.

Another critical issue with tax data is the problem of tax evasion, which definitely can bias the assessment of inequality. Alvaredo and Saez (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. 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.

III. Comparison of tax data and other data sources - advantages and shortcomings

To define a standard of measuring economic resources and related inequality we discussed different dimensions of demands the data needs to meet. To sum up, ideally we want to measure *disposable income* for *all* swiss *households* yearly for an extended period of time. Tax data has one weakness when it comes to the ideal income measure as we need to work with "taxable income" which is designed to serve taxation purposes. However, both concepts – what needs to be taxed and what can be spent – are related as both address a distinction between necessary and voluntary expenses. By subtracting taxes from taxable income we are confident to satisfyingly come close the theoretically ideal income measure of disposable income. The second most important drawback with tax data is that it does not adequately address households. There are few situations in which the tax unit equals the household, that is for individuals living alone, with married partner and/or deductible children and nobody else. This "classic" household setup however became less common in the last decades so we need to assume statements about household inequality based on taxed data became more and more biased. The swiss law seems to be slightly outdated in this context. Finally however, another advantage of tax data (over survey data) is indeed the observation period. While the FTA data we use range back to 1945, the earliest period of survey data is 2000 (HBS data). The striking feature of the long FTA time series is its consistency. Both population coverage and income measures are rather consistent for the complete observation period[12]. This can not be said for survey data as these are based on samples and therefore require an ideal sampling design or reweighting to be representative and comparable over time. We will see how well sampling and weighting is done in the results section.

IV. Outline of applied methods

In the last section we described the advantages and drawback of taxdata discussing five aspects, which we regard as crucial concering the assessment of inequality. To get a feeling of the importance of these aspects, we exploit the FTA-tax data as far as possible and perform several insightful calculations adressing four of the five mentioned aspects. No further investigation is possible regarding aspect (1) concepts on measuring economic ressource. This is a point which cannot be further addresses with FTA tax data. But for the other four thematic areas, we can provide deeper insight. In general our main strategy is to apply different possible concepts within one conceptional area (income, units, measurement, population coverage) while holding other conceptual differences constant. With this strategy we want to show, the sensitiveness the assessement of inequality is especially if one looks at time trens. In this section we describe which methods we use to produce the results in the next section.

Incomplete coverage of the population (left censored data.) What can be done about the not-taxed? Dell et al. (2007) impute for non-fillers the 20 percentage of the annual average income. This flattens the distribution on the left side, which is not a problem if you are interested in the top income shares, but it would surly affect overall measures of inequality. Furthermore the authors calculate the proportion of non-fillers by estimating the total of tax units out of the population records.

changes in taxation system (switch from annual to biannual taxation) In the mid-1990s a fundamental change in the Swiss tax system took place by switching form the two-years based praenumerando taxation to the one-year based postnumerando taxation. This change was enacted with a transitional period of several years, during which each canton could choose when to adopt the new system. This is why during the transitional period from 1995 to 2003 there is no uniform tax data published on the Swiss level but only data on the cantonal level (Foellmi and Martínez, 2013, 8f).

Estimating percentiles from bracket income tabulation Pareto interpolation

Missing of mean-tested benefits as part of the income -¿ imputation with recommendation for minimum level for basic needs defined by the SKOS.

deductions Dell et al. (2007, 477):" we can check with statistics for 1971-72 (as well as later years) presented both by size of income before deductions and income after deductions that adding back deductions does not introduce any significant error in our estimates." Schaltegger and Gorgas (2011, 5): "..., information on [...] deductions is provided in the tax statistics, thus, we could add the personal deductions to the income data to obtain a consistent series over time". Können wir das auch? Zumindest für gewisse Zeiträume? Das wäre noch gut.

Studies on income try to focus on the disposable income, which subtracts certain expenditures from the primary income. Deductions

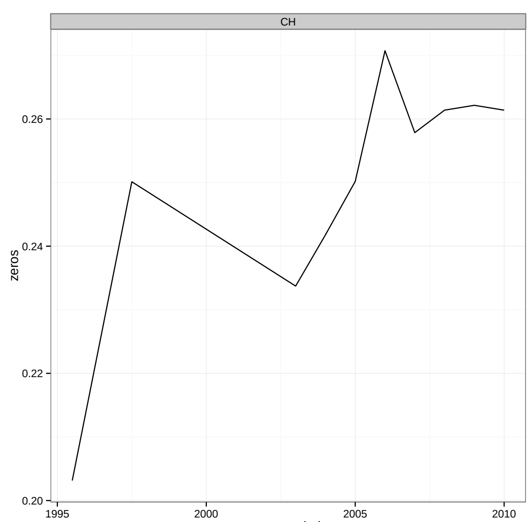
[12] Limitations might however exist due to tax evasion and major changes in the tax system, e.g. 1945. Furthermore, tax units with incomes too small to quality for federal taxation are not documented before 1995 and treated as zeros afterwards making it difficult to track income development in the lower percentiles.

reflect somehow compulsory expenditures and thus taxable income can be seen as a sort of pseudo disposable income. On the other hand deductions can affect the distribution. There are recent studies about the correlation of progressivity and deductions in Switzerland, which examines if deductions have a "perverse redistribution" effect by redistributing income from the lower middle class to the upper middle class (see Peters 2011 and **?**.

IV.1 Non-taxed

Some tax units are missing in the FTA data as they do not qualify for taxation below a certain income threshold. The income distribution is therefore cut at the left end. For most of the observed range it is not documented how many tax units fall into this unobserved category. Since 1995 however the FTA provides the number of people not being taxed (we will call these tax units "zeros"). From 1995 to 2010 we can therefore estimate the bias we introduce when ignoring the zeros. In this section we will try to analyze whether the bias is stable over time (and cantons). If the bias was stable over time, we could still analyze trends and changes of inequality no matter how severe the bias is.

First it is useful to look at the descriptives plotting the share of zeros over time.



The number of zeros increased slightly. Different interpretations seems possible. The increased share of zeros could be driven by the FTA's inflation adjustments or an increase in social deductions. Another plausible explanation is the increasing number of unmarried couples filling in two separate tax forms instead of one so we observe two "small" incomes instead of one "large" income, thus inflating the number of people falling below the taxation threshold.

As we want to report reliably inequality time series we pursue two strategies:

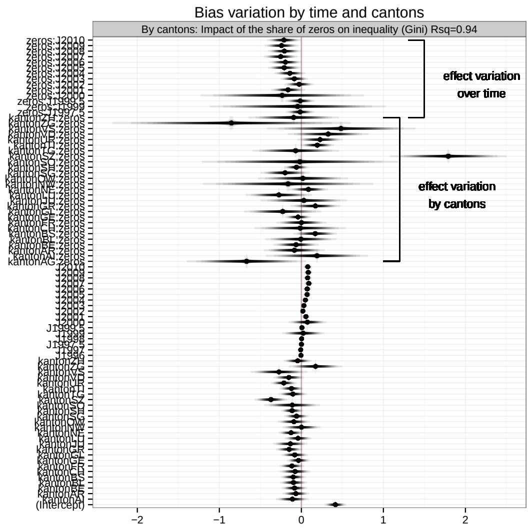
1. Add the zeros as a separate group
2. Carry out a statistical test whether the share of zeros alters the Gini coefficient

By adding the zeros back as a separate group we need to make an assumption about the income distribution within that group. We know these cases had incomes between zero and the threshold (usually around 15.000 CHF for the period 1995-2010). In figure WITHWITHOUTZEROSTIMESERIES we show Gini coefficients calculated under the assumption of the zeros being zero. This is problematic for two reasons: first we know these units might earn more (up to the threshold) so Ginis calculated on this assumptions will be too high. Second we assume the zero group to be perfectly homogenous which on the contrary has a lowering effect on the Gini coefficient. Therefore we must at least assume a distribution for the zeros that has some variance and ranges from zero to the threshold. We decided to use the most simple approach fulfilling both criteria which is to interpolate a uniform distribution between zero and each tax periods threshold.

The second approach investigates whether the FTA time series is suitable to analyze trends of inequality even if the inequality measures are biased. We can do this by regressing the inequality measure (here the Gini coefficient) on the share of zeros. Simply spoken we test if cutting a larger (left) tail of the distribution changes the shape of the distribution. We include Gini coefficient for all cantons in our

[13] Furthermore without cantonal data (but only data for aggregate Switzerland) it is impossible to distinguish whether

analyses to improve the power of our statistical test[13].



The model outputs a test statistic for each canton that tells us whether the variation of the zero-rate over time leads to a significant deviation from the typical "canton gini-level". As the model has a decent fit we are not in great danger of omitted variable bias. Using a joint F-Test we can now test if all canton interactions are zero.

In our case we can clearly reject the hypothesis that all interactions are zero ($p = 0$). This leads to the conclusion that Gini coefficients are biased by the variation of the zero-share. As a minimum a researcher using tax data should therefore control the share of unobserved people, while the best solution is to thoroughly the mechanisms behind a change in the zero share (e.g. this might be due to increased unemployment in one period and due to changes of the tax system in another period.). Fortunately the model coefficient of Switzerland as a whole is not significant suggesting that the cantonal biases cancel each other out. This seems plausible if the phenomenon is e.g. driven by tax competition because most of the "tax optimization" happens within Switzerland.

Figure COEFPLOT can be read as follows: A positive coefficient (e.g. Schwyz) states that we measure higher Gini coefficients in periods with many zeros. We can derive, that the distribution of incomes is more skewed for high incomes than for low incomes. Simply speaking, the contrast between low and middle class is less pronounced than the contrast between middle and top class. One possible explanation would be that incomes stem from two different distributions (populations): 1) local people of Schwyz who follow a less skewed distribution and 2) particularly rich people who moved to Schwyz (to avoid taxes).

A negative coefficient (e.g. Geneva) means the more zeros there are the smaller the Gini measure was compared to other tax periods within that canton (remember this is a fixed-effects model). This is the case we would usually expect: more zeros mask inequality that arises from the bottom.

Another information we can read from the coefficient plot is the (zero-)adjusted inequality development represented by the period dummies. However these values indicate an unweighted change of the Gini coefficient which relates to cantons (not aggregate Switzerland). Fortunately the overall picture of inequality development looks very similar to the plain time series of Gini coefficients.

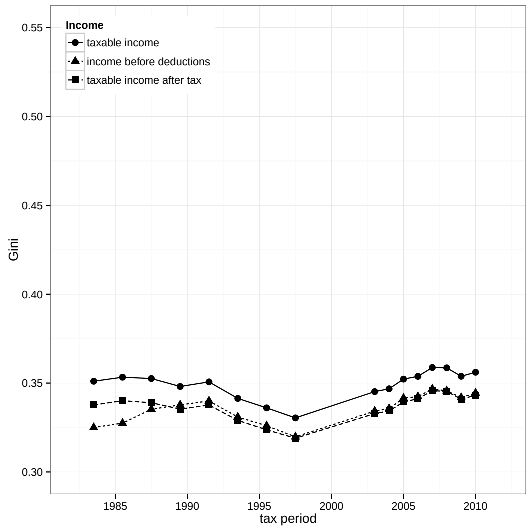
And there is even more to see from the model coefficients: The period-zero-interactions in the model state how an increase of zeros (cutting off a larger piece of the left tail) affects the Gini coefficient in each period compared to 1995. We can see that 2004 to 2010 the (negative) effect is significantly larger than 1995. To simplify: cutting off zeros increasingly leads to an underestimation of inequality measures, probably because the skewness in the left part of the distribution increased. To simplify: in the last decade tax units are increasingly pushed below the observable threshold concealing part of the inequality development. One less dramatic interpretation of this result is the increasing number of unmarried couples filling in two separate tax forms.

inequality changed due to the number of unobserved people or any other variable that changed over time)

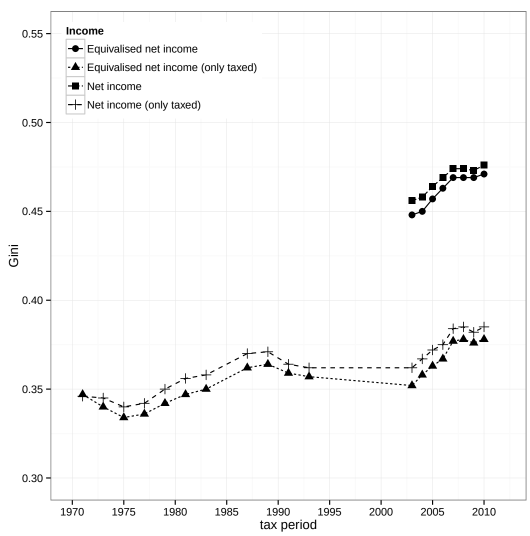
III. Results

I. Defining income

Gini coefficients for taxable income, income before deductions and taxable income after fedaral tax

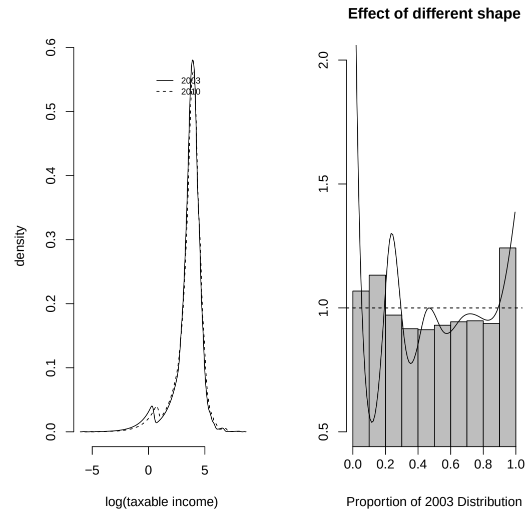


Equivalence scale



II. Measuring inequality

```
## Smoothing using 0.4
```




```
## Loading required package:  splines
```

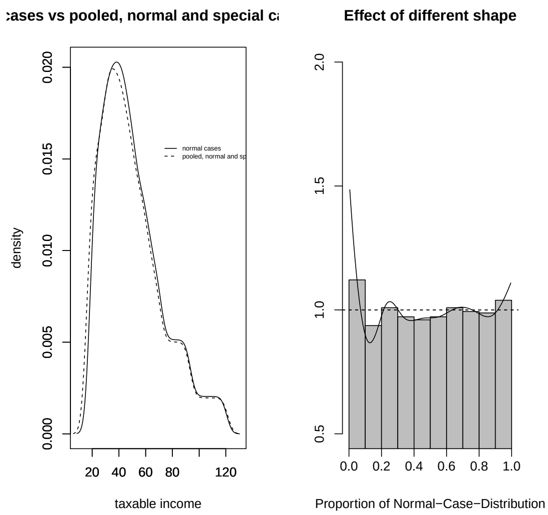
III. Population coverage

III.1 Normal versus special cases

The FTA stopped to publicly report data for special cases after tax period 1993/94. For more recent periods however we can compare already calculated percentiles provided by the FTA[14] which are based on both special and normal cases to normal cases alone. We will have a look at both, the 1993/94 period as the last period where both numbers were reported as well as 2010 where we compare data from the same source (FTA) but based on different reporting (aggregated data vs. percentiles).

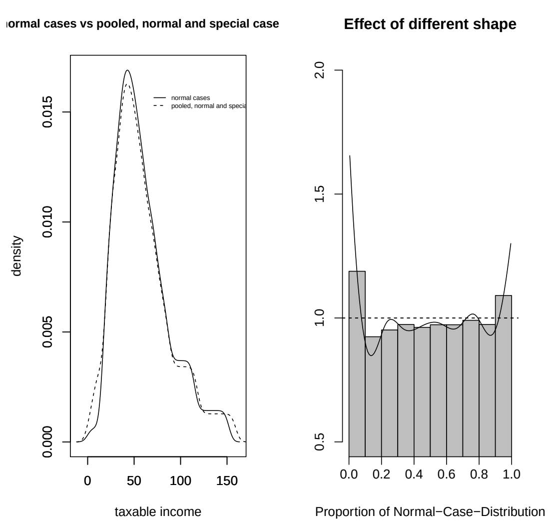
```
## Smoothing using 0.4
```

[14]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.



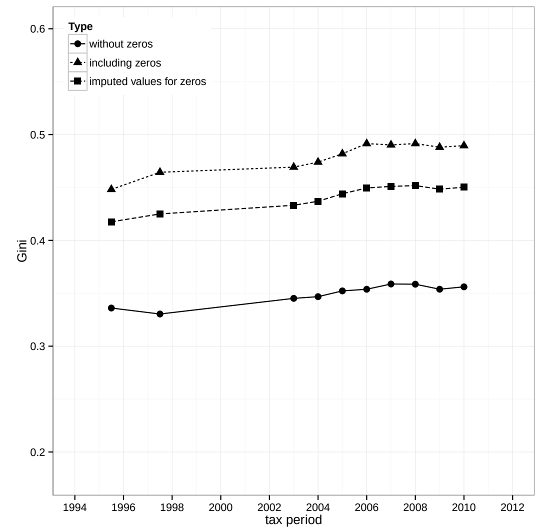
1993/94 including special cases changes the distribution slightly at both ends. Special cases have a slightly lower median income and therefore are more frequent in the lower percentiles and less frequent in the upper percentiles. In addition there is an effect of shape: the distribution of special cases is more skewed so that lower percentiles show lower incomes and higher percentiles show higher incomes compared to just normal cases.

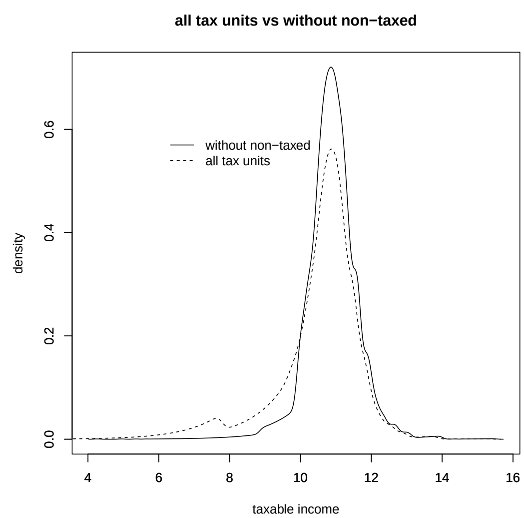
```
## Smoothing using 0.4
```



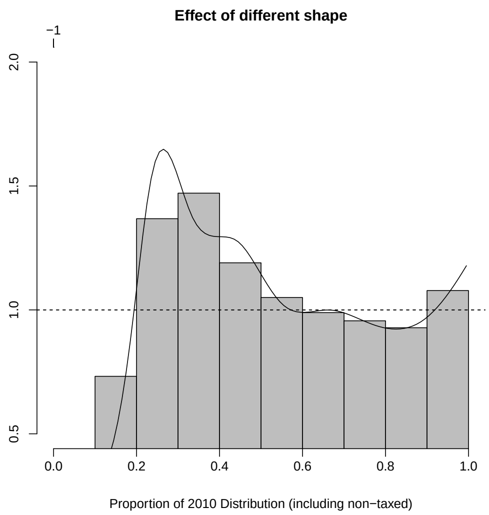
2010 the picture is similar but more apparent: There is a small effect of median shift lowering incomes in the lowest percentiles. The effect of shape however contributes most.

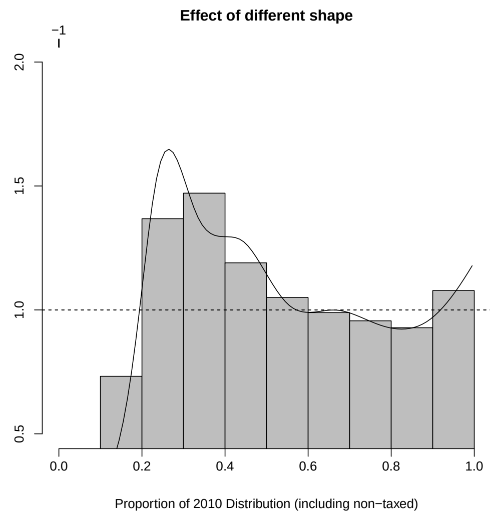
taxable income with and without non-taxed



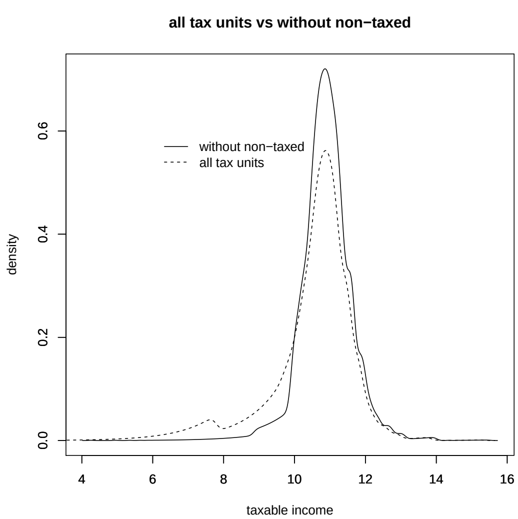



```
## Smoothing using 0.4
```






```
## Error: Objekt 'estvReink' nicht
gefunden
## Error: undefined columns selected
```

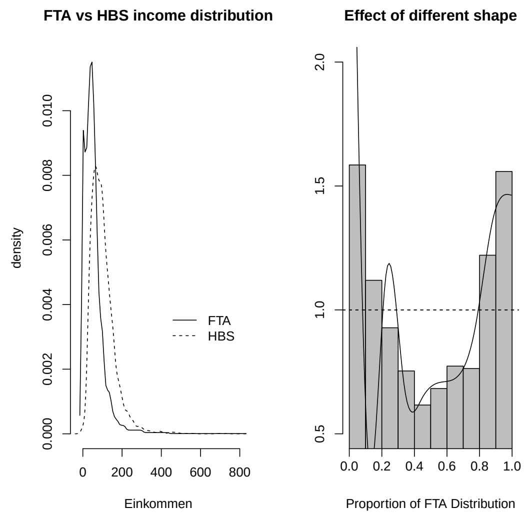


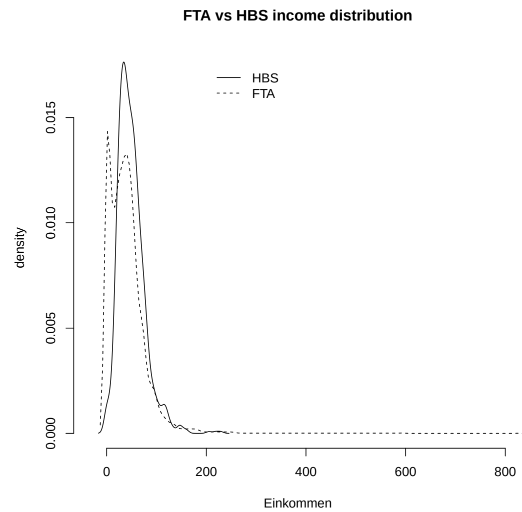
Including zeros leads to significantly higher gini coefficients. However we must keep in mind, that these might be artificially high values as we assume zero income for everyone in the zero group. We can conclude more from the graphic: the ratio between both measures seems to be quite constant although for aggregate Switzerland but there are minor deviations for multiple cantons as well as strong deviations for the cantons Geneva and Ticino. However the problems seem not to result from a shift in the zero-share over time but they are specific for the time-period when the tax system changed.

comparison of tax-data and survey data distribution

```
## Smoothing using 0.4
```

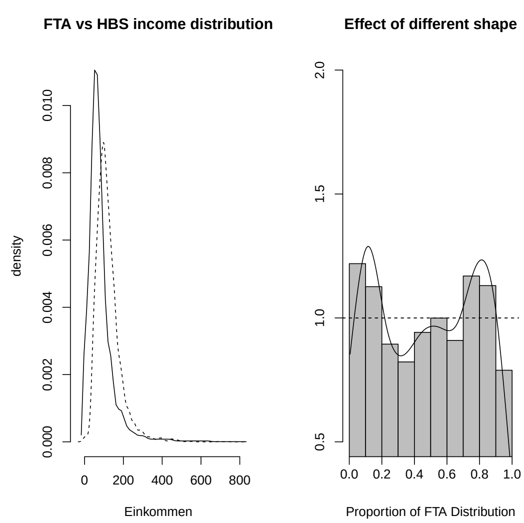
```
## Smoothing using 0.4
```

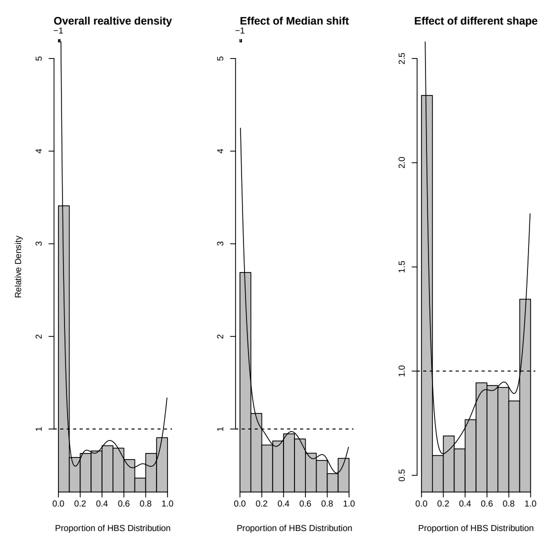




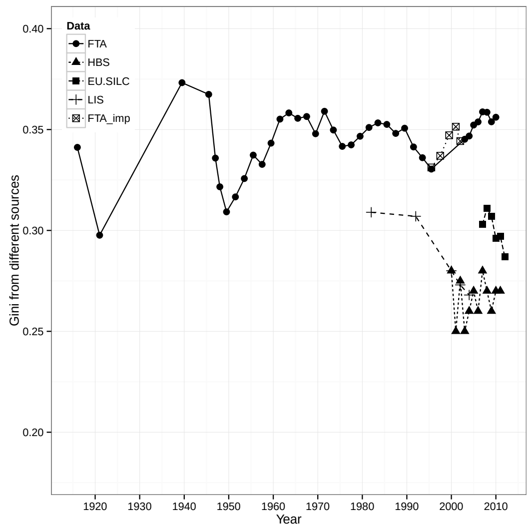

```
## Smoothing using 0.4
## Smoothing using 0.4
## Smoothing using 0.4
```

```
## Smoothing using 0.4
```





IV. Intertemporal comparison



IV. Discussion

V. Acknowledgements

We thank Ben Jann, Robert Fluder and Tobias Fritschi for helpful comments on the article. We also like to thank Stefan Ilic for the preparation of the data set.

References

Organisation for Economic Cooperation and OECD Development. Divided we stand. why inequality keeps rising. Technical report, OECD Publishing, 2011.

Janet C. Gornick and Markus Jäntti. *Income inequality: economic disparities and the middle class in affluent countries*. Stanford University Press, Stanford, 2013. ISBN 9780804778244 0804778248.

Richard G. Wilkinson and Kate E. Pickett. Income inequality and social dysfunction. *Annual Review of Sociology*, 35:493–511, 2009. URL `http://www.annualreviews.org/doi/pdf/10.1146/annurev-soc-070308-115926`.

World Economic Forum. Global risks 2013. eighth edition. Technical report, World Economic Forum., Geneva, 2013.

OECD. Growing unequal? income distribution and poverty in OECD countries. Technical report, OECD Publishing, 2008.

OECD. Divided we stand. why inequality keeps rising. Technical report, OECD Publishing, 2011.

Wiemer Salverda, Brian Nolan, Daniele Checchi, Ive Marx, Abigail McKnight, and Istvan Gyorgy Toth. *Changing Inequalities in Rich Countries: Analytical and Comparative Perspectives*. Oxford University Press, Oxford, 2014. ISBN 0199687439.

Arthur S. Alderson and Kevin Doran. How has income inequality grown? In Janet C Gornick and Markus Jäntti, editors, *Income inequality: economic disparities and the middle class in affluent countries*. Stanford University Press, Stanford, 2013. ISBN 9780804778244 0804778248.

Joseph Eugene Stiglitz. *The price of inequality*. W.W. Norton & Company, New York; London, 2012. ISBN 9780393088694 0393088693.

A. B Atkinson. Foreword. In Janet C Gornick and Markus Jäntti, editors, *Income inequality: economic disparities and the middle class in affluent countries*. Stanford University Press, Stanford, 2013. ISBN 9780804778244 0804778248.

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

Thomas Piketty. *Les hauts revenus en France au XXe siecle: inegalites et redistributions : 1901-1998*. Grasset, Paris, 2001. ISBN 2246616514 9782246616511.

Thomas Piketty. Income inequality in france, 1901-1998. *Journal of Political Economy*, 111(5):1004–1042, 2003. URL http://ideas.repec.org/a/ucp/jpolec/v111y2003i5p1004-1042.html.

Thomas Piketty and Emmanuel Saez. Income inequality in the united states, 1913–1998. *The Quarterly Journal of Economics*, 118(1):1–41, 2003. URL http://qje.oxfordjournals.org/content/118/1/1.short.

A. B Atkinson and Thomas Piketty. *Top incomes over the twentieth century a contrast between continental European and English-speaking countries*. Oxford University Press, Oxford; New York, 2007. ISBN 0191536865 9780191536861 9780199286881 0199286884.

A. B Atkinson and Thomas Piketty. *Top incomes: a global perspective*. Oxford University Press, Oxford, 2010. ISBN 9780199286898 0199286892.

Facundo Alvaredo, A. B Atkinson, T. Piketty, and E. Saez. The world top incomes databas, June 2014. URL http://topincomes.g-mond.parisschoolofeconomics.eu/.

Markus Grabka and Ursina Kuhn. The evolution of income inequality in germany and switzerland since the turn of the millennium. 2012. URL http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2131220.

Klaus Deininger and Lyn Squire. A new data set measuring income inequality. *The World Bank Economic Review*, 10(3):565–591, 1996. URL http://wber.oxfordjournals.org/content/10/3/565.short.

Brian Nolan, Wiemer Salverda, Daniele Checchi, Ive Marx, Abigail McKnight, and Istvan Gyorgy Toth. *Changing Inequalities and Societal Impacts in Rich Countries: Thirty Countries' Experiences*. Oxford University Press, Oxford, 2014. ISBN 0199687420.

F. Dell, Thomas Piketty, and E. Saez. Income and wealth concentration in switzerland over the 20th century. In *Top Incomes Over the Twentieth Century: A Contrast Between Continental European and English-Speaking Countries*, pages 472–500. Oxford University Press, Oxford, May 2007. ISBN 9780191608728.

Reto Foellmi and Isabel Martínez. Volatile top income shares in switzerland? St. Gallen, 2013. URL http://www2.unine.ch/files/content/sites/irene/files/shared/documents/SSES/Martinez.pdf.

Yves Flückiger, José Ramirez, Joseph Deutsch, and Jacques Silber. Inégalité des revenus et ouverture au commerce extérieur. (french). *Einkommensungleichheit und Öffnung des Aussenhandels. (German)*, 28(2):241–265, June 2002. ISSN 03793664.

Bruno Jeitziner and Rudi Peters. Regionale einkommens-und vermögensverteilung in der schweiz: Was sagen die steuerdaten. *Die Volkswirtschaft*, 12, 2007.

Bruno Jeitziner and Rudi Peters. Regionale einkommens-und vermögensverteilung in der schweiz. untersuchung der steuerdaten 2003 und der entwicklung seit 1995/1996. *Bern: Eidgenössische Steuerverwaltung ESTV. Available at: http://www. estv. admin. ch/bundessteuer/dokumentation/00882/00888/index. html*, 2009.

T. Bauer and Spycher T. *Verteilung und Besteuerung des Reichtums im Kanton Zürich: Eine Analyse der Steuerstatistiken 1945-1991*. Büro Bass, Bern, 1994.

Enrico Bolzani and Ramses H. Abul Naga. *La distribution des salaires en Suisse: Quelques observations sur la récession des années 90*. Ecole des HEC/DEEP, 2001. URL http://www.hec.unil.ch/deep/textes/01.04.pdf.

M. Buchmann and S. Sacchi. Zur entwicklung der einkommensungleichheit in der schweiz seit den siebziger jahren: eine forschungsnotiz. *Schweizerische Zeitschrift für Soziologie*, 21(1):177–185, 1995. ISSN 0379-3664. URL http://cat.inist.fr/?aModele=afficheN&cpsidt=3719069.

Carol Ernst, Michael Gerfin, and Robert E. Leu. Inequality trends in the swiss income distribution. *Swiss Journal of Economics and Statistics (SJES)*, 136(III):289–305, 2000. ISSN 0303-9692. URL http://econpapers.repec.org/article/sesarsjes/2000-iii-5.htm.

Caterina Modetta and Bettina Müller. Einkommensungleichheit und staatliche umverteilung. *Bundesamt für Statistik, Neuchâtel*, 2012.

Anthony B. Atkinson, Thomas Piketty, and Emmanuel Saez. Top incomes in the long run of history. Technical report, National Bureau of Economic Research, 2009. URL http://www.nber.org/papers/w15408.

Facundo Alvaredo. A note on the relationship between top income shares and the gini coefficient. *Economics Letters*, 110(3):274–277, 2011. URL http://www.sciencedirect.com/science/article/pii/S0165176510003460.

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

SNA. *System of national accounts 2008*. UN, New York, 2008. URL http://books.google.ch/books?hl=de&lr=&id=gQWmAdwdI8gC&oi=fnd&pg=PA1&dq=System+of+National+Accounts+2008&ots=qwGKDJUo-y&sig=xQ2tLN6Lm81QnkpHTRszi_NbP1w.

United Nations. *Canberra Group. Handbook on Household Income Statistics*. United Nations, Geneva, 2011.

International Labour Organisation (ILO). Final report of the 17th international conference of labour statisticians. Technical report, 2004. URL http://www.ilo.int/global/statistics-and-databases/meetings-and-events/international-conference-of-labour-statisticians/WCMS_087568/lang--en/index.htm.

UNECE and EUROSTAT. Conference of european statisticians recommendations for the 2010 censuses of population and housing. Technical report, United Nations, New York and Geneva, 2006. URL http://www.unece.org/fileadmin/DAM/stats/publications/CES_2010_Census_Recommendations_English.pdf.

Brigitte Buhmann, Lee Rainwater, Guenther Schmaus, and Timothy M. Smeeding. Equivalence scales, well-being, inequality, and poverty: Sensitivity estimates across ten countries using the luxembourg income study (lis) database. *Review of Income and Wealth*, 34(2):115–142, June 1988. ISSN 1475-4991. doi: 10.1111/j.1475-4991.1988.tb00564.x. URL http://onlinelibrary.wiley.com/doi/10.1111/j.1475-4991.1988.tb00564.x/abstract.

André Müller and Tobias Schoch. Vermögenslage der privaten haushalte. Technical report, Bundesamt für Statistik (BFS), Neuchatel, 2014.

Henriette Engelhardt. Modelle zur messung und erklärung personeller einkommensverteilungen. In Prof Dr Ulrich Mueller, Prof Dr Bernhard Nauck, and Prof Dr Andreas Diekmann, editors, *Handbuch der Demographie 2*, pages 1066–1091. Springer Berlin Heidelberg, January 2000. ISBN 978-3-642-63022-4, 978-3-642-57098-8. URL http://link.springer.com/chapter/10.1007/978-3-642-57098-8_9.

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

Arthur S. Alderson, Jason Beckfield, and François Nielsen. Exactly how has income inequality changed? patterns of distributional change in core societies. *International Journal of Comparative Sociology*, 46(5-6):405–423, December 2005. ISSN 0020-7152, 1745-2554. doi: 10.1177/0020715205059208. URL http://cos.sagepub.com/content/46/5-6/405.

F. A. Cowell. Chapter 2 measurement of inequality. In Anthony B. Atkinson and François Bourguignon, editor, *Handbook of Income Distribution*, volume Volume 1, pages 87–166. Elsevier, 2000. ISBN 1574-0056. URL http://www.sciencedirect.com/science/article/pii/S1574005600800056.

Lingxin Hao and Daniel Q. Naiman. *Assessing Inequality*. SAGE, California, New Delhi, London, Singapore, May 2010. ISBN 9781412926294.

Jelke Bethlehem, Fannie Cobben, and Barry Schouten. *Handbook of Nonresponse in Household Surveys*. Wiley, March 2011. ISBN 9780470542798.

Anton Korinek, Johan A. Mistiaen, and Martin Ravallion. Survey nonresponse and the distribution of income. *The Journal of Economic Inequality*, 4(1):33–55, April 2006. ISSN 1569-1721, 1573-8701. doi: 10.1007/s10888-005-1089-4. URL http://link.springer.com/article/10.1007/s10888-005-1089-4.

Andreas Diekmann. *Empirische Sozialforschung: Grundlagen, Methoden, Anwendungen*. Rowohlt-Taschenbuch-Verl., Reinbek bei Hamburg, 2009. ISBN 9783499556784 3499556782.

Carl-Erik Särndal, Bengt Swensson, and Jan Wretman. *Model Assisted Survey Sampling*. Springer, November 2003. ISBN 9780387406206.

Facundo Alvaredo and Emmanuel Saez. Income and wealth concentration in spain from a historical and fiscal perspective. *Journal of the European Economic Association*, 7(5):1140–1167, September 2009. ISSN 1542-4774. doi: 10.1162/JEEA.2009.7.5.1140. URL http://onlinelibrary.wiley.com/doi/10.1162/JEEA.2009.7.5.1140/abstract.

Lars P. Feld and Bruno S. Frey. Tax evasion in switzerland: The roles of deterrence and tax morale. In *Working paper series: Institute for Empirical Research in Economics*, volume 284, pages 123–153. Zürich, 2006.

Christoph A. Schaltegger and Christoph Gorgas. The evolution of top incomes in switzerland over the 20th century. CREMA Working Paper Series 2011-06, Center for Research in Economics, Management and the Arts (CREMA), 2011. URL http://ideas.repec.org/p/cra/wpaper/2011-06.html.

Rudi Peters. Steuerabzüge: Wer profitiert? Technical report, Eidgenössische Steuerverwaltung, Bern, 2011. URL http://z%C3%BCrichwww.mieterverband.ch/fileadmin/alle/Dokumente/Statistik_Studien/Steuerabz%C3%BCge_Zug.pdf.

VI. Appendix