Research in Social Stratification and Mobility 44 (2016) 136–150

www.sciencedirect.com

Research in Social Stratification and Mobility

journal homepage: http://www.elsevier.com/locate/rssm

The suitability of tax data to study trends in inequality—A theoretical and empirical review with tax data from Switzerland

Oliver Hümbelin [a,*], Rudolf Farys [b]

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

ARTICLE INFO

Article history:
Received 19 November 2015
Received in revised form 3 April 2016
Accepted 22 April 2016
Available online 26 April 2016

Keywords:
Tax data
Inequality trend
Income distribution
Switzerland

ABSTRACT

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

1. Introduction

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

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

* Corresponding author.
E-mail addresses: oliver.huembelin@bfh.ch (O. Hümbelin), Rudolf.farys@soz.unibe.ch (R. Farys).

http://dx.doi.org/10.1016/j.rssm.2016.04.004
0276-5624/

the distribution of personal incomes. More recently, Piketty (2001, 2003) and Piketty and Saez (2003) popularized the use of tax data. Following Piketty's approach, many top income studies have been conducted in several countries (Atkinson & Piketty, 2007, 2010). Today, all time series that are based on top income tax statistics are collected and accessible through the World Top Incomes Database (Alvaredo, Atkinson, Piketty, & Saez, 2015).

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

2. Standards in assessing economic inequality

2.1. Income concepts

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

With tax data, *concepts of economic resources and definitions of key measures* are strongly data-driven, because tax data are collected for administrative and not for scientific purposes. Tax statistics are often easily available in an aggregated form, showing tax units per taxable income/wealth brackets, but without any information on individuals. The missing link on the micro level implies therefore that there is no possibility of doing a conjoint analysis of income and wealth. Researchers therefore are only able to analyze the distribution of either income or wealth, but not both simultaneously. In addition, information on consumption is missing entirely. The definition of key measures is often restricted too, because only tax-relevant measures are reported. Taxable incomes in Switzerland for example include direct social transfers (e.g. rents), but no mean-tested benefits (e.g. social assistance) and taxes are not subtracted. Thus, a researcher using taxable income can look at neither a pre- nor a post-transfer measure (see Fig. 1). Taxable income is rather something in between. Furthermore, deductions impose changes to income measures, which can bias the result, when deductions change over time. Aside of transfers and deductions Atkinson et al. (2011) identify changes in taxation of capital income and capital gains that potentially hinder comparability over time especially for top income analysis. The situation is far better with micro tax data. If income and wealth are taxed, a complete conjoint distributional analysis is possible. Key measures can also be constructed quite flexibly, because individual tax data contain information on pre-tax income (before deductions) as well as most important expenditures like taxes. However, detailed information on consumption is still missing. Nonetheless, with respect to concepts of economic resources and definitions of key measures survey data are clearly superior, because concepts and measures can be tailored carefully to the needs of scientists.

2.2. Inequality measures

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

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

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