CONTRASTING METHODOLOGIES FOR THE ESTIMATION OF THE WEALTH DISTRIBUTION

Raun Van Ooijen   Eduard Suari-Andreu   Rob Alessie

Due to observed increasing trends in wealth-to-income ratios and wealth inequality in developed countries (Piketty and Zucman, 2014), it is becoming increasingly important to accurately estimate the wealth distribution. Measuring wealth, especially at the top of the distribution, is in many ways problematic, and there is no perfect method to do so. For this reason, and following the recent work by Bricker *et al* (2016) and Fagareng *et al.* (2016,) in this paper we aim at comparing the different available methods employed in the literature to estimate the wealth distribution, with the intention to help improving them in the future.

The most straightforward method to measure wealth consists of resorting to survey data, which is often the most available for researchers. However, it is well known in the literature that survey data recurrently have the problem of not capturing the top of the distribution accurately enough (Vermeulen, 2017). The very wealthy are usually too few to be properly sampled, and even oversampling them usually leaves out a relevant part of the picture (Meyer, 2015), which is especially a problem when one wants to compute top wealth shares.

An alternative means for measuring the wealth distribution is the so-called capitalization method. This method requires data on capital income for a representative sample of households, and the corresponding rates of return for each income-generating asset. The latter are then used to capitalize each asset category. This method has been employed in the past by Atkinson and Harrison (1978) among others, and has been recently revived by Saez and Zucman (2016). It has the advantage of being better able to capture the top of the distribution. However, it assumes that capitalization factors for each asset category are the same across the wealth distribution, which may lead to an overstatement of inequality if the wealthy enjoy higher rates of return than the poor. Furthermore, there are asset categories that do not generate capital income and thus cannot be directly dealt with using the capitalization method.

For this paper we have available a very unique administrative dataset consisting of the assets and liabilities of 200 thousand Dutch households for the period 2005-2015. The data are provided both by the tax authorities and by financial institutions, and are put together by the Dutch national statistical agency (Statistics Netherlands). Our data also include information on capital income coming from a wide range of assets composing household portfolios. Using the Dutch national accounts we can compute capitalization factors and apply the capitalization method in a similar fashion as in Saez and Zucman (2016). We can then compare the outcome of the capitalization method with the actual wealth data, and thus quantify the bias derived from the assuming homogenous capitalization factors across the distribution. In this aspect our paper is similar to Fagereng *et al.* (2016), who perform a similar comparison using Norwegian data, but only have access to two asset categories, namely stocks and bonds.

In addition we have access to a very high quality survey data, *i.e.* the Dutch National Bank Household Survey. These data provide self reported information on wealth for about 2000 households from 1993 until 2016. We can therefore use it

to estimate the wealth distribution and compare the outcome with the actual wealth data. We can then quantify the bias derived from under-sampling of the very wealthy in household surveys. In this aspect our paper is similar to Bricker *et al.* (2016) who compare the wealth distribution estimated using data from the Survey of Consumer Finances with the outcome of the capitalization method in Saez and Zucman (2016).

The objective of this paper is thus to compare the outcome of the different methods to estimate the wealth distribution with the actual data provided by Statistics Netherlands. The latter reflects the actual wealth distribution with the only flaw that it does not capture offshore wealth. Once the different versions of the wealth distribution are estimated, we can compare what are the implications for the computation of different inequality measures such as the Gini index and top wealth shares and discuss the advantages and limitations of each method. We also propose an alternative method that exploits the large persistence in wealth over time, which can be used to derive household wealth from data on income and assets when data on assets is only available in a few periods.

References:

Atkinson, A., Harrison, A., 1978. *Distribution of Personal Wealth in Britain.* Cambridge University Press.

Bricker, J., Henriques, A., Krimmel, J. and Sabelhaus, J., 2016. Measuring income and wealth at the top using administrative and survey data. *Brookings Papers on Economic Activity*, 2016(1), 261-331.

Fagereng, A., Guiso, L., Malacrino, D. and Pistaferri, L., 2016. Heterogeneity in returns to wealth and the measurement of wealth inequality. *The American Economic Review*, 106(5), 651-655.

Meyer, B.D., Mok, W.K. and Sullivan, J.X., 2015. Household surveys in crisis. *The Journal of Economic Perspectives*, 29(4), 199-226.

Piketty, T., and Zucman, G., 2014. Wealth and Inheritance in the Long Run. CEPR Discussion Paper, 10072.

Saez, E. and Zucman, G., 2016. Wealth inequality in the United States since 1913: Evidence from capitalized income tax data. *The Quarterly Journal of Economics*, 131(2), 519-578.

Vermeulen, P., 2017. How fat is the top tail of the wealth distribution?. *Review of Income and Wealth* (Forthcoming).