Joint distribution of income, consumption and wealth

Friderike Oehler, EUROSTAT, friderike.oehler@ec.europa.eu

ABSTRACT

Poverty indicators purely based on EU-SILC income statistics do not reflect the full picture of household's economic well-being. Consumption and wealth are two additional key dimensions that determine the economic opportunities of people or material inequalities. In recent years, the European Commission and National Statistical Institutes have stressed the need to bring these social dimensions on a par with macroeconomic indicators. Harmonised EU statistics covering the distributional aspects of households' income, consumption and wealth (ICW) in a joint data set could help to reach the goals of the European Union's economic governance framework monitoring economic trends in EU countries. It could further contribute to the impact analysis of fiscal policies.

Statistical matching methods were used to join consumption data from the Household Budget Survey (HBS) to the EU-SILC micro data. In a second step, micro data from the Household Finance and Consumption Survey (HFCS), conducted every three years by the ECB and National Central Banks of the Euro area, can be joint to produce a common distribution of income, consumption and wealth variables. In previous experiments, results produced by different matching methods (random hot-deck, rank hot-deck, distance hot-deck, conditional mean, mixed approach) had been compared. The random hot-deck method turned out to be best suited in most cases, although –like any of the non-parametric methods that do not require additional information- it has the drawback of relying on the Conditional Independence Assumption (CIA).

Experimental statistics produced using the random hot-deck method for matching EU-SILC, HBS and HFCS data for the reference year 2010 were published on Eurostat's webpage in 2017. The method was now applied on EU-SILC and HBS data of several EU countries for the reference year 2015. An ex-post harmonisation process ensured the comparability of potential 'matching variables' (mostly household characteristics) in both datasets. Subsequently, the matching variables were chosen based on their respective Hellinger distance between the recipient (EU-SILC) and the donor (HBS) data sets and their explanatory power of income and consumption. Thus, the final set of matching variables varies from country to country. Multiple imputations of the random hot-deck procedure allow the estimation of confidence intervals indicating the accuracy of the results. For each imputation, the original EU-SILC weights are re-calibrated to fit a number of margins of the matched consumption variable. Fréchet bounds are estimated for the contingency tables to relax the CIA.

The comparison of the original distribution of total consumption in the HBS data set with the matched data shows good results for most countries. The resulting joint distribution of income and consumption data allows, among other analysis, the calculation of median saving rates for households in different income percentiles and with different household characteristics. Preliminary results show insignificant or small differences in aggregate median saving rates of the reference years 2010 and 2015 for most countries, but there are exceptions. Also, differences tend to be much larger for the lowest income quintile. The latter underlines the importance to look not only at aggregate indicators, but also at how these indicators differ

between different households and income groups. In particular, where it comes to measuring the impact of tax or social benefit policies, the whole distribution needs to be considered.

Obviously, any conclusion drawn based on the joint distribution of income, consumption and wealth data should come along with sensitivity checks. A high quality of the original EU-SILC, HBS and HFCS data is the prerequisite for any further analysis.

Keywords: statistical matching, EU-SILC, HBS, HFCS