

# Pareto Tail Estimation in the Presence of Missing Rich in Compiling Distributional National Accounts

Jorrit Zwijnenburg  
OECD  
[Jorrit.Zwijnenburg@oecd.org](mailto:Jorrit.Zwijnenburg@oecd.org)

Joseph Grilli  
OECD

Recent developments on measuring inequality have combined micro-level survey data with national accounts macro statistic aggregates to produce distributional national accounts. However, micro survey data is often considered to underestimate the most wealthy households, despite oversampling methods, due to the limited number and higher non-response rates at the tops of distributions. Pareto distributions are often appended to survey data distributions to represent missing rich households, making assumptions about the presence and shape of the top tail of the distribution or relying on external information.

This paper utilises tests to identify the existence of Pareto distributions in the survey data at item level, whether the distribution has been truncated, and what proportion of the distribution is not covered by observations in the micro data. This is used to estimate the Pareto distribution with corrected parameter values, from which additional observations are drawn and merged with household observations in the survey data to append a Pareto distribution when supported by the data while also retaining a household-level survey dataset.

Using data from the Luxembourg Income Study, the paper tests for the existence of Type 1 and Generalized Pareto distributions for components of primary income, and whether the distribution in the survey is truncated. It then presents a novel sampling method to append the survey data with the estimated item-level Pareto distribution using synthetic households sampled from the unsampled region of the estimated top-tail distribution. This data is then matched to national accounts aggregates using proportional allocation so that measures of inequality can be calculated.

Using data for USA, Japan, and the Netherlands, results show that distributional national accounts produce higher measures of inequality than those found in the micro survey data. The distributional national accounts find greater concentrations of primary income at the top of the income distribution than is found in the unadjusted survey data, with the top 10% of households holding 50.2% in the USA, 33.3% in the Netherlands, and 33.4% in Japan compared to shares for the survey data of 40.7%, 30.9%, and 31.1% respectively. Appending the Pareto distribution is found to have varying sizes of impact across countries, increasing the Gini coefficient by 1.4 percentage points for the USA, 0.4 points for the Netherlands, and 1.7 points for Japan.