

# A Joint Top Income and Wealth Distribution

Junyi Zhu

Deutsche Bundesbank

[junyi.zhu@bundesbank.de](mailto:junyi.zhu@bundesbank.de)

Viktor Steiner

Deutsche Bundesbank

We use the marginal and bivariate parametric distributions to extrapolate both income and wealth distributions from German PHF data. The methodology developed recently in estimating the Pareto distribution incorporating top tail information is adopted in extrapolating the marginal distribution. We then fit a copula model in the top joint distribution observed in PHF. This fully parametric model can help us to perform out-of-sample prediction on the very top of the conditional tail distribution.

The literature in fitting top distribution has evolved to incorporate refined restrictions and top tail information (e.g. rich list; see Eckerstorfer et al. (2015) and Dalitz (2016)). We follow them by estimating the Pareto parameters using goodness-of-fit criteria after Kolmogorov-Smirnov (KS) and Cramer-van Mises (CM) as well as requiring the continuity of Pareto distribution with observed density function. Top rich list from the Manager Magazin for the wealth and fine top percentile list from the administrative income tax data can be taken potentially for checking the extrapolation validity and/or imposing estimating restrictions.

After determining a sensible starting values where Pareto distributions on both marginals become good approximates of the observed ones, we single this top joint distribution out for fitting a copula dependence. Using the empirical cumulative distributions, we select the copula structure among candidates using the information criteria. The estimated top joint distribution is then pinned down by this copula as well as the Pareto marginal distributions estimated above. Fitness is further examined by comparing the simulation from this full parametric model and the empirical copula by contour curves, scatterplot and so on.

Applying this estimated copula, we can answer the following questions:

1. By studying the income/wealth distribution conditional on wealth/income, we can infer how much the unobserved (very) top marginal distribution is driven by undercoverage/nonresponse of the (very) top distribution (ie. when the opposite distribution conditional on these unobserved parts is beyond the top observed value) and how much this is caused by underreporting (ie. when the opposite distribution conditional on these unobserved parts is under the top observed value). This inquiry can assist the improvement of future survey

design.

2. We can replicate the same fitting procedure to the other (lower) part of joint distribution. The dependence structure from the rest of the distribution can be compared with that of the top part. It can be appealing to see, for instance, how much degree the top households are living in a class society – rich income earners are separated from rich capitalists – relative to the other households, which is similar to Aaberge et al. (2013).

3. The joint top distributions between some component incomes (eg. labor vs capital) and wealth or component wealth (eg. nominal vs real and assets vs debts) are worth exploring. This examination can potentially answer the source of wealth inequality and its interaction with income inequality.

4. Last but not the least, this project helps to bridge the gap between survey aggregates and national account counterparts. The other dimension provides a consistency check for the investigation in this aspect since any answer is expected to respect the dependence structure. The results can be informative in guiding the efforts to distribute national accounts aggregate into various dimensions of household.