Estimation of Wealth on Spatially Disaggregated Levels in Germany based on the Household Finance and Consumption Survey

Ann-Kristin Kreutzmann, Philipp Marek, Nicola Salvati, Timo Schmid, Sylvia Harmening

Freie Universität Berlin, Deutsche Bundesbank, University of Pisa

1 Motivation

- Even more than 25 years after the reunification in Germany, differences in financial resources of households in the East and in the West are not ignorable
- However, for the investigation of financial resources it is not sufficient to focus on income since the differences in wealth are even stronger than for income
- While the disposable household income in the East is about 80% of the disposable income in the West, the average household net wealth in East Germany is still only about 43% of the West German level
- The evaluation of financial resources only in the East and West neglects regional differences on smaller geographical levels

How can estimates for German federal states and planning regions be obtained in order to improve the monitoring of household net wealth?

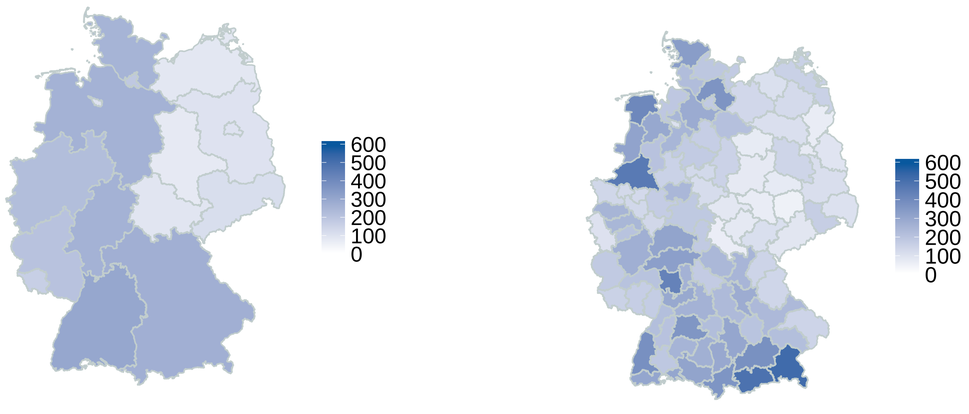

Figure: Map of the estimates for the federal states (left) and for the planning regions (right) of the mean of net wealth based on Forschungsdaten- und Servicezentrum (FDSZ) der Deutschen Bundesbank, Panel on Household Finances 2014, own estimations.

2 The Household Finance and Consumption Survey (HFCS)

The data

- The Eurosystem and several National Statistical Institutes initiated a joint survey called Eurosystem Household Finance and Consumption Survey (HFCS) as a consequence of the financial crisis
- HFCS is the first harmonized survey data across eurozone countries and thus it is unique in enabling cross-country comparisons on a micro-level
- The German part used in this analysis is the Panel on Household Finances (PHF) (FDSZ 2014)

Methodological issues to consider

- Overrepresentation of wealthy households and unit non-response are adjusted by weighting procedures
  ↪ Sampling weights are informative and need to be considered in the estimation
- Item non-response is addressed by multiple imputation (MI) such that the PHF consists of five imputed data sets
  ↪ The additional variance due to the imputation needs to be considered

Regional levels

- The regional levels of interest are the 16 federal states (BL) and the 96 regional planning regions (ROR) in Germany
- The latter allow to observe possible differences between urban and rural areas

Table: Sample sizes of federal states and regional planning regions

	No. domains	Min.	1st Qu.	Median	3rd Qu.	Max.
BL	16	32.00	98.25	188.50	357.50	925.00
ROR	96	9.00	28.00	40.00	65.00	340.00

3 The statistical method

Aims

- increase the reliability of the estimates,
- receive estimates for domains with zero sample size or confidentiality issues,
- be able to handle the complex survey design and the MI,
- return estimates that are consistent with the direct estimator on the regional levels East and West and the national level.

↪ Proposal: Adjusted Fay-Herriot (FH) model (Fay and Herriot 1979) that additionally accounts the MI

1. **Imputation**: Impute the missing values. In the case of the HFCS data set, the imputation is already conducted by the responsible institution in each country.
2. **Analysis (Direct)**: Obtain $\hat{\theta}_{d,m}^{Dir}$ and $var(\hat{\theta}_{d,m}^{Dir})$ for $m = 1, \ldots, M$ where M is the number of imputed data sets. For the PHF data set, M equals to 5.
3. **Pooling (Rubin's rule)**: Obtain $\hat{\theta}_d^{RRDir} = \frac{1}{M}\sum_{m=1}^{M} \hat{\theta}_{d,m}^{Dir}$ and

$$var(\hat{\theta}_d^{RRDir}) = \text{Within Variance} + \text{Between Variance},$$
$$= \frac{1}{M}\sum_{m=1}^{M} var(\hat{\theta}_{d,m}^{Dir}) + \left(1 + \frac{1}{M}\right)\frac{1}{M-1}\sum_{m=1}^{M}(\hat{\theta}_{d,m}^{Dir} - \hat{\theta}_d^{RRDir})^2.$$

4. **Analysis (FH)**: Obtain the Fay-Herriot estimator for multiple imputed data sets (FH-MI) expressed by:

$$\hat{\theta}_d^{FH-MI} = \hat{\gamma}_d \hat{\theta}_d^{RRDir} + (1 - \hat{\gamma}_d) x_d^T \hat{\beta},$$

where $\hat{\gamma}_d = \frac{\hat{\sigma}_u^2}{\hat{\sigma}_u^2 + \hat{\sigma}_{e,d}^2}$ using $\hat{\sigma}_{e,d}^2 = var(\hat{\theta}_d^{RRDir})$.

Analysis (log-transformed FH): Log-transform the direct estimator and modify the variance estimator:

$$\hat{\theta}_d^{RRDir*} = log(\hat{\theta}_d^{RRDir}),$$
$$var(\hat{\theta}_d^{RRDir*}) = \left(\hat{\theta}_d^{RRDir}\right)^{-2} var(\hat{\theta}_d^{RRDir}).$$

Obtain the FH estimator for multiple imputed data sets (FH-MI) in the transformed scale expressed by:

$$\hat{\theta}_d^{FH-MI*} = \hat{\gamma}_d \hat{\theta}_d^{RRDir*} + (1 - \hat{\gamma}_d) x_d^T \hat{\beta},$$

where $\hat{\gamma}_d^* = \frac{\hat{\sigma}_u^2}{\hat{\sigma}_u^2 + \hat{\sigma}_{e,d,*}^2}$ using $\hat{\sigma}_{e,d,*}^2 = var(\hat{\theta}_d^{RRDir*})$.
Back-transform the estimation results to the original scale.

5. **Benchmarking of the FH-MH estimates**: Obtain internal consistency of the FH-MI estimates with the regional estimates for East and West following Datta et al. (2011)

Further methodological considerations

- Small number of domains: Adjusted profile maximum likelihood approach (AMPL) proposed by Li and Lahiri (2010) or REML approach?
  ↪ REML is chosen since the variance estimate is far from 0 and more weight is put on the direct estimate
- Bias-corrections for the back-transformation: Slud and Maiti (2006), Chandra et al. (2017) or naive back-transformation?
  ↪ Since both bias-corrections are only suitable for in-sample domains and the difference to the results using a naive back-transformation is negiligible, the naive back-transformation is chosen

4 Discussion and Outlook

- For the mean it is possible to consider both the weighting procedures and the MI in the Fay-Herriot approach
- The method enables to monitor differences in household net wealth across smaller regions within the East and the West

Further research questions

- Is it possible to transfer the approach to other indicators like the median or the Gini?
- How to handle out-of-sample domains when using a transformed Fay-Herriot model?

References

[1] Fay, R.E. & Herriot, R.A. (1979). *Estimates of Income for Small Places: An Application of James-Stein Procedures to Census Data.*
[2] Forschungsdaten- und Servicezentrum (FDSZ) der Deutschen Bundesbank (2014). *Panel on Household Finances*
[3] Datta, G.S., Ghosh, M., Steorts, R. & Maples, J. (2011). *Bayesian benchmarking with applications to small area estimation* Test, 20(574), 574–588.
[4] Li, H. & Lahiri, P. (2010). *An adjusted maximum likelihood method for solving small area estimation problems* Journal of Multivariate Analysis, 101, 882–892.
[5] Slud, E.V. & Maiti, T. (2006). *Mean-squared error estimation in transformed Fay-Herriot models* Journal of the Royal Statistical Society, Series B, 68(Part 2), 239–257.
[6] Chandra, H., Aditya, K. & Kumar, S. (2017). *Small-area estimation under a log-transformed area-level model* Journal of Statistical Theory and Practice.

For further information

Ann-Kristin Kreutzmann,
Timo Schmid, Sylvia Harmening
ann-kristin.kreutzmann@fu-berlin.de,
timo.schmid@fu-berlin.de,
sylvia.harmening@fu-berlin.de
Department of Economics
Garystr. 21, D-14195 Berlin

Philipp Marek
philipp.marek@bundesbank.de
Deutsche Bundesbank
Wilhelm-Epstein-Str. 14,
60431 Frankfurt am Main

Nicola Salvati
nicola.salvati@unipi.it
University of Pisa
Via Cosimo Ridolfi 10
56124 Pisa