Estimating regional wealth in Germany: How different are west and east really?

To allow for blinded review: do NOT indicate author information or affiliation

Keywords: *Small area estimation, non-response, multiple imputation, HFCS*

1 Introduction

The increasing inequality of private income and wealth requires the redistribution of financial resources. Thus, several financial support schemes allocate budget across countries or regions. One compelling example in this context is the promoted catching-up process of East Germany after the German reunification in 1990. The disposable income of households living in East Germany has been stagnating at about 80% of the West German level [1], whereas private net wealth in the East has caught up only to about 40% of the West German level [2]. However, it is questionable if 25 years after the reunification differences only occur between the East and West or if an analysis on a lower regional level reveals a more diverse picture. This closer investigation of financial resources such as households wealth is of importance for institutions that are responsible to describe the wealth distribution – the Deutsche Bundesbank in Germany – and for policy decisions about the allocation of financial resources across countries or regions. In order to provide a data source for the estimation of private wealth, the European Central bank launched the Household Finance and Consumption Survey (HFCS) for all euro area countries in 2010. This work shows how to receive estimates based on the HFCS for low regional levels in Germany, namely the 16 federal states and 96 planning regions.
The estimates are obtained by the Fay-Herriot model [3]. In order to increase the accuracy of estimates on lower regional levels, direct estimates obtained from survey data are enriched with covariate information from other data sources like registers. The challenge of applying the Fay-Herriot model in this work is the consideration of the data structure while using the SAE approach. First, the skewness of the wealth distribution requires the usage of a log-transformation in the Fay-Herriot approach for the planning regions in order to fulfill the normality assumptions [4, 5]. Second, the present unit and item non-response needs to be taken into account. The unit non-response is adjusted by the data provider using weighting procedures. The produced sampling weights are considered in the Fay-Herriot model by using the weighted direct estimator in the model. The item non-response in the HFCS is handled with multiple imputation [6]. Therefore, our estimates are obtained by using a combination of Rubin's rule and the Fay-Herriot approach. From a theoretical perspective, this leads to a modified (transformed) Fay-Herriot that accounts for the additional uncertainty due to the multiple imputation. Third, for the reporting institution the internal consistency of the regional estimates with the estimate obtained for the national level needs to be ensured by benchmarking the model-based estimates.

2 Data

Since wealth is an important indicator for financial stability, the central banks of 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 in 2007. The survey provides detailed data on various aspects of

household balance sheets and related economic and demographic variables, including income, private pensions, employment and measures of consumption [7, 8]. The HFCS is the first harmonized survey data across eurozone countries and thus it is unique in enabling cross-country comparisons on a micro-level. This work uses the variable household net wealth from the German part of the HFCS, namely the Panel on Household Finances (PHF) [9], which is the data source that the Deutsche Bundesbank, the institution responsible to describe the wealth distribution in Germany, uses. However, for the levels of interest – the federal states and the planning regions – sample sizes are rather small such that these are not reported so far. Furthermore, some regions are either not sampled or the sample size is too small to obtain results not violating confidentiality issues by the Bundesbank. Thus, the application of a direct estimator for the mean of household net wealth bears two issues: First, the direct estimates might be unreliable due to large variances in small areas. Second, direct estimates cannot be reported for regions with zero sample size or for direct estimates violating confidentiality issues. Therefore, small area estimation methods are used in order to combine the sample data with other data sources and thus increase the reliability of the estimates. However, all other issues in the data such as a complex survey design and the conducted multiple imputation need to be considered.

3 Methods

In this work, a benchmarked Fay-Herriot (FH) estimator [3] that additionally accounts for the variability due to the MI is proposed. Furthermore, the skewness of the wealth distribution may lead to violations of the normality assumption of the error terms in the Fay-Herriot model. Thus, a log-transformed Fay-Herriot model is also considered. The steps of the analysis can be summarized as follows:

Step 1. Imputation: Impute the missing values. In the case of the PHF data set, the imputation is already conducted by the Deutsche Bundesbank [10].

Step 2. Analysis (Direct): Obtain the weighted direct estimator $\hat{\theta}^{\text{Dir}}_{d,m}$ and its variance $var(\hat{\theta}^{\text{Dir}}_{d,m})$ for $m = 1, \ldots, M$ where M is the number of imputed data sets. For the PHF data set, M equals to 5.

Step 3. Pooling (Rubin's rule): Obtain $\hat{\theta}^{\text{RRDir}}_d = \frac{1}{M}\sum_{m=1}^{M}\hat{\theta}^{\text{Dir}}_{d,m}$ and

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

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

$$\hat{\theta}^{\text{FH-MI}}_d = \hat{\gamma}_d\hat{\theta}^{\text{RRDir}}_d + (1-\hat{\gamma}_d)\boldsymbol{x}^T_d\hat{\boldsymbol{\beta}},$$

where $\hat{\gamma}_d = \frac{\hat{\sigma}^2_u}{\hat{\sigma}^2_u + \hat{\sigma}^2_{\epsilon_d}}$ is the ratio of the estimated variance of the random effects, $\hat{\sigma}^2_u$, and the total variance, $\boldsymbol{x}_d$ is a $p \times 1$ vector of area-level covariate information and $\boldsymbol{\beta}$ is the vector of regression parameters with dimension $p \times 1$.

Step 4*. Analysis (log-transformed FH): Log-transform the direct estimator and modify the variance estimator (here according to [5]):

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

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

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

where $\hat{\gamma}_d^* = \frac{\hat{\sigma}_u^2}{\hat{\sigma}_u^2 + \hat{\sigma}_{\epsilon_d,*}^2}$.
Back-transform the estimation results to the original scale.

Step 5. Computation of the MSE: Obtain the MSE estimate for $\hat{\theta}_d^{\text{FH-MI}}$. The choice of the MSE estimator depends on the chosen estimation method for $\hat{\sigma}_u^2$ in Step 4.
Step 5*. Computation of the back-transformed MSE: Obtain the MSE estimate for the back-transformed $\hat{\theta}_d^{\text{FH-MI}^*}$. The MSE estimator depends on the chosen bias-correction in Step 4*.

4 Results

Figure 1 shows the regional distribution of benchmarked FH-MI estimates for the states and the planning regions. The map for the federal states shows the fairly known pattern of a clear cut at the former border between East and West. All federal states in the East report an average private net wealth of TEUR 90, which is more than 50% below of the national mean reaching from about TEUR 75 in Saxony-Anhalt to TEUR 110 in Brandenburg. As outlined by the [2], the strong separation with respect to private net wealth can be attributed to differences in financial wealth (TEUR 30 in the East to TEUR 60 in the West), home-ownership (35% to 47%) as well as in the average value of owned dwellings (TEUR 145 to TEUR 250). Furthermore, the estimates also provide evidence for heterogeneity of private wealth across West German federal states.
The analysis on the level of the planning regions enables further insights. Our results provide evidence for heterogeneity in West Germany. The regions around economically prosperous cities in the West – namely Munich, Frankfurt and Hamburg – report the highest private wealth levels in Germany. The top two regions (Südostoberbayern and Oberland) are located in the South of Munich, where average private net wealth is around TEUR 520. In the East, the regions with highest private wealth are located in the South-West of Berlin (Havelland-Fläming: TEUR 130), at the Baltic Sea (Vorpommern: TEUR 143) as well as in the region around the city of Dresden (Oberes Elbtal/Osterzgebirge: TEUR 154).
The results for German planning regions show that wealth is geographically dispersed in both parts of the country. Furthermore, we can show that private wealth in all East German planning regions still remains far below the national average. However, the wealthiest planning regions in the East report higher private wealth figures than the West German regions with lowest private wealth estimates.

5 Conclusions

While the HFCS is, so far, used to report national estimates for private wealth, this work shows how to estimate average HH net wealth for low regional levels, namely the 16 federal states and 96 planning regions in Germany. We contribute to the literature by estimating the regional distribution of private wealth in Germany by means of a modified FH model, which

a) accounts for the skewness of the wealth distribution by means of a log-transformation in the estimation,

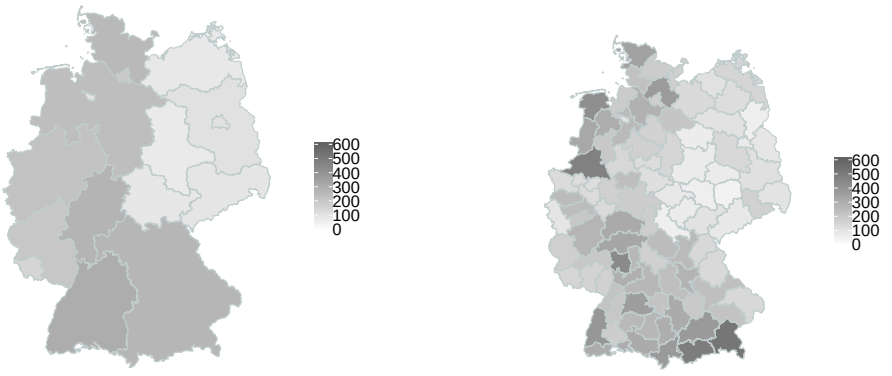

Figure 1: Map of the benchmarked FH-MI estimates for the federal states (left) and for the planning regions (right) of the mean of HH net wealth in TEUR based on Forschungsdaten- und Servicezentrum (FDSZ) der Deutschen Bundesbank [9], Panel on Household Finances (PHF) 2014, own estimations.

b) accounts for multiple imputation, and

c) ensures internal consistency of the estimates with a national benchmark.

The results of the estimation are very insightful and contribute to the discussion on the distribution of private wealth, which has strikingly gained attention in the scientific literature as well as in the public debate in recent years. Even 25 years after the German reunification, there is clear cut at the former border with respect to private wealth. However, the wealthiest planning regions in the East report higher private wealth figures than the West German regions with lowest private wealth estimates. This important finding is highly relevant in the context of the discussion of a prolongation of the support scheme, *Solidarity Pact II*, assigning support exclusively to regions located in the East.
Even though the application in this work concentrates on Germany, the theory is easily transferable to the data of other countries attending the HFCS as well as other surveys that use multiple imputation in order to account for item non-response and have a similar data structure.

References

[1] U. Blum, H. S. Buscher, H. Gabrisch, J. Günther, G. Heimpold, C. Lang, U. Ludwig, M. T. W. Rosenfeld, and L. Schneider. Ostdeutschlands Transformation seit 1990 im Spiegel wirtschaftlicher und sozialer Indikatoren. IWH-Sonderheft, Institut für Wirtschaftsforschung Halle, 2010.

[2] Deutsche Bundesbank. Vermögen und Finanzen privater Haushalte in Deutschland: Ergebnisse der Vermögensbefragung 2014. Monatsbericht, Deutsche Bundesbank, 2016.

[3] R. E. Fay and R. A. Herriot. Estimates of income for small places: An application of James-Stein procedures to census data. *Journal of the American Statistical Association*, 74(366):269–277, 1979.

[4] E.V. Slud and T. Maiti. Mean-squared error estimation in transformed Fay-Herriot models. *Journal of the Royal Statistical Society Series B*, 68(2):239–257, 2006.

[5] A. Neves, D. Silva, and S. Correa. Small domain estimation for the Brazilian service sector survey. *ESTADÍSTICA*, 65(185):13–37, 2013.

[6] D. B. Rubin. *Multiple imputation for nonresponse in surveys*. John Wiley & Sons, Hoboken, 1987.

[7] Eurosystem Household Finance and Consumption Network. The Eurosystem Household Finance and Consumption Survey - methodological report for the first wave. Statistics Paper Series, European Central Bank, 2013.

[8] Eurosystem Household Finance and Consumption Network. The Eurosystem Household Finance and Consumption Survey - results from the first wave. Statistics Paper Series, European Central Bank, 2013.

[9] Forschungsdaten- und Servicezentrum (FDSZ) der Deutschen Bundesbank. Panel on Household Finances (PHF). DOI: 10.12757/Bbk.PHF.02.02.01, 2014. Plus one additional attribute (district code).

[10] M. Eisele and J. Zhu. Multiple imputation in a complex household survey - the German Panel on Household Finances (PHF): Challenges and solutions. User Guide, Deutsche Bundesbank, 2013.