STATISTICS IN TRANSITION new series and *SURVEY METHODOLOGY*
Joint Issue: Small Area Estimation 2014
Vol. 17, No. 1, pp. 25–40

SMALL AREA ESTIMATION IN THE GERMAN CENSUS 2011

Ralf Münnich[1], Jan Pablo Burgard[2], Siegfried Gabler[3], Matthias Ganninger[4], Jan-Philipp Kolb[5]

ABSTRACT

In 2011, Germany conducted the first census after the reunification. In contrast to a classical census, a register-assisted census was implemented using population register data and an additional sample. This paper provides an overview of how the sampling design recommendations were set up in order to fulfil legal requirements and to guarantee an optimal but still flexible source of information. The aim was to develop a design that fosters an accurate estimation of the main objective of the census, the total population counts. Further, the design should also adequately support the application of small area estimation methods. Some empirical results are given to provide an assessment of selected methods. The research was conducted within the German Census Sampling and Estimation research project, financially supported by the German Federal Statistical Office.

Key words: register-assisted census, small area estimation, design optimisation, relative root mean squared error.

1. Introduction

The Census 2011 was the first after the German reunification. The last census in the Federal Republic of Germany was implemented in 1987, whereas the last census in the former German Democratic Republic (GDR) was conducted in 1981. For the first time in German Census history and for the first common census after the reunification, it was decided to conduct the Census in 2011 as a register-assisted census. The main sources of information are population registers. Additionally, a sample of approximately 10% of the population is drawn for two purposes. First, to assess the number of over- and under-counts in the registers aiming at deriving

[1] University of Trier. muennich@uni-trier.de.
[2] University of Trier. burgardj@uni-trier.de.
[3] GESIS Mannheim. siegfried.gabler@gesis.org.
[4] Roche Diagnostics. matthias@ganninger.de.
[5] GESIS Mannheim. jan-philipp.kolb@gesis.org.

the total census counts. Second, the sample information is used to estimate variables that were not included in the population registers. Certainly, the population register information can also be used as a source of auxiliary information.

The lowest level of official territorial division in Germany is the communities that have varying numbers of inhabitants. At the time of the Census 2011 the most populated community is Berlin but there are also 5 communities with less than 20 inhabitants. The most important target of the German Census was the determination of the official population sizes for each of the 11,399 communities. Due to the new census mode, adequate methodologies had to be developed, including sampling design and estimation strategies. Hence, a research project was granted by the Federal Ministry of the Interior and the German Federal Statistical Office to investigate an appropriate sampling design taking into account the German administration structure.

In addition to developing and recommending an optimal sampling design under the given circumstances, estimation strategies had to be developed that can be used in connection with this sampling design. In order to appropriately investigate the interplay of sampling design and estimation strategies, a close-to-reality universe of synthetic data had to be developed which was based on real register data. This universe was used as a sound basis for carrying out an extensive simulation. This article addresses the key findings of the sampling and estimation research project.

2. Objectives and frame of the German Census 2011

The Census 2011 sample had to be drawn to fulfil two main objectives:

Objective 1 Determination of the official population size for each community, i.e. estimating census over- and under-counts in order to derive the population sizes,

Objective 2 Estimation of key figures for additional variables.

Extensive planning preceded the realisation of the Census 2011. A census test, implemented in 2001, served as a preparation to gain initial information for the concept of a register-assisted census in Germany. The census law was launched in 2006. In the contract of the coalition, the reduction of burden and the use of modern methods were stipulated. The aim was to reduce costs without losing quality of important figures. The resulting figures should serve as a basis for administrative planning and decisions, and especially for financial adjustments between federal states. Therefore, it was necessary to reach a high level of quality.

The census law covered several important settings of the register-assisted census like variables of interest, rough description of the register-assisted structure, the sampling units, and quality margins. The quality constraints were especially important for objective 1 due to the importance of the population figures. The second objective was the estimation of variables not contained in the registers, e.g. on housing and living conditions. The relevant source of information for estimating over- and under-counts as well as for non-register variables is based on a sample of addresses drawn from an address register containing all buildings and dwellings. Here, an address is defined as an address with housing space. There are addresses with only one inhabitant but also addresses with several hundred flats and inhabitants. A detailed description of the frame is provided by Kleber et al. (2009) and Bechtold (2013).

Numerous legal and administrative criteria of constraints had to be considered for the development of the underlying sampling design. As sampling units, complete addresses had to be drawn from the address register, i.e. all persons and households living at the given address. The address register was built exclusively for the Census 2011. In Germany, in general, one house is considered as an address. Obviously, the sampling units differed in size considerably, i.e. the variation of the number of inhabitants was very high which may have yielded a clustering effect for sampling. As an upper bound, 7.9 million inhabitants were sampled which covers approximately 10% of the population, not necessarily of the addresses. The design had to be as efficient as possible while considering the accuracy objectives for the estimation of the population counts stated in the census law. Further, feasibility was an important criterion that had to be considered.

Finally, the estimation had to be carried out for small areas and domains. The main areas of interest were districts or communities with at least 10,000 inhabitants. As domains, the main population subgroups were of interest. One of the main tasks at the beginning of the project was to find a coherent way of defining areas for sampling that considered the hierarchical structure of 16 federal states, 412 districts, and 11,339 communities.

As already mentioned, the sizes of communities in Germany differed greatly. Within the census law, it was stated that communities with at least 10,000 inhabitants played a major role in administrative and planning processes such that a different kind of inspection of over- and under-counts had to take place. It was important to consider these differences in the sampling design. Therefore, the first step was to build the so-called sampling points (SMP). These sampling points should be units with at least 10,000 registered persons that yield a frame of areas from which samples were drawn, according to the following scheme:

- Type 0 (SDT): Parts of communities with more than 400,000 inhabitants,
- Type 1 (GEM): Communities with at least 10,000 inhabitants and not of type 0,
- Type 2 (VBG): Collection of small communities within districts that together covered 10,000 inhabitants and more,
- Type 3 (KRS): Collection of the rest of small communities within a district.

With these settings, Germany was completely split into regional structures that considered all administrative and legal constraints and which could be used directly for optimizing the sampling design. The distribution of the sampling point types in Germany is illustrated by Figure 1. Sampling points of type 0 are depicted in yellow, sampling points of type 1 are coloured red. Sampling points of type 2 are coloured green and sampling points of type 3 are blue in colour.

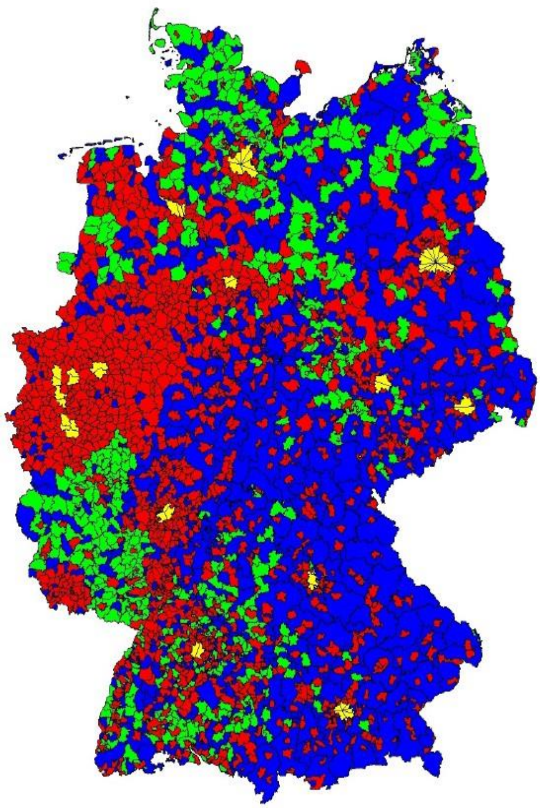

Figure 1. Map with sampling points in Germany (types 0 to 3 in yellow, red, green, and blue)

3. Design optimisation and small area estimation

The main objective, stated before as objective 1, was to accurately determine the population counts. Nevertheless, it was necessary to keep regional and substantive points resulting from objective 2 in mind. The task, after all, was to derive an appropriate sampling design and to allocate the total sample size in an appropriate way to the aforementioned sampling points to fulfil certain quality specifications laid down by law. In order to appropriately account for quality margins in terms of relative variances or related components, design-based (or model-assisted) methods should be considered. However, with respect to local area analysis it was important to ensure that model-based, and particularly small area estimation methods, could be employed and not be adversely affected by a sampling design which is too elaborate.

3.1. Design optimisation

The criteria for the evaluation of the possible designs resulted from different sources. The requirements imposed by the German government were formalised in the census law. The target of the Federal States was to get reliable regional estimates and tables. The interest of academia was empirical research using Census microdata. There was a strong interest in model building based on results of the Census in economic and social sciences.

The selected survey design should ensure that different precision requirements were met for the different hierarchical entities. Adequate estimators had to be chosen to meet the requirements of the sampling design. The first step was to define the accuracy objectives adequately. This was necessary to decide on the allocation and optimisation issues of the sampling design.

As a starting point, the relative root mean squared error (RRMSE) was chosen as a measure of accuracy. It allows for the comparison of design-based and model-based estimators in a design-based environment. This collapses to the coefficient of variation for design-unbiased estimators. Based on the RRMSE, the following accuracy requirements were formulated for the Census 2011. The first objective concerned only total estimates $\hat{t}_d$ of the size of the population U_d in communities d with more than 10,000 inhabitants:

$$\text{RRMSE}(\hat{t}_d) \leq 0,5\% \tag{1}$$

The same accuracy requirement was valid for parts of large towns with more than 400,000 inhabitants, the sampling points of type 0. As regards the second objective, the accuracy requirements depended on the type of sampling point and the variable of interest. In order to appropriately define the quality margins for objective 2 variables, the proportion p of the occurrence of an outcome of the variable of interest Y is used. The following rule was applied for all types of variables, whereas the proportion p varied across sampling point types. The proportion p of the variable of interest Y is given by

$$\frac{t_{dY}}{t_{dZ}} \approx p \quad \text{with} \quad p \geq \frac{1}{15} \tag{2}$$

where t_{dY} is the sum of inhabitants with property Y in area d and t_{dZ} is the total number of inhabitants in area d. Small proportions with $p < \frac{1}{15}$ were not considered under these settings. The accuracy requirement on the variable of interest is:

$$\text{RRMSE}(\hat{t}_{dY}) \leq \frac{1}{p} \tag{3}$$

The relevant proportions p of the variables of interest are depicted in Table 1.

Table 1. Maximum RRMSE dependent on p

Objective	1	2	2	2	2	2	2
p (in %):	100	80	50	30	20	10	6.7
Maximum RRMSE (in %):	0.5	1.25	2	3.33	5	10	15

As already stated, an optimised sampling design had to satisfy all administrative criteria as well as the accuracy margins given above. However, some additional issues became important. Before 2011, the annual Microcensus sample of 1% was often used as a gold standard. Any census estimate of households and persons should not, therefore, be based on a smaller sampling fraction than that of the Microcensus. Further, it was necessary to ensure that mean squared error estimation should be possible in a closed form, at least for objective 1 estimates. Additionally, the design had to be robust against the above settings. Furthermore, considerable dissimilarities in the treatment of different groups of persons had to be reduced as much as possible. And finally, statistical modelling, like sociometric, econometric and, of course, small area models, should be supported.

In the context of model building, which is of particular interest for economic and social sciences, Gelman (2007) illustrated difficulties of survey weighting for regression modelling and argued that *survey weighting is a mess*. The *Gelman bound* (GB), which is defined as the ratio of the largest to the smallest design weight, is aimed not to exceed 10 and is unacceptable beyond 100. The reason for this is that Bayesian model building may become complicated in the presence of highly varying survey weights.

On the basis of the exigencies defined above, a stratified sampling design was suggested. Information on variances and the numbers of persons within addresses (objective 1) within the strata were available from the population register. Note that a comparison of the accuracy of different sampling designs is presented in Section 4.1, which yielded the recommendation to apply a stratified design.

Maximal sampling fractions had to be chosen because sample sizes within strata should not exceed the population sizes. Minimal sampling fractions should guarantee reliable estimates in all relevant areas. The approach published by Gabler et al. (2012) takes into account all of the above criteria. An optimal allocation in the Neyman-Tschuprov sense was developed, which satisfied the upper and lower bounds of the sample sizes within each stratum and, hence, is called box-constraint optimal allocation. This approach also allows the optimization of the sample sizes amongst all sampling points simultaneously using a 2-norm of the RRMSE for all areas of interest:

$$\| RRMSE(\hat{t}) \|_2 = \sqrt[2]{\sum_d \mathrm{RRMSE}(\hat{t}_d)^2}\,. \tag{4}$$

A comparison of different algorithms for the box-constraint optimal allocation can also be found in Münnich et al. (2012b) and, as an integer problem, in Friedrich et al. (2015).

In order to achieve a stratified sampling routine that enables a considerable variance reduction, all sampling points were stratified into eight address size classes. The eight classes in each sampling point were constructed to contain approximately the same number of persons. The box constraints yielded a maximal Gelman bound of 25.

3.2. Design-based and model-based small area estimation

Different estimators for the total of persons living in Germany have been examined within the research project. The most important ones are briefly presented here. An extensive discussion on these estimators is given in Rao (2003). Further details about the implementation in the German Census can be found in Münnich et al. (2012a). Münnich et al. (2009) discussed the application of binomial mixed-models and spatial small area models in the context of the census. For the implementation in other research projects see the working papers of the EURAREA project (see for example The EURAREA Consortium, 2004, or Guiblin et al., 2004) and the DACSEIS project (cf. Münnich et al., 2004).

The following estimators are considered in this paper:

- **Horvitz-Thompson estimator (HT)** The HT was considered as a benchmark. However, for objective 1, the loss of efficiency was very high since the population register was a very strong auxiliary variable.

- **Generalized regression estimator (GREG)** With regards to the GREG, the question arose of the level at which the parameter estimation for the regression coefficients should take place. Two major results appeared. First, a separation with regards to the address size class yielded very unstable results, since in some cases extremely homogeneous numbers of individuals live in an address class. Second, using indirect estimates, i.e. using the regression information on higher than SMP level did not show significant differences in the quality of the estimates. With regards to the importance of objective 1, the community separate regression estimator was preferred for SMP 0 and 1.

- **EBLUP** The classical Battese-Harter-Fuller unit-level estimator (Battese et al., 1988) was considered as the main small area estimator.

- **Weighted EBLUP (YOURAO)** An extension of the EBLUP using design weights was proposed by You and Rao (2002). This estimator also fulfils the necessary benchmarking conditions to aggregate the small area estimates to the design-based national estimate.

In all cases where auxiliary variables could be included, the necessary demographic variables from the population register were applied, i.e. number of persons, gender, and age classes. In the census test in 2001, the correlation between register counts and real counts was estimated to the level of 0.993 (cf. Münnich et al., 2012a, p. 70) and, thus, the register count is a very efficient auxiliary variable in terms of objective 1.

The Fay and Herriot (1979) basic area-level estimator was also considered. However, due to the very highly correlated population register information, this estimator was generally outperformed by the unit-level estimator and, hence, was omitted in this overview. As well as normal distribution-based models, estimators based on the binomial or Poisson distributions have been applied that account for the count structure. These estimators are built on the best prediction (BP) approach of Jiang and Lahiri (2001) with a setup similar to the one used in González-Manteiga et al. (2007) and Münnich et al. (2009). The estimation was done based on the *R*-package *lme4*. Details can be found in the given references.

For reasons of coherence, we focus in the next section on the main findings on the impact of sampling designs and some selected results in terms of objective 2. Some additional results of the project are as follows:

- For objective 1, the community-separate GREG estimator yielded convincing results which could not be outperformed by small area estimators. The main reason is that the SMPs of type 0 and 1 are not sufficiently small, so that model-based methods cannot show their advantage. Additionally, accuracy estimation is much easier when applying design-based methods.
- Objective 2 is much more complicated. Here, in many cases the YOURAO estimator was the best solution. However, it seems very important to think further about additional sources of auxiliary information in the future in order to further improve model-based estimates. The information in the population registers in many cases is not very efficient.
- Further research needs to be done when there is interest in deriving high-dimensional tables or the *one-number-census*. A generalized calibration routine is under development, which at least allows implementing hierarchical information on areas and domains with different penalties.

4. Estimation results

Within the census sampling and estimation research project, a large number of Monte Carlo simulations have been conducted using sampling from the register dataset. This dataset was synthetically enlarged by some objective 2 variables using other sources like the Microcensus so that the final dataset was close to reality. The procedure is described in Münnich et al. (2012a) and Kolb (2013).

4.1. Results for classic estimation

In an early stage of the project, different sampling approaches have been evaluated using a classical design-based simulation study. As main measure for the evaluation the RRMSE was applied using the known true value from the above mentioned universe. The results for three estimators under different sampling designs are shown in Figure 2, i.e. the HT, the GREG, the EBLUP.

Every grid in Figure 2 shows the results for one estimator for the federal state of Saarland. As auxiliary variable, the register counts were used. The sampling designs are described in Table 2. The abbreviation BV in Table 2 denotes balancing variables in the case of balanced sampling (cf. Tillé, 2011). These were the register variables address size class (ADC) or nationality (NAT).

The different sampling designs are presented in rows. For each sampling design three rows of ticks are presented, which denote the RRMSEs of 52 communities in Saarland. The upper tick in red covers one town, Saarbrücken. The green ticks denote the RRMSEs of the large communities above 10,000 inhabitants. And finally, the yellow ticks present the RRMSEs of the smaller communities. The blue line yields the kernel density estimates of the RRMSEs from all communities.

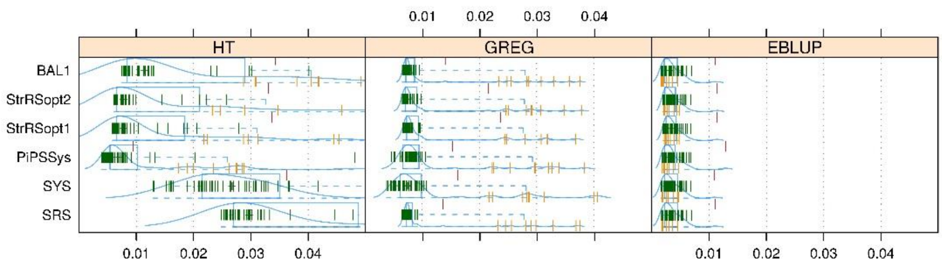

Figure 2. Comparisons of RRMSEs for various sampling designs for three total population estimates in Saarland

The names of the sampling designs in Table 2 refer to the classical designs within the SMPs. The allocation between the communities was drawn proportionally to the number of addresses.

Table 2. Different sampling designs - acronyms and their meanings

Acronym	description
BAL1	Balanced sampling; BV: ADC, NAT
StrRSopt2	Stratified random sample under optimal allocation (addresses)
StrRSopt1	Stratified random sample under optimal allocation (persons)
PiPSSys	π-PS systematic random sample
SYS	Systematic random sample
SRS	Simple random sample

As one can see in Figure 2, the EBLUP is fairly robust against the given sampling designs since the register counts allow for a strong model. However, one has to note that here no Neyman-based stratified design was applied, which would already negatively influence the results for the EBLUP. In terms of design-based methods, the design had considerable impact on the accuracy of the estimators. Simple random sampling, as expected, is very inefficient using the HT and much better for the GREG. Since the PiPSSys selects almost all large addresses, which seems hardly preferably in terms of a representative census sample, the accuracy of the results was considerably dependent on the address size structure of the communities. Balanced sampling did not yield very good results since the balancing variables seemed not to be so powerful and not many variables were available from the population register. The easier to implement stratified sampling designs seemed preferable with respect to accuracy, simplicity, and robustness against changes of settings as long as no specific allowance had to be made for smaller communities.

Finally, the results pointed to the use of stratified random sampling with some further optimization. The main gain in efficiency was stratification by address size. By the reasons given above, a box-constraint optimal allocation was introduced that guaranteed the necessary efficiency while still avoiding too much variation in the weights. Further, in any area, the census estimates were considerably more efficient than Microcensus estimates and no sub-population was drawn with a probability greater than 50%.

The results for the RRMSE under stratified sampling and box-constraint optimal allocation for different federal states are shown in Figure 3. The ordinate displays the different types of sampling points (from zero to three), while the abscissa indicates the RRMSE of the GREG estimator. Note that the results are theoretical results based on the address structure within the register and on the preassigned correlation of 0.993 between the register and true counts of people within addresses. Figure 3 shows that a-priori accuracy goals given in the census law were met in all SMPs of type 0 and 1, except for one community which failed slightly.

One has to note that even if the accuracy within SMPs 2 and 3 seems much lower, most SMPs of type 2 are still under 1% RRMSE which would have been the theoretical quality threshold using the objective 2 definition. Münnich et al. (2012a) showed that aggregating several estimates yields a RRMSE which is at least as good as the worst of the separate areas, which guarantees a hierarchical improvement by aggregation. Within the simulations, it turned out that this improvement, in general, is considerable.

Since the estimates of the total population (objective 1) were expected to be used for fiscal equalisation schemes between federal states and communities, special attention had to be paid to the estimator. Different possibilities were available for the estimation of the β -parameter of the GREG. Finally, the β -coefficient was estimated separately by SMP-type level, which in terms of the census law was separately by large community-level (above 10,000 inhabitants).

This was important to avoid a consideration of quality effects using indirect estimates employing information from other areas.

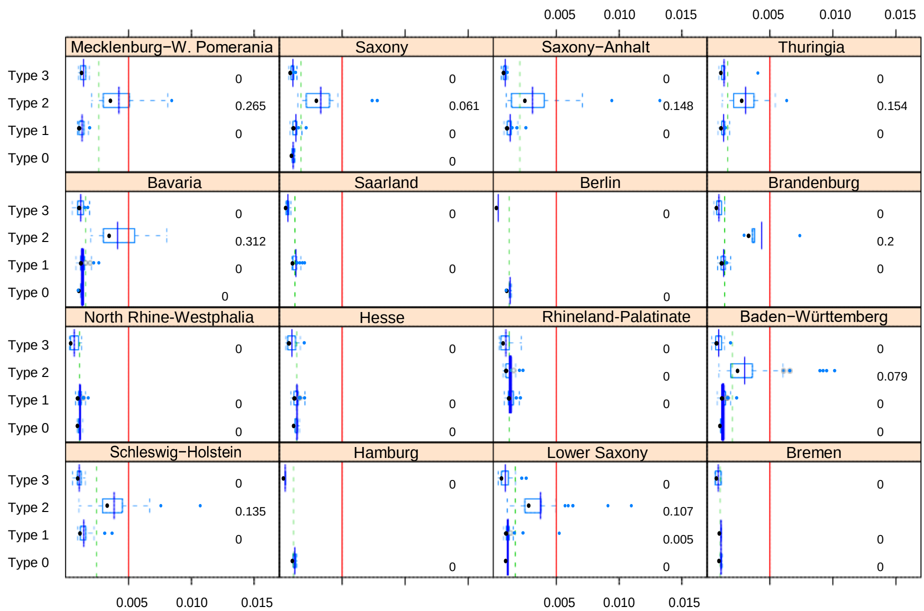

Figure 3. RRMSE of GREG estimates of population totals by federal state and sampling point type

4.2. Further ideas on small area applications

One of the objective 2 variables in the research project was the number of persons with foreign nationality (here the Turkish population) who had moved to Germany within a certain time span. Different estimators were applied to this estimation problem and the resulting RRMSEs are depicted in Figure 4 (Münnich et al. 2012a, pp. 104).

The descriptions of the headings in Figure 4 are provided in Table 3. In this Figure it is clear that YOURAO and EBLUP achieved the best results. However, for type 3 SMPs, the YOURAO estimator still yielded slightly better results, especially for earlier years of interest. Amazingly, the small area estimators also performed well in most cases of larger areas of type 0 and 1.

As a very important task in small area modeling, we have to consider vertical coherence, i.e. the aggregated small area estimates shall sum up to the national level estimates. It is well known that the GREG and YOURAO estimators fulfil this benchmarking condition. However, as a slightly more detailed assumption, we consider coherence to the next level, which in the German Census should also include coherence at the district level. This is, therefore, measured as the difference between the sum of the lower level total estimates and the higher total estimate.

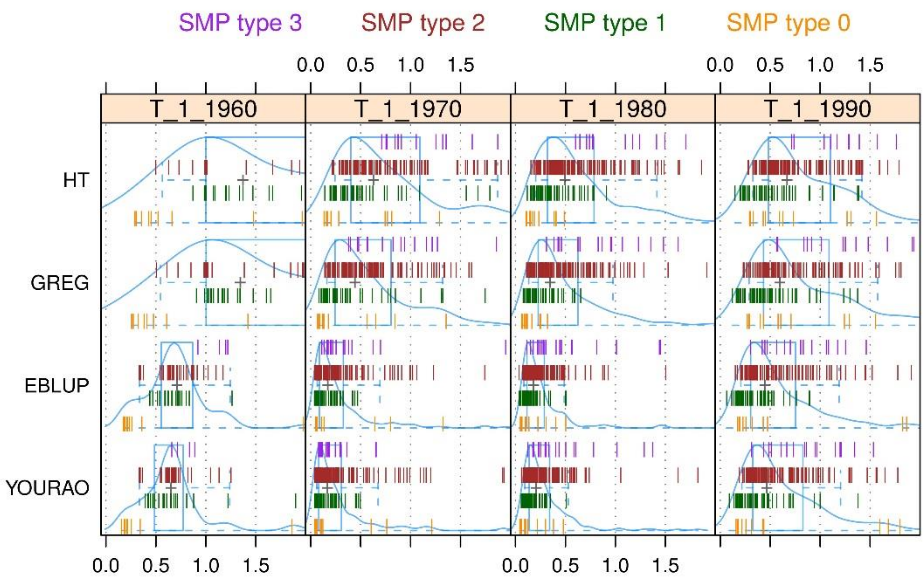

Figure 4. RRMSE for the estimation of selected years of moving in

Figure 5 indicates that the GREG is always coherent as long as the regression coefficient was estimated on the upper level which was the case in this example. The EBLUP suffers considerably from a lack in coherence. The YOURAO shows little deviations which may in fact be a result of a lack of the model in some districts. Further, the benchmarking condition holds only for the level on the β estimates, which here was the federal state level, which is higher than the district level. However, the deviation from perfect coherence is already small and much better than in the case of the EBLUP.

Table 3. Description of headings in Figure 4

Heading	description
	Number of persons, with Turkish nationality which
T_1960	moved to Germany between 1950 and 1960
T_1970	moved to Germany between 1960 and 1970
T_1980	moved to Germany between 1970 and 1980
T_1990	moved to Germany between 1980 and 1990

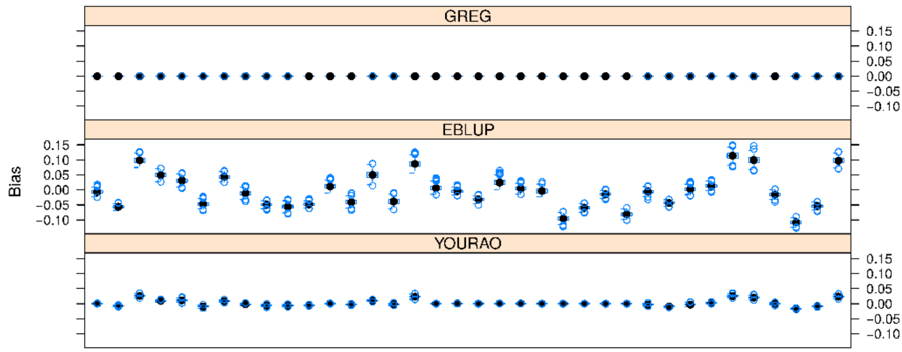

Figure 5. Coherence of aggregated SMPs estimates against district level estimates

To further force coherence of small area methods using different types of estimator the use of an extended calibration functional with penalties for further constraints based on regional small area estimates could be applied. By legal reasons, the first objective was exactly met on SMP level (0 and 1). Selected objective 2 estimates were met with high precision on the district level, whereas a lower preassigned precision was reached by other objective 2 estimates on SMP level. The variation of weights was constrained within the Gelman-bounds. It was possible to control the penalties separately on different levels and outcomes of covariates. Small area estimates for the totals can be used as benchmarks for objective 2 estimation. It is possible to apply different small area methods like, for example, the Battese-Harter-Fuller (Battese et al., 1988), the You-Rao (You and Rao, 2002), the Fay-Herriot and other estimators. The Lagrange multipliers provide a means to understand possible strains on area, domain or outcome of variables. This generalized calibration routine can be drawn from Münnich et al. (2012c) or Wagner (2013).

For a deeper overview of results from the entire study, we refer to Münnich et al. (2012a). The results suggested that the design recommendation still left enough space for applying small area methods. However, if a wider set of auxiliary variables was available from registers, e.g. by using matching methods, we would expect still a considerable improvement in the small area estimators.

5. Conclusions

As an outcome of the census sampling and estimation research project on the first German register-assisted census a recommendation was made for adopting a hierarchical SMP structure and a box-constraint optimal allocation for the sample sizes of addresses. For the first objective the use of a SMP-separate GREG was suggested. Either GREG or YOURAO estimators seemed adequate for the second objective depending on the target variable. An important consideration was that objective 1 estimates were used to construct new population figures. Furthermore, the coherence of estimates was a very important target. In the case of a mix of methods on different hierarchies, an application of the generalized calibration method may be considered in the future.

It had to be ensured that the chosen methods were computationally tractable. Multinomial small area estimates may be promising to be applied in the future but currently suffer from the computational effort. To achieve further improvements of model-based estimators, the use of linking and matching of several registers should be further analysed, e.g. using specialised matching routines.

More information about the German Census 2011 can be found on the official website www.zensus2011.de or in Münnich et al. (2012a).

Acknowledgements

We would like to thank the German Federal Statistical Office and the Statistical Offices of the Federal States for the opportunity to carry out the underlying research within the Census sampling and estimation research project. Further, we thank Professor Li-Chun Zhang for providing the opportunity to present the work in the invited session on census small area applications during the SAE 2014 in Poznan. Finally, special thanks go to the editors and reviewers who helped to improve the readability of the text considerably.

REFERENCES

BATTESE, G. E., HARTER, R. M., FULLER, W. A., (1988). An error-components model for prediction of county crop areas using survey and satellite data. Journal of the American Statistical Association 83 (401), pp. 28–36.

BECHTOLD, S., (2013). The new register-based census of Germany − a multiple source mixed mode approach. In: Presentation on the 59th World Statistics Congress (WSC), Hong Kong, August 2013,
URL http://www.statistics.gov.hk/wsc/IPS027-P2-S.pdf.

FAY, R. E., HERRIOT, R. A., (1979). Estimates of income for small places: An application of James-Stein procedures to census data. Journal of the American Statistical Association, Vol. 74, No. 366, pp. 269–277.

FRIEDRICH, U., MÜNNICH, R., DE VRIES, S., WAGNER, M., (2015). Fast integer-valued algorithms for optimal allocations under constraints in stratified sampling. Resubmit- ted.

GABLER, S., GANNINGER, M., MÜNNICH, R., (2012). Optimal allocation of the sample size to strata under box constraints. Metrika 75(2), pp. 151–161.

GELMAN, A., (2007). Struggles with survey weighting and regression modeling. Statistical Science, pp. 153–164.

GONZÁLEZ-MANTEIGA, W., LOMBARDÍA, M. J., MOLINA, I., MORALES, D., SANTAMARÍA, L., (2007). Estimation of the mean squared error of predictors of small area linear parameters under a logistic mixed model. Comput Stat Data Anal 51:2720–2733, DOI 10. 1016/j.csda.2006.01.012.

GUIBLIN, P., LONGFORD, N., HIGGINS, N., (2004). Standard estimators for small areas: Sas programs and documentation. Tech. rep., EURAREA − IST-2000-26290.

JIANG, J., LAHIRI, P., (2001). Empirical best prediction for small area inference with binary data. Annals of the Institute of Statistical Mathematics 53, pp. 217–243, dOI 10.1023/A:1012410420337.

KLEBER, B., MALDONADO, A., SCHEUREGGER, D., ZIPRIK, K., (2009). Aufbau des anschriften- und gebäuderegisters für den zensus 2011. Wirtschaft und Statistik 7, pp. 629–640.

KOLB, J. P., (2013). Methoden zur erzeugung synthetischer simulationsgesamtheiten. PhD thesis, Universität Trier.

MÜNNICH, R., MAGG, K., SOSTRA, K., SCHMIDT, K., WIEGERT, R., (2004). Workpackage 10: Variance estimation for small area estimates: Deliverables 10.1 and 10.2. URL http://www dacseis de-IST-2000-26057-DACSEIS Reports.

MÜNNICH, R., BURGARD, J. P., VOGT, M., (2009). Small area estimation for population counts in the German Census 2011. In: Proceedings of the Joint Statistical Meeting of the American Statistical Association. Washington.

MÜNNICH, R., BURGARD, P., VOGT, M., (2009). Small area estimation for population counts in the German Census 2011. In: Section on Survey Research Methods JSM 2009.

MÜNNICH, R., GABLER, S., GANNINGER, M., BURGARD, J. P., KOLB, J. P., (2012a). Stichprobenoptimierung und Schätzung im Zensus 2011. Statistisches Bundesamt.

MÜNNICH, R., SACHS, E. W., WAGNER, M., (2012b). Numerical solution of optimal allocation problems in stratified sampling under box constraints. AStA Advances in Statistical Analysis 96 (3), pp. 435–450.

MÜNNICH, R., WAGNER, M., SACHS, E. W., (2012c). Calibration benchmarking for small area estimates: An application to the German Census 2011. Symposium on the Anal- ysis of Survey Data and Small Area Estimation in Honour of the 75th Birthday of J. N. K Rao.

RAO, J. N. K., (2003). Small Area Estimation. Wiley series in survey methodology, John Wiley and Sons, New York.

THE EURAREA CONSORTIUM, (2004). Project reference volume vol. 2: Explanatory appendices. Tech. rep., EURAREA − IST-2000-26290.

TILLÉ, Y., (2011) Sampling algorithms. Springer.

WAGNER, M., (2013). Numerical optimization in survey statistics. PhD thesis, Univer- sität Trier, Universitätsring 15, 54296 Trier.

YOU, Y., RAO, J. N. K., (2002). A pseudo-empirical best linear unbiased prediction approach to small area estimation using survey weights. The Canadian Journal of Statistics / La Revue Canadienne de Statistique 30(3), pp. 431–439.