Sampling design for the German surveys on short-term statistics in the trade and service sectors

Keywords: multi source statistics, response burden, domain estimation.

1. Introduction

Currently, German short term statistics (STS) on the trade and service sectors are usually compiled as multi source statistics, where data on the smaller enterprises are obtained from administrative sources, while data on large enterprises are collected using a take-all survey. However, implementing the regulation on European business statistics regarding STS in this domain would lead to an inacceptable increase of the response burden under the current operating model. This is because large enterprises will be required to provide information on each kind-of-activity unit; the information has to be provided on a monthly instead of a quarterly basis for the STS in service sectors and the inclusion of new NACE sections. To prevent an unsustainable increase of the response burden, Destatis and the statistical offices of the Länder (states) decided to raise the thresholds above which primary survey data are collected in 2021, see Table 1.

Table 1. New thresholds for the primary survey data collection in different sectors

NACE classification	Threshold on turnover	Threshold on number of employees
Division 45	11 million Euro	250
Division 46	20 million Euro	250
Division 47	450 000 Euro	/
Sections H, J, L, M, N	15 million Euro	250
Section I	165 000 Euro	/

First analyses under the new thresholds indicate that while they help to reduce the increase of the response burden, their implementation alone is not sufficient to keep the burden at acceptable levels. Hence, the decision was taken to launch a complimentary project with the aim of replacing the take-all surveys for the larger enterprises by probability samples.

In this paper, we elaborate on different sampling design issues that we considered in our project. Specifically, we examine whether the sample allocation procedure should target estimates of levels or estimates of changes. Furthermore, we explore the choice of the final layer of stratification that helps to improve the precision of the estimates. For this purpose, we consider commonly employed size classes as well as a classification based on similarities of the structure of the underlying time series. Furthermore, we present results of simulation studies that describe the hypothetical performance of the different methods in the period between 2013 and 2017.

2. METHODS

Our surveys for STS indicators are designed to produce reliable estimates for many subgroups simultaneously, where the subgroups are indexed by $g, g = 1, \ldots, G$. Let Y_g denote the total of a variable in subgroup g, and $\hat{Y}_g$ its estimate based on the sample survey. In our application, the subgroups are so-called planned domains, within which the population is further stratified. In each stratum (indexed by $h, h = 1, \ldots, H$) we draw a simple random sample. Hence, choices on the sample allocation and stratification are crucially important for the quality of the estimates.

2.1. Sample allocation

Perhaps the most popular approach to sample allocation for domain estimation is the concept of power allocation due to [1], which minimizes the Euclidean norm of weighted coefficients of variation (CVs) of the domain estimates. Destatis has applied the procedure of [2] for a long time, where the allocation minimizes the maximum of weighted CVs of the domain estimates. Nowadays, Destatis employs the generalized power allocation developed by [3], which comprises the allocations due to Bankier and Schäffer as special cases. In this approach, the survey planner chooses the allocation minimizing the objective function

$$\sum_{g=1}^{G} \left(W_g^q \, CV(\hat{Y}_g) \right)^p \text{ with } W_g^q = Y_g^q C_g^{-1}. \tag{1}$$

under the constraints

$$\sum_{h=1}^{H} n_h = n \tag{2}$$

$$a_h \leq n_h \leq b_h \; \forall h. \tag{3}$$

Here, a_h and b_h refer to the lower and upper bound of the sample size allocated to stratum h respectively, whereas C_g denotes a further parameter at the domain level that can be chosen by the survey planner, where $C_g = 1$ is a popular choice. Hohnhold called his approach a generalized power allocation, because it extends Bankier's method to minimize p-norms of weighted CVs of domain estimates ($2 \leq p \leq \infty$). Note that while Bankier's method corresponds to $p = 2$, Schäffer's approach results as the limiting case $p = \infty$.

Most of the methods in the literature on sample allocation in stratified simple random sampling have been tailored for the estimation of levels. However, STS indicators aim to provide relevant information on economic trends and developments. Hence, in addition to levels, users of STS are frequently interested in changes, growth rates and alike. Therefore, we wanted to develop a sample allocation method tailored towards the estimation of growth rates. For this purpose, we modify the objective function of the generalized power allocation to accommodate minimizing weighted p-norms of standard errors and CVs of growth rates. Specifically, we will consider the allocation minimizing the objective function

$$\sum_{g=1}^{G} \left(W_g^{*,q} \sqrt{Var\left(\hat{Q}_g^{(t)}\right)} \right)^p \text{ with } W_g^{*,q} = Y_g^{(t-1),q} C_g^{-1}, \tag{4}$$

under the constraints (2) and (3) as the generalized power allocation for growth rates. In (4), $\hat{Q}_g^{(t)}$ denotes the estimator of $Q_g^{(t)} = Y_g^{(t)}/Y_g^{(t-1)}$, i.e. the gross growth rate between periods $t-1$ and t. We use the standard linearization approach to approximate the variance of $\hat{Q}_g^{(t)}$. If we choose $C_g = 1$, the solution to (4) minimizes the p-norm of weighted standard errors of the growth rates, while $C_g = Q_g^{(t)}$ leads to minimizing the p-norm of weighted coefficients of variation of the growth rates. Note that when allocating a sample for the purpose of the estimation of changes it is not intuitively clear whether we should prefer minimizing a norm composed of weighted standard errors or a norm composed of weighted CVs.

2.2. Stratification

A standard stratification in German business statistics is to stratify by states, business sector classification and size classes (say turnover, or number of employees). As the domains of interest usually emerge as a cross-classification of states and business sectors, we consider these two variables as given in our stratification scheme. An alternative to size classes is due to [4], who proposed to create strata within domains of interest based on similarities of the time series. Their approach comprises two steps. In the first step, input factors are produced. This may include components from a decomposition of the time series as well as other information such as the correlation between the number of trading days and the turnover of a unit. In the second step, the aim is to obtain homogeneous strata based on the multivariate structure of the input data using methods from cluster analysis.

3. RESULTS

We evaluate different options for sample allocation and stratification by means of simulation studies that cover the period from the beginning of 2013 to the end of 2017. For this purpose, we compare estimated STS indicator series using hypothetical probability survey samples against benchmark series, which are computed using the fully available (take-all) survey data. As parameters of interest, we consider yearly growth rates of monthly series based on overlapping units. For each method and survey, we ran 1000 Monte-Carlo replications. The sample sizes were restricted to include a maximum of 45 per cent of the legal units above the thresholds in Table 1. Regarding the parameters in the objective functions (1) and (4), we fixed $C_g = 1$ and $p = \infty$ in all studies. Sector specific choices have been taken for parameter, which ranged from $q = 0$ to $q = 0.3$.

A comparison of the two objective functions used in the sample allocation shows a mixed picture. Altogether, we noted small advantages for the allocation designed for growth rates (4) in comparison to the allocation designed for totals (1). Thus, we focus on allocation minimising objective function (4) under the constraints (2) and (3) in the following.

Our results indicate that both stratification schemes permit the production of reliable indicator series using probability surveys for the large enterprises. In fact, aggregate series are very similar (e.g. Figure 1). On a more granular level, we observed some differences between the different stratifications. However, our results do not indicate systematic advantages of any particular method.

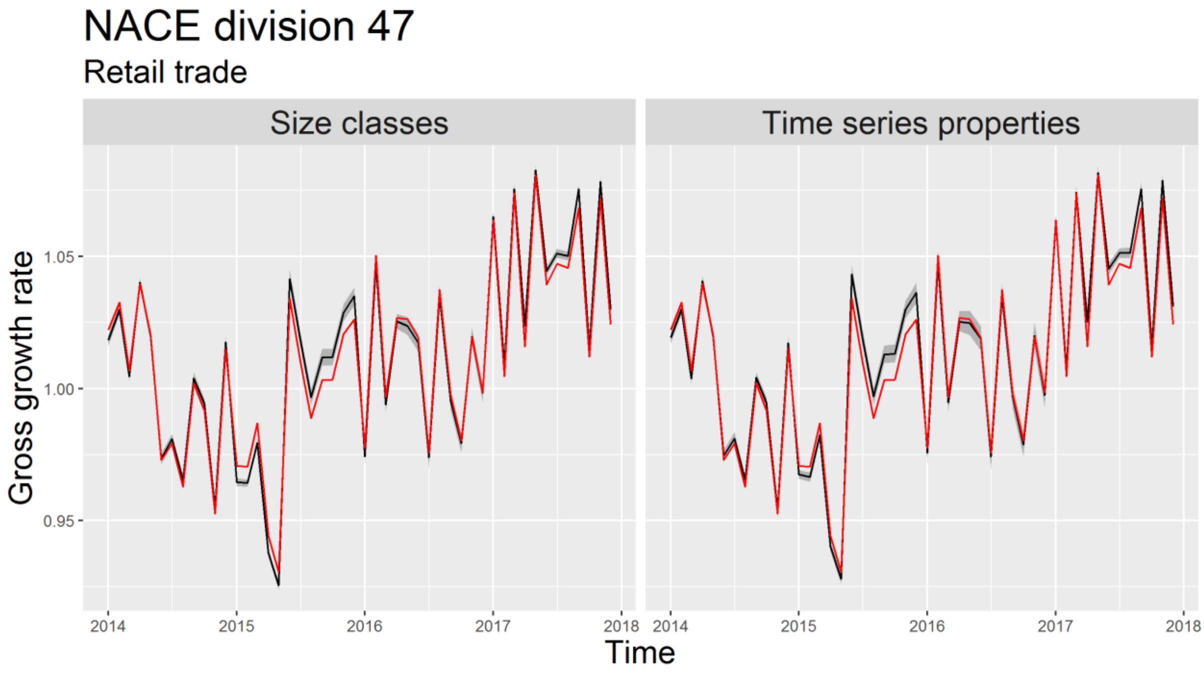

Figure 1. Comparison of estimated indicator series (black) and benchmark series (red) for a stratification with size classes (left panel) and a classification based on time series properties (right panel) in the retail trade sector. The shaded area in grey depicts the range in which 90 per cent of the estimates fall.

4. Conclusions

Our findings suggest that probability samples are able to replace take all surveys of large enterprises in the German surveys on STS indicators for the trade and service sectors. Combining sample survey data with administrative data for smaller enterprise yields a significant reduction of the response burden compared to the current operating model without major expected quality losses.

We note some advantages for a sample allocation procedure tailored towards the estimation of changes rather than levels. However, the differences compared to a standard allocation designed for the estimation of levels are of minor importance. We also found similar results under different stratification schemes. Both findings could be due to large sampling fractions in the group of large enterprises that are eligible for selection into the sample. Thus, one might expect larger effects of either component if a sample is designed to cover the whole population instead.

References

[1] Michael D. Bankier. Power Allocations: Determining Sample Sizes for Subnational Areas, The American Statistician (1988), 174-177.

[2] Karl-August Schäffer. Planung geschichteter Stichproben bei Vorgabe einer Fehlerabstufung, Allgemeines statistisches Archiv (1961), 350-361.

[3] Henning Hohnhold. Generalized Power Allocations, Unpublished Manuscript (2010).

[4] Katrin Mörixbauer and Elmar Wein. Extended usage of administrative data for German statistics -Final grant report - Grant Agreement Number: 822695 — 2018-DE-ESS-VIP-ADMIN.