**Do declining response rates negatively affect sample composition?
A longitudinal analysis using data from the German General Social Survey (ALLBUS)**

Michael Blohm, GESIS – Leibniz Institute for the Social Sciences, michael.blohm@gesis.org

Achim Koch, GESIS – Leibniz Institute for the Social Sciences, achim.koch@gesis.org

**Abstract**

*The German General Social Survey (ALLBUS) is a repeated cross-sectional multi–thematic survey conducted every other year since 1980. Like many other surveys ALLBUS has been facing the problem of declining response rates. Since 1994 ALLBUS response rates decreased steadily from 54% to 35% at present – despite exerting higher fieldwork efforts. Declining response rates give rise to the question whether this affects survey quality. Selective participation might threaten the representativeness of a survey in general or with regard to specific characteristics and questions. In our presentation we use ALLBUS data to analyze the relationship between the response rate and indicators of nonresponse bias over time. The ALLBUS provides a good basis to analyze this relationship, since it has used an almost identical study design since 1994. To investigate nonresponse bias we analyse a) survey participation in ALLBUS with sample frame information from the population registers and b) deviations of the net sample from external benchmark data (namely the German (micro-) census from official statistics). Our results indicate that decreasing response rates are not systematically associated with deterioration in demographic sample composition.*

**Keywords:** Response Rate, Nonresponse Bias, F2F Survey, ALLBUS

**1. Introduction**

Declining response rates are a continuing problem for household surveys in many Western countries (Atrostic et al. 2001; Beullens et al. 2018; de Leeuw and de Heer 2002; Dixon and Tucker 2010). The German General Social Survey (ALLBUS) – which has been conducted every other year since 1980 – also has been facing an increase in nonresponse in the past decades. Between 1994 and 2016 the response rate of ALLBUS decreased from 54% to 35%. The main reason for this decline was a rise in the number of refusals. The decrease in the response rate raised concerns whether it would be accompanied by an increase in nonresponse bias. In the following we investigate whether the decline in ALLBUS response rates comes along with an increase in nonresponse bias over time.

**2. Data and Methods**

ALLBUS is a face-to-face survey of the adult population in Germany, covering a wide range of topics and aiming at charting the long-term trends in attitudes and behaviour (http://www.gesis.org/en/allbus). ALLBUS provides a good opportunity to analyze the relationship between declining response rates and nonresponse bias, as ALLBUS has used the same basic survey design since 1994 (target population covering adults living in private households; samples of named individuals from a register; 3.500 completed interviews per round; interview duration around 70 minutes; consistent calculation of response rates). In order to analyse nonresponse bias we rely on two different sets of indicators. On the one hand, we use micro level data for respondents and nonrespondents from the ALLBUS sampling frame to calculate R -indicators (Schouten et al. 2011). On the other hand, we use aggregate level data from official statistics and analyse the deviations of the ALLBUS net sample from these external benchmark data.

2.1 R-indicators

R-indicators use frame information available both for respondents and nonrespondents of a survey to calculate response propensities. R-indicators reflect the variability of the estimated response propensities and are defined by
$R=1-2S(p)$. Conditional partial R-indicators allow to single out the effect of individual variables, controlling for the influence of the other variables included in the model. The frame information available for the ALLBUS gross sample stem from the community lists of residents. Five variables, measured in the same way over time, are available to estimate response propensities: Gender (male vs. female), age (6 age groups), nationality (German vs. Non-German), region (Western vs. Eastern Germany) and urbanicity (7 categories). Using these variables as independent variables, we estimate logistic regression models with response (y/n) as dependent variable.

2.2 Deviation from external benchmark data

In our second approach, we compare the distributions of several socio-demographic variables with the respective (aggregate) data from official statistics. This is a straightforward and frequently used method for analyzing nonresponse bias (see Hartmann 1990, Hartmann & Schimpl-Neimanns 1992, Koch 1998). Data from the German microcensus survey can be used as a valid external benchmark to evaluate the net samples from ALLBUS. The microcensus is a yearly survey of a 1% sample of the population in Germany. Participation in the microcensus is mandatory and its unit nonresponse rate is only about three percent. Comparisons of microcensus data with data from other official sources (like Census data, and data on educational achievement) confirm the good quality of the microcensus data (Hartmann, 1990). The advantage of this method is, compared to relying on frame information, a slightly wider range of variables can be analysed.

Seven variables are measured in a comparable way in ALLBUS and the microcensus. These are gender (male vs. female), age (10-year age categories), education (low – medium – high), marital status (married: Y vs. N), household size (1/2/3/4/5+ persons), work status (in paid work 17 hours per week or more: Y vs. N) and occupational status (5 categories). As a summary measure for the consistency of ALLBUS and microcensus variable distributions we rely on the index of dissimilarity (Duncan & Duncan, 1955). We calculate this index both a) for each variable separately and b) as a mean index across all seven variables:



a) Index of dissimilarity for variable j with i categories:

b) Mean index of dissimilarity for v variables with i categories:

**3. Results**

3.1 R-indicators and frame information

Using the available information from the sampling frames, R-indicators and conditional partial R-indicators were estimated for eleven ALLBUS surveys conducted between 1994 and 2016 (Figure 1). For the analyses we used the scripts provided by the RISQ-Project. Generally, R-indicators can vary between 0 and 1. A value of 1 indicates that response propensities do not vary. The lower the value of the R-indicator, the larger is the variation in response propensity (and as a consequence, also the degree of nonresponse bias).

Figure 1:

Source: ALLBUS 1994 -2016, own calculations

The R-indicators for the ALLBUS surveys are quite high; they vary between .82 and .91. This indicates a high level of representativity of the ALLBUS samples. Over time R-indicators tend to become larger, what indicates an increase in sample quality over time. Given the fact, that ALLBUS response rates decrease over the years, this result contradicts the usual expectation.

Figure 2 provides further details of this trend and depicts the results for the individual variables, i.e. the conditional partial R-indicators. For the sake of simplicity, only the linear trends for the five variables are shown in the figure.

Figure 2:

Source: ALLBUS 1994 -2016, own calculations

Figure 2 shows that the relevance of four variables for the variability of response propensities has declined. The variables ‘urbanicity’ and ‘age’ have the largest effect on response propensities. Although their effect has been decreasing over time, their effect is still largest in 2016. In addition, the effect of ‘region’ and ‘gender’ on response propensities has diminished. Solely the effect of ‘nationality’ became larger over time.

3.2 Deviations from the German microcensus (external benchmark)

Figure 3 shows the mean index of dissimilarity (D) between ALLBUS and microcencus variable distributions across the seven characteristics included in our analysis. The index of dissimilarity has a rather streightforward interpretation: it measures the percentage of respondents that would need to move between the categories of a variable to produce exactly the same distribution for the two surveys. The larger the discrepancy between the ALLBUS and microcensus variable distribution, the larger the value of D. The average values of D for the ALLBUS surveys between 1994 and 2016 are on a rather low level. (For surveys using a different survey design like random route, much higher values can be observed, see Blohm 2006). Over time, no clear-cut trend in the average value of D can be observed. Thus, with this approach too, we do not find that a decrease in the response rate over time comes along with a decrease in quality of the demographic sample composition.

Figure 3:

Source: ALLBUS 1994 -2016, german microcensus 1993 – 2015, own calculations

Figure 4 shows the indices of dissimilarity separately for the seven sociodemographic variables. For the sake of simplicity, only the linear trend for each variable is presented. The trend differs between the variables. The indices of D for ‘occupational status’ and ‘gender’ became smaller over time. The values for ‘work status’ and ‘marital status’ were quite stable over the years. The discrepancies between the ALLBUS and the microcensus of the remaining variables ‘age’, ‘education’ and ‘household size’, however, increased over time. As regards the variable ‘age’, this result seems to contradict the results in section 3.1 using the R-indicators. There we found a declining effect for the variable ‘age’. However, one has to take into account that in 3.1 the results refer to the effects of the variables in a multivariate model, whereas in section 3.2 we only analyse the univariate distributions of the variables. The largest increase pertains to the variable ‘education’. A closer look reveals, that in particular the underrepresentation of persons with a low school-leaving certificate became stronger over the years. That the variable ‘education’ is prone to be biased in sample surveys in Germany has already been discussed in the past (Hartmann & Schimp-Neimanns 1992).

Figure 4:

Source: ALLBUS 1994 -2016, german microcensus 1993 – 2015, own calculations

**4. Discussion**

The present paper investigated whether the decline in response in the German ALLBUS survey which could be observed in the past 20 years comes along with an increase in bias in terms of sociodemographic sample composition. Using both information available from the ALLBUS sampling frame and external benchmark data, we could show that this is – in general - not the case. Despite facing a decrease in the response rate of twenty percentage points, the current ALLBUS surveys exhibit – on average – either a better balanced sample composition (according to the sample frame information and the respective R-indicator) or an equally well balanced sample (according to the external benchmark data and the average value of D) than the ALLBUS surveys in the second half of the 1990s. Thus, these results are another example that response rates alone are an imperfect indicator of nonresponse bias. Various factors might have contributed to this result. First and foremost, one has to take into account that ALLBUS has implemented various measures to improve fieldwork and sample composition in the past 20 years. This includes more general measures like an increase in the number of contact attempts, and more specific procedures like the tracing of target persons who moved shortly before or during the ALLBUS fieldwork period.

The analyses also show, that response rates are an imperfect indicator of nonresponse bias, since nonresponse bias is a variable-specific phenomenon. Accordingly, we found in our analyses both examples of variables, the bias of which became smaller over time (despite a decline in the response rate) and examples of variables, the bias of which increased (as one usually expects, when the response rate decreases).

The ALLBUS surveys between 1994 and 2016 used a sample of named individuals from a register; interviewers were not involved in the selection of the sample. Comparison with surveys using other types of sample selection show, that this is an important precondition for achieving high quality of the demographic sample composition (Koch 1998; Blohm 2006). One has to be aware, that a similar analysis of surveys using a different, less strict and transparent sampling design might lead to different results. Furthermore, the broad range of topics included in the ALLBUS questionnaire has to be taken into account when evaluating the present results. This broad range of topics probably prevents that ALLBUS is relevant only for persons with specific characteristics and interests.

**5. References**

Atrostic, B. K., Bates, N., Burt, G., & Silberstein, A. (2001), Nonresponse in U.S. Government Household Surveys: Consistent Measures, Recent Trends, and New Insights. *Journal ofOfficial Statistics*, 17(2), pp. 209-226.

Beullens, K., Loosveldt G., Vandenplas C. & Stoop I. (2018), Response Rates in the European Social Survey: Increasing, Decreasing, or a Matter of Fieldwork Efforts? Survey Methods: Insights from the Field. Retrieved from <https://surveyinsights.org/?p=9673>

Blohm, Michael (2006), Datenqualität durch Stichprobenverfahren bei der Allgemeinen Bevölkerungsumfrage der Sozialwissenschaften - ALLBUS, in: Frank Faulbaum und Christof Wolf (Ed.), Stichprobenqualität in Bevölkerungsumfragen, Bonn: Informationszentrum Sozialwissenschaften: pp37 - 54.

de Leeuw, E., & de Heer, W. (2002). Trends in Household Survey Nonresponse. A Longitudinal and International Comparison. In R. M. Groves, D. A. Dillman, J. L. Eltinge, & R. J. A. Little (Eds.), *Survey Nonresponse* (pp. 41-54). New York: John Wiley & Sons.

Dixon, J., & Tucker, C. (2010). Survey Nonresponse. In P. V. Marsden & J. D. Wright (Eds.), *Handbook of Survey Research.* Second Edition (pp. 593-630). Bingley: Emerald.

Duncan, O. D. & Duncan, B. (1955). A Methodological Analysis of Segregation Indexes. American Sociological Review, 20, 210-217

Hartmann, Peter, (1990), „Wie repräsentativ sind Bevölkerungsumfragen? Ein Vergleich das ALLBUS und des Mikrozensus“. *ZUMA-Nachrichten*, 26, pp. 7- 30.

Hartmann, Peter H.; Schimpl-Neimanns, Bernhard (1992), Sind Sozialstruktur-analysen mit Umfragedaten möglich? Analysen zur Repräsentativität einer Sozialforschungsumfrage. Kölner Zeitschrift für Soziologie und Sozialpsychologie, 44, pp.315-340

Koch, Achim, (1998). „Wenn "Mehr" nicht gleichbedeutend mit "Besser" ist: Ausschöpfungsquoten und Stichprobenverzerrungen in Allgemeinen Bevölkerungsumfragen“. *ZUMA-Nachrichten*, 42, pp. 66-90.

[Schouten, B., Shlomo, N. & Skinner, C. (2011), Indicators for Monitoring and Improving Representativeness of Response. Journal of Official Statistics 27, pp. 1-24](http://hummedia.manchester.ac.uk/institutes/cmist/risq/schouten-shlomo-skinner-2011.pdf)