Catalogue no. 12-001-X
ISSN 1492-0921

Survey Methodology

Who's asking? Interviewer effects on unit nonresponse in the Household Finance and Consumption Survey

by Nicolas Albacete, Pirmin Fessler and Peter Lindner

Release date: December 23, 2025

Statistics Canada Statistique Canada

Canada

How to obtain more information

For information about this product or the wide range of services and data available from Statistics Canada, visit our website, www.statcan.gc.ca.

You can also contact us by

Email at infostats@statcan.gc.ca

Telephone, from Monday to Friday, 8:30 a.m. to 4:30 p.m., at the following numbers:

- Statistical Information Service 1-800-263-1136
- National telecommunications device for the hearing impaired 1-800-363-7629
- Fax line 1-514-283-9350

Standards of service to the public

Statistics Canada is committed to serving its clients in a prompt, reliable and courteous manner. To this end, the Agency has developed standards of service which its employees observe in serving its clients. To obtain a copy of these service standards, please contact Statistics Canada toll-free at 1-800-263-1136. The service standards are also published on www.statcan.gc.ca under "Contact us" > "Standards of service to the public."

Note of appreciation

Canada owes the success of its statistical system to a long-standing partnership between Statistics Canada, the citizens of Canada, its businesses, governments and other institutions. Accurate and timely statistical information could not be produced without their continued co-operation and goodwill.

Survey Methodology, December 2025
Vol. 51, No. 2, pp. 629-646
Statistics Canada, Catalogue No. 12-001-X

Who's asking? Interviewer effects on unit nonresponse in the Household Finance and Consumption Survey

Nicolas Albacete, Pirmin Fessler and Peter Lindner[1]

Abstract

This study examines interviewer effects on household nonresponse in three waves of the Household Finance and Consumption Survey (HFCS) in Austria using a multilevel model. Addressing nonresponse at its source is crucial for maintaining survey data quality and representativeness. Our findings indicate that the variation in response behavior explained by interviewer effects decreased from about one-third in the first wave to 7% in the third wave. Effective interviewers tend to have a university degree, be married, homeowners, and have a larger workload. Additionally, higher mean wages in the household's municipality negatively affect survey participation. These insights suggest targeted interviewer selection and training strategies to improve response rates.

Key Words: HFCS; Interviewer effects; Interviewer survey; Unit nonresponse.

1. Introduction

Survey data are crucial for economics and social sciences, often collected by statisticians and analyzed by economists who may not be familiar with the data collection process. Unit nonresponse is a significant and growing issue in survey collection. This study examines interviewer effects on unit nonresponse using data from the Household Finance and Consumption Survey (HFCS), a complex face-to-face interview survey.

Kreuter (2008) identifies four ways interviewers influence responses: presence, observable characteristics, verbal and nonverbal behavior, and potential errors. Related studies on item nonresponse include Blom and Korbmacher (2013).

Understanding how interviewers and sample units interact is important not only for interviewer selection and training, for the matching of interviewers with sample units and for the monitoring and rewarding of interviewers (Groves and Couper, 1998; De Leeuw, Hox and Dillman, 2008; Kennickell, 2006b; Kennickell, 2006a; Kennickell, 2008 and Kreuter, 2008), but also for the statistical analysis of survey results (e.g. in the construction of survey weights in the Austrian HFCS, see Albacete, Dippenaar, Lindner and Wagner, 2018). Despite its importance, little research has been conducted on this crucial part of data production.

Studies such as Durrant, Groves, Staetsky and Steele (2010) and Hox and de Leeuw (2002) focus on initial interviewer-sample unit contact, identifying successful behaviors and strategies. Beerten (1999) and Jäckle, Lynn, Sinibaldi and Tipping (2013) analyze the relationship between observable interviewer characteristics and survey response, finding that interviewers with fewer refusals also have fewer noncontacts (Pickery and Loosveldt, 2002; O'Muircheartaigh and Campanelli, 1999).

1. Nicolas Albacete, Pirmin Fessler and Peter Lindner, Oesterreichische Nationalbank, Research Section. E-mails: nicolas.albacete@oenb.at; pirmin.fessler@oenb.at; peter.lindner@oenb.at.

Response rates are higher among female (O'Muircheartaigh and Campanelli, 1999; Hox and de Leeuw, 2002) and older interviewers (Kennickell, 1999; O'Muircheartaigh and Campanelli, 1999; Hox and de Leeuw, 2002; Merkle and Edelman, 2002; Singer, Glassman and Frankel, 1983). Effects of other characteristics like voice (Schaeffer, Dykema and Maynard, 2010) and race (Merkle and Edelman, 2002; Singer et al., 1983) are inconclusive or insignificant. Experience, knowledge, and positive attitudes towards persuasion are positively related to response rates (Schaeffer et al., 2010), while personality measures show no strong effects (Groves and Couper, 1998).

Interviewer-respondent interaction, especially between introduction and participation decision, is crucial. Improvisation increases response rates (Houtkoop-Steenstra and van den Bergh, 2000; Morton-Williams, 1993). Techniques like "tailoring" and "maintaining interaction" reduce refusal rates (Cialdini, Couper and Groves, 1992; Schaeffer et al., 2010; Kennickell, 1999).

For extensive literature reviews, see Schaeffer et al. (2010) and Jäckle et al. (2013). Our study builds upon findings by West and Blom (2017) in the Journal of Survey Statistics and Methodology, extending their analysis of interviewer effects on unit nonresponse.

Our study contributes to the existing literature by modeling nonresponse in the HFCS using three data sources. By analyzing the effects of interviewer characteristics and the social environment, we highlight the significant role of interviewer selection and training. Our findings show that interviewers with higher education, marital status, homeownership, and larger workload are more effective (for wealth surveys). Consequently, we recommend targeted assignment of interviewers with these characteristics to the field and implementing tailored training programs to enhance their effectiveness. Additionally, we suggest ongoing monitoring and support to maintain high response rates.

We use data from a large-scale household survey on a sensitive topic, namely wealth, with selective nonresponse analyzed by including a large set of paradata and auxiliary information available for all sampled units and not only for respondents. The second source is administrative regional data on income. The third is a detailed interviewer survey including interviewer characteristics and personality traits. The combination of these three datasets allows us to use multilevel modelling to identify the amount of variation in response behaviour explained at the interviewer level. A likelihood ratio test of the variance components yields a rejection of the null hypothesis based on a mixture of χ^2 distributions at conventional significance levels and thus makes the multilevel approach warranted. It also allows us to analyse the effect of interviewer characteristics and personality traits on response behaviour while controlling for other important determinants which are neither interviewer nor sample unit characteristics but paradata and auxiliary information which proxy the social environment of the sample unit (found to be a crucial control in Groves and Couper, 1998 and Beerten, 1999), both for participating and non-participating sample units.

The paper is structured as follows: in Section 2, we provide an overview of our study design, Section 3 presents the results, and Section 4 concludes. For additional tables and figures, we refer to the online Appendix available at:

https://www.hfcs.at/dam/jcr:52e5e12d-8dfe-4083-983f-f7056bf07c48/AFL_OnlineAppendix.pdf.

2. Study design

Section 2.1 outlines the theoretical framework and hypotheses, Section 2.2 details the survey and interviewer data, Section 2.3 describes our empirical estimation strategy, and Section 2.4 presents the model specifications used in our analysis.

2.1 Theoretical considerations

The goal of our work is to understand the decision to participate in a survey and its interaction with various factors to improve future surveys. Three factors are relevant: the interviewer; the social environment in which the interview with the respondent is to take place; and the sample unit characteristics, which may themselves be shaped by the social environment. An interviewer cannot select and thus cannot influence the characteristics of the sample unit. However, the social environment might influence both the interaction between interviewer and sample unit, on the one hand, and the sample unit itself on the other.

At the social environment level, we consider various social factors at the dwelling, area, and region levels that might influence participation. At the interviewer level, we use characteristics, personality traits, and experience, which can significantly impact the interaction with the sample unit. Interviewer selection and training also play a crucial role (Groves and Couper, 1998). From psychology, we include the "big five" personality traits (openness to experience, conscientiousness, extraversion, agreeableness, and neuroticism) as they might influence participation (McCrae and John, 1992). At the sample unit level, we observe participation decisions and consider social and personal characteristics that influence survey readiness, combining refusals and non-contacts as non-participation.

Our main goal is to quantify the influence of the interviewer on the participation decision, controlling for the social environment. To this end, it is important to observe both respondents and non-respondents. Our secondary goal is to better understand which interviewer characteristics might play a role in the process and, more specifically, whether experience and personality traits measured by the big five might have an influence.

This understanding can help improve survey participation and enhance interviewer training and selection, matching the "right" interviewer with a respondent to maximize interaction quality. One could potentially consider characteristics such as the interviewer's experience, personality traits, or demographic factors like age, gender, and cultural background. While current collection systems are not equipped to

automatically facilitate such matching, it is something that could be integrated into future survey designs. Practical limitations such as increased costs or recruitment difficulties may dampen the potential gains, but the knowledge remains valuable. Even with these challenges, we can derive practical recommendations to enhance survey methodologies. By understanding the key characteristics and training needs of effective interviewers, we can implement targeted recruitment and training programs. These strategies could help improve the quality of interactions between interviewers and respondents, thereby increasing response rates and the overall quality of survey data.

2.2 Data and estimation strategy

This study utilizes the largest wealth survey in Europe, complemented by administrative data.

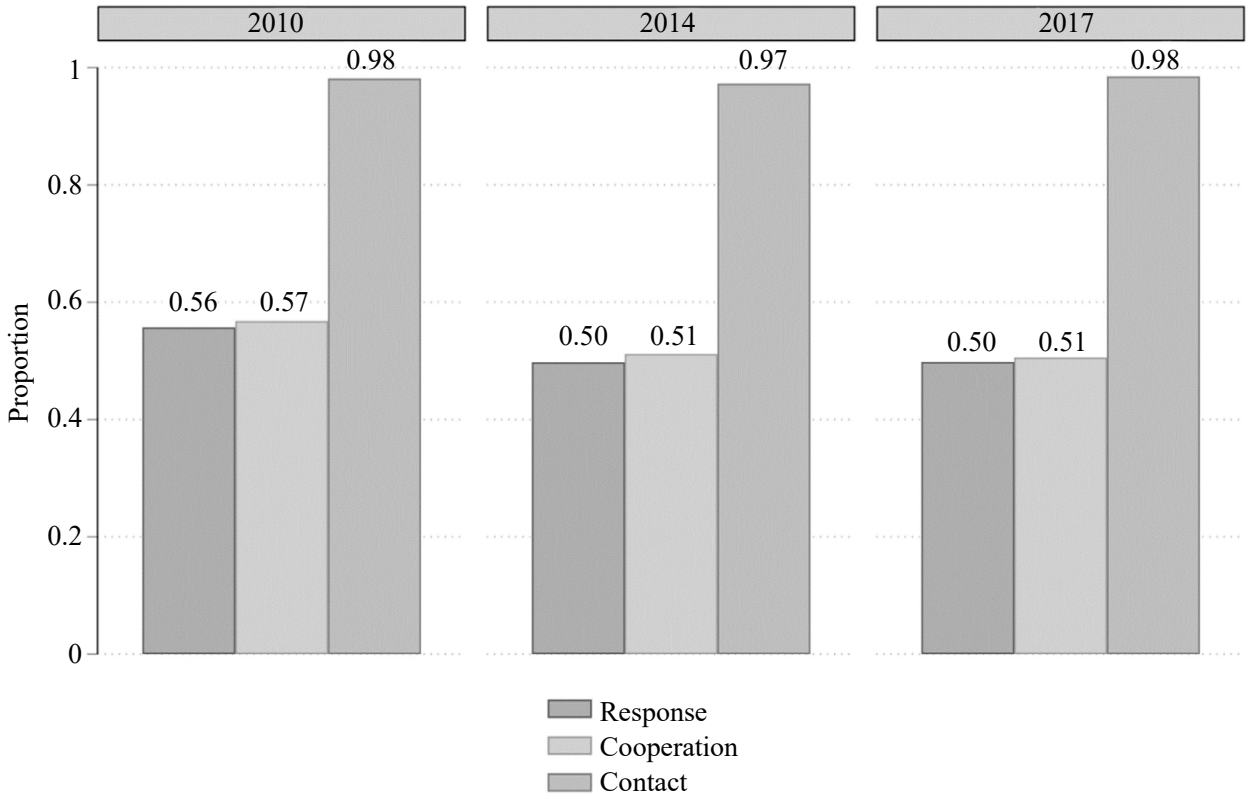
2.2.1 Household Finance and Consumption Survey

We use three waves – repeated cross-sections – of the HFCS in Austria conducted in 2010, 2014, and 2017 (Albacete, Lindner, Wagner and Zottel, 2012; Albacete, Lindner and Wagner, 2016 and Albacete et al., 2018, as well as Fessler, Mooslechner and Schürz, 2012; Fessler, Lindner and Schürz, 2016 and Fessler, Lindner and Schürz, 2018). The Austrian part of the HFCS uses stratified two-stage cluster sampling. Sample stratification is based on geographical information (NUTS-3 regions: Federal States and political districts) and municipality size classes yielding 170-185 (depending on the wave) very fine-grained strata (in the median only two clusters are randomly drawn per stratum). The two-stage sampling design entails, first, selecting a random sample of enumeration districts (the smallest geographical unit for which statistical data are available) from each stratum, yielding 422-619 (depending on the wave) clusters, and second, selecting a random sample of households from each sampled enumeration district. To the extent possible we control for sampling design in the estimation. The gross sample sizes were 4,436, 6,398, and 6,280, with response rates around 50-56% (see Figure 2.1). The number of interviewers decreased from 85 to 70, with all receiving a full day of training before fieldwork.

Interviewers were assigned households randomly throughout their Federal State and tended not to interview households in other Federal States to reduce costs. Apart from their own location interviewers had no influence over the characteristics of assigned households. In particular, the interviewer could not select "easy" households owing to (1) the exclusion of substitute households, thus interviewers had to use the strictly limited addresses; (2) a performance-related payment scheme and high set-up costs for the interviewer (the time and effort required to attend the interviewer training, to participate in the survey of interviewers, and to become familiar with the survey questionnaire, its implementation in Computer Assisted Personal Interview (CAPI), the interviewer documents and other supporting material, the contact strategies, or the processes such as the monitoring of interviewer performances or consistency checks), thus interviewers had a strong incentive to complete assigned households, as otherwise, in combination with (1),

it would have made their prior investment in training and preparation inefficient; (3) the requirement of area managers to allocate new households only if sufficient effort was made to assigned households; and (4) continuous post-interview expert analysis of interviewed households and those that refused to participate during field period (see Albacete et al., 2018 for more details). Only in rare cases (e.g. interviewer illness or drop-out) did reassignment of households take place.

Figure 2.1 Response behaviour indicators of households in each wave



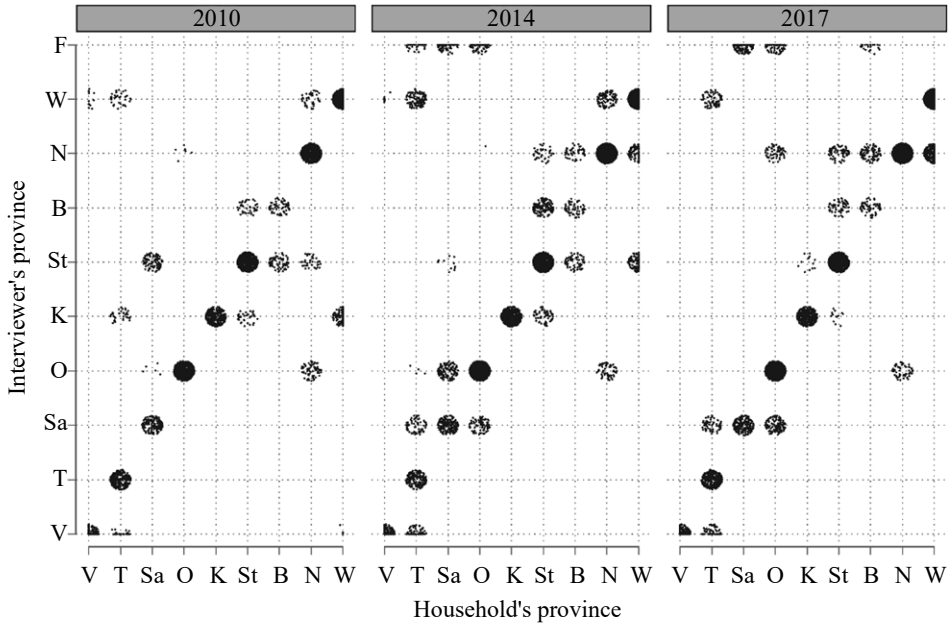
Note: This graph shows the proportion of households in the gross sample of each Household Finance and Consumption Survey (HFCS) wave which participated in the survey (response), the proportion of those which were contacted by the interviewer (contact) and the proportion of those which cooperated with the interviewer when contacted (cooperation).
Sources: HFCS Austria 2010, 2014 and 2017, Oesterreichische Nationalbank.

Figure 2.2 shows the Austrian Federal States where households and their interviewers live, including a few from Germany conducting interviews in neighboring Austrian Federal States, relevant for our model in Section 2.3.

The graphs at the top in Figure 2.3 show the increase in the number of households per interviewer, from a mean of 50 to 88, with the maximum increasing from 205 to 358.

The graphs at the bottom in Figure 2.3 display the distribution of household response rates per interviewer, showing decreased dispersion over time, indicating that performance has become less dependent on individual interviewers, likely due to increased experience and improved training.

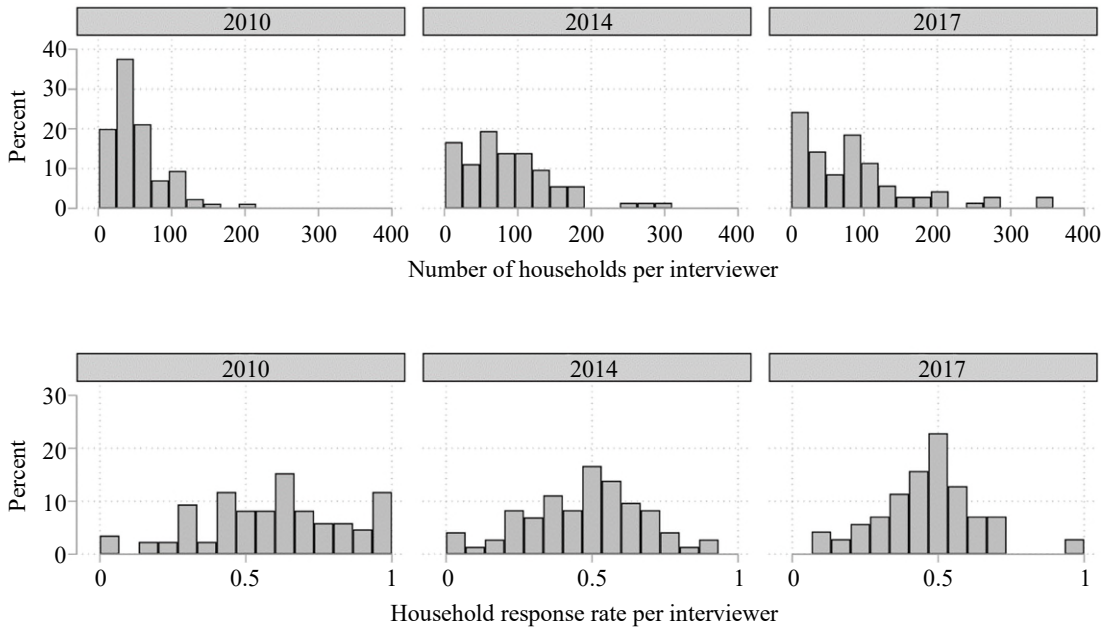
Figure 2.2 Province of households and their interviewers in each wave



Note: This graph relates the federal state of each household in the Household Finance and Consumption Survey (HFCS) gross sample to the federal state of the corresponding interviewer assigned to this household. The abbreviations stand for Vorarlberg (V), Tirol (T), Salzburg (Sa), Upper Austria (O), Carinthia (K),Styria (St), Burgenland (B), Lower Austria (N), Vienna (W) and foreign country, i.e. Germany, (F). As the data of this graph are categorical, and many of the points would be on top of each other, making it impossible to tell whether the plotted point represented one or 1,000 observations, spherical random noise has been added to the data in order to produce this graph and to avoid overprinting of the plotted points.

Sources: Household Finance and Consumption Survey (HFCS) Austria 2010, 2014 and 2017, Oesterreichische Nationalbank; Survey of Interviewers 2010, 2014 and 2017, Oesterreichische Nationalbank (data not publicly available).

Figure 2.3 Number and response rate of households per interviewer in each wave



Note: This graph shows the distribution of the number of households (at the top) and the response rates of households (at the bottom) per interviewer across Household Finance and Consumption Survey (HFCS) waves.

Sources: HFCS Austria 2010, 2014 and 2017, Oesterreichische Nationalbank.

2.2.2 Survey paradata and auxiliary information

The HFCS Austria includes various paradata and auxiliary information (Albacete and Schürz, 2014), available internally due to anonymization requirements. Table 2.1 shows a list of variables available in each wave for both respondents and non-respondents. Auxiliary information includes assessments of building type, location, condition, residential area, and security measures, collected without entering residences. Despite potential measurement errors (Sinibaldi, Durrant and Kreuter, 2013), housing type accuracy is over 95% for contacted households, giving us confidence in this information. HFCS interviewers in Austria are trained accordingly. Paradata variables include contact attempts (date, time, type, outcome) for each interview. Additional auxiliary information includes sample design data like NUTS-3 region, municipality size, and enumeration district.

Table 2.1
Paradata and auxiliary information for respondents and non-respondents in the Household Finance and Consumption Survey (HFCS) Austria

Type of paradata and auxiliary information	Details
Contact form	Number of contact attempts, type, date, time and outcome of each contact attempt.
Sample design variables	NUTS-3 region, municipality size class, enumeration district.
Interviewer assessments	Building, construction type of the building, geographical location of the building, condition of the building, condition of the building compared to other buildings in the neighbourhood, residential area, special security measures.

Note: This table shows the paradata and auxiliary information variables of the Household Finance and Consumption Survey (HFCS) Austria that are available for both respondents and non-respondents.
Sources: HFCS Austria 2010, 2014 and 2017, Oesterreichische Nationalbank (data not publicly available).

Descriptive statistics of paradata and auxiliary information are in the online Appendix (Tables A.1 to A.8). Over time, the oversampling of urban households increased: from 40% in municipalities with at least 50,000 inhabitants in the first wave to 47% in the third wave.

2.2.3 Administrative data

We also use an income database that is based on wage tax data for all Austrian municipalities, including the 23 districts of Vienna for the year 2011. This database includes the mean, median and 90th percentile of income taxpayers' gross wages (leaving out the self-employed), which are defined as all income received in a year including supplementary payments and social security contributions. The dataset links to the HFCS via municipality ID (for details see Moser and Schnetzer, 2014).

2.2.4 Interviewer survey

Each wave of the HFCS in Austria systematically collects interviewer information (Albacete and Schürz, 2013), as listed in Table 2.2. Table 2.2 shows a list of the interviewer data variables. Interviewer data includes socio-economic information, employment status, personality traits, and financial situation.

Interviewer participation was 72 of 85 in the first wave, 55 of 72 in the second, and 70 of 70 in the third wave, the latter conducted during training. Despite some nonresponse, basic variables were always available for all interviewers, suggesting minimal impact on results (see Sections 2.4 and 3). Another interviewer-level variable consistently available for all interviewers is workload, represented by the total number of sampled households assigned to each interviewer.

Table 2.2
Interviewer data in the Household Finance and Consumption Survey (HFCS) Austria

Type of interviewer data	Details
Socio-demographic characteristics	Gender, age, region, migration background, marital status, education, parental education.
Socio-economic characteristics	Real estate ownership, employment, occupation, experience as an interviewer, experience with similar surveys, income, wealth.
Assessments	Trust, big five psychological profile (25-question battery), opinions on redistribution of income and wealth.

Note: This table shows the interviewer variables of the Survey of Interviewers.
Sources: Survey of Interviewers 2010, 2014 and 2017, Oesterreichische Nationalbank (data not publicly available).

A degree of item nonresponse, i.e. some interviewers not answering on certain variables, has to be taken into account in the further analysis. Therefore, in the regressions we interact each regressor containing missing values with a dummy variable indicating whether or not the observation of the regressor is missing (also for interviewers who refused to participate). These dummy variables were included in the models to ensure that missing data did not bias the results. By interacting the dummy variables with the respective characteristics, we were able to control for the potential impact of missing data on our analysis, maintaining the robustness of our findings. In the case of the interviewer income and net wealth variables, if information about bounds was provided by the interviewer, we impute the mean between the lower and upper bound (if both bounds were provided) or we impute either the lower or the upper bound (if only one bound was provided). For these reasons, the impact of interviewer item nonresponse on the results of the survey of interviewers should also be limited.

Descriptive statistics of interviewer variables are in the online Appendix (Tables A.9 to A.16). Over the waves, interviewer experience increased from 83 to 96 months, despite a decrease in mean age. Interviewers became less open to experience but more agreeable and conscientious, which means that they are less inventive/curious but more friendly/compassionate and efficient/organised (Gerlitz and Schupp, 2005). The proportion of female interviewers increased from 49% to 64%.

2.3 Estimation strategy

As mentioned before, our main goals are (1) to identify the amount of variation in household response behaviour explained at the interviewer level and (2) to analyse the effect of interviewer characteristics and personality traits on household response behaviour. In this framework, the use of standard regression models is not appropriate, as the assumption of independence of all observations is violated by the fact that

observations from the same interviewer are generally more similar to each other than observations from different interviewers – for example, because of the use of regional criteria in the assignment of households to the interviewers (see Section 2.2). Therefore, we use multilevel regression models (for details see Hox, 1994; West and Blom, 2017, or Rabe-Hesketh and Skrondal, 2022; we employ the same notation).

In contrast to the standard logistic regression model estimated for all households i, we assume that each interviewer j has a different intercept coefficient $\boldsymbol{\beta}_{0j}$, whereas the slope coefficient vector $\boldsymbol{\beta}_p$ is assumed to be the same for each interviewer. Thus, imagine that, for each household i and for each interviewer j, there is a continuous latent variable $\mathbf{Y}_{ij}^*$ describing household response propensity that is distributed as follows:

$$\mathbf{Y}_{ij}^* = \boldsymbol{\beta}_{0j} + \boldsymbol{\beta}_p \mathbf{X}_{pij} + \mathbf{e}_{ij} \tag{2.1}$$

where $\mathbf{X}_{pij}$ is a vector of explanatory variables $(p=1,\ldots,P)$ at the respondent level and $\mathbf{e}_{ij}$ is assumed to have a standard logistic distribution with mean zero and variance $\sigma_e^2 = \frac{\pi^2}{3}$. The binary responses $\mathbf{Y}_{ij}$ are determined via the usual threshold model:

$$\mathbf{Y}_{ij} = \begin{cases} 1 & \text{if } \mathbf{Y}_{ij}^* > 0 \\ 0 & \text{otherwise} \end{cases}. \tag{2.2}$$

Furthermore, we explain the variation of the intercept coefficient by:

$$\boldsymbol{\beta}_{0j} = \boldsymbol{\gamma}_{00} + \boldsymbol{\gamma}_{0q} \mathbf{Z}_{qj} + \mathbf{u}_{0j} \tag{2.3}$$

where $\mathbf{Z}_{qj}$ is a vector of explanatory variables $(q=1,\ldots,Q)$ at the interviewer level, the corresponding slope coefficient vector $\boldsymbol{\gamma}_{0q}$, the intercept coefficient $\boldsymbol{\gamma}_{00}$, and $\mathbf{u}_{0j}$ is assumed to have a normal distribution with mean zero and variance $\sigma_{u_0}^2$. After substituting (2.3) into (2.1) we obtain a single complex regression equation ("random-intercept model") with a fixed and a random part:

$$\mathbf{Y}_{ij}^* = \left[\boldsymbol{\gamma}_{00} + \boldsymbol{\gamma}_{0q} \mathbf{Z}_{qj} + \boldsymbol{\beta}_p \mathbf{X}_{pij}\right] + \left[\mathbf{u}_{0j} + \mathbf{e}_{ij}\right] \tag{2.4}$$

with $\mathbf{u}_{0j}$ being assumed to be independent from $\mathbf{e}_{ij}$.

This model can also be used to produce an estimate to express the extent to which observations depend on interviewers (goal 1 from above). It indicates the proportion of the variance explained by the interviewer grouping structure and is called the intraclass correlation coefficient:

$$\rho = \frac{\sigma_{u0}^2}{\sigma_{u0}^2 + \frac{\pi^2}{3}} \tag{2.5}$$

To test for the null hypothesis that the intraclass correlation coefficient ρ is 0 we use a likelihood ratio-test for whether the residual between-interviewer variance σ_{u0}^2 is 0. Its statistic is not the usual chi-squared with 1 degree of freedom but is instead a 50:50 mixture of a chi-squared with no degrees of freedom and a chi-squared with 1 degree of freedom (see Zhang and Lin, 2007). The p-value of the LR test takes this into account.

2.4 Model specification

We estimate a two-level random-intercept logistic regression model with explanatory variables at the respondent and interviewer levels. Although regional income from the income database is measured at the municipality level, we include it as a respondent-level variable to focus on the interviewer and respondent levels. We investigate interviewer effects using the following nine specifications. In specification 1 we include only the constant. Specification 2 additionally makes use of basic respondent level (dwelling type, dwelling location, dwelling surrounding, state, municipality size and mean municipality wage) variables. Interviewer level information (gender, age, state, education, experience in months as an interviewer, experience with similar surveys and workload) extends the model further in specification 3. Starting from model 3, specifications 4a through to 4f include separately at the interviewer level the labour status, marital status, migration background, home ownership status, personality information (trust and big five personality traits) and economic resources (interviewer's household income and net wealth). The intention of this set-up is to display the stability of results over model specification.

We (partly) account for stratification effects in our models by including Federal States (NUTS2) and municipality size class as explanatory variables. As a robustness check, in a separate set of models we also account for cluster effects by additionally including clusters as explanatory variables. Note, that while we included clusters as fixed explanatory variables in our robustness checks, we acknowledge this differs from the conventional use of random effects. However, due to the extremely fine-grained stratification in our sampling design, cluster effects are likely minimal given the low geographic and demographic variability within strata. Additionally, Austria's low socio-economic segregation, even in urban areas, further reduces within-strata variability in response behavior. Lastly, the assignment of one or very few interviewers to each cluster inherently limits the ability to disentangle interviewer from cluster effects.

3. Results

Table 3.1 presents the main results for the pooled sample and all specifications, including wave dummies to control for survey year differences. These results are qualitatively robust against additionally accounting for cluster effects (see Table A.28 in the online Appendix), which is probably due to the very small strata size. Estimating a three-level random intercept model taking into account the Federal States shows that interviewers also account for a significant share of the response variance within those regions. Descriptive statistics of the mean response rates of households across interviewer and household characteristics in each wave and overall are in the online Appendix (see Tables A.17 to A.24), while the estimation results of these regression models for each wave are also in the online Appendix (see Tables A.25 to A.27).

Significant interviewer effects include homeownership, being married, having a university degree, and having a larger workload all increasing household response propensity at the 5% significance level. Interviewer personality, specifically lower openness, positively affects response, although only weakly statistically significant (p-value = 0.102). No significant effects were found for age, gender, or general

experience, but the sequential number of interviews conducted ("household interview order") significantly increases response propensity. No significant effects were found for the dummy variables indicating whether or not an observation of an interviewer-level regressor is missing (not shown in Table 3.1; see Section 2.2.4 for more details), except for missingness of interviewer education and interviewer household net wealth (negative impact), and missingness of interviewer openness (positive impact).

The social environment significantly affects response propensity. Higher mean wages in the municipality decrease response at the 1% level. Living in an individual house or in a town/city, rather than the countryside, also negatively impacts response at the 1% level. Similarly, living in Vienna, rather than in any other Federal State, negatively impacts response at the 1% level.

Table 3.1
Random-intercept logistic regression estimation of household response (all waves)

	(1)	(2)	(3)	(4a)	(4b)	(4c)	(4d)	(4e)	(4f)
Survey wave (Ref: 2010):									
2014	-0.600***	-0.561***	-0.572***	-0.535***	-0.531***	-0.549***	-0.509***	-0.556***	-0.497***
2017	-0.669***	-0.584***	-0.822***	-0.855***	-0.811***	-0.795***	-0.766***	-0.860***	-0.839***
Household level information:									
Dwelling type (Ref: apartment):									
Individual house		-0.225***	-0.227***	-0.227***	-0.226***	-0.226***	-0.226***	-0.226***	-0.225***
Semi-detached house		-0.145*	-0.147**	-0.148**	-0.148**	-0.147**	-0.149**	-0.148**	-0.147**
Other		0.283	0.273	0.271	0.272	0.272	0.274	0.263	0.275
Dwelling location (Ref: countryside):									
Downtown		-0.233***	-0.226***	-0.227***	-0.225***	-0.226***	-0.223***	-0.225***	-0.227***
In between		0.129**	0.142**	0.142**	0.142**	0.142**	0.146**	0.142**	0.141**
Town outskirts		-0.009	-0.003	-0.002	-0.001	-0.003	-0.002	-0.003	-0.003
Surrounding rating (Ref: Mid-range):									
Luxury		0.650***	0.658***	0.659***	0.656***	0.658***	0.660***	0.657***	0.655***
Upscale		0.298***	0.302***	0.301***	0.301***	0.302***	0.303***	0.303***	0.301***
Modest		0.020	0.027	0.028	0.027	0.028	0.027	0.027	0.030
Low-income		0.370**	0.373**	0.373**	0.380**	0.374**	0.378**	0.367**	0.375**
Very low-income		-0.527	-0.537	-0.534	-0.531	-0.538	-0.532	-0.550	-0.544
Household interview order		0.001***	0.001**	0.001**	0.001**	0.001**	0.001**	0.001**	0.001**
Austrian Federal State (Ref: Vienna):									
Vorarlberg		1.335***	1.574***	1.567***	1.583***	1.573***	1.565***	1.543***	1.563***
Tyrol		0.718***	0.698***	0.696***	0.704***	0.696***	0.695***	0.694***	0.695***
Salzburg		0.605***	0.548***	0.551***	0.556***	0.551***	0.552***	0.541***	0.552***
Upper Austria		0.868***	0.862***	0.859***	0.861***	0.863***	0.860***	0.857***	0.864***
Carinthia		1.185***	1.372***	1.375***	1.362***	1.376***	1.369***	1.368***	1.373***
Styria		1.148***	1.174***	1.177***	1.169***	1.174***	1.156***	1.141***	1.159***
Burgenland		1.113***	1.159***	1.163***	1.147***	1.161***	1.139***	1.135***	1.150***
Lower Austria		0.977***	1.119***	1.116***	1.109***	1.120***	1.118***	1.118***	1.119***
Municipality size (Ref : <2thd):									
2-3thd		0.083	0.080	0.080	0.081	0.080	0.079	0.078	0.081
3-5thd		0.020	0.009	0.010	0.009	0.009	0.009	0.009	0.011
5-10thd		0.039	0.048	0.049	0.049	0.049	0.047	0.049	0.050
10-20thd		0.270***	0.270***	0.272***	0.274***	0.271***	0.272***	0.270***	0.275***
20-50thd		0.425***	0.448***	0.448***	0.445***	0.449***	0.448***	0.432***	0.452***
50thd-1mill		0.011	0.015	0.017	0.012	0.015	0.014	0.020	0.018
>1mill		0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Mean wage in municipality (in thd)		-0.000***	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***

Notes: - This table shows the regression and intraclass correlation coefficient estimates of running a random-intercept logistic regression of household response. The following variables have item nonresponse at the interviewer level: experience as interviewer, neuroticism, extraversion, openness, agreeableness, conscientiousness, income, and wealth; The corresponding missing values were set to "0" and interacted with a missing dummy which is omitted from the output in this table.
- * p < 0.10, ** p < 0.05, *** p < 0.01.

Sources: Household Finance and Consumption Survey (HFCS) Austria 2010, 2014 and 2017, Survey of Interviewers 2010, 2014 and 2017, Oesterreichische Nationalbank (data not publicly available), Lohnsteuerstatistik 2011.

Table 3.1 (continued)
Random-intercept logistic regression estimation of household response (all waves)

	(1)	(2)	(3)	(4a)	(4b)	(4c)	(4d)	(4e)	(4f)
Interviewer level information:									
Female			-0.125	-0.129	-0.123	-0.124	-0.131	-0.124	-0.131
Age			0.004	0.007	-0.006	0.004	0.002	0.010	-0.002
Age squared			0.000	0.000	0.000	0.000	0.000	-0.000	0.000
Province (Ref: Vienna):									
Vorarlberg			-0.652*	-0.607*	-0.738**	-0.651*	-0.698*	-0.653*	-0.659*
Tyrol			0.148	0.174	0.089	0.180	0.114	0.091	0.134
Salzburg			0.147	0.061	0.102	0.140	0.081	0.285	0.170
Upper Austria			-0.203	-0.224	-0.208	-0.182	-0.232	-0.280	-0.176
Carinthia			-0.505*	-0.486	-0.552*	-0.492	-0.578*	-0.600**	-0.479
Styria			-0.147	-0.086	-0.223	-0.147	-0.236	-0.232	-0.113
Burgenland			-0.024	0.039	-0.070	0.003	-0.085	-0.036	0.035
Lower Austria			-0.600***	-0.560***	-0.600***	-0.586***	-0.668***	-0.611***	-0.609***
Foreign country			0.581	0.695	0.664	0.597	0.514	0.541	0.684
Education (Ref: University):									
Primary School			0.254	0.255	0.138	0.254	0.280	0.212	0.200
Vocational School/Apprenticeship			-0.208	-0.212	-0.291	-0.212	-0.207	-0.244	-0.198
Upper/Post Secondary			-0.374**	-0.408**	-0.395**	-0.369**	-0.323*	-0.317*	-0.368**
Experience as interviewer (in months)			0.000	0.000	0.000	0.000	0.000	0.000	0.000
No experience with similar surveys			0.273	0.218	0.318	0.260	0.310	0.319	0.261
Workload (number of addresses)			0.002**	0.002**	0.002**	0.002**	0.002**	0.002**	0.002**
Labour status (Ref: Employee):									
Self-employed				-0.110					
Unemployed				0.069					
Retired				-0.238					
Other				0.651					
Not married					-0.307**				
Not migrant						-0.034			
Not homeowner							-0.281**		
Trust								0.097	
Neuroticism (in points)								-0.003	
Extraversion (in points)								0.015	
Openness (in points)								-0.024	
Agreeableness (in points)								0.028	
Conscientiousness (in points)								0.005	
Household income (in thd)									0.000
Household net wealth (in thd)									0.000
Constant	0.476***	0.008	-0.305	-0.349	-0.075	-0.251	-0.377	-0.504	0.464
Random intercept variance	0.823***	0.736***	0.572***	0.564***	0.548***	0.571***	0.553***	0.531***	0.534***
Intraclass correlation coefficient	0.200	0.183	0.148	0.146	0.143	0.148	0.144	0.139	0.140
LR test vs. Logistic model (test statistic)	1,458.10***	1,060.88***	807.25***	790.14***	731.47***	806.22***	778.46***	740.62***	725.34***
Log-Likelihood	-10,656	-10,457	-10,434	-10,432	-10,431	-10,434	-10,432	-10,428	-10,428
AIC	21,319	20,975	20,970	20,977	20,968	20,974	20,970	20,981	20,967
BIC	21,350	21,214	21,363	21,408	21,377	21,382	21,378	21,467	21,391
N	16,465	16,465	16,465	16,465	16,465	16,465	16,465	16,465	16,465

Notes: - This table shows the regression and intraclass correlation coefficient estimates of running a random-intercept logistic regression of household response. The following variables have item nonresponse at the interviewer level: experience as interviewer, neuroticism, extraversion, openness, agreeableness, conscientiousness, income, and wealth; The corresponding missing values were set to "0" and interacted with a missing dummy which is omitted from the output in this table.
- * p < 0.10, ** p < 0.05, *** p < 0.01.
- AIC = Akaike Information Criterion; BIC = Bayesian Information Criterion.

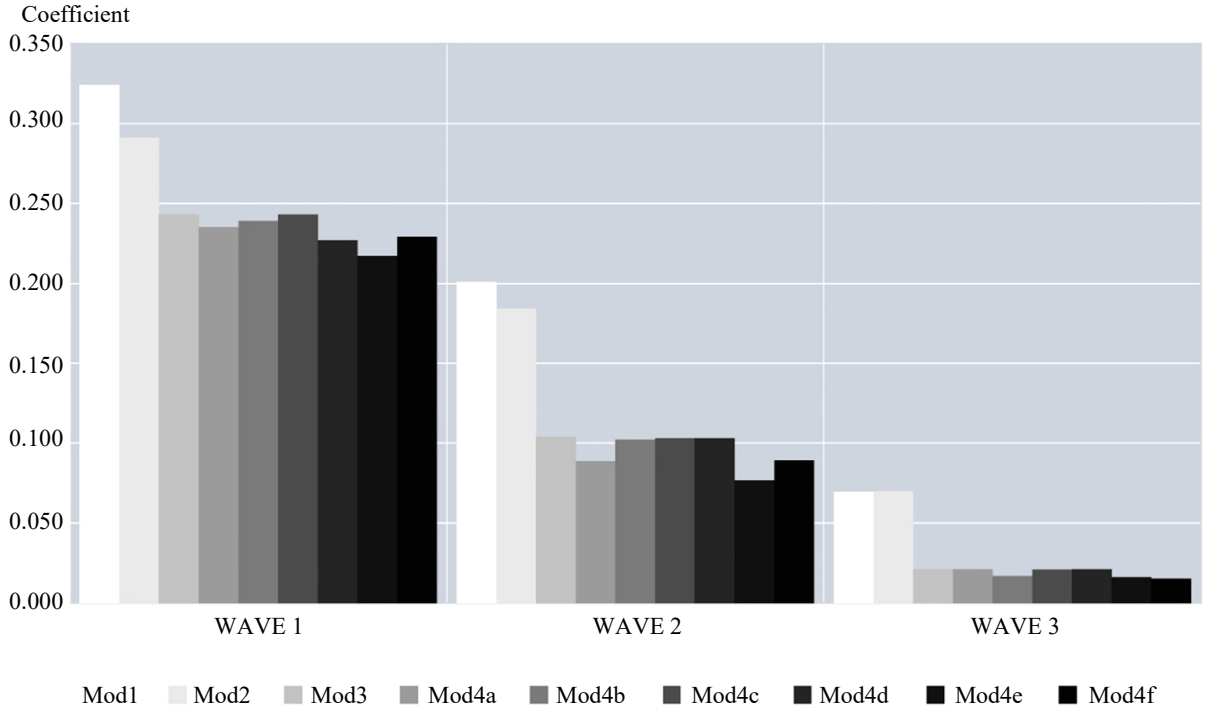
Sources: Household Finance and Consumption Survey (HFCS) Austria 2010, 2014 and 2017, Survey of Interviewers 2010, 2014 and 2017, Oesterreichische Nationalbank (data not publicly available), Lohnsteuerstatistik 2011.

Table 3.1 provides likelihood-ratio tests for the null hypothesis that the residual between-interviewer variance is equal to zero. They are based on mixtures of chi-square distributions as described in Section 2.3. For all specifications, the likelihood-ratio statistic is between 1,458 and 725, giving p-values < 0.001, which suggests that a multilevel model is required for all specifications.

Table 3.1 also shows the intraclass correlation coefficient (ICC), measuring the dependence of household responses on interviewers (see Section 2.3). The ICC, ranging from 0 (no dependence) to 1 (complete dependence), explains the variance in household response due to interviewer grouping. Using the pooled sample, the ICC shows interviewer dependence in all specifications (see Table 3.1). In specification 1, without explanatory variables, interviewer grouping explains one-fifth of the variance in household response. As more variables are added, the ICC decreases; for instance, in specification 4e, the remaining unexplained interviewer variance is 0.139.

The ICC decreases over the waves, from 0.324 in the first wave to 0.201 in the second and 0.0696 in the third wave (see Figure 3.1 and Tables A.25 to A.27 in the online Appendix). This indicates a significant reduction in the interviewer grouping's impact on variance in household response across waves.

Figure 3.1 Intraclass correlation coefficient estimate across model specifications in each wave



Note: This graph shows the intraclass correlation coefficient estimated with the help of the multilevel regression model for each specification and for each wave. The coefficient explains how much of the variance in household response can be explained by the interviewer grouping structure in the household sample and lies between 0 (no dependence) and 1 (complete dependence).

Sources: Household Finance and Consumption Survey (HFCS) Austria 2010, 2014 and 2017, Survey of Interviewers 2010, 2014 and 2017, Oesterreichische Nationalbank (data not publicly available).

Survey administration, interviewer training, and sample design remained stable, but interviewer experience increased over the waves (see Section 2.2). Low-experience interviewers may exhibit greater variance due to trying different strategies, whereas high-experience interviewers likely use a consistent, effective strategy.

We test this hypothesis by fitting an OLS model of mean household response rate per interviewer and plotting the residuals against interviewer experience. Chart A.1 in the online Appendix supports our hypothesis: variance for low-experience interviewers is greater than for high-experience interviewers. A likelihood ratio test for heteroscedasticity results in a $\chi^2(1)$ statistic of 6.64 which confirms this relationship as statistically significant. Re-running the OLS for a subsample of very experienced interviewers shows that even after a certain beginning period (80 months of experience) the extra variance continues to decrease significantly (in this case, a likelihood ratio test for heteroscedasticity results in a $\chi^2(1)$ statistic of 3.69). These results align with previous findings showing no impact of experience on mean response rates but a significant impact on variance.

Other omitted factors may explain the decline in the intraclass correlation over waves, which is beyond the scope of this paper and left for future research.

4. Conclusions

This section summarizes our key findings and provides practical recommendations based on our analysis.

4.1 Summary

Our multilevel model reveals a significant decrease in the variance of response behavior attributed to interviewer effects, from about one-third in the first wave to about 7% in the third wave of the HFCS. We suspect this decline is due to increased interviewer experience, leading to more consistent and effective survey strategies. Key interviewer characteristics positively influencing household response rates include higher education, being married, homeownership, and having a larger workload. Additionally, higher mean wages in a municipality and urban dwelling locations are negatively associated with response rates. These findings underscore the importance of considering interviewer effects and regional characteristics in survey analysis.

4.2 Recommendations for practice

Our findings suggest several key recommendations to improve survey response rates and data quality.

First, our analysis highlights the importance of targeted assignment of interviewers to the field. Interviewers with higher education, stable marital status, and homeownership are associated with higher response rates. This suggests that survey organizations should prioritize deploying individuals with these characteristics in the field to enhance respondent participation.

Second, the implementation of comprehensive training programs is crucial. Our study shows that interviewer experience significantly impacts response behavior (through the sequential number of interviews conducted during the field period) and its variance (through general interviewer experience in months). Training should focus on effective engagement strategies, including explaining the survey process,

motivating respondents, reading questions verbatim, and using non-directive probing techniques. These practices can help reduce variability in interviewer performance and improve overall response rates.

Third, our results suggest looking at socio-economic characteristics of survey areas to try to mitigate negative impacts on response rates. For instance, higher mean wages in a municipality and urban dwelling locations are linked to lower response rates. Survey organizations should adapt their strategies to address these specific regional challenges, such as by providing additional support or resources in areas with higher nonresponse risks.

Finally, incorporating personality assessments in the recruitment process can further enhance response rates. Our study finds weak evidence that interviewers with lower openness to experience are more effective in obtaining responses. By focusing on personality traits that positively influence respondent participation, organizations can select interviewers who are better suited to the demands of survey administration.

While survey nonresponse has been increasing over time, particularly for personal interviews, these methods remain crucial for high-quality data gathering. Addressing the nonresponse problem at its source is essential for maintaining the representativeness of survey data. Future research should focus on the interaction between interviewers and respondents, utilizing all available data to tailor solutions for nonresponse.

Acknowledgements

We would like to thank the participants of the HFCN Research Seminar in Krakow (June 2019), the editor and three anonymous referees for valuable comments and discussion. Additional to the usual disclaimer, the opinions expressed in this paper solely represent those of the authors and do not necessarily reflect the official viewpoint of the Oesterreichische Nationalbank or of the Eurosystem.

References

Albacete, N. and Schürz, M. (2013). *Interviewereffekte beim HFCS Austria 2010*: Statistiken.

Albacete, N. and Schürz, M. (2014). *Paradaten im HFCS Austria 2010 – Teil 1: Evaluierung von Non-Response-Fehlern*: Statistiken.

Albacete, N., Lindner, P. and Wagner, K. (2016). Eurosystem Household Finance and Consumption Survey 2014. Methodological Notes for Austria (second wave). *Monetary Policy & the Economy*, Q2/16 – Addendum.

Albacete, N., Dippenaar, S.T., Lindner, P. and Wagner, K. (2018). Eurosystem Household Finance and Consumption Survey 2017. Methodological Notes for Austria. *Monetary Policy & the Economy*, Q4/18 – Addendum.

Albacete, N., Lindner, P., Wagner, K. and Zottel, S. (2012). Eurosystem Household Finance and Consumption Survey 2010. Methodological Notes for Austria. *Monetary Policy & the Economy*, Q3/12 – Addendum.

Beerten, R. (1999). The effect of interviewer and area characteristics on survey response rates: An explanatory analysis. *Survey Methodology Bulletin*, 45, 7-15.

Blom, A.G. and Korbmacher, J.M. (2013). Measuring interviewer characteristics pertinent to social surveys: A conceptual framework. *Survey Methods: Insights from the Field*, 1-16.

Cialdini, R.B., Couper, M.P. and Groves, R.M. (1992). Understanding the decision to participate in a survey. *Public Opinion Quarterly*, 56, 4, 475-495.

De Leeuw, E.D., Hox, J.J. and Dillman, D.A. (2008). International handbook of survey methodology. *European Association of Methodology Series*.

Durrant, G.B., Groves, R.M., Staetsky, L. and Steele, F. (2010). Effects of interviewer attitudes and behaviors on refusal in household surveys. *Public Opinion Quarterly*, 74, 1-36.

Fessler, P., Lindner, P. and Schürz, M. (2016). Eurosystem Household Finance and Consumption Survey 2014: First results for Austria (second wave). *Monetary Policy & the Economy*, Q2/16, 34-95.

Fessler, P., Lindner, P. and Schürz, M. (2018). Eurosystem Household Finance and Consumption Survey 2017 for Austria. *Monetary Policy & the Economy*, Q4/18, 36-66.

Fessler, P., Mooslechner, P. and Schürz, M. (2012). Eurosystem Household Finance and Consumption Survey 2010. First results for Austria. *Monetary Policy & the Economy*, Q3/12, 23-62.

Gerlitz, J.-Y. and Schupp, J. (2005). Zur Erhebung der Big-Five-basierten Persönlichkeitsmerkmale im SOEP. *Research Notes*, 4, 1-36.

Groves, R.M. and Couper, M.P. (1998). Nonresponse in household interview surveys. *Wiley Series in Survey Methodology*, John Wiley & Sons, Inc.

Houtkoop-Steenstra, H. and van den Bergh, H. (2000). Effects of introductions in large-scale telephone survey interviews. *Sociological Methods and Research*, 28, 3, 251-280.

Hox, J.J. (1994). Hierarchical regression models for interviewer and respondent effects. *Sociological Methods & Research*, 22, 3, 300-318.

Hox, J.J. and de Leeuw, E. (2002). The influence of interviewers' attitude and behavior on household survey nonresponse: An international comparison. In *Survey Nonresponse*, (Eds., R.M. Groves, D.A. Dillman, J.L. Eltinge and R.J.A Little), New York: John Wiley & Sons, Inc.

Jäckle, A., Lynn, P., Sinibaldi, J. and Tipping, S. (2013). The effect of interviewer experience, attitudes, personality and skills on respondent co-operation with face-to-face survey. *Survey Research Methods*, 7, 1, 1-15.

Kennickell, A. (1999). Analysis of nonresponse effects in the 1995 Survey of Consumer Finances. *Journal of Official Statistics*, 15, 2, 283-303.

Kennickell, A. (2006a). How do we know if we aren't looking? An investigation of data quality in the 2004 SCF. *SCF Working Paper*, Federal Reserve Board, Washington.

Kennickell, A. (2006b). Who's asking? Interviewers, their incentives, and data quality in field surveys. *SCF Working Paper*, Federal Reserve Board, Washington.

Kennickell, A. (2008). Look and listen, but don't stop: Interviewers and data quality in the 2007 SCF. *SCF Working Paper*, Federal Reserve Board, Washington.

Kreuter, F. (2008). Interviewer effects. In *Encyclopedia of Survey Research Methods*, (Ed., P.J. Lavrakas), SAGE Publications, Inc, Thousand Oaks, California, 1, 369-371.

McCrae, R.R. and John, O.P. (1992). An introduction to the five-factor model and its applications. *Journal of Personality*, 60, 175-215.

Merkle, D. and Edelman, M. (2002). Nonresponse in exit polls: A comprehensive analysis. In *Survey Nonresponse*, (Eds., R.M. Groves, D.A. Dillman, J.L. Eltinge and R.J.A. Little), New York: John Wiley & Sons, Inc, 243-258.

Morton-Williams, J. (1993). *Interviewer Approaches*. Dartmouth Publishing Co Ltd, Dartmouth, United Kingdom.

Moser, M. and Schnetzer, M. (2014). The geography of average income and inequality. Spatial evidence from Austria. *Department of Economics Working Paper*, 191.

O'Muircheartaigh, C. and Campanelli, P. (1999). A multilevel exploration of the role of interviewers in survey non-response. *Journal of the Royal Statistical Society, Series A (Statistics in Society)*, 162, 3, 437-446.

Pickery, J. and Loosveldt, G. (2002). A multilevel multinomial analysis of interviewer effects on various components of unit nonresponse. *Quantity and Quality*, 36, 4, 427-437.

Rabe-Hesketh, S. and Skrondal, A. (2022). Multilevel and longitudinal modeling using Stata, 4th ed., StataCorp LP.

Schaeffer, N.C., Dykema, J. and Maynard, D.W. (2010). Interviewers and Interviewing. In *Handbook of Survey Research*, (Eds., P.V. Marsden and J.D. Wright), Emerald, Bingley, United Kingdom, 437-470.

Singer, E., Glassman, M.B. and Frankel, M.R. (1983). The effect of interviewer characteristics and expectations on response. *Public Opinion Quarterly*, 47, 1, 68-83.

Sinibaldi, J., Durrant, G.B. and Kreuter, F. (2013). Evaluating the measurement error of interviewer observed paradata. *Public Opinion Quarterly*, 77, S1, 173-193.

West, B.T. and Blom, A.G. (2017). Explaining interviewer effects: A research synthesis. *Journal of Survey Statistics and Methodology*, 5, 2, 175-211.

Zhang, D. and Lin, X. (2007). Variance component testing in generalized linear mixed models for longitudinal/clustered data and other related topics. In *Model Selection in Linear Mixed Models*, (Ed., D. Dunson), Berlin Heidelberg New York: Springer.