Consent when Linking Survey Data with Administrative Records: The Role of the Interviewer

Julie M. Korbmacher
MEA, Munich

Mathis Schroeder
DIW, Berlin

Linking survey data with administrative records is becoming more common in the social sciences in recent years. Regulatory frameworks require the respondent's consent to this procedure in most cases. Similar to non-response, non-consent may lead to selective samples and could pose a problem when using the combined data for analyses. Thus investigating the selectivity and the determinants of the consent decision is important in order to find ways to reduce non-consent.

Adapting the survey participation model by Groves and Couper (1998), this paper identifies different areas influencing the respondents' willingness to consent. In addition to control variables at the individual and household level, two further areas of interest are included: the interview situation and the characteristics of the interviewer. A multilevel approach highlights the importance of the interviewer for the consent decision: the empty model shows an intra-class correlation of 55%, which can be reduced to 35% in a full model including interviewer variables. An additional analysis including measures of interviewer performance shows that there are further unobserved interviewer characteristics that influence the respondents consent decision.

The results suggest that although respondent and household characteristics are important for the consent decision, a large part of the variation in the data is explained by the interviewers. This finding stresses the importance of the interviewers not only as an integral part in data collection efforts, but also as the direct link to gain a respondent's consent for linking survey data with administrative records.

Keywords: record linkage; interviewers

1 Introduction

The number of projects linking survey data with administrative records is increasing. At the conference of the European Survey Research Association (ESRA) in July 2011, sixteen papers dealt with challenges of, and findings from, combining survey data with administrative records, a sharp increase from four papers in the 2009 conference. While record linkage is becoming more popular in the social sciences, it is already common in other fields, especially in epidemiology. The enrichment of survey data through administrative records is the primary motivation for the linkage (Calderwood and Lessof 2009). The data quality in the resulting data sets provides excellent opportunities for research, but the linked data also help to reduce the burden for respondents and interviewers as well as survey cost (Sala, Burton and Knies 2010; Schnell 2012). Although administrative data are not primarily generated for research purposes, there are some advantages compared to survey data. For example, they usually cover the universe of the population of interest and are thought to be more accurate than survey data, because problems arising in surveys, such as recall error or misreporting, may not affect the quality of adminis-

trative data as severely (Calderwood and Lessof 2009).[1] On the other hand, administrative data are often collected for a specific purpose and only include standardized information, such as process data for a hospital visit. Unlike surveys, researchers have no, or only limited, influence on what data are collected (Hartmann and Krug 2009). Thus, using administrative data alone may restrict the selection of control variables. Therefore, researchers benefit if they can combine survey data with administrative records.

There are two common ways to establish the data linkage: one is to use statistical matching procedures based on distance measures, where respondents from a survey are matched to "similar" (in a statistical sense) people in the administrative records. The other way is to ask respondents directly for the permission to link their survey data to their administrative records, building a direct link between the two data sources (see Calderwood and Lessof 2009, for an overview). There are advantages and disadvantages to both procedures: the first approach heavily depends on the variables identical in both data sources and the smaller this overlap is, the harder it is to establish a match that is statistically sound.[2] As long as matching the datasets does not allow the

---

Contact information: Julie M. Korbmacher, MEA, Munich, e-mail: korbmacher@mea.mpisoc.mpg.de

[1] As Groen (2012) points out, administrative data could also suffer from errors related to imputation and editing, even though they may exceed survey data in terms of quality in many contexts.

[2] Specifically, the distance measure relies on the conditional independence assumption: conditionally on the variables identical in

specific and exact identification of respondents, data protection regulations usually do not require the respondents' permission to the matching procedure (Rasner 2012). For the direct linkage on the other hand, data quality is usually thought to be more promising, but, in most cases, the informed consent by the respondent is necessary (Lessof 2009; Schnell 2012).

There are some examples of surveys both in the United States and in Europe asking respondents to consent to linking their data to administrative records. In the US, a well-known example is the Health and Retirement Study (HRS), where respondents were asked for their consent to link their survey data to data from the Social Security Administration (Olson 1999). In the UK both the English Longitudinal Study of Ageing (Lessof et al. 2004) and the ISMIE-Survey, a subsample of the European Community Household Panel (ECHP), link their survey data to administrative records from social security and employer records (Jäckle et al. 2004).

The number of surveys that directly link survey and administrative data in Germany is increasing, mainly because the "Research Data Centers" ("Forschungsdatenzentren", FDZ), providing administrative records for research purposes, were established in 2001 (Gramlich et al. 2010). For example, the Panel Study "Labour Market and Social Security" (PASS) linked survey data with administrative records from the German Federal Employment Agency during its initial wave, conducted in 2006/2007 (Trappmann et al. 2009); the same records were linked with the ALWA-Survey ("Arbeiten und Lernen im Wandel"; Antoni and Seth 2011) as well as with the lidA-survey ("leben in der Arbeit"; Tisch and Tophoven 2011).

Even though the regulatory framework is different in different countries, there is one similarity to all these studies: without the respondents' explicit and informed consent (written or verbal), the linkage of a person's survey data with their administrative records is generally not possible. The consent decision, specifically a refusal to consent, leads to methodological complications, because the sample size of usable combined data decreases and consent bias may be an issue if there are systematic differences between those individuals who consent and those who do not. Consequently, understanding the mechanisms behind the consent decision is important for determining the sources of possible biases and reducing their influence in the future.

This paper investigates the determinants of consent to record linkage in the German part of the Survey of Health, Ageing and Retirement in Europe, SHARE, where in 2008/2009 a pilot study was conducted to link SHARE with administrative records of the "Deutsche Rentenversicherung" (DRV), the German Pension Fund. Our results show that while there are effects at the respondent level that determine consent, interviewers are important to the consent decision as well. Section 2 presents a brief overview of the previous literature on consent to data linkage, followed by a description of the linked data sets and the linkage procedure in section 3. Section 4 develops a model of consent and shows how determinants of consent are measured. The empirical results follow in section 5, while section 6 concludes with a brief summary and a discussion of the findings.

2 Previous Research

Systematic research on the differences between consenting and non-consenting respondents is not widespread (Sala, Burton and Knies 2010) and is typically found in medical and epidemiological studies (Jenkins et al. 2006). The majority of studies analyze respondent characteristics such as demographics (like age and gender), health status and socioeconomic factors (like education and income), finding some significant differences (e.g. Woolf et al. 2000; Dunn et al. 2004; Kho et al. 2009). Dunn et al. (2004) analyzed data from seven epidemiological mail surveys conducted in the UK, which all contained demographic, disease-specific, and generic items. They considered consent to follow up and/or the review of medical records and found effects of age, gender as well as the symptoms under investigation. In another meta-analysis, Kho et al. (2009) report statistical differences with respect to respondents characteristics (i.e. age, sex, race, income, education and health status) between consenters and non-consenters when using data from 17 unique medical studies, where the influences differ between the studies in direction and magnitude. There is also evidence for an impact of other socio-economic factors such as area effects (e.g. Huang et al., 2007). It is not clear whether the results from these medical studies can be easily transferred to consent questions covering different topics. In addition, the studies mentioned above are all limited to influences of respondents characteristics.

Early work in the social sciences about potential selection bias in linked data sets is based on the Health and Retirement Study (HRS), where data was linked to administrative records from the Social Security Administration (SSA) in 1992. The analyses provide evidence of a consent bias related to respondent characteristics like age, race, gender, income or education (e.g. Olson 1999; Gustman and Steinmeier 1999; Haider and Solon 2000).

More recent studies also take into account information about the survey design, the behavior of the respondent during the interview and the influence of the interviewers. Jenkins et al. (2006) analyzed two different consent questions in a large survey originally based on the British part of the ECHP, the ISMIE ("Improving Survey Measurement of Income and Employment"). The authors' findings confirm that there are differences in demographics between consenters and non-consenters, additionally showing that the interview situation is important for the consent decision. For example, respondents with problems during the previous interview are less likely to consent to the linkage with administrative data, while respondents with a longer interview in the previous wave are more likely to consent. When analyzing two different consent questions (consent to record linkage and consent to contact the employer), the authors find that the influences vary depending on the context of the consent question.

In two recent studies, Sakshaug et al. defined "resistance indicators", that are correlated to the consent decision not

---

both datasets, the remaining non-overlapping variables have to be independent (e.g. Rässler 2002; D'Orazio et al. 2006). The fewer variables overlap, the less likely is the assumption to hold.

just when asking for permission to link the data (Sakshaug et al. 2012), but also when asking for consent to take physical measurements in a survey (Sakshaug et al. 2010). Considering the link of the HRS with administrative data from the SSA, the authors found negative effects corresponding to the number of financial questions the respondents refused to answer both during the current and prior interviews. Respondents who expressed confidentiality concerns in the previous wave are less likely to consent as are those who were rated less cooperative or less attentive by the interviewer. The number of call attempts (current and previous wave) prior to the interview is negatively associated with the consent rate as well.

In addition to these resistance indicators, another important extension of these two studies is the inclusion of interviewer characteristics as well as an estimation of the interviewer level variation. The interviewers' education and experience as an interviewer are both negatively associated with the consent to record linkage. Consent for physical measurements is affected only by the interviewers' race. However, a significant interviewer variance term for both consent questions indicates that additional (not measured) interviewer characteristics are influencing the consent decision.

Sala et al. (2010) investigated the influences on consent in another study based on the British Household Panel Study. Using an interviewer survey, the authors were able to add information on the interviewer level and test its influence on the consent decision for linking survey data to health records and to social security benefit records. Respondents' demographics are not strongly associated with consent, while attitudes toward privacy and community-mindedness seem to be of greater importance. Respondents participating in the panel for a longer time are less likely to give consent, whereas the collected interviewer characteristics (attitudes and personality traits) do not have significant effects. However, the authors found "intra-household dynamics" such that each respondent's decision to consent is "located within the interaction between the individual, the interviewer and the wider household context" (Sala et al. 2010:19).

In Germany, the Institute for Employment Research (IAB) asks respondents for their consent to linkage with administrative records from the German Federal Employment Agency in several surveys. Hartmann and Krug (2009), Beste (2011), and Antoni (2011) analyzed the consent decision in different studies (the so-called "Mainzer Modell", PASS, and ALWA, respectively), reporting the influence of the interviewers. Beyond effects of respondent and interviewer characteristics as well as factors of the interview situation, Antoni (2011) finds a significant interaction effect of respondent and interviewer age: interviewers who are at least 10 years older than their respondents are less successful in obtaining consent.

This overview of studies investigating the determinants of the consent decision shows the growing number of surveys in the social sciences that ask for the respondents' consent to data linkage in various topics. When analyzing determinants of consent, it is important to not only take into account respondent characteristics, but also include indicators of the

interview situation as well as interviewer variables, as they are important parameters of the consent decision. There are some general results: in all studies, respondent characteristics turned out to be significant predictors of consent, thus evidencing the existence of a consent bias. Respondents who are more cooperative are also more likely to consent, while problems during the interview (also in previous interviews) reduce the likelihood of consent. Most studies control for interviewer characteristics to some degree, but the results are not definite.[3] Our study adds to the literature by providing additional evidence on the respondents' selectivity in the consent decision. The findings also stress the importance of the interviewer in obtaining consent. Moreover, we advance previous studies by testing different multi-level models to quantify and explain the interviewer's proportion of the variance. In addition, the consideration of interviewer quality and performance measures may help survey agencies in training and selecting interviewers who will increase consent rates.

3 Data Linkage in SHARE

The Survey of Health, Ageing and Retirement in Europe (SHARE) is a multidisciplinary biannual panel survey that collects micro data on health and socio-economic status as well as social and family networks. The initial face-to-face interviews were conducted in 2004 in eleven European countries. By the fourth wave of data collection in 2010/2011, more than 55,000 respondents from 20 European countries were interviewed. SHARE is based on probability samples in all participating countries, and represents the non-institutionalized population aged 50 and older (see Börsch-Supan and Jürges 2005, for methodological details of the first wave). SHARE's third wave in 2008/2009 (called "SHARELIFE") was different from other waves of SHARE as it was completely retrospective in nature, covering the respondents' lives from birth up to the current interview (for details on SHARELIFE, see Schroeder 2011). The linkage reported in this paper covers the German subsample of SHARELIFE, where a pilot study for other SHARE countries was conducted (for a project overview, see Czaplicki and Korbmacher 2010).

The administrative records of the German Pension Fund ("Deutsche Rentenversicherung", DRV) constitute – for the most part – the universe of all Germans paying into the social security system. People are not included in this database if they have always been self-employed, worked only as civil servants, or have never worked and have not accumulated any social security entitlements through other activities. For all others (nearly 90% of the German population, see Rehfeld and Mika 2006), the data contain monthly information about the respondents' work history beginning at the age of 14. In addition to some basic demographics, detailed information about the employment status (e.g. working, unemployed, in

---

[3] Antoni (2011) provides a nice tabular overview of the literature, which is replicated in this paper with the author's permission in Appendix Table 1.

care, disabled) as well as the personal retirement entitlements is included (Mika and Czaplicki 2010).

All German SHARE respondents were asked for consent to link their survey data with their DRV records. The linking procedure in Germany, conducted via the Social Security Number (SSN), is tied to strict data protection rules. Consent to linking SHARE data with DRV data must be given in written form by each respondent. There are two steps to consent: First, all German SHARE respondents are asked verbally for permission to link their data at the end of the CAPI interview. If the respondent gives her consent, the interviewer provides a consent form to collect the SSN, all information needed to check (and if necessary construct) the SSN, as well as the respondent's signature. The second step is completed by the respondent, who fills out the form and mails it back to the DRV.

Each step presents a hurdle along the way to the data linkage. Conditional on participating in the interview, a respondent may decline consent directly to the interviewer, she may fill in an incorrect SSN, omit it, or may not send in the consent letter at all.[4] The analyses of this paper only consider the initial step of consent, as it is similar to the decisions in many other studies in Germany (e.g. PASS, ALWA, LiDA). Further, as we are interested in the influence of the interviewer on the consent decision, the first step is more appropriate to use than the second. Based on release 1.0.0 of SHARELIFE, 1,350 (73%) of the 1,844 respondents with complete interviews gave their verbal consent to link their survey data to the DRV records, 21% (390) refused consent or claimed "don't know", and the remaining 6% (104) stated that they do not have any entitlements from the German Pension Fund.[5] This leads to a consent rate among the eligible of 77.6% (see also Table 1 below), which is lower than in ALWA (91.6%; Antoni and Seth 2011), but similar to PASS (79.8%; Beste 2011). In the BHPS, where only written consent is asked for, the rates are between 32 and 41% (Sala et al. 2010).

4 Models and Methods

Groves and Couper assume that only a "few householders have strongly preformed decisions about survey requests" (1998:32). The analyses in this paper base on the assumption that the same holds for the consent decision. Asking a respondent to answer survey questions is different from asking her for consent to data linkage. Although the respondent receives information on the data in her administrative records, she may feel insecure about what exactly she is asked to consent to. In addition, there is no possibility for her to release only a portion of the administrative records. Finally, the respondent has no control about what was collected in the administrative records – she may know the contents, but she cannot change them. These characteristics of the consent decision clearly differentiate the consent question from "regular" survey questions. In fact, the decision to give consent may be viewed as being similar to the decision to participate in a survey, where a respondent cooperates without knowing the exact questionnaire.

To model the consent decision, the "conceptual framework for survey cooperation", developed by Groves and Couper (1998, Figure 2.3), is slightly adapted in Figure 1, depicting aspects influencing the respondent's decision to cooperate or refuse when asking for consent. The respondent's consent decision is the result of several influences channeled through three different groups: social environment (such as the household settings) and respondent characteristics (such as age, gender or personality) in the left column, survey design (such as topic, length or mode) and interviewer characteristics (again age, gender or personality) in the right column, and the interaction between respondent and interviewer in the middle column. The consent decision is the result of influences of some areas on others, as depicted by the arrows. The conceptual framework highlights the fact that the interviewer is an important factor in the process, especially because she is "under researcher control".

Unit non-response analysis, a central topic in Groves and Couper (1998), usually lacks sufficient data to test the theoretical hypothesis of what influences participation behavior. SHARELIFE allows for using a full set of control variables from the interview in the third wave as well as from the previous two waves, to investigate the determinants of non-consent. The estimation models in this paper follow the three columns of the framework in Figure 1 in dividing the variables in three areas of influence (described in detail below):

- First, variables describing the respondent, including personal characteristics as well as household and environmental determinants.
- Second, as the respondent-interviewer interaction is not directly measurable, aspects of the interview situation including paradata[6] are used as proxies.
- Third, characteristics of the interviewer including both personal information and aspects related to the interviewer's quality. (Given that the survey design is constant for all respondents and interviewers in SHARE, any influence of the survey design on consent cannot be considered in the analyses.)

The choice of variables characterizing the respondents is mainly motivated by findings in other studies. Although contradictory results are reported for gender, age, and years of education, the influences are shown to be significant. Both age and age-square are included in order to control for any

---

[4] The DRV checks and if necessary corrects the SSN, and if missing, constructs it if all other information on the consent form is available.

[5] The 6% of respondents who claim that they do not have any entitlements at the DRV is lower than the expected 10% from the general population. However, given that SHARE is representative of the population 50+, it seems very likely that in this specific population the number of people not having paid into the social security system is lower.

[6] There is no exact definition of what the term "paradata" includes (Kreuter and Casas-Cordero 2010), but here paradata are defined as all data collected during the process of the interview like timestamps, interviewers' observations as well as the contact protocol.
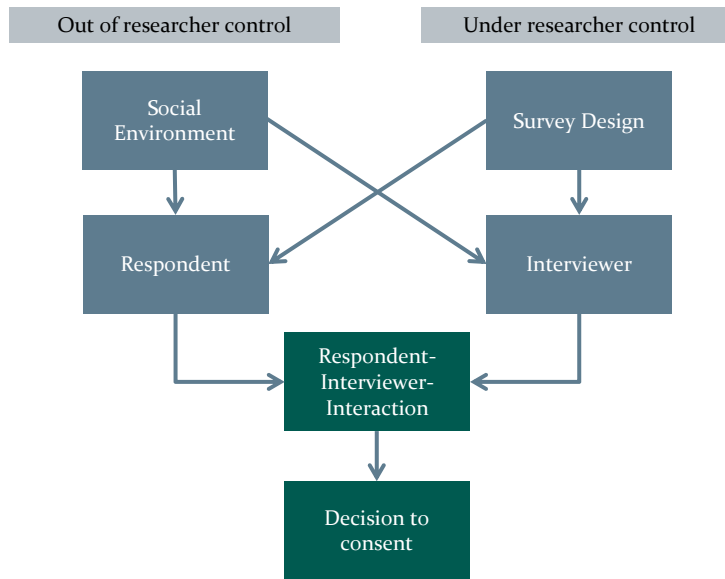
Figure 1. Decision to consent (adapted from figure 2.3 in Groves and Couper 1998)

non-linear relationship. The respondents' partner status influences the consent decision in the majority of studies controlling for it and is measured here with three indicators for currently living with a partner, having ever been married and having ever been divorced. The current work status is influential in three of six studies. The indicator variable used here differentiates between people currently working and all others (retirees, unemployed, housewives). The total number of jobs is added in order to capture the respondents' survey burden in the third wave, as they were asked detailed retrospective questions about each of their past jobs. Respondents with a large number of jobs may see the benefit of consent, as it potentially reduces the time spent on job related questions in the future. Income quartiles of the household's equivalence income (net income divided by the square root of household size), as well as an indicator if household income is not reported are added to capture differences in socio-economic status.[7] An indicator whether the respondent lived in East Germany during communist times is included, as these respondents may feel less confident in their government and hence may be less likely to consent. To control for some environmental influences, information about the household's location (urban or rural), the household's building type (more than two units or a 1- or 2-family home) and whether there is a foreign-born person in the household are used.

To describe the interview situation some variables originating from paradata are included. Whether the interviewer had been in the household during previous waves can play a role, if knowing the interviewer increases the respondent's trust in the confidentiality of the record linkage. In addition, the interviewer's assessment of how well the respondent understood the questions and needed clarifications is used, where a dummy variable with the value 1 is created if the respondent never asked for clarifications and always (to the

interviewer's knowledge) understood the questions. People with problems understanding questions may be less likely to consent because the decision costs are higher. The duration of the interview was found to significantly influence the consent decision in one previous study (Jenkins et al. 2006). To include interview time here, the average time per question was split into a respondent and an interviewer part. The average of how long a respondent takes to answer a question net of the interviewer's average time per question is included to measure the effect of slow or fast respondents independently of the interviewer. Jenkins et al. (2006) found a positive influence of interview time, indicating that using more time (and thus more effort) may be related to more committed respondents. On the other hand, people with more time per question may also be more skeptical and ask more questions, which could reduce the likelihood to consent. Therefore the "net" respondent time per question may influence the consent decision in both directions. The respondent's willingness to answer is measured by the proportion of missing answers (number of missing answers for every hundred questions). This value is split by questions directed at financial values and all other questions, as respondents may view financial questions to be more sensitive and hence more similar to the consent question.[8] To capture possible learning effects,

---

[7]At the beginning of the interview a "financial respondent" is determined to answer all questions about the household's financial situation. Therefore the two variables "equivalent income" and "missing income" are measured on the household level. Note that because the sample is restricted to the first person in each household (see below), using variables at the household level does not introduce a distortion of the variance in these variables.

[8]Differently to Antoni (2011), the missing answer rates combine "refusals" and "don't know"-answers because respondents may use "don't know"-answers to mask a refusal.

the interviewer's experience with SHARELIFE is included, where a simple counter from the first to the last interview the interviewer conducts was transformed into five categories.

The inclusion of all available interviewer characteristics reduces the number of observations (see Table 1 below), because the information on interviewer demographics (gender, age, and education) provided by the survey agency was collected after the data collection was finished. Information on age and education is missing for those interviewers, who had left the agency by the time the information was requested. The age of the interviewer enters as a second order polynomial, while education is provided in three mutually exclusive categories: completed high school (12th/13th grade), the highest degree; left school after 10th grade; and left school after 9th grade, the lowest category, which is selected as the reference group. For gender and education the findings from other studies are contradictory, while interviewer age shows a significant positive effect in most studies.

In addition to these demographics, two variables were constructed from the current wave for each interviewer. The first is an attempt to measure interviewer "quality": SHARE conducts grip strength measures in every wave, where each respondent is required to conduct two measures of their grip strength with each hand. Interviewers are asked to record these measures ranging from 0 to 100, with the explicit instruction not to round these numbers, because previous waves showed that multiples of 5 and 10 were recorded more than statistically expected. Based on the total number of grip strength measures an interviewer conducted, a 90% confidence interval was constructed around the 20.8%, which are expected if no heaping on multiples of 5 and 10 has occurred.[9] If the actual percentage for an interviewer lies outside this interval, the underlying assumption is that the interviewer is still rounding (if above the upper cutoff point of the confidence interval) or that she is overdoing the non-rounding (if below the lower cutoff point).[10] Two dummy variables – one for being above and one for being below the expected cutoffs of the confidence interval are added to the model. An additional variable of interviewer performance is the average time the interviewer took per question over all cases that she interviewed. Although longer does not necessarily mean better in this case, the assumption is that interviewers who have smaller values in this variable are less thorough when reading the question texts. As mentioned before, the inclusion of interviewers' characteristics reduces the number of cases of analysis.

Table 1 gives an overview of the stepwise reduction from the complete into the final sample used in the estimations. The initial sample (SHARELIFE, Release 1.0.0, Germany only) consists of 1,852 interviews, of which eight respondents did not complete the interview. 104 respondents claim that they do not have a record in the DRV data and are therefore excluded. To separate interviewer and "contagion effects" from previous consent decisions within the household, the sample only consists of the first respondent who is asked for consent in a household. Of the 64 interviewers working the sample, 12 (19%) dropped out after the wave, further reducing the sample by 122 (10%) cases. A final reduction by

5 cases is introduced through item-nonresponse. The final sample consists of 1,055 respondents and 51 interviewers, who each interviewed between 7 and 51 respondents in the estimation sample. There are some differences in the consent rates between the cases dropped and the respective remaining sample (shown in parentheses in Table 1), but none are significant.

The dependent variable is verbal consent to record linkage. Interviewers are requested to ask respondents for consent and record their answer in the CAPI program. The dependent variable is 1 if the respondent consents to the linkage of her survey data with administrative records and 0 if she refuses.

To take the dichotomous nature of the dependent variable and the hierarchical structure of the SHARE data into account, a multilevel logistic regression is used. Two different levels are distinguished: the respondents (level 1) who are nested within interviewers (level 2). The model is developed from an empty (or "null") model to the model using all variables. The first set of variables is taken solely from the respondent's side (characteristics of the respondent and of the respondent's household), then variables showing the respondent-interviewer interaction are included, and finally, variables describing the interviewer are added. Via the intra-class correlation (calculated as the proportion of the variance coming from the interviewer level compared to the overall variance) this procedure determines how much of the interviewer's proportion of the total variance can be explained in each step toward the full model.

5 Estimation Results

Table 2 shows the results from the multilevel estimations, depicting the intra-class correlation for all models along with the corresponding χ^2-statistic from a test of the estimated multilevel model against a regular logistic regression. In all cases, the test rejects the simple logistic regression model. In addition, a χ^2-statistic is provided from a likelihood-ratio test of the current model against the one in

---

[9]The underlying assumption regarding the distribution of digits from 0 to 9 is that it is uniform on the grip strength distribution's support from 0 to 100, which means an expectation of eleven "0"s and ten "5"s, i.e. a fraction of 21 of 101 numbers should be "0" or "5".

[10]As an example: Suppose an interviewer has done 25 interviews with 100 grip strength measures. This creates a confidence interval of: $20.8\% \pm 1.645 \sqrt{[20.8\%(1-20.8\%)]}/\sqrt{(100)}=[14.3\%; 27.7\%]$. Thus, if she has 14 or fewer multiples of 5 and 10, the indicator for "too few multiples of 5 and 10" is set to 1. If she has more than 27 multiples of 5 and 10, the indicator for "too many multiples of 5 and 10" is set to 1. More interviews and thus more grip strength measures reduce the confidence interval. The minimum number of conducted grip strength measures of the interviewers used in the analyses is 34, so even though the measure is less precise for interviewers with fewer interviews, the differences in confidence intervals are not large. Note that the underlying assumption is that the standard errors are not clustered on the respondent level (this would likely increase the standard error). Given that the interviewer's measurement is the variable of interest, this assumption is not far-fetched.

Table 1: Sample Size Development, Consent Rates and Number of Interviewers SHARELIFE 1.0.0, Germany

	Cases Dropped (% Consent Rate)	Reduced Sample (% Consent Rate)	Number of Interviewers
Full Sample		1,852	64
Incomplete Interview	8	1,844	64
	(-)	(-)	
Linkage not applicable	104	1,740	64
	(-)	(77.6)	
Second HH respondent	558	1,182	64
	(76.9)	(77.9)	
Missing interviewer information	122	1,060	52
	(85.3)	(77.1)	
Item nonresponse	5	1,055	51
	(60.0)	(77.2)	
Final sample	797	1,055	51

Notes: The table shows the development from the full sample to estimation sample. The consent rate percentages refer rate percentages refer to the dropped or remaining cases only. No consent rates are provided where some or all observations are not eligible for the consent question.

the previous column. As mentioned above and shown in Table 1, the sample is reduced by those cases where the interviewer information is missing. To test for sample selectivity and possible bias, all models are re-estimated using the sample that includes the cases with missing interviewer information.[11] The results of these estimations as well as all other robustness checks are referred to in the text and presented in the appendix.

The first column of Table 2 shows the multilevel model without any explanatory variables. The intra-class correlation of 55.2% provides evidence of a very large interviewer influence on the consent decision. Column 2 of Table 2 shows the model including indicators for fourteen federal states ("Bundesländer") to correct for potential region effects.[12] The intra-class correlation drops slightly to 50.9%, showing that some of the interviewer variation can be attributed to variation at the state level.

The model is then augmented in column 3 by variables that solely depend on the respondent and are not influenced by the interviewer. In this regard, it is not surprising that the intra-class correlation remains almost identical at 50.5%. Including the additional variables is important, as the likelihood-ratio test against the previous model shows. The respondent's age has a significant inversely U-shaped influence on the consent decision. The peak age (from calculating the marginal effects) is at about 65 years, which is right at the official retirement age for the sample under investigation. It is very likely at this age that individuals have obtained most information about their retirement entitlements and the German Pension Fund, while older and younger groups face more uncertainty that reduces the willingness to consent.[13] There are no significant differences between men and women, and neither education nor currently being employed have significant influences on the consent decision. The number of jobs a person had during her working life has a significantly positive influence, which may be related to the survey burden: because detailed information was asked about each of these jobs, individuals with more

jobs may be more likely to see the benefit of record linkage to reduce future survey burden. Partnership and marital status also matter for consent: respondents living with a partner have 88% higher odds to consent, while having ever been divorced has a significantly negative effect of about the same magnitude (calculated as 1/0.495). Ever being married does not show any significant effect. Respondents who have spent some time in East Germany during communist times are much more likely to consent (the odds are increased by 385%). This effect has to be interpreted with the state indicators in mind, which implicitly control for currently living in the East.[14] Almost all who ever lived in the GDR still live in that area (87%, or 258 of 296), so the GDR variable captures the effect of those who moved from the East into the West, showing that this is a selective group compared to those who stayed.

---

[11]Of the 122 cases dropped because of missing interviewer information shown in Table 1, five observations need to be removed because of item non-response. Hence the sample including those cases without interviewer information amounts to 1,172 observations.

[12]Because the states of Bremen and Saarland each only have few observations, they are joined with adjacent states: Bremen and Lower Saxony receive the same state indicator, as do Saarland and Rhineland-Palatinate.

[13]Some of the retirees may also have been in direct contact with the DRV at this age, because the DRV attempts to validate the pension account information directly with the employees ("Kontenklärung") to assure that the pension benefits payments are correct (Rasner 2012).

[14]With the re-unification five states (the so-called "neue Bundesländer") were joined with the former West Germany while keeping the old states unchanged in their boundaries. As a consistency check, the above estimation was conducted with a simple East/West indicator, which shows that the odds of not consenting are increased by 554% for those (still) living in the East. Leaving out the indicator for the change (ever lived in the GDR) shows that currently living in the East increases the odds of not consenting by 90%, although not significantly so (see Appendix Table 2).

Table 2: Multilevel Estimations of the Consent Decision

	(1) Null Modell	(2) State Effects	(3) Respondent Characteristics	(4) Interview Situation	(5) Interviewer	(6) Most parsimonious
Respondent Characteristics						
Age			1.401**	1.342**	1.336*	1.339**
Age^2			0.997***	0.998**	0.998**	0.998**
Female			1.122	1.127	1.133	
Years of Education			1.015	0.998	0.995	
Currently employed			0.820	0.859	0.831	
Number of jobs			1.122*	1.094	1.104	
Lives with Partner			1.875**	1.732**	1.739**	1.814***
Ever married			1.182	1.352	1.281	
Ever divorced			0.495**	0.544**	0.543**	0.606*
Ever lived in GDR			4.849***	4.003**	3.923**	3.913**
Household in urban area			0.636	0.554*	0.465**	0.496**
Household in 1- or 2-family house			1.078	1.096	1.102	
Foreigner in household			0.750	0.821	0.779	
Income is missing			0.247***	0.486*	0.502*	
1st income quartile			0.725	0.784	0.811	
2nd income quartile			0.524*	0.507 *	0.523*	
3rd income quartile			0.728	0.732	0.742	
Interview Situation						
Interviewer is known				0.827	0.767	
Respondent comprehension				1.810**	1.809**	1.828**
Seconds per question (net Interviewer)				1.007	1.014	
Missing rate: financial questions				0.986***	0.987**	0.984***
Missing rate: non-financial questions				0.769*	0.794	0.769*
Interviewer's experience: interview 6-10				0.698	0.712	0.684
Interviewer's experience: interview 11-20				0.715	0.740	0.709
Interviewer's experience: interview 21-50				0.364***	0.348***	0.336***
Interviewer's experience: interview 51+				0.230***	0.223***	0.214***
Interviewer Characteristics						
Interviewer age					0.309***	0.298***
Interviewer age^2					1.011***	1.011***
Interviewer Education: high					4.103*	
Interviewer Education: middle					4.208	
Interviewer is male					1.060	
Average seconds per question (I'wer)					1.167	
Quality: too few multiples of 5 and 10					0.075**	0.093**
Quality: too many multiples of 5 and 10					0.331	0.204**
State ("Bundesländer") fixed effects	No	Yes	Yes	Yes	Yes	Yes
Intra-Class Correlation	0.552	0.509	0.505	0.458	0.352	0.413
χ^2 (2) vs. Logistic Regression	266.72***	139.92***	119.61***	64.28***	47.11***	83.68***
χ^2 of LR-Test against previous model		14.99	74.54***	33.49***	21.33***	14.34
(degrees of freedom; p-value of LR-test)		(13; 0.308)	(17; 0.000)	(9; 0.000)	(8; 0.006)	(17; 0.643)

Notes: *, **, *** mark significance on the 10, 5, 1 per cent level, respectively.
Dependent variable in all models is the dichotomous variable "consent to record linkage". All models are estimated with 1,055 observations in a multilevel logistic regression with Stata's xtlogit command with a random intercept on the interviewer level. Coefficients are odds ratios. χ^2-values are the respective test statistics.
Reference categories: Income: 4th income quartile; Interviewer education: low; Experience: interview 1-5; Quality: rounding within confidence intervals (see text for details).

Except for household income, none of the other variables describing the household situation (urban/rural, building size, foreigner) show any significant influence on the consent decision. Compared to the fourth income quartile as the reference category all income groups have a negative effect on consent, where only the middle group (2nd quartile) shows a significant effect on the 10% level. As was expected, those who refuse to report their income (15% of the sample) are far less likely to consent to linking their data with administrative records.

Including those observations that have missing interviewer information does not change the results (see Appendix Table 3, column 1): even though some of the significant odds ratios change substantially, qualitatively the results

are identical to the model in Table 2's column 3. The test of the indicator for missing interviewer information (bottom in Appendix Table 3) shows that the sample is not selective with regard to respondent characteristics.

Column 4 of Table 2 shows the effects of the interview situation, which describe the interviewer-respondent interaction. The inclusion of these variables leads to a reduction in the intra-class correlation by five percentage points. Also, the likelihood-ratio test confirms that their inclusion is important. Knowing the interviewer from a previous interview is not significantly related to the consent probability. A positive interviewer assessment of the respondent's comprehension increases the odds of consent by 81% compared to those with a negative assessment. The respondent-specific time per question does not play a role in the consent decision, while – as expected – the rate of missing answers is negatively associated with the consent for both financial and non-financial questions. The categories of interviewer's field experience in SHARELIFE show that compared to the first five interviews, the 6th to 10th and 11th to 20th interview is less likely to lead to consent, although not significantly so. From the 21st interview onwards the effect becomes larger and significant. It is likely that the "experience" variables capture two effects working against each other: a "reluctance" effect, which is increasing with the time elapsed in the field work, where respondents who are more reluctant to participate in the survey are also less likely to consent to record linkage. On the other hand, a "learning" effect can be assumed such that the more experience the interviewer has asking the consent question, the more successful she should become. Here the positive learning effect is not larger than the negative reluctance effect for any measured level of experience.

Using the sample of all interviewers leaves the previous results almost unchanged, and the included indicator for missing interviewer information does also not show a significant influence (see Appendix Table 3, column 2). As a further robustness check, the rates of missing values and the interviewer's assessment of the respondent's comprehension are taken from previous waves to counter the possible endogeneity of using the same wave's variables. This reduces the sample by nine observations (0.9%) and leaves the results qualitatively identical, although some coefficients are no longer significant (see column 1 and 2 of Appendix Table 4). In addition, the assessment of how willing the respondent was to answer during the previous wave can be used (see column 3 of Appendix Table 4). The variable shows a highly significant effect on the consent decision for wave 3: a respondent with a high willingness to answer in the previous wave has 184% higher odds of agreeing to the record linkage.[15]

Turning to the explanatory power that the interviewer level variables provide, column 5 of Table 2 shows that not all considered variables turn out to have a significant effect on the consent decision. Overall, their inclusion is warranted (likelihood-ratio test statistic of 21.3 with eight degrees of freedom). Although the intra-class correlation drops by ten percentage points to 35.2%, the model cannot explain all of the interviewer variance. The age of the interviewer influ-

ences consent in a U-shaped way such that older interviewers are more effective in obtaining consent (the turning point calculated from marginal effects is at 55 years). The education of the interviewer affects consent positively, but the estimates for the indicators are not very precise. The interviewer's gender is not significant. Interaction effects between interviewer and respondent education as well as interviewer and respondent gender do not have significant effects on the consent decision (results not shown). The measure of the average time an interviewer needs per question is not significantly related to consent, while the variables on rounding (included to measure interviewer quality) show a negative and significant effect.

When including the cases missing the interviewer information, the variables on education and age of the interviewer must be dropped. Nevertheless, the remaining effects are similarly estimated, although the interviewer gender effect becomes much larger and significant and the negative effect of the rounding is attenuated (see Appendix Table 3, column 3). The ICC increases, showing that interviewer age and education explain part of the variation on the interviewer level. The effect of missing interviewer information, as shown by the indicator and the likelihood-ratio test, is negative, but not significant.

The most parsimonious model fitted to the data is shown in column 6 of Table 2. Included are variables (or groups of variables) if they are significant at least on the 10% level (for groups, the likelihood-ratio test against the model without the group has to be significant on the 10% level). The variables remaining in the model still come from all three areas mentioned in Figure 1 above. The coefficients of those variables remaining are very similar to those in the full model of column 5. Comparing the intra-class correlation across all models in Table 2 shows the importance of the interviewer level variables, as the ICC drops from an initial 55.2% to 41.3% in the final model, a reduction of 25%. Still, the component of the interviewer remains very large in the model, even with the inclusion of variables on that level. This is a strong indication that further unobserved heterogeneity among the interviewers matters in the consent decision.[16]

So far, the results show that interviewers are crucial to obtaining consent. To assess the influences of interviewer performance in the field, some additional variables are now considered, using the same estimation sample. These variables were not included in the analyses before, because they are endogenous to the consent decision to some degree. The considered variables are for each interviewer, (i) the consent decision of the last person visited before coming to the current household; (ii) the consent rate of all previously visited

---

[15]Note that due to endogeneity, the wave 3 version of this question is not used, as the interviewer's assessment comes after the consent question and is thus not independent of the decision. Variables from previous waves are not included as regular variables in the analyses because of the required panel setting that would limit the use of this study in other contexts.

[16]Further specifications of the model with random coefficients for respondent comprehension and the missing rates did not yield significant improvements of the shown model in column 6.

households; (iii) the response rate of all cases previously contacted; and (iv) the overall response rate of this interviewer. The first variable shows the immediate impact of having been successful in the previous household, capturing any boost in motivation to gain consent in the next household. The second variable captures the mix of convincing strategies, perseverance and other interviewer personality traits, which are unobserved but play a role in the consent decision. The third and fourth variables provide a measure of interviewer quality, which is not directly related to the consent decision but may affect interviewer motivation.[17] All of these variables, especially (i) and (ii), are endogenous, as the error term in the consent decision is likely to be correlated to the variables via unobserved interviewer characteristics. Still we believe that they increase the understanding of the problem at hand.

Table 3 shows the odds ratios of the four variables measuring interviewer performance estimated in four different models while leaving the rest of the model identical to column 6 in Table 2. An interviewer's experience in the previous household spills over into the next one visited: If the last decision in the previous household was positive, the odds of obtaining consent from the first person in the current household increase by 106%. Interviewers who were more successful up to the current interview are also more likely to gain consent in the present household: a one-percentage-point higher consent rate up to the current interview yields three percent higher odds of obtaining consent. The inclusion of this variable explains almost all of the interviewer variation in the model, as the intra-class correlation drops to 7.7% and is no longer significant (shown by χ^2-statistic). The consent rate picks up otherwise unobserved variation among interviewers, indicating which interviewers are good at obtaining respondent consent. This may be important information for fieldwork agencies, as this rather obvious relationship (high consent rate equals good interviewers) holds up controlling for a whole set of other variables. Assuming that respondent differences are controlled for, this variable allows survey agencies to identify and react to differences in the interviewers' abilities *during* the fieldwork period.

The response rates for individual interviewers lead to different results: both measures (response rate up to the current interview and total response rate) have a negative correlation with consent, where only the total response rate has a significant influence. The intra-class correlation in the two models remains significant, such that the inclusion of these variables does not explain much in the unknown interviewer variation determining consent. The coefficients could be taken as an overall performance measure for interviewers, where interviewers with a one-percentage-point higher response rate have odds to obtain consent decreased by almost four percent. However, one has to be careful: interviewers with a high response rates will have convinced more respondents than those with a low response rate. This will also include more reluctant respondents, such that gaining consent is more difficult in such a sample.

6 Summary and Discussion

The analyses in this paper provide insights on what determines respondents to consent to a linkage of their survey data to administrative records. Using a theoretical framework adapted from Groves and Cooper (1998), the results show that while some variables at the respondent level are important, the interviewer-respondent interaction and especially the interviewers are a main component in obtaining consent. Using multi-level estimations, the initial proportion the interviewers contribute to the whole variation can be reduced from 55% to 35% by including interviewer level variables such as age, education, and quality indicators. However, a large part of the interviewer variance remains unexplained, which is likely to be related to unobserved interviewer abilities, as additional analyses show.

As far as comparisons are possible, our findings fit well with the existent literature on explaining consent. For most of the respondent variables, the results are similar to the majority of studies, which find significant effects of age, little evidence of a gender or education bias, and a positive effect of being in a relationship. The rate of missing values in financial questions is almost always related to lower consent rates. The effect of the interviewer variables is similar: interviewer gender is not that relevant, while interviewer age is positively related to consent in most studies, which – according to the reported U-shape influence – holds true in the SHARE setting once the interviewer has reached a certain age. The interviewer's experience in the study has a negative effect on consent, which is similar to Sala et al. (2010). The estimated proportion of the interviewer variance is large in our paper, but similar to other studies which estimate it: Beste (2011) finds an intra-class correlation of 28%, from Sakshaug et al. (2010), a value of 34% can be calculated, while Sakshaug et al. (2012) implies an ICC of 32%.[18]

There are some limitations to this study. One is the lack of an interpenetrated sample, which would be necessary to estimate pure interviewer effects (Bailar 1983). The multistage clustered sampling in SHARE does not allow for distinguishing interviewer effects from sampling-point effects, because interviewers are not assigned at random to respondents (for details on sampling in SHARE, see Klevmarken et al. 2005). The inclusion of household and respondent characteristics as well as state fixed effects in the analyses minimizes the influence of sampling-point effects as much as possible. Additionally, two studies show that the interviewer-induced variance is greater than the variance component that comes from the different areas (Schnell and Kreuter 2005; O'Muircheartaigh and Campanelli 1999), so this paper's set-

---

[17]The response rate up to the interview is somewhat imprecise, as it is not clear if the interviewer will not contact a household again and "convert" formerly non-cooperating respondents. Nevertheless, interviewers may still be influenced by the success they had prior the current interview.

[18]Ideally, such a comparison would be done for the respective "empty" models, to compare the original degree of interviewer variation. However, neither of the papers provides such information, hence the full models have to be used for the comparison here.

Table 3: Assessing Interviewer Performance Indicators on the Consent Decision

	Odds Ratio	Intraclass Correlation	n	χ^2
(i) Last person visited before current household HH gave consent	2.060***	0.334	1004	32.0***
(ii) Consent rate when entering current household	1.031***	0.077	1004	1.0
(iii) Response rate when entering current household	0.994	0.395	1023	66.3***
(iv) Total response rate over all assigned cases	0.963*	0.366	1055	39.5***

Notes: *, **, *** mark significance on the 10, 5, 1 per cent level, respectively.

Dependent variable in all models is the dichotomous variable "consent to record linkage". All models are estimated in a multilevel logistic regression with Stata's xtlogit command with a random intercept on the interviewer level. Coefficients are odds ratios. χ^2-values refer to the test statistic against a regular logistic model. Variables in rows (i) to (iv) are added separately to the most parsimonious model in column 6 of Table 2.

Differences in sample size are due to construction of the variables: For variables (i) and (ii), each first observation per interviewer has to be discarded. This is similar for variable (iii), however, in 19 cases the interviewer had previously been in households not in this estimation sample.

ting may also "benefit" from smaller area and larger interviewer effects.

The results lead to the conclusion that there is some consent bias in the sample, as certain respondent characteristics are important determinants of consent. With the focus on the consent bias, this paper addresses only one part of the total survey error, and does not relate it to other sources of error. A comparison with the attrition bias in SHARE's previous waves would in principle be possible, but the setup is not easily adapted to an attrition analysis, because, due to the construction of SHARELIFE, most of the variables used here are not available in the previous waves. Investigating the size and direction of consent bias in relation to attrition bias is clearly an important path for future research. The study by Sakshaug and Kreuter (2012) suggests that non-response biases and measurement errors are generally larger than nonconsent biases, while the direction of the bias is ambiguous.

The attrition process may also have influenced the sample composition, which could question how well the analyses extend to other studies in different contexts or different consent questions. But even though the selection could be problematic on the respondent level, the interviewer's importance for the consent decision is unlikely to vary. The German SHARE sample consists of two parts – those who are interviewed since 2004 and those from a refreshment sample drawn in 2006. Additional tests do not show a difference in consent when considering the respondents' participation time in the SHARE panel.[19]

This paper focuses on "first consenters" in a household and, unlike Sala et al. (2010), does not consider intra-household dynamics. This restriction was applied to avoid contagion effects and measure the "pure" interviewer effect on the first consent decision in a household. Although the SHARE interviewer instructions call for interviews without additional persons present, there could be communication among household members during the first interview that influences the decision to consent. In such a case, the distinction between interviewer and household effects may not have been perfect. However, with the current data available,

any such communication cannot be detected. Future research should thus expand the multi-level approach to the context of intra-household dynamics and investigate how both interact in their effect on the consent decision.

The analyses show that the interviewers are a main source of differences in the consent decision, which highlights the importance of interviewer training in general. Future research should investigate how training could reduce the effect of the interviewer in such a setting. In an ideal world (from a researcher's perspective), all interviewers would be trained such that there are no detectable interviewer effects. As this state will never be achieved, both researchers and survey institutes need information about interviewers' abilities in order to be able to identify important drivers of not only the participation decision but also the consent to link data sources. Future research should thus focus on obtaining this information – possibly through interviewer questionnaires – to use resources more effectively in increasing consent rates, reducing consent bias and improving the overall quality of survey data.

Acknowledgements

---

[19]Specifically, a dummy variable indicating refresher status was added to the model in column 6 of Table 2. The estimated coefficient is positive, but with a p-value of 0.582 far from significant. The intra-class correlation also remains at 41%.

References

Antoni, M. (2011). *Linking survey data with administrative employment data: The case of the ALWA Survey.* Working Paper, accessed July 10th, 2012 at: http://www.norc.org/PDFs/October%202011%20Utilizing%20Administrative%20Data%20Conference/4.%20Antoni%20Linkage_October2011.pdf.

Antoni, M., & Seth, S. (2011). *ALWA-ADIAB – Linked individual survey and administrative data for substantive and methodological research.* FDZ Methodenreport 12/2011, Nuremberg: Institut für Arbeitsmarkt- und Berufsforschung.

Bailar, B. A. (1983). Interpenetrating subsamples. In N. L. Johnson & S. Kotz (Eds.), *Encyclopedia of Statistical Science* (p. 197-201). New York: Wiley.

Beste, J. (2011). *Selektivitätsprozesse bei der Verknüpfung von Befragungs- mit Prozessdaten. Record Linkage mit Daten des Panels "Arbeitsmarkt und soziale Sicherung" und administrativen Daten der Bundesagentur für Arbeit.* FDZ Methodenreport 09/2011, Nuremberg: Institut für Arbeitsmarkt- und Berufsforschung.

Börsch-Supan, A., & Jürges, H. (Eds.). (2005). *The Survey of Health, Ageing and Retirement in Europe – Methodology.* Mannheim: Mannheim Research Institute for the Economics of Aging (MEA).

Calderwood, L., & Lessof, C. (2009). Enhancing Longitudinal Surveys by Linking to Administrative Data. In P. Lynn (Ed.), *Methodology of Longitudinal Surveys.* New York: John Wiley & Sons.

Czaplicki, C., & Korbmacher, J. (2010). *SHARE-RV: Verknüpfung von Befragungsdaten des Survey of Health, Ageing and Retirement in Europe mit administrativen Daten der Rentenversicherung.* DRV-Schriften Band 55/2010.

D'Orazio, M., Di Zio, M., & Scanu, M. (2006). *Statistical matching: theory and practice* (M. D. Zio & M. Scanu, Eds.). Chichester, West Sussex, England; Hoboken, NJ: John Wiley & Sons.

Dunn, K. M., Jordan, K., Lacey, R. J., Shapley, M., & Jinks, C. (2004). Patterns of Consent in Epidemiologic Research: Evidence from Over 25,000 Respondents. *American Journal of Epidemiology*, *159*(11), 1087-1094.

Gramlich, T., Bachteler, T., Schimpl-Neimanns, B., & Schnell, R. (2010). Panelerhebungen der amtlichen Statistik als Datenquelle für die Wirtschafts- und Sozialwissenschaften. *Wirtschafts- und Sozialwissenschaftliches Archiv*, *4*(3), 153-183.

Groen, J. A. (2012). Sources of error in survey and administrative data: The importance of reporting procedures. *Journal of Official Statistics*, *28*, 173-198.

Groves, R. M., & Couper, M. P. (1998). *Nonresponse in Household Interview Surveys.* New York: John Wiley & Sons.

Gustman, A. L., & Steinmeier, T. L. (1999). *What People Don't Know About Their Pensions and Social Security: An Analysis Using Linked Data from the Health and Retirement Study.* NBER Working Paper No. 7368.

Haider, S. J., & Solon, G. (2000). *Nonrandom Selection in the HRS Social Security Earnings Sample.* RAND Labor and Population Program Working Paper Series 00-01, DRU-2254-NIA.

Hartmann, J., & Krug, G. (2009). Verknüpfung von personenbezogenen Prozess- und Befragungsdaten – Selektivität durch fehlende Zustimmung der Befragten? *Zeitschrift für ArbeitsmarktForschung*, *42*, 121-139.

Huang, N., Shih, S.-F., Chang, H.-Y., & Chou, Y.-J. (2007). Record linkage research and informed consent: who consents? *BMC Health Services Research*, *7*(18). Available from http://www.biomedcentral.com/1472-6963/7/18

Jäckle, A., Emanuela, S., Jenkins, S. P., & Lynn, P. (2004). *Validation of Survey Data on Income and Employment: The ISMIE Experience.* Working Papers of the Institute for Social and Economic Research, 2004-14. Colchester: University of Essex.

Jenkins, S. P., Cappellari, L., Lynn, P., Jäckle, A., & Sala, E. (2006). Patterns of consent: evidence from a general household survey. *Journal of the Royal Statistical Society*, *Series A, 169*(4), 701-722.

Kho, M. E., Duffett, M., Willison, D. J., Cook, D. J., & Brouwers, M. C. (2009). Written informed consent and selection bias in observational studies using medical records: systematic review. *British Medical Journal*, *338:b866*. Available from http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2769263/

Klevmarken, A., Hesselius, P., & Swensson, B. (2005). The SHARE sampling procedures and calibrated designs weights. In A. Börsch-Supan (Ed.), *The Survey of Health, Aging, and Retirement in Europe – Methodology* (p. 28-69). Mannheim: MEA.

Kreuter, F., & Casas-Cordero, C. (2010). *Paradata.* RatSWD Workingpaper Series, No. 136, Berlin.

Lessof, C. (2009). Ethical issues in longitudinal surveys. In P. Lynn (Ed.), *Methodology of Longitudinal Surveys.* New York: John Wiley & Sons.

Lessof, C., Banks, J., Taylor, R., Cox, K., & Philo, D. (2004). *Linking survey and administrative data in the English Longitudinal Study of Ageing.* Talk presented at ESRC Research Methods Programme Seminar: Linking Survey Response and Administrative Records, London. Retrieved December, 12, 2011, from http://www.ccsr.ac.uk/methods/events/linkage/Lessof.pdf.

Mika, T., & Czaplicki, C. (2010). SHARE-RV: Eine Datengrundlage für Analysen zu Alterssicherung, Gesundheit und Familie auf der Basis des Survey of Health, Ageing and Retirement in Europe und der Daten der Deutschen Rentenversicherung. *RVaktuell*, *12*, 396-401.

Olson, J. A. (1999). Linkages with Data from Social Security Administrative Records in the Health and Retirement Study. *Social Security Bulletin*, *62*(2), 73-85.

O'Muircheartaigh, C., & Campanelli, P. (1999). A Multilevel Exploration of the Role of Interviewers in Survey Non-Response. *Journal of the Royal Statistical Society*, *162*(3), 437-446. (Series A)

Rasner, A. (2012). *The distribution of pension wealth and the process of pension building – Augmenting survey data with administrative pension records by statistical matching.* Dissertation, Technische Universität, Berlin, Germany. http://opus.kobv.de/tuberlin/volltexte/2012/3384/.

Rässler, S. (2002). *Statistical Matching: A Frequentist Theory, Practical Applications, and Alternative Bayesian Approaches.* New York: Springer.

Rehfeld, U. G., & Mika, M. (2006). The Research Data Centre of the German Statutory Pension Insurance (FDZ-RV). *Schmollers Jahrbuch*, *126*, 121-127.

Sakshaug, J. W., Couper, M. P., & Ofstedal, M. B. (2010). Characteristics of Physical Measurement Consent in a Population-Based Survey of Older Adults. *Medical Care*, *48*(1), 64-71.

Sakshaug, J. W., Couper, M. P., Ofstedal, M. B., & Weir, D. (2012).

Linking survey and administrative records: Mechanisms of consent. *Sociological Methods & Research*, *41*(4), 535-569.

Sakshaug, J. W., & Kreuter, F. (2012). Assessing the Magnitude of Non-Consent Biases in Linked Survey and Administrative Data. *Survey Research Methods*, *6*(2), 113-122.

Sala, E., Burton, J., & Knies, G. (2010). *Correlates of Obtaining Informed Consent to Data Linkage: Respondent, Interview and Interviewer Characteristics.* Working Papers of the Institue for Social and Economic Research, 2010-28. Colchester: University of Essex.

Schnell, R. (2012). *Survey Interviews. Methoden standardisierter Befragungen*. Wiesbaden: VS Verlag.

Schnell, R., & Kreuter, F. (2005). Separating Interviewer and Sampling-Point Effects. *Journal of Official Statistics*, *21*(3), 389-410.

Schroeder, M. (2011). *Retrospective Data Collection in the Survey of Health, Ageing and Retirement in Europe. SHARELIFE Methodology*. Mannheim: MEA.

Tisch, L., & Tophoven, S. (2011). *The lidA Project.* Paper presented at the 4th Workshop of Panel Surveys in Germany. Nuremberg.

Trappmann, M., Christoph, B., Achatz, J., Wenzig, C., Mueller, G., & Gebhardt, D. (2009). *Design and Stratification of PASS: a New Panel Study for Research on Long-term Unemployment.* IAB Discussion paper 5/2009, Nuremberg: Institut für Arbeitsmarkt- und Berufsforschung.

Woolf, S. H., Rothemich, S. F., Johnson, R. E., & Marsland, D. W. (2000). Selections Bias from Requiring Patients to Give Consent to Examine Data for Health Services Research. *Archives of Family Medicine*, *9*, 1111-1118.

Appendix

Table 1: Tabular Literature Overview (replicated from Antoni, 2011, with permission of the author)

	Beste (2011)	Gustman and Steinmeier (1999)	Haider and Solon (2000)	Hartmann and Krug (2009)	Jenkins et al. (2006)	Olson (1999)	Sakshaug and Kreuter (2011)	Sala et al. (2010)	Singer et al. (2010)
Respondent									
Male	ns	ns	ns	+	ns	ns		+	ns
Foreign, ethnic minority	-	-	-	-		-	-	-	-
Native language									
Region of residence	ns			sig	sig	ns		sig	ns
Age	ns			ns	sig	ns	+	-	+
Qualification	ns	-	ns	ns	ns	-		+	ns
Cognitive skills									
Labor market status	ns		sig	ns	sig	ns	ns		
Income	+		+	+	ns	+		ns	+
Refused income information	-	-		-	ns			-	
Wealth, assets		-	-			-			
Existing relationship/marriage		+		ns	+	+		ns	
Children		ns			+			ns	
Interviewer									
Male	+			ns				ns	
Age	+			+				ns	
Qualification	-			+				ns	
Experience before study								ns	
Prior interviews within actual study	ns							-	
Similarity of respondent and interviewer									
Sex					ns				
Age					ns				
Qualification	ns								
Interview situation									
Weekday/time of interview									
Share of refused answers									
Share of answers like "don't know"									
Duration of interview				ns	+				
Disturbances/problems during interview					-				
Cooperation in other consent questions		+							

Notes: +/-/ns/sig denote significantly positive/significantly negative/no significant/overall significant influence on consent, respectively.
Note: "Sakshaug and Kreuter (2011)" refers to an earlier version of Sakshaug and Kreuter (2012).

Table 2: Multilevel Estimation of the Consent Decision: Using an East/West Indicator instead of State Indicators

	with GDR indicator	without GDR indicator
Age	1.340**	1.353**
Age^2	0.998**	0.998**
Female	1.091	1.115
Years of Education	1.012	1.012
Currently employed	0.799	0.817
Number of jobs	1.118*	1.134**
Lives with Partner	1.871**	1.919***
Ever married	1.234	1.195
Ever divorced	0.535**	0.585**
Ever lived in GDR	4.327***	
Household in urban area	0.686	0.661
Household in 1- or 2-family house	1.068	1.058
Foreigner in household	0.725	0.721
Income is missing	0.236***	0.242***
1st income quartile	0.701	0.703
2nd income quartile	0.512*	0.521*
3rd income quartile	0.712	0.731
Living in East	0.153**	0.526
Intra-Class Correlation	0.534	0.538

Notes: *, **, *** mark significance on the 10, 5, 1 per cent level, respectively.

Dependent variable in both models is the dichotomous variable "consent to record linkage".

Both models are estimated with 1,055 observations in a multilevel logistic regression with Stata's xtlogit command with a random intercept on the interviewer level. Coefficients are odds ratios. χ^2-values are the respective test statistics.

Reference category: Income: 4th income quartile.

Table 3: Multilevel Estimation of the Consent Decision: Testing the Influence of Missing Interviewer Information

	Respondent Characteristics	Interview Situation	Interviewer Characteristics
Age	1.261*	1.216	1.211
Age^2	0.998**	0.998	0.999
Female	1.059	1.077	1.093
Years of Education	1.023	1.010	1.007
Currently employed	0.794	0.827	0.820
Number of jobs	1.128**	1.103*	1.111*
Lives with Partner	1.797**	1.642**	1.649**
Ever married	1.147	1.283	1.290
Ever divorced	0.534**	0.569**	0.556**
Ever lived in GDR	4.800***	4.002***	3.917**
Household in urban area	0.637	0.554*	0.518**
Household in 1- or 2-family house	1.008	1.038	1.022
Foreigner in household	0.579	0.659	0.650
Income is missing	0.250***	0.440**	0.438**
1st income quartile	0.635	0.665	0.681
2nd income quartile	0.567*	0.535*	0.553*
3rd income quartile	0.766	0.761	0.767
Interviewer is known		0.877	0.882
Respondent comprehension		1.625**	1.642**
Seconds per question (net Interviewer)		1.006	1.010
Missing rate: financial questions		0.989**	0.989**
Missing rate: non-financial questions		0.751**	0.760 **
Interviewer's experience: interview 6-10		0.821	0.829
Interviewer's experience: interview 11-20		0.581*	0.580
Interviewer's experience: interview 21-50		0.354***	0.362***
Interviewer's experience: interview 51+		0.226***	0.225***
Interviewer is male			2.535*
Average seconds per question (Interviewer)			1.047
Quality: too few multiples of 5 and 10			0.186*
Quality: too many multiples of 5 and 10			0.487
Interviewer information missing	1.038	0.693	0.380
ICC	0.467	0.439	0.404
χ^2 (1) of LR-Test for interviewer information	0.003	0.230	1.520
Observations	1,172	1,172	1,172

Notes: *, **, *** mark significance on the 10, 5, 1 per cent level, respectively.
Dependent variable in all models is the dichotomous variable "consent to record linkage". All models are estimated with a multilevel logistic regression with Stata's xtlogit command with a random intercept on the interviewer level. All estimations include state fixed effects. The χ^2-values refer to the test statistics from a test of two nested models including the indicator for missing interviewer information.
Reference categories: Income: 4th income quartile; Experience: interview 1-5; Quality: rounding within confidence intervals (see text for details).

Table 4: Multilevel Estimation of the Consent Decision: Including Previous Wave Information on the Interview Situation

	Basic Model (similar Table 2, Column 4)	Check 1: previous waves	Check 2: previous waves
Interviewer is known	0.866	0.808	0.692
Respondent comprehension	1.848**		
Respondent comprehension (w1/w2)		1.265	0.973
Seconds per question (net Interviewer)	1.007	1.029	1.027
Missing rate: financial questions	0.986**		
Missing rate: non-financial questions	0.768*		
Missing rate: financial questions (w1/w2)		0.977***	0.977***
Missing rate: non-financial questions (w1/w2)		0.893	0.920
Respondent willingness to answer (w1/w2)			2.838***
Interviewer's experience: interview 6-10	0.678	0.765	0.729
Interviewer's experience: interview 11-20	0.712	0.845	0.881
Interviewer's experience: interview 21-50	0.368***	0.447**	0.442**
Interviewer's experience: interview 51+	0.233***	0.278***	0.260***
ICC	0.458	0.512	0.499
Observations	1,046	1,046	1,046

Notes: *, **, *** mark significance on the 10, 5, 1 per cent level, respectively.
Dependent variable in all models is the dichotomous variable "consent to record linkage". All models are estimated in a multilevel logistic regression with Stata's xtlogit command with a random intercept on the interviewer level. All estimations include state fixed effects and all variables on the respondent level (see Table 2, column 3). The coefficients represent odds ratios. "w1/w2" refers to data coming from previous waves of SHARE: from wave 2, if they were available there or otherwise from wave 1.
Reference categories: Experience: interview 1-5.