

Assessing the Magnitude of Non-Consent Biases in Linked Survey and Administrative Data

Joseph W. Sakshaug

Institute for Employment Research, Nuremberg
Ludwig Maximilians University of Munich

Frauke Kreuter

University of Maryland, College Park
Institute for Employment Research, Nuremberg
Ludwig Maximilians University of Munich

Administrative records are increasingly being linked to survey records to heighten the utility of the survey data. Respondent consent is usually needed to perform exact record linkage; however, not all respondents agree to this request and several studies have found significant differences between consenting and non-consenting respondents on the survey variables. To the extent that these survey variables are related to variables in the administrative data, the resulting administrative estimates can be biased due to non-consent. Estimating non-consent biases for linked administrative estimates is complicated by the fact that administrative records are typically not available for the non-consenting respondents. The present study can overcome this limitation by utilizing a unique data source, the German Panel Study “Labour Market and Social Security” (PASS), and linking the consent indicator to the administrative records (available for the entire sample). This situation permits the estimation of non-consent biases for administrative variables and avoids the need to link the survey responses. The impact of non-consent bias can be assessed relative to other sources of bias (nonresponse, measurement) for several administrative estimates. The results show that non-consent biases are present for few estimates, but are generally small relative to other sources of bias.

Keywords: administrative data linkage, non-consent bias, nonresponse, measurement error, data quality, record-linkage

1 Introduction

Survey organizations are constantly faced with the struggle of producing high quality survey data while keeping costs low. The demand for high quality survey data is ever increasing, but survey budgets do not always keep pace with the growing demands and increasing costs of data collections. The problem is heightened by methodological challenges, such as falling response rates (De Heer 1999; Curtin et al. 2005), noncoverage of the target population (Blumberg and Luke 2007; Eckman and Kreuter 2011), and measurement errors (Kreuter et al. 2008), which all contribute to survey cost and data quality concerns.

These challenges have pushed surveys to consider alternative data collection technologies that are less proven methodologically, but may offer significant cost savings in the long run. For example, administrative records have received substantial interest as a complement (or alternative) to traditional survey data collection (Calderwood and Lessof 2009). The increasing interest in administrative records has been noted by the Director of the United States Census Bureau, Robert M. Groves, who wrote in a recent blog post titled, “The Future of Producing Social and Economic Statistical Information, Part I”¹:

“The world is now producing large amounts of data without active participation of persons (e.g., data from Internet searches, credit card transactions, retail scanners, and social media). There also are more and more digital administrative data (e.g., tax records, social security records, Medicare/Medicaid records, food stamp records, HUD records). [...] While few of our surveys have used such data, many of our surveys are discovering that multiple modes of data collection (e.g., paper forms, internet, telephone interviews, face-to-face interviewers), employed in one survey, can address some of their participation problems within current budgets (with administrative records considered a “mode”). Indeed, there is a consensus among survey methodologists that multi-mode surveys will be a key component of the future of statistical information.”

The act of combining survey and administrative databases is not without challenges. The quality of linked databases can be adversely affected by a) inconsistencies between information collected from respondents in the survey and the information contained in administrative databases and b) errors in the record linkage process itself (Smith

Contact information: Joseph W. Sakshaug, Institute for Employment Research, Nuremberg, Germany, e-mail: joe.sakshaug@iab.de

¹ blogs.census.gov/directorsblog/2011/09/the-future-of-producing-social-and-economic-statistical-information-part-i.html

2011). Such errors can be due to erroneous linkage, imprecise matching variables, or the lack of respondent consent to link their survey responses to administrative data. It is often the case that the impact of these issues on the quality of the linked data sets is unknown. Evaluations of the procedures and criteria used in combining survey and administrative databases and assessments of their impact on data quality are clearly needed if linked administrative databases are to become a key source of statistical information in the future.

To date several studies have examined inconsistencies between survey and administrative data. In some instances administrative records have been used for the purpose of validating survey responses under the assumption that the records data are error-free and/or a suitable 'gold standard' against which survey responses can be compared (Davern et al. 2008; Kreuter et al. 2008, 2010; Sakshaug et al. 2010a). Other studies, particularly in the health survey literature, have questioned this assumption and made attempts to evaluate the quality of administrative data, in some cases finding that survey responses are more accurate than their administrative counterparts (Fowles et al. 1998; Hebert et al. 1999; Keating et al. 2003; Losina et al. 2003; Sakshaug et al. 2010b).

While research on identifying suitable linkage variables, efficient linkage algorithms, and statistical matching techniques is well developed (Scheuren and Winkler 1993; Lahiri and Larsen 2005; Winkler 2006; Chambers 2009; Schnell et al. 2009), studying the impact of linkage non-consent has received less attention. Prior to linkage, informed consent is usually needed to ensure that respondents are aware of the risks and benefits involved in releasing their administrative records for research purposes. Consent rates to linkage requests vary widely from study-to-study, study population and country, and across the social, economic, and health fields; with percentages ranging anywhere from the mid-20's to the high-80's (see Table 1).

Studies that examined correlates of linkage consent found by and large gender, age, education, and wealth to be strongly related to the likelihood of consent, and found health status, income item nonresponse, and prior-wave observations in longitudinal studies to also correlate with linkage consent (Tate and Calderwood 2005; Jenkins et al. 2006; Sala et al. 2010). The studies cited in Table 1 rely solely on survey data or other available information on respondents to find characteristics that are correlated with non-consent and may be indicative of non-consent bias in the administrative data. There are only a few studies in the medical field where administrative data were used to assess non-consent bias (Buckley et al. 2007; Kho et al. 2009).

Monitoring linkage consent rates and studying linkage propensities is valuable, but given that the characteristics found to be related to linkage propensity (see Table 1) are likely to be related to key outcomes in administrative data sets, one also needs to be concerned about non-consent biases in the administrative data. Studies focusing on rates have discussed this issue, but outside of the medical field they have not estimated consent biases on the administrative

data linked to the survey data. The current focus on rates over bias is reminiscent of the early nonresponse literature where the study of nonresponse rates and mechanisms of survey participation dominated the field for a long time before researchers started to systematically assess nonresponse bias (the latter was done in, for example, Schnell (2007); Curtin et al. (2000); Merkle and Edelman (2002); Groves (2006)).

Shifting the focus from consent rates to non-consent bias in linked survey and administrative data sets is needed to 1) validate suspicions of bias inferred from consent rate studies, 2) evaluate their possible impact on estimates obtained from linked data sets, and 3) determine whether data linkage yields fewer biases compared to asking respondents to self-report their administrative information. A critical barrier to estimating non-consent biases in linked survey and administrative studies is obtaining and analyzing the administrative records for both consenting *and* non-consenting respondents. Administrative records are typically not released for the non-consenting cases, which restricts the type of bias analysis that one can perform. This study overcomes this limitation by utilizing a unique data source that allows for the estimation of non-consent biases in a German Panel Study. Here administrative records are available for the entire respondent pool, including those who did not give linkage consent. Using indirect estimation procedures, estimates of non-consent, nonresponse, and measurement error biases can be obtained jointly without linking the survey data to the administrative data, but instead linking the consent indicator; thus, preserving the confidentiality and data protections. With this the paper address the following research questions:

1. How large are non-consent biases for key administrative estimates in linked data sets?
2. What is the contribution of non-consent error to the overall error relative to more traditional sources of error (nonresponse, measurement error)?
3. Given that non-consent could be seen as a form of refusal, is there a relationship between nonresponse and non-consent?

The answers to these questions have practical implications: If non-consent error is larger than measurement error than it might be beneficial to ask respondents to self-report the information that would otherwise be obtained from administrative records. If reluctant respondents show higher non-consent bias, survey designers might consider specific interventions to convince those to give their linkage consent as well.

2 Data

The data used for this study comprise survey data matched with personal data generated in labour administration and social security data processing.

2.1 Survey Data

The survey data come from the German study "Labour Market and Social Security" (PASS), conducted by the German Institute for Employment Research (Trappmann et al. 2010). PASS is an annual dual-frame mixed-mode (CATI

Table 1: List of surveys linking survey records to administrative data sources. Table adapted from Sakshaug et al. (2012) and expanded with additional studies, predictors of consent, and direction of effects.

Survey	Citation	Type of Admin Data	Response Rate (%)	Consent Rate (%)	Predictors of Consent
<i>Cross-sectional</i>					
Questionnaire Design and Evaluation Research Survey (USA)	Bates (2005)	Government records	42.4 (AAPOR RR2) 59.2 (AAPOR RR6)	Experiment 40.5 (SSN) 24.0 (SSN+expl) 50.6 (Last 4 SSN) 63.4 (Opt-out)	Male(+), education(-), age(-), non-White(-), income(-), income item refusal(-)
National Health Interview Survey Field Test (USA)	Dahlhamer and Cox (2007)	Government health records	68.0 (adult) 77.0 (child)	1) 66.2 (adult) 2) 60.7 (child)	1) Age(-), # health conditions(+), income item nonresponse(-), region-Midwest(+), non-MSA status(+), pre-interview concerns(-) 2) Age(-), # health conditions(+), income item refusal(-), region-Midwest(+), non-MSA status(+)
National Population Health Survey (Canada)	Finkelstein (2001)	Health insurance records	88.0	89.0	No differences based on socio-demographic characteristics
Office Worker Survey (Australia)	Silva et al. (2002)	Healthcare utilization records	53.0	60.0	Subgroup analysis not performed
<i>Panel</i>					
English Longitudinal Study of Ageing (Wave 1; UK)	Banks et al. (2005)	1) Hospital records; 2) Benefit and tax records	68.3	1) 81.5 (hospital) 2) 77.8 (benefits)	1) Male(+), age(-), income(+) 2) Age(-), income(+), wealth(-)
Health and Retirement Study (1992; USA)	Haider and Solon (2000)	Social Security records	81.7	74.9 (matched)	Male(-), earnings(+), wealth(-), poor health(-), unemployed(-), non-whites(-)
Improving Survey Measurement of Income and Employment Household Survey (UK)	Jenkins et al. (2006)	1) Tax credit and benefit records; 2) Employer records	89.0	1) 77.4 (benefits) 2) 58.5 (employer)	1) Age(+), coupled household(+), benefit reciprocity(+), problems w/ interview(-) 2) Income item nonresponse(-)
British Household Panel Survey (UK)	Knies et al. (2012)	Health data and NHS registration records	84.2	41.0	Living in England(+), age(-), male(+), British/Irish White(+), education(+), diabetes(+), used health services(+), general practitioner visits(-), hospitalizations incl. childbirth(+)

Table 1: Continued.

Survey	Citation	Type of Admin Data	Response Rate (%)	Consent Rate (%)	Predictors of Consent
<i>Panel</i>					
Health and Retirement Study (2008; USA)	Sakshaug et al. (2012)	Social Security records	88.4	67.8	Education(+), net worth(-), government income receipt(+), confidentiality concerns(-), interview resistance(-), wave nonresponse(-)
Australian Longitudinal Study on Women's Health	Young et al. (2001)	Medicare records	43.9	49.4	Age(+), education(+), private health insurance(+), mortality rate(-),

Notes: All predictors reported in this table are statistically significant at the 0.05 level. In Bates (2005) the experimental treatment group SSN refers to those asked for their full social security number without explanation for linking records; SSN+expl refers to those asked for their full social security number with explanation for linking records; Last 4 SSN refers to those asked for the last 4 digits of their social security number with explanation for linking records; Opt-out refers to those asked if they objected to record linkage (no SSN requested).

and CAPI) household panel survey that asks detailed questions about labour market, welfare state, and poverty characteristics in Germany. The first wave of PASS data collection occurred between the months of December 2006 and July 2007. A total of 49,052 households were sampled across both frames within the same primary sampling units using a two-stage stratified cluster design. About half ($n = 25,316$) were selected from a residential database and comprise a representative cross-section of the German population. The other half ($n = 23,736$) were sampled directly from administrative data records housed at the Federal Employment Agency, and are households that were receiving means-tested unemployment benefits at the time the sample was drawn.

At the onset of data collection, households where a telephone number was available were contacted by CATI. A CAPI interview was attempted for the remaining sample. A total of 12,794 households was contacted and interviewed across the benefit recipient and general population subsamples, yielding a response rate of 26.7 percent (28.7 percent for the recipient sample and 24.7 percent for the population sample; RR1 according to AAPOR standards²). An initial interview was carried out with heads of household, and a person-level questionnaire was administered thereafter for each household member aged 15 years and older. Further details of the PASS study design can be found in Trappmann et al. (2010) and Christoph et al. (2008).

2.2 Administrative Data

The administrative data come from the notification process of the German social security system and data generated in administrative processes at the Federal Employment Agency. These include employment spells for all employment subject to social security, benefit reciprocity, periods of job searching, and participation in employment and training measures. Information on status variables is available to the

day. An integrated file of these administrative data is housed at the German Institute for Employment Research.³ Studies have found these data to be highly reliable, in particular the data on employment spells, income, and benefit reciprocity (Bender and Haas 2002; Jacobebbinghaus and Seth 2007).

2.3 Selected Items of Interest

Our analyses focus on six variables that are recorded in the PASS survey and the administrative data. The selected items include an indicator variable of current Unemployment Benefit II (UB II) receipt, the respondent's current employment status, whether they have a registered disability, the respondent's income in the last calendar month (conditional on being employed), and two demographic variables – age and foreign citizenship status. With the exception of disability status and income, creating comparable measures between survey and administrative data is straightforward.

Disability status is somewhat subjective, although respondents were asked if they have any officially recognized disability, which is more precise than asking for any disabilities. All employed respondents were asked about their gross income in the previous month. In the administrative data, gross income in the previous month is not readily available and had to be generated for the purpose of this study. The income information in the register data reflects total income earned among the subset of employed individuals in the mandated notification period (usually one year). For the purpose of this analysis this total was averaged over the number of

² AAPOR Standard Definitions can be found at the shortened URL: <http://goo.gl/iCwFw>

³ A scientific use file of these data is available through the Research Data Center at the German Institute for Employment Research. For details of the "Integrated Employment Biographies" (http://fdz.iab.de/en/FDZ_Individual_Data/Integrated_Employment_Biographies.aspx)

days falling into the notification period and reported in the data as income on a daily basis. When generating the measure of gross income for comparison of the survey data, the sum over all days falling into the calendar month prior to the interview was taken. The lack of full comparability between survey and administrative income is not expected to affect the analysis of non-consent and nonresponse bias, as these two bias estimates rely solely on the administrative data. However, calculating measurement error for income in this way is likely to produce larger estimates than calculating measurement error from direct linkage where other information from the survey can be used to fine-tune the comparability of the two sources.

2.4 Analytic Sample

For the analyses in this paper, it is necessary to have good frame data for both nonresponding cases and responding cases, and among responders for those consenting to have their data linked to the administrative data and those that refused. Thus a few restrictions had to be made to create an analytic data set that meets those needs. First, the analytic sample was restricted to the benefit recipient subsample. Only for this subsample are frame data available for *all* sample units. To ensure a good match between the survey and administrative data on household-level information the sample was further restricted to persons designated as “head of the household” in the administrative data.⁴ Finally, the sample was restricted to cases pursued only by CATI. The CATI-only cases contained detailed paradata for every call attempt including the date of the last contact attempt. This information was needed to construct the administrative variables and ensure that the observation period of these variables was comparable to those collected in the survey. These restrictions yielded an overall sample size of 17,167 administrative records and 4,513 respondent survey records, among which 3,538 gave consent to data linkage. For the analyses below, including the joint assessment of non-consent, nonresponse, and measurement error bias, the administrative data are assumed to be error free.

2.5 Linkage Consent

Early in the person-level interview (after about 22 questions), interviewers asked respondents for verbal consent to link their survey data with their corresponding administrative data. The consent question was read as follows (English translation⁵):

[P23a] “To keep the interview as brief as possible, the Institute for Employment Research in Nuremberg could merge the study results with data about your times of employment, unemployment or participation in measures by the employment office (Arbeitsamt). For the results of this study it would be a great advantage.

For reasons of data protection this cannot be done without your agreement, which I kindly

ask you to provide. This is of course just as voluntary as the interview you are so kind as to give us. Of course, you may withdraw your consent at any time. It goes without saying that all rules of data protection and of the de-personalization of the results reported apply to these additional data as well.

So may I write down your answer: Do you agree to the use of this additional data?”

PASS yields a fairly high consent rate of about 79.8 percent for all respondent, 79.9 percent in the recipient sample, and 78.4 percent in the analysis file used here. Despite the high numbers, the risk of non-consent bias is still present. For the assessment of non-consent, nonresponse, and measurement error, it is not possible to link the survey and administrative data for the non-consenting cases. However, this problem can be circumvented by linking the consent indicator from the PASS survey to the administrative data. The linkage consent indicator is seen as part of the data production process variables and not as substantive survey responses and therefore can be linked to the administrative data. Approval of this procedure was obtained by the legal team of the Institute for Employment Research.

3 Methods

3.1 Estimation of Non-Consent, Nonresponse, and Measurement Error Bias

Each household has a unique key on the sampling frame and in the administrative data that enables us to link the consent indicator and identify whether an administrative record belongs to a consenting or non-consenting respondent. An estimate of non-consent bias for an administrative statistic of interest is then obtained by computing the difference between the statistic based on the consenting respondents and the statistic based on the full set of respondents:

$$\text{Non-Consent Bias } (\bar{y}_{\text{ADMIN}}) = \bar{y}_{\text{ADMIN, Consent}} - \bar{y}_{\text{ADMIN, Resps}} \quad (1)$$

Estimates of nonresponse and measurement error are estimable by linking the paradata (contact protocols and disposition codes) only to the administrative data. The contact protocols contain a unique key for each household that enables us to link those sources and identify which administrative records are associated with respondents and nonrespondents. Bias estimates were obtained for each component of nonresponse: noncontacts and refusals. The nonresponse bias due to noncontacts is estimated by computing the difference between the administrative statistic of interest based on the contacted cases and the statistic based on the full sample:

⁴ Similar measures were used in Kreuter et al. (2010) to ensure comparability.

⁵ http://fdz.iab.de/en/FDZ_Individual_Data/PASS/Working_Tools.aspx

$$\text{Noncontact Bias } (\bar{y}_{\text{ADMIN}}) = \bar{y}_{\text{ADMIN, Contacts}} - \bar{y}_{\text{ADMIN, Sample}} \quad (2)$$

The bias due to refusals is estimated as the difference between the respondents and contacted cases based on the administrative statistic of interest:

$$\text{Refusal Bias } (\bar{y}_{\text{ADMIN}}) = \bar{y}_{\text{ADMIN, Resps}} - \bar{y}_{\text{ADMIN, Contacts}} \quad (3)$$

Estimates of total nonresponse bias are then obtained by computing the difference between the administrative statistic of interest based on the respondents and the statistic based on the full sample:

$$\text{Total Nonresponse Bias } (\bar{y}_{\text{ADMIN}}) = \bar{y}_{\text{ADMIN, Resps}} - \bar{y}_{\text{ADMIN, Sample}} \quad (4)$$

Finally, estimates of measurement error bias are obtained by first estimating two versions of the same statistic: one using survey data, the other one using administrative data from the portion of respondents answering up to this point. With this estimates of measurement error bias can be obtained by computing the difference between the survey estimate (based on respondents) and the administrative estimate (based on the same respondents):

$$\text{Measurement Bias } (\bar{y}_{\text{PASS}}) = \bar{y}_{\text{PASS, Resps}} - \bar{y}_{\text{ADMIN, Resps}} \quad (5)$$

Given that these data are weighted and have a complex sample design, the standard errors need to be adjusted. Also, it is recommended to account for the overlapping and non-independent portions of the sample (contacts, respondents, consenters) when computing differences between dependent subsamples, especially when the proportion of overlap exceeds 10 percent (Hayes and Berry 2006) as is the case here. The random group method is one approach that can be used to obtain approximately valid standard errors in this setting. The method is known for its versatility in accommodating almost any estimator and almost any sampling design and for its low implementation costs (see Wolter (2007) for a detailed description of the procedure as well as a review of its analytic and administrative properties). Because many strata had only a small number of PSUs, a modified procedure proposed by (Wolter 2007:83) was implemented. First, the data were sorted by stratum and PSUs within stratum. Then multiple systematic random samples of PSUs were drawn, where each sample of PSUs comprised a random group. A total of 40 random groups were created. The variances of the biases were then estimated by computing the variability in the bias estimates across the random groups.

3.2 Analyzing the Relationship between Nonresponse and Non-Consent

A propensity score analysis was performed to examine the relationship between respondents' likelihood of response and consent and determine whether reluctant respondents are more likely to refuse linkage consent and introduce

non-consent bias. A logistic regression model was fit of response based on a limited set of covariates obtained from call records (number of contact attempts, time of day of last contact attempt) and administrative variables obtained from the sample frame (age, gender, household composition). All covariates were significantly related to response at the 0.10 level. Following suggestions from Rosenbaum and Rubin (1983), five roughly equal-sized propensity score groups are formed, ordered from low to high propensity of survey response. Respondents are assigned the value 1, indicating the lowest propensities of response, if their propensity to respond lay between 0.16 and 0.21; the value 2 if their predictive propensity fall within the interval of 0.21 and 0.24; the value 3 for propensities between 0.24 and 0.27, the value 4 if the estimated propensity is between 0.27 and 0.31, and the value 5, indicating the highest propensities of response, if they fall within the range of 0.31 to 0.50.

Consent rates and estimates of non-consent bias are computed within each response propensity quintile. A nonparametric trend test (Cuzick 1985) is used to determine whether the rates increase, and the magnitude of the bias estimates decrease, monotonically as the response propensity stratum increase from lowest to highest. Consent rates are expected to be lowest and the magnitude of the non-consent biases will be greatest for the respondents grouped in the lowest response propensity stratum. The opposite is expected to be true for respondents in the highest response stratum.

All Stata code used in this analysis is available for download in the online appendix.

4 Results

4.1 Survey and Administrative Estimates

Table 2 shows the distribution of the administrative variables for the full sample and for various subgroups of the sample at each stage of the survey process (getting contacted, completing the interview, and providing data linkage consent). For example, according to the administrative data 16.5 percent of the initial sample were foreign citizens; the corresponding estimates were 13.6 percent for the sample cases who were contacted and 11.0 percent for those who completed the interview. In addition, the table shows the distribution of reported values for each survey item for the full set of respondents and for those respondents who gave consent to link their survey responses and administrative records. The corresponding estimate of reported foreign citizenship was 8.5 percent for the full set of respondents and 7.6 percent for the consenting respondents.

By comparing the means and percentages for the various subgroups of the sample in the different columns of Table 2, the effect of non-consent, nonresponse, and measurement bias can be estimated. The differences between the estimates in the third and fourth columns reflect the impact of non-consent. For nonresponse, the differences between the estimates in the first and second columns reflect the impact of noncontact: the differences between those in the second and third columns reflect the impact of refusal nonresponse.

Table 2: Percentage/Mean in Each Subgroup (and Standard Errors), According to Administrative Data and Survey Data

	Administrative Data				Survey Reports	
	Sample	Contacts	Respondents	Consenters	Respondents	Consenters
	n = 17,167	n = 10,717	n = 4,513	n = 3,538	n = 4,513	n = 3,538
Age	39.5 (0.1)	40.3 (0.1)	39.5 (0.2)	39.3 (0.2)	39.5 (0.2)	39.2 (0.2)
Foreign	16.5 (0.4)	13.6 (0.5)	11.0 (0.8)	10.0 (0.7)	8.5 (0.6)	7.6 (0.6)
UB II	80.2 (0.3)	80.8 (0.4)	83.4 (0.6)	83.1 (0.6)	75.9 (0.7)	76.0 (0.8)
Disability	4.9 (1.7)	5.4 (0.2)	5.3 (0.3)	5.3 (0.4)	11.3 (0.5)	11.2 (0.5)
Employed	29.3 (0.4)	30.4 (0.5)	30.3 (0.8)	30.6 (0.8)	29.3 (0.8)	30.0 (0.8)
	n = 5,011	n = 3,234	n = 1,352	n = 1,070	n = 1,352	n = 1,070
Income	799.9 (11.2)	788.8 (14.0)	728.5 (21.3)	730.2 (24.8)	1130.9 (29.7)	1124.7 (32.9)

Note: Parenthetical entries are standard errors.

The impact of measurement error bias for all respondents (regardless of consent) can also be assessed by comparing the estimates in the third and fifth columns, which are based on administrative records and survey reports, respectively. Similarly, the impact of response error for the consenting respondents is assessed by comparing the fourth and sixth columns. Table 2 shows the resulting estimates for individual components that make up the non-consent, nonresponse, and measurement bias for each statistic.

4.2 Estimates of Non-Consent Bias for Administrative Estimates

It is apparent from Table 3 that the impact of non-consent bias is generally small for each estimate. Out of the six estimates, only two of them yield non-consent biases that are significantly different from zero: the non-consent bias estimates for age and foreign citizenship are -0.3 years and -0.9 percent, respectively; these estimates indicate that younger respondents are more likely and foreign citizens are less likely to consent to data linkage. Age is often cited as a significant predictor of consent (see Table 1) and suspicions of its susceptibility to bias are confirmed in this study. However, variables central to the topic, such as benefit reciprocity, employment status, disability status, and income, which have also been found to be related to consent in prior studies, were not subject to non-consent biases in the data examined here. For example, the bias estimate for mean income, an important statistic in economic research, is only 1.7 Euros; a reassuring result.

4.3 Relative Contributions of Non-Consent, Non-response, and Measurement Error Bias

Table 3 also allows an examination of the joint impact of non-consent, nonresponse, and measurement bias and assess their relative contributions to the overall bias in the estimates. First, one can compare the sizes of the non-consent biases and nonresponse biases. In nearly all cases, the nonresponse biases are significantly greater than the non-consent biases. The same pattern is true when non-consent biases are contrasted against the individual components of nonresponse due to noncontact and refusal. One could hypothe-

size that refusal to give linkage consent is influenced by the same mechanism as refusal nonresponse since both require an active refusal by the respondent. In this case, one would expect the non-consent and refusal biases to move the estimates in the same direction. This pattern is true for two of the estimates: age and foreign citizenship; both estimates yield negative non-consent and refusal biases.

If measurement error biases are smaller than non-consent biases, then this would suggest that asking respondents to self-report their administrative information may result in better data quality (less overall error) than asking respondents for linkage consent. The present study does not support this notion (see Table 3). In all cases, the measurement error biases for all respondents (regardless of consent) are larger than the non-consent biases; most of these differences are statistically significant as indicated by the double cross symbol in Table 3. Hence, the strategy of asking for linkage consent appears to yield fewer biases than asking respondents to self-report information contained in their administrative record for the variables considered here. Though one should note that reports on welfare benefit reciprocity is provided for the entire household, thus could be a proxy-report. Also, measurement error for income is difficult to assess as described above.

4.4 Relationship Between Nonresponse and Linkage Non-Consent

The previous results showed that non-consent bias is generally small compared to more traditional sources of survey bias (nonresponse, measurement). However, in some cases, the nonresponse and non-consent biases impacted the estimates similarly by pushing the estimates in the same direction. Thus, a key question to ask is whether there exists a relationship between nonresponse and non-consent, such that reluctant respondents are more likely to refuse linkage consent? Table 4 examines this issue, showing the consent rates by respondents' likelihood of survey response. The table also shows estimates of the non-consent bias for each statistic by respondents' likelihood of response.

A few observations can be made about Table 4. First, there is little consent rate variation across the response propensity groups; the consent rates range from 76.3 to 80.4

Table 3: Nonresponse, Non-Consent, and Measurement Error Bias Estimates (and Standard Errors), by Survey Statistic

	Nonresponse Bias			Measurement Bias		
	Noncontact	Refusal	Total Nonresponse	Non-Consent	Respondents	Consenters
Age	0.8 (0.1)*	-0.7 (0.1)*	0.1 (0.2)	-0.3 (0.1)*	-0.004 (0.01)	0.03 (0.02)
Foreign	-3.0 (0.2)*	-2.6 (0.4)*	-5.6 (0.5)*†	-0.9 (0.3)*†‡	-2.5 (0.3)*†	-2.5 (0.3)*†
UB II	0.6 (0.2)*	2.6 (0.4)*	3.2 (0.5)*†	-0.3 (0.3)†‡	-7.5 (0.4)*‡	-7.1 (0.4)*‡
Disability	0.5 (0.2)*	-0.1 (0.3)	0.4 (0.3)†	0.01 (0.2)†‡	6.1 (0.4)*‡	6.0 (0.5)*‡
Employed	1.0 (0.3)*	-0.1 (0.5)	1.0 (0.5)	0.3 (0.4)	-1.0 (0.6)	-0.6 (0.7)
Income	-11.1 (8.0)	-60.3 (15.4)*	-71.4 (16.1)*†	1.7 (9.3)†‡	402.4 (27.8)*‡	394.5 (30.5)*‡

* $p < 0.05$ † difference between non-consent bias and total nonresponse bias is statistically significant, $p < 0.05$ ‡ difference between non-consent bias and measurement error bias is statistically significant, $p < 0.05$

Table 4: Non-Consent Bias over Response Propensity Strata

	n	Consent rate, %	Non-Consent Bias					
			Age	Foreign, %	UB II, %	Disability, %	Employed, %	Income
Response Propensity Strata								
[1] Q1 (low)	619	78.5	-0.3	-1.0	-0.7	-0.4	0.8	-1.3
[2] Q2	857	76.3	-0.3	-1.0	0.5	0.1	-0.6	-10.7
[3] Q3	855	80.3	0.2	-1.7	-0.2	-0.0	0.3	15.9
[4] Q4	994	76.7	-0.6	-1.1	-0.7	0.2	1.3	-0.2
[5] Q5 (high)	1188	80.4	-0.2	-0.2	-0.3	0.0	-0.3	2.7
Trend Test								
$\Delta [1] - \Delta [5]$		$ z = 1.00,$ $p = 0.317$	$ z = 0.20,$ $p = 0.841$	$ z = 0.40,$ $p = 0.689$	$ z = 0.20,$ $p = 0.841$	$ z = 1.00,$ $p = 0.317$	$ z = 0.00,$ $p = 1.00$	$ z = 1.20,$ $p = 0.230$

percent. Second, the relationship between respondents' likelihood of response and the rate of consent does not seem to be consistent. That is, the consent rates do not increase monotonically as a function of respondents' likelihood of response (trend test, $p < 0.317$). One would expect the lowest response propensity cases to be the least likely to consent but this group yielded a consent rate of 78.5 percent, which was the midpoint among the consent rates observed across the five groups. The highest response propensity group did yield the highest consent rate (80.4 percent), but so did the third quartile (80.3 percent). Finally, just as there was no monotonic trend between the likelihood of response and consent rates, no consistent relationship between the likelihood of response and non-consent bias can be observed for the six variables. In conclusion, Table 4 suggests that respondents' likelihood of response are not related to linkage consent rates nor are they related to non-consent biases.

5 Discussion

The PASS survey provides a unique opportunity for survey researchers to estimate non-consent biases for key administrative variables. To our knowledge, this is the first study outside of the medical field to estimate non-consent biases on the actual administrative data rather than estimates of bias based on survey variables.

The study has three main findings that correspond to the research questions posed initially. First, small non-consent

biases were found for estimates of demographic variables (age, foreign citizenship), but not for substantive variables (disability status, employment, income, benefit reciprocity).

Second, estimates of non-consent bias were quite small relative to other sources of bias. In general, the biggest contributor of bias was due to measurement, followed by non-response and non-consent. To put it another way, the effort of obtaining linkage consent from respondents (and the possible non-consent bias consequences of doing so) seems to pay off in terms of better data quality (less bias) over asking respondents to report their administrative information during the survey interview.

Finally, given the fact that both non-consent and non-response can be caused by similar factors (both can be caused by refusal), no consistent relationship was found between the two. The most reluctant respondents, or those who possessed the lowest propensities of response, were not less likely to consent to data linkage. Also, no strong relationship was found between response propensities and non-consent biases.

In general, the results of this case study paint an optimistic picture for the practice of linking survey and administrative data. Despite numerous findings in the literature of differential consent rates among key population subgroups (see Table 1, last column), the results presented here suggest that the impact of non-consent bias on important labour-market variables is small and negligible in most cases. Furthermore, data linkage seems to produce better data quality than asking respondents to self-report their administrative in-

formation. Not only does data linkage reduce the length of the questionnaire (and, presumably, respondent burden), but it also appears to make sense from a total survey error perspective.

This study is not without limitations. The analysis presented here is based on a specific population under study, a sample of recent benefit recipients under the age of 65 in Germany. Results may differ for older and consistently employed populations. A sensitivity check was performed to examine differences between the benefit and general population samples based on a consent propensity model. The fact that both samples yielded similar correlates of consent, including the same correlates found in other studies, lends support to the generalizability of the findings.

An important caveat about administrative data is that they are not always free of errors. For this reason, we chose a set of administrative variables for our analyses that are known to be carefully checked as they are used to distribute benefit payments. It would take a significant error in the administrative data to reverse our finding that data linkage yields better data quality than asking respondents to self-report their administrative data.

It is conceivable that in the future non-consent bias studies will be mandated for federally funded surveys, just as nonresponse bias studies are mandated in some countries when surveys fail to achieve a predefined response rate (e.g., U.S. Office of Management and Budget (2006)). While not every survey has access to administrative records for the full sample, the methods employed in this study could potentially be used to assess the impact of non-consent bias and inform data users of any possible consequences. Linking the consent indicator to the administrative data is a viable approach to estimating non-consent biases and upholds the promise to respondents not to link their survey responses and administrative records if they so choose. This example could serve as a precedent for researchers and may be used as a stimulus to generate discussions with Institutional Review Boards and administrative data agencies to facilitate non-consent bias research.

Acknowledgements

We are very thankful to Gerrit Müller and Mark Trappmann for their time and assistance in helping us getting to know the data structure, the intricate details of the PASS data collection, and for providing us with pieces of analysis code. Without their help this paper would not have been possible. We also thank Jörg Heining and Stefan Bender for their assistance with data access and support, and Rainer Schnell as well as anonymous reviewers for their suggestions. Finally, we thank Rick Valliant and Brady West for their advice and suggestions regarding variance estimation procedures. This work was partially supported by the Alexander von Humboldt Foundation.

References

Banks, J., Lessof, C., Taylor, R., Cox, K., & Philo, D. (2005). *Linking survey and administrative data in*

- the English Longitudinal Study of Ageing*. Presented at the Meeting on Linking Survey and Administrative Data and Statistical Disclosure Control, Royal Statistical Society. <http://www.ccsr.ac.uk/methods/events/linkage/Lessof.pdf>.
- Bates, N. (2005). *Development and testing of informed consent questions to link survey data with administrative records*. Proceedings of the Survey Research Methods Section, American Statistical Association, 3786-3793.
- Bender, S., & Haas, A. (2002). *Die IAB-Beschäftigtenstichprobe*. IAB-Kompodium Arbeitsmarkt- und Berufsforschung. (Beiträge zur Arbeitsmarkt- und Berufsforschung, BeitrAB 250), 3-12.
- Blumberg, S. J., & Luke, J. V. (2007). Coverage bias in traditional telephone surveys of low-income and young adult. *Public Opinion Quarterly*, 71, 734-749.
- Buckley, B., Murphy, A. W., Byrne, M., & Glynn, L. (2007). Selection bias resulting from the requirement for prior consent in observational research: a community cohort of people with ischaemic heart disease. *Heart*, 93, 1116-1120.
- Calderwood, L., & Lessof, C. (2009). *Enhancing longitudinal surveys by linking to administrative data*. New York: Wiley.
- Chambers, R. (2009). *Regression analysis of probability-linked data*. Official Statistics Research Series, 4. <http://www.statisphere.govt.nz/further-resources-and-info/officialstatistics-research/series//media/statisphere/Files/official-statistics-research-series/osrseries-v4-2009-regression-analysis-probability-linked-data.pdf>.
- Christoph, B., Müller, G., Gebhardt, D., Wenzig, C., Trappmann, M., Achatz, J., et al. (2008). *Codebook and documentation of the Panel Study 'Labour Market and Social Security' (PASS). vol. 1: introduction and overview, wave 1 (2006/2007). Number 05*.
- Curtin, R., Presser, S., & Singer, E. (2000). Changes in telephone survey nonresponse over the past quarter century. *Public Opinion Quarterly*, 64, 413-428.
- Curtin, R., Presser, S., & Singer, E. (2005). Changes in telephone survey nonresponse over the past quarter century. *Public Opinion Quarterly*, 69, 87-98.
- Cuzick, J. (1985). A Wilcoxon-type test for trend. *Statistics in Medicine*, 4, 87-89.
- Dahlhamer, J. M., & Cox, C. S. (2007). *Respondent consent to link survey data with administrative records: results from a split-ballot field test with the 2007 National Health Interview Survey*. Proceedings of the Federal Committee on Statistical Methodology Research Conference.
- Davern, M., Call, K. T., Ziegenfuss, J., Davidson, G., Beebe, T. J., & Blewett, L. (2008). Validating health insurance coverage survey estimates: a comparison of self-reported coverage and administrative data records. *Public Opinion Quarterly*, 72, 241-259.
- De Heer, W. (1999). International response trends: results of an international survey. *Journal of Official Statistics*, 15, 129-142.
- Eckman, S., & Kreuter, F. (2011). Confirmation bias in housing unit listing. *Public Opinion Quarterly*, 75, 139-150.
- Finkelstein, M. M. (2001). Do factors other than need determine utilization of physician's services in Ontario? *Canadian Medical Association Journal*, 165, 565-570.
- Fowles, J. B., Fowler, E. J., & Craft, C. (1998). Validation of claims diagnoses and selfreported conditions compared with medical records for selected chronic diseases. *The Journal of Ambulatory Care Management*, 21, 24-34.
- Groves, R. (2006). Rates and nonresponse bias in household surveys. *Public Opinion Quarterly*, 70, 646-675.

- Haider, S., & Solon, S. (2000). *Non random selection in the HRS social security earnings sample*. Working Paper No. 00-01, RAND Corporation, Santa Monica, Calif.
- Hayes, L. J., & Berry, G. (2006). Comparing the part with the whole: should overlap be ignored in public health measures? *Journal of Public Health*, 28(3), 278-282.
- Hebert, P. L., Geiss, L. S., Tierney, E. F., Engelgau, M. M., Yawn, B. P., & McBean, A. M. (1999). Identifying persons with diabetes using medicare claims data. *American Journal of Medical Quality*, 14, 270-277.
- Jacobebbinghaus, P., & Seth, S. (2007). The german integrated employment biographies sample IEBS. *Schmollers Jahrbuch. Zeitschrift für Wirtschafts- und Sozialwissenschaften*, 127, 335-342.
- 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, 701-722.
- Keating, N. L., Landrum, M. B., Landon, B. E., Ayanian, J. Z., Borbas, C., & Guadagnoli, E. (2003). Measuring the quality of diabetes care using administrative data: is there bias? *Health Services Research*, 38, 1529-1545.
- Kho, M. E., Duggett, 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.
- Knies, G., Burton, J., & Sala, E. (2012). Consenting to health record linkage: evidence from a multipurpose longitudinal survey of a general population. *BMC Health Services Research*, 12.
- Kreuter, F., Müller, G., & Trappmann, M. (2010). Nonresponse and measurement error in employment research: making use of administrative data. *Public Opinion Quarterly*, 74, 880-906.
- Kreuter, F., Presser, S., & Tourangeau, R. (2008). Social desirability bias in CATI, IVR, and web surveys: the effects of mode and question sensitivity. *Public Opinion Quarterly*, 72, 874-865.
- Lahiri, P., & Larsen, M. D. (2005). Regression analysis with linked data. *Journal of the American Statistical Association*, 100, 222-230.
- Losina, E. J., Barrett, J., Baron, J. A., & Katz, J. N. (2003). Accuracy of medicare claims data for rheumatologic diagnoses in total hip replacement recipients. *Journal of Clinical Epidemiology*, 56, 515-519.
- Merkle, D., & Edelman, M. (2002). Nonresponse in exit polls: a comprehensive analysis. In R. Groves, D. Dillman, J. Eltinge, & R. Little (Eds.), *Survey Nonresponse* (p. 243-258). New York: Wiley.
- Rosenbaum, P. R., & Rubin, D. R. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70, 41-55.
- Sakshaug, J. W., Couper, M. P., Ofstedal, M. B., & Weir, D. (2012). Linking survey and administrative records: Mechanisms of consent. *Sociological Methods and Research*. (Forthcoming)
- Sakshaug, J. W., Weir, D., & Nicholas, L. H. (2010a). *Validating diabetes disease status in claims and survey data*. Proceedings of the Statistics Canada Symposium on the Interplay among Censuses, Surveys, and Administrative Data.
- Sakshaug, J. W., Yan, T., & Tourangeau, R. (2010b). Nonresponse error, measurement error, and mode of data collection: trade-offs in a multi-mode survey of sensitive and non-sensitive items. *Public Opinion Quarterly*, 74, 907-933.
- Sala, E., Burton, J., & Knies, G. (2010). *Correlates of obtaining informed consent to data linkage: respondent, interview and interviewer characteristics*. ISER Working Paper Series, No 2010-28. <http://econpapers.repec.org/RePEc:ese:iserwp:2010-28>.
- Scheuren, F., & Winkler, W. E. (1993). Regression analysis of data files that are computermatched. *Survey Methodology*, 19, 39-58.
- Schnell, R. (1997). *Nonresponse in Bevölkerungsumfragen*. Opladen: Leske und Budrich.
- Schnell, R., Bachteler, T., & Reiher, J. (2009). Privacy-preserving record linkage using bloom filters. *BMC Medical Informatics and Decision Making*, 9(41).
- Silva, M. S., Smith, W. T., & Bammer, G. (2002). The effect of timing when seeking permission to access personal health services utilization records. *Annals of Epidemiology*, 12, 326-330.
- Smith, T. (2011). The report of the international workshop on using multi-level data from sample frames, auxiliary databases, paradata and related sources to detect and adjust for nonresponse bias in surveys. *International Journal of Public Opinion Research*, 23, 389-402.
- Tate, A. R., & Calderwood, L. (2005). Mother's consent to linkage of survey data with her child's birth registration records in a multi-ethnic national cohort study. *International Journal of Epidemiology*, 35, 294-298.
- The American Association for Public Opinion Research. (2011). *Standard definitions: final dispositions of case codes and outcomes rates for surveys*.
- Trappmann, M., Gundert, S., Wenzig, C., & Gebhardt, D. (2010). PASS: a household panel survey for research on unemployment and poverty. In: Schmoller's Yearbook. *Journal of Economic and Social Science*, 130, 609-622.
- U.S. Office of Management and Budget. (2006). *Standards and guidelines for statistical surveys*. http://www.whitehouse.gov/sites/default/files/omb/inforeg/statpolicy/standards_stat_surveys.pdf.
- Winkler, W. E. (2006). *Overview of record linkage and current research directions*. U.S. Census Bureau, Research Report Series (Statistics #2006-2). <http://www.census.gov/srd/papers/pdf/rrs2006-02.pdf>.
- Wolter, K. M. (2007). *Introduction to Variance Estimation* (2nd ed.). New York: Springer.
- Young, A. F., Dobson, A. J., & Byles, J. E. (2001). Health services research using linked records: who consents and what is the gain? *Australian and New Zealand Journal of Public Health*, 25, 417-420.