Bruckmeier, Kerstin; Riphahn, Regina T.; Wiemers, Jürgen

Article — Published Version

Misreporting of program take-up in survey data and its consequences for measuring non-take-up: new evidence from linked administrative and survey data

Empirical Economics

Provided in Cooperation with:
Springer Nature

Suggested Citation: Bruckmeier, Kerstin; Riphahn, Regina T.; Wiemers, Jürgen (2020) : Misreporting of program take-up in survey data and its consequences for measuring non-take-up: new evidence from linked administrative and survey data, Empirical Economics, ISSN 1435-8921, Springer, Berlin, Heidelberg, Vol. 61, Iss. 3, pp. 1567-1616,
https://doi.org/10.1007/s00181-020-01921-4

This Version is available at:
https://hdl.handle.net/10419/288381

Empirical Economics (2021) 61:1567–1616
https://doi.org/10.1007/s00181-020-01921-4

Misreporting of program take-up in survey data and its consequences for measuring non-take-up: new evidence from linked administrative and survey data

Kerstin Bruckmeier[1] · Regina T. Riphahn[2] · Jürgen Wiemers[1]

Received: 25 July 2019 / Accepted: 29 July 2020 / Published online: 7 October 2020

Abstract
The international literature studies non-take-up behavior of eligible populations to evaluate the effectiveness of government programs. A major challenge in this literature is the measurement error regarding benefit take-up. In our data, we observe both actual welfare receipt and respondents' survey information on their program take-up. This allows us to observe the measurement errors that other researchers must estimate. We describe survey misreporting and investigate how it biases the estimates of the magnitude and patterns of benefit take-up among eligible households. Our findings suggest that the extent of measurement error can be substantial. It varies with the characteristics of the misreporting population and is associated with the drivers of misreporting. This indicates that survey-based analyses of take-up behavior are likely subject to severe biases.

Keywords Take-up · Welfare · Misreporting · Survey data · Administrative data · Data linkage

JEL Classification I32 · H75 · C81

1 Introduction

The non-take-up of benefit programs is an important aspect of state support programs; if eligible households do not take up transfers, such programs are ineffective and the basic needs of population groups may remain unaddressed. Eurofound (2015) shows that benefit non-take-up is internationally pervasive and frequently affects more than

✉ Regina T. Riphahn
regina.riphahn@fau.de

1 IAB, Nuremberg, Germany

2 Friedrich-Alexander-University Erlangen-Nürnberg, Lange Gasse 20, 90403 Nuremberg, Germany

40% of the eligible population. This has motivated a long-standing discussion of the extent and determinants of take-up behavior in the international literature (e.g., Moffitt 1983; Blundell et al. 1988; Hernandez et al. 2007). We contribute to that literature and address the research question of how measurement error in reported benefit receipt biases survey-based analyses of take-up behavior. Our unique data provide unusually precise information on true program participation.

A key challenge in the empirical analysis of non-take-up is its measurement. A correct measurement of non-take-up requires valid information on both program eligibility and program take-up. Most non-take-up studies must rely on survey data to measure both concepts. However, several factors can generate measurement errors in survey data and lead to imprecise estimations of program eligibility as well as program take-up. Several studies discussed the relevance of misreporting for the reliability of survey data, see, e.g., Meyer et al. (2015), Meyer and Goerge (2011), Taeuber et al. (2004), or Card et al. (2001). Regarding mismeasurement in program take-up, Meyer et al. (2015) pointed out that data from household surveys missed the measurement of approximately half of all welfare and food stamp payments in major household surveys in the USA. Given the importance of mismeasurement in reported program take-up, our paper focuses on its consequences for the analysis of benefit non-take-up.

Several factors contribute to the mismeasurement of benefit receipt (Bound et al. 2001). Surveys often ask respondents whether they have received benefits during a certain period in the past. Respondents may have completely forgotten past benefit receipt (recall bias), or they may not remember the exact dates of receipt. For example, events can be reported as more recent than they actually occurred, which is known as the “forward telescoping bias” in the survey literature (Bradburn et al. 1994). This form of bias could lead to mismeasurement in the form of misreporting of benefit take-up. Additionally, if different benefits are available or benefits can be claimed simultaneously, beneficiaries might report incorrectly, claiming specific benefit(s) that they did not receive while inadvertently omitting the benefit(s) that they did receive, which ultimately leads to misreporting (Hancock and Barker 2005; Krafft et al. 2015). Another source of misreporting is the “social desirability bias” (Bound et al. 2001). In particular, the receipt of means-tested social welfare benefits is often perceived as stigmatizing and thus respondents may underreport their receipt of these benefits.

The literature on program non-take-up and program participation has proposed several approaches to deal with potential measurement biases. We contribute to the set of studies that apply external validation samples for the study of measurement errors in benefit receipt. Other approaches rely on structural modeling (e.g., Duclos 1995, 1997; Pudney 2001; Bollinger and David 2001; Hernandez and Pudney 2007) or focus on careful data handling (Hancock and Barker 2005).

External validation samples can be used to assess the extent of measurement errors in survey data and to mitigate their effects in analyses. Studies that employ this approach are scarce because this requires linking survey data to administrative data. Linked data are often not available or only available for specific groups, periods, or regions. One example of this approach is Bollinger and David (1997). They take advantage of data with information on true participation and survey responses. Information on response errors in the validation sample is then considered in the likelihood function for the primary sample. The authors find that modeling response errors generates large differences in the estimation of program participation even when the validation data are gathered on a sample that differs from the survey data. Other examples are Mittag (2016) and Meyer and Mittag (2017a). Both investigate different methods to account for misclassification in survey data if linked data are not available to the researcher. They use validation data to evaluate the effectiveness of their formulas for bias reduction. The studies closest to our approach are Meyer and Mittag (2017b) and Meyer et al. (2018). The former use linked data to correct survey data from the Current Population Survey (CPS) on reported benefit receipt from different programs and cover the state of New York. They find that the poverty-reducing effect of benefit programs is nearly doubled using the corrected data but do not study take-up behavior. In contrast, Meyer et al. (2018) only focus on the food stamp program, but link data from three different surveys (American Community Survey, Current Population Survey, and Survey of Income and Program Participation) with administrative data covering the states of Illinois and Maryland. The authors find underreporting rates of between 23% and 50% depending on the survey. They study the reasons for misreporting and investigate the impact of misreporting on program take-up estimations. The authors find that particularly single parents, non-whites, and the elderly understate participation which generates biases in program receipt regressions: Underreporting is part of the explanation for low benefit receipt among the elderly.

In this study, we similarly link survey data to administrative data, which informs us about the true program take-up of survey respondents. This allows us to determine precisely when survey information differs from actual benefit receipt. Thus, we can determine the presence of measurement errors and misreporting directly. We consider a general income support program that is available for the working-age population in Germany (Unemployment Benefit II). This general benefit is less subject to the risk of benefit confusion than specific transfers available, e.g., for retirees only, which have been discussed in the literature (e.g., Duclos 1997; Hancock and Barker 2005). We investigate whether the extent and pattern of program non-take-up differ after correcting for the misreporting of survey respondents. To the best of our knowledge, our study is the first to use linked data to investigate the impact of mismeasurement on program non-take-up.[1]

Relative to the extant literature, the value added by our contribution relates to several aspects. First, we offer evidence from a European country where social norms and the cost–benefit setting with respect to public transfers may differ from that in the USA, the origin of the previous literature. Second, we can take advantage of a nationwide program with nationwide validation data from administrative sources. Prior studies in the USA are typically based on evidence from regional units. Third, we refer to prior studies on take-up behavior and their common empirical specification and investigate to what extent the results may be biased. This differs from the existing literature in that we limit our analysis to the eligible population identified using simulation techniques. Thus, our analysis does not examine program participation within the general population, i.e., program targeting, but rather the claiming behavior of the

[1] The setting of our analysis lends itself to an application of the literature on classification errors in binary dependent variables (see Hausman et al. 1998; Hug 2010).

eligible population. Finally, we offer evidence on the direction of the bias introduced by misreporting and study the correlation patterns of misreporting itself.

We find that correcting for misreporting in the data modifies the results of take-up regressions. The marginal effects of characteristics associated with benefit take-up often deviate by more than 30% after correcting for measurement error in the outcome. These results are robust to various changes in procedure. Furthermore, we find evidence that the patterns determining misreporting of program take-up are reflected in the sensitivity of marginal effects to the data correction. In particular, we observe that reporting behavior varies with age, income, and immigrant status as in Meyer et al. (2018); those who are more likely to rely on benefits in a less permanent way are most likely to underreport. Therefore, uncorrected estimations may misleadingly lead us to overestimate the stability of welfare receipt and suggest that welfare is less a protection against income shocks and volatility and more a source of subsistence income.

Our results are important for several reasons. First, they show that survey data can yield biased results in the study of take-up behavior based on self-reported information; this confirms prior studies which investigated the role of misreporting on program take-up in the overall population (e.g., Meyer et al. 2018). Second, our findings are more reliable than prior contributions that are not based on linked survey and administrative data. Third, we cover a general and well-known nationwide benefit program. Fourth, we show that the patterns of misreporting and the estimation biases can be related. The coefficients that are the most biased in take-up equations are those associated with the misreporting groups' characteristics. Thus, even though misreporting rates may be lower than in other settings studied, the biases are nevertheless sizeable.

We structure our paper in six sections. In the next section, we briefly characterize the benefit program considered in our analysis. We lay out our empirical approach in Sect. 3. Section 4 describes the nature of our data and provides descriptive statistics. Section 5 presents our empirical results and robustness tests. We draw conclusions in Sect. 6.

2 Institutional background

We study the take-up of the German minimum income support program Unemployment Benefit II (UB II). The transfer is available for working-age individuals who are able to work and their families. Alternative programs cover persons who have reached retirement age or are unable to work. UB II eligibility exists if a household's net income is below the legally determined minimum; in this situation, the benefit covers the difference. The benefits are means tested. Almost all types of income are considered in the means test. The minimum income deemed sufficient to guarantee an acceptable minimum living standard for a household is calculated based on the number of household members and—for minors—their age. In 2018, the standard benefit for an adult is 416 euros per month. Expenses for rent, heating, and health care are paid in addition to the standard benefit; benefits can be higher in special circumstances (e.g., for single-parent families, pregnant women, or those with special food requirements). Households with more than a maximum amount of wealth are not eligible; wealth comprises financial assets plus the value of owned property minus mortgage liabilities. Self-occupied residential property is ignored in the wealth check in most cases. Eligibility is not conditional on unemployment. In 2016, approximately 41% of regular benefit recipients were unemployed and 28% received the benefit to top up (insufficient) earnings from employment. Others were temporarily unable to work, e.g., because of child care obligations (Statistik der Bundesagentur für Arbeit 2018).[2]

The UB II program follows federal regulations and is administered either by the employment services of the Federal Employment Agency or by the municipality. In 2016, the program covered approximately 6.2 million individuals in 3.3 million households and paid out approximately 35.2 billion euros (STBA 2018). Thus, UB II is a well-known program that is generally available to the entire working-age population which is able to work. As a large part of the German population receives UB II, the program is comparable to programs such as Medicaid and SNAP (see Table 14), which are the subject of recent studies on benefit misreporting based on linked data (Meyer et al. 2018; Davern et al. 2019). In contrast to these programs, UB II provides a means-tested basic income that is the most important source of income for many households. The UB II benefit level is thus also significantly higher than, for example, for SNAP (Table 14). This may lead to a relatively small measurement error for UB II in survey data if the household is well informed about the benefit.

Recent studies on take-up of UB II using survey data from the German Socioeconomic Panel (SOEP) show that based on monthly data, between 46 and 58% of eligible households did not take up the benefit in the years 2005–2007 (e.g., Bruckmeier and Wiemers 2012). The authors find that take-up varies with the potential benefit amount and the expected duration of eligibility expressed in proxy variables such as education and region of residence.

3 Empirical approach

In recent decades, a large number of empirical studies on the determinants of (non-)take-up have been conducted for a wide range of means-tested benefits (see, e.g., Blundell et al. 1988; Blank and Ruggles 1996; Riphahn 2001; Wilde and Kubis 2005; Whelan 2010; Bruckmeier and Wiemers 2012). All survey-based studies of take-up behavior have to address the problem that the data do not provide information about benefit eligibility. The studies therefore simulate welfare eligibility for every household in the dataset using a microsimulation model. Then, given a model of welfare eligibility, the literature typically defines benefit non-take-up as being eligible according to the simulation model while reporting non-receipt of the benefit in the survey data.

Following Blundell et al. (1988), we model the take-up decision in a discrete choice framework to analyze determinants of take-up behavior. This approach assumes that benefits are taken up if the household's net utility from claiming exceeds the utility when not claiming the benefit. The net utility takes into account that claiming a benefit is associated with non-monetary costs, which occur, e.g., because of insufficient knowledge of entitlement rules, the claiming process, the administrative procedures,

[2] For more information on details of the program, see Bruckmeier and Wiemers (2018).

and stigma costs, i.e., the fear of stigmatization and negative societal attitudes toward welfare dependence (see van Oorschot 1991).

Since neither the utility nor the costs associated with claiming a benefit can be directly observed, empirical analyses of take-up behavior use proxy variables to capture both. We build on the existing empirical literature for choosing the proxy variables that influence the take-up decision (Riphahn 2001; Wilde and Kubis 2005; Frick and Groh-Samberg 2007; Whelan 2010; Bruckmeier and Wiemers 2012) to enhance the external validity of our analysis. The most obvious factor affecting utility when claiming UB II is the household's benefit entitlement (see, e.g., Blundell et al. 1988). We follow the standard procedure in the literature and simulate the level of benefit entitlements for all households in the survey data, i.e., for households reporting benefit receipt and for households not reporting benefit receipt. This ensures that the level of entitlement for both types of households is determined according to a uniform measurement concept.[3]

In addition, utility and costs of claiming may vary by household type, i.e., whether it is a single or a couple household and whether or not children are present. Therefore, we consider indicators of household type. As general sociodemographic indicators, we account for age, education, an indicator for first- or second-generation immigrants, and disability of the household head. We also control for actual rent paid and home ownership. We expect higher take-up rates in Eastern Germany due to higher unemployment there and thus control for residence in Eastern Germany.

We follow the literature (e.g., Blundell et al. 1988; Bollinger and David 1997, 2001; Duclos 1995; Pudney 2001) in assuming an i.i.d. standard normal error term to capture the unobserved part of the take-up decision.[4] Furthermore, we account for the potential endogeneity of the simulated benefit. The simulated benefit amount is endogenous if unobserved factors that influence the take-up decision are correlated with unobserved factors determining the household's labor supply. In this case, the simulated benefit amount would be endogenous because it depends on the household's simulated net income, which is a function of earned income.

Thus, our first specification for the empirical take-up equation is an IV probit estimation (see, e.g., Whelan 2010 and Bruckmeier and Wiemers 2012). The estimation of the IV probit requires the choice of credible instruments for the level of entitlement. Following Bruckmeier and Wiemers (2012), we use (a) the level of the regular benefit entitlement before own income is deducted, (b) non-earned public transfers, i.e., income independent of current labor supply including public and private pensions, and (c) private transfers from persons outside of the household. These are important determinants of the simulated UB II entitlement, and thus, these instruments satisfy the requirement that the instrument must be (strongly) partially correlated with the endogenous variable.

[3] Clearly, simulation may generate measurement error which we abstract from in this paper. The eligibility simulation is described in greater details in Appendix 1.

[4] As a robustness check, we also consider a logistic distribution for the error term. We find that the results are robust with respect to the distributional assumption (results available upon request).

In a second specification, we utilize the panel structure of our data by adding random effects to the IV probit model (RE IV); this accounts for unobserved heterogeneity at the household level.[5]

The endogeneity of the benefit level for the take-up decision is often ignored in the literature on empirical models of take-up behavior, either because of the lack of credible instruments or because accounting for endogeneity complicates the estimation of the model substantially. Thus, to demonstrate the impact of misreporting on take-up models that do not consider the benefit level as endogenous, we also estimate a pooled probit model and a RE probit model (see Table 9).

4 Data and sample

4.1 Data

We use data from the household panel study "Labour Market and Social Security" ("Panel Arbeitsmarkt und soziale Sicherung," PASS), a survey designed for research on unemployment and poverty (Trappmann et al. 2010, 2013; Berg et al. 2014 for technical documentation). The first survey of this study interviewed more than 12,000 respondents in 2006–2007. The seventh survey wave was completed in 2013 (for interview periods, see Table 5). Because the survey instruments and interview program were revised after the first wave (Gebhardt et al. 2009), we only use surveys 2–7.

The data consist of two subsamples. The first subsample considers UB II recipients, while the second subsample covers the overall German population, oversampling those with low socioeconomic status. The UB II sample is randomly drawn from the administrative records of the Federal Employment Agency. To retain a representative character for the population of UB II recipients, subsample one is refreshed each year to include new recipients of UB II (benefit-inflow-sample).

The general population sample is a random draw from a database of addresses of private households in Germany. It is provided by a commercial provider in wave one and is taken from municipality population registers in wave five (refreshment sample). For a detailed description of the sampling design, see Gebhardt et al. (2009).[6] The final weights we use in the descriptive analysis balance distortions arising from the sample design and reflect the entire German population. All regression analyses use unweighted data; a robustness test discussed in Sect. 5.3 evaluates whether this affects the results.

The PASS data are particularly suitable for our analyses because they focus on potential beneficiaries living in low-income households. Beste et al. (2018) find that

[5] All models are estimated with the Stata command "gsem."

[6] The sampling in wave five involved several steps. Step one draws 300 postcodes (regions) as primary sampling units, i.e., households from both populations—UB II recipients and private households—within each postcode. Based on the number of benefit-receiving households (sample 1) and the number of private households (sample 2) in a postcode, each household receives a uniform selection probability. Design weights for the gross sample reflect the selection probability. Logit models for panel participation are the basis to account for the participation probability and to adjust design weights in the second step (see Gebhardt et al. 2009). Finally, both samples were calibrated to official statistics on UB II recipients and private households in Germany.

the income distribution in the PASS data (starting with wave 2) is similar to that of two other data sources (SOEP and Mikrozensus). Furthermore, PASS interviews respondents about their current welfare receipt, and it allows us to link survey data with administrative records on welfare receipt. The interviewers ask the head of household whether and for how long the household has received UB II since the last survey or in the last two years, and whether the household currently still receives UB II. Interviewers determine the head of the household during the household's first participation in the survey as the person who is best informed about the household finances. The PASS gathers information on UB II receipt via "dependent interviewing," i.e., interviewers remind the head of the household of the answer in the previous interview prior to asking about current receipt (Berg et al. 2012). This form of interviewing should result in reduced misreporting of program take-up (Lynn et al. 2012).

In our analysis sample, we consider household observations with realized personal interviews. We drop household observations whose heads are above age 65 years or in receipt of retirement benefits, students, and individuals pursuing apprenticeship training because these groups benefit from alternative transfer programs. We require that the household responds to the question on current welfare receipt, that there is only one benefit-receiving unit ("community of need") in the household, and that there is valid information on earnings (see Appendix 1 for details). Across waves 2–7, our sample covers 30,878 annual household-level observations overall and approximately 5000 observations per year. For each household-year observation, we simulate UB II benefit eligibility and calculate potential benefit amounts. This yields 17,585 UB II eligible household-year observations.

4.2 Data linkage

The opportunity to link survey with administrative data is rare in the literature. In particular, we are able to link the PASS survey data to the administrative records of the Federal Employment Agency. The data, originally collected at local job agencies ("job centers"), contain information on claims for UB II. The data perfectly reflect official payments. The Institute for Employment Research (IAB) and the Research Data Center (FDZ) of the German Federal Employment Agency (BA) at the IAB have access to these administrative data and are responsible for processing, anonymizing, and providing it for empirical research. For our analysis, we link the survey data to administrative data of the "Unemployment Benefit II Recipient History" ("Leistungshistorik Grundsicherung," LHG, version 11.01.01-150220) of the IAB (Antoni et al. 2016). The administrative UB II data contain information on socioeconomic variables of eligible individuals and regional variables. To measure UB II over- and underreporting, we compare the information the respondents provide on UB II receipt at the month of the interview with the information on UB receipt in this month gained from the LHG.

Because of legal constraints, the survey information can only be linked to the administrative data if the participant consented to linkage in the survey. Therefore, interviewers ask participants during the course of the interview for consent to merge their survey data to their administrative data that are available at the IAB (for details, see

Appendix 2). The consent rate in the PASS is approximately 80%, which is comparable to other survey studies (Berg et al. 2014; Sakshaug and Kreuter 2012). In our sample of simulated eligible households, we have a consent rate of 83.4% (see Table 6). Because respondents who do not agree to the data linkage are asked again in the next wave, the proportion of observations for which an approval is available is significantly higher. Overall, we could not use only 4% of all household-year observations of eligible households because of missing consent to data linkage. This leaves us with 16,874 household-year observations of simulated eligible respondents who agreed to the linkage.

Next, we merge these 16,874 observations with a key file generated by the German Record Linkage Center (Antoni and Schnell 2019). This key file is based on the identification of the PASS respondents in administrative data of the IAB. To identify respondents in the administrative data, harmonized information on addresses and personal characteristics from different administrative data sources collected by the Federal Employment Agency are used. Individuals who never worked in dependent employment, who are exempt from social security contributions (e.g., civil servants), or who have never been registered as unemployed or benefit recipients are not in the data. The record linkage is based on multilevel deterministic and probabilistic methods for linking datasets (see Sakshaug et al. 2017 for a detailed description and Appendix 2). From our sample of 16,874 household-year observations of respondents who agreed to data linkage, we identified 15,925 observations in the administrative data, which amounts to a linkage rate of 94%. Of the 15,925 matches, 15,095 were unique matches and 830 were duplicates, which were corrected following a procedure described in Appendix 2. As a robustness check, we verify our main findings for a sample without duplicates in Sect. 5.3. From our sample of 15,925 linked observations, we keep 14,834 observations with no missing values in the covariates for our descriptive results and the regression analysis. Finally, we drop 270 observations for whom the gender and age information of survey data and administrative data in the LHG does not match. After these steps, the analysis sample comprises 14,564 observations, which represents 83% of the simulated eligible population sample (17,585).

A potential problem of the data linkage is that results may be biased because of selectivity in either non-consent or non-identifiability in the administrative data. With respect to non-consent, misreporting of benefit receipt might be biased downwards if non-consent to the data linkage is positively correlated with the misreporting of benefit receipt: Households who do not want to admit to receiving UB II might also be reluctant to agree to data linkage if they fear that their misreporting might be discovered. Column 1 of Table 8 indicates correlation patterns underlying the probability of not giving consent to the data linkage. We find some statistically significant and small effects: Immigrants, younger, non-disabled persons, those living in single households, or those living in Western Germany have a higher probability to refuse consent. This suggests that our analysis is a rather conservative estimate of misreporting because individuals with low data linkage probability and who are underrepresented in the data generally tend to have a high propensity to misreport.

With respect to non-identifiability in administrative data, column 2 of Table 8 shows the correlation patterns behind the probability that a household cannot be linked to the administrative data for the sample of simulated eligible households with consent to data linkage. Here, we find no significant marginal effects in most sociodemographic and household characteristics except for some age groups, migration background, home owners, and the subsample two indicator.

Overall, these results indicate small systematic effects; thus, we conservatively underestimate misreporting and its correction. We will provide two robustness checks in Sect. 5.3 concerning the potential selectivity of data linkage.

4.3 Over- and underreporting

Next, we compare UB II program take-up as reported in the survey, with the information in the administrative data. In our data, 11,106 respondents reported benefit receipt in the survey and—based on administrative records—actually received benefits in the month of the interview (take-up households). Additionally, 2277 respondents reported not claiming the benefit, which is confirmed by the information from the administrative data (non-take-up households). A group of 851 respondents did not indicate receipt in the survey, but actually received benefits in the month of the interview based on administrative data (underreporting households). This results in an underreporting rate of 7.1 (7.6)% relative to all true recipients without (with) survey weights being applied. Compared to other benefit programs, this rate is rather low. For example, Meyer et al. (2018) find underreporting rates for food stamp program take-up between 23 and 50% in US surveys. Parolin (2019) confirms high underreporting of the US Food Stamp Program (SNAP) of about 38%: The share of SNAP recipients in the CPS is 11.1% compared to almost 18% in the administrative data. Comparing EU Statistics on Income and Living Conditions (EU-SILC) with administrative data, Tasseva (2016) reports that the number of benefit recipients of important programs in Bulgaria is underreported by between 3 and 40%. For the UK, Brewer et al. (2017) show that particularly, tax credits, means-tested benefits, as well as disability or health benefit have low coverage rates in the Living Costs and Food Survey (LCFS).

In addition to underreporting, 330 benefit-eligible respondents claimed to receive UB II in the survey, but they did not receive benefits according to the administrative records (overreporting households). Various mechanisms may cause overreporting. First, it may result if respondents are mistaken about the period when they actually received the benefit. The administrative data show that for 83 observations (25%) the household received UB II not during the month of the interview, but in the previous or in the following month. Hence, the telescoping bias could be one important source of overreporting in our data. Second, the method of dependent interviewing can also contribute to overreporting when former recipients incorrectly confirm the preloaded information on UB II receipt in the last year (Eggs and Jäckle 2015). Third, overreporting could result from incorrect linkage. However, a closer look at the match types used in the data linkage across the three groups does not reveal large differences: While 96.5% of the observed take-up households are linked by "gold-standard" or deterministic match,[7] this share is 96.4% for the underreporting households and 94.2% for the

[7] The gold-standard matches use an exact match of the household identifier, name, sex, and date of birth. Observations, which cannot be matched by the gold-standard linkage, are matched based on "deterministic

overreporting households. This could indicate that inaccurate answers may be more common in the survey than errors in the matching process.

Table 1 shows descriptive statistics of our explanatory variables used in the regression analysis for the full sample and separately for households with and without benefit program take-up and for those underreporting and overreporting benefit program take-up. The four subgroups differ in their characteristics. Interestingly, we find some similarities between non-take-up households (column 3) and the underreporting households (column 4). Compared to take-up households, the latter two groups have significantly lower simulated benefit entitlements, younger household heads, a lower (higher) share of household heads with lower secondary education (upper secondary education), and a larger share of families with children. This similarity suggests that a take-up regression erroneously classifying underreporting households as non-take-up households overestimates the heterogeneity between the take-up and non-take-up groups, i.e., after correcting underreporting, we expect that the take-up regression yields coefficients of smaller magnitude. An exception are the differences concerning the immigration status of the head of the household: While we find a significantly lower share of immigrants among the non-take-up households compared to take-up households, significantly more household heads who are first-generation immigrants belong to the underreporting households. For the overreporting households, we find several significant differences compared to take-up households. However, the differences in their characteristics make them comparable neither with underreporting households nor with non-take-up households.

The next section describes our analysis results. First, we describe the extent to which benefit program take-up as reported in survey data must be corrected based on information from administrative data. Then, we look at the effect of correcting the dependent variable on the correlation patterns of non-take-up behavior and investigate the robustness of these results. We describe the characteristics of those misreporting benefit receipt in the last step.

5 Results

5.1 Descriptive effects of data correction

In Table 2, we report the simulated group-specific UB II non-take-up rates for the sample that could be linked to administrative records. Column 1 shows the shares before considering corrections for UB II program take-up misreporting, column 2 shows the rates after correction for underreporting only, and column 3 shows the rates after corrections for under- and overreporting. Initially, we observe an overall weighted non-take-up rate of 40% (see bottom of column 1 of Table 2) with substantial heterogeneity across subgroups: We observe the highest rate of non-take-up for couples without children (64%), while single-parent households feature the lowest rates of benefit non-take-up (30%). The size of the non-take-up rate and the variation over

Footnote 7 continued
linkage," which uses first name, last name, zip code, city, street name, house number, sex, and the birth cohort indicator. Both gold-standard and deterministic linkage should result in highly reliable results.

Table 1 Descriptive statistics: covariate means. *Source*: Own calculation based on PASS waves 2–7

	(1) All	(2) Take-up households	(3) Non-take-up households	(4) Misreporting households	(5)
				Underreporting	Overreporting
Simulated entitlement/100 EUR	6.18	7.14	4.67***	5.51***	6.40***
Female hh	0.53	0.48	0.61***	0.55	0.44
Hh is no immigrant	0.77	0.76	0.82***	0.67*	0.78
Hh is immigrant (1.gen.)	0.17	0.18	0.13***	0.26*	0.17
Hh is immigrant (2.gen.)	0.06	0.06	0.05	0.06	0.05
Age of hh: 15–24 years	0.06	0.05	0.07*	0.07**	0.04
Age of hh: 25–34 years	0.19	0.21	0.16**	0.22	0.24
Age of hh 35–44 years	0.23	0.22	0.24	0.29	0.29
Age of hh: 45–54 years	0.28	0.27	0.29	0.28	0.16***
Age of hh: >=55 years	0.24	0.25	0.24	0.15***	0.26
Hh is disabled	0.15	0.15	0.16	0.11	0.25*
Hh holds no sec. degree	0.10	0.10	0.11	0.08	0.19
Hh holds lower sec. degree	0.37	0.41	0.30***	0.34*	0.46
Hh holds intermediate sec. degree	0.33	0.32	0.35	0.33	0.25
Hh holds upper sec. degree	0.20	0.17	0.24***	0.24*	0.10**
Eastern Germany	0.31	0.35	0.24***	0.31	0.36
Household owns home	0.10	0.05	0.19***	0.05	0.03
Monthly rent/100 EUR	3.37	3.35	3.38	3.63	3.24
Young children in household (age <=4 years)	0.08	0.09	0.07*	0.11	0.05**
Single person	0.62	0.66	0.55***	0.61	0.66
Family without children	0.08	0.05	0.14***	0.05	0.03***
Single parents	0.18	0.21	0.12***	0.17	0.23
Family with children	0.12	0.08	0.18***	0.16**	0.09
Subsample two	0.38	0.17	0.77***	0.29***	0.23
N	14,564	11,106	2277	851	330

Asterisks */**/*** denote significantly different means compared to the group of take-up households (column 2) at the significance level of 0.1/0.05/0.01. Hh stands for head of household. "Subsample two" indicates whether an observation belongs to the second, nationally representative subsample. "Take-up households" report UB II in the survey and actually receive UB II according to the administrative data. "Non-take-up households" report non-receipt in the survey and actually do not receive UB II according to the administrative data. "Underreporting households" report UB II non-receipt in the survey and actually receive UB II according to the administrative data. "Overreporting households" report UB II receipt in the survey and actually do not receive UB II according to the administrative data. Weighted values using cross-sectional sample weights the subgroups is in line with findings based on other data (see, e.g., Bruckmeier and Wiemers 2012 and the literature cited there).[8]

We use our administrative data on actual benefit receipt to correct for misreporting of program take-up in the survey.[9] The correction of the underreporting results in a non-take-up rate of 35%, i.e., a reduction by five percentage points (see column 2 of Table 2). If we additionally reclassify the overreporting households as non-take-up households, this increases the non-take-up rate again to 37% (see column 3 of Table 2). Hence, program take-up misreporting reduces the overall non-take-up rate from approximately 40% to approximately 37%. Thus, misreporting caused us to overestimate the UB II non-take-up rate by approximately 3% points, or 8%. The extent of the correction in the non-take-up rate varies across subgroups (see the last two columns of Table 2). The relative decline in non-take-up rates ranges from 2% for disabled heads of the household to approximately 17% for households whose head is a first-generation immigrant.

Figure 1 depicts the distribution of simulated benefit entitlements for those taking up the benefit according to administrative data and reporting it (take-up households) or not reporting it (underreporting household), and those who did not take up the benefit according to administrative data and did not report it in the survey (non-take-up households) or reported it in the survey (overreporting households). As expected, we observe the highest benefits among households who claim their benefits with a median value of 722 euros (see the top panel of Fig. 1). The distributions of benefit entitlements for the non-take-up and underreporting households yield a large share of households with small claims and median claims of 415 and 563 euros, respectively. The median of the overreporting households lies in between with 658 euros.

5.2 Patterns of benefit take-up and the effects of data correction

Table 3 presents the estimation results of our take-up model. We regress a binary indicator of benefit take-up on household characteristics in the sample of 14,564 pooled observations of benefit-eligible households. In column 1, we present the estimated marginal effects of an IV probit estimation with cluster-robust standard errors; in column 5, we show the estimates of a random effects IV probit (RE-IV) estimation; and in columns 2 and 6, both estimation approaches are repeated, now using the corrected dependent variable.

First, the correlations between the first-stage and the second-stage error terms (“$\rho_{1,2}$”) are statistically highly significant for all estimated models. Accordingly, the corresponding Wald tests (“Wald test of exogeneity: $\chi^2(1)$”) strongly reject the null

[8] Although the UB II non-take-up rate is quite similar to the non-take-up rate found in other studies, comparability is limited due to the selection steps described in Appendix 1 and Sect. 4.2, which make our sample no longer representative of the entire population.

[9] Note that the simulation may erroneously predict benefit eligibility for households that underreport other income than benefits. Because these households are not actual benefit recipients, we would overestimate the non-take-up rate. However, their reported benefit receipt would also not be corrected based on linked administrative data. Therefore, these observations are irrelevant to the evaluation of the take-up correction.

Table 2 Group-specific non-take-up rates before and after correction of misreporting. *Source*: Own calculation based on PASS waves 2–7

	(1) Before corrections	(2) After correcting underreporting	(3) After correcting misreporting (under- and overreporting)	(4) After correcting misreporting absolute change (in % points)	(5) Relative change (in %)
Male hh	0.34	0.30	0.32	− 2.5	− 7.3
Female hh	0.46	0.41	0.42	− 3.7	− 8.0
Hh is native	0.42	0.37	0.39	− 2.5	− 6.0
Hh is immigrant (1.gen.)	0.35	0.28	0.29	− 5.8	− 16.6
Hh is immigrant (2.gen.)	0.37	0.32	0.33	− 3.3	− 9.0
Age of hh: 15–24 years	0.51	0.45	0.46	− 4.5	− 8.8
Age of hh: 25–34 years	0.35	0.30	0.32	− 2.6	− 7.4
Age of hh 35–44 years	0.43	0.37	0.39	− 4.0	− 9.4
Age of hh: 45–54 years	0.42	0.37	0.38	− 3.8	− 9.0
Age of hh: >=55 years	0.37	0.34	0.36	− 1.6	− 4.2
Hh not disabled	0.40	0.35	0.37	− 3.5	− 8.7
Hh is disabled	0.41	0.37	0.40	− 0.9	− 2.1
Hh holds no sec. degree	0.41	0.37	0.39	− 1.7	− 4.2
Hh holds lower sec. degree	0.34	0.29	0.31	− 2.2	− 6.4
Hh holds intermediate sec. degree	0.42	0.37	0.39	− 3.1	− 7.3
Hh holds upper sec. degree	0.50	0.43	0.44	− 5.7	− 11.4
Western Germany	0.44	0.39	0.41	− 3.2	− 7.2
Eastern Germany	0.32	0.27	0.29	− 3.0	− 9.3
No young children in household (age <=4 years)	0.41	0.36	0.38	− 3.0	− 7.4
Young children in household (age <=4 years)	0.36	0.30	0.31	− 4.4	− 12.5
Single person	0.36	0.32	0.33	− 2.8	− 7.7
Family without children	0.64	0.60	0.60	− 3.9	− 6.0
Single parents	0.30	0.25	0.27	− 2.5	− 8.4
Family with children	0.60	0.54	0.55	− 5.0	− 8.4
All	0.40	0.35	0.37	− 3.1	− 7.7

Hh stands for head of household. Sample of 14,564 households with simulated UB II entitlements. The total non-take-up rate is reduced from 40.1% (21.5% unweighted) to 37.1 (17.9% unweighted) after correcting misreporting (over- and underreporting of benefit program take-up). Weighted values using cross-sectional sample weights

Table 3 Take-up regression: marginal effects before and after correction of misreporting. *Source*: Own calculation based on PASS waves 2–7

Dependent variable: take-up of UB II	(1) IV probit, uncor.	(2) IV, corrected for misreporting	(3) Abs. diff. (2)-(1)	(4) Rel. diff. (2)/(1) (%)	(5) RE-IV probit, uncor.	(6) RE-IV probit, corrected for misreporting	(7) Abs. diff. (6)-(5)	(8) Rel. diff. (6)/(5) (%)
Simulated entitlement/100 EUR	0.051*** (0.004)	0.043*** (0.004)	– 0.008*** (0.003)	– 15.69	0.047*** (0.004)	0.031*** (0.004)	– 0.016*** (0.003)	– 34.04
Female hh	– 0.008 (0.009)	– 0.004 (0.009)	0.004 (0.006)	– 50.00	– 0.008 (0.009)	– 0.005 (0.006)	0.004 (0.006)	– 37.50
Hh is immigrant (1.gen.)	– 0.014 (0.011)	0.003 (0.011)	0.016* (0.009)	– 121.43	– 0.007 (0.011)	0.013* (0.008)	0.020** (0.008)	– 285.71
Hh is immigrant (2.gen.)	0.030** (0.014)	0.028** (0.014)	– 0.003 (0.011)	– 6.67	0.030** (0.013)	0.021** (0.010)	– 0.009 (0.010)	– 30.00
Age of hh: 25–34 years (ref.: 15–24 years)	0.096*** (0.018)	0.048*** (0.018)	– 0.048*** (0.015)	– 50.00	0.095*** (0.021)	0.038** (0.017)	– 0.058*** (0.016)	– 60.00
Age of hh: 35–44 years (ref.: 15–24 years)	0.119*** (0.019)	0.077*** (0.018)	– 0.042*** (0.015)	– 35.29	0.117*** (0.022)	0.063*** (0.017)	– 0.054*** (0.017)	– 46.15
Age of hh: 45–54 years (ref.: 15–24 years)	0.143*** (0.019)	0.096*** (0.018)	– 0.047*** (0.015)	– 32.87	0.135*** (0.022)	0.073*** (0.017)	– 0.062*** (0.017)	– 45.93
Age of hh: >=55 years (ref.: 15–24 years)	0.192*** (0.020)	0.128*** (0.018)	– 0.065*** (0.015)	– 33.33	0.182*** (0.022)	0.094*** (0.017)	– 0.088*** (0.017)	– 48.35
Hh is disabled	0.000 (0.011)	– 0.022* (0.012)	– 0.022*** (0.008)	–	0.007 (0.010)	– 0.008 (0.008)	– 0.015* (0.008)	– 214.29
Hh holds lower sec. degree (ref. no sec. degree)	– 0.012 (0.014)	– 0.001 (0.015)	0.011 (0.012)	– 91.67	– 0.019 (0.013)	– 0.012 (0.010)	0.007 (0.010)	– 36.84
Hh holds interm. sec. degree (ref. no sec. degree)	– 0.041*** (0.015)	– 0.021 (0.015)	0.020* (0.012)	– 48.78	– 0.049*** (0.013)	– 0.031*** (0.010)	0.017* (0.010)	– 36.73

Table 3 continued

Dependent variable: take-up of UB II	(1) IV probit, uncor.	(2) IV, corrected for misreporting	(3) Abs. diff. (2)-(1)	(4) Rel. diff. (2)/(1) (%)	(5) RE-IV probit, uncor.	(6) RE-IV probit, corrected for misreporting	(7) Abs. diff. (6)-(5)	(8) Rel. diff. (6)/(5) (%)
Hh holds upper sec. degree (ref. no sec. degree)	− 0.076*** (0.016)	− 0.043*** (0.017)	0.032*** (0.012)	− 43.42	− 0.090*** (0.016)	− 0.048*** (0.011)	0.042*** (0.012)	− 46.67
Eastern Germany	0.043*** (0.008)	0.038*** (0.008)	− 0.005 (0.006)	− 11.63	0.042*** (0.008)	0.028*** (0.006)	− 0.014** (0.006)	− 33.33
Household owns home	− 0.088*** (0.020)	− 0.083*** (0.019)	0.005 (0.013)	− 5.68	− 0.102*** (0.022)	− 0.091*** (0.018)	0.011 (0.016)	− 10.78
Monthly rent/100 EUR	− 0.031*** (0.003)	− 0.025*** (0.003)	0.006*** (0.002)	− 19.35	− 0.030*** (0.003)	− 0.021*** (0.002)	0.010*** (0.002)	− 30.00
Young children in household (age <=4 years)	0.049*** (0.012)	0.042*** (0.012)	− 0.006 (0.010)	− 14.29	0.047*** (0.011)	0.044*** (0.008)	− 0.004 (0.009)	− 6.38
Family without children (ref. single person)	− 0.028* (0.016)	− 0.022 (0.015)	0.006 (0.010)	− 21.43	− 0.027 (0.016)	− 0.022* (0.013)	0.005 (0.011)	− 18.52
Single parents (ref. single person)	0.043*** (0.011)	0.036*** (0.011)	− 0.007 (0.008)	− 16.28	0.044*** (0.011)	0.031*** (0.008)	− 0.013 (0.008)	− 29.55
Family with children (ref. single person)	− 0.065*** (0.018)	− 0.058*** (0.018)	0.007 (0.012)	− 10.77	− 0.062*** (0.020)	− 0.051*** (0.016)	0.012 (0.013)	− 17.74
Subsample two	− 0.203*** (0.019)	− 0.216*** (0.020)	− 0.013 (0.011)	6.40	− 0.273*** (0.026)	− 0.289*** (0.027)	− 0.016 (0.016)	5.86
N	14,564	14,564			14,564	14,564		
Log likelihood	− 40,394	− 39,906			− 40,018	− 39,293		
$\rho_{1,2}$	0.34***	0.30***			0.47***	0.52***		

Table 3 continued

Dependent variable: take-up of UB II	(1) IV probit, uncor.	(2) IV, corrected for misreporting	(3) Abs. diff. (2)-(1)	(4) Rel. diff. (2)/(1) (%)	(5) RE-IV probit, uncor.	(6) RE-IV probit, corrected for misreporting	(7) Abs. diff. (6)-(5)	(8) Rel. diff. (6)/(5) (%)
Wald test of exogeneity: $\chi^2(1)$	75.93**	52.33***			119.8***	136.32***		
Panel variance share ρ					0.54***	0.71***		

Asterisks */**/*** denote statistically significant results (standard errors in parentheses) using cluster-robust standard errors at the significance level of 0.1/0.05/0.01. Hh stands for head of household. The correlation of the error terms of the first- and second-stage equations is denoted "$\rho_{1,2}$" and "Wald test of exogeneity:$\chi^2(1)$" shows the test statistic for the null hypothesis $\rho_{1,2} = 0$. For the RE-IV model, "Panel variance share ρ" denotes the share of the total variance contributed by the panel-level variance component. "Subsample two" indicates whether an observation belongs to the second, nationally representative subsample. Survey wave indicators are included in all estimation. Unweighted results

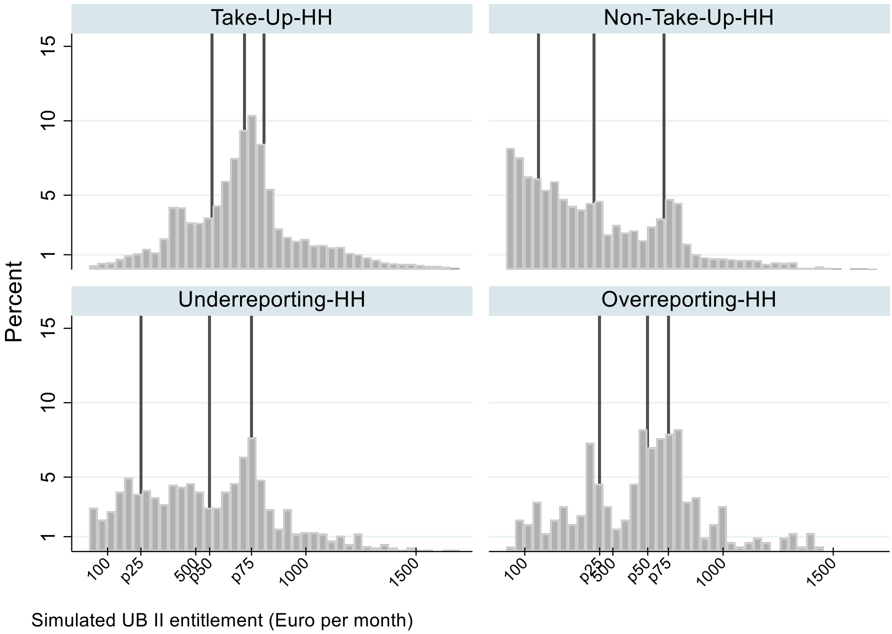

Fig. 1 Distribution of simulated monthly benefit entitlements by type of household (HH). *Notes*: HH stands for household. Fifty-seven outlier observations with monthly entitlements above 1700 euros excluded (51 take-up-HH, 3 non-take-up-HH, 2 underreporting-HH, 1 overreporting-HH). Weighted values using cross-sectional sample weights

hypothesis of exogeneity for the simulated benefit.[10] Thus, we prefer the IV models in Table 3 to the non-IV models presented in Table 9 for all specifications (uncorrected and corrected models). Second, a comparison of the pooled and random effects probit models in columns 1 and 2 versus 5 and 6 reveals the importance of controlling for unobserved heterogeneity at the household level. The results of the random effects estimation in columns 5 and 6 allow us to reject the pooled model of take-up in columns 1 and 2: The share of the total variance ("panel variance share ρ") contributed by the household-level variance reaches 54 and 71% in columns 5 and 6, respectively. It is highly statistically significant at the one percent level in both cases. Therefore, the

[10] In order to examine the validity of the instruments used in the IV estimations (IV probit and RE-IV probit), we also estimate the take-up models using pooled 2SLS, because some validity tests (underidentification and weak instrument tests, in particular) are only available for the linear probability model. All instruments turn out to be highly statistically significant in the first-stage regression ($p < 0.001$). The first-stage R^2 is 0.31, and the partial R^2 for the instruments is 0.12. The Hansen J statistic of overidentifying restrictions has a value of 0.65 with a corresponding p value of 0.23. Thus, the null hypothesis that the instruments are valid (uncorrelated with the error term and correctly excluded from the estimated equation) cannot be rejected. The underidentification test shows that all excluded instruments are relevant in the sense of being correlated with the endogenous regressor. Since we use cluster-robust standard errors, the appropriate test is the Kleibergen and Paap rk LM statistic. With a value of 434.96, the null hypothesis of no correlation with the endogenous regressor is strongly rejected ($p < 0.001$). Finally, the Kleibergen–Paap Wald rk F statistic is 615.04. This result rejects the null hypothesis of weak instruments, since it strongly exceeds the critical value of 13.91 for a maximal test size of 5%. The results for the 2SLS estimations are available from the authors upon request.

RE-IV probit is our overall preferred model when interpreting the marginal effects of the determinants of take-up.

We start by briefly discussing the estimated marginal effects for the models prior to correcting the dependent variable for UB II program take-up misreporting (see columns 1 and 5). In general, the results across the two estimation approaches are somewhat similar and the signs of the estimated marginal effects meet our expectations. The marginal effect of the simulated benefit entitlement in the uncorrected IV probit model (column 1) implies that raising the entitlement by 100 euros per month increases the probability of program take-up by 5.1% points. The RE probit (column 5) results in a smaller marginal effect (4.7% points). The propensity of benefit program take-up is positively correlated with lower education, the advanced age of the head of household, or the presence of young children in the household. The size of the effects is broadly in line with other studies on the determinants of taking up UB II (e.g., Bruckmeier and Wiemers 2012, 2017). Those residing in East Germany also have a higher propensity to take-up benefits than their counterparts in the West, which may reflect the relatively poor labor market situation and lower earnings expectations. Bruckmeier and Wiemers (2017) find similar relationships based on data from the SOEP.

Interestingly, disability status is not correlated with benefit program take-up. We observe significant marginal effects of household composition. Ceteris paribus, the likelihood of claiming the benefit is significantly lower for couples with or without children compared to single person households, i.e., the reference group. This result might reflect the importance of a potential second earner for the program take-up decision in couple households. Additionally, single parents have a higher probability of program take-up.[11]

Next, we study the results when we consider an indicator of benefit take-up corrected for misreporting. We now recode all eligible households who report no receipt in the survey data but who actually receive benefits according to administrative records as take-up households. Conversely, households reporting the benefit in the survey but who do not receive UB II according to the administrative data are reclassified as non-take-up households.

Columns 2 and 6 of Table 3 present the estimated marginal effects, and columns 3 and 7 show the differences between the corrected and uncorrected marginal effects and their statistical significance[12] and columns 4 and 8 present the relative change in

[11] We investigated the marginal effects for various subgroups (e.g., by gender, migration background and for East and West Germany) and did not find major heterogeneities. The results are available upon request.

[12] We test for the significance of the difference between the marginal effects across specifications by estimating a joint covariance matrix of the "stacked" system of equations, i.e., the joint system consisting of the uncorrected and the corrected probit models. Fundamentally, this is a non-standard application of the robust Huber–White covariance estimator. The exact technical implementation of the test differs with respect to our model specifications. For the models which do not include random effects on the household level (pooled IV probit and pooled probit), we use the Stata command "suest" to estimate the joint covariance matrix. For the models including a random effect (RE-IV probit and RE probit), we perform an ML estimation of the uncorrected and corrected RE probit models as a joint system of equations using the Stata command "gsem." In the final step, the difference in the marginal effects and their standard errors are estimated using Stata's "margins" command. The default prediction of the "margins" command does not account for the endogeneity of the simulated benefits, which leads to biased estimates of the marginal effects of the uncorrected and corrected models as well as biased estimates of the differences of the marginal effects. Therefore, we correct for this bias as suggested by Skeels and Taylor (2015).

individual marginal effects when the corrected instead of the original take-up measure is used.

In the pooled IV and the RE-IV estimation, some of the marginal effects change in economically and statistically significant ways when we correct the dependent variable (columns 3 and 7). For both estimators, the largest absolute change results for the indicators of advanced age versus young age and for highly educated versus low educated heads of households. The majority of the marginal effects of the RE-IV estimation that are statistically significant in the uncorrected as well as the corrected model change by more than 30%. Meyer and Mittag (2017a) study the impact of misclassification corrections on probit coefficient estimates in food stamp take-up regressions. They find that estimates are mostly attenuated due to non-random misreporting. We do not confirm this pattern: The marginal effects in columns 1 and 5 of Table 3 appear to be generally larger than the marginal effects obtained after data correction.

Thus, correcting for misreporting affects not only the level of non-take-up but also the impact of the correlates of the take-up decision. Therefore, we empirically corroborate the findings of Pudney (2001) Monte Carlo simulations that even moderate measurement errors in reported benefit receipt can lead to strong biases in coefficient estimates.

5.3 Robustness tests

We offer several types of robustness tests. First, we repeat the estimations in Table 3 using sample weights. The direction of all statistically significant marginal effects is robust to adding sample weights (Table 10). For most variables, the extent of the impact and the significance also remain at a similar level. Our key interest is the effect of correcting the dependent variable. The comparison of the changes in the marginal effects in Table 10 (see columns 3 and 7) with Table 3 shows that most of the results are robust.

Second, we address concerns about the potential selectivity due to (a) missing consent to the data linkage and (b) failed linkage because not all households that agreed to the data linkage could actually be linked (see Sect. 4.2 and Appendix 2). We reestimate our take-up equation for the full sample of eligible households before data linkage and thus without corrections. Table 11 shows the take-up estimation results for the full sample of households that are simulated as eligible to UB II before data linkage. Thus, the results in Table 11 are independent of the linkage procedure and are comparable to the typical situation in which administrative data are not available. Columns 1 and 3 show the results for all eligible households, while columns 2 and 4 again show the results of the linked sample from Table 3. With the exception of the marginal effect for the family without children indicator in the RE-IV model, the significance of the marginal effects does not change. Additionally, the magnitude of all marginal effects changes only slightly. Thus, we find no evidence that there is selection into linkage based on reported receipt.

In two further robustness checks, we analyze whether false matches in the linked data may bias our estimates. In the third robustness check, we only keep observations that were identified in the matching procedure based on the gold-standard or deterministic procedure, i.e., the most reliable matches (see Appendix 2). False matches based on these two procedures are very unlikely. We show the results for the take-up regression for the restricted sample of gold-standard/deterministic matches in Table 12. The comparison with our main findings depicted in Table 3 shows that the statistical significances, the signs, and the magnitude of the marginal effects and the effects of the data correction change only slightly.

In the fourth robustness check, we dropped all non-unique matches, which were corrected after matching (see Sect. 4.2 and Appendix 2). We present the results of our take-up estimation for the sample without these observations in Table 13. Again, we find only minor differences compared to the main results, which are based on all matched observations (Table 3). Thus, the inclusion of the corrected duplicates does not affect our results.

5.4 Patterns of misreporting

Finally, we offer a multivariate characterization of those households who failed to report their benefit program take-up going beyond the descriptive statistics of Tables 1 and 2. Using our baseline specification, we estimate probit models for the outcomes "underreporting" and "overreporting" with pooled and random effect models. Patterns of underreporting are analyzed within the sample of households with UB II program take-up according to the administrative data. For the determinants of overreporting, we rely on observations without registered UB II take-up in the administrative data only.

Table 4 presents the marginal effects of pooled (column 1 and 3) and RE models (column 2 and 4), respectively. We find under- as well as overreporting to be correlated with several characteristics. Particularly, for the indicators of the age of the household head the effects are large and statistically significant, where younger heads are most likely to underreport and less likely to overreport program take-up. Families with children are more likely to underreport and have a lower probability to overreport than individuals in single-person households. Furthermore, households with smaller entitlements or whose head of household is higher educated are more likely to underreport. In contrast, within the sample of households without registered UB II receipt, households with smaller simulated entitlements and a higher educated heads of the household are less likely to overreport. The amount of the monthly rent is positively (negatively) correlated with the probability of underreporting (overreporting). Finally, we find a lower probability of underreporting benefit program take-up for households from East Germany and a higher probability of underreporting for households whose head of household is a first-generation immigrant. Overall, we find that misreporting is not random. This agrees with the finding of Meyer et al. (2018) who show that misreporting is systematically related to household characteristics.

A comparison of the characteristics of underreporting and overreporting households (Table 4) and of the misreporting-induced bias of marginal effects of given characteristics (Table 3) yields similar patterns. We find both strong age dependence in misreporting and a strong response of the misreporting correction on the marginal effect of age. Similarly, we observe statistically significant changes in the marginal

Table 4 Regression of misreporting on household characteristics: marginal effects. *Source*: Own calculation based on PASS waves 2–7

	(1) Pooled probit, underreporting	(2) RE-probit, underreporting	(3) Pooled probit, overreporting	(4) RE-probit, overreporting
Monthly simulated UB II entitlement (in 100 €)	− 0.017*** (0.001)	− 0.018*** (0.001)	0.024*** (0.002)	0.020*** (0.002)
Female hh	− 0.001 (0.006)	− 0.001 (0.006)	− 0.020 (0.017)	− 0.024 (0.016)
Hh is immigrant (1.gen.)	0.030*** (0.008)	0.029*** (0.008)	0.026 (0.024)	0.011 (0.020)
Hh is immigrant (2.gen.)	− 0.006 (0.009)	− 0.007 (0.009)	0.022 (0.030)	0.022 (0.030)
Age of hh: 25–34 years (ref.: 15–24 years)	− 0.050*** (0.015)	− 0.050*** (0.016)	0.077*** (0.021)	0.063*** (0.020)
Age of hh: 35–44 years (ref.: 15–24 years)	− 0.057*** (0.015)	− 0.057*** (0.016)	0.063*** (0.020)	0.061*** (0.020)
Age of hh: 45–54 years (ref.: 15–24 years)	− 0.068*** (0.015)	− 0.068*** (0.016)	0.081*** (0.021)	0.070*** (0.021)
Age of hh: >=55 years (ref.: 15–24 years)	− 0.093*** (0.015)	− 0.094*** (0.016)	0.135*** (0.027)	0.126*** (0.027)
Joint signif. of age (*p* values)	0.0000	0.0000	0.0000	0.0000
Hh is disabled	− 0.011* (0.007)	− 0.011 (0.007)	0.034 (0.022)	0.021 (0.019)
Hh holds lower sec. degree (ref. no sec. degree)	0.010 (0.008)	0.008 (0.009)	− 0.022 (0.038)	− 0.030 (0.035)
Hh holds interm. sec. degree (ref. no sec. degree)	0.015* (0.008)	0.014 (0.009)	− 0.069* (0.037)	− 0.075** (0.034)
Hh holds upper sec. degree (ref. no sec. degree)	0.026*** (0.010)	0.027** (0.010)	− 0.112*** (0.037)	− 0.119*** (0.034)
Joint signif. of educ. (*p* values)	0.0380	0.0460	0.0000	0.0000
Eastern Germany	− 0.013** (0.005)	− 0.014** (0.006)	0.019 (0.017)	0.023 (0.016)
Household owns home	0.024 (0.015)	0.026* (0.016)	− 0.033 (0.025)	− 0.021 (0.028)
Monthly rent/100 EUR	0.013*** (0.002)	0.013*** (0.002)	− 0.016*** (0.005)	− 0.012** (0.005)
Young children in household (age<=4 years)	− 0.011 (0.008)	− 0.011 (0.009)	0.069* (0.037)	0.044 (0.031)
Family without children (ref. single person)	0.011 (0.011)	0.011 (0.011)	− 0.045* (0.024)	− 0.029 (0.025)
Single parents (ref. single person)	− 0.011 (0.007)	− 0.010 (0.007)	0.012 (0.026)	0.009 (0.025)
Family with children (ref. single person)	0.045*** (0.013)	0.044*** (0.013)	− 0.091*** (0.020)	− 0.079*** (0.020)
Joint signif. of hh (*p* values)	0.0000	0.0000	0.0000	0.0000
Subsample two (ref.: SA-recipients-sample)	0.010 (0.010)	0.012 (0.011)	− 0.108*** (0.014)	− 0.108*** (0.013)
Observations	11,957	11,957	2607	2607

Table 4 continued

	(1) Pooled probit, underreporting	(2) RE-probit, underreporting	(3) Pooled probit, overreporting	(4) RE-probit, overreporting
(Pseudo)log-likelihood	− 2780	− 2745	− 819	− 754
Panel variance share ρ		0.37***		0.79***

Asterisks */**/*** denote statistically significant results (standard errors in parentheses) using cluster-robust standard errors at the significance level of 0.1/0.05/0.01. Hh stands for head of household. "Panel variance share ρ" denotes the share of the total variance contributed by the panel-level variance component. "Subsample two" indicates whether an observation belongs to the second, nationally representative subsample. Survey wave indicators are included in all estimation. Unweighted results

effects of the monthly simulated entitlement, monthly rent, first-generation immigrant, upper secondary education, and East Germany after correction for mismeasurement of UB II program take-up. These outcomes are also directly correlated with the misreporting of benefit take-up. Clearly, the estimation of take-up regressions is more reliable for those groups for whom the outcome is measured correctly. Thus, the bias in survey-based estimations of take-up equations varies depending on the extent to which an analysis framework correlates with the propensity to misreport program take-up.

6 Conclusions

This study contributes to the literature on benefit program non-take-up behavior. Because this literature relies on survey data, it suffers from measurement error if respondents do not reveal their true use of welfare benefits. We inspect this issue for the case of a general welfare program using linked representative survey and administrative data. Given that our welfare receiving respondents are aware of the administrative origin of the sampling, misreporting might be smaller here than in other data settings. For households with linked survey and administrative data, we simulate a non-take-up rate based on survey information of 40%, which is in line with results found for comparable benefit programs in other countries (see, e.g., Eurofound 2015). The data linkage shows under- as well as overreporting of benefit program take-up in the survey, whereas we find an underreporting rate of 7.6% in the survey data. Correcting the survey responses for mismeasurement of benefit program take-up (under- and overreporting) reduces the simulated non-take-up rate to 37%.

We use the information on misreporting on program take-up to test whether the results of take-up regressions differ depending on the treatment of misreporting households, i.e., depending on whether we recode them as take-up (underreporting) or non-take-up (overreporting) households in our outcome measure. We estimate pooled and panel instrumental variable probit models and calculate marginal effects. When we compare the estimation results obtained with corrected and uncorrected dependent variables, we find that the absolute difference in marginal effects is often statistically significant and large. In relative terms, many marginal effects change by at least 30%

after the correction. These results hold up to various changes in the empirical approach: We used different estimators (e.g., with and without random effects), and offer a variety of robustness tests to account for potential selectivity problems and sensitivities of the data linkage procedures.

We find that the patterns of the changed marginal effects mostly agree with the correlation patterns underlying misreporting behaviors: Households, whose head of the household is a first-generation immigrant, with small benefit claims and high monthly rents, or with young or highly educated household heads are particularly likely to underreport their benefit receipt. In terms of overreporting, we find an opposite pattern: Most characteristics that positively correlate with underreporting have a negative effect on overreporting and vice versa.

The marginal effects of the characteristics which are correlated with misreporting also changed in a statistically significant way when we corrected our take-up outcome measure for misreporting. This agrees well with the literature showing that households close to the labor market or with a risk of benefit confusion tend to misreport or to not take up their benefit (Bruckmeier and Wiemers 2012, 2017; Bruckmeier et al. 2014). The mechanisms of these groups' underreporting program take-up are different and individually plausible. Because German naturalization rules require that applicants should be able to support themselves and do not rely on social transfers or means-tested benefits, there may be a high perceived cost connected to admitting benefit receipt for immigrants (see, e.g., Riphahn and Saif 2019). Also, households who are close to the labor market, e.g., higher educated and younger persons, may suffer from (perceived) stigma effects and work the hardest to avoid transfer dependence. Those receiving several social transfers at the same time may not be able to keep track of the specific transfer programs from which they benefit, particularly if transfer eligibility changes in short intervals. Overreporting, in contrast, would be more likely among households who confuse different benefit programs or accidentally err on the timing of benefit receipt.

In sum, the analyses based on our linked data suggest that research concerned with take-up and its determinants needs to account for potential misreporting of benefit receipt. The marginal effects in regressions of benefit take-up may well be biased unless the outcome measures can be corrected for misreporting. The bias in survey-based estimations of take-up equations depends on the extent of mismeasurement and the correlates of the propensity to misreport program take-up with the characteristic of the analyzed population.

Take-up analyses are often determined by an interest in distributional effects of government transfers. If the indicators of certain parts of the distribution are more likely mismeasured due to misreporting of benefit receipt, the results for these groups tend to be biased. We are among the first to show such patterns empirically, which are important for the correct interpretation of distributional analyses of government benefits.

Acknowledgements We thank two anonymous referees and the associate editor as well as Mark Trappmann and Jonas Beste for helpful comments on earlier versions of the manuscript.

Funding Open Access funding provided by Projekt DEAL.

Appendix 1: Simulation of welfare benefit eligibility/simulation of entitlement

We base our simulation on waves 2 to 7 of the Panel Study "Labour Market and Social Security" (PASS). In these six waves, 55,069 household interviews were realized. In general, the simulation of welfare benefits using survey data requires several sample selection steps (Table 5).

We first restrict the sample to households for which both household and personal interviews are available. We further drop households headed by individuals over the age of 65 years because they are not eligible according to the rules of the German minimum income support program Unemployment Benefit II (UB II). If households do not answer the question on benefit receipt, we are not able to distinguish between take-up and non-take-up. Therefore, we omit household observations with missing information on benefit receipt. Next, we drop students and apprentices because they must apply for different benefits. Furthermore, we drop households in which not all members participated in the survey. Households in which more than one "community of need" exists are also omitted. The community of need is the legally relevant unit in the means test (for a definition, see Social Code II § 7(2, 3)); in most cases, it is identical with the household. The community of need consists of singles and their partner and their children up to age 24 years. Since we cannot assign household incomes and household wealth across different communities of need, we keep only those households in the sample that consist of exactly one such community.

For the computation of household income, which is relevant to the means test, we need information on household incomes as reported in the household questionnaire and on personal incomes of household members as reported in individual questionnaires. Thus, we must drop households with missing information on these variables. After these sample selection steps, the simulation sample is composed of 30,878 household-year observations.

Our simulation procedure implements the eligibility rules for UB II benefits. A household is eligible if the household's total need exceeds the income and the household's wealth remains below the household-specific maximum. Additionally, we simulate eligibility for means-tested housing allowance and supplementary child allowance ("Kinderzuschlag"), which must be claimed prior to claiming UB II benefits. Including these prioritized benefits in the simulation procedure is important to assess UB II eligibility correctly: The household can only claim UB II if benefits from the housing allowance (possibly in combination with the supplementary child allowance) do not cover its basic needs.

Table 5 Sample selection for the simulation of unemployment benefit II (number of households). *Source*: Own calculation based on PASS waves 2–7

	Wave 2	Wave 3	Wave 4	Wave 5	Wave 6	Wave 7	All
All households	8429	9535	7848	10,235	9513	9509	55,069
Only household interview available	140	190	109	50	58	49	596
Interviews of persons over 65 years	847	939	1164	1766	1781	1840	8337
Missing information on receipt of unemployment benefit II	58	45	68	15	13	17	216
Households of students and apprentices	368	434	352	468	421	414	2457
Households, in which not all individuals responded	1854	2332	1292	1851	1525	1562	10,416
More than one "community of needs" in the household	168	179	164	231	226	203	1171
Missing values in income variables	346	351	257	18	17	9	998
Final simulation sample	4648	5065	4442	5836	5472	5415	30,878

Respondents were interviewed between December and July (waves 2 and 3) and between February and September (waves 4 to 7). Unweighted results

Table 6 Data linkage consent by wave. *Source*: Own calculation based on PASS waves 2–7

Wave	Number of respondents with simulated UB II entitlements who were asked for consent to the data linkage	Number of respondents with simulated UB II entitlements who agreed to the data linkage	Share of respondents who agreed to the data linkage (in percent)
Wave 1	3622	3022	83.4
Wave 2	1161	906	78.0
Wave 3	908	742	81.7
Wave 4	676	595	88.0
Wave 5	1565	1355	86.6
Wave 6	828	686	82.9
Wave 7	589	488	82.9
All	9349	7794	83.4

Respondents who agreed to linkage in wave t were linked in all subsequent waves. Respondents who did not agree to linkage in wave t where asked once again in the subsequent wave. If they declined again, they were not asked about linkage again. Unweighted results

In the first step of our simulation procedure, we calculate the total needs of each household. We determine total needs as the legally defined regular personal needs of household members, additional needs, and housing costs. Next, we consider an additional national standardized benefit for single parents, which varies with the age of their children. The PASS data provide information on the household type and the age of children in the household. Households also report their housing costs. We employ the reported monthly rent for tenants and the reported monthly repayment of mortgage loans for homeowners. Furthermore, we consider reported housing costs for all households. We do not consider additional needs that can be claimed by households in "special circumstances." Special circumstances include certain disabilities, pregnancy, and special dietary needs for health reasons. The incidence of these special benefits is very low: In 2007, approximately 5% of all households received one of these benefits, and the average benefit was 50 euros per month (Statistik der Bundesagentur für Arbeit 2008). Therefore, we can assume that ignoring them will not alter our results considerably.

In the second step of our simulation, we determine the household income relevant to the means test. Household income consists of the sum of all the individual incomes of household members. We apply the earned income exemption rules, which depend on family type and vary with gross wages. All other reported types of income—capital income, rental income, public, and private transfers—reduce the benefit by 100% of their amount.

In the third step, we check whether the household's wealth exceeds the allowable wealth. Each household reports its total financial wealth in each survey wave. The answers are coded in wealth brackets. We assume that the mean of the reported wealth bracket represents the household's wealth and compare it with the individual maximum allowable wealth. The calculation of total allowable maximum wealth depends on household structure and age only. We are not able to consider housing wealth in the simulation due to missing information.

In step four, we calculate entitlements to the housing allowance and the supplementary child allowance and compare the results with those calculated for UB II. If the combined amount of the former two benefits exceeds the UB II benefit, the household is ineligible for UB II.

Following these steps, our simulation classifies 17,585 out of 30,878 households as eligible for UB II. To assess the quality of the simulation we calculate the type II or beta error, i.e., the share of households simulated as ineligible among all households who do report benefit receipt (see, e.g., Bargain et al. 2012, Frick and Groh-Samberg 2007). The share of misclassified households provides an upper bound of the simulation error because beta errors can also be caused by administrative errors in the assessment of eligibility or false answers provided by the respondents in the survey. Our simulation yields a beta error rate of 3.8% (weighted) or 4.6% (unweighted). We interpret these small beta error rates as indicative of a high simulation quality. For comparison, based on the German Socioeconomic Panel, Frick and Groh-Samberg (2007) report a beta error rate of 12.6%.

In addition to quantifiable errors, our simulations may be subject to non-quantifiable errors. These may result if the simulations mispredict transfer eligibility or if the information on household income and wealth is incorrect. If, e.g., respondents underreport household finances, then a simulated benefit eligibility may be wrong and the observed non-take-up rate is overestimated. However, this will not affect our evaluation of household misreporting on transfer receipt, since these households will not appear in the administrative data as benefit recipients.

Appendix 2: Data linkage

In the person interview, the PASS asks respondents aged 15 to 65 years for their consent to link administrative data from the Federal Employment Agency to their survey data. The question reads as follows (English translation, see Sakshaug and Kreuter 2012):

> 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 these additional data?

Overall, in the PASS, the share of respondents who agree to merge their data is approximately 80% (Berg et al. 2014). Our sample of 17,585 simulated UB II eligible household-year observations consists of 8318 different individual households. Table 6 shows the sample of 8318 UB II eligible households how often the consent question was asked in each wave and how many respondents agreed. The number of times the consent question was asked (9349) exceeds the number of eligible household observations, as participants who did not agree to the data linkage during their initial interview are asked once more in the next wave. We find a high average consent rate of approximately 83% in our sample.

In the sample that we use for the analysis, we can link 16,874 household-year observations of simulated eligible respondents to the administrative data, i.e., 96% of our 17,585 UB II eligible household-year observations. Thus, our rate of linkage is substantially higher than the average consent rate. This happens for two reasons: First, a consent given once holds for all future and past waves of the PASS. Second, participants who do not agree to the data linkage are asked again in the next wave. Only if households refuse to give their consent in two consecutive waves is the question no longer repeated in future waves.

In the next step, we merge the observations of respondents who agreed to the linkage to a key file, which identifies respondents in the administrative data. The German Record Linkage Center provides this file, which utilizes several administrative data sources collected by the Federal Employment Agency (BA) (Antoni et al. 2016).

The matching variables used in the linkage are a person's first name, last name, zip code, city, street name, house number, sex, and an indicator for the birth cohort (Antoni et al. 2017). These variables are available in the sampling data and in the administrative data. For the PASS sample drawn from UB II recipients, an additional household identifier is available. The linkage follows a stepwise procedure with variation across the number of matching variables and record linkage processes. Antoni et al. (2017) and Sakshaug et al. (2017) describe the linkage processes. They label a match "gold-standard linkage" if it is based on an exact match of the household identifier, name, sex, and date of birth. This highest quality match is possible only for households in the UB II sample. Observations that cannot be matched by the gold-standard linkage are matched based on "deterministic linkage." This procedure uses first name, last name, zip code, city, street name, house number, sex, and the birth cohort indicator. Both gold-standard and deterministic linkage should result in highly reliable results. For observations that could not be linked using these two procedures, distance-based and probabilistic linkage procedures are used, which match based on comparison functions using first name, last name, zip code, city, street name, house number, sex, and birth cohort.

Table 7 shows the frequency of linkage procedures for our sample of simulated UB II eligible households. From our 15,925 matched observations, 13,089 observations (82%) are linked by the gold-standard match. Adding the 2073 observations which are linked by the deterministic match, our overall share of highly reliable matches (gold standard and deterministic) exceeds 95%. Because of this high share of reliable matches, we consider the overall match quality to be excellent.

Only 574 observations, mainly from the population sample, are matched based on the distance-based procedure. Finally, 189 observations are valid matches, but the type of match is recorded as missing in the data. Since there might be concerns about the reliability of these latter two types of matches, we provide a robustness check

Table 7 Linkage procedures by sample and reported benefit receipt (household-year observations). *Source*: Own calculation based on PASS waves 2–7

	Observations		Sample		UB II receipt	
			UB II	Population	Reported	Not reported
Gold-standard match	13,089	82.2%	13,089	0	2521	10,568
Deterministic match	2073	13.0%	1021	1052	628	1445
Distance-based/probabilistic match	574	3.6%	132	442	247	327
Valid match, unknown match type	189	1.2%	142	47	51	138
All	15,925	100.0%	14,384	1541	3447	12,478

Linkage procedures for the sample of 15,925 simulated UB II eligible household-year observations with consent to data linkage by sample type (columns 3 and 4) and reported UB II receipt (columns 5 and 6). Unweighted results

(Table 12) in which we keep only the gold-standard and deterministic matches in our estimation sample.

In some instances, the matching procedure generated duplicate matches, i.e., a survey observation can have more than one valid match in the administrative data and vice versa. In our linked data, duplicate matches are resolved by choosing one of the duplicate observations based on gender, year of birth, and highest level of education. This affected 830 cases for which two survey respondents were assigned to the same person in the administrative data and 77 cases for which two persons in the administrative data were assigned to one survey respondent. As a robustness check, we reestimated the key results presented in this paper based on a sample in which we dropped all observations with ambiguous, i.e., duplicate matches. All results proved to be robust against this selection step (see Table 13).

One potential problem with the data linkage is that results may be biased because of selectivity in non-consent and non-identifiability in the administrative data. For the sample of simulated eligible households with consent to data linkage, Table 8 shows the correlates of the probability of not giving consent to data linkage and the probability that a household cannot be linked to the administrative data. The results indicate only a minor selection bias concerning the composition of simulated eligible households.

Table 8 Regression of UB II eligible non-consent (1) and non-linkable (2) household-year observations: marginal effects. *Source*: Own calculation based on PASS waves 2–7

Dependent variable	(1) No consent to data linkage	(2) Consent to data linkage but not linkable
Model	Pooled probit	Pooled probit
Monthly simulated SA-entitlement (in 100 €)	− 0.000 (0.001)	− 0.000 (0.001)
Female hh	0.004 (0.005)	0.008 (0.007)
Hh is immigrant	0.016*** (0.006)	0.013* (0.007)
Age of hh: 25–34 years (ref.: 15–24 years)	− 0.019* (0.010)	− 0.023 (0.015)
Age of hh: 35–44 years (ref.: 15–24 years)	− 0.012 (0.010)	− 0.033** (0.016)
Age of hh: 45–54 years (ref.: 15–24 years)	− 0.023** (0.011)	− 0.027* (0.016)
Age of hh: >=55 years (ref.: 15–24 years)	− 0.021** (0.011)	− 0.029* (0.016)
Hh is disabled	− 0.014*** (0.005)	0.001 (0.008)
Hh holds lower sec. degree (ref. no sec. degree)	− 0.007 (0.008)	− 0.016 (0.012)
Hh holds interm. sec. degree (ref. no sec. degree)	− 0.002 (0.008)	− 0.007 (0.012)
Hh holds upper sec. degree (ref. no sec. degree)	0.014 (0.009)	− 0.006 (0.013)
Eastern Germany	− .0115** (0.005)	− 0.010 (0.007)
Household owns home	0.009 (0.011)	0.026** (0.012)
Young children in household (age <=4 years)	− 0.004 (.007)	− 0.010 (.009)
Family without children (ref. single person)	− .023*** (0.007)	0.019 (0.012)
Single parents (ref. single person)	− 0.022*** (0.006)	0.000 (0.009)
Family with children (ref. single person)	− 0.023*** (0.007)	0.003 (0.010)
Subsample two (ref.: SA-recipients-sample)	− 0.006 (0.007)	0.051*** (0.012)

Table 8 continued

Dependent variable	(1) No consent to data linkage	(2) Consent to data linkage but not linkable
Model	Pooled probit	Pooled probit
N	16,357	15,707
(Pseudo)log-likelihood	− 2659	− 3290

Asterisks */**/*** denote statistically significant results (standard errors in parentheses) using cluster-robust standard errors at the significance level of 0.1/0.05/0.01. Hh stands for head of household. "Subsample two" indicates whether an observation belongs to the second, nationally representative subsample. Survey wave indicators are included in all estimation. Column (1) shows the estimates for the consent to data linkage (with 0 = consent (N = 15,707), 1 = no consent (N = 650)) for the sample of 17,585 eligible household-year observations reduced by 1228 observations with missing values in some covariates. Column (2) shows the estimates for being a non-linkable household-year observation (with 0 = linkable (N = 14,834) and 1 = not linkable (N = 873)), for the sample of 16,874 household-year observations that agreed to the data linkage reduced by 1167 observations with missing values in some covariates. Unweighted results

Appendix 3: Regression results—robustness checks

See Tables 9, 10, 11, 12, and 13.

Table 9 Take-up regression: marginal effects before and after correction of misreporting (pooled probit and random effects (RE) probit). *Source*: Own calculation based on PASS waves 2–7

Dependent variable: Take-up of UB II	(1) Pooled probit, uncor.	(2) Pooled probit, corrected for misreporting	(3) Abs. diff. (2)-(1)	(4) Rel. diff. (2)/(1) (%)	(5) RE probit, uncor.	(6) RE probit, corrected for misreporting	(7) Abs. diff. (6)-(5)	(8) Rel. diff. (6)/(5) (%)
Simulated entitlement/100 EUR	0.048*** (0.001)	0.041*** (0.001)	− 0.007*** (0.001)	− 14.58	0.045*** (0.001)	0.031*** (0.001)	− 0.014*** (0.001)	− 31.11
Female hh	− 0.005 (0.008)	− 0.001 (0.008)	0.004 (0.006)	− 80.00	− 0.005 (0.009)	− 0.002 (0.007)	0.003 (0.006)	− 60.00
Hh is immigrant (1.gen.)	− 0.018* (0.011)	− 0.001 (0.011)	0.017* (0.009)	− 94.44	− 0.012 (0.011)	0.010 (0.009)	0.022** (0.009)	− 183.33
Hh is immigrant (2.gen.)	0.029** (0.014)	0.026* (0.014)	− 0.003 (0.011)	− 10.34	0.029** (0.014)	0.021* (0.011)	− 0.009 (0.011)	− 27.59
Age of hh: 25–34 years (ref.: 15–24 years)	0.095*** (0.018)	0.047*** (0.018)	− 0.048*** (0.014)	− 50.53	0.095*** (0.021)	0.039** (0.018)	− 0.056*** (0.016)	− 58.95
Age of hh: 35–44 years (ref.: 15–24 years)	0.116*** (0.019)	0.075*** (0.018)	− 0.041*** (0.014)	− 35.34	0.115*** (0.022)	0.065*** (0.018)	− 0.050*** (0.017)	− 43.48
Age of hh: 45–54 years (ref.: 15–24 years)	0.145*** (0.018)	0.098*** (0.018)	− 0.047*** (0.014)	− 32.41	0.139*** (0.022)	0.080*** (0.018)	− 0.059*** (0.016)	− 42.45
Age of hh: >=55 years (ref.: 15–24 years)	0.199*** (0.019)	0.134*** (0.018)	− 0.066*** (0.014)	− 32.66	0.190*** (0.021)	0.106*** (0.018)	− 0.085*** (0.016)	− 44.21
Hh is disabled	0.003 (0.011)	− 0.019* (0.012)	− 0.022*** (0.008)	− 733.33	0.010 (0.011)	− 0.007 (0.009)	− 0.017** (0.008)	− 170.00
Hh holds lower sec. degree (ref. no sec. degree)	− 0.009 (0.014)	0.001 (0.015)	0.010 (0.012)	− 111.11	− 0.017 (0.013)	− 0.011 (0.011)	0.006 (0.010)	− 35.29

Table 9 continued

Dependent variable: Take-up of UB II	(1) Pooled probit, uncor.	(2) Pooled probit, corrected for misreporting	(3) Abs. diff. (2)-(1)	(4) Rel. diff. (2)/(1) (%)	(5) RE probit, uncor.	(6) RE probit, corrected for misreporting	(7) Abs. diff. (6)-(5)	(8) Rel. diff. (6)/(5) (%)
Hh holds interm. sec. degree (ref. no sec. degree)	– 0.038*** (0.014)	– 0.019 (0.015)	0.020* (0.012)	– 50.00	– 0.047*** (0.014)	– 0.031*** (0.011)	0.016 (0.011)	– 34.04
Hh holds upper sec. degree (ref. no sec. degree)	– 0.073*** (0.016)	– 0.042** (0.016)	0.032*** (0.012)	– 42.47	– 0.089*** (0.016)	– 0.049*** (0.013)	0.040*** (0.012)	– 44.94
Eastern Germany	0.042*** (0.008)	0.037*** (0.008)	– 0.004 (0.006)	– 11.90	0.042*** (0.008)	0.029*** (0.007)	– 0.013** (0.006)	– 30.95
Household owns home	– 0.098*** (0.020)	– 0.091*** (0.018)	0.007 (0.012)	– 7.14	– 0.114*** (0.023)	– 0.103*** (0.020)	0.011 (0.016)	– 9.65
Monthly rent/100 EUR	– 0.031*** (0.002)	– 0.024*** (0.002)	0.006*** (0.002)	– 22.58	– 0.030*** (0.002)	– 0.021*** (0.002)	0.009*** (0.002)	– 30
Young children in household (age<=4 years)	0.049*** (0.012)	0.042*** (0.012)	– 0.007 (0.010)	– 14.29	0.048*** (0.011)	0.044*** (0.008)	– 0.004 (0.009)	– 8.33
Family without children (ref. single person)	– 0.050*** (0.016)	– 0.040*** (0.015)	0.010 (0.010)	– 20.00	– 0.052*** (0.017)	– 0.041*** (0.014)	0.011 (0.012)	– 21.15
Single parents (ref. single person)	0.027** (0.010)	0.023** (0.010)	– 0.004 (0.008)	– 14.81	0.027*** (0.010)	0.021** (0.008)	– 0.006 (0.008)	– 22.22
Family with children (ref. single person)	– 0.157*** (0.018)	– 0.135*** (0.018)	0.022* (0.013)	– 14.01	– 0.165*** (0.022)	– 0.133*** (0.020)	0.032** (0.016)	– 19.39
Subsample two	– 0.206*** (0.017)	– 0.219*** (0.017)	– 0.013 (0.009)	6.31	– 0.278*** (0.023)	– 0.294*** (0.025)	– 0.017 (0.014)	5.76
N	14,564	14,564			14,564	14,564		
Log likelihood	– 5907	– 5405			– 5520	– 4792		

Table 9 continued

Dependent variable: Take-up of UB II	(1) Pooled probit, uncor.	(2) Pooled probit, corrected for misreporting	(3) Abs. diff. (2)-(1)	(4) Rel. diff. (2)/(1) (%)	(5) RE probit, uncor.	(6) RE probit, corrected for misreporting	(7) Abs. diff. (6)-(5)	(8) Rel. diff. (6)/(5) (%)
Panel variance share rho					0.60***	0.74***		

Asterisks */**/*** denote statistically significant results (standard errors in parentheses) using cluster-robust standard errors at the significance level of 0.1/0.05/0.01. Hh stands for head of household. “Panel variance share ρ” denotes the share of the total variance contributed by the panel-level variance component. “Subsample two” indicates whether an observation belongs to the second, nationally representative subsample. Survey wave indicators are included in all estimation. Unweighted results

Table 10 Take-up regression: marginal effects before and after correction of misreporting (weighted results). *Source*: Own calculation based on PASS waves 2–7

Dependent variable: Take-up of UB II	(1) IV probit, uncor.	(2) IV, corrected for misreporting	(3) Abs. diff. (2)-(1)	(4) Rel. diff. (2)/(1) (%)	(5) RE-IV probit, uncor.	(6) RE-IV probit, corrected for misreporting	(7) Abs. diff. (6)-(5)	(8) Rel. diff. (6)/(5) (%)
Simulated entitlement/100 EUR	0.053*** (0.010)	0.045*** (0.010)	− 0.008* (0.004)	− 15.09	0.045*** (0.007)	0.029** (0.013)	− 0.016* (0.010)	− 35.56
Female hh	− 0.044** (0.021)	− 0.033 (0.021)	0.010 (0.010)	− 25.00	− 0.043** (0.020)	− 0.026* (0.014)	0.017 (0.012)	− 39.53
Hh is immigrant (1.gen.)	0.023 (0.024)	0.059** (0.026)	0.035** (0.015)	156.52	0.031 (0.022)	0.060*** (0.020)	0.029* (0.017)	93.55
Hh is immigrant (2.gen.)	0.045 (0.038)	0.044 (0.038)	− 0.001 (0.013)	− 2.22	0.054 (0.039)	0.044 (0.031)	− 0.009 (0.019)	− 18.52
Age of hh: 25–34 years (ref.: 15–24 years)	0.116*** (0.033)	0.090*** (0.034)	− 0.026 (0.018)	− 22.41	0.106*** (0.031)	0.058** (0.026)	− 0.048** (0.024)	− 45.28
Age of hh: 35–44 years (ref.: 15–24 years)	0.121*** (0.033)	0.111*** (0.034)	− 0.011 (0.020)	− 8.26	0.124*** (0.031)	0.077** (0.033)	− 0.047 (0.032)	− 37.90
Age of hh: 45–54 years (ref.: 15–24 years)	0.131*** (0.035)	0.121*** (0.036)	− 0.010 (0.019)	− 7.63	0.114*** (0.033)	0.067** (0.032)	− 0.047* (0.028)	− 41.23
Age of hh: >=55 years (ref.: 15–24 years)	0.183*** (0.037)	0.148*** (0.038)	− 0.035* (0.019)	− 19.13	0.158*** (0.034)	0.086** (0.037)	− 0.073** (0.031)	− 45.57
Hh is disabled	− 0.027 (0.023)	− 0.045* (0.024)	− 0.018 (0.012)	66.67	− 0.006 (0.021)	− 0.014 (0.014)	− 0.008 (0.014)	133.33
Hh holds lower sec. degree (ref. no sec. degree)	0.042 (0.034)	0.053 (0.037)	0.011 (0.017)	26.19	0.016 (0.033)	0.016 (0.022)	− 0.000 (0.018)	0.00
Hh holds interm. sec. degree (ref. no sec. degree)	− 0.005 (0.034)	0.016 (0.037)	0.020 (0.017)	− 420.00	− 0.033 (0.033)	− 0.020 (0.022)	0.012 (0.019)	− 39.39
Hh holds upper sec. degree (ref. no sec. degree)	− 0.046 (0.036)	− 0.010 (0.041)	0.036* (0.021)	− 78.26	− 0.080** (0.035)	− 0.040 (0.025)	0.040* (0.023)	− 50.00
Eastern Germany	0.057*** (0.019)	0.056*** (0.020)	− 0.001 (0.011)	− 1.75	0.063*** (0.021)	0.051*** (0.020)	− 0.012 (0.017)	− 19.05

Table 10 continued

Dependent variable: Take-up of UB II	(1) IV probit, uncor.	(2) IV, corrected for misreporting	(3) Abs. diff. (2)-(1)	(4) Rel. diff. (2)/(1) (%)	(5) RE-IV probit, uncor.	(6) RE-IV probit, corrected for misreporting	(7) Abs. diff. (6)-(5)	(8) Rel. diff. (6)/(5) (%)
Household owns home	− 0.113*** (0.038)	− 0.112*** (0.040)	0.001 (0.018)	− 0.88	− 0.121*** (0.040)	− 0.095** (0.043)	0.027 (0.036)	− 21.49
Monthly rent/100 EUR	− 0.033*** (0.006)	− 0.025*** (0.006)	0.008*** (0.003)	− 24.24	− 0.028*** (0.005)	− 0.019*** (0.007)	0.010* (0.006)	− 32.14
Young children in household (age <=4 years)	0.096*** (0.029)	0.111*** (0.032)	0.015 (0.021)	15.63	0.098*** (0.025)	0.109*** (0.026)	0.011 (0.021)	11.22
Family without children (ref. single person)	− 0.110*** (0.038)	− 0.112*** (0.039)	− 0.002 (0.013)	1.82	− 0.110*** (0.041)	− 0.087** (0.039)	0.023 (0.030)	− 20.91
Single parents (ref. single person)	0.071*** (0.025)	0.041 (0.027)	− 0.030* (0.017)	− 42.25	0.058* (0.030)	0.020 (0.024)	− 0.037** (0.017)	− 65.52
Family with children (ref. single person)	− 0.155*** (0.040)	− 0.157*** (0.039)	− 0.002 (0.023)	1.29	− 0.154*** (0.046)	− 0.118*** (0.044)	0.035 (0.038)	− 23.38
Subsample two	− 0.406*** (0.062)	− 0.417*** (0.070)	− 0.011 (0.034)	2.71	− 0.584*** (0.057)	− 0.711*** (0.125)	− 0.127 (0.099)	21.75
N	14,564	14,564			14,564	14,564		
Log likelihood	− 55,439,604	− 55,463,567			− 54,565,917	− 54,188,814		
$\rho_{1,2}$	0.42***	0.30**			0.65***	0.65***		
Wald test of exogeneity: $\chi 2(1)$	20.32***	8.27***			99.35***	48.31***		
Panel variance share ρ					0.61***	0.81***		

Asterisks */**/*** denote statistically significant results (standard errors in parentheses) using cluster-robust standard errors at the significance level of 0.1/0.05/0.01. Hh stands for head of household. The correlation of the error terms of the first- and second-stage equations is denoted "$\rho_{1,2}$" and "Wald test of exogeneity: $\chi^2(1)$" shows the test statistic for the null hypothesis $\rho_{1,2} = 0$. For the RE-IV model, "Panel variance share ρ" denotes the share of the total variance contributed by the panel-level variance component. "Subsample two" indicates whether an observation belongs to the second, nationally representative subsample. Survey wave indicators are included in all estimation. Weighted results

Table 11 Take-up regression: marginal effects of linked sample versus unlinked sample (uncorrected take-up). *Source*: Own calculation based on PASS waves 2–7. Unweighted results

Dependent variable Take-up of UB II	(1) IV probit, linked	(2) IV probit, not linked	(3) RE-IV probit, linked	(4) RE-IV probit, not linked
Simulated entitlement/100 EUR	0.051*** (0.004)	0.050*** (0.003)	0.047*** (0.004)	0.047*** (0.003)
Female hh	− 0.008 (0.009)	− 0.010 (0.008)	− 0.008 (0.009)	− 0.012 (0.008)
Hh is immigrant (1.gen.)	− 0.014 (0.011)	− 0.006 (0.010)	− 0.007 (0.011)	0.003 (0.010)
Hh is immigrant (2.gen.)	0.030** (0.014)	0.030** (0.014)	0.030** (0.013)	0.033*** (0.013)
Age of hh: 25− 34 years (ref.: 15–24 years)	0.096*** (0.018)	0.088*** (0.017)	0.095*** (0.021)	0.086*** (0.020)
Age of hh: 35–44 years (ref.: 15–24 years)	0.119*** (0.019)	0.114*** (0.018)	0.117*** (0.022)	0.112*** (0.020)
Age of hh: 45–54 years (ref.: 15–24 years)	0.143*** (0.019)	0.135*** (0.018)	0.135*** (0.022)	0.127*** (0.020)
Age of hh: >=55 years (ref.: 15–24 years)	0.192*** (0.020)	0.187*** (0.019)	0.182*** (0.022)	0.176*** (0.020)
Hh is disabled	0.000 (0.011)	0.002 (0.011)	0.007 (0.010)	0.010 (0.010)
Hh holds lower sec. degree (ref. no sec. degree)	− 0.012 (0.014)	− 0.011 (0.013)	− 0.019 (0.013)	− 0.016 (0.012)
Hh holds interm. sec. degree (ref. no sec. degree)	− 0.041*** (0.015)	− 0.043*** (0.014)	− 0.049*** (0.013)	− 0.048*** (0.013)
Hh holds upper sec. degree (ref. no sec. degree)	− 0.076*** (0.016)	− 0.081*** (0.015)	− 0.090*** (0.016)	− 0.094*** (0.015)
Eastern Germany	0.043*** (0.008)	0.048*** (0.008)	0.042*** (0.008)	0.049*** (0.007)
Household owns home	− 0.088*** (0.020)	− 0.102*** (0.019)	− 0.102*** (0.022)	− 0.118*** (0.022)
Monthly rent/100 EUR	− 0.031*** (0.003)	− 0.030*** (0.003)	− 0.030*** (0.003)	− 0.029*** (0.003)
Young children in household (age <=4 years)	0.049*** (0.012)	0.049*** (0.012)	0.047*** (0.011)	0.049*** (0.011)
Family without children (ref. single person)	− 0.028* (0.016)	− 0.036** (0.015)	− 0.027 (0.016)	− 0.037** (0.016)
Single parents (ref. single person)	0.043*** (0.011)	0.040*** (0.011)	0.044*** (0.011)	0.041*** (0.010)
Family with children (ref. single person)	− 0.065*** (0.018)	− 0.068*** (0.017)	− 0.062*** (0.020)	− 0.065*** (0.019)
Subsample two	− 0.203*** (0.019)	− 0.223*** (0.019)	− 0.273*** (0.026)	− 0.304*** (0.025)

Table 11 continued

Dependent variable Take-up of UB II	(1) IV probit, linked	(2) IV probit, not linked	(3) RE-IV probit, linked	(4) RE-IV probit, not linked
N	14,564	16,357	14,564	16,357
Log likelihood	− 40,394	− 45,540	− 40,018	− 45,091
$\rho_{1,2}$	0.34***	0.33***	0.47***	0.45***
Wald test of exogeneity: $\chi 2(1)$	75.93**	82.94***	119.8***	124.15***
Panel variance share ρ			0.54***	0.56***

Asterisks */**/*** denote statistically significant results (standard errors in parentheses) using cluster-robust standard errors at the significance level of 0.1/0.05/0.01. Hh stands for head of household. The correlation of the error terms of the first- and second-stage equations is denoted "$\rho_{1,2}$" and "Wald test of exogeneity: $\chi^2(1)$" shows the test statistic for the null hypothesis $\rho_{1,2} = 0$. For the RE-IV model, "Panel variance share ρ" denotes the share of the total variance contributed by the panel-level variance component. "Subsample two" indicates whether an observation belongs to the second, nationally representative subsample. Columns 1 and 3 correspond to columns 1 and 5 in Table 3 and are repeated for convenience. Columns 2 and 4 show the estimates for the 17,585 household-year observations simulated as eligible for UB II and reduced by 1228 observations with missing values in covariates. Unweighted results

Table 12 Take-up regression: marginal effects when using only gold-standard/deterministic links. *Source*: Own calculation based on PASS waves 2–7

Dependent variable: Take-up of UB II	(1) IV probit, uncor.	(2) IV, corrected for misreporting	(3) Abs. diff. (2)-(1)	(4) Rel. diff. (2)/(1) (%)	(5) RE-IV probit, uncor.	(6) RE-IV probit, corrected for misreporting	(7) Abs. diff. (6)-(5)	(8) Rel. diff. (6)/(5) (%)
Simulated entitlement/100 EUR	0.050*** (0.004)	0.042*** (0.004)	− 0.008*** (0.003)	− 16.00	0.046*** (0.004)	0.030*** (0.004)	− 0.016*** (0.003)	− 34.78
Female hh	− 0.002 (0.009)	0.002 (0.009)	0.004 (0.007)	− 200.00	− 0.003 (0.009)	− 0.001 (0.006)	0.002 (0.006)	− 66.67
Hh is immigrant (1.gen.)	− 0.017 (0.011)	− 0.000 (0.011)	0.016* (0.010)	− 100.00	− 0.011 (0.011)	0.010 (0.008)	0.021** (0.009)	− 190.91
Hh is immigrant (2.gen.)	0.028** (0.014)	0.031** (0.013)	0.003 (0.010)	10.71	0.026** (0.013)	0.021** (0.010)	− 0.006 (0.010)	− 19.23
Age of hh: 25–34 years (ref.: 15–24 years)	0.094*** (0.019)	0.046*** (0.018)	− 0.048*** (0.015)	− 51.06	0.091*** (0.021)	0.033** (0.017)	− 0.058*** (0.016)	− 63.74
Age of hh: 35–44 years (ref.: 15–24 years)	0.119*** (0.019)	0.075*** (0.018)	− 0.043*** (0.015)	− 36.97	0.114*** (0.022)	0.059*** (0.017)	− 0.055*** (0.017)	− 48.25
Age of hh: 45–54 years (ref.: 15–24 years)	0.142*** (0.019)	0.095*** (0.018)	− 0.046*** (0.015)	− 33.10	0.131*** (0.022)	0.069*** (0.017)	− 0.062*** (0.017)	− 47.33
Age of hh: >=55 years (ref.: 15–24 years)	0.193*** (0.020)	0.126*** (0.018)	− 0.067*** (0.015)	− 34.72	0.178*** (0.022)	0.090*** (0.017)	− 0.088*** (0.017)	− 49.44
Hh is disabled	0.006 (0.011)	− 0.015 (0.012)	− 0.021** (0.009)	− 350.00	0.009 (0.010)	− 0.006 (0.008)	− 0.015* (0.008)	− 166.67
Hh holds lower sec. degree (ref. no sec. degree)	− 0.007 (0.014)	0.002 (0.015)	0.009 (0.012)	− 128.57	− 0.015 (0.013)	− 0.011 (0.009)	0.005 (0.010)	− 26.67
Hh holds interm. sec. degree (ref. no sec. degree)	− 0.036** (0.015)	− 0.017 (0.016)	0.018 (0.012)	− 52.78	− 0.042*** (0.013)	− 0.028*** (0.010)	0.014 (0.010)	− 33.33

Table 12 continued

Dependent variable: Take-up of UB II	(1) IV probit, uncor.	(2) IV, corrected for misreporting	(3) Abs. diff. (2)-(1)	(4) Rel. diff. (2)/(1) (%)	(5) RE-IV probit, uncor.	(6) RE-IV probit, corrected for misreporting	(7) Abs. diff. (6)-(5)	(8) Rel. diff. (6)/(5) (%)
Hh holds upper sec. degree (ref. no sec. degree)	− 0.072*** (0.017)	− 0.040** (0.017)	0.031** (0.013)	− 44.44	− 0.085*** (0.016)	− 0.045*** (0.011)	0.040*** (0.012)	− 47.06
Eastern Germany	0.037*** (0.008)	0.034*** (0.008)	− 0.003 (0.006)	− 8.11	0.036*** (0.008)	0.024*** (0.006)	− 0.012** (0.006)	− 33.33
Household owns home	− 0.061*** (0.020)	− 0.060*** (0.018)	0.001 (0.013)	− 1.64	− 0.068*** (0.021)	− 0.063*** (0.016)	0.005 (0.015)	− 7.35
Monthly rent/100 EUR	− 0.031*** (0.003)	− 0.025*** (0.003)	0.006*** (0.002)	− 19.35	− 0.030*** (0.003)	− 0.020*** (0.002)	0.010*** (0.002)	− 33.33
Young children in household (age <=4 years)	0.047*** (0.013)	0.041*** (0.012)	− 0.006 (0.010)	− 12.77	0.044*** (0.011)	0.041*** (0.007)	− 0.003 (0.009)	− 6.82
Family without children (ref. single person)	− 0.023 (0.016)	− 0.017 (0.015)	0.007 (0.011)	− 26.09	− 0.020 (0.016)	− 0.015 (0.012)	0.005 (0.012)	− 25.00
Single parents (ref. single person)	0.040*** (0.011)	0.034*** (0.011)	− 0.006 (0.008)	− 15.00	0.039*** (0.011)	0.029*** (0.008)	− 0.011 (0.008)	− 25.64
Family with children (ref. single person)	− 0.059*** (0.019)	− 0.055*** (0.018)	0.004 (0.012)	− 6.78	− 0.053*** (0.020)	− 0.047*** (0.016)	0.007 (0.014)	− 11.32
Subsample two	− 0.197*** (0.022)	− 0.207*** (0.022)	− 0.011 (0.013)	5.08	− 0.258*** (0.029)	− 0.266*** (0.030)	− 0.007 (0.018)	3.10
N	13,878	13,878			13,878	13,878		
Log likelihood	− 38,242	− 37,742			− 37,913	− 37,200		
$\rho_{1,2}$	0.35***	0.29***			0.49***	0.51***		
Wald test of exogeneity: $\chi 2(1)$	77.02***	48.25***			130.69***	120.43***		

Table 12 continued

Dependent variable: Take-up of UB II	(1) IV probit, uncor.	(2) IV, corrected for misreporting	(3) Abs. diff. (2)-(1)	(4) Rel. diff. (2)/(1) (%)	(5) RE-IV probit, uncor.	(6) RE-IV probit, corrected for misreporting	(7) Abs. diff. (6)-(5)	(8) Rel. diff. (6)/(5) (%)
Panel variance share ρ					0.53***	0.70***		

Asterisks */**/*** denote statistically significant results (standard errors in parentheses) using cluster-robust standard errors at the significance level of 0.1/0.05/0.01. Hh stands for head of household. The correlation of the error terms of the first- and second-stage equations is denoted "$\rho_{1,2}$" and "Wald test of exogeneity: $\chi^2(1)$" shows the test statistic for the null hypothesis $\rho_{1,2} = 0$. For the RE-IV model, "Panel variance share ρ" denotes the share of the total variance contributed by the panel-level variance component. "Subsample two" indicates whether an observation belongs to the second, nationally representative subsample. Survey wave indicators are included in all estimation. Unweighted results

Table 13 Take-up regression: marginal effects for sample without corrected duplicates. *Source*: Own calculation based on PASS waves 2–7

Dependent variable: Take-up of UB II	(1) IV probit, uncor.	(2) IV, corrected for misreporting	(3) Abs. diff. (2)-(1)	(4) Rel. diff. (2)/(1) (%)	(5) RE-IV probit, uncor.	(6) RE-IV probit, corrected for misreporting	(7) Abs. diff. (6)-(5)	(8) Rel. diff. (6)/(5) (%)
Simulated entitlement/100 EUR	0.050*** (0.004)	0.043*** (0.004)	− 0.007*** (0.003)	− 14.00	0.047*** (0.004)	0.032*** (0.004)	− 0.015*** (0.003)	− 31.91
Female hh	− 0.009 (0.009)	− 0.003 (0.009)	0.006 (0.007)	− 66.67	− 0.010 (0.009)	− 0.004 (0.007)	0.006 (0.006)	− 60.00
Hh is immigrant (1.gen.)	− 0.010	0.007 (0.011)	0.017* (0.009)	− 170.00	− 0.004 (0.011)	0.016** (0.008)	0.020** (0.008)	− 500.00
Hh is immigrant (2.gen.)	0.027* (0.014)	0.028** (0.014)	0.001 (0.011)	3.70	0.028** (0.013)	0.021** (0.010)	− 0.007 (0.010)	− 25.00
Age of hh: 25–34 years (ref.: 15–24 years)	0.095*** (0.019)	0.048*** (0.018)	− 0.047*** (0.015)	− 49.47	0.094*** (0.021)	0.036** (0.017)	− 0.058*** (0.016)	− 61.70
Age of hh: 35–44 years (ref.: 15–24 years)	0.120*** (0.020)	0.076*** (0.018)	− 0.044*** (0.015)	− 36.67	0.117*** (0.022)	0.061*** (0.017)	− 0.057*** (0.017)	− 47.86
Age of hh: 45–54 years (ref.: 15–24 years)	0.145*** (0.020)	0.095*** (0.018)	− 0.050*** (0.015)	− 34.48	0.137*** (0.022)	0.071*** (0.017)	− 0.067*** (0.017)	− 48.18
Age of hh: >=55 years (ref.: 15–24 years)	0.194*** (0.020)	0.128*** (0.019)	− 0.066*** (0.015)	− 34.02	0.184*** (0.022)	0.096*** (0.017)	− 0.088*** (0.017)	− 47.83
Hh is disabled	0.000 (0.011)	− 0.022* (0.012)	− 0.022** (0.009)	–	0.007 (0.011)	− 0.008 (0.008)	− 0.015* (0.008)	− 214.29
Hh holds lower sec. degree (ref. no sec. degree)	− 0.009 (0.014)	− 0.003 (0.015)	0.006 (0.011)	− 66.67	− 0.018 (0.013)	− 0.014 (0.010)	0.004 (0.010)	− 22.22
Hh holds interm. sec. degree (ref. no sec. degree)	− 0.040*** (0.015)	− 0.023 (0.015)	0.017 (0.012)	− 42.50	− 0.047*** (0.014)	− 0.032*** (0.010)	0.015 (0.010)	− 31.91

Table 13 continued

Dependent variable: Take-up of UB II	(1) IV probit, uncor.	(2) IV, corrected for misreporting	(3) Abs. diff. (2)-(1)	(4) Rel. diff. (2)/(1) (%)	(5) RE-IV probit, uncor.	(6) RE-IV probit, corrected for misreporting	(7) Abs. diff. (6)-(5)	(8) Rel. diff. (6)/(5) (%)
Hh holds upper sec. degree (ref. no sec. degree)	− 0.074*** (0.017)	− 0.043*** (0.017)	0.030** (0.012)	− 41.89	− 0.087*** (0.016)	− 0.048*** (0.012)	0.039*** (0.012)	− 44.83
Eastern Germany	0.041*** (0.008)	0.036*** (0.008)	− 0.006 (0.006)	− 12.20	0.042*** (0.008)	0.027*** (0.006)	− 0.015** (0.006)	− 35.71
Household owns home	− 0.090*** (0.020)	− 0.084*** (0.019)	0.006 (0.013)	− 6.67	− 0.104*** (0.023)	− 0.091*** (0.018)	0.013 (0.016)	− 12.50
Monthly rent/100 EUR	− 0.031*** (0.003)	− 0.024*** (0.003)	0.007*** (0.002)	− 22.58	− 0.030*** (0.003)	− 0.020*** (0.002)	0.010*** (0.002)	− 33.33
Young children in household (age <=4 years)	0.048*** (0.013)	0.044*** (0.012)	− 0.004 (0.010)	− 8.33	0.046*** (0.012)	0.043*** (0.008)	− 0.003 (0.009)	− 6.52
Family without children (ref. single person)	− 0.024 (0.016)	− 0.016 (0.015)	0.009 (0.010)	− 33.33	− 0.023 (0.016)	− 0.017 (0.013)	0.006 (0.012)	− 26.09
Single parents (ref. single person)	0.047*** (0.011)	0.041*** (0.011)	− 0.006 (0.008)	− 12.77	0.047*** (0.011)	0.034*** (0.008)	− 0.013 (0.008)	− 27.66
Family with children (ref. single person)	− 0.070*** (0.019)	− 0.067*** (0.019)	0.003 (0.012)	− 4.29	− 0.066*** (0.021)	− 0.059*** (0.017)	0.008 (0.014)	− 10.61
Subsample two	− 0.211*** (0.020)	− 0.227*** (0.020)	− 0.016 (0.011)	7.58	− 0.283*** (0.027)	− 0.303*** (0.028)	− 0.020 (0.017)	7.07
N	13,940	13,940			13,940	13,940		
Log likelihood	− 38,625	− 38,141			− 38,273	− 37,571		
$\rho_{1,2}$	0.33***	0.30***			0.45***	0.51***		

Table 13 continued

Dependent variable: Take-up of UB II	(1) IV probit, uncor.	(2) IV, corrected for misreporting	(3) Abs. diff. (2)-(1)	(4) Rel. diff. (2)/(1) (%)	(5) RE-IV probit, uncor.	(6) RE-IV probit, corrected for misreporting	(7) Abs. diff. (6)-(5)	(8) Rel. diff. (6)/(5) (%)
Wald test of exogeneity: $\chi 2(1)$	67.88***	50.45***			102.96***	125.45***		
Panel variance share ρ					0.54***	0.70***		

Asterisks */**/*** denote statistically significant results (standard errors in parentheses) using cluster-robust standard errors at the significance level of 0.1/0.05/0.01. Hh stands for head of household. The correlation of the error terms of the first- and second-stage equations is denoted "$\rho_{1,2}$" and "Wald test of exogeneity: $\chi^2(1)$" shows the test statistic for the null hypothesis $\rho_{1,2} = 0$. For the RE-IV model, "Panel variance share ρ" denotes the share of the total variance contributed by the panel-level variance component. "Subsample two" indicates whether an observation belongs to the second, nationally representative subsample. Survey wave indicators are included in all estimation. Unweighted results

Appendix 4: basic characteristics of different social programs

See Table 14.

Table 14 Basic characteristics of Food Stamp Program (FSP)/Supplement Nutrition Assistance Program (SNAP), Medicaid Program, and Unemployment Benefit II (UB II). *Sources*: https://www.fns.usda.gov/snap/supplemental-nutrition-assistance-program, https://www.ers.usda.gov/webdocs/publications/94849/err-270.pdf?v=963.1, https://www.medicaid.gov/medicaid, Statistik der Bundesagentur für Arbeit: Tabellen, Strukturen der Grundsicherung SGB II (Zeitreihe Monats- und Jahreszahlen ab 2005), Nürnberg, Januar 2020, www.destatis.de

	FSP/SNAP (US)	Medicaid (US)	UB II (GER)
Target group	Needy families/households	Needy individuals under age 65 years, qualified pregnant women and children. State-specific rules to cover additional groups	Persons able to work and their families in need
Objective	Supplement the food budget	Provide health coverage	Guarantee a minimum income
Benefit	Electronic Benefit Transfer Card	Various healthcare services (nationwide standard). Further services possible at state level	Financial benefits to cover housing and living costs, plus health insurance
Calculation of benefit	30% of calculated net income up to a household-specific maximum	–	Nationally uniform benefit per household member plus expenditures for rent and heating
Income test	Gross monthly income below the legally defined threshold of 130% of the federal poverty level and net monthly income below 100% of the federal poverty line. Higher-income limits for elderly or disabled persons	Modified adjusted gross income below a state-specific percentage of the federal poverty level	Household net income below legally defined household minimum income (includes living and housing costs)

Table 14 continued

	FSP/SNAP (US)	Medicaid (US)	UB II (GER)
Income deductions	Deductions in special circumstances (e.g., care deduction needed for education). Shelter costs that are more than half of the household's income after other deductions	Deductions for certain expenses (e.g., contributions to an individual retirement account, payment of student loan interest)	Almost all incomes are considered in the means test
Earned income deductions	20% of earned income	–	First €100 of net wages free, then 20% (10%) of net wage income for gross wages between €101 (€1001) and €1000 (€1200 for households without children/€1500 for households with children) remains
Number of recipients	40.4 million individuals (12% of total population, annual average 2018)	64.7 million individuals (19.6% of total population, October 2019)	6.1 million individuals (7.3% of total population, annual average 2018)
Average benefit 2018	$126 per person per month (annual average 2018)	–	€511 per person per month (annual average 2018)
Wealth test	$2250 ($3500) per household (with at least one member aged >=60 years). Exemptions for resources of SSI and TANF recipients, retirement and pension plans, house and land	–	€150 per year of life for each full-aged individual, minimum: €3100, maximum: €10,500. €3100 for each child under 18 years
Work requirements	Yes, with exemption for certain groups	In most states no work requirements, but varying rules across states	Yes, with exemption for certain groups

References

Antoni M, Ganzer A, vom Berge P. Sample of integrated labour market biographies (SIAB) 1975-2014. FDZ-Datenreport, 04/2016 (en)), Nürnberg, 2016.

Antoni M, Schnell R. The past, present and future of the German Record Linkage Center (GRLC). J Econ Stat. 2019;239(2):1–13.

Antoni M, Dummert S, Trenkle S. PASS-Befragungsdaten verknüpft mit administrativen Daten des IAB (PASS-ADIAB) 1975-2015. FDZ-Datenreport, 06/2017 (de), Nürnberg, 2017.

Bargain O, Immervoll H, Viitamäki H. No claim, no pain. Measuring the non-take-up of social assistance using register data. J Econ Inequal. 2012;10:375–95.

Berg M, Cramer R, Dickmann C, Gilberg R, Jesske B, Kleudgen M, Bethmann A, Fuchs B, Huber M, Trappmann M. Codebook and documentation of the panel study "Labour market and social security" (PASS). Datenreport wave 7. Nürnberg, 2014.

Berg M, Cramer R, Dickmann C, Gilberg R, Jesske B, Kleudgen M, Bethmann A, Fuchs M, Trappmann M, Wurdack A. Codebook and documentation of the panel study "Labour market and social security" (PASS). Datenreport wave 5. Nürnberg, 2012.

Beste J, Grabka MM, Goebel J. Armut in Deutschland: Ein Vergleich zwischen den beiden Haushaltspanelstudien SOEP und PASS. AStA AStA Wirtsch Sozialstat Arch. 2018;12:27–62.

Blank R, Ruggles P. When do women use aid to families with dependent children and food stamps? The dynamics of eligibility versus participation. J Hum Resour. 1996;31(1):57–89.

Blundell R, Fry V, Walker I. Modelling the take-up of means-tested benefits: the case of housing benefits in the UK. Econ J. 1988;98:58–74.

Bollinger CR, David MH. Modeling discrete choice with response error: food stamp participation. J Am Stat Assoc. 1997;92(439):827–35.

Bollinger CR, David MH. Estimation with response error and nonresponse: food-stamp participation in the SIPP. J Bus Econ Stat. 2001;19(2):129–41.

Bound J, Brown C, Mathiowetz N. Measurement error in survey data. In: Heckman JJ, Leamer E, editors. Handbook of econometrics, vol. 5. Amsterdam: Elsevier Science; 2001. p. 3705–843.

Bradburn NM, Huttenlocher J, Hedges L. Telescoping and temporal memory. In: Schwarz N, Sudman S, editors. Autobiographical memory and the validity of retrospective reports. New York: Springer; 1994. p. 203–15.

Brewer M, Etheridge B, O'Dea C. Why are households that report the lowest incomes so well-off? Econ J. 2017;127:F24–49.

Bruckmeier K, Müller G, Riphahn RT. Who misreports welfare receipt in surveys? Appl Econ Lett. 2014;21(12):812–6.

Bruckmeier K, Wiemers J. A new targeting: a new take-up? Non-take-up of social assistance in Germany after social policy reforms. Empir Econ. 2012;43(2):565–80.

Bruckmeier K, Wiemers J. Differences in welfare take-up between immigrants and natives: a microsimulation study. Int J Manpow. 2017;38(2):226–41.

Bruckmeier K, Wiemers J. Benefit take-up and labor supply incentives of interdependent means-tested benefit programs for low-income households. Comp Econ Stud. 2018;60(4):583–604.

Card D, Hildreth AKG, Shore-Sheppard LD. The Measurement of Medicaid Coverage in the SIPP: Evidence from California 1990-1996, NBER Working Paper No. 8514, Cambridge, Mass, 2001.

Davern ME, Meyer BD, Mittag N. Creating improved survey data products using linked administrative-survey data. J Surv Stat Methodol. 2019;7(3):440–63.

Duclos J-Y. Modelling the take-up of state support. J Public Econ. 1995;58(3):391–415.

Duclos J-Y. Estimating and testing a model of welfare participation: the case of supplementary benefits in Britain. Economica. 1997;64(253):81–100.

Eggs J, Jäckle A. Dependent Interviewing and sub-optimal responding. Surv Res Methods. 2015;9(1):15–29.

Eurofound. Access to social benefits: reducing non-take-up. Luxembourg: Publications Office of the European Union; 2015.

Frick JR, Groh-Samberg O. To claim or not to claim: estimating non-take-up of social assistance in Germany and the role of measurement error. SOEP Papers on Multidisciplinary Panel Data Research, No. 53, DIW, Berlin, 2007.

Gebhardt D, Müller G, Bethmann A, Trappmann M, Christoph B, Gayer C, Müller B, Tisch A, Siflinger B, Kiesl H, Huyer-May B, Achatz J, Wenzig C, Rudolph H, Graf T, Biedermann A. Codebook and

documentation of the panel study “Labour Market and Social Security” (PASS), Vol. I: Introduction and overview. Wave 2 (2007/2008). FDZ-Datenreport, 06/2009, 2009.
Hancock R, Barker G. The quality of social security benefit data in the British family resources survey: implications for investigating income support take-up by pensioners. J R Stat Soc A. 2005;168:63–82.
Hausman JA, Abrevaya J, Scott-Morton FM. Misclassification of the dependent variable in a discrete-response setting. J Econom. 1998;87(2):239–69.
Hernandez M, Pudney S. Measurement error in models of welfare participation. J Public Econ. 2007;91(1–2):327–41.
Hernandez M, Pudney SE, Hancock R. The welfare cost of means-testing: pensioner participation in income support. J Appl Econom. 2007;22(3):581–98.
Hug S. The effect of misclassifications in probit models: Monte Carlo simulations and applications. Polit Anal. 2010;18(1):78–102.
Krafft C, Davis EE, Tout K. The problem of measurement error in self-reported receipt of child-care subsidies: evidence from two states. Soc Serv Rev. 2015;89(4):686–726.
Lynn P, Jäckle A, Jenkins SP, Sala E. The impact of questioning method on measurement error in panel survey measures of benefit receipt: evidence from a validation study. J R Stat Soc A. 2012;175:289–308.
Meyer BD, Mittag N. Misclassification in binary choice models. J Econom. 2017;200(2):295–311.
Meyer BD, Mittag N. Using linked survey and administrative data to better measure income: implications for poverty, program effectiveness and holes in the safety net, IZA Discussion Papers, No. 10943, Institute of Labor Economics (IZA), Bonn, 2017b.
Meyer BD, Goerge R. Errors in survey reporting and imputation and their effects on estimates of food stamp program participation, CES (Center for Economics Studies) Working Paper 11-14, U.S. Census Bureau, Washington, 2011.
Meyer BD, Wallace KCM, Sullivan JX. Household Surveys in Crisis. J Econ Perspect. 2015;29(4):199–226.
Meyer BD, Mittag N, Goerge RM. Errors in survey reporting and imputation and their effects on estimates of food stamp program participation, NBER Working Paper 25143, National Bureau of Economic Research, Cambridge MA, 2018.
Mittag N. Correcting for Misreporting of Government Benefits, IZA Discussion Papers No. 10266, IZA Bonn, 2016.
Moffitt R. An economic model of welfare stigma. Am Econ Rev. 1983;73(5):1023–35.
Parolin Z. The effect of benefit underreporting on estimates of poverty in the United States. Soc Indic Res. 2019;144(2):869–98.
Pudney S. The impact of measurement error in probit models of benefit take-up (mimeo). Leicester: University of Leicester; 2001.
Riphahn RT. Rational poverty of poor rationality? The take-up of social assistance benefits. Rev Income Wealth. 2001;47(3):379–98.
Riphahn RT, Saif S. Naturalization and the labor market performance of immigrants in Germany. Labour: Rev Labour Econ Ind Relat 2019;33(1):48–76.
Sakshaug JW, Antoni M, Sauckel R. The quality and selectivity of linking federal administrative records to respondents and nonrespondents in a general population survey in Germany. Surv Res Methods. 2017;11(1):63–80.
Sakshaug JW, Kreuter F. Assessing the magnitude of non-consent bias in linked survey and administrative data. Surv Res Methods. 2012;6(2):113–22.
Skeels CL, Taylor LW. Prediction in linear index models with endogenous regressors. Stata J. 2015;15(3):627–44.
Statistik der Bundesagentur für Arbeit. Grundsicherung für Arbeitsuchende: Bedarfe, Leistungen und Haushaltsbudget. Bericht der Statistik der Bundesagentur für Arbeit, Nürnberg, Juli 2008.
Statistik der Bundesagentur der Arbeit. Tabellen, Strukturen der Grundsicherung SGB II (Zeitreihe Monats- und Jahreszahlen ab 2005), Nürnberg, Mai 2018.
STBA (Statistisches Bundesamt). Statistisches Jahrbuch 2018, Wiesbaden, 2018.
Taeuber C, Resnick DM, Love SP, Staveley J, Wilde P, Larson R. Differences in the estimates of food stamp program participation between surveys and administrative records, Working Paper, U.S. Census Bureau, Washington, 2004.
Tasseva IV. Evaluating the performance of means-tested benefits in Bulgaria. J Comp Econ. 2016;44(4):919–35.
Trappmann M, Beste J, Bethmann A, Müller G. The PASS panel survey after six waves. J Labour Mark Res. 2013;46:275–81.

Trappmann M, Gundert S, Wenzig C, Gebhardt D. PASS: a household panel survey for research on unemployment and poverty. Schmollers Jahrbuch (J Appl Soc Sci Stud). 2010;130(4):609–22.
van Oorschot W. Non-take-up of social security benefits in Europe. J Eur Soc Policy. 1991;1(1):15–30.
Whelan S. The take-up of means-tested income support. Empir Econ. 2010;39(3):847–75.
Wilde J, Kubis A. Nichtinanspruchnahme von Sozialhilfe: Eine empirische Analyse des Unerwarteten. Jahrbücher für Nationalökonomie und Statistik (J Econ Stat). 2005;225(3):347–73.