Who misreports welfare receipt in surveys?

Kerstin Bruckmeier[a], Gerrit Müller[a] and Regina T. Riphahn[b]

Words total: 2079
Words text only: 1574

We match survey and administrative data and determine the extent of misreporting on welfare receipt. In our data, 10.5 % of German welfare recipients underreport and 1 % overreport benefit receipt. The analysis shows that particularly households who are close to the labor market, without children, and with relatively high household incomes and savings are prone to underreport their welfare receipt. This information is important for the study of transfer programs based on survey data.

JEL Code: I32, C81

[a] IAB, Regensburger Str.104, 90478 Nuremberg, Germany
[b] *Corresponding author:* Univ. Erlangen-Nuremberg; Economics Department; Lange Gasse 20; 90403 Nuremberg; Germany; E-mail: regina.riphahn@fau.de

I. Introduction

It is important to study and understand the workings of social transfer programs. Ideally, one would use precise, rich, and recent data. Unfortunately, often we can only use information from survey data, which are known to suffer from misreporting with respect to benefit receipt. In this article we study the characteristics of individuals who fail to indicate their welfare receipt.

Underreporting of transfer receipt can only be studied if survey data can be matched to administrative records, which is rarely possibly. We are able to combine self reported information on welfare participation from survey data with precise administrative data on true program participation for a large welfare program in Germany. This unique data situation allows us to characterize individuals and households with misreported transfer receipt. The information from micro-level administrative data is more reliable and richer than the evidence that can be drawn from comparisons of weighted survey and aggregate program participation patterns which are often used in this literature (e.g., Meyer *et al.* 2009).

So far, only few studies have looked into the determinants of misreporting based on matched micro-data. Bollinger and David (1997) estimate parsimoniously specified models for false negative (error of omission) and false positive (error of commission) benefit receipt information. They use data from the Survey of Income and Program Participation (SIPP) on the US Food Stamp program and report 22 and 12 % underreporting at the individual and household level, respectively. The authors find high respondent income to be correlated with underreporting. Card *et al.* (2001) study Medicaid receipt and find an average rate of underreporting of 15 % and higher underreporting among those with a lower probability of benefit coverage. Taeuber *et al.* (2004) study the reporting of Food Stamp receipt in the American Community Survey and suggest that underreporting is connected to the reference period of the question: individuals reliably describe the current situation but make mistakes if the benefit receipt took place in the past and was of only short duration. Some authors studied interview-specific determinants of misreporting. They confirm the importance of misremembering program names, the timing of benefits, the continuity of receipt, and of survey characteristics such as proxy interviews (e.g., Sears and Rupp 2003, Meyer *et al.* 2009). The

only paper which investigated the correlation of personal characteristics with misreporting more generally is Meyer and Goerge (2011). They find that false negative statements of Food Stamp receipt in two US states are correlated with old age, high education, high income, male sex, nonwhite ethnicity, and a short duration of receipt.

We contribute to this literature by applying rich data from a different country, Germany. Our data match individual-level survey and administrative information based on information on date of birth, residence ZIP code, and name. We investigate the patterns of misreporting welfare receipt using a broader set of individual characteristics than is usually available.

Already Bollinger and David (1997, 2001) pointed out that precise information is important for studies on the participation in public programs, the effectiveness of such programs, and for the analysis of benefit nontakeup.

II. Data and Approach

We apply data taken from the fifth survey wave of PASS, the household panel study "Labour Market and Social Security." The survey started in 2006/07 and was designed for research on unemployment and poverty (Trappmann *et al.* 2010, 2013). It features a dual sampling frame which combines a subsample of welfare benefit recipients with a random population sample that oversamples households with low socio-economic status; overall the population sample accounts for 41 % of all observations. Due to this sampling frame our analyses focuses on weighted data. A major advantage of the PASS survey is that it asks respondents about current welfare receipt. This circumvents any type of recall error. The survey started out with about 6,000 households in each of the two subsamples; over time some households attrited from the panel while refreshment samples where added. The fifth survey was gathered in 2011 and administered to 10 235 households. We omit observations of respondents above age 65 and keep those 8576 households with valid information on welfare receipt. About 91 % of these households agreed to match their administrative records to the survey data and of those who agreed again about 91 % (N = 7031) could be matched based on the available information.

Based on information from the administrative data 3184 of these households received welfare at the time of the survey. We can now code an error of omission, i.e., a nonreporting of actual welfare receipt, and an error of commission, i.e., an erroneous (over-)reporting of welfare receipt for the date of the interview. Given the nature of the sample we apply survey weights to determine the population average rate of misreporting. Row 1 in Table 1 shows that with weighted data 10.5 % of welfare recipients underreport their actual welfare receipt, while 1.3 % overreport a welfare receipt which is not recorded in the administrative data.

Given that the precise timing of their ongoing welfare payments may not be completely transparent to recipients we evaluated whether respondents actually received benefits within a time window of plus or minus 15 or 30 days around the interview or at any other time within the calendar month of the interview. The resulting rates of misreporting were rather similar (see Table 1). As the question posed to respondents asked about the duration of welfare spells in terms of calendar months we consider the entries in the last rows of Table 1 as the most appropriate measures of misreporting; they refer to the ongoing month and are almost identical to the entries based on the exact date.

In our analysis, we investigate the correlation patterns between the propensity to underreport welfare receipt and observable characteristics of the household. Following the literature we consider characteristics of the head of household such as sex, age, health, education, and immigrant status, the household structure, indicators of the household economic situation, and regional information.

Table 2 describes our explanatory variables for the weighted and unweighted samples of 2949 households for which we have complete information and who are welfare recipients based on the administrative records.

We estimate three models to investigate the characteristics of underreporters of benefit receipt: first, we use a simple probit without survey weights (model 1). In the focus of our interpretation is the second model which uses a weighted regression (model 2). As the quality of interview responses and the propensity of underreporting may vary across interviewers we finally add a set of interviewer fixed effects to the specification (model 3).

III. Estimation results

Table 3 presents the estimated coefficients with robust standard errors in three columns for the three models.[1] Based on these results the quintessential underreporting welfare recipient is a female household head below age 25 with an intermediate secondary school degree. The person heads a household without children, quite likely has someone in the household who is 'regularly employed,' which may be the household head him- or herself. The equivalent per capita household income and savings of underreporting households are at the higher end of the distribution, and welfare receipt has not lasted for more than 16 months so far. These general patterns are confirmed in all three estimations.

While interviewer fixed effects jointly and statistically significantly improve the model fit, holding them constant increases the estimated standard errors somewhat but does not shift the association between characteristics and underreporting. We find the single largest correlation between holding a 'regular employment' and misreporting; 'regular employment' describes those full- or part-time workers who earn more than 400 Euro per month and who mandatorily contribute to the social insurances. So our typical underreporting household head is close to the labour market and most likely does not have a substantial benefit claim to the welfare system, given the relatively high household income and savings and the short-term nature of the benefit receipt. These results agree well with prior findings in the literature.

IV. Conclusions

We find a rate of underreporting of 10.5 % among welfare recipients. This is much below the rates of 30-50 % which Meyer and Goerge (2011) find for the Food Stamp Program in two U.S. states. However, the rate agrees with the 12 % reported by Bollinger and David (1997, 2001) for the Food Stamp program at the household level. Those who underreport the receipt of German

[1] We lose about one third of our sample when estimating the fixed interviewer effects model. As a robustness test we also estimated the fixed interviewer effect model using the Stata "asis" option which does not omit observations for which the interviewer fixed effect cannot be identified. The results are qualitatively similar to those presented and available upon request. The main results also hold up, when we reestimate model 2 with the sample used in column 3.

welfare benefits tend to be young, without children, close to the labor market, and most likely with low benefit amounts. Long term recipients with characteristics that may obstruct exit from welfare receipt, e.g., advanced age or small children in the household, appear to be more likely to truthfully report their recipient status.

These findings are important for unbiased analyses of welfare take-up and program effectiveness: studies based on self-reported program participation are at risk of underestimating the benefit receipt of those at the margin of welfare dependence. This can be of particular relevance when evaluating program effectiveness and households' propensity to exit the transfer program.

References

Bollinger, Christopher R. and Martin H. David, 1997, Modeling discrete choice with response error: food stamp participation, *Journal of the American Statistical Association* 92(439), 827-835.

Bollinger, Christopher R. and Martin H. David, 2001, Estimation with response error and nonresponse: food-stamp participation in the SIPP, *Journal of Business and Economic Statistics* 19(2), 129-141.

Card, David, Andrew K. G. Hildreth and Lara D. Shore-Sheppard, 2001, The Measurement of Medicaid Coverage in the SIPP: Evidence from California 1990-1996, *NBER Working Paper* No. 8514, Cambridge, Mass.

Meyer, Bruce D., Wallace K.C. Mok and James X. Sullivan, 2009, The under-reporting of transfers in household surveys: its nature and consequences, *NBER Working Paper* No. 15181, Cambridge, Mass.

Meyer, Bruce D. and Robert M. Goerge, 2011, 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.

Sears, James and Kalman Rupp, 2003, Exploring Social Security Payment History Matched with the Survey of Income and Program Participation, Working Paper, Social Security Administration, Washington.

Taeuber, Cynthia, Dean M. Resnick, Susan P. Love, Jane Staveley, Parke Wilde and Richard Larson, 2004, Differences in the Estimates of Food Stamp Program Participation Between Surveys and Administrative Records, Working Paper, U.S. Census Bureau, Washington.

Trappmann, Mark, Stefanie Gundert, Claudia Wenzig and Daniel Gebhardt, 2010, PASS - A Household Panel Survey for Research on Unemployment and Poverty, *Schmollers Jahrbuch (Journal of Applied Social Science Studies)* 130(4), 609-622.

Trappmann, Mark, Jonas Beste, Arne Bethmann and Gerrit Müller, 2013, The PASS panel survey after six waves, forthcoming: *Journal of Labor Market Research.*

Table 1 Errors of omission and commission by window of observations

	Error of omission	Error of commission
Weighted data, exact date	0.105	0.013
Unweighted data, exact date	0.107	0.036
Weighted data, +/- 15 days	0.112	0.012
Unweighted data, +/- 15 days	0.117	0.032
Weighted data, +/- 30 days	0.138	0.012
Unweighted data, +/- 30 days	0.129	0.032
Weighted data, same calendar month	0.105	0.013
Unweighted data, same calendar month	0.108	0.036

Source: Own calculations using PASS (2011) data matched to administrative records. The above shares are calculated on a total 7033 observations.

Table 2 Descriptive statistics

	Mean	Std. Dev.	Mean	Std. Dev.
	Unweighted		Weighted	
Characteristics of head of household (hh)				
Male hh	0.46	0.50	0.52	0.50
Age of hh: 25-34 years (ref.:15-24 years)	0.19	0.39	0.22	0.41
Age of hh 35-44 years (ref.:15-24 years)	0.23	0.42	0.22	0.41
Age of hh: 45-54 years (ref.:15-24 years)	0.28	0.45	0.28	0.45
Age of hh: >=55 years (ref.:15-24 years)	0.23	0.42	0.24	0.43
Hh is disabled	0.14	0.35	0.12	0.32
Hh is foreign national	0.20	0.40	0.20	0.40
Hh holds no sec. degree (ref.: lower secondary)	0.08	0.27	0.10	0.29
Hh holds intermediate sec. degree (ref.: lower sec.)	0.33	0.47	0.31	0.46
Hh holds upper sec. degree (ref.: lower secondary)	0.17	0.38	0.17	0.38
Hh holds other sec. degree (ref.: lower secondary)	0.01	0.12	0.01	0.11
Vocational education (ref.: no vocational qualification)	0.56	0.50	0.55	0.50
Tertiary degree (ref.: no vocational qualification)	0.13	0.33	0.11	0.31
Household structure				
Young children in household (age<=4 years)	0.12	0.32	0.11	0.31
Family without children (ref.: single person)	0.13	0.33	0.07	0.26
Single parents (ref.: single person)	0.25	0.43	0.20	0.40
Family with children (ref.: single person)	0.16	0.37	0.13	0.33
Others (ref.: single person)	0.007	0.08	0.004	0.06
Household economic situation				
Regular employed person in household	0.15	0.36	0.14	0.35
Household income: 500-749 € (ref.: <500 €)	0.42	0.49	0.42	0.49
Household income: 750-999 € (ref.: <500 €)	0.33	0.47	0.31	0.46
Household income: >=1000 € (ref.: <500 €)	0.09	0.29	0.09	0.29
Household savings: <1000 € (ref.: no savings)	0.31	0.46	0.30	0.46
Household savings: <2500 € (ref.: no savings)	0.05	0.22	0.04	0.20
Household savings: <5000 € (ref.: no savings)	0.03	0.18	0.03	0.18
Household savings: >=5000 € (ref.: no savings)	0.03	0.17	0.04	0.20
Household owns home	0.06	0.24	0.06	0.24
Benefit receipt < 37 months (ref.: < 17 months)	0.25	0.43	0.18	0.39
Benefit receipt >=37 <71 months (ref.: < 17 months)	0.25	0.43	0.24	0.43
Benefit receipt >70 months (ref.: < 17 months)	0.25	0.43	0.27	0.45
Regional information				
Rural area (ref.: intermediate area)	0.19	0.39	0.23	0.42
Metropolitan area (ref.: interm. area)	0.39	0.49	0.39	0.49
Eastern Germany	0.34	0.47	0.34	0.47

Note: The full set of covariates is observed for 2949 observations. All variables are 0/1 indicators.

Table 3 Estimation results: coefficients of probit on error of omission (underreporting)

	(1)	(2)	(3)
Characteristics of head of household (hh)			
Male hh	-0.0271***	-0.0291**	-0.0210
	(0.0104)	(0.0122)	(0.0175)
Age of hh: 25-34 years (ref.:15-24 years)	-0.0567*	-0.0839*	-0.0883*
	(0.0300)	(0.0444)	(0.0524)
Age of hh 35-44 years (ref.:15-24 years)	-0.0926***	-0.0894*	-0.152**
	(0.0301)	(0.0464)	(0.0744)
Age of hh: 45-54 years (ref.:15-24 years)	-0.113***	-0.144***	-0.180**
	(0.0302)	(0.0446)	(0.0857)
Age of hh: >=55 years (ref.:15-24 years)	-0.108***	-0.150***	-0.187**
	(0.0311)	(0.0455)	(0.0896)
Hh is disabled	0.00173	0.0268	0.0316
	(0.0160)	(0.0209)	(0.0271)
Hh is foreign national	0.0185	0.0296	0.0359
	(0.0142)	(0.0199)	(0.0287)
Hh holds no sec. degree (ref.: lower secondary)	0.00289	0.0230	0.0548
	(0.0199)	(0.0262)	(0.0417)
Hh holds intermediate sec. degree (ref.: lower sec.)	0.0155	0.0546***	0.0395
	(0.0120)	(0.0167)	(0.0241)
Hh holds upper sec. degree (ref.: lower secondary)	-0.000792	-0.0105	-0.00522
	(0.0148)	(0.0154)	(0.0196)
Hh holds other sec. degree (ref.: lower secondary)	0.0152	0.0350	0.126
	(0.0394)	(0.0476)	(0.0888)
Vocational education (ref.: no vocational qualification)	-0.00902	0.0222	0.0168
	(0.0127)	(0.0160)	(0.0180)
Tertiary degree (ref.: no vocational qualification)	-0.00143	0.0257	0.0240
	(0.0182)	(0.0209)	(0.0295)
Household structure			
Young children in household (age<=4 years)	-0.0308*	-0.0598***	-0.0637*
	(0.0157)	(0.0169)	(0.0354)
Family without children (ref.: single person)	-0.00574	-0.0109	-0.0169
	(0.0173)	(0.0217)	(0.0231)
Single parents (ref.: single person)	-0.0579***	-0.0681***	-0.0685*
	(0.0126)	(0.0161)	(0.0372)
Family with children (ref.: single person)	0.00563	0.0191	0.0168
	(0.0186)	(0.0216)	(0.0246)
Others (ref.: single person)	-0.0260	-0.0275	0.00210
	(0.0593)	(0.0599)	(0.0654)

(continued next page)

	(1)	(2)	(3)
continued			
Household economic situation			
Regular employed person in HH	0.243***	0.248***	0.298**
	(0.0240)	(0.0350)	(0.124)
Household income: 500-749 € (ref.: <500 €)	-0.0178	-0.0162	0.0114
	(0.0144)	(0.0171)	(0.0187)
Household income: 750-999 € (ref.: <500 €)	0.00332	0.0254	0.0533
	(0.0160)	(0.0197)	(0.0344)
Household income: >=1000 € (ref.: <500 €)	0.123***	0.0911**	0.168**
	(0.0269)	(0.0356)	(0.0850)
Household savings: <1000 € (ref.: no savings)	0.00213	0.00861	0.00259
	(0.0106)	(0.0137)	(0.0163)
Household savings: <2500 € (ref.: no savings)	0.00683	0.0235	0.0109
	(0.0205)	(0.0279)	(0.0306)
Household savings: <5000 € (ref.: no savings)	0.0179	0.0115	-0.0127
	(0.0275)	(0.0314)	(0.0268)
Household savings: >=5000 € (ref.: no savings)	0.0915***	0.0510	0.0689
	(0.0336)	(0.0381)	(0.0597)
Household owns home	0.0101	0.0242	0.0336
	(0.0212)	(0.0277)	(0.0329)
Benefit receipt < 37 months (ref.: < 17 months)	-0.0604***	-0.0699***	-0.0619
	(0.0154)	(0.0215)	(0.0382)
Benefit receipt >=37 <71 months (ref.: < 17 months)	-0.0899***	-0.101***	-0.105*
	(0.0145)	(0.0192)	(0.0604)
Benefit receipt >70 months (ref.: < 17 months)	-0.0978***	-0.125***	-0.127*
	(0.0152)	(0.0193)	(0.0731)
Regional information			
Rural area (ref.: intermediate area)	-0.0176*	-0.00190	-0.0365
	(0.0106)	(0.0143)	(0.0272)
Metropolitan area (ref.: intermediate area)	0.0223	0.0277	0.0363
	(0.0147)	(0.0178)	(0.0311)
Eastern Germany	0.00157	0.00845	0.0318
	(0.0106)	(0.0148)	(0.0268)
Interviewer fixed effects	no	no	yes

Note: We use 2949 observations in columns 1 and 2, and 1926 observations in column 3; ***, **, and * represent statistical significance at the 1, 5, and 10 % level. Standard errors are heteroscedasticity robust. Model (1) uses unweighted, models (2) and (3) use weighted data.