



Working Paper Series

#13

Stefan ANGEL
Franziska DISSLBACHER
Stefan HUMER
Matthias SCHNETZER

What did you really earn last year? Explaining measurement error in survey income data

Original Citation:

Angel, Stefan and Disslbacher, Franziska and Humer, Stefan and Schnetzer, Matthias (2019)
What did you really earn last year? Explaining measurement error in survey income data.
INEQ Working Paper Series, 13. WU Vienna University of Economics and Business, Vienna.

The final version of this paper is published in Royal Statistical Society.
doi: <https://doi.org/10.1111/rssa.12463>



What did you really earn last year? Explaining measurement error in survey income data*

Stefan Angel^a, Franziska Disslbacher^a, Stefan Humer^b, and Matthias Schnetzer^c

^a*Department of Socioeconomics, Vienna University of Economics and Business, Vienna, Austria.*

^b*Research Institute Economics of Inequality, WU Vienna, Austria.*

^c*Department of Economics, Federal Chamber of Labor, Vienna, Austria.*

Abstract

This paper analyses the sources of income measurement error in surveys with a unique dataset. We use the Austrian 2008–2011 waves of EU-SILC which provide individual information on wages, pensions and unemployment benefits from survey interviews and officially linked administrative records. Thus, we do not have to fall back on complex two-sample matching procedures like related studies. We empirically investigate four sources of measurement error, namely (i) social desirability, (ii) socio-demographic characteristics of the respondent, (iii) the survey design, and (iv) the presence of learning effects. We find strong evidence for a social desirability bias in income reporting, while the presence of learning effects is mixed and depends on the income type under consideration. An Owen value decomposition reveals that social desirability is a major explanation of misreporting in wages and pensions, whereas socio-demographic characteristics are most relevant for mismatches in unemployment benefits.

Keywords: Income measurement error; Register data; Response error; Survey data; Survey error; Survey methods

JEL Classifications: C2, C81, C82, C83, D3

*This manuscript has been published as *Angel, S. , Disslbacher, F. , Humer, S. and Schnetzer, M. (2019), What did you really earn last year?: Explaining measurement error in survey income data. J. R. Stat. Soc. A. doi:10.1111/rssa.12463*. We thank Wilfried Altzinger, Stephen P. Jenkins, Arthur Kennickell and participants of various seminars for helpful comments on earlier versions of the paper. Stefan Humer's research is supported by funds of the Oesterreichische Nationalbank (OeNB Anniversary Fund, project number: 16728).

1 Introduction

The rapidly increasing availability of survey and administrative micro data has created immense possibilities for contemporary empirical economics. However, this progress has also raised questions of data quality and validation in order to ensure the reliability and accuracy of information. In the literature, the analysis of differences between responses in traditional surveys and administrative records has been a fruitful method to assess quality of micro data. Particular attention has been paid to income data since it is essential for a variety of welfare indicators and policy questions. Moreover, public interest in questions of income distribution has been growing considerably in recent years and research on income inequality has rapidly gained momentum. The underlying income information is usually obtained either from household surveys or from administrative records whereby both data sources have their idiosyncratic merits and drawbacks. While policy recommendations are frequently based on survey data, the accuracy of survey responses is often questioned and issues of measurement error are raised. Accordingly, there is still no consensus upon what is the best way to collect income data at the micro-level (Hansen and Kneale, 2013).

Potential differences between survey responses and administrative records have already been addressed in the literature on data quality and measurement error (Moore et al., 2000; Lohmann, 2011). The identification of measurement errors requires by definition a point of reference to judge the accuracy of the information. In validation studies, administrative records are frequently used as benchmark even though this assumption has also been relaxed to a varying extent in the recent literature (Kapteyn and Ypma, 2007; Abowd and Stinson, 2013; Bingley and Martinello, 2017). The choice of the benchmark data reflects the researchers' confidence in the accuracy of a specific dataset. We rely on evidence that the quality of Austrian administrative data is very high (Asamer et al., 2016; Schnetzer et al., 2015; Berka et al., 2012).

This paper focuses on two research questions. What are the reasons for misreporting income in surveys and do these reasons differ with respect to specific types of income, i.e. wages, pensions, and unemployment benefits? The causes why survey responses may deviate from administrative records are manifold (Bound, Brown, and Mathiowetz, 2001; Tourangeau, Rips, et al., 2000; Kapteyn and Ypma, 2007; Abowd and Stinson, 2013; Paulus, 2015). Specifically, we focus on four error sources emphasised in the literature. These are (i) the presence of a social desirability bias in survey responses, (ii) specific socio-demographic characteristics of the respondent, and (iii) the survey design. Finally, we test for (iv) the presence of learning effects in the response behaviour. These reasons may also vary with respect to the income component. For instance, social desirability may have different directions for earned income and for unemployment benefits.

The data requirements for an empirical evaluation of these error sources are extensive. Most previous studies have been forced to combine survey and register records via statistical matching techniques, based on either register data or error-prone self-reported identifiers, such as social security numbers. Additionally, consent of individuals to match survey and register data is usually needed. Since participation in the matched sample often is voluntarily, the sample is biased towards individuals giving more accurate responses (Bollinger, 1998). Briefly speaking, there is evidence of a consent bias for matched data sources. Furthermore, samples based on optional matching are often found to be non-representative for the whole population (Jenkins, Cappellari, et al., 2006; Jenkins,

Lynn, et al., 2008; Sakshaug and Eckman, 2016; Sakshaug, K Couper, et al., 2012).

Fortunately, the Austrian 2008–2011 waves of the European Union Statistics on Income and Living Conditions (EU-SILC) do not suffer from such shortcomings. In this dataset, the survey responses are linked directly to administrative records by the National Statistical Institute via unique personal identifiers. Compared to probabilistic methods, this procedure ensures that the administrative income information is linked exactly to the corresponding survey respondent and thus reduces matching uncertainty drastically. Hence, our dataset provides both survey and register income data for the exact same observational units and offers a unique opportunity to evaluate the drivers of income measurement error.

We thus are able to compare survey and administrative data within one dataset which is a considerable advantage compared to most existing research. Even the most prominent studies about income measurement error are based on sophisticated statistical matching procedures of survey and administrative datasets. Other validation studies are restricted to individuals working at a single company (Pischke, 1995). The dataset used in this paper is not subject to any of these limitations. Furthermore, in contrast to papers studying measurement errors in total household income, we are able to analyse different single income components including unemployment benefits, which has rarely been done yet.

In summary, we contribute to the literature in multiple ways. We exploit a unique dataset providing both administrative records and survey responses to analyse differences in income reports. We are able to conduct a comparative analysis of three different income concepts with a panel structure over the period of four years. Contrary to other studies that are limited to single companies, our analysis is representative for the whole population. Finally, we apply an Owen value decomposition to quantitatively assess the importance of four specific reasons for misreporting.

The remainder of this article is structured as follows. Section 2 provides a description of potential sources of measurement errors elaborated in the existing literature. In section 3, we discuss the specifics of our EU-SILC dataset and provide descriptive statistics for the structure of errors in our data. We then apply a multinomial logit model to evaluate the effects of the above-mentioned error sources (i) through (iv) in section 4. For a more detailed analysis of error sources (i) and (iv), we conduct panel regressions to detect changes of response behaviour over time. We conclude our analysis quantifying the relative importance of the four error sources based on an Owen value decomposition. Finally, section 5 provides a summary of our results.

2 Reasons for Measurement Error in Survey Income Data

While the use of administrative data in empirical research has rapidly gained momentum (Einav and Levin, 2014), the accuracy of survey information has increasingly been contested during the last years (Meyer et al., 2015). Erroneous information can arise from misreporting by respondents and decrease the overall quality of survey data. Misreporting in surveys is particularly grave when the affected data is, like income, the basis for policy-making.

Following the psychological literature on cognitive processes when answering a survey question, misreporting of income can arise from, first, troubles related to the interpretation and understand-

ing of the question asked, second, problems in retrieving and judging the relevant information as well as its placement in time and, third, difficulties in formulating a response in the required format (Tourangeau, Bradburn, et al., 2010). The theoretical and empirical literature on income measurement errors emphasises four main reasons for misreporting: (i) social desirability, (ii) socio-demographic characteristics of the respondent, (iii) specifics of the survey design, and (iv) the presence of learning effects. However, the existing literature has ignored that those four reasons for misreporting can vary with the income type under consideration.

For instance, Angel et al. (2017) analyse reporting errors for total disposable household income in Austria, which is not observed directly but aggregated ex-post based on a comprehensive inquiry of single income components. Misreporting based on total household income thus disregards potential heterogeneity of the error-generating process between household members and income components. Further, misreporting of income can go into two directions: respondents can overreport or underreport a particular income type. In what follows, we discuss how (i) to (iv) can result in over- or underreporting of wages and then, why (i) to (iv) might lead to different expectations for the misreporting of pensions and unemployment benefits, the two other income components considered in this paper.

First, social desirability bias is probably the most important reason for income misreporting in surveys. Due to the sensitivity of questions about income, social desirability might lead to deliberate misreporting (Moore et al., 2000). It is widely documented that sensitive questions elicit patterns of overreporting (underreporting) for socially (un)desirable behaviour, attitudes, and characteristics (Bound, Brown, and Mathiowetz, 2001). For wages, the resulting hypothesis is that reported values are biased towards the mean, hence reporting errors are expected to be mean-reverting. Respondents at the lower tail of the wage distribution overreport as they feel ashamed of their actual economic conditions, while respondents at the upper tail of the distribution underreport since they do not want to disclose their high wages to an (unknown) interviewer. Such a reporting pattern is consistent with a desire for social comfort in the sense that households tend to locate themselves in the middle of the distribution. Related micro-level validation studies typically confirm the mean-reverting nature of the error in reported earnings (Kreiner et al., 2015; Kim and Tamborini, 2014; Pischke, 1995; Bound, Brown, Duncan, et al., 1994; Bound and Krueger, 1991). As an exception, Hariri and Lassen (2017) find for the Netherlands that respondents at the top of the income distribution overreport their income. In their study, however, income comprises earnings, employers' pension contributions, transfer and capital income, which were collected exclusively via telephone interviews with a one-shot recall question. These results are thus not easily comparable with most other studies that focus on earnings and derive income data from surveys with more complex interview modes.

Second, misreporting of income might vary with socio-demographic characteristics of the respondent (Kreiner et al., 2015; Kim and Tamborini, 2014; Bound, Brown, and Mathiowetz, 2001; Tamborini and Kim, 2013). We expect to find a higher propensity to overreport wages for males, due to a desire to demonstrate social status and to comply with the male-breadwinner model. Existing research suggests different misreporting patterns by sex, where males are found to overreport earnings more often than females (Bound and Krueger, 1991; Bollinger, 1998; Micklewright and Schnepf, 2010; Pedace and Bates, 2000; Kim and Tamborini, 2014). Kim and Tamborini (2014) and Bound, Brown, Duncan, et al. (1994) find that higher educated workers report earnings more

correctly. Respondents with higher education might display lower rates of misreporting as this group is potentially more likely to be familiar with the purpose and relevance of households surveys. The positive correlation between education and financial literacy documented in the literature (Lusardi and Mitchell, 2014) may also explain some of these findings.

The relationship between misreporting and age is a priori unclear. On the one hand, cognitive abilities to answer the questionnaire decrease with age. On the other hand, older respondents receive wages for a longer period of time and are in more stable employment. Since the vast majority of validation studies has found a negative relationship between misreporting and age, we adopt these findings for the expectations for our Austrian sample (Kim and Tamborini, 2014; Kim and Tamborini, 2012; Bound and Krueger, 1991). Two additional socio-demographic characteristics potentially contributing to the accuracy of responses are the number of changes in the employment status during the income reference period, related to receiving income from different sources and second, the number of months spent in a specific employment status during the reference period. Changes in the employment status can be associated with telescoping errors, which refer to misplacing the receipt of a particular income source in time. A stable employment status is associated with less variation in the level of income received. With respect to wages, it is reasonable to assume that changes in the employment status increase reporting errors while longer employment periods lead to less misreporting, since it is easier to recall the remuneration. Kim and Tamborini (2014) find that occupation or industry switchers are more prone to misreport earnings while Bound, Brown, Duncan, et al. (1994) document a negative relationship between job tenure (years with current employer) and response error.

What the literature has largely ignored is the relationship between misreporting and health, the degree of urbanisation, and the country of birth. Healthier individuals may give more accurate answers because they are in better mental conditions and therefore are less likely to make recall or response errors. Further, we expect to find less misreporting, the higher the degree of urbanisation at the respondent's place of residence is. The rationale of this argument is rooted in the anonymity of cities, whereas in rural areas, mistrust in unknown interviewers might be more pronounced. Additionally, misreporting is related to the respondent's country of birth. Being foreign-born can serve as an indicator of language skills and familiarity with institutional settings. As both factors are relevant for the comprehension of the questions and the correct allocation of total income across income types, we expect to find more misreporting of wages for foreign-born respondents.

Third, and regarding the survey design and setting, a wide range of variables is likely to influence the response behaviour. We focus on the role of the interview mode (Lynn et al., 2012), the time span between the income reference period and the interview, and proxy responses. Regarding the interview mode, telephone interviews are considered to be more susceptible for misreporting of wages than a face to face setting. Fessler et al. (2017) find that households interviewed via telephone report higher incomes on average than those interviewed personally. Furthermore, it is crucial whether the respondent provides the required income information personally or via an entitled third party. While some studies have found little proxy bias (Bound and Krueger, 1991; Mellow and Sider, 1983), others suggest a downward bias (Tamborini and Kim, 2013; Reynolds and Wenger, 2012). We expect more misreporting of wages for proxy responses, resulting from a lack of information. Finally, we hypothesise that the larger the time span between the interview and the income reference

period, the larger the error in reported wages because of recall and memory problems.

The fourth, and last, central issue with reporting errors is the presence of learning effects. If present, reporting errors are supposed to attenuate with cumulated survey experience. Learning effects are related to recall and retrieval strategies of respondents and are best explained in the context of panel surveys, where the same households are interviewed repeatedly. In the first wave of participation, respondents are unexperienced regarding the survey setting and unprepared to answer the questionnaire. In the follow-up waves, however, respondents might be equipped with income tax documents and other relevant files. Therefore, we expect to find misreporting of wages to decrease with accumulated survey experience. Likewise, a learning effect can also be expected for pensions and unemployment benefits.

While some variables might have similar effects on wages, pensions, and unemployment benefits, we anticipate diverging effects for others. For instance, wages (and pensions) are attached to the labor market and tied to (past) individual effort, whereas unemployment benefits are often stigmatised as charity despite actually being an insurance. The socially desirable behaviour is thus to downplay the level received leading to a general underreporting of unemployment benefits. Regarding (ii) the socio-demographic characteristics of the respondent, we expect males to underreport unemployment benefits more often and stronger than females since receiving benefits contrasts the common male bread-winner model. Being unemployed might be associated with a higher social stigma for the better educated. Therefore, we expect underreporting (overreporting) of unemployment benefits to be an increasing (decreasing) function of education. While being foreign-born might lead to a higher misreporting of wages and pension income, we expect to find less overreporting of unemployment benefits since those born abroad might be particularly prone to downplay the level of received state transfers.

With respect to the number of changes in the employment status, some particularities have to be mentioned. Typically, respondents only retire once and thus might be better informed about the level of pension they are going to receive. Precisely for this reason, we expect to find less misreporting of pensions, if the employment status for those receiving pension income has changed in the reference period. The relationship between the number of months spent in unemployment and misreporting of unemployment benefits is ambiguous. Short unemployment spells might be associated with less misreporting since respondents are better informed about the actual transfer due to the singularity of the situation. On the other hand, respondents might care less about the level of received unemployment benefits, the shorter the time spent in unemployment.

Table 2 summarises our expectations regarding the direction of the effect of the relevant variables related to social desirability, socio-demographic characteristics, the survey setting, and the learning effect, for over- and underreporting wages, pension income and unemployment benefits.

Table 1: Expected effects of misreporting by survey income types

	Wages		Pension		Unemp. ben.	
	S<A	S>A	S<A	S>A	S<A	S>A
(i) Social Desirability						
Income Decile (<i>increasing</i>)	+	-	+	-	+	-
(ii) Socio-demographic characteristics						
Gender (<i>Ref: Female</i>)	-	+	-	+	+	-
Education (<i>Ref: Compulsory school</i>)	-	-	-	-	-	-
Age (<i>increasing</i>)	-	-	-	-	-	-
Country of birth (<i>Ref: Austria</i>)	+	+	+	+	+	-
Health status (<i>Ref: Very bad</i>)	-	-	-	-	+	+
Degree of urbanisation (<i>Ref: <10,000 inhabitants</i>)	-	-	-	-	-	-
Changes in employment status (<i>Ref: None</i>)	+	+	-	-	+	+
Months in corresponding emp. status (<i>Ref: 12 months</i>)	+	+	+	+	~	~
(iii) Survey setting						
Mode of interview (<i>Ref: CAPI</i>)	+	+	+	+	+	+
Type of interview (<i>Ref: Personal</i>)	+	+	+	+	+	+
Month of interview (<i>Ref: March-May</i>)	+	+	+	+	+	+
(iv) Learning effect						
Wave of interview (<i>Ref: First</i>)	-	-	-	-	-	-

Notes: This table summarises our expectations about the likelihood to over- and underreport income conditional on different sources of errors. S<A: Underreporting in Survey; S>A: Overreporting in Survey; + stands for an increasing probability to fall in that specific reporting category, - symbols a decreasing probability, ~ implies an ambiguous relationship.

3 Data, variables, and method

3.1 Data and variables

For the assessment of income measurement errors in surveys, we make use of the Austrian EU-SILC (European Union Statistics on Income and Living Conditions). EU-SILC is a rotational household panel with a quarter of respondents being exchanged each year. The sample is drawn from all private households with a main residence in Austria according to the central population register. The main questionnaire is aimed at household members aged 16 or older and is conducted partly with computer-assisted personal interviews (CAPI) and partly with telephone interviews (CATI). In certain cases, proxy interviews with other household members were carried out instead of personal interviews (e.g. roughly 11% in 2011).

The primary motivation to utilise Austrian EU-SILC data is a unique feature compared to other household surveys: for four consecutive years, it provides combined income information from personal interviews and administrative sources. Up to the year 2011, incomes on both personal and household levels were obtained via conventional survey interrogation. From 2012 on, income data gathered from administrative records have replaced survey data for certain components of disposable household income, like wages, pensions, and unemployment benefits. Fortunately, Statistics Austria was able to merge administrative income data to the full EU-SILC survey sample from 2011 back to 2008. The detailed merging process of register and survey data is described further below.

The rotational character of EU-SILC allows to track households for as long as four consecutive years. With the data at hand, this maximum period of observation applies to the survey cohort that

first participated in 2008 and remained in the survey until 2011. The 2007 and 2009 cohorts are each covered in three waves, the 2006 and 2010 cohorts in two waves respectively. In contrast to Eurostat’s EU-SILC User Database (UDB), the Austrian national SILC provided by Statistics Austria delivers cross-sectional and longitudinal information in one single database. In this integrated dataset, there are one cross-sectional and three longitudinal weights for those households in the rotational sample that have already been interviewed repeatedly. By virtue of permanent household and individual identifiers, we are able to track the changes in household responses compared to changes in the administrative records over time. It has to be noted in this respect that, in contrast to the first-time interrogation, follow-up interviews were predominantly accomplished via telephone (CATI).

Survey income data are collected retrospectively in EU-SILC and correspond to the calendar year prior to the interview. For income components with unequal net and gross values, respondents were asked to report either one or both values. When refusing to deliver a precise figure, interviewees could also report an income bandwidth. In the latter case, the specific income is estimated based on the empirical distribution of the corresponding income component. In the event of item non-response for single income categories, missing values are derived partly from socio-economic characteristics like sex, education, and age in an econometric exercise, and partly from statutory regulations like collective wage agreements. While the amount of imputations is generally rather low in EU-SILC, the application of such estimation methods may be an important source of measurement error. For instance, roughly 0.5% of the records had to be completely imputed in 2011 (Statistics Austria, 2014b). Since we focus on error generating processes in personal survey responses, we exclude imputed values from the analysis. We thus restrict our study on net income from wages, pension and unemployment benefits for which both survey responses and administrative data are available in EU-SILC. Table 2 provides descriptive statistics for the number of observations, the interview mode, and the share of imputations for the selected income types.

Table 2: Observations, Interview Mode, Share of Imputations – SILC 2008-2011

	Obs.		Int. Mode (%)		Wages (%)		Pensions (%)		Unemp. (%)	
	HH	Persons	CAPI	CATI	Svy.	Adm.	Svy.	Adm.	Svy.	Adm.
2008	5,707	10,946	70.0	30.0	4.6	2.6	1.5	0.5	0.2	0.8
2009	5,878	11,056	57.1	42.9	5.9	1.9	2.0	0.3	0.2	0.8
2010	6,188	11,493	59.6	40.4	4.8	2.4	1.5	0.4	0.3	0.9
2011	6,187	11,475	57.3	42.7	3.4	2.6	0.5	0.5	0.3	1.1

Notes: Columns 5–10 show the share of imputed values for selected income types. Source: SILC 2008-2011, own calculations.

Register information for the income types used in our study is obtained from various administrative sources whereof the most important are the Austrian Social Security Database (ASSD), the Wage Tax Register, and the Transfers Dataset by the Public Employment Service Austria (AMS). We briefly describe these datasets below while a detailed documentation of administrative sources used in EU-SILC is presented in the data appendix A and in Statistics Austria (2014a). The ASSD provides the social security status, e.g. whether individuals are employees, pensioners or eligible for unemployment benefits. The Wage Tax Register contains information on all taxable earnings of employees and pensioners. This dataset includes wages, public pensions (retirement benefits), paid maternity leave and sickness benefits. The Transfers Dataset contains the beginning and ending

date of unemployment spells and the respective benefits on a daily basis.

In Austria, employers have to report wages directly to the tax authorities, retain payroll taxes and social security contributions, and transfer the residual net income to the employees. Since this is a standardized electronic procedure and there are regular tax inspections, the data quality is significantly higher than of self-reported income tax returns. Concerning flaws in administrative data due to tax avoidance, there is empirical evidence that this is comparatively less of an issue in Austria than in other OECD countries (Alm and Torgler, 2006; Hassan and Schneider, 2016). Information on pensions and unemployment benefits are provided straight by public authorities and correspond with actual payments to the entitled individuals. The probability of measurement errors in this dataset is very low since reports are effectively linked to payments that are also administrated by the data holders. All in all, recently carried out quality reports on Austrian register data certify high confidence in its reliability (Asamer et al., 2016; Statistics Austria, 2014a).

The merging process between survey and administrative data is accomplished reliably with a branch-specific personal identification number for official statistics (bPIN OS) which serves as a unique ID in both data sources. These 172-digit PINs were introduced to protect privacy in the communication among public authorities via e-government. The PINs are created by the Austrian Data Protection Commission and used to identify individuals in the EU-SILC survey and in the administrative sources. Unlike previous studies, we thus do not have to fall back on two-sample matching processes or the like, since survey responses and retrospective administrative information are already combined for the years 2008 to 2011. Studies on measurement errors usually also depend on consent to link survey responses to administrative records which often leads to small sample sizes (Kreuter et al., 2010). In our study, between 95.6% (2008) and 99.4% (2011) of the respondents in the EU-SILC survey could be identified with a PIN in order to assign the register information (Statistics Austria, 2014a). The residual population in EU-SILC could not be found in the administrative sources and thus no PIN was available. These individuals are mainly younger than 40 years old, non-Austrian citizens, and not registered at their main residence. Another major advantage of the accurate linkage is that the income reference periods for the survey and the administrative records overlap exactly and no adjustments to ensure comparability between the data sources had to be made.

Table 3: Reporting of Income Types – SILC 2008-2011

	Wages (%)			Pensions (%)			Unemp. ben. (%)		
	Svy.	Adm.	Diff.	Svy.	Adm.	Diff.	Svy.	Adm.	Diff.
2008	53.8	56.6	-2.8	24.9	24.1	0.8	7.1	10.2	-3.1
2009	54.8	58.0	-3.2	25.0	24.7	0.4	7.4	10.2	-2.8
2010	55.4	57.9	-2.4	25.9	25.3	0.6	9.1	12.6	-3.5
2011	55.8	59.0	-3.2	24.0	24.6	-0.6	9.2	12.7	-3.5

Notes: This table shows the share of respondents reporting a specific income type.
Source: SILC 2008-2011, own calculations.

Reporting income is a two-stage process in EU-SILC. At the first stage, respondents have to indicate whether a certain income component was received during the reference period. Only in a second step, the amount of income received from a particular source has to be reported. Consequently, a mismatch between survey and register data can result at either stage. Table 3 shows the percentage

of respondents reporting the three involved income types. Wages are consistently underreported in the survey data by 2.4 to 3.2 percentage points. By contrast, the number of individuals reporting pension income is slightly higher in the survey data compared to the administrative records with the exception of 2011. The share of survey respondents with declared unemployment benefits is generally lower than indicated by the official statistics. The deviations range between 2.8 and 3.5 percentage points. Overall, we notice a prevailing underreporting of income receipt in survey responses with the exception of old-age benefits from 2008 to 2010.

With regard to the level of reported income, we distinguish the following possible cases of mismatch. Respondents can report a particular source of income in the survey even though it was not received according to register data ($S \neq A$) or vice versa, not report a particular type of income even though it was received according to register data ($S = A$). Further, the survey report can positively or negatively deviate from the register data, which corresponds to overreporting ($S > A$) and underreporting ($S < A$), respectively. No mismatch occurs if a respondent reports the amount that corresponds to the register entry within a narrow range of $\pm 5\%$ ($S = A$) or if a specific type of income was not received according to both survey and register data ($S = 0$). Since reporting an amount in the survey that corresponds exactly to the register value is almost impossible, we allow the survey report to deviate marginally from the register entry in order to fulfil our operational definition of correct answers. We define a categorical variable $Pr(Y_{i,k} = j)$ for mismatch types j , individuals $i = 1, \dots, N$, and income component $k \in [1, 3]$ as

$$Pr(Y_{i,k} = j) = \begin{cases} Pr(Y_{i,k} = 0 \times 0) & \text{if no income in survey and admin} \\ Pr(Y_{i,k} = 0 \times A) & \text{if only admin income report (false negative)} \\ Pr(Y_{i,k} = S < A) & \text{if survey underreporter} \\ Pr(Y_{i,k} = S = A) & \text{if survey corresponds to admin} \\ Pr(Y_{i,k} = S > A) & \text{if survey overreporter} \\ Pr(Y_{i,k} = S \times 0) & \text{if only survey income report (false positive)} \end{cases}$$

Table 4 gives an overview of the structure of mismatch in Austrian EU-SILC data from 2008 to 2011. In this summary, we display the shares of observations in the respective reporting categories and the median of the absolute and relative deviation between survey and administrative responses. For all income types, the shares of over- and underreporters are very stable over the years. There are roughly 17% overreporters and 23% underreporters for wages which is a high number compared to only 11 to 13% of respondents with matching values (apart from the 40% with no wage income in both data sources). The median relative deviation from the administrative information lies around -17% for underreporters and roughly 26% for overreporters.

Concerning pensions, we see stable shares of reporting types except for 2011 where the percentage of correct responses takes a sudden jump to roughly 20%. Between 2008 and 2010, the share of underreporters exceeds the number of overreporters, however, the median deviation is considerably higher for the latter group. With regard to unemployment benefits, the share of correct survey answers is very low and consistently smaller than the shares of over- and underreporters. The median respondents with lower survey than administrative values report approximately 30% less income. The median relative deviation for overreporters ranges between 34 and 48%.

Table 4: Structure of mismatch in income reporting – SILC 2008-2011

	Wages			Pensions			Unemp. benefits		
	Obs. (%)	Abs. (P50, €)	Rel. (P50, %)	Obs. (%)	Abs. (P50, €)	Rel. (P50, %)	Obs. (%)	Abs. (P50, €)	Rel. (P50, %)
2008									
0x0	40.3	0.0	0.0	73.5	0.0	0.0	88.6	0.0	0.0
0xA	5.9	-1566.2	-100.0	1.6	-4474.5	-100.0	4.3	-1703.1	-100.0
S<A	22.5	-3257.7	-18.1	8.0	-2334.2	-13.5	2.6	-882.4	-32.6
S=A	11.0	-78.7	-0.7	8.8	-110.6	-0.9	0.7	-38.3	-1.4
S>A	17.3	3314.3	27.0	5.6	3474.9	23.8	2.5	1137.4	48.1
Sx0	3.0	8459.3	<i>Inf</i>	2.4	10907.1	<i>Inf</i>	1.2	4200.0	<i>Inf</i>
2009									
0x0	39.4	0.0	0.0	73.4	0.0	0.0	88.6	0.0	0.0
0xA	5.8	-1386.2	-100.0	1.6	-4666.3	-100.0	4.0	-1398.7	-100.0
S<A	22.9	-3255.7	-17.8	8.3	-2580.0	-14.4	2.6	-833.8	-29.4
S=A	11.7	-56.7	-0.5	8.9	-81.9	-0.6	0.7	-31.4	-0.7
S>A	17.7	3177.2	25.9	5.9	3046.7	21.9	2.9	908.2	42.5
Sx0	2.6	7420.0	<i>Inf</i>	1.9	10904.8	<i>Inf</i>	1.2	4200.0	<i>Inf</i>
2010									
0x0	39.3	0.0	0.0	72.5	0.0	0.0	86.7	0.0	0.0
0xA	5.2	-1368.6	-100.0	1.5	-4860.6	-100.0	4.2	-1802.5	-100.0
S<A	22.9	-3197.6	-16.8	9.1	-2465.8	-13.4	4.0	-1003.7	-32.1
S=A	12.4	-86.1	-0.7	9.0	-114.0	-1.0	1.2	-1.2	0.0
S>A	17.3	3093.5	26.0	5.6	3497.2	25.7	3.2	923.2	34.6
Sx0	2.8	5716.7	<i>Inf</i>	2.2	11237.4	<i>Inf</i>	0.7	2800.0	<i>Inf</i>
2011									
0x0	38.5	0.0	0.0	74.2	0.0	0.0	86.6	0.0	0.0
0xA	5.7	-1281.0	-100.0	1.9	-5632.1	-100.0	4.2	-1589.2	-100.0
S<A	22.7	-3083.0	-16.9	1.1	-3455.7	-17.6	3.8	-1020.9	-30.8
S=A	13.3	-121.2	-0.8	20.6	0.0	0.0	1.1	-10.6	-0.5
S>A	17.2	3178.2	26.3	1.0	5249.3	84.9	3.6	1019.9	34.3
Sx0	2.5	4419.5	<i>Inf</i>	1.2	11540.6	<i>Inf</i>	0.6	5220.0	<i>Inf</i>

Notes: This table shows the proportion and the median difference (both in absolute and relative terms) of reported survey income and the income recorded in administrative sources per income type and survey year. (1) 0x0: no income in survey and administrative data. (2) 0xA: only income in administrative data. (3) S<A: reported income is below the value in administrative data (4) S=A: survey income corresponds to administrative data. (5) S>A: reported income is above the value in administrative data. (6) Sx0: only income in survey data. Source: SILC 2008-2011, own calculations.

As we have shown that differences between administrative records and survey responses are relevant for various income components, we aim to ascribe the occurring mismatches to the above-mentioned reasons for misreporting. Thus, we are interested in the effect of (i) social desirability, (ii) socio-demographic characteristics, (iii) aspects of the survey design, and (iv) learning effects on the presence, direction and extent of misreporting of three components of total disposable household income (wages, pensions, and unemployment benefits).

In the empirical analysis, we use different approaches to shed light on these issues. When studying the reasons (i) to (iv) for the *direction* of misreporting, we distinguish between those with practically correct information, over-, and underreporters, and consider both the positive and negative mismatch. In contrast, when analysing the impact of these reasons on the *extent* of

misreporting, we focus on the metric *difference* (survey minus register).

The right hand side in our econometric specification comprises the variables describing reasons (i) to (iv) for misreporting. The presence of (i) social desirability bias is indicated by the respondent’s position in the respective income distribution specified by the income decile in the register data. The explanatory variables referring to (ii) the socio-demographic characteristics of the respondent are gender, educational attainment according to the ISCED classification, age, and a categorical variable referring to the health status. We also include the country of birth, the degree of urbanisation at the place of residence, and the employment status with the following options: full-time employed, part-time employed, full-time entrepreneur, part-time entrepreneur, unemployed, retired, domestic worker, student or other. Additionally, dummy variables indicating the number of changes in the employment status during the income reference period should capture the stability of employment. Finally, depending on the type of income investigated, we include the number of months being either full- or part-time employed, retired, or unemployed. The distribution of over- and underreporters across the socio-demographic characteristics is shown in Table B.1.

The explanatory variables related to (iii) the survey setting comprise a dummy variable for the interview mode (CAPI/CATI), a dummy variable related to the reporting status (proxy vs. self-reported income), and a categorical variable specifying the month of the interview (March to May, June to August, September to November). The motivation for the latter variable is straightforward: the earlier the interview took place, the shorter is the time span between the income reference period and the income reporting. The distribution of the response categories across the survey setting variables is given in the lower part of Table B.1. The indicator for (iv) the learning effect is a dummy variable for the interview wave ranging from one to four. Individuals participating in more than one waves are a priori expected to have more experience with income surveys and tend to provide more reliable responses.

3.2 Method

Our empirical strategy is a three-step procedure. First, we apply a multinomial logit regression to assess the impact of the single reasons on the direction of mismatch and thus the probability to over- or underreport income. Second, we enrich the analysis with panel regressions to estimate the effect of the single reasons on the extent of misreporting while controlling for unobserved individual characteristics. Third, we determine the relative importance of the error sources for misreporting with an Owen decomposition. In the following, we describe each of the three steps briefly.

Multinomial logit. With a multinomial logit model, we search for factors that help us to better understand why self-reported incomes are above or below their corresponding administrative records. Our dependent variable is the mismatch category $Pr(Y_{i,k} = j)$ and we calculate the probabilities for reporting less ($S < A$), the same ($S = A$), or more ($S > A$) than his or her true income. Although strictly speaking, respondents who do not report income in both sources (0×0) have no mismatch, this group of observations is not of interest for our main research questions and therefore discarded from the analysis. Additionally, due to the very low number of observations without income information either in the survey ($0 \times A$) or in the administrative register ($S \times 0$), we restrict our attention to

overreporters, underreporters, and the consistent group.¹

By estimating a multinomial logit model via maximum likelihood, we explicitly allow the estimated coefficients to vary across the mismatch categories. Thereby, we are able to identify the determinants of mismatch separately for overreporters and underreporters. This is a considerable advantage compared to OLS regression, since it may be very hard to defend the assumption that the variables influencing underreporting are similarly affecting the probability of overreporting. Even more, in the standard OLS framework all reporting errors are pooled together. Consequently, assuming parameter homogeneity across mismatch categories could not only lead to misleading interpretations, but positive and negative errors could potentially cancel out and leave us with statistically insignificant estimates. In contrast, the effect of e.g. gender or education on the probability of misreporting is allowed to differ between over- and underreporters in the multinomial logit model. This flexibility enables us to draw a more detailed picture of the factors that influence the response behaviour of individuals in income surveys.

Panel regression. Although we consider an extensive and diverse set of control variables, we cannot rule out that our estimation lacks relevant but unobservable determinants. To check the robustness of our findings, we thus make use of the longitudinal dimension of EU-SILC from 2008 to 2011 and employ fixed-effect estimations. The focus on within-individual changes makes it possible to control for individual characteristics that are unobserved but supposedly constant over time, such as the cognitive ability to answer interviews or past experience with surveys. To purge these unobservable characteristics in two related specifications, we apply panel OLS regression models with individual and time fixed effects.²The dependent variable is the difference between survey and administrative records for each person-year.³

In the first panel specification, our primary interest is the influence of social desirability on the difference between the survey report and the administrative value. We expect a clear pattern across income deciles, with a positive mismatch (i.e. overreporting) in the lower parts of the distribution and a growing negative mismatch (i.e. underreporting) as we approach the top income earners. In the second panel specification, we are particularly interested in the learning effect where only the absolute mismatch is relevant. The question is whether individuals participating in multiple survey waves tend to decrease reporting errors and repeated interrogations are associated with a statistically significant learning effect over time. The set of explanatory variables resembles the multinomial models apart from gender and the educational level, which both show negligible within-variation among adults.

¹We provide estimates for linear probability models in table C.1 in the online appendix. For each income type, three linear probability models were estimated. One for those reporting less ($S < A$), one for those reporting the same ($S = A$) and one for those reporting more ($S > A$) income in the survey. In each model the dependent variable is a dummy variable indicating if a certain mismatch type is observed.

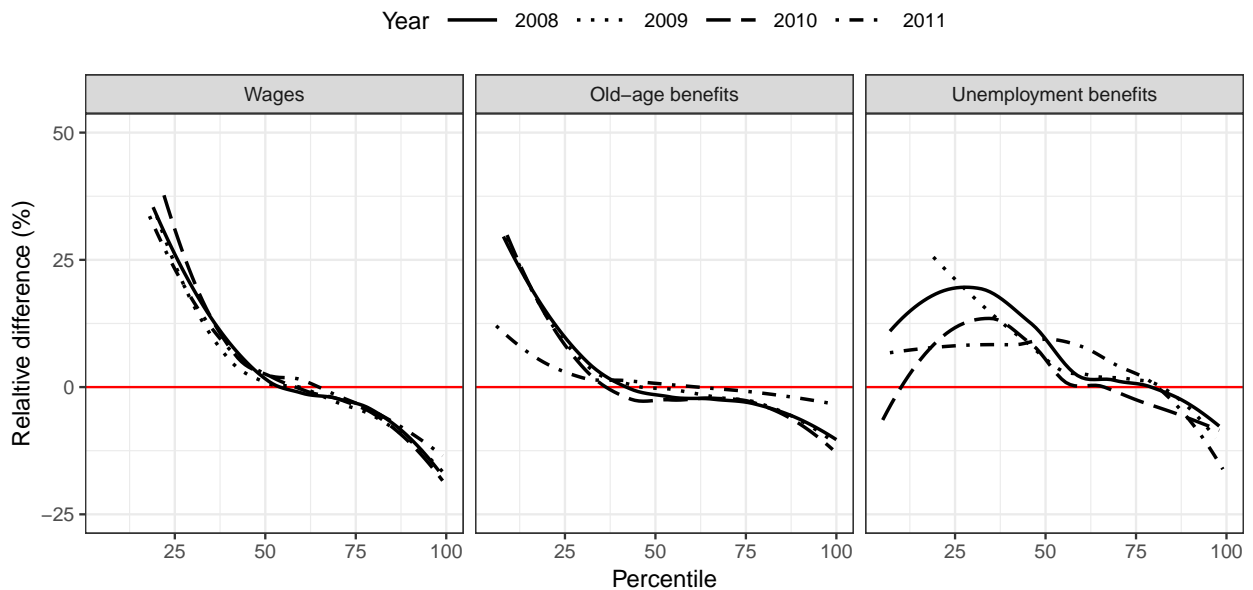
²In table B.9 in the online supplementary materials we report the number and the proportion of observations in our unbalanced panel sample that jumped over a given number of income deciles. For wages we find that 3.348 individuals (37.9 %) experienced a change of at least one decile, in the case of old age benefits 1.174 persons (25.2 %) and in case of unemployment benefits 491 (76.8 %) units changed at least on decile across the years. We conclude that the observed variation is sufficient to identify the corresponding coefficients.

³We prefer a linear OLS specification as this provides a clear interpretation of marginal effects on the original scale. In contrast, for longitudinal (nonlinear) binary and multinomial logit response models with fixed effects, the intuitive interpretation of estimates as predicted probabilities (or various types of marginal effects) is not a viable option because the unobserved heterogeneity vector of person fixed effects is not estimated (see e.g. Pfarr, 2014 for a more detailed discussion)

Owen value decomposition. Finally, we are interested in the relative importance of error sources (i) to (iv) and apply an Owen value decomposition of the explained variance (R^2) in a pooled cross-sectional regression. This procedure allows to estimate the marginal contribution of each group of explanatory variables to the total R^2 . The Owen value decomposition is a generalisation of the Shapley value decomposition and is suitable to assert the relative importance of groups of regressors for explaining the variance of the dependent variable (Huettner, 2012). To the best of our knowledge, such an assessment has not yet been done in the literature on income measurement error. To assess both time-varying and time-constant explanatory variables, the decomposition is based on pooled cross-sectional models for all four years with the difference between survey and register data as the dependent variable.

4 Results

Before turning to the results of the econometric exercise, we study the unconditional existence of a social desirability bias and a learning effect in the data.



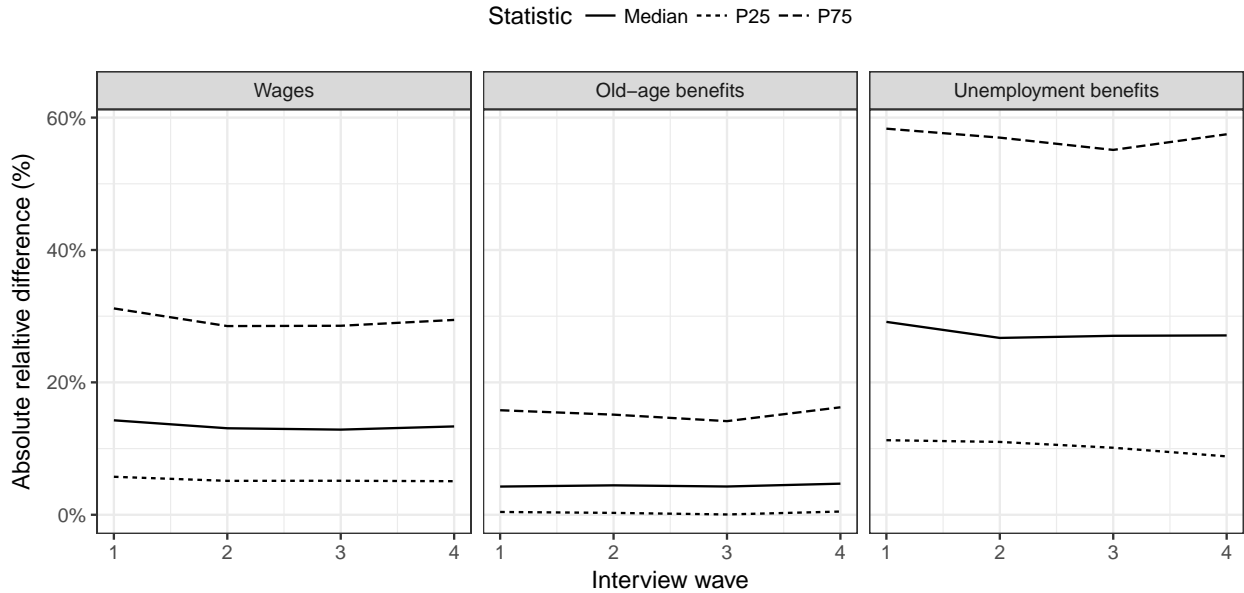
Notes: This graph shows the mean relative difference between survey responses and register information on wages, old-age and unemployment benefits by percentile of corresponding income distribution and year of the interview. Positive values in the lower half of the distribution imply structural over-reporting, whereas negative values above the median indicate under-reporting in surveys at the top. Source: SILC 2008-2011, own calculations.

Figure 1: Social desirability – SILC 2008-2011

Figure 1 shows the average relative difference between the survey response and the register entry per percentile of the distribution of wages, pensions, and unemployment benefits. Overall, the typical pattern of a mean-reverting error is visible. Thus, respondents in lower income percentiles of a specific income type report values that are higher than their register records and vice versa for respondents at the upper part of the distribution.⁴ The pattern is most distinctive for wages and least pronounced for unemployment benefits, but in general corresponds to the expected social desirability behaviour.

⁴The figure is cut between a relative mismatch of +50% and -25%.

Figure 2 illustrates the learning effect over the four survey waves. The lines depict the absolute logarithmic difference between the two data sources in the mean, the median, the 25th, and the 75th percentile. For wages, we find a slight reduction in the differences for all observed income quantiles after the first wave. A small error reduction over time is also present in the 25th quantile and the median of unemployment benefits. In contrast, for pension income, the error does not seem to decrease over time. Thus, the data displays no or, if anything, a very small learning effect.⁵



Notes: This graph shows the median as well as the 25th and 75th percentile of the absolute values of relative difference between survey responses and register information on wages, old-age and unemployment benefits by wave of the interview (i.e. negative values have been multiplied by -1). A reduction of absolute errors over time would be evidence for a learning effect in repeated interrogations, however the data display no strong support for this hypothesis. Source: SILC 2008-2011, own calculations.

Figure 2: Learning effect – SILC 2008-2011

In what follows, we will test whether the observed patterns of mean-reverting errors and the learning effect remain valid in a multivariate model. Given that we control for a variety of variables capturing the complexity of the annual income stream (e.g. employment status changes), a significant effect of mean-reverting errors would emphasise the role of social desirability as an important error source.

4.1 Likelihood of reporting more or less

Social desirability. For ease of interpretation, table 5 displays average marginal effects derived from the multinomial logit models.⁶ For all three types of personal income, the estimates for the corresponding income deciles—with the 5th decile as reference category—confirm the mean-reverting error. Thus, compared to the middle of the wage distribution, the likelihood of underreporting is significantly lower in the bottom deciles and significantly higher in the top deciles. Vice versa, the probability of overreporting increases by up to 51 percentage points in the first decile and is significantly lower in higher deciles.

⁵Figure B.1 in the Appendix replicates figure 2, however, using only observations that remained in the panel in all four waves, i.e. from 2008 to 2011. With this sample the outcome for pensions and unemployment benefits resembles that of the unbalanced sample. For wages, the learning effect almost disappears.

⁶Full tables with the estimates expressed in log(odds) are available from the authors upon request.

For pensions, this pattern is very similar although less pronounced since the average marginal effects of income deciles on both underreporting and particularly overreporting are smaller. Finally, recipients of unemployment benefits are no exception from the general pattern. Higher unemployment benefits correspond to a higher likelihood of underreporting and a lower likelihood of overreporting. Effects are statistically significant at the 5% level for almost all deciles. If we look at the probability of roughly similar income levels in both sources ($S = A$), lower income groups tend to report correct wages less often than middle-income groups. To a smaller extent, this also applies to the highest deciles. For pensions, changes in the probability of $S = A$ are more symmetrically spread around the middle whereas for unemployment benefits, income deciles generally do not have strong statistically significant effects.

Summing up, we interpret these results as evidence of an income-mean-reverting type of social desirability. The estimations also reveal that the mean-reverting pattern does not exclusively apply to wages but seems to be present also for non-market income and transfer payments.

Socio-demographic characteristics. Males have a significantly higher/lower tendency of overreporting/underreporting all three types of income. Overall, the gender-specific effect is largest for wages. Related to that, men display a slightly lower propensity to report matching values for wages. These results could reflect some underlying male breadwinner/masculinity norm which *ceteris paribus* renders men to overreport market income more often. Based on this argument, we also expected rather men to conceal receipt of non-market transfers. However, our estimates for unemployment benefits do not provide support for the latter hypothesis. Concerning education, we find significant differences between respondents with higher educational attainments compared to compulsory education for wages and pensions. Underreporting decreases whereas overreporting increases with the educational level. For the point estimates it also does not make much difference whether respondents hold a post-secondary or a tertiary degree. Contrary to what we have expected for unemployment benefits, there is no evidence that underreporting is an increasing function of education. All education dummies are statistically insignificant in this model. In line with the results for social desirability (income percentiles) and gender, it seems that there is no big social stigma related to levels of unemployment transfers (assuming that the other control variables capture cognitive errors sufficiently). Instead, it is possible that these transfers are generally regarded as legally acquired insurance payments. Age only exerts a statistically significant, albeit very small negative effect on the likelihood of underreporting unemployment benefits.

Table 5: Multinomial logit regressions

	Wages			Pensions			Unemp. benefits		
	SA	SA	SA
(i) Social desirability									
<i>Relative income (Ref: 5. Decile)</i>									
1. Decile	-28.3	-21.5	49.8	-15.2	-16.8	32.0	-22.4	-5.0	27.4
2. Decile	-17.8	-17.0	34.7	-4.3	-5.8	10.1	-11.0	-3.0	14.0
3. Decile	-11.2	-12.1	23.3	-4.2	-1.5	5.7	-0.3	<i>-6.5</i>	6.7
4. Decile	-2.0	-7.7	9.7	0.9	-3.9	3.0	-1.0	<i>-5.5</i>	6.5
6. Decile	8.5	-0.8	-7.7	8.7	-6.5	-2.2	<i>10.4</i>	1.8	-12.1
7. Decile	14.2	-1.3	-12.9	8.1	-7.6	-0.5	<i>8.9</i>	-0.5	-8.4
8. Decile	18.6	-3.3	-15.3	9.5	-7.9	-1.7	17.1	1.0	-18.1
9. Decile	27.1	-7.4	-19.6	12.7	-8.0	-4.7	20.8	<i>7.6</i>	-28.3
10. Decile	38.6	-14.7	-23.9	24.2	-19.1	-5.1	31.4	5.6	-37.1
(ii) Sociodemographic characteristics									
<i>Gender (Ref: Female)</i>									
Male	-10.7	-3.7	14.5	-4.8	-0.8	5.6	<i>-5.3</i>	-2.3	7.6
<i>Age</i>									
Age	0.0	0.0	0.0	-0.2	0.0	0.2	<i>-0.2</i>	0.2	0.0
<i>Education (Ref: Compulsory)</i>									
Upper secondary	-8.1	0.7	7.5	-5.4	<i>2.5</i>	2.8	4.2	-0.4	-3.8
Post-secondary	-8.8	-0.3	9.0	-6.9	2.6	<i>4.2</i>	-5.6	-0.5	6.1
1 st stage tertiary	-11.0	1.2	9.8	-4.4	-0.5	4.9	2.3	-5.2	2.9
2 nd stage tertiary	-4.1	-3.8	<i>7.9</i>	1.4	1.8	-3.2	-7.2	7.3	-0.2
<i>Country of birth (Ref: AUT)</i>									
EU15	-4.1	0.2	3.9	-2.8	-8.5	11.3	-0.4	0.6	-0.2
CEE	-2.7	0.4	2.3	0.4	-1.3	0.9	10.5	-7.4	-3.1
Turkey	<i>6.6</i>	-1.0	-5.6	4.1	-0.8	-3.2	-4.7	0.4	4.4
Yugosphere	5.1	-5.1	-0.1	4.9	-6.1	1.2	<i>6.7</i>	<i>-5.9</i>	-0.8
other	7.9	-5.2	-2.7	0.1	2.4	-2.6	-1.5	-1.1	2.5
<i>Health status (Ref: Very bad)</i>									
Bad	-5.1	2.5	2.7	-2.5	<i>5.7</i>	-3.2	10.7	0.5	-11.2
Fair	-4.2	0.1	4.1	-2.7	7.8	<i>-5.1</i>	2.9	4.8	-7.7
Good	-8.0	1.7	6.3	-4.3	9.9	<i>-5.6</i>	6.8	1.7	-8.5
Very good	<i>-12.9</i>	1.8	<i>11.1</i>	-8.8	10.8	-2.0	5.1	5.0	-10.1
<i>Degree of urbanisation (Ref: < 10 000 inhabitants)</i>									
> 10000 & < 100000	-1.6	<i>2.1</i>	-0.5	<i>-4.1</i>	<i>3.5</i>	0.5	-3.7	1.6	2.0
>100 000 inhabitants	-4.6	1.6	3.1	-4.8	<i>3.1</i>	1.7	-8.7	3.1	<i>5.6</i>
<i>Changes in employment status (Ref: None)</i>									
Once	<i>4.6</i>	<i>-4.4</i>	-0.2	30.8	-24.0	-6.8	8.6	-5.3	-3.2
Twice	<i>4.7</i>	-0.7	-4.0	30.5	-29.1	-1.4	9.2	<i>-6.2</i>	-3.0
Thrice or more	-0.6	-6.3	<i>6.9</i>	15.7	-9.7	-6.1	<i>13.1</i>	<i>-8.1</i>	-5.0
<i>Months in corresponding employment status (Ref: 12 months)</i>									
<6 months	27.4	-8.1	-19.3	4.0	<i>-6.5</i>	2.5	40.8	3.5	-44.3
6-8 months	11.7	<i>-5.6</i>	-6.1	-9.1	2.2	6.9	16.4	3.4	-19.8
9-11 months	3.5	-6.6	3.1	4.0	-12.0	8.0	<i>10.7</i>	3.6	<i>-14.4</i>
(iii) Survey setting									
<i>Mode of interview (Ref: CAPI)</i>									
CATI	-2.2	2.9	-0.7	-2.1	2.8	-0.7	-5.6	<i>7.5</i>	-1.8
<i>Interviewer</i>									
Same interviewer	0.9	1.1	-2.0	0.1	1.6	-1.7	-8.4	6.9	1.5
<i>Type of interview (Ref: Personal)</i>									
Proxy	7.9	-5.8	-2.2	3.8	-5.6	1.8	0.3	-0.7	0.4
<i>Month of interview (Ref: March to May)</i>									
June to Aug.	2.3	-2.9	0.6	1.3	<i>-2.5</i>	1.2	-0.2	0.2	0.0
Sept. to Nov.	2.2	-5.7	<i>3.5</i>	8.6	-9.6	1.0	0.2	0.6	-0.8
<i>Year of interview (Ref: 2008)</i>									
2009	-0.2	-0.8	1.0	0.3	-2.7	2.4	-2.3	-1.7	4.0
2010	0.5	0.1	-0.6	2.8	<i>-3.7</i>	0.9	2.1	3.6	-5.7
2011	-0.7	-0.2	0.9	-30.4	46.9	-16.5	1.1	1.5	-2.6
(iv) Learning effect									
<i>Wave of interview (Ref: 1st)</i>									
2 nd	-1.4	0.4	1.0	<i>3.0</i>	-3.1	0.1	<i>7.4</i>	-5.8	-1.5
3 rd	-1.8	0.8	1.0	2.0	<i>-4.5</i>	2.5	4.2	<i>-5.9</i>	1.8
4 th	-2.8	2.0	0.7	2.1	-2.3	0.2	4.0	-1.7	-2.3

Notes: This table shows the estimated average marginal effects of multinomial regressions per income type (wages, pensions and unemployment benefits) in three categories: (1) S<A: reported income is below the value in administrative data (2) S=A: survey income corresponds to administrative data. (3) S>A: reported income is above the value in administrative data. Estimates with an associated p-value below 0.01 are depicted in bold, values below 0.05 in italics. Insignificant estimates (p-value above 0.05) are given in gray. Source: SILC 2008-2011, own calculations.

Being born in a CEE⁷ country does not make a significant difference for the likelihood of misreporting wages compared to Austria, whereas we find a higher probability of underreporting wages for the Yugosphere⁸, Turkey and other countries of birth (and a corresponding lower propensity to report matching wages). For the remaining two income components, the country of birth is less relevant with two exceptions: being born in the EU15 significantly raises the likelihood of overreporting pensions by 11%-points. Furthermore, natives from the Yugosphere bear a higher probability of underreporting unemployment benefits. Summing up, the evidence hints to remaining problems of correctly understanding the data collection process by non-natives but this is primarily an issue for the underreporting of wages.

Health problems are expected to hinder income reporting. The strongest evidence for a decreasing probability of misreporting with improved health is found for pensions. Related to that, respondents tend to provide correct pension incomes the healthier they are. Better health also reduces the likelihood of underreporting wages but this is only significant when we compare those in *very good* health to those with *very bad* health. However, this pattern is completely reversed for overreporting wages. Moreover, there is no evidence for a relationship between health and the misreporting of unemployment benefits.

For all three types of income, underreporting is less prevalent, if the respondent's place of residence is in a highly urbanised region. In the case of overreporting, this relation is insignificant for pensions and inverted for wages and unemployment benefits. As we expected *any* kind of misreporting to be reduced with rising degree of urbanisation, the latter result could e.g. be due to differences in types of jobs and the associated wage structures in urbanised regions.

As an indicator for income stability, we include variables capturing the number of status changes and the number of months spent in a specific employment status. As expected, Table 5 exhibits that status changes correspond to higher probabilities of misreporting and to lower chances to report correct incomes. In general, the effects show significantly higher probabilities for underreporting, particularly for pensions and unemployment benefits. Furthermore, we find that overreporting unemployment benefits is 41%-points more likely if the spell was shorter than 6 months and 16%-points more likely if the spell lasted between 6 and 8 months during the income reference period. An explanation for this finding could be that particularly short spells of unemployment are associated with recall and telescoping errors. There are no significant average marginal effects of the number of months on the probability of a match between the two data sources. Overreporting unemployment benefits is 44%-points less likely for short-term recipients than for long-term recipients. This general pattern is similar for underreported wages but almost nonexistent for pension incomes.

Survey setting. Against what we have expected, average marginal effects of telephone interviews (CATI) on both types of misreporting are generally negative, however, statistically not significant. For wages, responses from CATI are 3%-points more likely to match administrative records than from CAPI. Proxy interviews increase the probability of underreporting wages and pensions. For wages, there is also a marginally negative effect on overreporting. Proxy interviews neither have an effect on overreporting of pensions nor on any type of misreporting of unemploy-

⁷Bulgaria, Cyprus, the Czech Republic, Estonia, Hungary, Latvia, Lithuania, Malta, Poland, Romania, Slovakia, and Slovenia

⁸Bosnia and Herzegovina, Croatia, Serbia and Montenegro, and Macedonia

ment benefits. Moreover, having the same interviewer as in the previous year does not have a major impact on misreporting. Only unemployment benefits are slightly less likely to be underreported.

Looking at the month of the interview, we find that reporting errors for wages are more likely, the later in the year the interview is conducted, i.e. the larger the time span between the income reference period and the interview. For pensions, however, this applies only to more distant interview months (>8 months) and is only statistically significant for underreporting. For unemployment benefits, there is no significant relationship at all. Finally, the interview year generally does not play any significant role. The large and significant time effects for pensions in 2011 partly appear by design because (only) for this year, Statistics Austria already derived a great share of pension incomes from registers before they back-calculated the Austrian SILC.

Learning effect. As already indicated by the descriptive illustration in figure 2, the multinomial logit model based on pooled cross-sections does not yield strong evidence for learning effects. In fact, we find slightly lower probabilities to provide correct answers for pensions and unemployment benefits for respondents in advanced survey waves. In the next step, we will apply panel regressions which additionally purge unobservable individual fixed effects, for this question and test whether this preliminary result holds.

To sum up the results of the logit models, the average marginal effects suggest that the income level and socio-demographic characteristics are far more relevant than the interview context for explaining reporting errors. Relative importance, however, will be investigated more systematically further below. Moreover, the general patterns found for social desirability (income percentiles) and for gender, age, education among the sociodemographic factors are quite similar for all three income types. Concerning the survey setting, effects for wages and pensions are close, whereas unemployment benefits are hardly influenced by this set of variables.

4.2 Extent of misreporting

In a next step, we employ individual fixed effects OLS panel regressions to gain further insights on the presence of social desirability bias and learning effects. The results of the panel estimation for the two subjects are displayed in tables 6 and 7, respectively. Although we present only the estimated coefficients relevant for these two errors sources, the estimations were done using the full set of controls, corresponding to table 5. The remainder of the estimates is contained in tables B.2 and B.3 in the Appendix.

Social desirability. The dependent variable in the panel regression focusing on the evaluation of social desirability is the negative or positive absolute difference between the survey report and the administrative record. We find that the difference between questionnaire and register wages increases with rising distance from the middle of the distribution. This effect is more pronounced at higher percentiles as compared to lower quantiles. For instance, in the lowest decile the average overreporting is 7,400 Euros above the error in the 5th decile, whereas the difference amounts to 11,600 Euros in the top decile. A similar pattern is present for differences in pensions, although only statistically significant at both ends of the distribution. In the case of unemployment benefits, point estimates roughly indicate a mean-reverting pattern in the longitudinal perspective. Only some quantiles show statistically significant differences from the average error in the 5th decile and

Table 6: Panel regression results – Social desirability

	Wages	Pensions	Unempl. Benefits
(i) Social desirability			
<i>Relative income (Ref: 5. Decile)</i>			
1. Decile	7387.25 (1020.64)	3536.36 (764.73)	192.52 (1048.41)
2. Decile	5765.24 (776.71)	1775.13 (684.73)	516.38 (441.90)
3. Decile	3524.27 (698.17)	1098.89 (471.01)	–33.88 (834.57)
4. Decile	1631.32 (362.68)	–136.74 (372.11)	182.32 (471.82)
6. Decile	– 1449.64 (509.38)	–349.92 (330.50)	–625.39 (661.51)
7. Decile	– 2972.27 (653.50)	–1389.06 (982.90)	–799.96 (937.23)
8. Decile	– 4425.72 (841.99)	–2143.49 (1184.84)	– 1387.76 (701.87)
9. Decile	– 8118.60 (1319.65)	– 3420.77 (773.14)	– 1554.50 (705.41)
10. Decile	– 11598.02 (1696.99)	– 6968.90 (1769.55)	– 2839.05 (831.64)
Other controls			
Individual fixed effects	yes	yes	yes
(ii) Sociodemographic characteristics	yes	yes	yes
(iii) Survey setting	yes	yes	yes
(iv) Learning effect	yes	yes	yes
Num. obs.	20372	10105	2470
R ² (full model)	0.70	0.72	0.81
Adj. R ² (full model)	0.32	0.41	0.30

Notes: This table shows the results of unbalanced panel regressions with positive and negative reporting errors as dependent variable (i.e. negative values correspond to under-reporting and positive values to over-reporting). Cluster robust standard errors are given in parentheses. As identification in fixed effect models relies on sufficiently large within variation, the variables gender, education and country of birth have been removed from the baseline specification. See full list of estimated coefficients in Table B.2. Source: SILC 2008-2011, own calculations.

indicate underreporting at the top.

Learning effect. Table 7 presents additional evidence on the presence of learning effects. Note that we define the learning effect as a decline in the absolute reporting errors over multiple survey waves. In this case, we do not distinguish between over- and underreporting and thus negative values of the dependent variable are multiplied by minus one. We find mixed evidence for learning effects, which crucially depend on the type of income under consideration. For wages and unemployment benefits, there is no statistically significant reduction of reporting errors with increasing panel duration, whereas such a pattern clearly emerges for pension incomes. For wages, our results are in line with previous literature finding a positive, although not statistically significant, serial correlation of misreporting.

Robustness checks. We applied two robustness checks for our baseline panel specification: (1) estimations based on the 4-wave balanced sample using longitudinal weights (see tables B.4 & B.5), (2) specifications with the difference in log income between register and survey data as dependent variable (see tables B.6 & B.7). The panel regressions for the balanced sample broadly confirm the social desirability effects although with slightly smaller point estimates. Surprisingly, the already small learning effect for pension income disappears completely in the balanced 4-wave panel. In the second check, we test whether our conclusions remain valid when studying the *relative* deviations of survey answers from administrative records. Overall, the models with the difference of the natural logs of incomes have a higher model fit. Mean-reverting errors found in the baseline model are observed again and have a similar pattern of statistical significance. Again, learning effects are only present for pensions where the model predicts a reduction of the difference between survey and

Table 7: Panel regression results – Learning effect

	Wages	Pensions	Unempl. Benefits
(iv) Learning effect			
<i>Wave of interview (Ref: 1st)</i>			
2 nd	–164.75 (631.32)	– 1622.45 (264.10)	191.71 (300.39)
3 rd	–353.89 (770.74)	– 2546.06 (520.99)	231.58 (433.76)
4 th	–478.33 (959.25)	– 3397.04 (821.01)	144.55 (568.08)
Other controls			
Individual fixed effects	yes	yes	yes
(i) Social desirability	yes	yes	yes
(ii) Sociodemographic characteristics	yes	yes	yes
(iii) Survey setting	yes	yes	yes
Num. obs.	20372	10105	2470
R ² (full model)	0.74	0.71	0.82
Adj. R ² (full model)	0.40	0.39	0.31

Notes: This table shows the results of unbalanced panel regressions with absolute values of the reporting errors as dependent variable (i.e. negative values (under-reporting) have been multiplied by -1). Cluster robust standard errors are given in parentheses. As identification in fixed effect models relies on sufficiently large with-in variation, the variables gender, education and country of birth have been removed from the baseline specification. See full list of estimated coefficients in Table B.3. Source: SILC 2008-2011, own calculations.

register data of approximately 19% in the last panel wave.

4.3 Relative importance of error sources

Finally, we apply an Owen value decomposition in order to assess the relative importance of the four error sources under consideration. The decomposition aims to assign a proportion of the explained variance to groups of the explanatory variables. We consider two connected settings. First, reporting errors and income variables enter the regressions transformed via the inverse hyperbolic sine function. This transformation is closely related to the well known logarithmic transformation, however, it is also defined for negative and zero values. In the context of a significant mass of zeros and negative values among the reporting errors, this is a desirable property as it allows us to consider the same number of observations as in the preceding calculations. The results of this exercise are given in table 8.

The first row contains the adjusted R^2 for each cross-sectional model. The total variance explained is clearly highest for reporting mismatch in unemployment benefits, where 46% can be traced back to the model variables. In contrast, for wage and pension differences the corresponding figures amount to 16% and 11% respectively. Noteworthy, especially for unemployment benefits, but also in the case of wages and pensions, the magnitude of explained variance in our regression is comparatively high (Kim and Tamborini, 2012). The remaining rows show the group sums of Owen values as percentage of the overall R^2 . While the patterns for wages and pensions are rather similar, unemployment benefits show a quite different picture. For the former two, around 30% of the explained variance can be attributed to the group of socio-demographic variables, whereas social desirability turns out to be of the highest relative importance (around 60%). For wage differences, the survey setting and variables measuring the panel participation (and thus learning effects) virtually do not contribute to the total R^2 at all. Learning effects also play a minor role for pensions whereas the survey setting contributes roughly 15%.

Outcomes are considerably different for unemployment benefits. With a share of 65%, the group of socio-demographic variables is most relevant for overall R^2 . Compared to the models for wages and pensions, social desirability is substantially less important whereas the socio-demographic characteristics gain relevance. Thus, misreporting of unemployment benefits does not so much depend on the level of unemployment benefits but rather on socio-demographic characteristics of the recipients and is also more sensitive to the interview context and mode.

Table 8: Decomposition of explained variance – SILC 2008-2011

	Wages	Pensions	Unempl. Benefits
Proportion of variance explained	15.8	11.4	46.4
<i>Relative importance:</i>			
(i) Social desirability	65.1	58.8	23.7
(ii) Sociodemographic characteristics	34.0	24.5	64.5
(iii) Survey setting	0.8	14.5	9.2
(iv) Learning effect	0.0	2.2	2.6

Notes: This table shows the goodness-of-fit of OLS regressions and its decomposition to four error sources, i.e. four groups of explanatory variables. We quantify the relative importance of (i) social desirability, (ii) socio-demographic characteristics, (iii) aspects of the survey design, and (iv) learning effects on the basis of separate regressions for wages, pensions and unemployment benefits. Reporting errors are regressed on the same set of explanatory variables that were used before (see section 3), using all available pooled cross-sections. Error and income variables are transformed via the inverse hyperbolic sine function, which facilitates a log-log interpretation in the context of a significant mass of zeros and negative values among the errors. Source: SILC 2008-2011, own calculations.

Additionally, we repeat the same procedure without transforming the input variables (see table B.8 in the online appendix). Whereas the results on old-age and unemployment benefits are almost identical, the R^2 of the wage regression drops by two thirds. However, our estimates on the relative importance of the four error sources are hardly affected, which strengthens our confidence in the robustness of our findings.

5 Conclusions

Income is very likely one of the most pervasive information in micro datasets, since it plays an essential role for a wide range of welfare indicators and policy questions. The traditional way of collecting income information are household surveys, however, the accuracy of survey data has increasingly been contested during the last years. A main factor behind this critique is the suspected presence of measurement error in surveys, resulting from (un)intentional misreporting. The identification of data errors requires by definition some point of reference. We follow the traditional literature and check survey data against administrative records using a unique dataset: the Austrian 2008-2011 waves of EU-SILC. We make use of the fact that due to a legal initiative, the Austrian SILC provides both survey and register income data for the same observational units for four consecutive years.

While the vast majority of existing research assesses measurement error in income data for US households, there is virtually no research using European panel data for various income types. EU-SILC is a key dataset for social policy issues since it provides the main indicators for evaluating the Europe 2020 strategy. Given its importance as reference source for comparative statistics on

income distribution and social inclusion, data quality is a crucial matter. Compared to previous literature, using the Austrian EU-SILC for assessing income measurement error has two main advantages. First, we do not have to fall back on two-sample matching processes since agreement from respondents concerning data linkage was not legally required. This advantage helps to avoid selection bias and, given that the Austrian EU-SILC is representative for the national population residing in private households, ensures high external validity. Second, we are able to evaluate income measurement error for various components of total disposable household income in the very same dataset.

We elaborate four major reasons for misreporting discussed in the literature: social desirability, socio-demographic characteristics of the respondent, specifics of the survey design and the presence of learning effects for three types of personal income (wages, pensions and unemployment benefits).

The main findings are the following. For personal income and in line with the existing literature, statistically significant mean-reverting errors are revealed in both, cross-sectional and panel regression models. We find significantly lower probabilities of underreporting at the bottom of the wage distribution and vice versa, higher likelihoods of overreporting at the top tail. By including a broad range of control variables in order to capture the complexity of the annual income stream, we interpret this result as evidence of social desirability in reporting wages. Although the effects are generally less pronounced and sometimes even statistically insignificant, similar patterns occur for pension income and unemployment benefits.

Concerning socio-demographic characteristics, males are found to have a significantly higher tendency of overreporting wages, pensions, and unemployment benefits. Additionally, there is a significant relationship between health and misreporting, for instance, good health conditions correlate with more correct survey responses particularly for pensions. The higher chances of underreporting for respondents being born outside the EU-15 hint to the presence of language and comprehension problems, despite the fact that interviews were also conducted in other languages if requested by respondents. We find consistent evidence that a complex income and employment context with many changes during the income reference period hinders recall, and thus increases misreporting. Multiple changes in the employment status have a strong effect on reporting errors of pensions, a shorter status duration increases errors particularly for unemployment benefits and wages.

For survey designers, the Owen value decomposition might be of particular interest since it reveals that social desirability is a major explanation for misreporting wages and pension income. For unemployment benefits, socio-demographic characteristics of the respondents seem to play the major role for reporting errors. The survey setting is relatively less important for explaining misreporting, while learning effects are hardly noticeable. Our findings from the Owen value decomposition suggest that data producers should be even more aware of social desirability when constructing interview questionnaires. The order and wording of questions on income could incite or inhibit erroneous income reports. It is crucial that survey responses should be validated by actual income proofs, such as pay slips. If wages of employees are available online in web portals of the financial authorities (like in Austria), data producers could possibly push on cross-checking with these sources. Moreover, proxy interviews show significantly higher probabilities of misreporting than personal interviews. Similarly, the larger the time span between the income reference period and the interview, the more likely is misreporting. Evidence on the presence of a learning effects is

mixed and depends crucially on the type of income under consideration.

Survey users may be interested in the various types of bias that result from measurement error in linear models. In table B.10 of the supplementary materials, we present estimates of three types of bias: bias in OLS estimators due to measurement error in the explanatory variables, bias in IV estimators due to measurement error in the explanatory variables, and bias due to measurement error in the dependent variable. Quantifying the bias that is associated to misreporting gives data users and practitioners some guidance on what to expect when they use (or read publications using) income data from SILC in regression models. For instance, we find that all results estimating returns (e.g. of education) on wages are underestimated by about 25%. However, we advise against interpreting these results as general attenuation bias of survey income data. The reasons for and structure of misreporting might vary considerably across countries.

For policy makers, our results point to the fact that for some socio-demographic groups — including those most relevant for policy makers — survey income data may potentially be an infirm ground for decision-making. Thus, investments in the development, maintenance, advancement, and accessibility of public administrative data, as an alternative, may pay-off for better targeted policies. Having both survey and register data at hand for the same units opens up several perspectives for further research. For instance, a next step could be to look at the joint distribution of errors for respondents that have more than one of the three income components at once in a given year and check if those who underreport one source are also more likely to underreport another. It could also be worthwhile to replicate our analysis based on gross income values. Another fruitful avenue for further inquiries could be to train a statistical model that allows survey producers to correct the income information based on other unobservable characteristics of the respondents. Moreover, income is a key explanatory variable in social science research. As in Hariri and Lassen (2017), one could therefore check how conclusions for prominent regressions where income is used as right-hand-side variable (health, well-being, etc.) are altered when survey data is replaced by register data.

References

- Abowd, J. M. and M. H. Stinson (2013): “Estimating Measurement Error in Annual Job Earnings: A Comparison of Survey and Administrative Data”. In: *Review of Economics and Statistics* 95 (5), pp. 1451–1467.
- Alm, J. and B. Torgler (2006): “Culture differences and tax morale in the United States and in Europe”. In: *Journal of Economic Psychology* 27 (2), pp. 224–246.
- Angel, S., N. Lamei, and R. Heuberger (2017): “Differences between household income from surveys and registers and how these affect the poverty headcount. Evidence from the Austrian SILC”. In: *Social Indicators Research (forthcoming)*.
- Asamer, E., F. Astleithner, P. Cetkovic, S. Humer, M. Lenk, M. Moser, and H. Rechta (2016): “Quality Assessment for Register-based Statistics - Results for the Austrian Census 2011”. In: *Austrian Journal of Statistics* 45 (2), pp. 3–14.
- Berka, C., S. Humer, M. Moser, M. Lenk, H. Rechta, and E. Schwerer (2012): “Combination of Evidence from Multiple Administrative Data Sources: Quality Assessment of the Austrian Register-Based Census 2011”. In: *Statistica Neerlandica* 66 (1), pp. 18–33.
- Bingley, P. and A. Martinello (2017): “Measurement Error in Income and Schooling and the Bias of Linear Estimators”. In: *Journal of Labor Economics* 35 (4), pp. 1117–1148.
- Bollinger, C. R. (1998): “Measurement Error in the Current Population Survey”. In: *Journal of Labor Economics* 16 (3), pp. 576–594.

- Bound, J., C. Brown, G. J. Duncan, and W. L. Rodgers (1994): “Evidence on the Validity of Cross-Sectional and Longitudinal Labor Market Data”. In: *Journal of Labor Economics* 12 (3), pp. 345–368.
- Bound, J., C. Brown, and N. Mathiowetz (2001): “Measurement Error in Survey Data”. In: *Handbook of Econometrics*. Ed. by J. Heckman and E. Leamer. Vol. 5. Elsevier, pp. 3705–3843.
- Bound, J. and A. B. Krueger (1991): “The Extent of Measurement Error in Longitudinal Earnings Data: Do Two Wrongs Make a Right?” In: *Journal of Labor Economics* 9 (1), pp. 1–24.
- Einav, L. and J. Levin (2014): “Economics in the age of big data”. In: *Science* 346 (6210), p. 715.
- Fessler, P., M. Kasy, and P. Lindner (2017): *Survey mode effects on income inequality measurement*. Working Paper 48766. Harvard University.
- Hansen, K. and D. Kneale (2013): “Does How You Measure Income Make a Difference to Measuring Poverty? Evidence from the UK”. In: *Social Indicators Research* 110 (3), pp. 1119–1140.
- Hariri, J. G. and D. D. Lassen (2017): “Income and Outcomes”. In: *Public Opinion Quarterly* 81 (2), pp. 564–576.
- Hassan, M. and F. Schneider (2016): “Size and Development of the Shadow Economies of 157 Worldwide Countries: Updated and New Measures from 1999 to 2013”. In: *Journal of Global Economics* 4 (3), pp. 1–14.
- Huettner Frank; Sunder, M. (2012): “Axiomatic arguments for decomposing goodness of fit according to Shapley and Owen values.” In: *Electronic Journal of Statistics* 6), pp. 1239–1250.
- Jenkins, S. P., L. Cappellari, P. Lynn, and E. Sala (2006): “Patterns of Consent: Evidence from a General Household Survey”. In: *Journal of the Royal Statistical Society. Series A* 169 (4), pp. 701–722.
- Jenkins, S. P., A. J. Lynn, and E. Sala (2008): “The Feasibility of Linking Household Survey and Administrative Record Data: New Evidence for Britain”. In: *International Journal of Social Reserach Methodology* 11 (1), pp. 29–43.
- Kapteyn, A. and J. Y. Ypma (2007): “Measurement Error and Misclassification: A Comparison of Survey and Administrative Data”. In: *Journal of Labor Economics* 25 (3), pp. 513–551.
- Kim, C. H. and C. Tamborini (2012): “Do Survey Estimate Earnings Inequality Correctly? Measurement Errors Among Black and White Male Workers”. In: *Social Forces* 90 (4), pp. 1157–1181.
- Kim, C. and C. R. Tamborini (2014): “Response Error in Earnings An Analysis of the Survey of Income and Program Participation Matched With Administrative Data”. In: *Sociological Methods & Research* 43 (1), pp. 39–72.
- Kreiner, C. T., D. D. Lassen, and S. Leth-Petersen (2015): “Measuring the Accuracy of Survey Responses using Administrative Register Data: Evidence from Denmark”. In: *Improving the Measurement of Consumer Expenditures*. Ed. by C. D. Carroll, T. F. Crossley, and J. Sabelhaus. National Bureau of Economic Research Studies in Income and Wealth 74.
- Kreuter, F., G. Muller, and M. Trappmann (2010): “Nonresponse and Measurement Error in Employment Research: Making Use of Administrative Data”. In: *Public Opinion Quarterly* 74 (5), pp. 880–906.
- Lohmann, H. (2011): “Comparability of EU-SILC Survey and Register Data: The Relationship among Employment, Earnings and Poverty”. In: *Journal of European Social Policy* 21 (1), pp. 37–54.
- Lusardi, A. and O. S. Mitchell (2014): “The Economic Importance of Financial Literacy: Theory and Evidence”. In: *Journal of Economic Literature* 52 (1), pp. 5–44. URL: <http://www.aeaweb.org/articles?id=10.1257/jel.52.1.5>.
- Lynn, P., S. P. Jenkins, and E. Sala (2012): “The Impact of Questioning Method on Measurement Error in Panel Survey Measures of Benefit Receipt: Evidence from a Validation Study”. In: *Journal of the Royal Statistical Society. Series A* 175 (1), pp. 289–308.
- Mellow, W. and H. Sider (1983): “Accuracy of Response in Labor Market Surveys: Evidence and Implications.” In: *Journal of Labor Economics* 1), pp. 331–344.
- Meyer, B. D., K. M. Wallace, and J. Sullivan (2015): “Household Surveys in Crisis”. In: *Journal of Economic Perspectives* 29 (4), pp. 199–226.
- Micklewright, J. and S. V. Schnepf (2010): “How Reliable are Income Data Collected with a Single Question?” In: *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 173 (2), pp. 409–429.
- Moore, J. C., L. L. Stinson, and E. J. Welniak Jr. (2000): “Income Measurement Error in Surveys: A Review”. In: *Journal of Official Statistics* 16 (4), pp. 331–362.
- Paulus, A. (2015): *Tax evasion and measurement error: An econometric analysis of survey data linked with tax data*. Technical Paper ISER Working Paper Series 2015-10. Institute for Social and Economic Research (ISER), University of Essex.
- Pedace, R. and N. Bates (2000): “Using Administrative Data to Assess Earnings Reporting Error in the Survey of Income and Program Participation”. In: *Journal of Economic and Social Measurement* 26 (3,4), pp. 173–192.

- Pfarr, K. (2014): “femlogit - Implementation of the multinomial logit model with fixed effects”. In: *The Stata Journal* 14 (4), pp. 847–862.
- Pischke, J. (1995): “Measurement Error and Earnings Dynamics: Some Estimates from the PSID Validation Study”. In: *Journal of Business & Economic Statistics* 13 (3), pp. 305–314.
- Reynolds, J. and J. Wenger (2012): “He Said, She Said: The Gender Wage Gap According To Self and Proxy Reports in the Current Population Survey.” In: *Social Science Research* 41 (2), pp. 392–411.
- Sakshaug, J. W. and S. Eckman (2016): “Are Survey Nonrespondents Willing to Provide Consent to use Administrative Records?” In: *Public Opinion Quarterly* 81 (2), pp. 495–522.
- Sakshaug, J. W., M. P. K Couper, M. B. Ofstedal, and D. R. Weir (2012): “Linking Survey and Administrative Records: Mechanisms of Consent”. In: *Sociological Methods and Reserach* 41 (4), pp. 535–569.
- Schnitzer, M., F. Astleithner, P. Cetkovic, S. Humer, M. Lenk, and M. Moser (2015): “Quality Assessment of Imputations in Administrative Data”. In: *Journal of Official Statistics* 31 (2), pp. 231–247.
- Statistics Austria (2014a): *Methodenbericht zur Rückrechnung von EU-SILC 2008-2011 auf Basis von Verwaltungsdaten*. Tech. rep. Statistics Austria.
- (2014b): *Standard-Dokumentation Metainformationen. (Definitionen, Erläuterungen, Methoden, Qualität) zu EU-SILC 2011*. Tech. rep. Statistics Austria.
- Tamborini, C. R. and C. Kim (2013): “Are Proxy Interviews Associated with Biased Earnings Reports? Marital Status and Gender Effects of Proxy”. In: *Social Science Research* 42 (2), pp. 499–512.
- Tourangeau, R., N. M. Bradburn, P. V. Marsden, and J. D. Wright (2010): “The Psychology of Survey Response”. In: *Handbook of Survey Research* 2), pp. 315–346.
- Tourangeau, R., L. J. Rips, and K. Rasinski (2000): *The Psychology of Survey Response*. Cambridge University Press.

A Data appendix

a) Administrative sources of income variables

i) Wages (variable `py010`):

The starting point for the wage information in the administrative data are the pay slips that employers have to transmit to the fiscal authorities administrated by the Austrian Ministry of Finance. Further, the Austrian Federal Computing Center (BRZ) receives the raw pay slips from the Ministry and forwards them electronically to Statistics Austria. To avoid processing errors, the institute then undertakes a broad range of quality checks and discards invalid records. Full details on the process of data transmission by employers and data cleaning by the public authorities are available in sections 2.2. and 3. of the [Standard-Dokumentation Metainformationen zur Lohnsteuerstatistik](#) (only in German unfortunately).

Three different datasets are used to construct the wage variable (see also [Statistik Austria 2014](#))

1. The main source is the Austrian Wage Tax Data (Lohnsteuerstatistik) which contains information on all taxable earnings of employees and pensioners. The dataset includes wages, public pensions (retirement benefits), paid maternity leave (8 weeks before and after expected birth date) and sickness benefits. The data comprises information on gross income, paid social security contributions and retained payroll taxes. Moreover, care allowances are available in the data, although they are not taxable in Austria.
2. The second source is the Austrian Social Security Data. While this source does not supply income information, it indicates the social security status, e.g. whether individuals are eligible for unemployment benefits, disability benefits etc. Specifically, it is used to calculate the income from compulsory military (6 months) or civil service (9 months) for male citizens. These earnings are tax free and thus not included in the wage tax data. Statistics Austria calculate the monthly income based on their social security status and the average payment in the standardized scheme.
3. The third source is the Austrian Tax Adjustment Dataset, which includes total repayments or receipts for tax adjustment in a given year for employee earnings. This data is important for employees living in Austria but working abroad (cross-border commuters). This group is not comprised in the wage tax data; however, these employees have to declare their income in tax returns. Finally, Statistics Austria identifies cross-border commuters from these tax returns and imputes their incomes. Note, that we exclude fully imputed observations from our sample.

In Austria, wages are contractually agreed between employers and employees and underlie certain controls by supervising bodies (like social security authorities or financial police). When the employer registers an employee for social security, the wage has to be stated. At this point, tax avoidance is practically impossible. Employers are obliged to transmit the pay slips directly to the tax authorities each month, while employees only receive the net wage after taxes and social contributions.

In EU-SILC, there are efforts to capture unregistered income (typically cash-in-hand jobs) that are not reported in administrative data. After running through all income component questions, the survey contains two final questions that ask for “further income”. However, concerning such untaxed income of private individuals, e.g. which they generate by doing chores and selling their skills, we argue that this is rather income from self-employment (py050) than income from employment. It is very likely that labor for such activities is sold outside a principal-agent setup. Exactly this reasoning is reflected by Eurostat: the wage income definition of Eurostat (and Statistics Austria) does NOT include such types of income.

ii) Old-age benefits (variable py100):

Statistics Austria uses multiple data sources to record old-age benefits. The main origin of administrative data is the above-mentioned wage tax register which contains all old-age pensions subject to taxation including second pillar pensions. However, the variable py100 also captures transitional payments for long-term unemployed persons shortly before retirement and survivor’s as well as accident benefits for persons older than the official retirement age. The first variable is delivered by the Public Employment Service Austria (AMS), the latter two are provided by the “Dataset for benefits to accident victims and surviving dependents”. The numbers in these datasets correspond to actual payments and thus are practically error-free. ([Statistik Austria, 2014, p 598](#))

Tax avoidance is not an issue for old-age benefits as they are managed and paid out by public authorities. There is no incentive to report lower pension payments than actually are paid out.

iii) Unemployment benefits (variable py090)

For this variable, only one public dataset is used: the so-called “transfers dataset on unemployment benefits” from the Public Employment Service Austria (AMS). It contains the beginning and ending date of unemployment spells and thus the benefits on a daily basis. The AMS also manages the payment of these benefits and thus there should be virtually no errors in the data. Based on a correspondence with data managers from the AMS, the data generating process for the unemployment benefits can be summarized as follows:

The AMS IT generates a clearly identifiable entry for each client (defined as those being eligible for unemployment benefits). Numerous validity checks with other public IT systems and databases (e.g. with data from social security) ensure that neither double entries (for the same person) nor identifier confusion occur. The AMS then calculates the level of payments for each beneficiary and transfers these figures to the Austrian Federal Computing Centre (BRZ). The BRZ recalculates benefit claims and sends the final results for each client back to the AMS IT. The latter data transfer (fully automatic without human intervention) constitutes the final data entry that is also relevant for our analysis. Furthermore, the BRZ also makes the final money transfer to the beneficiaries. There is thus an incentive to avoid errors as these cause costs and client complaints. Similar to old-age pensions, there is no incentive for public authorities to report lower unemployment benefits than actually are paid out.

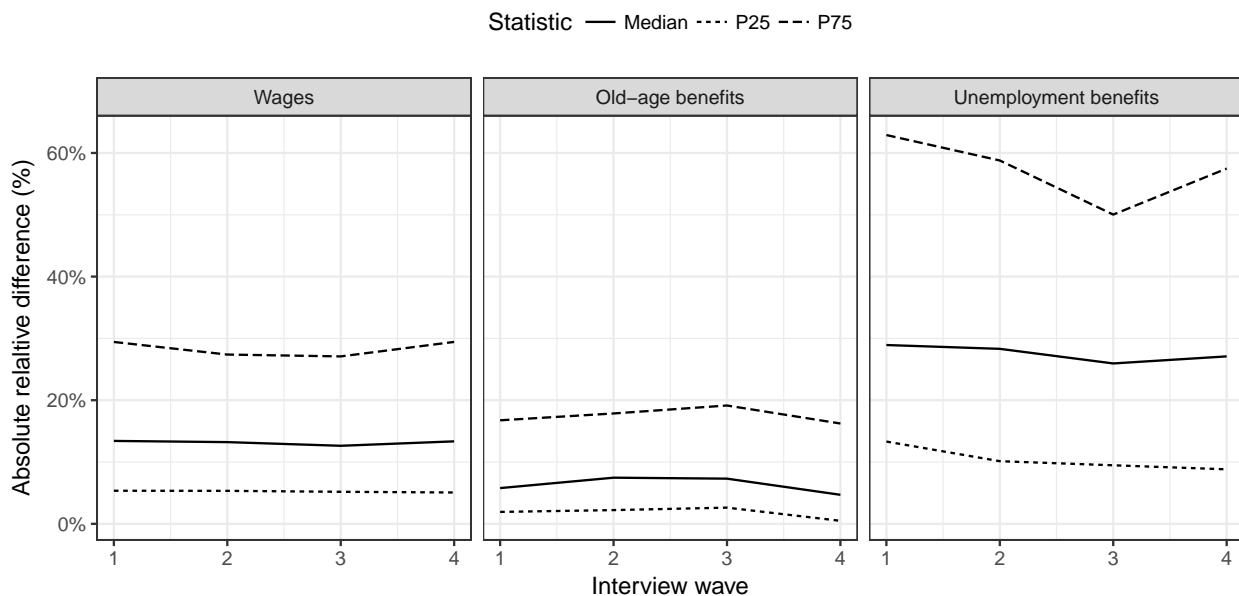
b) Identifier linking process

Austrian law stipulates that the linkage of personal micro data from surveys with registers be done using an anonymized personal identifier (bPIN). These 172-digit PINs are created by the Austrian Data Protection Commission. The identifier then is used as a cryptographically secure way to link survey and register data since it is available for (almost) every person and all registers (without the need to seek consent of the respondents).

Naturally the linkage to administrative data requires valid personal identifiers for the sample population covered by the survey. Usually, this information is already collected as part of the sampling procedure. However, there are always individuals who turn out to be actual household members but are not covered by the sampling frame (mainly because they are not officially registered at a household's address). Their PIN is missing *ex ante* and has to be retrieved by a procedure involving the Federal Ministry for the Interior. The share of available identifiers for the total population in the survey is high and ranges between 96 % (2008) and 99 % (2011). For 2008–2011 missing keys most often occur among younger people, persons living in Vienna (capital) and persons with non-Austrian citizenship. With the exception of EU-SILC 2011, the proportion of missing keys for women was higher than for men among all age groups (by about 1 percentage point). Therefore, using register data results in an under-reporting of income data for those groups.

Further details on the link between register and survey data can also be found in section 10.2. of [Heuberger et al. \(2013\)](#).

B Supplementary material



Notes: This graph shows the median as well as the 25th and 75th percentile of the absolute values of relative difference between survey responses and register information on wages, old-age and unemployment benefits by wave of the interview (i.e. negative values have been multiplied by -1). The underlying sample is a balanced panel of individuals that participated in the survey four times. A reduction of absolute errors over time would be evidence for a learning effect in repeated interrogations, however the data display no strong support for this hypothesis. Source: SILC 2008-2011, own calculations.

Figure B.1: Learning effect - Participated in all four waves– SILC 2008-2011

Table B.1: Distribution of response categories – SILC 2008-2011

	Wages			Pensions			Unemp. benefits		
	SA	SA	SA
(ii) Sociodemographic characteristics									
<i>Gender</i>									
Female	0.46	0.50	0.43	0.54	0.57	0.59	0.45	0.43	0.39
Male	0.54	0.50	0.57	0.46	0.43	0.41	0.55	0.57	0.61
<i>Education</i>									
Compulsory	0.16	0.14	0.17	0.34	0.35	0.37	0.26	0.30	0.32
Upper secondary	0.63	0.67	0.67	0.53	0.54	0.52	0.64	0.59	0.57
Post-secondary	0.06	0.05	0.05	0.07	0.07	0.07	0.04	0.06	0.05
1 st stage tertiary	0.12	0.12	0.10	0.04	0.03	0.03	0.06	0.05	0.05
2 nd stage tertiary	0.02	0.01	0.01	0.02	0.01	0.01	0.00	0.01	0.01
<i>Age</i>									
Age	40.05	39.27	35.99	70.11	70.63	70.42	38.10	42.08	37.88
<i>Country of birth</i>									
AUT	0.83	0.85	0.81	0.91	0.91	0.86	0.71	0.76	0.68
EU15	0.03	0.03	0.04	0.02	0.03	0.06	0.03	0.03	0.03
EU10	0.02	0.02	0.03	0.03	0.03	0.03	0.04	0.02	0.03
Yugosphere	0.06	0.06	0.07	0.02	0.02	0.03	0.12	0.07	0.10
Turkey	0.02	0.02	0.02	0.00	0.00	0.00	0.05	0.07	0.08
other	0.03	0.02	0.03	0.01	0.01	0.01	0.06	0.05	0.07
<i>Health status</i>									
Very bad	0.01	0.00	0.00	0.05	0.04	0.06	0.02	0.03	0.03
Bad	0.03	0.03	0.02	0.14	0.14	0.16	0.09	0.11	0.10
Fair	0.15	0.13	0.12	0.37	0.37	0.36	0.21	0.31	0.27
Good	0.44	0.43	0.38	0.34	0.34	0.30	0.37	0.29	0.34
Very good	0.38	0.41	0.46	0.10	0.11	0.11	0.30	0.26	0.26
<i>Degree of urbanisation</i>									
<10 000 inhabitants	0.56	0.55	0.55	0.57	0.56	0.56	0.52	0.45	0.43
>10 000 and <100 000 inhabitants	0.17	0.17	0.15	0.16	0.16	0.16	0.17	0.20	0.19
>100 000 inhabitants	0.27	0.29	0.30	0.26	0.28	0.28	0.31	0.36	0.38
<i>Changes in employment status</i>									
None	0.81	0.87	0.75	0.95	0.99	0.96	0.23	0.41	0.38
Once	0.11	0.07	0.14	0.05	0.01	0.03	0.40	0.33	0.33
Thrice or more	0.01	0.01	0.03	0.00	0.00	0.00	0.11	0.06	0.06
Twice	0.06	0.05	0.08	0.00	0.00	0.00	0.26	0.20	0.23
<i>Months in corresponding employment status</i>									
<6 months	0.11	0.07	0.14	0.06	0.03	0.06	0.72	0.52	0.51
6-8 months	0.04	0.03	0.06	0.01	0.00	0.01	0.10	0.11	0.13
9-11 months	0.06	0.05	0.09	0.01	0.00	0.00	0.05	0.08	0.06
12 months	0.79	0.86	0.70	0.92	0.96	0.92	0.13	0.30	0.30
(iii) Survey setting									
<i>Mode of interview</i>									
CAPI	0.60	0.56	0.60	0.63	0.61	0.63	0.61	0.60	0.67
CATI	0.40	0.44	0.40	0.37	0.39	0.37	0.39	0.40	0.33
<i>Type of interview</i>									
Personal	0.81	0.86	0.82	0.82	0.88	0.81	0.86	0.90	0.87
Proxy	0.19	0.14	0.18	0.18	0.12	0.19	0.14	0.10	0.13
<i>Month of interview</i>									
March to May	0.45	0.50	0.45	0.37	0.60	0.37	0.50	0.52	0.47
June to Aug.	0.49	0.45	0.48	0.55	0.37	0.56	0.43	0.43	0.46
Sept. to Nov.	0.06	0.05	0.07	0.08	0.03	0.06	0.07	0.05	0.07

Notes: This table shows the distribution of socio-economic and interview related statistics per income type (wages, pensions and unemployment benefits) in three categories: (1) S<A: reported income is below the value in administrative data (2) S=A: survey income corresponds to administrative data. (3) S>A: reported income is above the value in administrative data.

Source: SILC 2008-2011, own calculations.

Table B.2: Full panel regression results – Social desirability

	Wages	Pensions	Unempl. Benefits
(i) Social desirability			
<i>Relative income (Ref: 5. Decile)</i>			
1. Decile	7387.25 (1020.64)	3536.36 (764.73)	192.52 (1048.41)
2. Decile	5765.24 (776.71)	1775.13 (684.73)	516.38 (441.90)
3. Decile	3524.27 (698.17)	1098.89 (471.01)	–33.88 (834.57)
4. Decile	1631.32 (362.68)	–136.74 (372.11)	182.32 (471.82)
6. Decile	– 1449.64 (509.38)	–349.92 (330.50)	–625.39 (661.51)
7. Decile	– 2972.27 (653.50)	–1389.06 (982.90)	–799.96 (937.23)
8. Decile	– 4425.72 (841.99)	–2143.49 (1184.84)	– 1387.76 (701.87)
9. Decile	– 8118.60 (1319.65)	– 3420.77 (773.14)	– 1554.50 (705.41)
10. Decile	– 11598.02 (1696.99)	– 6968.90 (1769.55)	– 2839.05 (831.64)
(ii) Sociodemographic characteristics			
Age	1041.45 (865.28)	–699.76 (1444.45)	290.60 (697.79)
Age ²	–12.07 (9.63)	5.53 (10.17)	–3.15 (9.69)
<i>Health status (Ref: Very bad)</i>			
Bad	652.46 (1011.68)	24.96 (399.12)	–450.05 (804.55)
Fair	429.82 (894.88)	–210.10 (396.13)	–348.42 (822.66)
Good	387.15 (984.01)	–177.94 (400.96)	–378.85 (594.77)
Very good	748.99 (1061.87)	189.34 (414.51)	–123.89 (708.52)
<i>Degree of urbanisation (Ref: <10 000 inhabitants)</i>			
>10 000 & <100 000	–2491.82 (2885.74)	3930.92 (1754.64)	–164.98 (1639.16)
>100 000 inhabitants	–6859.11 (11983.43)	1405.62 (4139.35)	–223.44 (1191.05)
<i>Changes in employment status (Ref: None)</i>			
Once	–1016.75 (1146.34)	– 2498.45 (1036.00)	– 1587.53 (349.37)
Twice	–812.94 (912.02)	–1702.99 (1079.29)	– 1421.28 (524.98)
Thrice or more	–800.33 (717.63)	725.26 (4035.64)	– 1153.15 (576.20)
<i>Months in corresponding employment status (Ref: 12 months)</i>			
<6 months	– 4456.66 (1048.08)	–280.56 (1450.14)	– 5114.40 (1440.46)
6-8 months	–2098.93 (1309.92)	–380.23 (2632.00)	–4918.24 (2918.45)
9-11 months	–1093.98 (1028.51)	232.64 (2295.04)	–5103.07 (5069.19)
(iii) Survey setting			
<i>Mode of interview (Ref: CAPI)</i>			
CATI	108.23 (596.98)	–186.16 (532.93)	74.67 (1043.64)
<i>Type of interview (Ref: Personal)</i>			
Proxy	253.76 (596.98)	–19.23 (670.13)	–44.80 (497.64)
<i>Interviewer</i>			
Same interviewer	–365.98 (363.92)	–106.27 (516.43)	–107.55 (929.09)
<i>Month of interview (Ref: March to May)</i>			
June to Aug.	330.08 (318.84)	– 353.17 (97.35)	–91.83 (204.94)
Sept. to Nov.	580.43 (469.05)	–632.85 (412.93)	–122.50 (694.32)
(iv) Learning effect			
<i>Wave of interview (Ref: 1st)</i>			
2 nd	176.62 (607.60)	169.67 (480.78)	–295.35 (928.96)
3 rd	146.15 (688.70)	219.37 (514.02)	–287.21 (912.64)
4 th	125.91 (943.23)	66.61 (501.63)	–421.60 (966.44)
Other controls			
Individual fixed effects	yes	yes	yes
Num. obs.	20372	10105	2470
R ² (full model)	0.70	0.72	0.81
Adj. R ² (full model)	0.32	0.41	0.30

Notes: This table shows the results of unbalanced panel regressions with positive and negative reporting errors as dependent variable (i.e. negative values correspond to under-reporting and positive values to over-reporting). Cluster robust standard errors are given in parentheses. As identification in fixed effect models relies on sufficiently large with-in variation, the variables gender, education and country of birth have been removed from the baseline specification. Source: SILC 2008-2011, own calculations.

Table B.3: Full panel regression results – Learning effect

	Wages	Pensions	Unempl. Benefits
(i) Social desirability			
<i>Relative income (Ref: 5. Decile)</i>			
1. Decile	1969.83 (860.87)	759.50 (542.99)	–763.05 (637.59)
2. Decile	1598.28 (1027.30)	–130.17 (604.22)	– 483.91 (228.98)
3. Decile	1252.88 (428.70)	–463.58 (606.05)	–632.23 (577.18)
4. Decile	836.38 (371.02)	–627.86 (348.83)	222.77 (181.61)
6. Decile	–256.38 (293.93)	192.10 (266.99)	–24.88 (393.59)
7. Decile	–166.95 (492.04)	352.91 (586.66)	–20.74 (385.66)
8. Decile	648.48 (733.64)	434.66 (762.01)	378.98 (453.58)
9. Decile	1104.73 (1273.15)	727.65 (1090.20)	72.25 (532.29)
10. Decile	3364.05 (1487.87)	4057.68 (1392.37)	484.78 (533.07)
(ii) Sociodemographic characteristics			
Age	554.52 (656.51)	1289.72 (1233.35)	390.57 (495.37)
Age ²	–5.74 (7.11)	–5.53 (9.28)	–4.78 (6.98)
<i>Health status (Ref: Very bad)</i>			
Bad	643.14 (1223.03)	280.25 (468.71)	80.49 (578.47)
Fair	905.05 (1016.26)	42.28 (412.83)	–249.82 (313.03)
Good	696.00 (954.01)	463.50 (591.92)	127.48 (522.65)
Very good	1022.07 (924.85)	38.76 (423.92)	23.68 (656.28)
<i>Degree of urbanisation (Ref: <10 000 inhabitants)</i>			
>10 000 inhabitants	–2520.54 (2787.27)	2953.45 (1362.68)	–1308.86 (1399.41)
>100 000 inhabitants	–7006.61 (10352.24)	597.81 (4176.38)	–531.65 (1211.78)
<i>Changes in employment status (Ref: None)</i>			
Once	44.53 (982.68)	2081.64 (635.37)	–179.53 (560.39)
Twice	–503.86 (727.33)	552.73 (937.39)	–189.54 (624.62)
Thrice or more	–498.78 (1071.04)	–40.26 (2197.32)	–62.50 (815.18)
<i>Months in corresponding employment status (Ref: 12 months)</i>			
<6 months	–915.70 (916.26)	1167.80 (1745.76)	–908.60 (3480.81)
6-8 months	285.80 (972.59)	–2333.25 (3281.31)	–1533.19 (4068.05)
9-11 months	577.78 (897.73)	–3116.79 (2101.97)	1172.29 (4244.70)
(iii) Survey setting			
<i>Mode of interview (Ref: CAPI)</i>			
CATI	–204.82 (597.24)	288.37 (259.12)	–250.47 (290.43)
<i>Type of interview (Ref: Personal)</i>			
Proxy	390.59 (533.37)	547.18 (370.68)	10.83 (208.57)
<i>Interviewer</i>			
Same interviewer	–37.90 (493.80)	301.58 (420.77)	–336.48 (432.79)
<i>Month of interview (Ref: March to May)</i>			
June to Aug.	401.76 (305.70)	412.98 (102.85)	–50.16 (180.07)
Sept. to Nov.	617.17 (549.12)	901.50 (320.56)	42.88 (573.35)
(iv) Learning effect			
<i>Wave of interview (Ref: 1st)</i>			
2 nd	–164.75 (631.32)	– 1622.45 (264.10)	191.71 (300.39)
3 rd	–353.89 (770.74)	– 2546.06 (520.99)	231.58 (433.76)
4 th	–478.33 (959.25)	– 3397.04 (821.01)	144.55 (568.08)
Other controls			
Individual fixed effects	yes	yes	yes
Num. obs.	20372	10105	2470
R ² (full model)	0.74	0.71	0.82
Adj. R ² (full model)	0.40	0.39	0.31

Notes: This table shows the results of unbalanced panel regressions with absolute values of the reporting errors as dependent variable (i.e. negative values (under-reporting) have been multiplied by –1). Cluster robust standard errors are given in parentheses. As identification in fixed effect models relies on sufficiently large with-in variation, the variables gender, education and country of birth have been removed from the baseline specification. Source: SILC 2008-2011, own calculations.

Table B.4: Full panel regression results – Social desirability – Balanced panel

	Wages	Pensions	Unempl. Benefits
(i) Social desirability			
<i>Relative income (Ref: 5. Decile)</i>			
1. Decile	6971.50 (1313.50)	2304.94 (1218.85)	–570.89 (904.81)
2. Decile	5946.43 (1280.64)	1167.47 (970.35)	215.02 (617.07)
3. Decile	4242.16 (1298.96)	369.92 (1186.05)	–343.08 (890.27)
4. Decile	2481.70 (679.79)	–420.96 (1195.65)	–495.25 (1066.41)
6. Decile	–916.78 (751.95)	–1174.41 (1044.57)	–898.94 (812.39)
7. Decile	– 2443.69 (573.49)	–2114.36 (1864.04)	–416.95 (1016.83)
8. Decile	– 2713.74 (886.54)	–2222.16 (2774.73)	–536.64 (698.26)
9. Decile	– 6633.62 (1563.00)	–4638.03 (2658.24)	– 2293.42 (629.81)
10. Decile	– 6935.37 (2286.25)	– 10015.24 (3544.56)	– 4123.79 (1087.79)
(ii) Sociodemographic characteristics			
Age	–91.24 (555.65)	85.93 (1075.20)	140.26 (795.41)
Age ²	–3.98 (7.74)	3.88 (7.35)	–1.08 (11.55)
<i>Health status (Ref: Very bad)</i>			
Bad	3191.78 (2637.40)	90.69 (842.57)	–502.13 (958.26)
Fair	2157.16 (2684.82)	95.38 (869.85)	–1311.19 (682.15)
Good	2416.33 (2783.00)	–184.89 (780.95)	–424.30 (1277.91)
Very good	3441.87 (2693.37)	1121.60 (845.30)	–200.50 (804.32)
<i>Degree of urbanisation (Ref: <10 000 inhabitants)</i>			
>10 000 & <100 000	1057.53 (1214.26)		–320.37 (554.14)
>100 000 inhabitants	2889.36 (3430.07)	–1339.05 (1800.97)	– 4378.07 (1448.65)
<i>Changes in employment status (Ref: None)</i>			
Once	224.92 (2393.76)	–2869.33 (1539.73)	– 1417.25 (455.06)
Twice	–291.20 (1678.45)	–833.09 (2000.12)	– 1309.01 (571.35)
Thrice or more	–28.50 (1660.38)	–243.54 (3708.31)	– 1137.83 (427.13)
<i>Months in corresponding employment status (Ref: 12 months)</i>			
<6 months	– 4949.70 (1520.14)	–71.61 (647.63)	
6-8 months	–2229.81 (2117.91)		3987.32 (1452.19)
9-11 months	–2087.31 (1973.44)	1045.42 (1893.75)	– 8092.31 (1641.11)
(iii) Survey setting			
<i>Mode of interview (Ref: CAPI)</i>			
CATI	1272.76 (819.10)	175.68 (1588.95)	166.76 (1042.71)
<i>Type of interview (Ref: Personal)</i>			
Proxy	–166.28 (431.34)	–404.47 (539.96)	–405.06 (527.97)
<i>Interviewer</i>			
Same interviewer	680.92 (442.31)	72.43 (1678.76)	–323.48 (1092.34)
<i>Month of interview (Ref: March to May)</i>			
June to Aug.	418.46 (319.76)	–212.91 (316.33)	–48.63 (330.03)
Sept. to Nov.	1577.45 (1524.67)	–602.96 (691.62)	–107.91 (627.76)
(iv) Learning effect			
<i>Wave of interview (Ref: 1st)</i>			
2 nd	–1617.49 (923.46)	–687.35 (1308.66)	–332.03 (1014.64)
3 rd	–690.54 (766.69)	–1651.54 (1145.86)	–643.69 (1268.31)
4 th	202.75 (930.00)	–1548.11 (1073.00)	–553.60 (1425.47)
Other controls			
Individual fixed effects	yes	yes	yes
Num. obs.	3736	1861	437
R ² (full model)	0.53	0.66	0.67
Adj. R ² (full model)	0.31	0.50	0.19

Notes: This table shows the results of balanced panel regressions with positive and negative reporting errors as dependent variable (i.e. negative values correspond to under-reporting and positive values to over-reporting). Cluster robust standard errors are given in parentheses. As identification in fixed effect models relies on sufficiently large with-in variation, the variables gender, education and country of birth have been removed from the baseline specification. Source: SILC 2008-2011, own calculations.

Table B.5: Full panel regression results – Learning effect – Balanced panel

	Wages	Pensions	Unempl. Benefits
(i) Social desirability			
<i>Relative income (Ref: 5. Decile)</i>			
1. Decile	1868.99 (1021.66)	-1136.84 (1118.39)	-808.56 (566.24)
2. Decile	2189.35 (1217.60)	-1167.98 (832.27)	-718.11 (595.36)
3. Decile	1384.01 (802.00)	-1512.10 (755.57)	-390.53 (390.55)
4. Decile	1287.84 (811.26)	-1248.46 (904.36)	-711.36 (678.22)
6. Decile	151.92 (439.57)	-1076.42 (1171.20)	-87.61 (390.48)
7. Decile	234.32 (673.32)	-1229.92 (1215.15)	-246.06 (315.67)
8. Decile	1692.38 (1178.99)	-1392.66 (1353.55)	365.76 (582.68)
9. Decile	1280.60 (2001.19)	-1310.33 (2856.39)	585.48 (712.85)
10. Decile	2619.83 (2789.93)	1612.56 (3873.60)	497.41 (985.76)
(ii) Sociodemographic characteristics			
Age	394.94 (602.54)	492.00 (609.37)	1035.11 (604.77)
Age ²	-6.16 (8.29)	-1.22 (3.28)	-13.89 (8.80)
<i>Health status (Ref: Very bad)</i>			
Bad	653.47 (803.95)	71.74 (500.55)	-1014.31 (1455.56)
Fair	366.75 (823.02)	234.88 (538.00)	-1222.66 (1006.02)
Good	-150.33 (880.81)	443.86 (509.18)	-850.89 (1238.92)
Very good	886.01 (878.31)	-214.52 (500.48)	-1254.88 (1333.87)
<i>Degree of urbanisation (Ref: <10 000 inhabitants)</i>			
>10 000 & <100 000	-1490.76 (1374.83)		-1120.65 (759.05)
>100 000 inhabitants	451.32 (2104.55)	-1721.60 (760.44)	2829.69 (779.45)
<i>Changes in employment status (Ref: None)</i>			
Once	1132.60 (2370.60)	3370.75 (1141.89)	308.38 (658.66)
Twice	402.27 (1648.84)	772.80 (1902.82)	288.78 (452.71)
Thrice or more	760.89 (2176.22)	3082.17 (514.56)	269.06 (617.99)
<i>Months in corresponding employment status (Ref: 12 months)</i>			
<6 months	-1999.35 (1308.20)	-476.34 (330.28)	
6-8 months	-674.07 (2581.14)		1524.54 (1184.56)
9-11 months	-365.75 (2376.15)	-5023.99 (1860.77)	3529.37 (965.18)
(iii) Survey setting			
<i>Mode of interview (Ref: CAPI)</i>			
CATI	1606.78 (1096.31)	-1232.58 (601.10)	951.70 (457.75)
<i>Type of interview (Ref: Personal)</i>			
Proxy	51.11 (467.45)	464.39 (367.94)	-11.73 (499.10)
<i>Interviewer</i>			
Same interviewer	1316.17 (671.17)	-1360.22 (999.51)	691.92 (760.42)
<i>Month of interview (Ref: March to May)</i>			
June to Aug.	428.54 (277.48)	48.17 (223.24)	-142.38 (273.23)
Sept. to Nov.	1201.99 (1440.25)	145.20 (702.48)	-321.45 (289.05)
(iv) Learning effect			
<i>Wave of interview (Ref: 1st)</i>			
2 nd	-2300.95 (1364.41)	743.14 (644.20)	-978.94 (603.76)
3 rd	-1986.53 (1262.53)	781.45 (995.95)	-1283.02 (812.28)
4 th	-2038.26 (1192.46)	-1723.86 (1269.63)	-1096.72 (583.18)
Other controls			
Individual fixed effects	yes	yes	yes
Num. obs.	3736	1861	437
R ² (full model)	0.49	0.68	0.70
Adj. R ² (full model)	0.25	0.52	0.27

Notes: This table shows the results of balanced panel regressions with absolute values of the reporting errors as dependent variable (i.e. negative values (under-reporting) have been multiplied by -1). Cluster robust standard errors are given in parentheses. As identification in fixed effect models relies on sufficiently large with-in variation, the variables gender, education and country of birth have been removed from the baseline specification. Source: SILC 2008-2011, own calculations.

Table B.6: Full panel regression results – Social desirability – Log differences

	Wages	Pensions	Unempl. Benefits
(i) Social desirability			
<i>Relative income (Ref: 5. Decile)</i>			
1. Decile	1.51 (0.10)	0.54 (0.16)	1.43 (0.46)
2. Decile	0.74 (0.09)	0.19 (0.06)	0.70 (0.22)
3. Decile	0.34 (0.04)	0.11 (0.04)	0.17 (0.28)
4. Decile	0.13 (0.03)	–0.00 (0.04)	0.12 (0.14)
6. Decile	– 0.11 (0.04)	–0.02 (0.02)	–0.21 (0.25)
7. Decile	– 0.21 (0.04)	–0.09 (0.05)	–0.33 (0.29)
8. Decile	– 0.30 (0.03)	–0.14 (0.07)	–0.40 (0.25)
9. Decile	– 0.44 (0.04)	– 0.20 (0.07)	–0.34 (0.21)
10. Decile	– 0.54 (0.05)	– 0.29 (0.09)	– 0.53 (0.25)
(ii) Sociodemographic characteristics			
Age	0.12 (0.06)	–0.09 (0.12)	0.04 (0.21)
Age ²	– 0.00 (0.00)	0.00 (0.00)	–0.00 (0.00)
<i>Health status (Ref: Very bad)</i>			
Bad	0.05 (0.10)	–0.01 (0.03)	0.03 (0.34)
Fair	0.02 (0.11)	–0.01 (0.03)	0.14 (0.33)
Good	0.02 (0.10)	–0.01 (0.03)	0.03 (0.33)
Very good	0.04 (0.10)	–0.00 (0.04)	0.14 (0.34)
<i>Degree of urbanisation (Ref: <10 000 inhabitants)</i>			
>10 000 & <100 000	0.01 (0.11)	0.32 (0.18)	–0.16 (0.40)
>100 000 inhabitants	0.03 (0.11)	0.06 (0.13)	–0.07 (0.45)
<i>Changes in employment status (Ref: None)</i>			
Once	– 0.15 (0.04)	– 0.25 (0.08)	–0.23 (0.17)
Twice	– 0.17 (0.05)	–0.03 (0.23)	–0.24 (0.17)
Thrice or more	–0.03 (0.09)	0.01 (0.26)	–0.14 (0.18)
<i>Months in corresponding employment status (Ref: 12 months)</i>			
<6 months	– 0.55 (0.12)	–0.03 (0.07)	– 0.72 (0.26)
6-8 months	–0.12 (0.08)	–0.08 (0.16)	–0.56 (0.53)
9-11 months	–0.00 (0.07)	–0.08 (0.16)	–1.01 (1.23)
(iii) Survey setting			
<i>Mode of interview (Ref: CAPI)</i>			
CATI	0.02 (0.02)	–0.02 (0.03)	0.02 (0.27)
<i>Type of interview (Ref: Personal)</i>			
Proxy	–0.00 (0.02)	–0.00 (0.04)	–0.06 (0.14)
<i>Interviewer</i>			
Same interviewer	–0.01 (0.02)	–0.01 (0.04)	0.07 (0.24)
<i>Month of interview (Ref: March to May)</i>			
June to Aug.	0.01 (0.01)	–0.01 (0.01)	–0.01 (0.06)
Sept. to Nov.	0.03 (0.03)	–0.04 (0.02)	0.00 (0.19)
(iv) Learning effect			
<i>Wave of interview (Ref: 1st)</i>			
2 nd	–0.02 (0.04)	0.01 (0.04)	–0.15 (0.22)
3 rd	–0.03 (0.09)	0.01 (0.04)	–0.21 (0.19)
4 th	–0.06 (0.13)	0.00 (0.06)	–0.27 (0.23)
Other controls			
Individual fixed effects	yes	yes	yes
Num. obs.	20372	10105	2470
R ² (full model)	0.82	0.88	0.84
Adj. R ² (full model)	0.59	0.75	0.38

Notes: This table shows the results of unbalanced panel regressions with positive and negative logarithmic differences as dependent variable (i.e. negative values correspond to under-reporting and positive values to over-reporting). Cluster robust standard errors are given in parentheses. As identification in fixed effect models relies on sufficiently large with-in variation, the variables gender, education and country of birth have been removed from the baseline specification. Source: SILC 2008-2011, own calculations.

Table B.7: Full panel regression results – Learning effect – Log differences

	Wages	Pensions	Unempl. Benefits
(i) Social desirability			
<i>Relative income (Ref: 5. Decile)</i>			
1. Decile	1.03 (0.13)	0.38 (0.13)	0.59 (0.47)
2. Decile	0.54 (0.05)	0.12 (0.07)	0.16 (0.14)
3. Decile	0.24 (0.03)	0.03 (0.05)	–0.09 (0.11)
4. Decile	0.12 (0.03)	–0.00 (0.03)	0.08 (0.14)
6. Decile	– 0.06 (0.02)	0.01 (0.02)	–0.09 (0.15)
7. Decile	– 0.08 (0.03)	0.01 (0.04)	–0.17 (0.12)
8. Decile	–0.07 (0.04)	–0.01 (0.05)	–0.09 (0.12)
9. Decile	–0.07 (0.04)	–0.05 (0.05)	– 0.22 (0.09)
10. Decile	–0.06 (0.06)	0.04 (0.04)	– 0.24 (0.09)
(ii) Sociodemographic characteristics			
Age	0.08 (0.05)	0.04 (0.09)	0.06 (0.16)
Age ²	–0.00 (0.00)	–0.00 (0.00)	–0.00 (0.00)
<i>Health status (Ref: Very bad)</i>			
Bad	–0.00 (0.11)	0.02 (0.02)	0.13 (0.14)
Fair	0.00 (0.09)	0.01 (0.03)	0.06 (0.20)
Good	–0.02 (0.08)	0.02 (0.03)	0.15 (0.19)
Very good	–0.01 (0.08)	0.00 (0.02)	0.11 (0.20)
<i>Degree of urbanisation (Ref: <10 000 inhabitants)</i>			
>10 000 & <100 000	–0.03 (0.06)	0.28 (0.20)	–0.26 (0.18)
>100 000 inhabitants	–0.05 (0.08)	0.02 (0.13)	–0.07 (0.45)
<i>Changes in employment status (Ref: None)</i>			
Once	–0.02 (0.04)	0.09 (0.03)	0.02 (0.15)
Twice	–0.03 (0.04)	0.09 (0.19)	0.03 (0.18)
Thrice or more	–0.01 (0.11)	–0.01 (0.12)	0.00 (0.20)
<i>Months in corresponding employment status (Ref: 12 months)</i>			
<6 months	– 0.12 (0.04)	0.03 (0.04)	–0.32 (0.44)
6-8 months	0.01 (0.05)	–0.32 (0.26)	–0.56 (0.53)
9-11 months	0.04 (0.04)	– 0.40 (0.15)	0.33 (0.81)
(iii) Survey setting			
<i>Mode of interview (Ref: CAPI)</i>			
CATI	–0.01 (0.02)	0.02 (0.02)	0.07 (0.10)
<i>Type of interview (Ref: Personal)</i>			
Proxy	–0.00 (0.02)	0.02 (0.02)	0.10 (0.04)
<i>Interviewer</i>			
Same interviewer	–0.00 (0.02)	0.00 (0.04)	0.06 (0.10)
<i>Month of interview (Ref: March to May)</i>			
June to Aug.	0.02 (0.02)	0.04 (0.01)	0.00 (0.04)
Sept. to Nov.	0.01 (0.03)	0.07 (0.03)	0.02 (0.13)
(iv) Learning effect			
<i>Wave of interview (Ref: 1st)</i>			
2 nd	–0.07 (0.04)	– 0.08 (0.03)	–0.09 (0.08)
3 rd	–0.12 (0.08)	– 0.15 (0.04)	–0.13 (0.17)
4 th	–0.17 (0.12)	– 0.19 (0.05)	–0.20 (0.21)
Other controls			
Individual fixed effects	yes	yes	yes
Num. obs.	20372	10105	2470
R ² (full model)	0.81	0.88	0.81
Adj. R ² (full model)	0.56	0.74	0.29

Notes: This table shows the results of unbalanced panel regressions with absolute logarithmic differences as dependent variable (i.e. negative values (under-reporting) have been multiplied by –1). Cluster robust standard errors are given in parentheses. As identification in fixed effect models relies on sufficiently large with-in variation, the variables gender, education and country of birth have been removed from the baseline specification. Source: SILC 2008-2011, own calculations.

Table B.8: Decomposition of explained variance – SILC 2008-2011

	Wages	Pensions	Unempl. Benefits
Proportion of variance explained	5.9	13.3	49.2
<i>Relative importance:</i>			
(i) Social desirability	69.5	75.0	22.5
(ii) Sociodemographic characteristics	28.9	21.8	67.7
(iii) Survey setting	0.8	2.9	6.2
(iv) Learning effect	0.7	0.3	3.6

Notes: This table shows the goodness-of-fit of OLS regressions and its decomposition to four error sources, i.e. four groups of explanatory variables. We quantify the relative importance of (i) social desirability, (ii) socio-demographic characteristics, (iii) aspects of the survey design, and (iv) learning effects on the basis of separate regressions for wages, pensions and unemployment benefits. Absolute Reporting errors are regressed on the same set of explanatory variables that were used before (see section 3), using all available pooled cross-sections. Source: SILC 2008-2011, own calculations.

Table B.9: Changes of Income Deciles (Income Mobility)

Deciles changed	Wages		Pension		Unemployment	
	%	Obs	%	Obs	%	Obs
0	62.1	5478	74.8	3492	23.2	148
1	27.2	2402	18.2	849	31.1	199
2	6.0	528	2.8	132	18.5	118
3	2.3	207	1.6	74	11.1	71
4	1.2	104	1.0	46	7.2	46
5	0.6	57	1.0	45	3.8	24
6	0.4	32	0.2	9	3.1	20
7	0.1	11	0.2	8	1.4	9
8	0.1	7	0.2	10	0.6	4
9	0.0	0	0.0	1	0.0	0

B.1 Misreporting and attenuation bias in linear models

Quantifying the attenuation bias due to misreporting gives data users and practitioners some guidance on the expected bias when they use income data from SILC in their own regression models (or see it being used in existing publications). Second, it allows for a comparison of the magnitude of attenuation bias in the Austrian SILC with similar quantifications for other countries and in particular with Bingley and Martinello (2017) (table 1 & table 2, column 1).

To show how the error model translates into attenuation bias, we have replicated table 2, col. 1 of Bingley and Martinello (2017) for our data and for all income variables. Thus, we estimated ρ (i.e. the coefficient indicating mean reversion if negative and significant) under the assumption that the administrative data are error-free. Based on that we calculated various measures for attenuation bias resulting from the error model.

See equation 1 in Bingley and Martinello (2017): "We are interested in estimating the relationship between a dependent variable y and an explanatory variable x . However, we observe not one or both variables in the model, but a measure m_s "

$$m_s = \kappa_s + \rho(m + \mu_m) + \varepsilon_s$$

where $m \in \{y, x\}$ and the subscript $_s$ indicates that we are validating survey information; thus m_s is the measure observed in the survey data. According to this formulation, measurement error consists of three components: κ is a constant representing a nonzero mean error; ε_s is in *iid* distributed error term with mean zero and constant variance σ_s^2 and ρ_s represents the dependence between measurement error and the quantity of interest. Ultimately, for a quantification of bias in different settings, we are interesting in estimating ρ_s . Under the assumption of classical measurement error in the survey data, $\rho_s = \kappa_s = 0$. In this case measurement error $m_s - m$ has mean zero and is independent of the quantity of interests.

We base our assessment of the size of the bias on a regression of the measurement error (defined as $m_s - m$, where m_s is the log of income according to the survey and m is the log if income according to the registers) on m . The residual variance of that model is σ_s , the variance of m is given by σ_x and ρ_s is the slope coefficient. Using this notation, the estimate bias in OLS estimators due to measurement error in the explanatory variables is given by

$$\text{bias OLS} = \frac{(1 + \rho_s)\sigma_x^2}{(1 + \rho_s)^2\sigma_s^2}$$

For IV models, the bias is given by

$$\frac{1}{1 + \rho_s}$$

and the bias resulting from measurement error on the left hand side variable is given by

$$1 + \rho_s$$

As the following table shows, we find a negative correlation (mean reversion), significantly different from zero, between measurement error in all 3 income types and its register value. The effect is stronger than in Bingley and Martinello (2017) and in the classical studies of Bound and Krueger (1991) and Bound, Brown, Duncan, et al. (1994). However, direct comparisons have to be treated with caution as the exact definition of ‘income’ varies between studies.

This apparent nonclassicality e.g. suggests that all results estimating returns (e.g. of education) on wages are biased (LHS bias) and are underestimated by about 25%, whereas Bingley and Martinello (2017) observe a bias of only 10%. If error-prone pension income is used as dependent variable this bias increases to 35%

Table B.10: Replication of Bingley/Martinello (2017), table 2 col. 1 based on SILC

	Bingley (2017)		SILC	
	1	2	3	4
Log(Income)	Gross total income	Wages	Pensions	Unemp. Benefits
ρ_s	-.103**	-.256***	-0.355***	-0.283***
	(.0375)	(.0082)	(0.0201)	(0.0201)
OLS bias	.479	.905	.931	.741
IV bias	1.115	1.344	1.550	1.394
LHS bias	.897	.744	.645	.717
N	1649	23372	10850	2794

** p<.05; *** p<0.001 Robust standard errors are in parenthesis. OLS estimates. Columns (2) to (4) based on pooled cross-sections 2008-2011. (1): To account for classification conventions in the Danish Tax system and to increase sample size, Bingley and Martinello (2017) use gross income defined as the sum of both wages and pensions.

C Robustness checks

Table C.1: Results: Linear Probability Model for Mismatch Groups

	<i>Dependent variable:</i>								
	Wages			Unemp.			Pensions		
	<i>S < A</i>	<i>S = A</i>	<i>S > A</i>	<i>S < A</i>	<i>S = A</i>	<i>S > A</i>	<i>S < A</i>	<i>S = A</i>	<i>S > A</i>
(i) Social desirability									
<i>Relative income (Ref: 5. Decile)</i>									
1. Decile	-0.38*** (0.02)	-0.17*** (0.02)	0.55*** (0.02)	-0.28*** (0.05)	-0.04 (0.04)	0.32*** (0.06)	-0.18*** (0.02)	-0.13*** (0.02)	0.31*** (0.02)
2. Decile	-0.19*** (0.02)	-0.15*** (0.02)	0.34*** (0.02)	-0.14*** (0.05)	-0.02 (0.03)	0.16*** (0.05)	-0.05** (0.02)	-0.05* (0.02)	0.09*** (0.02)
3. Decile	-0.11*** (0.02)	-0.12*** (0.01)	0.23*** (0.02)	-0.01 (0.05)	-0.06** (0.02)	0.07 (0.05)	-0.04** (0.02)	-0.01 (0.02)	0.05*** (0.02)
4. Decile	-0.02 (0.02)	-0.08*** (0.02)	0.10*** (0.02)	-0.01 (0.05)	-0.06** (0.03)	0.06 (0.05)	0.01 (0.02)	-0.04* (0.02)	0.03* (0.02)
6. Decile	0.09*** (0.02)	-0.01 (0.02)	-0.08*** (0.02)	0.12** (0.05)	0.01 (0.03)	-0.12** (0.05)	0.09*** (0.02)	-0.07*** (0.02)	-0.02 (0.02)
7. Decile	0.15*** (0.02)	-0.01 (0.02)	-0.14*** (0.02)	0.09* (0.05)	-0.01 (0.03)	-0.08* (0.05)	0.09*** (0.02)	-0.08*** (0.02)	-0.01 (0.02)
8. Decile	0.19*** (0.02)	-0.02 (0.02)	-0.17*** (0.02)	0.17*** (0.05)	0.00 (0.03)	-0.17*** (0.05)	0.10*** (0.02)	-0.08*** (0.02)	-0.02 (0.02)
9. Decile	0.28*** (0.02)	-0.06*** (0.02)	-0.22*** (0.02)	0.21*** (0.04)	0.09** (0.04)	-0.30*** (0.05)	0.13*** (0.03)	-0.08*** (0.02)	-0.06*** (0.02)
10. Decile	0.42*** (0.02)	-0.13*** (0.02)	-0.28*** (0.02)	0.32*** (0.05)	0.10** (0.04)	-0.42*** (0.05)	0.25*** (0.02)	-0.19*** (0.02)	-0.06*** (0.02)
(ii) Sociodemographic characteristics									
<i>Gender (Ref: Female)</i>									
Male	-0.11*** (0.01)	-0.04*** (0.01)	0.15*** (0.01)	-0.06*** (0.02)	-0.02 (0.02)	0.08*** (0.02)	-0.06*** (0.01)	0.00 (0.01)	0.06*** (0.01)
<i>Age</i>									
Age	-0.01*** (0.00)	0.00*** (0.00)	0.00 (0.00)	-0.00 (0.01)	0.00 (0.01)	-0.00 (0.01)	-0.02 (0.01)	0.08*** (0.01)	-0.06*** (0.01)
Age ²	0.00*** (0.00)	-0.00** (0.00)	-0.00 (0.00)	-0.00 (0.00)	0.00 (0.00)	-0.00 (0.00)	0.00 (0.00)	-0.00*** (0.00)	0.00*** (0.00)
<i>Education (Ref: Compulsory)</i>									
Upper secondary	-0.08*** (0.01)	-0.00 (0.01)	0.09*** (0.01)	0.04 (0.03)	-0.00 (0.02)	-0.04 (0.03)	-0.05*** (0.01)	0.02* (0.01)	0.03*** (0.01)
Post-secondary	-0.09*** (0.02)	-0.01 (0.02)	0.10*** (0.02)	-0.05 (0.05)	-0.00 (0.04)	0.06 (0.06)	-0.07*** (0.02)	0.02 (0.02)	0.04** (0.02)
1 st stage tertiary	-0.11*** (0.02)	0.00 (0.02)	0.11*** (0.02)	0.03 (0.05)	-0.06** (0.03)	0.03 (0.05)	-0.05 (0.03)	-0.00 (0.03)	0.06** (0.03)
2 nd stage tertiary	-0.04 (0.03)	-0.05* (0.03)	0.09*** (0.03)	-0.08 (0.15)	0.07 (0.14)	0.00 (0.16)	0.01 (0.05)	0.01 (0.04)	-0.02 (0.04)
<i>Country of birth (Ref: AUT)</i>									
EU15	-0.04 (0.02)	0.00 (0.02)	0.04 (0.03)	-0.01 (0.07)	0.00 (0.05)	0.00 (0.05)	-0.02 (0.03)	-0.09*** (0.03)	0.12*** (0.03)
CEE	-0.03 (0.03)	0.00 (0.03)	0.02 (0.03)	0.11 (0.07)	-0.08* (0.04)	-0.03 (0.07)	0.00 (0.03)	-0.01 (0.04)	0.01 (0.03)
Yugosphere	0.05** (0.02)	-0.05** (0.02)	0.00 (0.02)	0.06* (0.03)	-0.06** (0.03)	-0.01 (0.03)	0.05 (0.04)	-0.07 (0.04)	0.02 (0.04)
Turkey	0.07** (0.03)	-0.01 (0.03)	-0.06** (0.03)	-0.04 (0.04)	0.00 (0.04)	0.04 (0.04)	0.02 (0.06)	0.01 (0.10)	-0.03 (0.09)
Other	0.07*** (0.02)	-0.05** (0.02)	-0.03 (0.03)	-0.02 (0.04)	-0.01 (0.04)	0.03 (0.05)	-0.02 (0.07)	0.03 (0.08)	-0.01 (0.08)
<i>Health status (Ref: Very bad)</i>									
Bad	-0.05 (0.07)	0.02 (0.05)	0.03 (0.05)	0.09 (0.07)	0.01 (0.06)	-0.10 (0.08)	-0.02 (0.02)	0.05* (0.03)	-0.03 (0.02)
Fair	-0.04 (0.07)	-0.01 (0.05)	0.05 (0.05)	0.02 (0.07)	0.06 (0.06)	-0.08 (0.08)	-0.02 (0.02)	0.07*** (0.02)	-0.05** (0.02)
Good	-0.08 (0.07)	0.01 (0.05)	0.07 (0.05)	0.06 (0.06)	0.03 (0.05)	-0.09 (0.07)	-0.04* (0.02)	0.09*** (0.02)	-0.06** (0.02)
Very Good	-0.13* (0.07)	0.01 (0.05)	0.12** (0.05)	0.04 (0.06)	0.06 (0.06)	-0.10 (0.08)	-0.08*** (0.02)	0.10*** (0.03)	-0.02 (0.03)
<i>Degree of urbanisation (Ref: <10 000 inhabitants)</i>									
>10 000 & <100 000	-0.02 (0.01)	0.02* (0.01)	-0.00 (0.01)	-0.04 (0.03)	0.01 (0.02)	0.03 (0.03)	-0.04** (0.02)	0.04** (0.02)	0.00 (0.01)
>100 000	-0.05*** (0.01)	0.01 (0.01)	0.03*** (0.01)	-0.09*** (0.02)	0.03 (0.02)	0.05** (0.02)	-0.05*** (0.01)	0.03** (0.01)	0.02 (0.01)

Table continues on next page

	<i>Dependent variable:</i>								
	Wages			Unemp.			Pensions		
	<i>S < A</i>	<i>S = A</i>	<i>S > A</i>	<i>S < A</i>	<i>S = A</i>	<i>S > A</i>	<i>S < A</i>	<i>S = A</i>	<i>S > A</i>
Table C.1 cont.									
<i>Changes in employment status (Ref: None)</i>									
Once	0.04** (0.02)	-0.03** (0.02)	-0.01 (0.02)	0.10** (0.04)	-0.04 (0.03)	-0.05 (0.04)	0.30*** (0.05)	-0.22*** (0.04)	-0.08* (0.05)
Twice	0.05** (0.02)	0.00 (0.02)	-0.05** (0.03)	0.10** (0.05)	-0.05 (0.03)	-0.05 (0.04)	0.25** (0.11)	-0.27*** (0.10)	0.02 (0.07)
Thrice or More	-0.03 (0.03)	-0.05* (0.03)	0.07** (0.04)	0.14** (0.06)	-0.07* (0.04)	-0.07 (0.06)	0.22 (0.14)	-0.10 (0.15)	-0.12 (0.15)
<i>Months in corresponding employment status (Ref: 12 moths)</i>									
<6	0.29*** (0.02)	-0.03* (0.02)	-0.26*** (0.02)	0.39*** (0.05)	0.03 (0.04)	-0.42*** (0.05)	0.03 (0.02)	-0.06** (0.03)	0.03 (0.02)
6-8	0.11*** (0.03)	-0.05* (0.02)	-0.06** (0.03)	0.14** (0.06)	0.02 (0.04)	-0.16*** (0.06)	-0.11 (0.07)	0.02 (0.07)	0.09 (0.07)
9-11	0.03 (0.02)	-0.07*** (0.02)	0.05* (0.03)	0.09 (0.07)	0.02 (0.05)	-0.12* (0.06)	0.05 (0.09)	-0.10 (0.07)	0.05 (0.06)
(iii) Survey setting									
<i>Mode of interview (Ref: CAPI)</i>									
CATI	-0.02 (0.01)	0.03*** (0.01)	-0.01 (0.01)	-0.05 (0.03)	0.06** (0.03)	-0.01 (0.04)	-0.02 (0.02)	0.03* (0.02)	-0.01 (0.01)
<i>Interviewer</i>									
Same interviewer	0.01 (0.02)	0.01 (0.02)	-0.02 (0.02)	-0.08* (0.04)	0.06* (0.03)	0.02 (0.04)	-0.00 (0.02)	0.02 (0.02)	-0.02 (0.02)
<i>Type of interview (Ref: Personal)</i>									
Proxy	0.08*** (0.01)	-0.06*** (0.01)	-0.02** (0.01)	-0.00 (0.03)	-0.01 (0.02)	0.01 (0.03)	0.04** (0.02)	-0.06*** (0.02)	0.02 (0.01)
<i>Month of interview (Ref: March to May)</i>									
June to Aug.	0.02*** (0.01)	-0.03*** (0.01)	0.01 (0.01)	-0.00 (0.02)	-0.00 (0.02)	0.00 (0.03)	0.01 (0.01)	-0.03** (0.01)	0.01 (0.01)
Sept. to Nov.	0.02 (0.02)	-0.06*** (0.02)	0.04* (0.02)	0.01 (0.05)	-0.00 (0.03)	-0.01 (0.05)	0.10*** (0.03)	-0.11*** (0.03)	0.01 (0.02)
<i>Year of interview (Ref: 2008)</i>									
2009	-0.00 (0.01)	-0.01 (0.01)	0.01 (0.01)	-0.02 (0.03)	-0.02 (0.02)	0.04 (0.03)	0.00 (0.01)	-0.03* (0.01)	0.02* (0.01)
2010	0.01 (0.01)	-0.00 (0.01)	-0.01 (0.01)	0.02 (0.03)	0.03 (0.02)	-0.06* (0.03)	0.03* (0.01)	-0.04*** (0.01)	0.01 (0.01)
2011	-0.01 (0.01)	-0.00 (0.01)	0.01 (0.01)	0.01 (0.04)	0.01 (0.02)	-0.02 (0.04)	-0.30*** (0.02)	0.46*** (0.02)	-0.16*** (0.01)
(iv) Learning effect									
<i>Wave of interview (Ref: 1st)</i>									
2 nd	-0.01 (0.02)	0.00 (0.01)	0.01 (0.01)	0.07* (0.04)	-0.05* (0.03)	-0.02 (0.04)	0.03 (0.02)	-0.04* (0.02)	0.00 (0.01)
3 rd	-0.02 (0.02)	0.01 (0.01)	0.01 (0.01)	0.04 (0.04)	-0.05* (0.03)	0.01 (0.04)	0.02 (0.02)	-0.05** (0.02)	0.03* (0.02)
4 th	-0.03 (0.02)	0.02 (0.01)	0.01 (0.02)	0.04 (0.04)	-0.01 (0.02)	-0.03 (0.04)	0.02 (0.02)	-0.03 (0.02)	0.01 (0.02)
Constant	0.68*** (0.09)	0.27*** (0.06)	0.05 (0.07)	0.04 (0.17)	0.04 (0.12)	0.91*** (0.18)	1.00*** (0.35)	-2.28*** (0.41)	2.28*** (0.35)
Observations	20,365	20,365	20,365	2,467	2,467	2,467	10,104	10,104	10,104
Log Likelihood	-14,612.23	-12,427.26	-12,729.54	-1,693.76	-948.46	-1,697.32	-5,781.48	-6,271.54	-4,293.09
Akaike Inf. Crit.	29,314.45	24,944.52	25,549.09	3,477.52	1,986.93	3,484.64	11,652.96	12,633.08	8,676.18

This tables shows the marginal effects from linear probability models estimated for each mismatch category (*S<A*: survey reported income is below the value in the administrative database; *S=A*: survey reported income corresponds to administrative database within a bandwidth of $\pm 5\%$; *S>A*: survey reported income is above the value in the administrative data) per income type (wages, pensions and unemployment benefits). The respective dependent variable corresponding to columns (2) to (9) is a binary indicator, equal to one if the mismatch type is observed for the corresponding income type and zero otherwise. Standard errors based on 500 bootstrap replicates are reported in parenthesis.

Table C.2: Results: Linear Probability Model for Mismatch Groups- Exclusive of Income Deciles

	<i>Dependent variable:</i>								
	Wages			Unemp.			Pensions		
	<i>S < A</i>	<i>S = A</i>	<i>S > A</i>	<i>S < A</i>	<i>S = A</i>	<i>S > A</i>	<i>S < A</i>	<i>S = A</i>	<i>S > A</i>
(ii) Sociodemographic characteristics									
<i>Gender (Ref: Male)</i>									
Male	-0.01 (0.01)	-0.04*** (0.01)	0.05*** (0.01)	-0.02 (0.02)	-0.01 (0.02)	0.03 (0.03)	0.01 (0.01)	-0.01 (0.01)	0.01 (0.01)
<i>Age</i>									
Age	0.01*** (0.00)	0.01*** (0.00)	-0.02*** (0.00)	0.01 (0.01)	0.00 (0.01)	-0.01* (0.01)	-0.01 (0.01)	0.09*** (0.01)	-0.07*** (0.01)
Age ²	-0.00 (0.00)	-0.00*** (0.00)	0.00*** (0.00)	-0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	-0.00*** (0.00)	0.00*** (0.00)
<i>Education (Ref: Compulsory)</i>									
Upper secondary	-0.02 (0.01)	0.01 (0.01)	0.01 (0.01)	0.05* (0.03)	0.00 (0.02)	-0.05* (0.03)	-0.01 (0.01)	0.01 (0.01)	-0.00 (0.01)
Post secondary	0.02 (0.02)	-0.01 (0.02)	-0.01 (0.02)	-0.03 (0.05)	0.01 (0.04)	0.02 (0.06)	-0.01 (0.02)	0.01 (0.02)	0.01 (0.02)
1 st stage tertiary	0.04** (0.02)	-0.00 (0.02)	-0.04** (0.02)	0.08* (0.05)	-0.04 (0.03)	-0.04 (0.04)	0.07** (0.03)	-0.06* (0.03)	-0.01 (0.03)
2 nd stage tertiary	0.19*** (0.03)	-0.09*** (0.03)	-0.11*** (0.03)	-0.06 (0.15)	0.08 (0.15)	-0.03 (0.18)	0.13*** (0.05)	-0.07* (0.04)	-0.07 (0.04)
<i>Country of birth (Ref: AUT)</i>									
EU15	-0.07*** (0.03)	0.01 (0.02)	0.07** (0.03)	-0.00 (0.07)	0.00 (0.05)	0.00 (0.06)	-0.05* (0.03)	-0.08** (0.03)	0.14*** (0.04)
CEE	-0.11*** (0.03)	0.00 (0.03)	0.10*** (0.03)	0.09 (0.07)	-0.08* (0.04)	-0.00 (0.07)	-0.01 (0.04)	-0.01 (0.04)	0.02 (0.02)
Yugosphere	-0.00 (0.02)	-0.04* (0.02)	0.04* (0.02)	0.07* (0.04)	-0.05** (0.03)	-0.01 (0.03)	0.01 (0.04)	-0.06 (0.04)	0.05 (0.04)
Turkey	0.03 (0.04)	0.00 (0.03)	-0.03 (0.03)	-0.05 (0.05)	-0.00 (0.04)	0.05 (0.05)	-0.02 (0.07)	0.01 (0.11)	0.01 (0.11)
Other	-0.01 (0.02)	-0.05** (0.02)	0.06** (0.03)	-0.02 (0.04)	-0.02 (0.04)	0.04 (0.05)	-0.07 (0.08)	0.05 (0.08)	0.02 (0.07)
<i>Health status (Ref: Very bad)</i>									
Bad	-0.01 (0.07)	0.03 (0.05)	-0.01 (0.05)	0.10 (0.07)	0.01 (0.06)	-0.10 (0.08)	-0.03 (0.02)	0.05** (0.03)	-0.03 (0.02)
Fair	0.00 (0.07)	0.00 (0.05)	-0.00 (0.05)	0.01 (0.07)	0.05 (0.06)	-0.06 (0.09)	-0.03* (0.02)	0.08*** (0.02)	-0.05* (0.02)
Good	-0.02 (0.06)	0.01 (0.05)	0.01 (0.05)	0.05 (0.07)	0.02 (0.06)	-0.07 (0.08)	-0.03 (0.02)	0.09*** (0.02)	-0.06*** (0.02)
Very Good	-0.06 (0.06)	0.01 (0.05)	0.05 (0.05)	0.04 (0.07)	0.05 (0.06)	-0.09 (0.09)	-0.07*** (0.02)	0.10*** (0.03)	-0.03 (0.03)
<i>Degree of urbanisation (Ref: <10 000 inhabitants)</i>									
>10 000 & <100 000	-0.00 (0.01)	0.02* (0.01)	-0.02* (0.01)	-0.04 (0.04)	0.02 (0.02)	0.02 (0.03)	-0.02 (0.02)	0.03** (0.02)	-0.01 (0.01)
> 100000	-0.03*** (0.01)	0.01 (0.01)	0.02** (0.01)	-0.08*** (0.02)	0.03* (0.02)	0.04* (0.02)	-0.03* (0.01)	0.03** (0.01)	-0.00 (0.01)
<i>Changes in employment status(Ref: None)</i>									
Once	0.05** (0.02)	-0.03* (0.02)	-0.02 (0.02)	0.09** (0.04)	-0.06* (0.03)	-0.04 (0.04)	0.23*** (0.05)	-0.20*** (0.04)	-0.03 (0.05)
Twice	0.06** (0.03)	0.00 (0.02)	-0.06** (0.03)	0.09* (0.05)	-0.07** (0.03)	-0.02 (0.04)	0.18 (0.12)	-0.26*** (0.09)	0.08 (0.08)
Thrice or More	-0.03 (0.04)	-0.06** (0.03)	0.09** (0.04)	0.17*** (0.06)	-0.09** (0.04)	-0.08 (0.06)	0.18 (0.15)	-0.08 (0.14)	-0.10 (0.13)
<i>Months in corresponding employment status (Ref: 12 months)</i>									
<6	0.01 (0.02)	-0.11*** (0.01)	0.10*** (0.02)	0.24*** (0.04)	-0.02 (0.04)	-0.22*** (0.05)	-0.02 (0.03)	-0.07** (0.03)	0.09*** (0.03)
6-8	-0.05* (0.03)	-0.09*** (0.02)	0.14*** (0.03)	0.09 (0.05)	0.01 (0.04)	-0.10 (0.06)	-0.06 (0.07)	-0.01 (0.07)	0.07 (0.07)
9-11	-0.08*** (0.02)	-0.08*** (0.02)	0.16*** (0.03)	0.08 (0.07)	0.03 (0.05)	-0.11* (0.07)	0.16* (0.09)	-0.15* (0.07)	-0.01 (0.07)

Table continues on next page

	<i>Dependent variable:</i>								
	<i>S < A</i>	Wages <i>S = A</i>	<i>S > A</i>	<i>S < A</i>	Unemp. <i>S = A</i>	<i>S > A</i>	<i>S < A</i>	Pensions <i>S = A</i>	<i>S > A</i>
Table C.2 cont.									
(iii) Survey setting									
<i>Mode of interview (Ref: CAPI)</i>									
CATI	-0.01 (0.02)	0.03** (0.01)	-0.02 (0.01)	-0.04 (0.04)	0.07** (0.03)	-0.02 (0.04)	-0.01 (0.02)	0.03 (0.02)	-0.02 (0.01)
<i>Interviewer</i>									
Same Interviewer	0.00 (0.02)	0.01 (0.02)	-0.02 (0.02)	-0.07* (0.04)	0.06* (0.03)	0.02 (0.04)	0.00 (0.02)	0.02 (0.02)	-0.02 (0.02)
<i>Type of interview (Ref: Personal)</i>									
Proxy	0.07*** (0.01)	-0.06*** (0.01)	-0.02 (0.01)	-0.01 (0.03)	-0.01 (0.02)	0.02 (0.04)	0.03* (0.02)	-0.06*** (0.02)	0.03** (0.01)
<i>Months of interview (Ref: March to May)</i>									
June to Aug.	0.02** (0.01)	-0.03*** (0.01)	0.00 (0.01)	0.00 (0.03)	0.00 (0.02)	-0.00 (0.03)	0.01 (0.01)	-0.02** (0.01)	0.01 (0.01)
Sept. to Nov.	0.02 (0.02)	-0.05*** (0.02)	0.04* (0.02)	0.02 (0.05)	0.00 (0.03)	-0.02 (0.05)	0.09*** (0.03)	-0.11*** (0.03)	0.02 (0.03)
<i>Year of interview (Ref: 2008)</i>									
2009	0.00 (0.01)	-0.01 (0.01)	0.01 (0.01)	-0.02 (0.04)	-0.02 (0.02)	0.05 (0.04)	0.00 (0.02)	-0.02* (0.01)	0.02 (0.01)
2010	0.02 (0.01)	0.00 (0.01)	-0.02 (0.01)	0.02 (0.03)	0.03 (0.02)	-0.05 (0.03)	0.03* (0.01)	-0.04*** (0.01)	0.01 (0.01)
2011	0.01 (0.01)	0.00 (0.01)	-0.01 (0.01)	0.01 (0.04)	0.01 (0.02)	-0.02 (0.04)	-0.30*** (0.02)	0.46*** (0.02)	-0.16*** (0.01)
(iv) Learning effect									
<i>Wave of interview (Ref: 1st)</i>									
2 nd	-0.02 (0.02)	0.00 (0.01)	0.01 (0.02)	0.06 (0.04)	-0.05* (0.03)	-0.01 (0.04)	0.02 (0.02)	-0.03 (0.02)	0.01 (0.01)
3 rd	-0.02 (0.02)	0.01 (0.01)	0.01 (0.01)	0.04 (0.04)	-0.05* (0.03)	0.01 (0.04)	0.01 (0.02)	-0.04** (0.02)	0.03** (0.02)
4 th	-0.03* (0.02)	0.02 (0.01)	0.01 (0.02)	0.04 (0.04)	-0.01 (0.02)	-0.03 (0.04)	0.01 (0.02)	-0.02 (0.02)	0.01 (0.02)
Constant	0.28*** (0.08)	0.13** (0.06)	0.60*** (0.07)	-0.07 (0.17)	0.04 (0.11)	1.03*** (0.17)	0.79** (0.39)	-2.57*** (0.41)	2.78*** (0.35)
Observations	20,366	20,366	20,366	2,468	2,468	2,468	10,105	10,105	10,105
Log Likelihood	-15,389.93	-12,559.88	-13,810.54	-1,779.31	-968.37	-1,833.13	-6,015.67	-6,335.31	-4,560.57
Akaike Inf. Crit.	30,851.86	25,191.76	27,693.08	3,630.61	2,008.75	3,738.26	12,103.33	12,742.63	9,193.13

This tables shows the marginal effects from linear probability models estimated for each mismatch category ($S < A$: survey reported income is below the value in the administrative database; $S = A$: survey reported income corresponds to administrative database within a bandwidth of $\pm 5\%$; $S > A$: survey reported income is above the value in the administrative data) per income type (wages, pensions and unemployment benefits). The respective dependent variable corresponding to columns (2) to (9) is a binary indicator, equal to one if the mismatch type is observed for the corresponding income type and zero otherwise. Standard errors based on 500 bootstrap replicates are reported in parenthesis. In contrasted to the original specification corresponding to table 5 in the paper and table C.1 in the appendix, income deciles are not included in this specification.

Table C.3: Multinomial logit regressions: p-values (see table 5)

	Wages			Pensions			Unemp. benefits		
	SA	SA	SA
(i) Social desirability									
<i>Relative income (Ref: 5. Decile)</i>									
1. Decile	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.117	0.000
2. Decile	0.000	0.000	0.000	0.071	<i>0.031</i>	0.000	0.005	0.315	0.001
3. Decile	0.000	0.000	0.000	0.060	0.529	0.006	0.951	<i>0.017</i>	0.140
4. Decile	0.164	0.000	0.000	0.643	0.057	0.129	0.795	<i>0.021</i>	0.123
6. Decile	0.000	0.648	0.000	0.000	0.003	0.254	<i>0.010</i>	0.567	0.008
7. Decile	0.000	0.459	0.000	0.000	0.001	0.752	<i>0.018</i>	0.869	0.051
8. Decile	0.000	0.056	0.000	0.000	0.000	0.344	0.000	0.753	0.000
9. Decile	0.000	0.000	0.000	0.000	0.001	0.003	0.000	<i>0.042</i>	0.000
10. Decile	0.000	0.000	0.000	0.000	0.000	0.005	0.000	0.096	0.000
(ii) Sociodemographic characteristics									
<i>Gender (Ref: Female)</i>									
Male	0.000	0.000	0.000	0.000	0.522	0.000	<i>0.030</i>	0.187	0.001
<i>Age</i>									
Age	0.584	0.690	0.810	0.101	0.914	0.051	<i>0.032</i>	0.001	0.874
<i>Education (Ref: Compulsory)</i>									
1 st stage tertiary	0.000	0.422	0.000	0.133	0.883	0.094	0.626	0.061	0.534
Post-secondary	0.000	0.879	0.000	0.003	0.199	<i>0.035</i>	0.299	0.887	0.343
2 nd stage tertiary	0.239	0.173	<i>0.015</i>	0.767	0.687	0.530	0.617	0.594	0.993
Upper secondary	0.000	0.545	0.000	0.000	<i>0.044</i>	0.009	0.095	0.836	0.140
<i>Country of birth (Ref: AUT)</i>									
EU10	0.355	0.867	0.385	0.900	0.669	0.696	0.086	0.087	0.614
EU15	0.068	0.877	0.097	0.357	0.003	0.000	0.945	0.880	0.960
other	0.001	0.004	0.183	0.983	0.730	0.698	0.774	0.838	0.620
Turkey	<i>0.046</i>	0.696	0.071	0.570	0.936	0.665	0.326	0.912	0.304
Yugosphere	0.008	0.003	0.975	0.265	0.111	0.751	<i>0.040</i>	<i>0.020</i>	0.809
<i>Health status (Ref: Very bad)</i>									
Bad	0.420	0.657	0.615	0.304	<i>0.014</i>	0.162	0.152	0.903	0.111
Fair	0.497	0.985	0.410	0.260	0.001	<i>0.021</i>	0.694	0.165	0.260
Good	0.179	0.737	0.212	0.083	0.000	<i>0.012</i>	0.308	0.594	0.204
Very good	<i>0.033</i>	0.724	<i>0.029</i>	0.001	0.000	0.407	0.436	0.151	0.119
<i>Degree of urbanisation (Ref: < 10 000 inhabitants)</i>									
<10 000 inhabitants	0.216	<i>0.044</i>	0.631	<i>0.011</i>	<i>0.022</i>	0.683	0.222	0.442	0.502
>100 000 inhabitants	0.000	0.107	0.005	0.000	<i>0.017</i>	0.169	0.000	0.084	<i>0.038</i>
<i>Changes in employment status (Ref: None)</i>									
Once	<i>0.030</i>	<i>0.029</i>	0.912	0.000	0.000	0.009	0.095	0.074	0.525
Thrice or more	0.858	0.060	<i>0.032</i>	0.407	0.586	0.597	<i>0.012</i>	<i>0.017</i>	0.358
Twice	<i>0.050</i>	0.777	0.054	0.005	0.000	0.833	0.060	<i>0.047</i>	0.556
<i>Months in corresponding employment status (Ref: 12 months)</i>									
<6 months	0.000	0.000	0.000	0.181	<i>0.014</i>	0.224	0.000	0.178	0.000
6-8 months	0.000	<i>0.023</i>	0.009	0.209	0.833	0.344	0.000	0.343	0.000
9-11 months	0.166	0.001	0.204	0.639	0.219	0.342	<i>0.047</i>	0.338	<i>0.027</i>
(iii) Survey setting									
<i>Mode of interview (Ref: CAPI)</i>									
CATI	0.139	0.005	0.577	0.166	0.105	0.656	0.085	<i>0.026</i>	0.585
<i>Interviewer</i>									
Same interviewer	0.603	0.430	0.139	0.960	0.382	0.296	<i>0.043</i>	0.067	0.716
<i>Type of interview (Ref: Personal)</i>									
Proxy	0.000	0.000	<i>0.013</i>	<i>0.015</i>	0.000	0.143	0.927	0.765	0.882
<i>Month of interview (Ref: March to May)</i>									
June to Aug.	0.006	0.000	0.471	0.246	<i>0.046</i>	0.224	0.926	0.922	0.986
Sept. to Nov.	0.226	0.000	<i>0.045</i>	0.000	0.000	0.611	0.961	0.857	0.865
<i>Year of interview (Ref: 2008)</i>									
2009	0.872	0.366	0.353	0.860	0.057	0.061	0.481	0.413	0.260
2010	0.670	0.906	0.569	0.060	<i>0.014</i>	0.444	0.517	0.119	0.094
2011	0.603	0.860	0.434	0.000	0.000	0.000	0.777	0.590	0.519
(iv) Learning effect									
<i>Wave of interview (Ref: 1st)</i>									
2 nd	0.295	0.736	0.393	<i>0.030</i>	0.077	0.946	<i>0.043</i>	0.059	0.661
3 rd	0.209	0.516	0.417	0.243	<i>0.019</i>	0.103	0.253	<i>0.040</i>	0.644
4 th	0.069	0.103	0.600	0.199	0.242	0.885	0.339	0.621	0.581

Source: SILC 2008-2011, own calculations. This table shows the estimated average marginal effects of multinomial regressions per income type (wages, pensions and unemployment benefits) in three categories: (1) S<A: reported income is below the value in administrative data (2) S=A: survey income corresponds to administrative data. (3) S>A: reported income is above the value in administrative data. Estimates with an associated p-value below 0.01 are depicted in bold, values below 0.05 in italics. Insignificant estimates (p-value above 0.05) are given in gray.