DIW Berlin

German Institute
for Economic Research

Discussion Papers

736

Joachim R. Frick
Markus M. Grabka

Item Non-response and Imputation of Annual Labor Income in Panel Surveys from a Cross-National Perspective

Berlin, October 2007

Opinions expressed in this paper are those of the author and do not necessarily reflect views of the institute.

IMPRESSUM

DIW Berlin
German Institute for Economic Research
Mohrenstr. 58
10117 Berlin
Tel. +49 (30) 897 89-0
Fax +49 (30) 897 89-200
http://www.diw.de

ISSN print edition 1433-0210
ISSN electronic edition 1619-4535

Available for free downloading from the DIW Berlin website.

DIW Berlin

Discussion Papers 736

Joachim R. Frick*

Markus M. Grabka**

Item non-response and Imputation of Annual Labor Income in Panel Surveys from a Cross-National Perspective

Berlin, October 2007

* DIW Berlin, SOEP, and TU Berlin; jfrick@diw.de

** DIW Berlin, SOEP; mgrabka@diw.de

Abstract

Using data on annual individual labor income from three representative panel datasets (German SOEP, British BHPS, Australian HILDA) we investigate a) the selectivity of item non-response (INR) and b) the impact of imputation as a prominent post-survey means to cope with this type of measurement error on prototypical analyses (earnings inequality, mobility and wage regressions) in a cross-national setting. Given the considerable variation of INR across surveys as well as the varying degree of selectivity build into the missing process, there is substantive and methodological interest in an improved harmonization of (income) data production as well as of imputation strategies across surveys. All three panels make use of longitudinal information in their respective imputation procedures, however, there are marked differences in the implementation.

Firstly, although the probability of INR is quantitatively similar across countries, our empirical investigation identifies cross-country differences with respect to the factors driving INR: survey-related aspects as well as indicators accounting for variability and complexity of labor income composition appear to be relevant. Secondly, longitudinal analyses yield a positive correlation of INR on labor income data over time and provide evidence of INR being a predictor of subsequent unit-non-response, thus supporting the "cooperation continuum" hypothesis in all three panels. Thirdly, applying various mobility indicators there is a robust picture about earnings mobility being significantly understated using information from completely observed cases only. Finally, regression results for wage equations based on observed ("complete case analysis") vs. all cases and controlling for imputation status, indicate that individuals with imputed incomes, ceteris paribus, earn significantly above average in SOEP and HILDA, while this relationship is negative using BHPS data. However, once applying the very same imputation procedure used for HILDA and SOEP, namely the "row-and-column-imputation" approach suggested by Little & Su (1989), also to BHPS-data, this result is reversed, i.e., individuals in the BHPS whose income has been imputed earn above average as well. In our view, the reduction in cross-national variation resulting from sensitivity to the choice of imputation approaches underscores the importance of investing more in the improved cross-national harmonization of imputation techniques.

Keywords: Item non-response, imputation, income inequality, income mobility, panel data, SOEP, BHPS, HILDA

JEL-code: J31, C81, D33

Acknowledgement: We thank conference participants at ISA 2006 in Durban/South Africa, MOLS 2006 at the University of Essex, IARIW 2006 in Joensuu/Finland, SOEP 2006 in Berlin, BHPS 2007 in Colchester/UK, HILDA 2007 in Melbourne/AUS and especially Robert Breunig, Ian Plewis, Peter Lynn and Stephen P. Jenkins for helpful comments on earlier versions of this paper.

Content

1 Motivation **1**

2 Data and Incidence of INR **3**

2.1 The three panels 3

2.1.1 BHPS 3

2.1.2 HILDA 3

2.1.3 SOEP 4

2.2 Incidence of INR and the “cooperation continuum” 5

2.3 Selectivity of INR 7

3 Imputation rules in the three surveys **11**

4 Empirical application on the impact of imputation **14**

4.1 Imputation and the analysis of earnings inequality and mobility 14

4.2 Imputation and wage regressions 19

5 Conclusion **24**

References **27**

Appendix: **30**

Appendix A: Detailed Information on Item Non-Response and Imputation by Survey 30

Appendix B: Exact Wording of Earnings-Related Questions in Original Survey Instruments 35

Content of Tables and Figures

Table 1: Estimating the probability for INR on labor income – Results from random effects probit models ... 9

Table 2: Income Inequality and Income Mobility by Imputation Status ... 17

Table 3: Results from Fixed-Effects Wage Regression; Dependent Variable: Log Annual Labor Income ... 22

Figure 1: Observations with INR on labor income by post-imputation deciles (in %) ... 5

Figure 2: INR in a longitudinal Perspective: The Case of individual labor earnings ... 7

Figure A-1: Observations with INR on Labor Income by Deciles (in %) Based on Original and Alternative Imputation in the BHPS and Incorporating a “Fresh” Panel (SOEP-F) ... 30

Figure A-2: Item Non-Response from a Longitudinal Perspective: The Case of Individual Labor Earnings ... 31

Figure A-3: Inequality of Individual Labor Earnings (based on observed cases only) ... 32

Table A-1: Results from Quantile Wage Regressions; Dependent Variable: Log Annual Labor Income (normalized) ... 33

1 Motivation

A common phenomenon in population surveys is the failure to collect complete information due to respondent's unwillingness or lacking capability to provide a requested piece of information. This non-response behavior is referred to as item non-response (INR) and may be caused by a respondent's reservation to answer to a question that appears to be too sensitive, or that affects confidentiality and privacy or it may simply arise from the fact that the correct answer is not known (given the underlying complexity of the surveyed construct). While in general, simple demographic information such as sex, age or marital status is not very sensitive to ask for, thus leading to low incidence of INR, wealth or income questions, however, are typically associated with higher rates of INR (e.g., Riphahn and Serfling 2005, Hawkes & Plewis 2006). There is increasing literature that explicitly acknowledges this phenomenon in micro-economic research as a specific form of measurement error (e.g., Cameron & Trivedi 2005). Most importantly, INR on income questions has been found to be selective with respect to inequality as well as to mobility (e.g., Jarvis & Jenkins 1998, Biewen 2001; Frick & Grabka 2005, Watson & Wooden 2006). Although there is growing awareness of the risk of selectivity inherent in item non-response (at least since Ferber 1966), much of the literature on non-response behavior in longitudinal studies focuses on unit non-response and on the possible bias arising from selective attrition in such surveys (see, e.g., Groves 2006, Groves & Couper 1998, Lepkowski & Couper 2002, Watson & Wooden 2006). A minority of studies (e.g., Lee, Hu & Toh 2004; Serfling 2006; Burton, Laurie & Moon 1999) have argued that the two types of non-response should be analysed in a common framework and have proposed that respondents be arranged on a cooperation continuum ranging from (a) those who will (always) be willing to participate in surveys and also to provide valid answers (b) those who will be more or less willing to cooperate (i.e., who will take part in the survey as such but who may refuse to answer certain items, causing INR) and finally (c) those who will not take part at all (causing unit-non response, UNR). Above and beyond these basic traits, there will most likely also be situational factors that interfere with the individual's basic willingness or ability to cooperate. These may include severe illness, exceptional events such as the death of a relative, or an unpleasant relationship with the interviewer.

All these arguments will apply to any national (panel) survey. But how do they relate to internationally comparative research? In recent years, a large body of empirical literature has emerged focusing on cross-national comparisons. Databases such as the European Community Household Panel (ECHP) provide the empirical basis for such studies across countries or welfare regimes with harmonized (or functionally equivalent defined) micro-data (e.g., Nicoletti & Peracchi 2006). A typical welfare economics application arises from the need to empirically monitor the harmonization of social politics in the EU by using, for example, harmonized pre- or post-government income measures to assess national redistribution policies. For optimal comparability, the harmonization of micro-data (e.g., income measures) is obviously a crucial issue in this context, but the same is true for other methodologically relevant decisions in the pre- and post-data collection phase regarding the definition of relevant population, the choice of data collection method (e.g., interview or register data), and the management of attrition-related phenomena.

This paper deals with the handling of missing (annual gross) labor income information caused by INR in three major national panel data sets, the British Household Panel Study (BHPS), the German Socio-Economic Panel Study (SOEP), and the Survey of Household, Income and Labour Dynamics in Australia (HILDA). When the underlying missing process is not MCAR (missing completely at random, see Rubin, 1976), INR is often dealt with by imputation, the strategy applied in all three datasets considered here. However, while all three surveys take advantage of the longitudinal character of the underlying panel data, the actual implementation of the respective imputation strategies differs. This aspect might be of particular importance for cross-national comparability. Following the postulates of the “Canberra Group on Household Income Measurement” for harmonized national household income statistics (Canberra Group, 2001), we present evidence in the following that it is important to harmonize not only income measurement but also the procedures for handling and possibly also imputing observations affected by INR.

The paper is organized as follows: Chapter 2 outlines the basic characteristics of the three panel surveys including the incidence of INR (with respect to labor income), demonstrating the selectivity entailed by INR and investigating the longitudinal relationship between INR and subsequent UNR. Chapter 3 describes the imputation methods applied in the three surveys. Based on rather typical empirical research questions using labor income, Chapter 4 demonstrates the impact of imputation on earnings inequality and mobility, as well as on wage regressions. Finally, Chapter 5 concludes from the perspective of cross-nationally comparative research.

2 Data and Incidence of INR

2.1 The three panels

The following section briefly describes the underlying panel datasets, all of which are included in the Cross-National Equivalent File as of 2007 (CNEF; see Burkhauser et al. 2001). The annual labor income information as well as the accompanying information on imputation status (flag) which is used in this paper is included as a standard variable in the CNEF.

2.1.1 BHPS

The British Household Panel Survey (BHPS) is carried out by the Institute for Social and Economic Research (ISER) at the University of Essex (see Taylor 2005; http://www.iser.essex.ac.uk/ulsc/bhps/doc/vola/contentsI.php). It was started in 1991 with about 5,500 households and roughly 10,300 individuals surveyed in England. The sample was extended in 1999 with about 1,500 households in each, Scotland and Wales. In 2001 a further sample of 2,000 households in Northern Ireland was added, supporting panel research for all of the UK. However, the following analyses are based on the original sample only, including data for waves 1991 through 2004. In 1999, the interview mode was entirely changed for the whole sample from Paper and Pencil to CAPI. Annual gross labor income in the BHPS is surveyed in principle by means of a single question where the amount of the last gross pay including any overtime, bonuses, commission, tips or tax refund is asked (see appendix B). Apparently, such a "one-shot" question targeting at a rather complex construct, namely the aggregation of a variety of income sources over a period of twelve months, bears a high risk of measurement error following from understating, rounding, omitting, and non-responding.

2.1.2 HILDA

The "Household, Income and Labour Dynamics in Australia" (HILDA) Survey started in 2001 with about 7,700 participating households (Watson 2005; http://www.melbourneinstitute.com/hilda/). HILDA, compiled by the Melbourne Institute of Applied Economic and Social Research, provides information on living conditions of private households in Australia. By and large, the panel design used in HILDA resembles the one of

BHPS. The sampling unit is the private household, and only original members of those households are to be tracked in case of residential mobility.

Annual gross labor income in HILDA comes from three sources of information. Firstly, all respondents are asked for their total wages and salaries from *all* jobs over the last financial year (July 1st of the previous year to June 30st of the survey year). Secondly, income from own business or farming from incorporated businesses were added and finally the total share of profit or loss from unincorporated businesses or farms are summed-up (see appendix B). One time payments and irregular payments are not explicitly surveyed. Data from the first five waves, covering the period 2001 to 2005, is used in this paper.

2.1.3 SOEP

The German SOEP is the longest running household panel study in Europe (cf. Haisken-DeNew and Frick 2005; Wagner et al 2007; http://www.diw.de/gsoep). All household members aged 17 and over are surveyed individually each year, and an additional household interview is conducted with the head of household. Interviews usually take place face-to-face with the interviewer filling in the questionnaire. Although Computer Assisted Personal Interviewing (CAPI) was introduced in 1998, paper and pencil interviews are still a most relevant interview mode. In order to keep the survey sample representative, various new subsamples have been incorporated since the initial start in 1984. In 1990 and 1995 new samples were introduced to capture the effects of unification with East Germany and recent immigrants, respectively. A major "refreshment sample" (called Sample F) was started in 2000. In this paper, we will show results based on the entire SOEP sample (survey years 1992 to 2004) as well as separately for the new sample F (survey years 2000 to 2004), in order to control for eventual panel effects in the old sample. Moreover, sample F may be more comparable to the rather young HILDA sample which was started in 2001, while the results based on the overall SOEP-sample may be better comparable to the BHPS results which capture the period 1991-2004. The SOEP sample as of 2004 includes about 11,800 households, thereof 4,200 in Sample F.

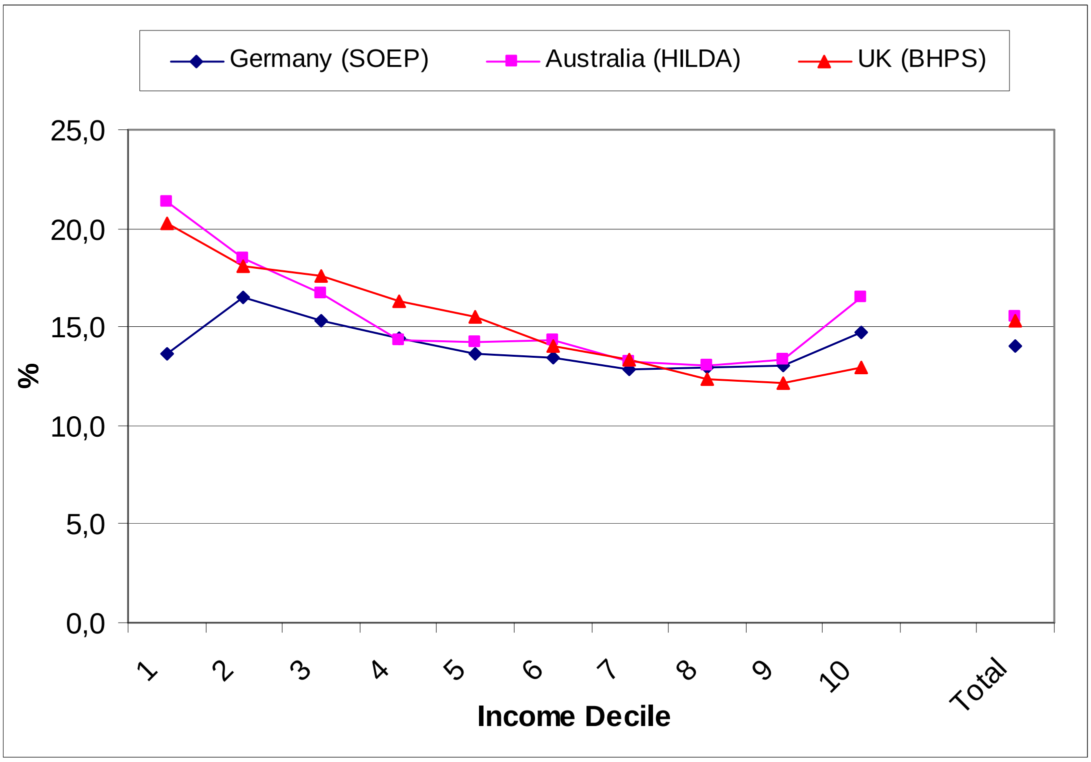
Information about gross annual labor income is gathered from 10 different single questions. In principle, from each individual labor income for the previous calendar year is asked separately for dependent employment as well as for self-employment. In each case, the average monthly amount is collected as well as the number of months with receipt of this income type.

Additionally, one time or irregular payments like 13th or 14th monthly salary, holiday pay or bonuses are separately asked for and added together (see appendix B for the exact wording of the respective income questions in the SOEP).

2.2 Incidence of INR and the "cooperation continuum"

Given the apparent differences among the three panels in the means used to collect annual labor earnings data, we find surprisingly little cross-national variation in the incidence of item non-response (Figure 1).

Figure 1: Observations with INR on labor income by post-imputation deciles (in %)



Note: Contingent on the imputation as provided in the original datasets described in Section 3.

Source: SOEP survey years 1992-2004; HILDA survey years 2001-2005; BHPS survey years 1991-2004.

While about 16% of all observations in the relatively young HILDA survey suffer from INR, SOEP and BHPS have shares of about 14% and 15%, respectively. Compared to an INR rate of only around 8% in SOEP for the question on "current monthly net household income", the high share of missing data might be related to the high number of different income items collected (up to ten), which raises the odds of at least one missing component. In the case of the BHPS this finding is rather unexpected, however, given that merely one question is asked[1]. On the other hand, the HILDA and BHPS questioning offers a "Don't know" category, which may as well tempt respondents to refrain from giving a positive value instead (see Burton et al 1999; Schräpler 2003b). Finally, one should note that any seemingly valid observed income information may be affected by measurement error as well, e.g., by rounding or rough estimation (see e.g., Hanisch 2005).

Depending on the imputation procedures used (to be described below), the incidence of INR in annual labor earnings appears to follow a somewhat u-shaped pattern over the labor income distribution, with INR generally more prominent among the lower incomes (see also Biewen 2001).[2]

Given our substantive analytical interest in inequality and mobility analysis, there is an inherent need to control for possible time dependence of INR. Separating individual observations by imputation status at time t_0 (i.e., "valid" observed income [Obs. in t_0] vs. INR [Imp. in t_0])[3], Figure 2 differentiates four potential outcomes at time t_1, namely "valid earnings information", "INR with subsequent imputation", "zero labor income due to leaving the labor force" and "attrition".

In all panel studies, we not only find indications of state dependence of INR, but also clear support for the "cooperation continuum" hypothesis (see Burton et al 1999, Loosveldt et al 2002, Schräpler 2004), according to which INR is a valid predictor of subsequent unit-non-response, namely attrition. Figure A-1 in the appendix provides a more differentiated picture of these processes across the income distribution. There is a stable finding of higher unit non-response at all income levels among those with INR in the previous wave than among those with observed income information. However, for the former group, we find that INR increases with income in the current wave. [4]

---

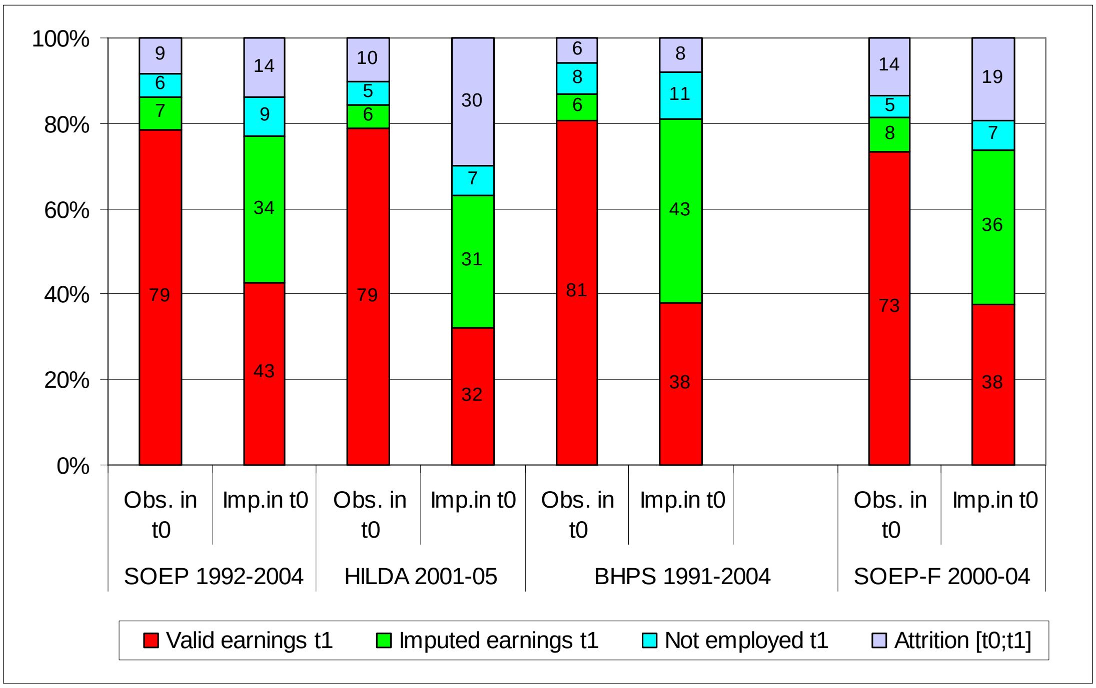
[1] However, there is information available on earnings received on September 1 of the previous as well as of the current year in case of any variation across time. This may have been considered in the generation of the annual income measure used here.

[2] An exception is the relatively new SOEP Sample F (see Figure A-1 in the Appendix). Here an undulated distribution can be observed with the highest decile showing the highest share of INR.

[3] Leaving out observations out of the labor force, i.e., those with zero labor earnings.

[4] Separating "refusals" from "don't knows" as the underlying reason for INR, Schräpler (2003b) finds attrition in the subsequent wave to be significant only among refusers, while no such strong relationship is found among those who "don't know".

Figure 2: INR in a longitudinal Perspective: The Case of individual labor earnings



Source: SOEP survey years 1992-2004; HILDA survey years 2001-2005; BHPS survey years 1991-2004.

2.3 Selectivity of INR

As mentioned above, INR may be a function of various factors such as the respondent's unwillingness to answer questions that are perceived as highly sensitive or in violation of confidentiality and privacy, the fact that the information requested is too complex or simply that the answer is not known (e.g., Schräpler 2003, 2004). The specific formulation of questions and the complexity of the construct being measured may also play a role (Hill & Willis 2001). One strand of research has shown that the interview situation, the survey mode, the presentation of the question with a "don't know" answer option, and possible interviewer effects including a change of interviewers in panel studies, are relevant determinants of INR (e.g., Rendtel 1995, Pickery et al. 2001, Dillmann et al 2002, Riphahn & Serfling 2005, Groves 2006, Watson & Wooden 2006).

For the sake of cross-national comparability it is most important to control for whether the missing mechanisms coincide for the datasets considered here. For each of the panels specifically and utilizing the panel nature of the underlying data, we specify a random effects model estimating the probability of INR on our measure of annual labor earnings.[5] Based on currently employed individuals (including the self-employed) aged 20 to 65 years, we control for socio-demographic characteristics, the interview situation, the survey experience of the respondent, as well as for the complexity of the income receipt. The latter is operationalized by various dummy-variables indicating changes in an individual's labor market career over the previous (calendar or financial) year by identifying experience of unemployment and exit from education (see Table 1).

In brief, INR on previous year's labor income is clearly more frequent among the self-employed, while it becomes less likely with an increasing number of months in (full- or part-time) employment. As expected, one finds a higher probability of INR in SOEP and HILDA among those who were unemployed at some point within the last year, but the opposite effect is seen in BHPS. Inconsistent findings are also found with respect to gender (SOEP and BHPS showing more INR among men, while women in HILDA provide more often a seemingly valid answer to labor income questions). In HILDA and BHPS, there is a negative education effect—i.e., more highly educated individuals are less likely to show non-response—while there is no such effect in the SOEP. Controlling for long-term employment patterns, it appears that INR is reduced with tenure, but at a reduced pace. *Ceteris paribus*, foreigners in SOEP are more likely to provide income data, while there is no significant immigrant/citizenship effect in HILDA or the BHPS. The UK and Australian panels do, however, confirm our expectation of higher response likelihood among public servants. In Germany, there is a pronounced negative probability of INR among East Germans. The results for the INR-reducing effect of survey experience, here measured by the number of interviews, are consistent across all panels.

[5] All empirical results presented in this paper are based on calculations using Stata (version 9.0), including the ado-modules INEQUAL7, INEQDECO, IMOBFOK, FOKMOB, SHORMOB authored by Stephen P. Jenkins and Philippe van Kerm, respectively.

Table 1: Estimating the probability for INR on labor income – Results from random effects probit models

	Germany (SOEP)		Australia (HILDA)		UK (BHPS)		Germany (SOEP) – F	
Age	-0.000	*(0.006)*	-0.013	*(0.011)*	0.001	*(0.007)*	-0.011	*(0.015)*
Age squared	0.000	*(0.000)*	0.000	*(0.000)*	0.000+	*(0.000)*	0.000	*(0.000)*
Male	0.044*	*(0.021)*	-0.237**	*(0.034)*	0.075*	*(0.030)*	-0.012	*(0.047)*
Education level = Low	0.058*	*(0.026)*	-0.104*	*(0.042)*	-0.073*	*(0.033)*	-0.052	*(0.069)*
Education level = Intermediate	0.052	*(0.033)*	-0.236**	*(0.048)*	-0.142**	*(0.032)*	-0.105	*(0.085)*
Education level = University	-0.003	*(0.032)*	-0.250**	*(0.067)*	-0.266**	*(0.081)*	-0.055	*(0.079)*
Disability status	0.031	*(0.039)*	0.106**	*(0.040)*	-0.019	*(0.030)*	-0.020	*(0.088)*
Married	-0.005	*(0.021)*	-0.149**	*(0.039)*	-0.009	*(0.025)*	0.030	*(0.049)*
# HH members aged 0-14	-0.001	*(0.011)*	0.026	*(0.017)*	0.005	*(0.009)*	-0.002	*(0.026)*
Metrop. area	-0.036	*(0.028)*	-0.145**	*(0.042)*	-0.087**	*(0.028)*	-0.191**	*(0.068)*
Remote area	0.040*	*(0.019)*	0.092+	*(0.051)*	-0.107**	*(0.037)*	0.091*	*(0.043)*
Tenure	-0.003	*(0.003)*	-0.027**	*(0.005)*	-0.068**	*(0.004)*	0.009	*(0.006)*
Tenure squared	0.000+	*(0.000)*	0.001**	*(0.000)*	0.002**	*(0.000)*	-0.000	*(0.000)*
Foreigner	-0.064*	*(0.032)*	0.058	*(0.043)*	0.020	*(0.040)*	-0.281**	*(0.094)*
Public service	0.027	*(0.021)*	-0.221**	*(0.083)*	-0.092**	*(0.022)*	-0.007	*(0.046)*
Firm size: small	-0.013	*(0.019)*	0.257**	*(0.037)*	0.015	*(0.022)*	0.014	*(0.042)*
Firm size: large	-0.008	*(0.021)*	-0.045	*(0.062)*	-0.012	*(0.036)*	0.019	*(0.048)*
East Germany	-0.216**	*(0.024)*	-	-	-	-	-0.228**	*(0.055)*
Months full-time (last year)	-0.022**	*(0.003)*	-0.048**	*(0.008)*	-0.015**	*(0.004)*	-0.015*	*(0.008)*
Months part-time (last year)	-0.020**	*(0.003)*	-	-	-	-	-0.025**	*(0.008)*
Months in unemployment (last year)	0.090**	*(0.031)*	0.237**	*(0.065)*	-0.083*	*(0.036)*	0.099	*(0.076)*
Left educ. system during last year	0.007	*(0.033)*	-0.028	*(0.047)*	-0.127+	*(0.075)*	-0.151+	*(0.090)*
Self employed	0.468**	*(0.028)*	1.328**	*(0.041)*	1.053**	*(0.029)*	0.624**	*(0.062)*
Problems during Interview	0.212**	*(0.016)*	-0.206*	*(0.098)*	0.117	*(0.074)*	0.133**	*(0.036)*
# Interviews = 2	-0.125+	*(0.065)*	-0.221**	*(0.081)*	-0.014	*(0.084)*	-0.161	*(0.103)*
# Interviews = 3+	-0.353**	*(0.047)*	-0.385**	*(0.062)*	-0.294**	*(0.060)*	-0.301**	*(0.075)*
Constant	-1.297**	*(0.132)*	-0.545+	*(0.303)*	-0.841**	*(0.139)*	-0.800**	*(0.310)*
Obs.	120818		35238		72696		22456	
N	24178		10722		11134		7063	
-2 Log-Likelihood	-36493.31		-5151.03		-24036.69		-8807.08	
Pseudo-R-squared	.1254		.1609		.2120		0.1261	

Note: Time effects controlled, but not reported. Standard errors in parentheses;
Significance level: + significant at 10%; * significant at 5%; ** significant at 1%.
Source: SOEP survey years 1992-2004; HILDA survey years 2001-2005; BHPS survey years 1991-2004.

Summing up the results of this section, we observe a similar incidence of INR across surveys despite the fact that the surveys use rather different ways of asking for labor income. With respect to the selectivity of INR in the three panels, we do find some cross-country similarities, however, there emerge country-specific reasons for INR as well.[6] Even the two relatively young panels (HILDA and the SOEP Sample F) are not congeneric, which supports the importance of such cross-country analyses.

[6] In light of such selection problems, it seems surprising that simply dropping observations with missing data on a variable of interest—thus assuming the missing mechanism to be completely at random—still appears to be common practice (see eg. Gebel & Pfeiffer 2007), not only in labor economics research.

3 Imputation rules in the three surveys

Imputation is a most prominent way to handle INR in micro-data. An exhaustive description of such procedures other than the one used in SOEP, BHPS and HILDA is beyond the scope of this paper. However, it should be noted that even a very sophisticated approach of substituting for non-response may not completely eliminate any bias resulting from it. As such, the choice of the adequate imputation technique is a problem in itself. Potential bias due to imputation may creep in due to "regression-to-the-mean effects" and a potential change in total variance—most likely a decline—may occur.[7]

Annual individual labor income in the BHPS is imputed using a regression based predictive mean matching (PMM) procedure proposed by Little (1988) also known as regression hot deck. The basic idea of the PMM is the use of observed predictor variables from a linear regression to predict variables with missing values. The advantage of this method is, that a possible real value is imputed and that a random error component is added to preserve variance. The PMM method adopted in the BHPS also considers longitudinal information from a shifting three-year window. Depending on the availability of observed information about labor income in previous and subsequent waves as well as eventual job changes, either forward or backward imputation is applied resulting in 14 different regression models (ISER 2002). An indication for the imputation quality is given by the corresponding R-squares of the underlying regression estimations. In the first three waves of the BHPS, the share of explained variance of gross usual pay – which is the main income component for annual individual labor income – varies between 0.78 and 0.94 (ISER 2002: A5-27).

HILDA and SOEP are both using a two-step procedure to impute any income information missing due to INR. The primary method is based on the "row-and-column-imputation", described by Little & Su (1989) (hereafter L&S). The row-and-column-imputation takes advantage of cross-sectional as well as individual longitudinal information – using income data avai-

7 See Rubin (1987) for a discussion of imputation methods and the advantages of multiple imputation that allow us to assess the degree of variation added to parameter estimates as a result of imputation. Most producers of micro-data (including those of the three panel datasets used in this paper) do not, however, provide multiply imputed information at this time. One exception is the US Survey of Consumer Finances (Kennickell & McManus 1994). Multiple imputation is also used to correct for item non-response in the wealth data collected in the 2002 wave of the German SOEP (Frick, Grabka, and Marcus 2007). For an evaluation of alternative treatments of INR by means of weighting see, e.g., Rässler & Riphahn (2006) or Little & Rubin (2002).

lable from the entire panel duration – by combining row (unit) and column (period/trend) information and adds a stochastic component resulting from a nearest neighbor matching, i.e.,

*imputation = (row effect) * (column effect) * (residual).*

Using an exemplary panel with 20 waves of data, the column effects are given by

$$(1) \qquad c_j = (20 * \bar{Y}_j) / \sum_{k=1}^{20} \bar{Y}_k$$

and are calculated for each of the 20 waves of data, where j = 1, ..., 20 and $\bar{Y}_j$ is the sample mean income for year j. The row effects are given by:

$$(2) \qquad r_i = m_i^{-1} * \sum_{j=1}^{20} (Y_{ij} / c_j)$$

and are computed for each sample member. Y_{ij} is the income for individual i in year j and m_i is the number of recorded periods. Sorting cases by r_i and matching the incomplete case i with information from the nearest complete case, say l, yields the imputed value

$$(3) \qquad i = [r_i] * [c_j] * [Y_{lj} / (r_l * c_j)]$$

The three terms in brackets represent the *row*, *column*, and *residual* effects. The first two terms estimate the predicted mean, and the last term is the stochastic component of the imputation stemming from the matching process. While the SOEP applies this L&S-procedure to the entire population (Grabka & Frick 2003) as described above, HILDA uses a modification of this technique by matching donors and recipients within imputation classes defined by seven age groups (Starick 2005).

A secondary method is needed whenever longitudinal information is lacking. This includes not only first time respondents, but all those observations for whom a given income variable has been surveyed for the very first time. Hence, a purely cross-sectional imputation method needs to be applied. In the case of HILDA a nearest neighbor regression method (similar to that used by the BHPS) is deployed. In the SOEP, this is accomplished by means of a hot-deck regression model supplemented by a residual term retrieved from a randomly chosen donor with observed income information in the regression model.[8]

In an evaluation of various imputation methods, Starick (2005) argues that "in a longitudinal sense, the Little & Su methods perform much better when compared to the nearest neighbor regression method. Evidence shows that the Little & Su methods preserve the distribution of income between waves. Furthermore, the Little & Su method perform better in maintaining cross-wave relationships and income mobility" (Starick 2005: 31). This finding is also confirmed by Frick and Grabka (2005) for the SOEP by showing that L&S imputation performs better in terms of preserving the distribution than a regression based imputation strategy.[9]

To check for robustness and to control for possible effects of the choice of imputation strategy on the inequality and mobility measures, we use the methodology of Little & Su (1989) for the BHPS data as well. It must be noted that we do not impute the single income components but only the aggregated "annual labor earnings" measure here. About 80% of individuals with missing labor earnings can be imputed with the L&S method, while for the remaining 20% we use the original BHPS regression results. In other words, there are no *longitudinal* earnings data available for the latter group.

In the following we will compare results obtained from the imputation techniques as given by the various original data providers: HILDA, BHPS, and SOEP, where we will also look at results obtained from a "fresh" panel, SOEP Sample F. In the case of the BHPS, we will provide a point of comparison using the alternative imputation method of Little & Su (1989).

[8] An indication for the quality of the secondary imputation in SOEP is given by the R-squares of gross annual labor income, which varies between 0.48 and 0.66.

[9] In a simulation study, Frick and Grabka (2005) use a random sample of approx. 1,000 observations for which a positive value of "labor income from first job" has been observed and who provide longitudinal information as a prerequisite for the L&S procedure. While the L&S procedure overstates inequality by about 9%, the cross-sectional approach understates the Gini by about 18%. This finding is in line with the results of Spiess and Goebel (2003) based on survey and register data for Finland.

4 Empirical application on the impact of imputation

Keeping in mind the above findings on incidence and selectivity of INR across panels as well as the differences and commonalities in the respective imputation process, the following analyses focus on the impact of imputation on prototypical applications. We will first concentrate on distributional aspects (measured by various income inequality indicators) and on earnings mobility derived from wave-to-wave comparisons (again applying various mobility indicators in order to control for robustness of our results (section 4.1)). In section 4.2, we investigate whether imputed observations "behave" differently in a wage regression model, i.e., whether correct inferences can be drawn from a dataset excluding observations with INR.

4.1 Imputation and the analysis of earnings inequality and mobility

Accepting the applied imputation strategies, i.e., assuming that these correctly identify the underlying missing mechanism, obviously any increase in selectivity of non-response will be reflected in the deviation of empirical results based on truly observed cases ("complete case analyses") from those derived on the basis of all observations (i.e., observed plus imputed cases).

A comparison of basic statistics of annual gross labor income (top panel of Table 2) shows income levels (given by mean and median) to be clearly lower among the population with imputed values in the case of BHPS and HILDA, while in the SOEP a reverted tendency can be observed.[10] The result for the overall population ("all cases") thus deviates from the one for the observed cases, e.g., the overall median in HILDA is about 2.2% lower than the value resulting from "observed cases" only. Extending our perspective to cross-sectional measures of inequality, a rather robust picture of understated inequality appears when using "complete case" analysis. For selected indicators, we find statistically significant differences after including imputed values. For example, the 90:10 decile ratio as well as the MLD (mean logarithmic deviation) for the observed cases in HILDA understate inequality by about 5%, while in Germany the top-sensitive HSCV (half-squared coefficient of variation) increases by alm-

[10] The analysis of income inequality is based on pooled, deflated income data for all available years as described in section 2. In case of Australia inequality is rather stable over the 5-year period, whereas in Germany we observe an increase in earnings inequality over the recent years. Finally, the results for Britain show an increase in inequality in most years since 2002 after a period of slightly declining inequality. The development of the top-sensitive HSCV measure appears rather erratic since the late 1990s (see Appendix, Table A-1).

ost 4% and even the change in the rather robust Gini coefficient indicates rising inequality when considering the imputed cases as well. The results obtained from the row-and-column imputation of missing income data in the BHPS instead of the originally provided hot-deck imputation yields somewhat higher imputed values, but inequality among the L&S-imputed observations is less pronounced here. Following from this, the deviation between "all" and "observed" is not significant in any of the measures employed.

As shown above, the missing mechanisms for INR on labor income point towards selectivity with respect to characteristics found more often among attriters. Given that attrition is controlled for in most panel surveys through weighting factors that represent the inverse probability of being selected into and dropping out of the sample, one may assume that the use of population weights in the present context will increase the percentage of the population showing INR. Indeed, the weighted population share containing imputed labor income data is as high as 15% in HILDA, 13% in the overall SOEP sample (but 20% in Sample F), and 18% in BHPS.

With respect to labor income *mobility*, as is true for any longitudinal analyses, one can expect the impact of imputation to be even more relevant because INR may be an issue in at least one of the waves under consideration. For matter of simplification in this application, we just use a series of two-wave balanced panels (pooled across all available waves in each survey), i.e., the effects shown below would be even more pronounced in any multi-wave analyses (see lower panel of Table 2).

Above and beyond the general finding of inequality being understated among the "observed cases", clearly more distinct and statistically significant differences can be found for labor income mobility—conditional on the applied imputation techniques. Depending on the mobility measure used as well as depending on the population share affected by imputation, the results between "observed" and "all" cases (including the imputed ones) deviate in case of the original BHPS by as much as 27% to 47%, using the alternative imputation this change in mobility is somewhat less pronounced (between 19% and 43%). In the SOEP (as well as in SOEP-sample F) the corresponding shares are between 10% and 30% and in HILDA this range is from 15% to 31%.

Focusing only on "complete cases" would yield an even higher loss in statistical power or efficiency due to the massive reduction in the number of observations. The last row in Table 2 indicates that the (weighted) population share containing imputed data in at least one of the two waves considered is as high as 20% in HILDA, 31% in SOEP, 38% in BHPS and even 43% in SOEP’s recent Subsample F.

Table 2: Income Inequality and Income Mobility by Imputation Status

	Germany (SOEP)				Australia (HILDA)			
	Imputation status			*Deviation: "All" vs. "Observed" (%)*	Imputation status			*Deviation: "All" vs. "Observed" (%)*
	"All cases"	"Observed cases"	"Imputed cases"		"All cases"	"Observed cases"	"Imputed cases"	
Basic statistics*								
Mean	24408	24401	24455	*+0.0*	27349	27630	25826	*-1.0*
Median	21940	22077	21010	*-0.6*	23375	23916	21231	*-2.3*
Income inequality								
Theil 0 (Mean log deviation)	0.40964	0.40563	0.43416	*+1.0*	0.45871	0.43896	0.56398	*+4.5*
Gini	0.41405	0.41006	0.43769	*+1.0*	0.42728	0.41922	0.47024	*+1.9*
Half-SCV (top-sensitive)	0.34880	0.33692	0.42106	*+3.5*	0.44557	0.43896	0.53454	*+1.5*
Decile ratio 90:10	13.71	13.66	14.17	*+0.4*	14.916	14.174	23.375	*+5.2*
Decile ratio 90:50	2.13	2.11	2.27	*+1.0*	2.203	2.154	2.465	*+2.3*
Decile ratio 50:10	6.45	6.49	6.25	*-0.6*	6.757	6.579	9.523	*+2.7*
Average N per cross-section	***10773***	***9501***	***1272***	***+13.4***	***9082***	***7876***	***1206***	***+15.3***
Income mobility								
Quintile matrix mobility: Average jump	0.448	0.376	0.713	*+19.1*	0.530	0.459	0.857	*+15.5*
Quintile matrix mobility: Normalized average jump	0.179	0.150	0.285	*+19.3*	0.212	0.184	0.343	*+15.2*
Fields & Ok: Percentage income mobility	24.38	18.89	42.94	*+29.1*	28.81	24.51	49.56	*+17.5*
Fields & Ok: Non-directional	0.333	0.301	0.460	*+10.6*	0.447	0.384	0.733	*+16.4*
Shorrocks: Using Gini Coefficient	0.0290	0.0242	0.0465	*+19.8*	0.0445	0.0340	0.0837	*+30.8*
Average N per 2-wave balanced panel	***9878***	***7554***	***2324***	***+30.8***	***7474***	***6236***	***1238***	*+19.9*

… *contd.*

contd. Table 2

	UK (BHPS)				UK (BHPS) – "L & S"				Germany (SOEP) - F			
	Imputation status			*Deviation: "All" vs. "Observed" (%)*	Imputation status			*Deviation: "All" vs. "Observed" (%)*	Imputation status			*Deviation: "All" vs. "Observed" (%)*
	"All cases"	"Observed cases"	"Imputed cases"		"All cases"	"Observed cases"	"Imputed cases"		"All cases"	"Observed cases"	"Imputed cases"	
Basic statistics*												
Mean	13621	13872	12237	*-1.8*	13849	13872	13727	*-0.2*	24695	24309	26504	*+1.6*
Median	11360	11677	9713	*-2.7*	11553	11677	10956	*-1.1*	21781	21774	22245	*+0.0*
Income inequality												
Theil 0 (Mean log deviation)	0.44248	0.40733	0.63063	*+8.6*	0.42109	0.40733	0.4972	*+3.4*	0.44672	0.44613	0.44630	*+0.1*
Gini	0.42804	0.42086	0.4654	*+1.7*	0.42681	0.42086	0.4590	*+1.4*	0.43357	0.43012	0.44727	*+0.8*
Half-SCV (top-sensitive)	0.47092	0.44888	0.61306	*+4.9*	0.46516	0.44888	0.5571	*+3.6*	0.38577	0.36422	0.46560	*+5.9*
Decile ratio 90:10	12.688	11.959	16.348	*+6.1*	12.427	11.959	14.872	*+3.9*	15.43	15.40	14.89	*+0.2*
Decile ratio 90:50	2.333	2.292	2.532	*+1.8*	2.316	2.292	2.438	*+1.1*	2.20	2.17	2.33	*+1.4*
Decile ratio 50:10	5.439	5.218	6.452	*+4.3*	5.376	5.218	6.098	*+3.0*	7.00	7.10	6.39	*-1.4*
Average N per cross-section	***5002***	***4235***	***767***	***+18.1***	***5002***	***4235***	***767***	***+18.1***	***6790***	***5641***	***1149***	***+20.4***
Income mobility												
Quintile matrix mobility: Average jump	0.456	0.349	0.859	*+30.7*	0.457	0.349	0.836	*+31.0*	0.455	0.371	0.677	*+22.6*
Quintile matrix mobility: Normalized average jump	0.183	0.140	0.344	*+30.7*	0.183	0.140	0.335	*+30.7*	0.182	0.149	0.271	*+22.2*
Fields & Ok: Percentage income mobility	25.42	17.29	52.75	*+47.0*	24.74	17.29	48.30	*+43.1*	26.78	20.49	42.81	*+30.7*
Fields & Ok: Non-directional	0.348	0.273	0.641	*+27.5*	0.326	0.273	0.533	*+19.4*	0.348	0.316	0.447	*+10.1*
Shorrocks: Using Gini Coefficient	0.0279	0.0199	0.0562	*+40.3*	0.0254	0.0199	0.0444	*+27.8*	0.0302	0.0239	0.0472	*+26.4*
Average N per 2-wave balanced panel	***4389***	***3187***	***1202***	***+37.7***	***4389***	***3187***	***1202***	***+37.7***	***4928***	***3453***	***1475***	***+42.7***

* Germany in 2000 Euro; UK in 1996 GBP; Australia in 1989/90 AUD. Shaded cells indicate statistically significant deviations (for HILDA and SOEP based on a random group approach; in case of BHPS bootstrapping with 200 replications was applied).

Source: SOEP survey years 1992-2004; HILDA survey years 2001-2005; BHPS survey years 1991-2004.

4.2 Imputation and wage regressions

Obviously, there is convincing evidence for selectivity in INR on labor income questions in all three considered panel datasets. Concluding from this, it stands to reason that coefficients derived from (simple) wage regressions will be biased as well. Potential ways of dealing with such phenomena could be given by estimating a Heckman selection model (Heckman 1979) where the selection function would focus on the INR and the wage regression would be based only on the "observed" values. Even if this would allow for a perfect correction, there remains the problem of a loss in efficiency (caused by the loss in observations).

Following we will try to shed some light on this issue by comparing the results of fixed effects wage regressions based on the "observed" cases (column 1 in Table 3) to those based on the entire population including the imputed ones (columns 2). Finally, in column 3 we repeat the estimation from column (2), however we add a dummy-variable identifying the imputed observations. Table 3 gives those results separately for the three panels (as well as the alternative BHPS-imputation and separately for SOEP-sample F) controlling for usual covariates relating to human capital, socio-demographics, regional agglomeration, health status and (changes in) labor market participation over the last year. We refrain from including covariates focusing on the current employment situation in order to be able to include individuals currently not employed but who did receive earnings over the observation period (e.g., those who recently retired or who are unemployed).

In general, the findings based on "observed cases" are widely consistent for SOEP and BHPS with respect to direction and significance of most parameter estimates as well as with respect to the overall degree of explained variance (about 50%). Contrary results are given in case of the unemployment experience in the previous year, which is found to be significantly positive in the BHPS and significantly negative in SOEP.[11]

For HILDA, however, the specified model performs rather poor with an exceptionally low R-squared (approx. 22%) for such kind of an analysis.[12] Nevertheless, the estimated coefficients show into the expected direction, although sometimes lacking statistical significance.

[11] Results for the recent SOEP Subsample F are, by and large, in line with those of the entire SOEP sample.

[12] However, this finding is confirmed by Watson 2005.

More important for the sake our paper, however, is the effect of the additional consideration of imputed observations (see columns 2): In all three panels, this yields a pronounced reduction in the degree of explained variance: This decline is most prominent for HILDA with a reduction in R-Squared by about 23% to only 0.1723. Obviously, this effect is driven by the consideration of a group of less homogenous individuals following the above mentioned selectivity of INR. This may be exemplified by the fact, that “all” observations (see column 2) include significantly more self-employed in all three datasets. Other striking differences are given in case of the BHPS by under-representing individuals who retired, in SOEP and HILDA by those who experienced at least one month of unemployment in the previous year. Observations from the first waves of BHPS and HILDA are also underreported among the observed cases, while this is not the case in the more mature panel population in SOEP.[13] In addition, it is worthwhile considering whether the size of a given estimated coefficient varies once we include observations with imputed earnings. Bearing in mind a 95% confidence interval around the estimators, we find the effect of self-employment to significantly deviate in the two estimations (columns 1 and 2, respectively) in HILDA, while the strong effect of number of months in employment is even different in all three panels. Comparing such findings for the original BHPS imputation to the alternative row-and-column-imputation, it appears that the deviations between the coefficients derived from the observed cases and from the overall sample are not always perfectly in line. For example, while the age effect due to the original BHPS imputation does not change significantly, the alternative imputation method yields a reduced age effect. Although the two methods do not show any explicit contradictions, the coefficients for “remote area” and “disabled” decrease in statistical significance when using the Little & Su (1989) method. Such variations, however, may simply result from the selection of controls in the PMM regression model underlying the original BHPS imputation.

Finally, column 3 contains the repetition of the estimation in column 2, however, controlling for imputation status. The corresponding effect indicates that individuals with imputed incomes, ceteris paribus, earn significantly above average in SOEP and HILDA (about 5% to 6% more), while they earn 4% less in BHPS-data. However, changing the imputation strategy for the BHPS again yields a significant change in the “behavior” of the imputation flag: with “row-and-column” imputation we also find a positive effect of similar size (almost 5%).

[13] These figures are not reported in a table but are available from the authors on request.

For each set of panel data separately, we estimated quantile regressions (at the 25th, 50th and 75th percentiles), controlling for potential regression-to-the-mean effects emerging from the imputation process across the earnings distribution (see Appendix, Table A-1). Consistently for all estimations, the results for the imputation dummy is smallest at the 25th percentile, intermediate at the median, and finally, strongest at the 75th percentile. Using an appropriate F-test confirms that this effect is statistically different between the 25th and the 75th percentile.

For the BHPS, in line with the changing effect of the imputation flag when changing the imputation strategy in the fixed effects wage regressions, we find an almost identical effect across the UK earnings distribution. The Little & Su (1989) imputation method also produces a significant negative effect for imputed observations at the 25th percentile, which becomes insignificant at the median, and finally positive and significant at the income threshold to the upper quartile. We interpret these findings as an indication that the imputation techniques applied did not produce a relevant regression-to-the-mean effect. [14]

[14] In order to control for possible endogeneity, we excluded the covariates “disabled” and “retired” from the BHPS estimations. This resulted in a minor decrease in the R-squared, but there was little change in the remaining results except for the “unemployment” effect, which reversed sign and significance at the 25th and 75th percentile.

Table 3: Results from Fixed-Effects Wage Regression; Dependent Variable: Log Annual Labor Income

	Germany (SOEP)			Australia (HILDA)			UK (BHPS)			UK (BHPS) – "L&S"			Germany (SOEP) – Sample F		
	(1) Population: obs. cases	(2) Population: all cases	(3) Population: all cases	(1) Population: obs. cases	(2) Population: all cases	(3) Population: all cases	(1) Population: obs. cases	(2) Population: all cases	(3) Population: all cases	(1) Population: obs. cases	(2) Population: all cases	(3) Population: all cases	(1) Population: obs. cases	(2) Population: all cases	(3) Population: all cases
Age	0.050**	0.050**	0.049**	0.150**	0.160**	0.160**	0.064**	0.066**	0.066**	0.064**	0.047**	0.047**	0.097**	0.084**	0.084**
	(0.002)	(0.002)	(0.002)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.010)	(0.009)	(0.009)
Age squared	-0.000**	-0.000**	-0.000**	-0.002**	-0.002**	-0.002**	-0.001**	-0.001**	-0.001**	-0.001**	-0.001**	-0.001**	-0.001**	-0.001**	-0.001**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Female with kid(s)*	-0.162**	-0.159**	-0.159**	-0.337**	-0.325**	-0.325**	-0.336**	-0.339**	-0.339**	-0.336**	-0.300**	-0.300**	-0.054*	-0.057**	-0.056**
	(0.007)	(0.007)	(0.007)	(0.021)	(0.022)	(0.022)	(0.008)	(0.009)	(0.009)	(0.008)	(0.009)	(0.009)	(0.022)	(0.021)	(0.021)
Male with kid(s) *	0.030**	0.028**	0.028**	-0.051**	-0.052**	-0.052**	0.018*	0.021*	0.020*	0.018*	0.019*	0.020*	0.008	0.002	0.002
	(0.006)	(0.006)	(0.006)	(0.019)	(0.020)	(0.020)	(0.008)	(0.009)	(0.009)	(0.008)	(0.008)	(0.008)	(0.022)	(0.020)	(0.020)
Disability Status *	-0.006	-0.013*	-0.013*	-0.020+	-0.019+	-0.019+	-0.014+	-0.017*	-0.017*	-0.014+	-0.011	-0.010	-0.011	-0.016	-0.017
	(0.006)	(0.006)	(0.006)	(0.011)	(0.012)	(0.012)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.034)	(0.031)	(0.031)
Married *	-0.018	-0.012	-0.013	0.049**	0.049**	0.050**	0.031**	0.037**	0.037**	0.031**	0.029**	0.029**	0.004	-0.017	-0.017
	(0.011)	(0.011)	(0.011)	(0.016)	(0.016)	(0.016)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.024)	(0.022)	(0.022)
Metrop. area *	0.031**	0.036**	0.036**	0.041+	0.045+	0.045+	0.091**	0.106**	0.106**	0.091**	0.087**	0.088**	-0.030	-0.003	-0.001
	(0.011)	(0.011)	(0.011)	(0.024)	(0.024)	(0.024)	(0.016)	(0.017)	(0.017)	(0.016)	(0.017)	(0.017)	(0.044)	(0.044)	(0.044)
Remote area*	-0.001	0.000	0.000	0.008	-0.001	-0.001	0.035+	0.045*	0.045*	0.035+	0.024	0.024	0.018	0.029	0.028
	(0.007)	(0.007)	(0.007)	(0.025)	(0.026)	(0.026)	(0.020)	(0.022)	(0.022)	(0.020)	(0.021)	(0.021)	(0.028)	(0.026)	(0.026)
Intermed. education*	-0.020*	-0.016*	-0.016*	0.157**	0.108*	0.109*	-0.020	-0.018	-0.019	-0.020	-0.016	-0.016	0.039	0.009	0.007
	(0.008)	(0.008)	(0.008)	(0.043)	(0.043)	(0.043)	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)	(0.047)	(0.045)	(0.045)
Upper education*	0.013	0.010	0.008	0.544**	0.498**	0.500**	0.040**	0.054**	0.053**	0.040**	0.055**	0.056**	0.076	0.068	0.065
	(0.011)	(0.011)	(0.011)	(0.052)	(0.053)	(0.053)	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)	(0.062)	(0.059)	(0.059)
Highest educ. level*	0.307**	0.289**	0.287**	0.571**	0.528**	0.529**	0.247**	0.269**	0.269**	0.247**	0.215**	0.215**	0.405**	0.357**	0.354**
	(0.014)	(0.013)	(0.013)	(0.064)	(0.067)	(0.067)	(0.032)	(0.035)	(0.035)	(0.032)	(0.035)	(0.035)	(0.065)	(0.062)	(0.062)
East Germany*	-0.101**	-0.088**	-0.088**	---	---	---	---	---	---	---	---	---	-0.144*	-0.148**	-0.146**
	(0.017)	(0.017)	(0.017)	---	---	---	---	---	---	---	---	---	(0.058)	(0.054)	(0.054)
Self employed*	-0.019*	-0.007	-0.010	-0.069**	0.029*	0.018	-0.285**	-0.254**	-0.245**	-0.285**	-0.177**	-0.187**	-0.139**	-0.067**	-0.070**
	(0.009)	(0.008)	(0.008)	(0.016)	(0.015)	(0.015)	(0.010)	(0.009)	(0.009)	(0.010)	(0.009)	(0.009)	(0.027)	(0.024)	(0.024)
Became retired*	-0.020	-0.008	-0.010	0.205**	0.156**	0.157**	-0.244**	-0.260**	-0.259**	-0.244**	-0.213**	-0.214**	0.020	0.074+	0.073+
	(0.014)	(0.013)	(0.013)	(0.038)	(0.038)	(0.038)	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)	(0.049)	(0.043)	(0.043)
Left education *	-0.065**	-0.056**	-0.055**	-0.024*	-0.025*	-0.025*	-0.248**	-0.265**	-0.266**	-0.248**	-0.270**	-0.269**	-0.057*	-0.066**	-0.066**
	(0.007)	(0.007)	(0.007)	(0.011)	(0.011)	(0.011)	(0.017)	(0.019)	(0.019)	(0.017)	(0.018)	(0.018)	(0.022)	(0.021)	(0.021)
Unempl. (last year) *	-0.068**	-0.065**	-0.065**	-0.023**	-0.022**	-0.022**	0.141**	0.110**	0.110**	0.141**	0.075**	0.075**	-0.070**	-0.069**	-0.069**
	(0.001)	(0.001)	(0.001)	(0.004)	(0.004)	(0.004)	(0.009)	(0.010)	(0.010)	(0.009)	(0.010)	(0.010)	(0.003)	(0.003)	(0.003)
Months FT (last year)	0.118**	0.113**	0.113**	0.120**	0.104**	0.104**	0.186**	0.175**	0.174**	0.186**	0.145**	0.145**	0.108**	0.102**	0.102**
	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)
Months PT (last year)	0.066**	0.064**	0.065**	---	---	---	---	---	---	---	---	---	0.066**	0.064**	0.064**
	(0.001)	(0.001)	(0.001)	---	---	---	---	---	---	---	---	---	(0.002)	(0.002)	(0.002)

... contd. …

... *contd.* … Table 3

Imputed Labor Y*	**---**	**---**	**0.064****	---	---	**0.052****	-	-	**-0.042****			**0.047****	**---**	**---**	**0.042****
	---	---	(0.005)	---	---	(0.014)	-	-	(0.006)			(0.006)	---	---	(0.010)
Constant	7.515**	7.543**	7.533**	5.247**	5.240**	5.223**	5.959**	6.071**	6.093**	5.959**	6.805**	6.781**	6.291**	6.614**	6.605**
	(0.042)	(0.042)	(0.042)	(0.172)	(0.176)	(0.176)	(0.262)	(0.271)	(0.271)	(0.262)	(0.270)	(0.270)	(0.202)	(0.191)	(0.191)
Observations	119030	134337	134337	35661	38681	38681	62049	72729	72729	62049	72904	72904	20355	24392	24392
N (Persons)	24183	25487	25487	11097	11522	11522	10352	11137	11137	10352	11138	11138	6797	7448	7448
R-squared	**0.4869**	**0.4474**	**0.4484**	**0.2228**	**0.1723**	**0.1727**	**0.5169**	**0.4368**	**0.4372**	**0.5169**	**0.3661**	**0.3666**	**0.3849**	**0.3393**	**0.3400**

* indicates dummy variables.

Population: working age: 20-60 (Germany), 20-65 (Australia and UK)

Note: Time effects controlled, but not reported. Standard errors in parentheses; Significance level: + significant at 10%; * significant at 5%; ** significant at 1%.

Source: SOEP survey years 1992-2004; HILDA survey years 2001-2005; BHPS survey years 1991-2004.

5 Conclusion

This study deals with item non-response (INR) on annual labor income questions as a specific type of measurement error in three large panel surveys (the German SOEP, the British BHPS and the Australian HILDA). We provide empirical evidence for considerable cross-country variation with respect to incidence and selectivity of INR. Longitudinal imputation is the preferred way to handle INR in all three panels, with HILDA and SOEP using in principle the same strategy as suggested by Little & Su (1989), and the BHPS making use of a hot-deck regression approach.[15] Applying the approach used in HILDA and SOEP to the BHPS as well provides an empirical basis for robustness and sensitivity checks with respect to the choice of imputation technique.

The selectivity of item non-response and hence, the imputation of such missing observations, appears to have a significant impact on both, the distribution of earnings and earnings mobility. Results on *inequality* suggest that using observed values only, i.e., “case-wise deletion”, produces downward biased estimates. Likewise, analyses of earnings *mobility* based only on cases with observed information significantly understate income variability over time. Additionally, our analyses provide evidence for a positive inter-temporal correlation between item non-response and any kind of subsequent (item- and unit-) non-response, including permanent refusals.

Estimating wage regressions based on observed vs. all cases and controlling for imputation status, indicates that individuals with imputed incomes, *ceteris paribus*, earn significantly above average in SOEP and HILDA, while this relationship is negative using BHPS data.[16] However, using the same imputation technique for all three surveys produces remarkably similar BHPS results to those found for the other two surveys using the Little & Su imputation approach.

[15] The single imputation techniques currently applied in all three panels probably underestimate variance, and as such there may be demand for more complex variance estimation methods (e.g., jackknife estimators). However, the L&S imputation technique used in case of SOEP and HILDA may also be extended to a multiple imputation procedure by matching any non-respondent to more than one neighboring case (see Little & Su 1989: 415). Such a multiple imputation would more appropriately acknowledge the uncertainty embedded in the imputation as such.

[16] In any case, we find that selected estimated coefficients are subject to change when considering the entire population instead of the more homogenous population with observed income data.

Although a proper imputation is certainly preferable to simply ignoring cases with missing data by assuming MCAR mechanisms, thus reducing efficiency due to the reduced sample size, even this may yield biased results (see e.g., Nicoletti & Peracchi 2006). Nevertheless, the question of whether to use imputation for the treatment of missing values and if so, which imputation techniques and control variables are used, may depend heavily on the specific question under analysis. A particular type of imputation may fit the needs of cross-sectional data for the purposes of inequality analysis, for example, while the same imputed data may cause biased results in mobility analysis, especially if the imputation technique did not adequately consider the panel character of the underlying data. For the later type of analysis, one may consider imputing the value of a mobility index of interest instead of imputing two time-dependent income values (for t and t-1) as the basis for calculating the mobility index. However, this procedure may again yield conflicting results when compared to inequality analyses based on the same dataset.

Summing up, we are well aware that no definitive, "one-size-fits-all" imputation method exists, and our evidence underscores the possible variety of imputation methods that may be used by data providers for problems arising from item non-response. Data users should therefore not view the imputations produced by data providers as a panacea but should keep the potential shortcomings of the various methods in mind. When using imputed data, one should control for whether or not a particular piece of data was imputed.[17] A potentially superior, but also clearly more cumbersome approach would be to model INR and UNR in light of the individual research question (see, e.g., DeLuca & Peracchi 2007). This may, however, lead to more heterogeneous results among the various users of the same dataset.

The most important lesson to be learned from the present study is that the cross-national variation in INR presented here—variations in scope and selectivity, in strategies used, and consequences for prototypical labor income analyses—emphatically confirms the importance of further harmonizing the methods used to handle missing (income) data in (panel) surveys. Cross-national research relies crucially on homogeneously defined, functionally equivalent information. As such researchers should be aware of decisions made during the entire data production process: be it prior to data collection (e.g., regarding the wording of questions on

[17] The introduction of such imputation flags in our wage regressions clearly shows that—taking into account all other controls—the imputed observations are significantly "different" from those with truly observed information, whether because of different characteristics or because of the specific treatment of data in the imputation process.

annual income) or in the process of post-data collection treatment (e.g., when controlling for non-response through weighting and imputation). Given the importance of knowing all assumptions embedded in the imputation process, it is critical that survey data providers carefully document their imputation strategies and flag the imputed values in all microdata available to external users, thus allowing to differentiate imputed from observed data. In so doing, they will enable data users to conduct sensitivity tests to determine the impact of imputation, which—as shown in this paper—may be even more significant in the case of cross-national analyses. In the long run, this kind of methodological feedback from the user community may help to improve the quality of the imputation methods used by data collection, production, and dissemination agencies.

References

Biewen, M. (2001). Item non-response and inequality measurement: Evidence from the German earnings distribution. Allgemeines Statistisches Archiv 85 409-425.

Burkhauser, R.V., Butrica, B., Daly, M.C., Lillard, D.R. (2001). “The Cross-National Equivalent File: A product of cross-national research”. In Becker, I., Ott, N., Rolf, G. eds., Soziale Sicherung in einer dynamischen Gesellschaft. Festschrift für Richard Hauser zum 65. Geburtstag, Campus, Frankfurt/New York, 354-376.

Burton, J., Laurie, H. and Moon, N. (1999): Don't ask me nothin' about nothin', I just might tell you the truth: The interaction between unit non-response and item non-response”, Paper presented at the International Conference on Survey Nonresponse, Portland/USA, October 1999

Cameron, A. C. and P. K. Trivedi (2005): Micro-econometrics. Methods and Applications. New York: Cambridge University Press.

Canberra Group, Expert Group on Household Income Statistics: Final Report and Recommendations, Ottawa, 2001

DeLuca, G. and F. Peracchi (2007): A sample selection model for unit and item nonresponse in cross-sectional surveys, CEIS Tor Vergata - Research Paper Series, Vol. 33 , No. 99 , March 2007.

Dillman, D.A., Eltinge, J.L., Groves, R.M., Little, R.J.A. (2002). Survey nonresponse in design, data collection, and analysis. In: R.M. Groves, D.A. Dillman, J.L. Eltinge, R.J.A. Little, (eds.): Survey Nonresponse, Wiley, New York.

Ferber, Robert (1966): Item non-response in a consumer survey, Public Opinion Quarterly, 30: 399-415.

Frick, J. R. and Grabka, M. M. (2005): Item non-response on income questions in panel surveys: incidence, imputation and the impact on inequality and mobility. Allgemeines Statistisches Archiv, 89: 49-61

Gebel, M. and Pfeiffer, F. (2007), Educational Expansion and its Heterogeneous Returns for Wage Workers, ZEW Discussion Paper 07-10.

Groves R.M., and M.P. Couper (1998), Non-response in Household Interview Surveys, Wiley, New York.

Groves R.M., D.A. Dillman, J.L. Eltinge, and R.J.A. Little (2002), Survey Nonresponse, Wiley, New York.

Haisken-DeNew, J. and J. R. Frick (2005): DTC Desktop Companion to the German Socio-Economic Panel (SOEP).Version 8.0 - Dec 2005 Updated to Wave 21 (U).

Hanisch, J. U. (2005): Rounded responses to income questions. Allgemeines Statistisches Archiv, 89: 39-49.

Hawkes, D. and Plewis, I. (2006): Modelling Non-response in the National Child Development Study, Journal of the Royal Statistical Society (Series A), 169(3): 479-492.

Heckman J. (1979): Sample Selection Bias as a Specification Error, Econometrica, 47: 153—161.

Hill, D., Willis, R.J. (2001). Reducing Panel Attrition: A Search for Effective Policy Instruments. Journal of Human Resources 36: 416-438.

Institute for Social and Economic Research (ISER) (2002): ‘Weighting, imputation and sampling errors’, Chapter 5 in British Household Panel Survey (BHPS) User Documentation.

Jarvis, S. & Jenkins, S. P. (1998): How much income mobility is there in Britain?, Economic Journal 108(447), March 1998, 428-443

Kennickell, A. B. and McManus, D. (1994): Multiple Imputation of the 1983 and 1989 Waves of the SCF. American Statistical Association, 1994 Proceedings of the Section on Survey Research Methods, 523-528, Vol. I, American Statistical Association, Alexandria, VA.

Lee, E.; Hu, M.Y. and Toh, R.S. (2004): Respondent non-cooperation in surveys and diaries: an analysis of item non-response and panel attrition. International Journal of Market Research, 46(3): 311-326.

Little, R. J. A. & Su, H.-L. (1989):. Item Non-Response in Panel Surveys. In D. Kasprzyk, G. Duncan and M. P. Singh (Eds.), Panel Surveys. New York: John Wiley.

Little, R. J. A. (1988): “Missing-Data Adjustments in Large Surveys,” Journal of Business and Economic Statistics, 6, 287–296.

Little, R. J. A., & Rubin, D. B. (2002). Statistical analysis with missing data. 2. edition, Wiley: New Jersey.

Loosveldt, G., Pickery, J., Billiet, J. (2002). Item non-response as a predictor of unit non-response in a panel survey. Journal of Official Statistics 18: 545-557.

Nicoletti, C. and Peracchi, F. (2006). The effects of income imputation on microanalyses: evidence für the European Community Household Panel. Journal of the Royal Statistical Society Series A, 169 (3): 625-646.

Pickery, J., Loosveldt, G., Carlton, A. (2001). The Effects of Interviewer and Respondent Characteristics on Response Behavior in Panel Surveys. Sociological Methods & Research 29 509–523.

Riphahn, R. T. and Serfling, O. (2005): Item Non-response on Income and Wealth Questions. In: Empirical Economics 30(2), 521-538.

Rässler, S. and Riphahn, R. T. (2006): Survey Item Nonresponse and its Treatment, Allgemeines Statistisches Archiv, 90(1), 213-228.

Rubin, D. B. (1976). Inference with missing data. Biometrika, 63, 581-592.

Rubin, D. B. (1987): Multiple imputation for nonresponse in surveys. Wiley: New York.

Schräpler, J.-P. (2003). Gross Income Non-response in the German Socio-Economic Panel - Refusal or Don’t Know? Schmollers Jahrbuch 123: 109-124.

Schräpler, J.-P. (2003b). Income-Nonresponse in Household Surveys - An international comparison using the GSOEP and the BHPS. Paper presented at the “Workshop of Item non-response and data quality in panel studies”, University of Basel, 9-12 October 2003.

Schräpler, J.-P. (2004). Respondent Behavior in Panel Studies - A Case Study for Income-Nonresponse by means of the German Socio-Economic Panel (SOEP). Sociological Methods & Research, 33(1): 118-56.

Serfling, O. (2005): The interaction between Item, Questionnaire and Unit Nonresponse in the German SOEP. Schmollers Jahrbuch – Journal of Applied Social Science Studies, 125 :195-205.

Shoemaker, P. J.; Eichholz, M. Skewes, E. (2002): Item-Nonresponse: Distinguishing between Don’t know and Refuse. International Journal of Public Opinion Research 14: 193-201.

Spiess, M., Goebel, J. (2003). A comparison of different imputation strategies with respect to income related questions. Paper presented at the CHINTEX final conference "Harmonization of Surveys and Data Quality", Wiesbaden, 26 and 27 May 2003.

Starick, R. (2005): Imputation in Longitudinal Surveys: The Case of HILDA. Research Paper of the Australian Bureau of Statistics. ABS Catalogue no. 1352.0.55.075.

Taylor, M. F. (ed). with J. Brice, N. Buck and E. Prentice-Lane (2005) British Household Panel Survey User Manual Volume A: Introduction, Technical Report and Appendices. Colchester: University of Essex.

Wagner, G. G., Frick, J. R., Schupp, J. (2007): The German Socio-Economic Panel Study (SOEP) - Evolution, Scope and Enhancements. Schmoller's Jahrbuch - Journal of Applied Social Science Studies. 127(1): 139ff.

Watson, I. (2005). The earnings of casual employees: The problem of unobservables. Paper presented at the 2005 HILDA Survey Research Conference. The University of Melbourne, 29 September 2005

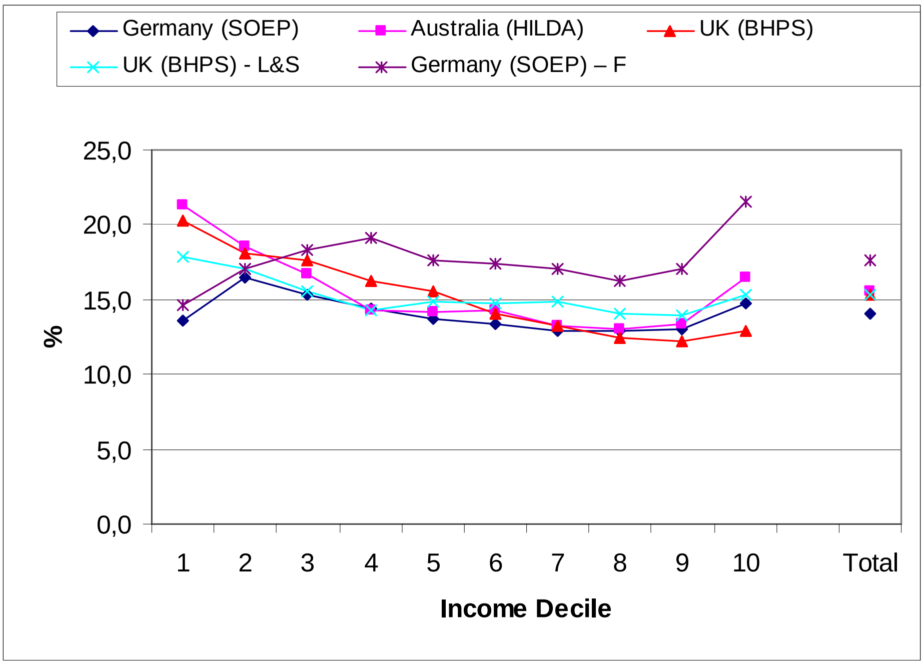
Watson, N. and Wooden, M. (2006): ‚Identifying factors affecting longitudinal survey response. Paper presented at the Methodology of Longitudinal Surveys (MOLS) International Conference, University of Essex, Colchester/UK, July 12-14, 2006.

Watson, N. (ed) (2005) HILDA User Manual – Release 3.0, Melbourne Institute of Applied Economic and Social Research, University of Melbourne. (http://www.melbourneinstitute.com/hilda/manual/userman_overview.html, accessed 28 February 2006)

Appendix:

Appendix A: Detailed Information on Item Non-Response and Imputation by Survey

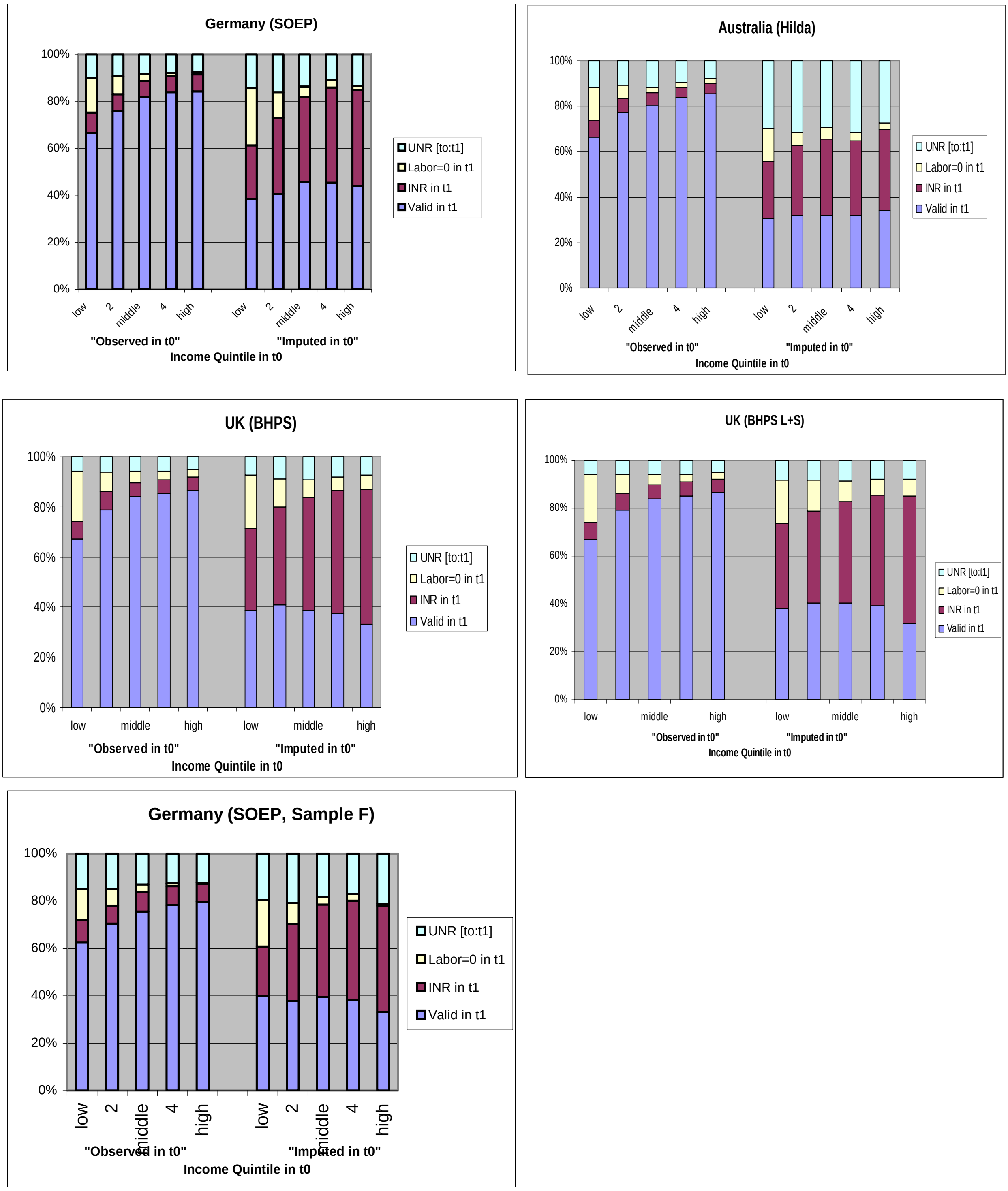
Figure A-1: Observations with INR on Labor Income by Deciles (in %) Based on Original and Alternative Imputation in the BHPS and Incorporating a "Fresh" Panel (SOEP-F)



Note: Contingent on the original and alternative imputation procedures as described in Section 3.

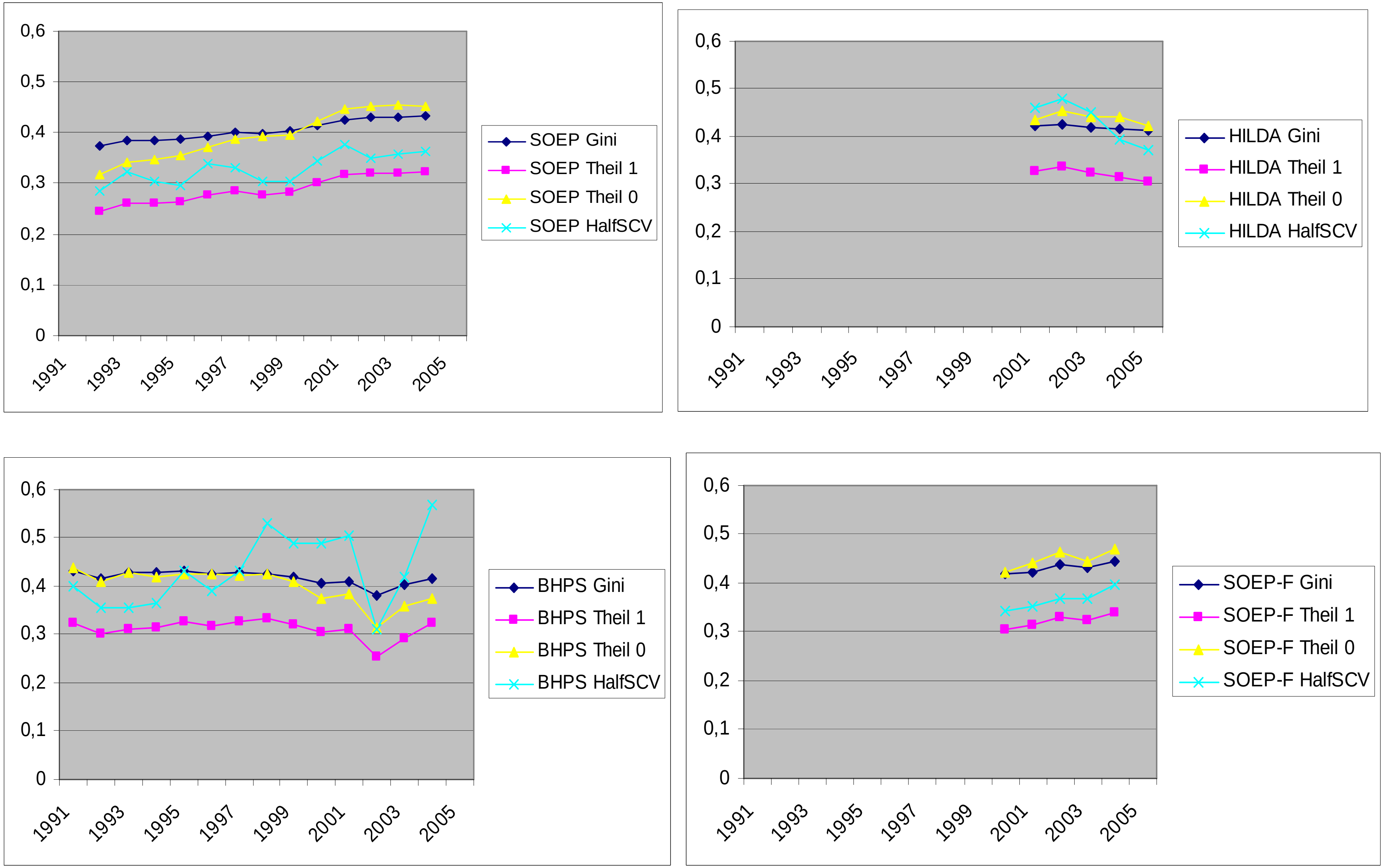
Source: SOEP survey years 1992-2004; HILDA survey years 2001-2005; BHPS survey years 1991-2004.

Figure A-2: Item Non-Response from a Longitudinal Perspective: The Case of Individual Labor Earnings



Source: SOEP survey years 1992-2004; HILDA survey years 2001-2005; BHPS survey years 1991-2004.

Figure A-3: Inequality of Individual Labor Earnings (based on observed cases only)



Source: **SOEP survey years 1992-2004; HILDA survey years 2001-2005; BHPS survey years 1991-2004.**

Table A-1: Results from Quantile Wage Regressions; Dependent Variable: Log Annual Labor Income (normalized)

	Germany (SOEP)			Australia (HILDA)			UK (BHPS)			UK (BHPS) – “L&S”			Germany (SOEP) – Sample F		
	p25	**p50**	**p75**	**p25**	**p50**	**p75**	**p25**	**p50**	**p75**	**p25**	**p50**	**p75**	**p25**	**p50**	**p75**
Age	0.0333**	0.034**	0.0324**	0.110**	0.083**	0.077**	0.074**	0.070**	0.074**	0.078**	0.071**	0.076**	0.035**	0.039**	0.043**
	(0.0012)	(0.0013)	(0.0013)	(0.004)	(0.002)	(0.002)	(0.003)	(0.002)	(0.002)	(0.004)	(0.002)	(0.002)	(0.004)	(0.003)	(0.002)
Age squared	-0.0003**	-0.000**	-0.0002**	-0.001**	-0.001**	-0.001**	-0.001**	-0.001**	-0.001**	-0.001**	-0.001**	-0.001**	-0.000**	-0.000**	-0.000**
	(0.0000)	(0.0000)	(0.0000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Female with kid(s)	-0.1754**	-0.152**	-0.1181**	-0.661**	-0.460**	-0.358**	-0.833**	-0.592**	-0.397**	-0.834**	-0.591**	-0.400**	-0.192**	-0.192**	-0.140**
	(0.0086)	(0.0058)	(0.0047)	(0.017)	(0.008)	(0.009)	(0.014)	(0.009)	(0.008)	(0.010)	(0.010)	(0.009)	(0.019)	(0.011)	(0.013)
Male with kid(s)	0.1377**	0.122**	0.1286**	0.157**	0.145**	0.160**	0.111**	0.087**	0.091**	0.116**	0.089**	0.092**	0.160**	0.142**	0.136**
	(0.0026)	(0.0030)	(0.0031)	(0.010)	(0.008)	(0.009)	(0.008)	(0.008)	(0.008)	(0.007)	(0.005)	(0.008)	(0.010)	(0.009)	(0.008)
Disability Status	-0.0080	-0.009	-0.0225**	-0.176**	-0.108**	-0.093**	-0.174**	-0.145**	-0.117**	-0.170**	-0.148**	-0.116**	-0.019	0.011	-0.013
	(0.0074)	(0.0059)	(0.0054)	(0.014)	(0.007)	(0.008)	(0.017)	(0.013)	(0.012)	(0.013)	(0.011)	(0.011)	(0.021)	(0.021)	(0.016)
Married	0.0025	0.074*	0.0085**	0.092**	0.067**	0.073**	0.322**	0.265**	0.220**	0.333**	0.272**	0.218**	0.010	0.020*	0.022*
	(0.0024)	(0.0032)	(0.0026)	(0.010)	(0.008)	(0.009)	(0.007)	(0.005)	(0.007)	(0.007)	(0.005)	(0.007)	(0.011)	(0.008)	(0.010)
Metrop. area	0.0505**	0.043**	0.0548**	0.068**	0.059**	0.064**	0.140**	0.145**	0.148**	0.141**	0.148**	0.152**	0.040*	0.036**	0.042**
	(0.0066)	(0.0041)	(0.0046)	(0.011)	(0.006)	(0.008)	(0.009)	(0.005)	(0.006)	(0.007)	(0.007)	(0.006)	(0.016)	(0.011)	(0.011)
Remote area	-0.0325**	-0.029**	-0.0202**	-0.075**	-0.056**	-0.024*	-0.027**	-0.034**	-0.051**	-0.031**	-0.042**	-0.053**	-0.039**	-0.033**	-0.020+
	(0.0027)	(0.0028)	(0.0033)	(0.018)	(0.012)	(0.010)	(0.010)	(0.006)	(0.006)	(0.010)	(0.009)	(0.009)	(0.008)	(0.007)	(0.010)
Educational level	0.1573**	0.174**	0.1878**	0.194**	0.188**	0.183**	0.226**	0.233**	0.234**	0.234**	0.238**	0.238**	0.170**	0.187**	0.199**
	(0.0030)	(0.0014)	(0.0014)	(0.004)	(0.003)	(0.004)	(0.004)	(0.004)	(0.003)	(0.005)	(0.003)	(0.002)	(0.005)	(0.003)	(0.003)
East Germany	-0.3443**	-0.331**	-0.3342**	---	---	---	---	---	---	---	---	---	-0.308**	-0.306**	-0.286**
	(0.0045)	(0.0032)	(0.0038)	---	---	---	---	---	---	---	---	---	(0.009)	(0.008)	(0.011)
Self employed	-0.2093**	-0.017*	0.1905**	-0.365**	-0.125**	0.018	-0.396**	-0.202**	-0.017	-0.350**	-0.187**	0.002	-0.196**	0.001	0.183**
	(0.0119)	(0.0069)	(0.0109)	(0.028)	(0.013)	(0.014)	(0.018)	(0.013)	(0.016)	(0.017)	(0.014)	(0.013)	(0.027)	(0.017)	(0.020)
Became retired	-0.1925**	-0.132**	-0.0612**	-0.016	0.154+	0.075	-0.121**	-0.092**	-0.025	-0.099**	-0.065**	-0.004	-0.442**	-0.602**	-0.190**
	(0.0143)	(0.0175)	(0.0114)	(0.134)	(0.085)	(0.083)	(0.038)	(0.019)	(0.021)	(0.025)	(0.016)	(0.016)	(0.064)	(0.080)	(0.058)
Left education	-0.0960**	-0.102**	-0.0871**	-0.071**	-0.051**	-0.064**	-0.280**	-0.228**	-0.265**	-0.268**	-0.267**	-0.288**	-0.156**	-0.190**	-0.156**
	(0.0125)	(0.0102)	(0.0088)	(0.013)	(0.011)	(0.009)	(0.036)	(0.023)	(0.026)	(0.047)	(0.032)	(0.031)	(0.034)	(0.032)	(0.017)
Months UE (last yr)	-0.0525**	-0.081**	-0.0917**	-0.033**	-0.030**	-0.049**	0.118**	0.082**	-0.012	0.095**	0.054*	-0.051**	-0.063**	-0.098**	-0.112**
	(0.0017)	(0.0012)	(0.0013)	(0.013)	(0.010)	(0.009)	(0.015)	(0.015)	(0.012)	(0.022)	(0.022)	(0.017)	(0.004)	(0.004)	(0.004)
Months FT (last yr)	0.1721**	0.120**	0.0915**	0.229**	0.176**	0.101**	0.226**	0.200**	0.162**	0.216**	0.185**	0.139**	0.175**	0.117**	0.084**
	(0.0015)	(0.0012)	(0.0007)	(0.007)	(0.006)	(0.006)	(0.003)	(0.003)	(0.002)	(0.002)	(0.003)	(0.002)	(0.003)	(0.003)	(0.003)
Months PT (last yr)	0.0374**	0.039**	0.0288**	---	---	---	---	---	---	---	---	---	0.039**	0.032**	0.019**
	(0.0016)	(0.0012)	(0.0008)	---	---	---	---	---	---	---	---	---	(0.002)	(0.003)	(0.002)
Imputed Income	**-0.0274****	**0.004**	**0.0457****	**-0.195****	**-0.017**	**0.078****	**-0.119****	**-0.076****	**-0.036****	**-0.139****	**-0.004**	**0.066****	**-0.016**	**0.013+**	**0.052****
	(0.0068)	(0.0054)	(0.0041)	(0.030)	(0.013)	(0.021)	(0.010)	(0.006)	(0.007)	(0.012)	(0.010)	(0.008)	(0.012)	(0.008)	(0.009)
Constant	1.4541**	2.155**	2.6652**	-0.960**	0.446**	1.652**	0.532**	1.141**	1.720**	0.537**	1.251**	1.935**	1.239**	1.996**	2.447**
	(0.0327)	(0.0274)	(0.0284)	(0.115)	(0.086)	(0.071)	(0.055)	(0.043)	(0.042)	(0.073)	(0.055)	(0.043)	(0.084)	(0.053)	(0.045)

contd.

contd. … Table A-1

Observations	139351			38681			72729			72904			25634		
R-squared	0.477	0.395	0.349	0.264	0.208	0.168	0.331	0.279	0.243	.307	.256	.222	0.470	0.396	0.345
Test on significant differences of imputation effect between the 25th and 75th percentile:															
	F(1,139321) = 94.41 Prob > F = 0.0000			F(1, 38661) = 69.15 Prob > F = 0.0000			F(1, 72700) = 46.77 Prob > F = 0.0000			F(1, 72875) = 241.16 Prob > F = 0.0000			F(1, 25612) = 27.19 Prob > F = 0.0000		

Population of working age: 20-60 (Germany), 20-65 (Australia and UK)

Note: Time effects controlled, but not reported. Standard errors in parentheses; Significance level: + significant at 10%; * significant at 5%; ** significant at 1%.

Source: SOEP survey years 1992-2004; HILDA survey years 2001-2005; BHPS survey years 1991-2004.

Appendix B: Exact Wording of Earnings-Related Questions in Original Survey Instruments

BHPS:

The last time you were paid, what was your gross pay - that is including any overtime, bonuses, commission, tips or tax refund, but before any deductions for tax, national insurance or pension contributions, union dues and so on?

IF `DON'T KNOW / CAN'T REMEMBER' PROBE: `Can you give me an approximate amount?'
ENTER TO NEAREST £: ASK E21 ***IPAYGL***
Don't know................ 8 **GO TO E22**
Refused .... 9 **GO TO E31 (page 43)**

RESPONDENT TO CHECK PAY SLIP IF POSSIBLE

HILDA:

F19 ***Last financial year*****, what was your total wage and salary income from** ***all*** **jobs** ***before*** **tax or anything else was deducted?**
Do not include income from businesses. This should be gathered at F24, rather than here.
Enter **annual** amount
(whole $) $____ → **F22**
Don't know.................................999999 → F20

F22 During the last financial year did you, at any time:
work in your own business or farm; or were a silent partner in a partnership; or were a beneficiary of a trust (excluding those that are used just for investment purposes)?
Yes..1
No... 2 → F28a

F24 Excluding dividends, in the last financial year, what was ***your*** **total income from wages and salary from these incorporated businesses** ***before*** **income tax was deducted? Please exclude wages and salary already reported.**
This includes trusts from F22
Enter amount (whole $) $
Recorded elsewhere.................................9999998
Don't know..9999999

F26a In the last financial year, did you have any unincorporated businesses?
Yes..1
No... 2 → F28a

Note: respondents cannot answer NO to both F26a and F23.
If they do, query.

F26b What was your total share of profit or loss from your unincorporated businesses or farms before income tax but after deducting business expenses in the last financial year?

Enter amount (whole $) ……. → F27
Don't know 9999999→ F28a

SOEP:

Q76. We have already asked for your current income. In addition, please state what sources of income you received in the past calendar year 2001, independent of whether the income was received all year or only in certain months. Look over the list of income sources and check all that apply. For all sources that apply please indicate how many months you received this income in 2001 and how much this was on average per month.
(Please state the gross amount which means not including deductions for taxes or social security).

	Source of income in 2001	**Received Months in 2001**	**Gross amount per month EURO**
Wages or salary as employee (including wages for training, "Vorruhestand", wages for sick time ("Lohnfortzahlung")			
Income from self-employment, free-lance work			
Additional employment			
Pay for compulsory military service, community service in place of military service ("Zivildienst")			

Q77. Did you receive any of the following additional payments from your employer last year (2001)? If yes, please state the gross amount.
13th month salary .... in total EURO
14th month salary .... in total EURO
Additional Christmas bonus ....in total EURO
Vacation pay .... in total EURO
Profit-sharing, premiums, bonuses .... in total EURO
Other .... in total EURO
No, I received none of these ....