Comparing estimation strategies in the presence of panel attrition. Empirical results based on the ECHP[*]

Andreas Behr[**]

JEL: I 22, J 31, C 8

Keywords: Panel attrition, earnings determinants, ECHP

Abstract

Since attrition in the European Community Household Panel (ECHP) has cumulated to a considerable extent, there is concern that attrition biases empirical analysis. In this paper we compare empirically the performance of four different estimation strategies in the presence of panel attrition. The example we analyze is the estimation of earnings equations. By splitting the completely observed sample according to the response behaviour of the following wave, we assess empirically the bias of an un-weighted, an inverse probability weighted, a Heckman and a matching estimator through bootstrap methods.

Our findings lead us to several conclusions. First, for the example of Mincerian earnings equations, attrition is no matter of great concern when using the ECHP data. Second, the three different estimation strategies, which correct for attrition based on estimated response probabilities, reduce the number of significantly biased parameters. Third, the correction strategies strongly increase the variance of the estimates through relying on estimated response probabilities and increase the relative bias. Hence, the reduction of significant biases is rather due to increased variance than due to lower biases. This result is confirmed when comparing the mean square error of the different estimation techniques. Therefore for the estimation of income equations the uncorrected estimation based on respondents is suggested as the superior estimation strategy.

[*] This work was partly financed by the European Commission under the 5th framework research programme in the Information Society Technology Programm (IST), Contract Number IST-1999-11101.

[**] PD Dr. Andreas Behr, Institut für Ökonometrie und Wirtschaftsstatistik, Wilhelms-Universität Münster, Am Stadtgraben 9, 48143 Münster.

1. Introduction

In this empirical paper we explore the performance of four different estimation strategies in the presence of panel attrition. The empirical application is the estimation of earnings-equations for male employees based on the European Community Household Panel (ECHP). The income equation we use for the empirical analysis is of Mincerian type based on the human capital approach. The considerable extent of panel attrition in the ECHP is documented in detail by Nicoletti and Peracchi (2002) and Behr, Bellgardt and Rendtel (2002). There is some concern that the extent of panel attrition is biasing the results of empirical socioeconomic analysis.

There are several estimation strategies suggested in the presence of attrition, ranging from the simplest alternative of OLS based on respondents only to elaborate estimators as the Heckman correction, which rest strongly on assumptions perhaps not present in the data. We assess empirically the bias of an un-weighted, an inverse probability weighted, a Heckman- and a matching estimator.

The effects of attrition on these estimators will be analyzed by transferring the observed attrition behaviour in the subsequent wave to the wave under consideration, which facilitates the comparison of the estimated income process for attriters and respondents as well as respondents only. This method has been used by Fitzgerald, Gottschalk and Moffitt to analyze the effect of attrition in the Michigan Panel Study of Income Dynamics (1998). For Finland this method has been validated through the use of register data available for respondents as well as attriters (Rendtel 2004).

In the following section we give a brief description of the extent of attrition in the ECHP and estimate cross sectional response probabilities by logit models. Section 3 contains the empirical results for the estimation strategies under consideration. Section 4 concludes.

2. Extent and determinants of panel attrition in the ECHP-UDB[1]

The first wave of the ECHP in 1994 covered about 130,000 individuals above 16 years living in about 60,000 households. In the first wave 12 countries participated: Belgium, Denmark, France, Germany, Greece, Ireland, Italy, Luxembourg, The Netherlands, Portugal Spain and UK. While Austria took part from the second wave on in 1995, Finland entered the panel in 1996 and Sweden in 1997.[2]

In the following we concentrate on active males. A detailed description of participation patterns based on the household as the relevant unit is given by Peracchi (2002). At the individual level the attrition in the ECHP is studied by Nicoletti and Peracci (2002) and Behr, Bellgardt and Rendtel (2002).

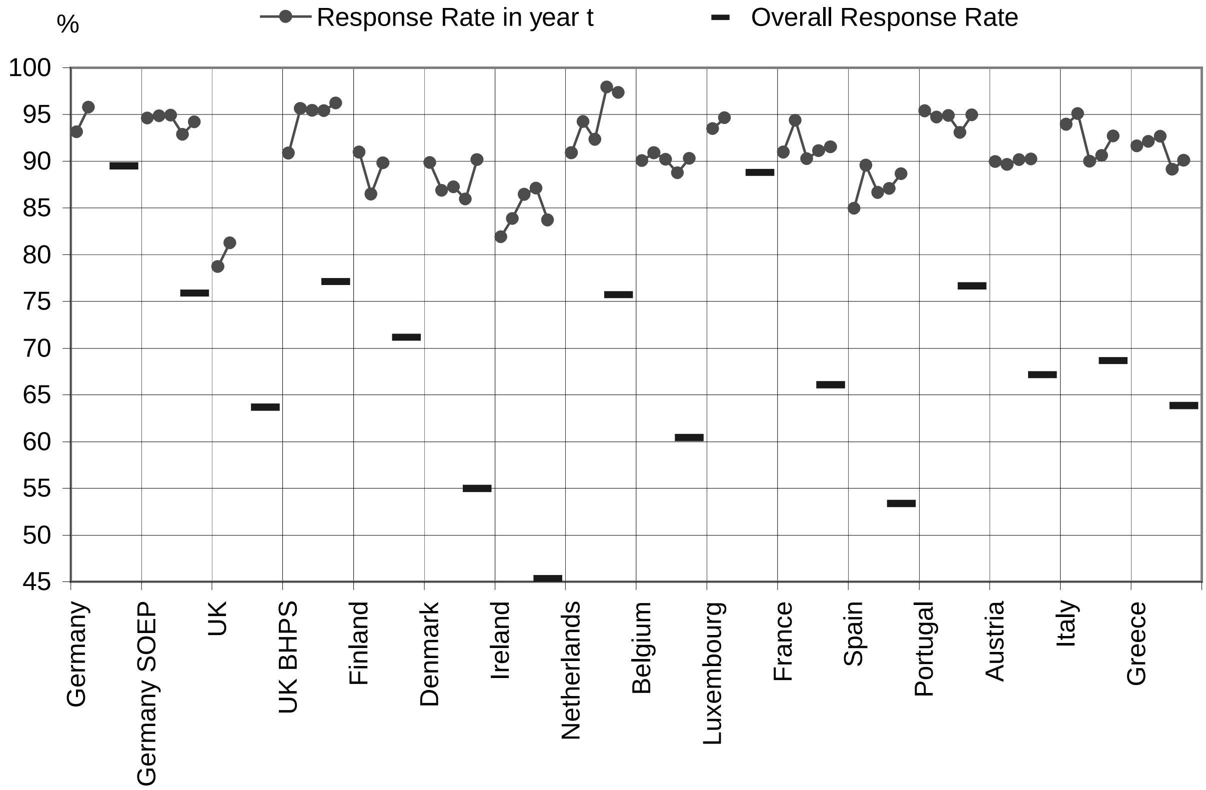
In the following figure we display the response rates in wave 2 up to wave 6 as well as the overall response rate in the latest wave for active males.[3]

1 The analysis is based on the 2002 release of the ECHP-User Data Base (UDB).

2 Because we assume effects of the number of waves being more important compared to effects of a given year, all data are ordered by country and wave. This means that data of wave 1 will include mainly data from 1994, but also from 1995 (Austria), 1996 (Finland). Because for Sweden only one wave is available (1997), we do not include Sweden in the following analysis.

3 For all persons the response rates resemble. See Behr et al. (2002)

Fig. 1: Response rates across countries for wave 2 to wave 6, active males



The horizontal bar indicates the ratio of respondents in the last wave to respondents in wave 1 (overall response rate). Especially for the overall response rate we find considerable differences across countries resulting from cumulated attrition rate differences in the individual years. The ECHP is strongest affected by attrition in Ireland where the remaining share of respondents dropped to 46% but also strongly in Spain and Denmark. High response rates were attained in Germany, the UK-BHPS (which started already in 1991), Luxembourg and Portugal. The figure shows no clear tendency in the response rates to rise or fall across all countries.

Next we estimate response probabilities using a logit model. Hence, we assume that the latent variable R_t^* is distributed according to the logistic distribution conditional on its expected value, which we model as a linear function of a set of variables X_{t-1} and a set variables V_t. While X is observable only prior to the attrition period, e.g. income and employment status, the set of additional (field-) variables V is observable in the period of attrition. Field variables are known to be rather important for attrition behaviour, in particular information about a move in the attrition period and a change of interviewer tend to increase the probability to attrit. Both variables were found highly significant in empirical studies, see e.g. Rendtel (1995), Behr, Bellgardt and Rendtel (2002)).

In the logistic model it is assumed that the observed response R takes the value of 1, if the logistic error term is below 0.

$$R_t^* = \gamma_1' X_{t-1} + \gamma_2' V_t + \delta_t \qquad (1)$$

$$R_t = \begin{cases} 1 & if \quad R_t^* = \gamma_1' X_{t-1} + \gamma_2' V_t + \delta_t > 0 \\ 0 & else \end{cases} \tag{2}$$

We use the following variables for explaining the attrition behaviour in the logit analysis:[4]

- log-earnings (ln *w*), lagged one period

- *tenure*, lagged one period

- *age* as well as age raised to the power of two (age^2), lagged one period

- dummy variable *edu*1 less than second stage education (second level education we use as the base category), lagged one period

- dummy variable *edu*3 to indicate third level education, lagged one period

- variable *cumemp* denotes the cumulated unemployment time in month, lagged one period

- dummy variable *married* indicates if the person is married (all other marital stati (divorced, widowed, not married) we use as the base category), lagged one period

- move of a household in the wave under analysis[5]

- interviewer change (if available) in the wave under analysis

To explain the response behaviour, all explanatory variables contained in the income equation are included in our logit analysis. The logit equation is additionally supplemented by the lagged log-earnings and the contemporaneous dummy variables indicating whether the person has moved and whether the interviewer changed.

---

[4] In our analysis we only consider variables that vary at the individual level. For an analysis including country characteristics and further information of the data collecting process, like whether the interview was by phone or personal interview, see Nicoletti/Peracchi (2002).

[5] These field-variables, move of the household and change of the interviewer were not available for all countries in the ECHP-User Data Base. This indicated by a missing value in the table (.). For some waves the inclusion of the variables indicating a move and the change of the interviewer caused numerical problems in the estimation procedure due to almost perfect dependence of variables (attrition, moved and change of interviewer). In these cases the variables moved and change of interviewer were dropped.

Table 1: Logit results by country and by wave

	wave	n	lnw	Tenure	Age	Age2	Education - third level	less than 2ndstage of	Unempl.time	Married	HH has moved	Interviewer change
Germany	1	1640								+	.	.
	2	1505							-			
Germany SOEP	1	1990			-	+		-			-	
	2	1923						-			-	
	3	1890									-	-
	4	1671								+	-	-
	5	1593	+								-	
United-Kingdom	1	1800							-	+	.	.
	2	1260		+							.	.
United-Kingdom BHPS	1	1724	+									
	2	1665			+							
	3	1627										
	4	1612			+							
	5	1650	-								.	.
Finland	3	1110									.	.
	4	1067			+	-		-			.	.
	5	1069									.	.
Denmark	1	1199								+	.	.
	2	1101										-
	3	1006	+									-
	4	938						-				-
	5	870										-
Ireland	1	1286								+	.	.
	2	1037									-	-
	3	849									-	-
	4	791									-	-
	5	778									-	-
The Netherlands	1	1446								+	.	.
	2	1365								+	.	.
	3	1323			+	-				+	.	.
	4	1269						-			.	.
	5	1314					+		-		.	.
Belgium	1	401									-	
	2	681			+	-						-
	3	632			+	-			-			-
	4	588									-	-
	5	640									-	-
Luxembourg	1	360						+			.	.
	2	328								+	.	.
France	1	1773			+	-						
	2	1331	-								-	
	3	1316		+								
	4	1656									-	-
	5	1099	-	+	+						.	.
Spain	1	2058							-	+	.	.
	2	1861								+	.	.
	3	1727								+	.	.
	4	1659			-	+				+	.	.
	5	1610								+	.	.
Portugal	1	1425						+			.	.
	2	1385	-							+	.	.
	3	1382	-	+							.	.
	4	1367		+							.	.
	5	1354									.	.
Austria	2	1185									.	.
	3	1057			-					+	.	.
	4	952								+	.	.
	5	869							-	+	.	.
Italy	1	1656									-	-
	2	1589			-							-
	3	1453							+		.	.
	4	1456		+							-	-
	5	1477									-	-
Greece	1	1194	-					+		+	.	.
	2	1039									.	.
	3	953									.	.
	4	914						+			.	.
	5	841						+			.	.

+ pos. sign. at 5%-level, - neg. sign. at 5%-level

To ease the readability and to allow for an overview of the results of the 68 logistic regressions, we present only summarized information in table 1.[6] The table contains for each country and each wave the number of observations (*n*) and the information whether the covariate was significant according to a *t*-test. We note significant positive and negative parameters at the five percent level by (+) and (-).

The overall finding is that the majority of parameters are insignificant. For log-wage we find a mixed picture with different directions of influence. While in the northern countries log wage exerts a positive influence, the opposite effect is present in southern European countries. The altogether rather weak influence of wage on the response probability could be seen as indicating a case of missing at random (MAR). But, first we include the log wage lagged one period and second, we should expect changes rather than levels of wage to influence the respond behaviour. Age and the age raised to the power of two exerts only in few waves significant influence. The finding of different effects for south and north European countries holds also for educational levels. While lower education reduces the respond probabilities in northern countries[7] the opposite is found for southern countries. Persons who gained the highest level of education show no significant tendency to respond differently.[8]

Cumulated unemployment time is found to decrease the response probability. The marital status was reflected using two different categories only. Compared to the reference category (not married) we find that married males have an almost unambiguous tendency to higher response rates.[9]

While the influence of lagged personal variables is rather weak, we find the field variables - move of the household and interviewer change - to exert a significant negative influence on the response behaviour, almost whenever included in the logistic regression. This result for active males confirms strongly the findings for all persons of Rendtel (1995) for the German SOEP and of Behr, Bellgardt and Rendtel (2002) for the ECHP.

---

6 The detailed results are given in the appendix. Beside the parameters the table gives some model information. The overall chi-square-test of model significance clearly rejects the hypothesis of no combined explanatory power of the model except four United Kingdom (BHPS), Finland and Austria. McFadden's Likelihood-Ratio-Criterion (LRI) has a rather low value about 0.04 in most countries. The R^2 (R2MZ) suggested by McKelvey and Zavoina (1975) has a slightly higher value in average (0.07). Both measures, which are defined between the range of 0 and 1, 1 in the case of perfect model fit, are indicating an unsatisfying model fit. This is somewhat in contrast to the Chi-Square test proposed by Hosmer and Lemeshow (1980), which indicates according to the high *p*-value, a satisfactorily model fit.

7 Fitzgerald/Gottschalk/Moffitt (1998) found the same pattern of decreasing risk of attrition with higher educational levels in the PSID.

8 This tendency was also found in the PSID, see e.g. Lillard and Panis (1998).

9 The finding of higher response probabilities for married persons corresponds to the findings of Lillard and Panis (1998) for PSID.

3. Empirical results and attrition effects

3.1. The empirical findings ignoring attrition

We restrict our analysis to males and to earnings of labour in the context of the human capital approach[10] and estimate cross-sectional log-earnings equations of the following type by country (j) and by wave (t):

$$\begin{aligned} \ln w_i &= \beta_0 + \beta_1 \cdot tenure_i + \beta_2 \cdot age_i + \beta_3 \cdot age_i^2 + \beta_4 \cdot edu3_i \\ &\quad + \beta_5 \cdot edu1_i + \beta_6 cumemp_i + \beta_7 married_i + u_i \end{aligned} \tag{3}$$

While after schooling job experience is the sum of time worked altogether, the job *tenure* measures the time the person is working within the same firm. Hence, *tenure* might be closer related to firm-specific knowledge compared to job experience. Concerning education, we use second level education as the standard category and the dummy variables *edu*3 and *edu*1 to indicate third level education and less than second stage education respectively. The variable *cumemp* denotes the cumulated unemployment time in month and the dummy variable *married* indicates whether the person is married while all other marital stati (divorced, widowed, not married) are the base category.

To ease readability, we give only summarized results in table 2. We indicate whether the effect was positive significant at the five (+) percent level (negative significant (-) respectively) by signs.[11]

We find an almost uniformly result concerning the sign and the level of significance for all countries and all waves. Tenure is significantly influencing the log wage in almost all countries and waves. There is evidence for a concave shape in the wage profile according to age. This is implied by the positive linear and negative quadratic age effect, which is evident in all countries. The positive effect of third level education as well as the negative effect of less than second stage education is highly significant in all countries and almost all waves, the third wave in The Netherlands being the only exception.[12] For all countries there is a strong negative effect on cumulated unemployment time present.

The uniformly negative effect of less than second stage, as well as the uniformly positive effect of third level education, shows considerable variation across countries. The premium for third level education is by far highest in Portugal, but also considerable in Luxembourg and France with about 40%. Males with less than second stage education suffer the highest losses compared to second stage education in Portugal and Austria.[13]

---

10 Historical functional approaches of the Ricardian type or Smith's compensatory principle to explain income differentials are not taken up in this analysis. See Mincer (1958, 1970, 1997).

11 Detailed results are given in the appendix.

12 As we mentioned in section 3.2, the measured effect will include schooling as well as ability effects. For an attempt to correct for a selectivity bias and to isolate the schooling effect in the case of three choices see Garen (1984).

13 For attempts to estimate social rates of return to schooling, which includes the difficult measure of public costs, see e.g. Hines/Tweeten/Redfern (1970) an Vaillancourt (1995).

Table 2: Log-earnings equations by country and by wave, summarized findings

Country	wave	n	Tenure	Age	Age2 /100	Third level education	Basic Education	Unempl.time	Married
Germany	1	1,640	+	+	-	+	-		+
	2	1,564	+	+	-	+	-	-	
Germany SOEP	1	2,261	+	+	-	+	-	-	+
	2	2,202	+	+	-	+	-	-	+
	3	2,108	+	+	-	+	-	-	+
	4	2,016	+	+	-	+	-	-	+
	5	1,914	+	+	-	+		-	+
United-Kingdom	1	1,800	+	+	-	+	-	-	+
	2	1,260	+	+	-	+	-	-	+
United-Kingdom BHPS	1	1,872	+	+	-	+	-	-	+
	2	1,738		+	-	+	-	-	+
	3	1,696	+	+	-	+	-	-	+
	4	1,680	+	+	-	+	-	-	+
	5	1,650		+	-	+		-	+
Finland	3	1,110	+	+	-	+	-	-	+
	4	1,067	+	+	-	+	-	-	
	5	1,069	+	+	-	+	-	-	
Denmark	1	1,199	+	+	-	+	-	-	
	2	1,113	+	+	-	+	-	-	
	3	1,019	+	+	-	+	-	-	
	4	942	+	+	-	+	-	-	
	5	871	+	+	-	+	-	-	+
Ireland	1	1,286	+	+	-	+	-	-	+
	2	1,052	+	+	-	+	-	-	+
	3	854	+	+	-	+	-	-	+
	4	799	+	+	-	+	-	-	+
	5	829	+	+	-	+	-	-	+
The Netherlands	1	1,446	+	+	-	+	-	-	+
	2	1,365	+	+	-	+	-	-	
	3	1,323	+	+	-	+		-	+
	4	1,269	+	+	-	+	-	-	+
	5	1,314		+	-			-	
Belgium	1	750	+	+	-	+	-	-	+
	2	688	+	+	-	+	-	-	+
	3	644	+	+	-	+	-	-	+
	4	593	+	+	-	+	-	-	+
	5	643	+	+		+	-		+
Luxembourg	1	360	+	+	-	+	-	-	
	2	328	+	+	-	+	-	-	
France	1	1,833	+	+	-	+	-	-	+
	2	1,718	+	+	-	+	-	-	+
	3	1,550	+	+	-	+	-	-	+
	4	1,242	+	+	-	+	-	-	+
	5	1,099	+	+	-	+	-	-	+
Spain	1	2,058	+	+	-	+	-	-	+
	2	1,861	+	+	-	+	-	-	+
	3	1,727	+	+	-	+	-	-	+
	4	1,659	+	+	-	+	-	-	+
	5	1,610	+	+	-	+	-	-	+
Portugal	1	1,425	+	+	-	+	-	-	+
	2	1,385	+	+	-	+	-	-	+
	3	1,382	+	+	-	+	-	-	+
	4	1,367	+	+	-	+	-	-	+
	5	1,354	+	+	-	+	-	-	+
Austria	2	1,185	+	+	-	+	-	-	
	3	1,057	+	+	-	+	-	-	+
	4	952	+	+	-	+	-	-	
	5	869		+	-	+	-	-	
Italy	1	1,657	+	+	-	+	-	-	+
	2	1,601	+	+	-	+	-	-	+
	3	1,453	+	+	-	+	-	-	+
	4	1,464	+	+	-	+	-	-	+
	5	1,482	+	+	-	+	-	-	+
Greece	1	1,194	+	+	-	+	-		+
	2	1,039	+	+	-	+	-	-	+
	3	953	+	+	-	+	-		+
	4	914	+	+	-	+	-	-	+
	5	841	+	+	-	+	-		+

+ pos. sign. at 5%-level, - neg. sign. at 5%-level

3.2. Analyzing the effect of attrition on the earnings equations

While in the preceding section we did not take into account a possible biasing effect of attrition on the income findings, we now analyze explicitly attrition effects. First, we analyze whether estimates for one period change significantly if the estimation includes only persons responding in the following wave, compared to estimates based on the full sample. Second, we apply three different estimation strategies to correct for an attrition bias and assess whether this methods improve the estimation results compared to the un-weighted approach.

In this section we apply four different estimation strategies: Using respondents only (OLS), inverse probability weighting (IPW), the estimator using the Heckman correction (Heck), and an imputation strategy (Imp). To assess the performance of the four different estimation strategies, we use the estimation results based on the full sample (respondents ***and*** attriters) as reference.

Hence, we consider the following estimators, subscript 0 refers to attriters, 1 to respondents:

$\beta_{0,1}$ the OLS-estimator, obtained using wave 2 respondents and attriters (*reference*)

β_1 the OLS-estimator, obtained using wave 2 respondents

$\beta_{1,IPW}$ the IPW-estimator, obtained using wave 2 respondents, inverse probability weighted

$\beta_{1,Heck}$ the estimator, obtained when including the Mills-ratio according to Heckman

$\beta_{1,Imp}$ the estimator, obtained duplicating the respondents that mimic attriters

3.2.1. The Inverse Probability Weighted Estimator (IPW)

In this section we describe the inverse response probability weighting approach, which potentially reduces the panel attrition bias (Robins, Rotnitzky and Zhao (1995), Neukirch (2002)). We assume that the individual log-earnings can be modelled as a linear equation containing an error term and that the response behaviour can be modelled by a logistic response equation:

Income equation: $$Y_t = X_t\beta + \varepsilon_t \quad if \quad R_t = 1 \tag{4}$$

Attrition equation: $$R_t^* = X_{t-1}\gamma_1 + V_t\gamma_2 + \delta_t \qquad \delta \sim \text{Logistic} \tag{5}$$

Observable Response: $$R_t = \begin{cases} 1 & if \quad R_t^* = X_{t-1}\gamma_1 + V_t\gamma_2 + \delta_t > 0 \\ 0 & else \end{cases} \tag{6}$$

Here Y is log-wage, R is the observed response variable and X contains explanatory variables common to both equations, while V contains additional variables considered as influencing the attrition behaviour only.

If the log-wage equation is estimated making use of available respondents (R = 1), only in the case of missing at random (MAR) the parameter vector of interest β will be estimated unbiased.[14] This means that one has to rely on the assumption that, given the set of explanatory variables, the missingness mechanism is independent of contemporaneous Y_t:

$$P\left(R_t = 1 \mid X_{t-1}, V_t, Y_{t-1}, Y_t\right) = P\left(R_t = 1 \mid X_{t-1}, V_t, Y_{t-1}\right). \tag{4}$$

This is equivalent to the desirable condition that $P\left(Y_t \mid X_t, R_t = 1\right) = P\left(Y_t \mid X_t\right)$ and hence estimates making use of respondents only will mirror the relation between Y and X for attriters as well as respondents.

If we denote the estimated response probabilities $\hat{\pi}$, and the diagonal matrix of estimated probabilities $\hat{\Pi}$, the IPW-estimator can be written as

$$\hat{\beta}_{IPW} = \left(X_t' \, \hat{\Pi}_{t+1}{}^{-1} X_t\right)^{-1} X_t' \, \hat{\Pi}_{t+1}{}^{-1} Y_t \tag{5}$$

with

$$prob(\hat{R}_{t+1} = 1) = \hat{\pi}_{t+1} = \frac{e^{X_t \gamma_1 + V_{t+1} \gamma_2}}{1 + e^{X_t \gamma_1 + V_{t+1} \gamma_2}} \tag{6}$$

The two-step procedure is rather intuitive. Because the observable sample contains respondents only, each available observation is weighted with the root of its inverse response probability. Because the observable observations will resemble the observations lost due to attrition the more the lower their response variability, the increased weight given to these observations (through dividing by the root of the low response probability) should improve the resemblance of the observable sample to the full sample.

Since the weights $\hat{\pi}$ are estimated and therefore contain random variation, this has to be taken care of when doing inference on the estimated coefficients of the income equation. Hence, we do not only use the potentially misleading standard *t*-statistics obtained from the income model, but rather carry out a non-parametric bootstrap simulation to assess the significance of the parameters.

3.2.2. The Heckman Estimator

A second strategy to correct for possible attrition bias is the estimator suggested by Heckman (1976, 1979). The basic idea is to take account of possible correlation between income and attrition error terms, which renders the income parameters biased. The estimation rests on the assumption that the unobservable error term in the unobservable attrition equation (11) is normal.

Income equation: $$Y_t = X_t \beta + \varepsilon_t \quad \text{if} \;\; R_t = 1 \qquad \varepsilon \sim N\left(0, \sigma_\varepsilon^2\right) \tag{10}$$

Attrition equation: $$R_t^* = X_{t-1}\gamma_1 + \delta_t \qquad \delta \sim N\left(0, \sigma_\delta^2\right) \tag{11}$$

Observable Response: $$R_t = \begin{cases} 1 & \text{if} \;\; R_t^* = X_{t-1}\gamma_1 + \delta_t > 0 \\ 0 & \text{else} \end{cases} \tag{12}$$

---

14 See Rendtel (1995, 2002) for a theoretical overview of panel attrition models.

Having obtained first step estimates of the observable attrition equation (12), the income equation is supplemented by the Mills-Ratio. If model assumptions hold, the regression coefficient corresponds to the correlation of the error terms of the income and attrition equation.

Supplemented income equation: $$Y_t = X_t\beta + \sigma_{\varepsilon\delta}\frac{\varphi(X_{t-1}\hat{\gamma}_1)}{\Phi(X_{t-1}\hat{\gamma}_1)} + \eta_t \quad if \;\; R_t = 1 \tag{13}$$

The coefficient vector $\hat{\beta}$ based on respondents will then be an unbiased estimator

$$E\left(Y_t \mid \beta, X_t, R_t = 1, \gamma\right) = X_t\beta + \sigma_{\varepsilon\delta}\frac{\varphi(X_{t-1}\gamma_1)}{\Phi(X_{t-1}\gamma_1)} \tag{14}$$

3.2.3. The Imputation-Estimator

The imputation strategy we use is based on Rubin (1973) and Rosenbaum and Rubin (1983). The basic idea is to replace the number of attriters by the same number of respondents who have similar attrition probabilities as the attriters. The error term of the unobservable attrition equation is assumed logistic and the estimated response probabilities are obtained through estimating the observable attrition equation (6):

$$prob(\hat{R}_t = 1) = \hat{\pi}_t = \frac{e^{X_{t-1}\,\hat{\gamma}_1 + V_t\,\hat{\gamma}_2}}{1 + e^{X_{t-1}\,\hat{\gamma}_1 + V_t\,\hat{\gamma}_2}} \tag{15}$$

Based on these probabilities, the data set is completed through duplicating *K* respondents where *K* is the number of attriters and *L* the number of respondents (*k*=1, 2, …, *K*), (*l*=1, 2, …, *L*). The replacement rule is as follows: chose respondents *l* that mimic mostly attriters according to $|\hat{\pi}_k - \hat{\pi}_l| = min$ for each k with $prob(\hat{R}_t = 1) = \hat{\pi}_t$. We duplicate respondents once at the most.

3.3. Assessing the significance of the bias

The estimate of the bias is

$$\hat{b}(\beta) = \hat{\beta}_{1,j} - \hat{\beta}_{0,1}, \tag{16}$$

where *j* indicates the different estimators under consideration based on respondents. As reference for all estimators we use the estimator $\hat{\beta}_{0,1}$ obtained using the full sample of respondents and attriters. To assess the significance of the estimated bias we test the hypothesis $b(\beta) = 0$ against the alternative $b(\beta) \neq 0$.

One way is making use of the asymptotic result that the covariance matrix of the difference (Σ_{diff}) between a consistent estimator under the null-hypothesis ($\hat{\beta}_{1,j}$) and an efficient estimator ($\hat{\beta}_{0,1}$) is given by the difference of the covariance matrix of the consistent estimator (Σ_{con}) and the covariance matrix of the efficient estimator (Σ_{eff}):

$$\Sigma_{diff} = \Sigma_{con} - \Sigma_{eff}\,. \tag{17}$$

The Hausman-test statistic is then calculated as:

$$t_{Haus} = \left(\hat{\beta}_{1,j} - \hat{\beta}_{0,1}\right)' \Sigma_{diff}^{-1} \left(\hat{\beta}_{1,j} - \hat{\beta}_{0,1}\right) \sim \chi_k^2 \tag{18}$$

In our analysis we face the problem, that empirically $\Sigma_{con}(\hat{\beta}_{1,j})$ sometimes smaller than $\Sigma_{eff}(\hat{\beta}_{0,1})$ in which case the Hausman-test statistic is not defined. A second reason for using an alternative to the Hausman-test is the fact that the IPW-estimators are based on estimated response probabilities. This additional source of variation is not taken into account by the Hausman-test. Hence, we apply alternatively a non-parametric bootstrap procedure to assess the significance of the bias. We draw 8,000 bootstrap replications with replacement (*B*=8,000) to simulate the distribution of the bias $\hat{b}_j(\beta) = \hat{\beta}_{1,j} - \hat{\beta}_{0,1}$ based on the realizations of $\hat{\beta}_{1,j}^*$ and $\hat{\beta}_{0,1}^*$, the bootstrap versions of $\hat{\beta}_{1,j}$ and $\hat{\beta}_{0,1}$. To assess the significance of the estimated bias, we use the 2.5% and 97.5% quintiles of the simulated bootstrap distributions. At this point we have to remember that the bias analysis is made possible through replacing $R_t^* = X_{t-1}\gamma_1 + \delta_t$ by $R_{t+1}^* = X_t\gamma_1 + \delta_{t+1}$. Therefore, we "observe" in period *t* respondents ($R_{t+1} = 1$) as well as attriters ($R_{t+1} = 0$).

3.4. The empirical results of the bias analysis

Due to the richness of our database, we estimate 30 cross sections and obtain 8 parameters in each cross section. Therefore, we have to evaluate 240 parameters for each estimation procedure and the need to summarize the results. Table 3 summarizes the results for the simple OLS estimator based on respondents, through indicating whether the estimated bias is significant (* at 10%, ** at 5% and *** at 1% level) or insignificant. Only 25 parameters are significantly biased at the five percent level according to the Hausman-test. This number is about twice the number of significant biases when applying a bootstrap test (13). The results of the bootstrap test for the uncorrected OLS estimation is given in table 4. Nevertheless, relative to the number of 240 parameters, the share of significantly biased parameters is small in either case.

Table 3: Hausman-significance pattern of ***biases*** ***,*** *OLS (β_1)*

Country	wave	n	Tenure	Age	Age2	Third level education	Basic education	Unempl. time	Married
Germany	2	1,505					*	*	
Germany SOEP	1	1,990	***						**
	2	1,923					*	**	
	3	1,890				*			
	4	1,671							
	5	1,593	/	***	***	/		/	***
United-Kingdom BHPS	1	1,724	/	/	/	/	/	/	/
	2	1,665							
	3	1,627						/	
	4	1,612							
Denmark	2	1,101			*				
	3	1,006	/	/	/	/	***	***	***
	4	0,938	***					**	
	5	0,870		**	**				
Ireland	2	1,037							
	3	0,849							
	4	0,791						*	
	5	0,778							
Belgium	1	0,401		*	*				
	2	0,681							
	3	0,632				***	***		*
	4	0,588	***	**	*			**	
	5	0,640	**						
France	1	1,773		**	***				
	3	1,331	/	/	/	***	*	/	/
	4	1,316				**	**	***	*
Italy	1	1,656		*	*		*		
	2	1,589							
	4	1,456							
	5	1,477							

*** pos. sign. at 1%-level, ** pos. sign. at 5%-level, * pos. sign. at 10%-level, / not available

Table 4: Bootstrap-Significance pattern of ***biases*** ***,*** *OLS (β_1)*

Country	wave	n	Tenure	Age	Age2	Third level education	Basic education	Unempl. time	Married
Germany	2	1,640							
Germany SOEP	1	1,564	+						
	2	2,261							
	3	2,202							
	4	2,108							
	5	2,016	-						
United-Kingdom BHPS	1	1,914						-	
	2	1,800							
	3	1,260							
	4	1,872							
Denmark	2	1,738							
	3	1,696							
	4	1,680	-						
	5	1,650							
Ireland	2	1,110							
	3	1,067							
	4	1,069						+	
	5	1,199							
Belgium	1	1,113							
	2	1,019						+	
	3	942					-		
	4	871							
	5	1,286							
France	1	1,052		+	-				
	3	854						+	
	4	799				+		+	
Italy	1	829							
	2	1,446							
	4	1,365						-	
	5	1,323							

+ pos. sign. at 5%-level, - neg. sign. at 5%-level

Table 5 displays the frequencies of significant biases of the four different estimation strategies under consideration. We find that the Heckman as well as the imputation procedure doing about equally well in reducing the bias. The number of 13 significant parameters in the case of the OLS is reduced to 1 and 0 respectively.

Table 5: Summary of bootstrap-significance pattern of ***biases***

	Respondents	IPW	Heckman	Imputation
pos. sign.	7	4	1	0
neg. sign.	6	8	0	0
significant	13	12	1	0
insignificant	227	228	239	240

But, as is often the case in empirical analysis, taking only the significance into account can be far misleading. For all three estimation procedures, which try to correct for attrition, the final estimation rests on the results of the first step estimation of the attrition equation. This first step regression introduces additional variation into the final estimation of the income equation. Hence, there is the possibility that the reduced number of significant biases is rather due to increased variation in the estimators and therefore in the variance of the bias than due to the reduced bias. One way to disentangle these two sources of the reduced number of significant biases is to regard the relative bias of the estimates. Since we are ultimately interested in coming close to the unbiased estimate, we regard the relative bias more appropriate as the significance to assess the performance of the estimators. Table 6 contains summary findings for the extent of the relative biases in the case of the uncorrected OLS estimates.

Table 6: Pattern of ***relative biases ,*** *OLS (β_1)*

Country	wave	n	Tenure	Age	Age2	Third level education	Basic education	Unempl. time	Married
Germany	2	1,505					+	++	
Germany SOEP	1	1,990							- -
	2	1,923					++		+
	3	1,890							
	4	1,671					++		-
	5	1,593	-	-	-				-
United-Kingdom BHPS	1	1,724							-
	2	1,665							
	3	1,627							
	4	1,612						+	
Denmark	2	1,101	- -					+	++
	3	1,006	-				- -	- -	-
	4	0,938	- -	-	-		-	-	++
	5	0,870	++	- -	- -			++	
Ireland	2	1,037	-	-	-	+		- -	
	3	0,849					-		
	4	0,791					- -	- -	+
	5	0,778	+		-			+	+
Belgium	1	0,401	+	- -	- -		+		++
	2	0,681						-	
	3	0,632			-	-	++	+	- -
	4	0,588	- -	++	++		++	- -	- -
	5	0,640	++	+	++		++	- -	-
France	1	1,773	-	++	++				
	3	1,331	+			-		-	
	4	1,316						- -	+
Italy	1	1,656		-	-		+		
	2	1,589							
	4	1,456						+	-
	5	1,477			-				

++ pos. bias >10%, + pos. bias >10%, -- neg. bias <-10%, - neg. bias <5%,

Table 7 shows the frequencies of the relative biases for the four different estimation strategies under consideration.

Table 7: Summarizing the results of the pattern of ***relative bias***

	Respondents	IPW	Heckman	Matching
|rel. bias|<5%	130	139	93	117
5%<|rel. bias|<10%	45	32	49	51
|rel. bias|>10%	35	39	68	42

The main finding is that the Heckman correction considerably increases the relative bias. We find that strong relative biases of 10% or even larger biases occur almost twice as often when applying the Heckman correction compared to the simple OLS estimation. The simple OLS as well as the IPW-estimator show the smallest number of relative biases. This finding clearly illustrates that the decrease in significant biases for the Heckman- and the imputation estimator is merely due to an inflation of the variance. This is confirmed in Table 8, which shows a summary of the variance inflation of the different estimators. The standard error is estimated based on the bootstrap replications:

$$SE(\hat{\beta}) = \sqrt{\frac{1}{B-1}\sum_{b=1}^{B}\left(\hat{\beta}_b^* - \frac{1}{B}\sum_{b=1}^{B}\hat{\beta}_b^*\right)^2} \tag{19}$$

It is evident that the Heckman procedure increases the standard error uniformly. The imputation estimator and the OLS estimator are performing about equally well. The inverse probability estimator also increases the variance, but less so compared to the Heckman estimator.

Table 8: Comparison of the variance inflation ($\sigma_{\hat{\beta}_{1,j}} / \sigma_{\hat{\beta}_{01}}$) of the different procedures

	Respondents	IPW	Heckman	Imputation
< 90 %	1	1	1	1
90 - 100 %	16	6	2	10
100 - 110 %	206	182	39	207
>110 %	17	51	198	22

One way to regard the bias as well as the variance is to estimate empirically the mean square error (MSE) based on the bootstrap replications:

$$MSE(\hat{\beta}) = \sqrt{\frac{1}{B}\sum_{b=1}^{B}\left(\hat{\beta}_b^* - \hat{\beta}_{01}\right)^2} \tag{20}$$

Table 9: Comparison of the MSE-inflation ($MSE(\hat{\beta}_{1,j}) / MSE(\hat{\beta}_{01})$) of the different procedures

	Respondents	IPW	Heckman	Imputation
< 90 %	2	2	1	2
90 - 100 %	5	2	1	4
100 - 110 %	96	80	8	96
>110 %	137	156	230	138

Table 9 contains the comparison of the mean square errors of the four different estimation procedures to the mean square error resulting when using the full sample of respondents and attriters in each bootstrap replication. The findings strongly support the result of the variance comparison. The result for the imputation estimator resembles the result for the OLS estimator. The inverse probability weighting and especially the Heckman correction strongly increase the mean square.

The main finding therefore is that the number of significant and strong attrition biases is in general small and that procedures to alleviate the effects of attrition tend to increase the uncertainty in the estimation rather than reducing the bias.

4. Conclusion

The aim of the paper was to assess the extent and significance of attrition biases when estimating Mincerian type income equations based on the ECHP and to compare the performance of different correcting estimation techniques. The significance pattern of the income determinants was found to resemble across countries. Nevertheless, the parameter values (partial effects) vary strongly across countries. Rates of return for highest education vary between 6% in Denmark and 32% in Luxembourg. The income loss for basic education is highest (30%) in Germany and Luxembourg.

Our findings indicate that the effects of attrition on income equations in general are very mild. Only few parameters were found to be estimated with significant bias when analysing 30 cross sections. Hence we conclude that the problem of attrition is no matter of great concern when estimating income equations of the Mincerian type based on the ECHP data.

The empirical comparison of different estimators revealed that procedures to correct for attrition, especially the Heckman-procedure, add considerable variance to the estimation. The inverse probability estimator was found to reduce the strongest biases but also to slightly increase the variance. Hence, for the case of estimating income equations the simple OLS estimation without correcting for attrition is the suggested best practice.

References

Behr A, Bellgardt E, Rendtel U (2002) Extent and determinants of panel attrition in the European Community Household Panel, Working Paper, URL: www.destatis.de /chintex.

Fitzgerald J, Gottschalk P, Moffitt R (1998) An Analysis of Sample Attrition in Panel Data. The Michigan Panel Study of Income Dynamics. The Journal of Human Resources 33:251-299.

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

Heckman, J I (1976), The common structure of statistical models of truncation, sample selection and limited dependent variables, a simple estimator for such models, Ann. Econ. Soc. Meas. 5: 475-492.

Hines F, Tweeten L/Redfern M (1970) Social and Private Rates of Return to Investment in Schooling, by Race-Sex Groups and Regions. The Journal of Human Resources 5:318-340.

Hosmer D W, Lemeshow S (1980) A Goodness-of-Fit Test for Multiple Logistic Regression Model. Communications in Statistics A10:1043-1069.

Lillard L A, Panis C (1998) Panel Attrition from the Panel Study of Income Dynamics. The Journal of Human Resources 33:437-457.

McKelvey R, Zavoina W (1975), A statistical model for the analysis of ordinal level dependent variables, Journal of Mathematical Sociology, 4, 103-120.

Mincer J (1958) Investment in Human Capital and Personal Income Distribution. The Journal of Political Economy 66:281-302.

Mincer J (1970) The Distribution of Labour Incomes: A Survey, Journal of Economic Literature 8:1-26.

Mincer J (1997) The Production of Human Capital and the Life Cycle of Earnings: Variations on a Theme. Journal of Labour Economics 15:S26-S47.

Neukirch T (2002) Nonignorable attrition and selectivity biases in the Finnish subsample of the ECHP. Working Paper, URL: www.destatis.de/chintex.

Nicoleti C, Peracchi F (2002) A Cross-Country Comparison of Survey Nonparticipation in the ECHP. Working Paper.

Peracchi F (2002) The European Community Houshold Panel: A Review. Empirical Economics 27:63-90.

Rendtel U (1995) Lebenslagen im Wandel: Panelausfälle und Panelrepräsentativität. Campus Verlag, Frankfurt/M., New York.

Rendtel U (2002) Attrition in Household Panels: A Survey. Working Paper, URL: www.destatis.de/chintex.

Rendtel U (2004) Report on panel effects. Working Paper, URL: www.destatis.de/chintex.

Robins J, Rotnitzky A, Zhao L (1995) Analysis of Semi-parametric Regression Models for Repeated Outcomes in the Presence of Missing Data. Journal of the American Statistical Association 90:106-121.

Rosenbaum P R, Rubin D B (1983), The central role of propensity score in observational studies for causal effects, Biometrika, Vol. 70, pp. 41-55.

Rubin D B (1973), The use of matched sampling and regression adjustment to remove bias in observational studies, Biometrics, Vol. 29, pp. 185-203.

Vaillancourt F (1995) The Private and Total Returns to Education in Canada 1985. The Canadian Journal of Economics. 28:532-554.

Appendix

Table A1: Log-earnings equations by country, first wave results

Country Wave	Germany 1	Germany SOEP 1	United-Kingdom 1	United-Kingdom BHPS 1	Finland 3	Denmark 1	Ireland 1	The Nether-lands 1
Intercept	4.8840 (36.53)	4.8848 (39.68)	4.7936 (37.24)	4.7471 (41.89)	5.4715 (32.19)	5.4389 (40.09)	5.2979 (34.25)	4.7145 (27.29)
Tenure	0.0199 (6.24)	0.0294 (9.81)	0.0297 (8.13)	0.0098 (2.72)	0.0202 (5.96)	0.0128 (3.59)	0.0314 (8.07)	0.0165 (4.98)
Age	0.1293 (16.96)	0.1246 (17.86)	0.1287 (17.92)	0.1401 (22.19)	0.0873 (9.09)	0.1136 (15.35)	0.0827 (8.83)	0.1371 (14.1)
Age2 /100	-0.1545 (-16.02)	-0.1498 (-17.18)	-0.1550 (-17.82)	-0.1685 (-21.78)	-0.0989 (-8.05)	-0.1341 (-14.84)	-0.0997 (-8.5)	-0.1608 (-12.97)
Third level education	0.2331 (8.58)	0.2281 (9.24)	0.3482 (11.04)	0.1917 (5.72)	0.3133 (11.19)	0.2070 (7.07)	0.2838 (7.73)	0.2760 (9.73)
Basic Education	-0.1864 (-6.28)	-0.1676 (-6.26)	-0.1410 (-4.68)	-0.1491 (-4.48)	-0.1844 (-4.99)	-0.1750 (-5.24)	-0.1411 (-4.56)	-0.0913 (-2.57)
Unempl.time	-0.0020 (-1.85)	-0.0019 (-3.07)	-0.0067 (-5.62)	-0.0069 (-3.44)	-0.0058 (-5.08)	-0.0085 (-5.37)	-0.0056 (-6.22)	-0.0066 (-5.13)
Married	0.0692 (2.5)	0.0547 (2.27)	0.0869 (2.8)	0.0993 (3.58)	0.0659 (2.27)	0.0456 (1.55)	0.2833 (7.3)	0.0899 (3.09)
n	1,640	2,261	1,800	1,872	1,110	1,199	1,286	1,446

Country Wave	Belgium 1	Luxem-bourg 1	France 1	Spain 1	Portugal 1	Austria 2	Italy 1	Greece 1
Intercept	6.1507 (29.3)	5.6706 (17.95)	5.9044 (38.48)	5.3092 (50.03)	5.1056 (47.13)	5.4067 (37.36)	-1.0588 (-10.04)	5.1343 (39.15)
Tenure	0.0093 (2.58)	0.0119 (2.06)	0.0157 (5.58)	0.0256 (10.32)	0.0129 (4.38)	0.0091 (2.51)	0.0139 (6.43)	0.0168 (5.68)
Age	0.0568 (4.7)	0.1089 (5.82)	0.0609 (6.95)	0.0665 (11)	0.0558 (9.19)	0.1017 (11.36)	0.0569 (9.26)	0.0506 (6.73)
Age2 /100	-0.0654 (-4.11)	-0.1443 (-5.66)	-0.0639 (-5.62)	-0.0751 (-9.87)	-0.0692 (-9.33)	-0.1176 (-9.63)	-0.0694 (-8.73)	-0.0570 (-6.14)
Third level education	0.2737 (9.07)	0.3869 (6.59)	0.4451 (17.24)	0.2749 (10.21)	0.7178 (13.05)	0.2341 (4.63)	0.2046 (6.89)	0.1663 (6.22)
Basic Education	-0.1048 (-3)	-0.2176 (-4.37)	-0.1871 (-7.4)	-0.1728 (-7.3)	-0.2846 (-8.85)	-0.2705 (-7.55)	-0.1342 (-7.86)	-0.1527 (-5.85)
Unempl.time	-0.0038 (-3.09)	-0.0250 (-3.26)	-0.0108 (-8.81)	-0.0035 (-5.85)	-0.0034 (-2.4)	-0.0093 (-3.47)	-0.0023 (-8.04)	-0.0017 (-1.59)
Married	0.0692 (2.33)	0.0116 (0.24)	0.1197 (5.07)	0.0876 (3.77)	0.1186 (4.04)	0.0182 (0.55)	0.0864 (4.12)	0.1885 (6.88)
n	750	360	1,833	2,058	1,425	1,185	1,657	1,194

Table A2: First wave logit results by country

Country	Germany	Germany SOEP	United-Kingdom BHPS	Denmark	Ireland	Belgium	France	Portugal	Italy
Wave	2	1	1	2	2	1	1	3	1
Intercept	3.43550	4.86830	-0.78140	-0.68810	0.49830	-5.07550	-1.17030	12.62270	3.41480
	(1.2)	(2.65)	(-0.39)	(-0.38)	(0.33)	(-1.32)	(-0.64)	(3)	(2.59)
lnw	-0.10740	0.44570	0.97880	0.47020	0.05200	0.52730	-0.20910	-1.41120	-0.08660
	(-0.28)	(1.91)	(3.41)	(1.88)	(0.27)	(1.23)	(-0.94)	(-2.57)	(-0.28)
Tenure	0.0685	0.0261	-0.0257	-0.0165	0.0012	0.0445	-0.0327	0.0819	0.0492
	(1.25)	(0.66)	(-0.53)	(-0.66)	(0.05)	(0.83)	(-1.18)	(1.35)	(1.73)
Age	0.02210	-0.29190	-0.15120	-0.03770	0.04160	0.19480	0.30950	-0.03250	-0.03580
	(0.16)	(-2.93)	(-1.62)	(-0.58)	(0.66)	(1.34)	(4.07)	(-0.19)	(-0.47)
Age^2	0.00010	0.00380	0.00180	0.00040	-0.00030	-0.00280	-0.00390	0.00070	0.00030
	(0.04)	(2.87)	(1.58)	(0.54)	(-0.37)	(-1.5)	(-4.1)	(0.31)	(0.3)
Education - third level	-0.00540	-0.01180	0.03050	0.17230	0.22540	-0.51570	0.33380	0.41420	0.01760
	(-0.01)	(-0.04)	(0.07)	(0.75)	(0.91)	(-1.21)	(1.13)	(0.4)	(0.05)
less than 2nd stage of	-0.11260	-0.78870	0.16800	-0.17930	0.34110	-0.25150	-0.33540	0.03990	-0.16870
	(-0.25)	(-2.74)	(0.38)	(-0.74)	(1.62)	(-0.52)	(-1.46)	(0.07)	(-0.77)
Unempl.time	-0.02380	0.01170	0.03430	0.00920	-0.00470	0.02980	-0.01050	0.00800	0.00950
	(-3.12)	(1.21)	(0.73)	(0.68)	(-0.77)	(0.76)	(-1)	(0.31)	(1.76)
Married	0.44940	0.53530	0.09240	0.28530	-0.06410	0.62450	-0.01540	0.80910	0.46150
	(1.03)	(1.89)	(0.24)	(1.29)	(-0.24)	(1.52)	(-0.06)	(1.41)	(1.77)
HH has moved	-0.91820	-1.33640	-0.56730	-0.58900	-1.36960	-1.51770	0.38430	-2.54730	-1.03400
	(-1.38)	(-3.47)	(-0.87)	(-1.44)	(-4.58)	(-2.08)	(0.51)	(-5.63)	(-2.61)
Interviewer change	-0.14840	0.10380	-0.00070	-0.97160	-1.42290	-0.27900	-0.31770	-0.12170	-1.03520
	(-0.23)	(0.25)	(0)	(-5.07)	(-6.58)	(-0.62)	(-1.26)	(-0.28)	(-5.19)
CHI	18.350	34.040	12.350	37.530	88.850	20.050	24.820	52.170	45.880
p	0.050	0.000	0.260	0.000	0.000	0.030	0.010	0.000	0.000
LRI	0.060	0.050	0.030	0.050	0.100	0.080	0.030	0.200	0.060
R2MZ	0.110	0.110	0.070	0.080	0.130	0.140	0.060	0.260	0.110
CHI,HL	6.950	4.080	9.530	10.180	6.080	6.800	8.060	13.750	4.980
p,HL	0.540	0.850	0.300	0.250	0.640	0.560	0.430	0.090	0.760
n	1,505	1,990	1,724	1,101	1,037	0,401	1,773	1,331	1,656

Country	Finland	The Nether-lands	Luxem-bourg	Spain	Austria	Greece
Wave	3	1	1	1	2	1
Intercept	0.31760	-0.45030	-4.30340	0.45380	2.86990	5.19220
	(0.16)	(-0.31)	(-1.14)	(0.43)	(1.88)	(3.05)
lnw	-0.10720	-0.02550	0.65620	0.09580	0.02970	-0.57480
	(-0.42)	(-0.13)	(1.29)	(0.63)	(0.14)	(-2.38)
Tenure	0.0302	2.0077	3.7237	2.478	1.5454	1.9657
	(0.01)	(0.71)	(0.63)	(1.34)	(0.59)	(0.75)
Age	0.12780	0.11420	-0.00330	0.03130	-0.05750	0.00270
	(1.54)	(1.5)	(-0.02)	(0.76)	(-0.86)	(0.04)
Age	-0.13260	-0.12250	0.08850	-0.06610	0.06250	-0.01270
	(-1.24)	(-1.24)	(0.29)	(-1.3)	(0.69)	(-0.15)
Education - third level	-0.23070	0.39740	0.74370	-0.08920	-0.20180	-0.27050
	(-0.94)	(1.51)	(1.29)	(-0.48)	(-0.58)	(-1.26)
less than 2nd stage of	0.31570	-0.32840	1.08160	0.25790	-0.24150	0.49700
	(0.88)	(-1.39)	(2.54)	(1.55)	(-0.96)	(2)
Unempl.time	-0.00640	0.01130	-0.07590	-0.00820	-0.02560	0.01010
	(-0.83)	(0.91)	(-1.36)	(-2.26)	(-1.81)	(0.96)
Married	0.12020	0.58060	0.03220	0.42930	0.28560	0.58070
	(0.48)	(2.61)	(0.07)	(2.64)	(1.22)	(2.52)
HH has moved						
Interviewer change						
CHI	11.850	48.770	18.140	28.910	6.970	30.750
p	0.160	0.000	0.020	0.000	0.540	0.000
LRI	0.020	0.050	0.080	0.020	0.010	0.030
R2MZ	0.040	0.090	0.170	0.030	0.010	0.070
CHI,HL	16.270	4.040	14.750	12.420	15.420	6.720
p,HL	0.040	0.850	0.060	0.130	0.050	0.570
n	1,110	1,446	0,360	2,058	1,185	1,194