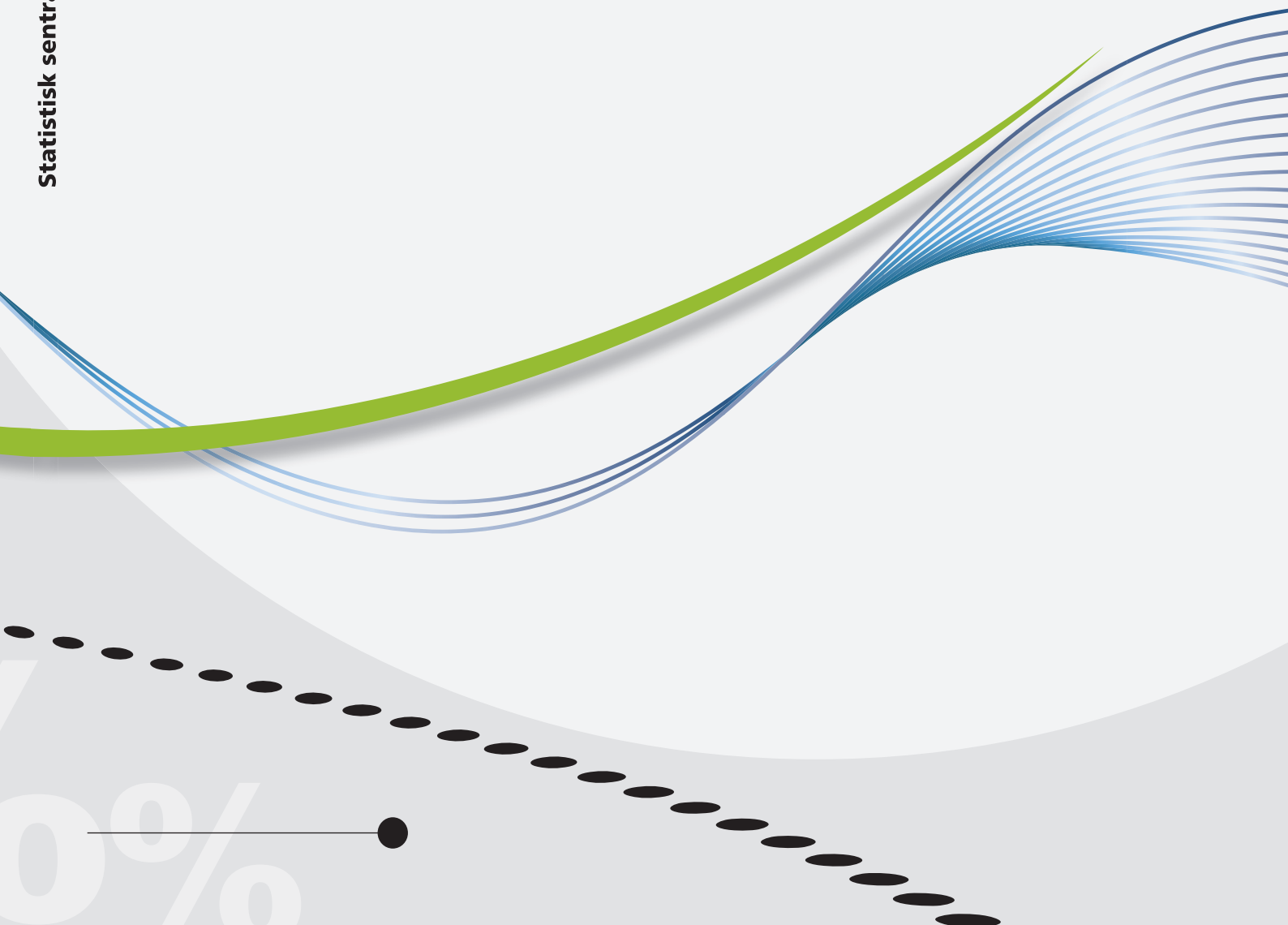




Marit Wilhelmsen

Effects of attrition in the Norwegian Survey on statistics on income and living conditions (EU-Silc)



Marit Wilhelmsen

**Effects of attrition in the Norwegian Survey on
statistics on income and living conditions
(EU-Silc)**

© Statistics Norway When using material from this publication, Statistics Norway shall be quoted as the source.	Symbols in tables	Symbol
	Category not applicable	.
	Data not available	..
	Data not yet available	...
	Not for publication	:
	Nil	-
ISBN 978-82-537-8391-8 Printed version	Less than 0.5 of unit employed	0
ISBN 978-82-537-8392-5 Electronic version	Less than 0.05 of unit employed	0.0
ISSN 1891-5906	Provisional or preliminary figure	*
Subject: 00.90	Break in the homogeneity of a vertical series	—
Published May 2012	Break in the homogeneity of a horizontal series	
Print: Statistics Norway	Decimal punctuation mark	.

Preface

In this publication we analyze the effect of attrition in the Norwegian Survey on statistics on income and living conditions, EU-Silc. Attrition is defined as unit nonresponse of eligible cases that occurs after the first wave of the panel survey. The reason to be concerned about attrition is that loss of respondents over time will reduce the sample size, and this may lead to less precise estimates. Also, attrition may be of a selective nature and thereby increase the bias of certain estimates.

To study the scope of attrition we followed the participants from a rotational group drawn in 2003 until 2008. To get an idea of the characteristics of attrition, we have used information from administrative registers on age, income gender and education, as well as information from the survey on health and employment status.

The paper was presented at Q2010 – the European Conference on Quality in Official Statistics, in Helsinki, Finland.

This paper has required a lot of help, and there are many people to thank: Gustav Haraldsen, Frode Berglund and Ib Thomsen have helped with literature about attrition and vital comments and suggestions during the work on the paper. Tor Morten Normann has given important feedback on the hypotheses and helped with accessing and analyzing data from the survey. Øyvind Kleven has been an indispensable collaborator in writing this paper, and finally Mari Sanderlien helped with the data files.

Statistics Norway,
Oslo/Kongsvinger, 8 May 2012

Hans Henrik Scheel

Abstract

In this paper we investigate attrition by using the Norwegian Survey on statistics on income and living conditions, EU-Silc, as a case. We define attrition as unit nonresponse of eligible cases that occurs after the first wave of the panel survey. There are two issues we investigate:

- What characteristics are associated with attrition?
- What is the effect of attrition?

The reason we are concerned about attrition is that loss of respondents over time will reduce the sample size, and may lead to less precise estimates. Also, attrition may be of a selective nature and thereby increase the bias of certain estimates.

By using relevant theory, we discuss causes for attrition, and try to identify relevant socio-demographic variables that are related to attrition. We are concerned if nonresponse in later waves differs in significant ways from initial nonresponse, and therefore introduces an unknown bias in survey estimates. We theoretically argue that the socio-demographic variables household income, education, household, age, employment and material status are important when it comes to attrition.

In the paper we demonstrate that household income, education and age influence the response rate in the Norwegian EU-silc. When it comes to attrition, we also find that these variables are important; we have also added two survey variables we think are important for attrition. These are employment status and health. We find that attrition follows many of the same patterns as nonresponse, and that employment status and health affect the likelihood to continue in the panel. Still, the most important characteristics for attrition are low education and a low household income. Since these variables are highly correlated we use a statistical model to adjust for other background variables so we can find the propensity to participate in previous waves. The model indicates that there is a much lower propensity to continue participating if you are in older age groups and have low education.

Contents

Preface	3
Abstract	4
1. Introduction	6
2. The Norwegian EU-Silc	7
2.1. Response rates in the Norwegian EU-Silc	8
3. Theoretical background	9
4. Characteristics of attrition in the Norwegian EU-Silc	10
5. Conclusions	13
References	14
Appendix	16
List of figures	18
List of tables	18

1. Introduction

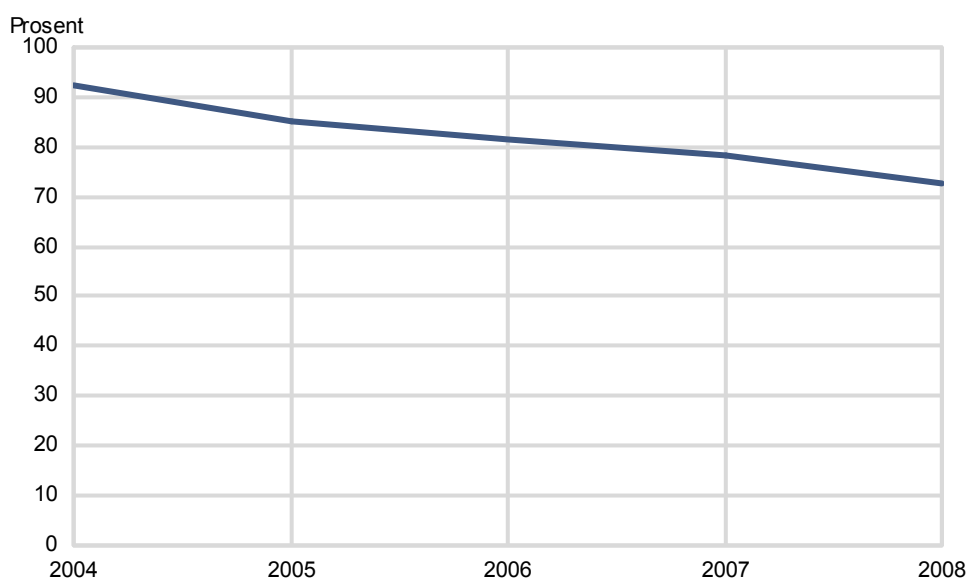
In this paper we use the Norwegian Survey on statistics on income and living conditions, EU-Silc, to study the effects and characteristics of attrition. Longitudinal studies offer the opportunity to examine trends over time, but they are subject to nonresponse or attrition in subsequent waves. We define attrition as unit nonresponse of eligible cases that occur after the first wave of the panel survey.

There are two main reasons to be concerned about attrition:

- 1) Loss of respondents from wave to wave will reduce the sample size over time. This can reduce the precision of estimates.
- 2) Attrition may be of a selective nature and thereby increase the bias of certain estimates. This occurs when non respondents and respondents are systematically different in characteristics that are relevant for the survey (Fitzgerald, Gottschalck and Moffitt 1998, Menard 2008, Rendel 2002).

To study the scope of attrition we have followed one rotational group drawn in 2003 until 2008¹. Figure one shows the participants from 2003, and who continued to participate from 2004 until 2008.

Figure 1. Response rate for participants in T1 (2003) in 2004 – 2008



The figure clearly displays that attrition is present in this survey. In 2008 we had lost 27 per cent of the original participants from 2003. The greatest decline was in the earliest waves; after that the effect was still present, but somewhat weaker.

In this paper we investigate the following questions:

Research questions

- What characteristics are associated with attrition? Which socio-demographic variables are associated with panel attrition?
-
- What is the effect of attrition? Does attrition differ from initial nonresponse and might it therefore lead to new biases in the survey?

¹ See appendix one for more information about the sampling design.

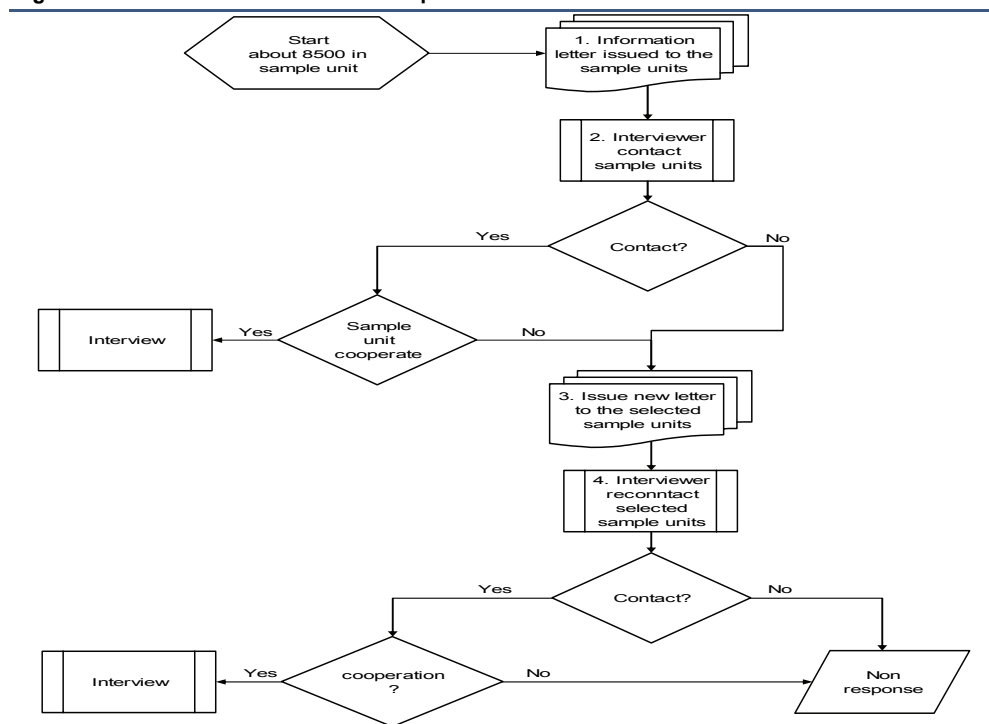
2. The Norwegian EU-Silc

Since 1997, Statistics Norway has carried out annual panel surveys on income and living conditions. The EU-Silc is expected to deliver comparative statistics on income distribution and social exclusion on a European level, and provides two types of data:

- Cross-sectional data of a given time or a certain time period with variables on income, poverty, social exclusion and other living condition.
- Longitudinal data on both household and individual-level changes over time, observed periodically over a four year period.

In figure two we present the data collection process.

Figure 2. Flowchart. The data collection process in EU-Silc



1. Each year the sample receives a Christmas card with greetings. Enclosed is a letter with results from the survey. We also ask if the recipient has changed contact information.
2. First the respondent is notified that Statistics Norway wishes to get in contact and perform an interview. The letter contains formal information about data protection and the use of administrative registers. The envelope also includes a brochure explaining the use of the survey and why it is important to participate.
3. The interviewer contacts the respondent to set up an interview. They contact the respondent by telephone. For respondents we have contacted in previous waves, we use the same contact information. For new respondents or respondents whom we were not able to get in contact with in previous years, telephone numbers from different registers are used. The interviewers also have information about the household, such as address, household composition, age and so on.
4. New notifications are sent out to respondents who are initially non-respondents and people we have not been able to get in contact with. The non-respondent receives a letter where we explain why it is important to participate. The people we have not been able to contact with also get a

letter where they can fill out contact information and return it in a prepaid envelope.

5. The interviewer recontacts the respondents who are initially non-respondents and tries to get a new interview. If the interviewer is still not able to contact the respondent by telephone, he or she will look up the address.

2.1. Response rates in the Norwegian EU-Silc

Table 2 presents the response rate distributed on different register variables from 2003 until 2008. We have selected variables that we know are highly correlated with nonresponse in the EU-Silc.

Table 2. Response rates 2003 – 2008. Distributed on relevant register variables. Per cent.

	2003	2004	2005	2006	2007	2008
All	70,9	73,1	72,1	68,5	70,4	64,3
Gender						
Male	70,3	73,1	73,0	69,5	72,2	66,8
Female	71,4	72,5	71,2	67,5	68,6	61,8
Age groups						
16 – 24 years	70,7	73,9	71,4	66,1	69,8	60,1
25 - 44 years	74,6	78,4	76,6	73,1	72,2	66,4
45 – 66 years	71,6	73,5	73,4	70,0	73,1	66,4
67 -79 years	65,7	64,9	67,3	64,3	66,8	63,3
80 years and older	50,8	46,4	42,1	40,2	48,6	48,8
Education						
Low	57,8	57,3	58,4	59,2	58,7	51,4
Medium	72,3	71,6	72,4	70,1	72,1	65,7
High	80,1	80,7	83,3	77,8	82,8	76,0
Household income						
1. quartile (low)	37,2	38,6	40,0	37,7	35,9	36,7
2. quartile	58,4	58,9	56,7	55,0	49,7	46,3
3. quartile	90,8	90,8	88,9	87,0	81,8	78,7
4. quartile (high)	96,6	97,1	97,3	97,1	97,3	97,0

The response rate has been relatively stable until 2006, where we see a decline. This decline is corrected in 2007, but the response rate continued to fall in 2008.

Household income influences the response rate, and people with high household income participate more compared to people with low household income. Looking at the response rate over the years, we find that the rate is falling for the two middle income quartiles, but is remaining stable for the lowest and highest quartile.

Studying education, we find the same pattern: People with low education participate less than people with high education. The fact that these two variables indicate the same pattern might be because there is a high positive correlation between them. The response rate is descending for all educational groups, but in contrast to income the distribution between the groups has stayed quite stable over the years.

We also find that age affects the response rate. People between 25 and 66 participate more than younger and older persons. For the oldest age group, the response rate is just around 50 per cent. This is about 20 percentage points lower than the average response rate. Even though the response rate is descending, the decline is lower than for the other age groups. So even though the response rate is low, it has been more stable than for other age groups.

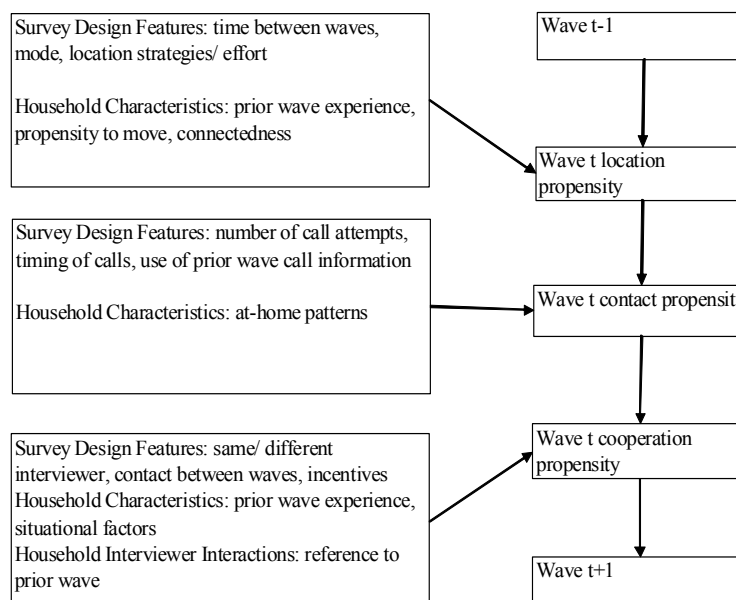
Finally, when studying gender we find that in 2003 women participated more than men. In 2008 this has been reversed, and men participate considerably more than women.

3. Theoretical background

If we are to understand some potential causes of panel attrition, it might be helpful to understand why attrition occurs and which variables are related to this event.

Figure 2 displays a model for nonresponse in longitudinal surveys (Lepowski and Cooper 2002, Rendel 2002).

Figure 2. Model for nonresponse in longitudinal surveys



Groves et al. 2002: 264

The figure divides the nonresponse process into three conditional processes: Location, cooperation and contact. Within each of the processes, survey design features and household characteristics will affect the outcome (Lepowski and Cooper 2002).

Location is a straightforward process in the EU-silc. We get information on addresses from the central population register, and if a person moves, the register has to be updated with new address information. *Contact* difficulty is a significant nonresponse factor in the first wave, but it is normally relatively small in later waves, given a successful location. The interviewer has prior knowledge about the likely at-home patterns and the best times to call. Given this, the number of contact attempts should be lower than in later waves. A change of interviewer or modus might affect the contact propensity (Lepowski and Cooper 2002). Still we find in the EU-silc that nonresponse because of contact difficulties is increasing. There are several reasons for this: (i) our register on telephone numbers do not have the same quality as for addresses. Generally, when we draw a new sample we are unable to find telephone numbers of about 20 per cent. Attrition might occur because a person has changed a number and this is not updated in the register. (ii) The better part of this survey is done by telephone interviewing. The amount of face-to-face interviews has gradually declined over the years. This might also affect the contact rate, because we might have incorrect information on telephone number, or the person is difficult to reach on telephone. (iii) The person might be unwilling to answer the telephone because the he or she can read on the display who is calling, and they know what the inquiry is about. This is a hidden form of non-cooperation. The main challenge in the EU-silc is *cooperation*. The sample person's willingness to cooperate in later waves is influenced by a variety of situational factors and recollection of the prior wave experience. This depends on several factors: The times between waves, whether there is an interwave contact and whether the same interviewer is used or not (Lepowski and Cooper 2002). There is also a cumulative

respondent burden over time, and a possible consequence might be higher attrition as the panel ages. The participant's general health might also influence the propensity to cooperate (Rendtel 2002). A classical hypothesis about cooperation in surveys assumes that social habits have an impact on participation rates in general. This is also relevant for panel studies. A social event might happen that the respondent classifies as private. For example the loss of a household member, change of partner and so on. The respondent might be reluctant to inform about this in a new interview and therefore decides to not participate. There might also have been social events that reveal a decline in social prestige, such as the loss of a job or drawing social benefits (Rendtel 2002).

Erosion of a panel over time might not be a problem if it is evenly spread across all socio- demographic subgroups. Unfortunately, in practice the pattern of attrition is not like this: Particular subgroups (the elderly, the uninterested and so on) are lost in disproportionately large numbers (Waterton and Livesley 1987). Fitzgerald, Gottschalck and Moffitt (1998) also found that "attritions tend to have lower earnings, lower education, lower marriage propensities, and appear to be generally drawn to the lower tail of the socioeconomic distribution" (page 295). Waterton and Livesley's study from 1987 on attrition in a panel study of attitudes in Great Britain, found that the demographic items with a significant differential in attrition were: Employment status, tenure, gross household income, age, party identification and household type. The subgroups that were underrepresented in the continued panel consisted of: Retired, sick, students (employment status), local authority renters (tenure), lower income (gross household income), over 60 years (age), non-aligned (party identification) and pensioner's household (household type). Many of these characteristics are closely associated with initial nonresponse, but Waterton and Livesley (1987) also investigated the effect of attrition on survey variables. At each wave, panel respondents were compared with panel non-respondents on the basis of their answers given in the initial wave of the survey. Some general points were found: i) In attitudes traditionally connected to party identification (liberal and conservative) they found a difference. The continuing panel is much more tolerant, particularity in their views on freedom of speech and sexual behaviour. They are also less likely to register disapproval. This may be partly explained by the fact that older age groups – who might have more stringent moral views – are lost. ii) Those who continue are more likely to hold firmer views. This is based on the fact that those who answer "don't know" are more likely to drop out. iii) Their findings on what demographic and attitudinal characteristics are associated with attrition are consistent with findings from other studies that have focused on non-attitudinal information (Waterton and Livesley 1987).

4. Characteristics of attrition in the Norwegian EU-Silc

As we discussed in the theoretical review, attritions can be associated with specific socio-demographic patterns, and these might influence bias in survey estimates. In this chapter we want to find out what characterises attrition in the EU-silc. Already, we have seen that household income, education and age influence the response rate, so we continue to use these variables when studying attrition. In addition we are adding two variables: Self-reported health (good or bad) and if you are currently employed or not. We argued from the theoretical review that these variables may be associated with behaviour that causes attrition.

In table 3 we present the response rate for the participants in 2003 distributed over relevant demographic variables. In the last column we display the difference between 2003 and 2008.

Table 3. Response rate, rotational group 8 from 2003 to 2008. Distribution of different demographic and survey variables.

	2003	2004	2005	2006	2007	2008	Difference 2003 - 2008
All	100	92,2	85,1	81,4	78,2	72,5	27,5
Gender							
Male	100	91,8	87,0	82,4	80,2	73,9	26,1
Female	100	92,6	83,3	80,4	76,5	71,2	28,8
Age groups							
16 – 24 years	100	86,3	81,7	74,8	74,0	63,4	36,6
25 - 44 years	100	95,6	86,7	85,3	81,9	76,8	23,2
45 – 66 years	100	94,1	88,7	84,8	82,8	78,9	21,1
67 -79 years	100	92,3	84,6	82,2	75,7	69,0	31,0
80 years and older	100	68,0	56,0	40,0	28,0	38,9	61,1
Education							
Low	100	88,3	86,7	75,8	64,1	60,2	39,8
Medium	100	92,6	80,4	80,0	77,6	70,4	29,6
High	100	96,2	90,9	88,2	89,2	85,5	14,5
Income							
1. quarter (low)	100	87,9	76,9	67,6	64,8	58,3	41,7
2. quarter	100	90,7	85,8	80,6	76,8	72,0	28,0
3. quarter	100	93,4	86,7	84,1	79,3	75,0	25,0
4. quarter (high)	100	93,9	86,8	85,8	85,3	77,6	22,4
Health							
Good	100	91,8	86,7	82,8	80,9	75,0	25,0
Bad	100	92,6	80,4	75,7	71,8	64,9	35,1
Employment							
Employed	100	94,4	82,8	85,9	82,7	77,3	22,7
Unemployed	100	87,3	75,7	71,5	68,9	63,4	36,6

Table 3 clearly indicates that the effect of attrition is relatively even between the gender. For both men and women, we find that over the years about three out of ten fall out of the survey, although the attrition rate is somewhat higher among women. When we compare these results to the response rate for all rotational groups (cf. table two), we find that the decline in response rate was higher among women, but that we recruited more women than men initially.

When it comes to age groups, the effect is most severe among the youngest and the oldest age groups. Especially in the older age group, the attrition rate is over 60 percent. This can be explained by age impairment. This group was 80 years or older in 2003, and in 2008 they would be six years older. We also find that this group has a low response rate (cf. table two). But the response rate has been consistently low over the years; as for the other age groups there has been a decline in response rate. People between 25 to 66 are most willing to stay in the panel. The attrition rate for these age groups is lower than for the average.

Also for education, we find the same pattern for attrition as for nonresponse. The higher the education, the better is the chance for a person to continue to participate. When it comes to household income we see the same pattern, with higher household income the likelihood to participate increases.

Finally, the two survey variables self-reported health and if you are currently employed affect the propensity to participate. Among people who reported that their health was bad in 2003, the attrition rate is ten percentage points higher than for people who reported it was good. We also find that persons employed in 2003 are more likely to stay in the panel compared to people not employed. Still the attrition rate is not as high as for people with low income or low education.

This table indicates that attrition follows many of the same patterns as nonresponse reported in table two. We find that the new survey variables affect the likelihood to continue in the panel. Still, the most important characteristics for attrition are low education and a low household income. These are also the variables that affect the response rate most of all.

The last condition we set out to investigate was if attrition has the same characteristics as nonresponse, or if some demographical variables have a stronger effect on the attrition rate than on nonresponse. In table 3, we observed that the demographical variables household income, education and age do affect the attrition rate, as well as the survey variables health and education. But these characteristics are highly correlated, so it is difficult to argue that older people are less likely to continue to participate because of age - or is it correlated with education, income and health? We also set out to investigate if attrition has the same pattern as nonresponse and therefore leads to an intensification of the existing under-representation of some demographic groups.

By using an advanced statistical model that adjusts for other background variables, we can find the propensity to participate in previous waves. We have applied two models: In the first model we have used demographic variables associated with nonresponse (gender, age, education and household income), in the second model we added the interview variables (self reported health and employment). The models consist of people who participated in the first wave (T1), and the dependent variable is if you were interviewed in 2008 (yes or no). The result can be read in table 4, and the significant values are in bold.

Table 4. Logistic regression². Interviewed in 2003. Dependent variable: interview in 2008

	Model 1			Model 2		
	Odds ratio estimates	95% Wald Confidence limits		Odds ratio estimates	95% Wald Confidence limits	
Gender						
Men v. women	1,150	0,817	1,619	1,098	0,777	1,552
Age group						
45 - 66 v. 16 - 24	1,808	1,086	3,011	1,762	1,028	3,019
45 - 66 v. 25 - 44	1,171	0,752	1,832	1,241	0,789	1,953
45 - 66 v. 67 - 79	1,250	0,667	2,343	1,230	0,624	2,424
45 - 66 v. 80 +	6,814	2,573	18,042	6,244	2,258	17,271
Education						
Low v. high	2,801	1,565	5,014	2,679	1,477	4,856
Medium v. high	2,096	1,299	3,318	1,969	1,214	3,191
Household income						
3 quartile v. 1 quartile (low) ..	1,451	0,883	2,387	1,346	0,809	2,241
3 quartile v. 2 quartile	1,188	0,764	1,846	1,123	0,716	1,760
Health						
Good v. bad				1,256	0,815	1,937
Employment						
Employed v. unemployed				1,119	0,714	1,755

The first model indicates that the likelihood to participate is lower for people in the oldest age group when compared to the reference group. Education is also an important factor for participation. The likelihood to participate is much greater among people with higher education than people with low education. In table 3, we found that income influences the propensity to participate, but in the model this variable is not relevant. This is probably because it is highly correlated with education, and the effect of income therefore disappears.

In the second model, we have also applied two interview variables to investigate if they affect the results. In table 3, we found that the two variables had an impact on participation in future waves. As we can see from the results of the model, the two variables do not affect the results. It is still age and education that influence if a respondent continues to participate.

² See appendix for Logistic regression

5. Conclusions

In this paper we set out to answer two questions. First; what socio-demographic variables are associated with attrition. In the theoretical review, we argued that variables connected to income, household type (marital status), employment status and education were associated with attrition. In the EU-Silc, we found that that household income, education and age influence the overall response rate. Also, we found that attrition in the EU-Silc follows many of the same patterns as nonresponse. But, it is the characteristics low education and low household income that are most affected by the effect of attrition.

Second; does attrition follows the same pattern as initial nonresponse. In the theoretical review, we discussed relevant socio-demographic variables that are related to attrition. We theoretically argue that the socio-demographic variables household income, education, household, age, employment and material status are important when it comes to causes of attrition.

In the paper, we demonstrated that household income, education and age influence the response rate, and for attrition we also find that these variables are important. In addition, we added two survey variables we think are important for attrition (employment status and self-reported health). We find that attrition follows many of the same patterns as nonresponse, and that employment status and health affect the likelihood to continue in the panel. Still, the most important variables connected to attrition were low education and a low household income. Since these variables are highly correlated, we use a statistical model that adjusts for other background variables to find the propensity to participate in previous waves. The model indicates that attrition is connected with low education and old age. We have indicated in this paper that these variables are associated with initial nonresponse. Therefore we argue that initial nonresponse and attrition pattern in the EU-silc share many of the same characteristics.

References

- Behr, A., Bellgardt, E. and Rendtel, U. (2005). *Extent and Determinants of Panel Attrition in the European Community Household Panel*. *European Sociological Review*. 21, 489-512.
- Contoyannis, P., Jones, A.M. and Rice, N. (2004). *The Dynamics of Health in the British Household Panel Survey*. *Journal of Applied Econometrics*. 19, 473-503.
- Ervasti, H., Fridberg, T., Hjerm, M. and Ringdal, K. (2008). *Nordic Social Attitudes in a European Perspective*. Edward Elgar Publishing Limited. UK.
- Eurostat (2004). *Description of Target Variables: Cross-Sectional and Longitudinal*. Unit E-2: Living Conditions. EU-Silc 065/04.
- Fitzgerald, J., Gottschalk, P. and Moffitt, P. (1998). *An Analyse of Sample Attrition in Panel Data: The Michigan Panel Study of Income Dynamics*. *The Journal of Human Resources*. 33, 251-299.
- Keulenar, F.D. (2009). *Using Process Data to Predict Attrition from a Panel Survey: a Case Study*. Paper presented at the annual meeting of the American Association For Public Opinion Association, Fontainebleau Resort, Miami Beach, FL. URL: http://www.allacademic.com/meta/p17153_index.html
- Laurie, H. (2008). *Minimizing panel attrition*. In *Handbook of Longitudinal Research. Design, Measurement and Analysis*. Ed Menard, S. Academic Press publications. USA.
- Lepkowski, M., J. and Couper, P., M. (2002). *Nonresponse in the Second Wave of Longitudinal Household Surveys*. in Groves, R., M., Dillman, D., A., Eltinge, J., L., and Little, R., J., A. Ed (2002). *Survey Nonresponse*. Wiley and Sons, New York, USA.
- McGonagle, K.A., Cooper, M.P. and Schoeni, R.F. (2011). *Keeping Track of Panel Members: An Experimental of Between-Wage Contact Strategy*. *Journal of Official Statistics*, 27, 319-338.
- Rendtel, U. (2002). *Attrition in Household Panels: A Survey*. CHINTEX. Working Paper nr 4. J.W. Goethe-Universitat Frankfurt/Main.
- Ringdal, K. (2007). *Enhet og mangfold. Samfunnsvitenskaplig forskning og kvantitativ metode*. Fagbokforlaget, Oslo
- Thomsen, I., Kleven, Ø., Wang, J.E. and Zhang, L.C. (2006). *Coping with decreasing response rates in Statistics Norway. Recommended practice for reducing the effect of nonresponse*. Report 2006/29 Oslo: Statistics Norway.
- Thornberry, T.P., Bjerregaard, B. and Miles, W. (1993). *The Consequences of Respondent Attrition in Panel Studies: A Simulation Based on the Rochester Youth Development Study*. *Journal of Quantitative Criminology*. 9, 127-158.
- Statistic Norway (2008). *Final Quality Report. EU-Silc 2008: Norway*.
- Waterton, J. and Lievesley, D (1987). *Attrition in a Panel Study of Attitudes*. *Journal of Official Statistics*, 3, 267- 282.

Wiggen, K. (2011). *Dokumentasjonsrapport for levekår panel 2003 – 2007*. Documents. Statistics Norway. Oslo.

Wiggen, K., Paulsen, K.A.S., (2011). *Dokumentasjonsrapport for levekår panel 2008 -2011*. Documents. Statistics Norway. Oslo.

Wilhelmsen, M. (2011). *Mediebruk 2010. Dokumentasjonsrapport*. Documents 38/2011. Statistics Norway. Oslo.

Wilhelmsen, M. (2011). *Holdninger til norsk bistand*. Reports 13/2011. Statistics Norway. Oslo.

Appendix

Sampling in the Norwegian EU-Silc

The sample units are persons aged 16 or more registered in the Central Population register (inhabitants). From 2003 to 2007, the primary stratification for the sample was age. The design chosen implies that age was a central criterion for representivity. The sample was drawn from a proportion p of the population within one-year groups. In addition, the population register was ordered to ensure geographical representivity. At the initialisation of the EU-Silc, the sample size was set to 8,500 persons, each representing a household. In 2003 (t) 8,500 persons constitutes a proportion $p = 0.0024$ of the total population (inhabitants aged 16 years or more). This proportion is meant to be identical each year of the survey, and thus the size of the gross sample will change according to changes in the population. In 2007, the sample design was changed slightly concerning the composition of the rotational groups.

At the initialisation of the EU-Silc in 2003, the sample consisted of both an existing panel from the EPCH survey and the part drawn to implement the sampling plan for EU-Silc. It was not until 2008 that the sample consisted of persons merely drawn for the sampling plan for the EU-Silc (see table 1)

Rotational design

A rotational design was chosen to ensure both cross-sectional and longitudinal data.

This design rotates a part of the sample from one year to the next, retaining the other part unchanged.

In the Norwegian design, each respondent (sample unit) is part of the sample for eight years. Each year 1/8 of the sample will be replaced. From 2007, the sample design was changed slightly and each rotational group is representative for the population of the year it has been drawn. Each rotational group in 2003 consists approximately of 1,060 – 1,070 persons. Since the sample is drawn from a proportion of the population, the groups will get larger as the population increases.

Table one shows how the groups were constituted in 2003.

Appendix table 1. Rotational groups in 2003.

Rotational group	N	Last year in sample
Group 1	1062	2003
Group 2	1062	2004
Group 3	1062	2005
Group 4	1062	2006
Group 5	1061	2007
Group 6	1067	2008
Group 7	1066	2009
Group 8	1066	2010

Logistic Regression Analyses

A logistic regression analysis is an analytical method used to analyse the linear functional relationship between one continuous dependent variable (Y) and one or more explanatory variables (X).

Logistic regression is suitable when the dependent variable (Y) has the values 1 and 0 with the probabilities $p(x_1, x_2, \dots, x_n)$ that $y = 1$. Instead of writing the Y on the left side of the equal sign, we write the log of the odds, the logit:

$$\log \frac{p(x_1, x_2, \dots, x_n)}{1 - p(x_1, x_2, \dots, x_n)} = \alpha + \beta_1 + \beta_2 + \dots + \beta_n$$

The antilogarithm of the logistic regression coefficient may be interpreted as odds ratios. For 0 – 1 variable, such as gender, the antilogarithm is the estimate of the odds ratio between men and women (or vice versa). For continuous explanatory X , the antilogarithm is the odds ratio between persons with one unit of measurement difference on X .

List of figures

Figure 1.	Response rate for participants in T1 (2003) in 2004 – 2008	6
Figure 2.	Flowchart. The data collection process in EU-Silc.....	7
Figure 2.	Model for nonresponse in longitudinal surveys	9

List of tables

Table 2.	Response rates 2003 – 2008. Distributed on relevant register variables. Per cent.....	8
Table 3.	Response rate, rotational group 8 from 2003 to 2008. Distribution of different demographic and survey variables.	11
Table 4.	Logistic regression. Interviewed in 2003. Dependent variable: interview in 2008.....	12
Appendix table 1.	Rotational groups in 2003.....	16

B Returadresse:
Statistisk sentralbyrå
NO-2225 Kongsvinger

Statistics Norway

Oslo:

PO Box 8131 Dept

NO-0033 Oslo

Telephone: + 47 21 09 00 00

Telefax: + 47 21 09 00 40

Kongsvinger:

NO-2225 Kongsvinger

Telephone: + 47 62 88 50 00

Telefax: + 47 62 88 50 30

E-mail: ssb@ssb.no

Internet: www.ssb.no

ISBN 978-82-537-8391-8 Printed version

ISBN 978-82-537-8392-5 Electronic version

ISSN 1891-5906

ISBN 978-82-537-8391-8



9 788253 783918



Statistisk sentralbyrå
Statistics Norway