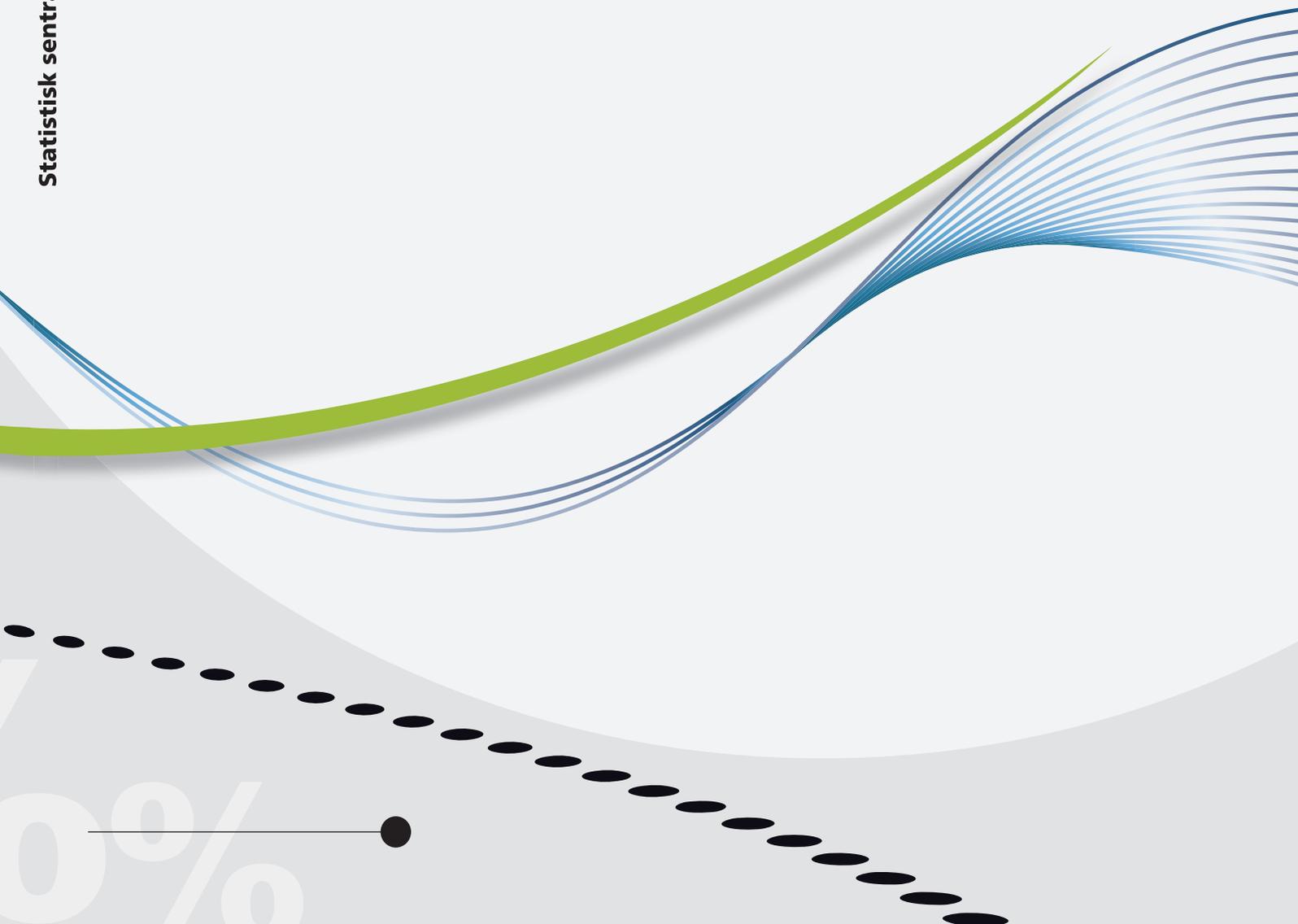




Lisa Li

Proxy Interviews in the Norwegian Labour Force Survey

Analysis of self-respondent and proxy interviews, and their impact on estimates



Lisa Li

Proxy Interviews in the Norwegian Labour Force Survey

Analysis of self-respondent and proxy interviews,
and their impact on estimates

In the series Documents, documentation, method descriptions, model descriptions and standards are published.

© Statistics Norway

When using material from this publication, Statistics Norway shall be quoted as the source.

Published 17 December 2018

ISBN 978-82-537-9861-5 (electronic)

Symbols in tables	Symbol
Category not applicable	.
Data not available	..
Data not yet available	...
Not for publication	:
Nil	-
Less than 0.5 of unit employed	0
Less than 0.05 of unit employed	0.0
Provisional or preliminary figure	*
Break in the homogeneity of a vertical series	—
Break in the homogeneity of a horizontal series	
Decimal punctuation mark	.

Preface

The aim of this study is to investigate selection and measurement effects of proxy interviews in the Norwegian Labour Force Survey.

Thanks to Susie Jentoft and Jørn Ivar Hamre for support and guidance.

Statistics Norway, 11 December 2018

Christian Thindberg

Abstract

In 2018 there will be a test-run of a multi-mode data collection for the labour force survey with the use of CAWI. This pilot does not permit proxy interviews, and it is therefore important to get a better understanding on who is responding with proxy interviews and how this may affect the labour force survey statistics. The sampling in the future may change from family selection to person selection.

This document summarises trends in the Norwegian Labour Force Survey between 2006 and 2017, and it summarises the impact proxy interview gives compared to self-respondent interview for persons aged 15 to 74 years. A closer look at background such as gender, age, register and survey employment data, student status and marital status etc. can help us define which group is represented most in proxy interviews. We found that younger persons who mostly are students have higher chance of proxy interviews. They are usually single persons, who haven't settled yet. The youngest age group from 15 to 19 years has more proxy interviews than self-respondent interviews. The distribution for females and males swapped for proxy interviews during this period. Until around 2012, men were more likely to have proxy interviews, but in the recent years, women are more likely to have proxy interviews.

Odds ratio for certain groups and sub-groups can compare the odds in self-respondent and proxy interviews. The results from odds ratio replicate the tables and plots from the trends between 2006 and 2017. We also compare the odds ratios for register employment and survey employment data for self-respondent interviews compared to proxy interviews, for students and non-students at the age of 15 to 29-year olds in 2000, 2006 and 2017. There is no significant difference for students and non-students by comparing the quarters for each year. However, the odds ratios to employments, (register and survey employment), for self-respondent compared to proxy interviews have increased by the years.

In a sample survey, each person in the net sample is been given a weight of how many persons they represent in the population, and the weights estimate the employment status for the total population. Originally, the weights include both self-respondents and proxies. However, going from family selection to person selection, the proxy interviews are assumed as part of the non-response. This means that the weights are only given to self-respondents. Overall there is an increase in the labour force, employed and unemployed, and a decrease outside the labour force, when we go from the original labour force survey to assume the proxies as non-response. However, the youngest age group 15 to 24 years has a 1 to 3.5 percentage points increase in the labour force, employed and unemployed and a decrease of 3.5 percentage points outside the labour force when we exclude the proxy interviews.

Contents

Preface	3
Abstract	4
1. Introduction	6
1.1. Overview of document	6
2. Main background variables	7
2.1. Genders	7
2.2. Age groups.....	8
2.3. Marital status.....	8
2.4. Student status	9
2.5. Other interesting background variables.....	10
2.6. Proportion of proxy interview for student status	10
3. Investigating interactions between variables	11
3.1. Age and gender.....	11
3.2. Age and student status	12
3.3. Age and marital status	14
4. Odds ratio for proxy interview	15
5. Non-response	17
6. Weighting for all respondents and for self-respondents only	17
7. Discussion	20
7.1. Final remarks	22
References	23
A. General description of odds ratio	24
List of figures	25
List of tables	26

1. Introduction

The main purpose of the Labour Force Survey is to give information on the development of employment status, such as employment and unemployment, in different population groups. The demographics could be age, gender, education, profession, industry etc.

The Norwegian Labour Force Survey (NLFS) covers all persons from 15 to 74 years who have their register residence in Norway for more than 6 months.

A random sample based on family unit is drawn, with a total of 14 500 households or about 24 000 persons interviewed every quarter. Each family member within the target age group gives an interview regarding their labour force activity for a specific reference week. The survey is a panel design and each person participates in total eight consecutive quarters. Post-interview, survey sample data is linked to register data at a micro level, using personal identification numbers found in both sources. The register data for employment is retrieved from the Norwegian Labour and Welfare Administration (NAV) and is called A-ordningen.

Data collection occurs through telephone interviews (CATI) and is normally self-respondent. However, close family members can respond on another person's behalf, which is called a proxy interview. This responding family member must be a spouse, or a parent or guardian.

In 2018, the pilot of a multimode NLFS with the use of web interview (CAWI) will not permit proxy interviews. In addition, the pilot survey will sample persons instead of family units. If this type of sampling and data collection is to be implemented on a larger scale, it is important to have a good understanding of who normally responds via proxy interview (the selection effect) and how this may affect the NLFS statistics.

1.1. Overview of document

Chapter 2 summarises the development of proxy interviews and their selection effects in the NLFS between 2006 and 2017 based on background variables from the survey or register data sources. These variables include age, gender, employment status, marital status and student status. Chapter 3 investigates the interactions between these variables.

Chapter 4 gives odds ratio for employment statuses, student statuses and genders for self-respondent and proxy interviews. Quarterly data is used from 2000, 2006 and 2017, with focus on the age group 15 to 29, but also in general.

Chapter 5 summarises the non-response for employment status in different population groups, including variables such as age, gender and the interaction of age and gender. The samples are based on auxiliary information from register data.

In Chapter 6, the employment status is divided in the same population groups as in Chapter 5. In the future, if we move from family to person selection, proxy interviews may not be allowed. We investigate this effect using new, calibrated weights for only self-respondents, and proxy interviews are assumed as non-response. The original weights and new weights are compared to see any affection to the employment status.

The following tables have a percentage for self-respondent and proxy interview for one whole year, and are an average of the four quarters within this year, to

summarise the distribution. This may give higher or lower than 100 per cent in total.

2. Main background variables

We compare those self-respondent and proxy interviews with variables including gender, age, employment status, student status, and marital status over the period from 2006 to 2017, with a total of 48 quarters. On average during this time, around 85 per cent are self-respondent and 15 per cent are proxy-respondent for approximately 20 000 observations every quarter, see Table 2.1.

Table 2.1 Distribution on average for interview type of approximately 20 000 observations per quarter

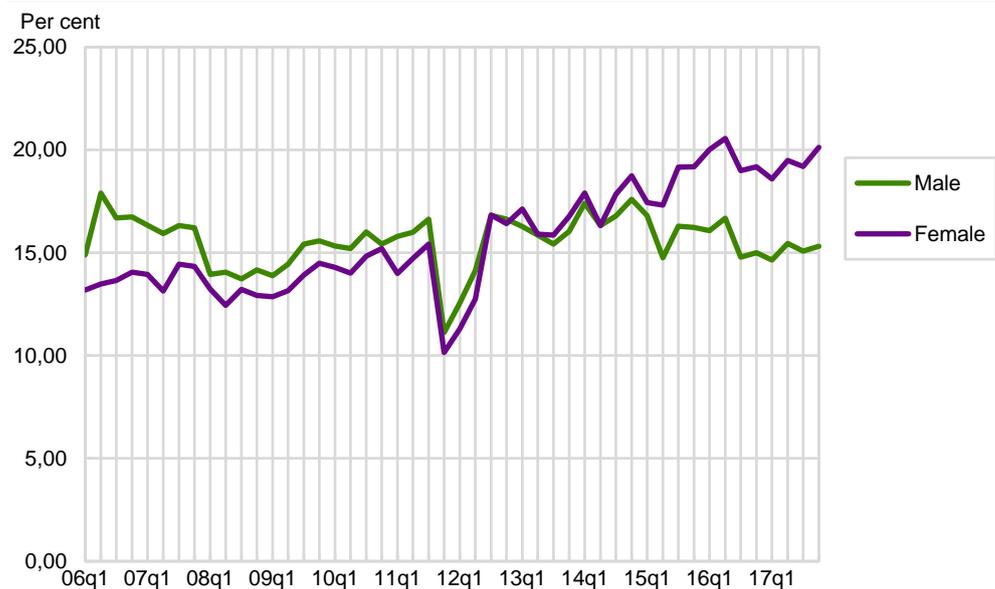
Interview type	Self-respondent interview	Proxy interview
Total observation	16 977	3 072
Percentage	84.66	15.34

Source: Statistics Norway

2.1. Genders

The development for females and males from 2006 until 2017 are shown in Figure 2.1. Both genders have approximately half of the total interviews. In 2006, there were no big differences between the two genders for self-respondent interviews. However, in recent years men self-respond more than women. At the end of 2017, female self-respondents have decreased by 2.83 percentage points from 2006, while male self-respondents have increased by 0.68 percentage points, see Table 2.2. The distribution of proxy interviews for females and males has swapped from 2006 to 2017 with an approximately even gender distribution around 2012, see Figure 2.1. Until around 2012, men were a bit more likely to have proxy interview, but in the recent years, women are more likely to have proxy interviews.

Figure 2.1 Comparison of proportion of proxy interviews by gender as a per cent



Source: Statistics Norway

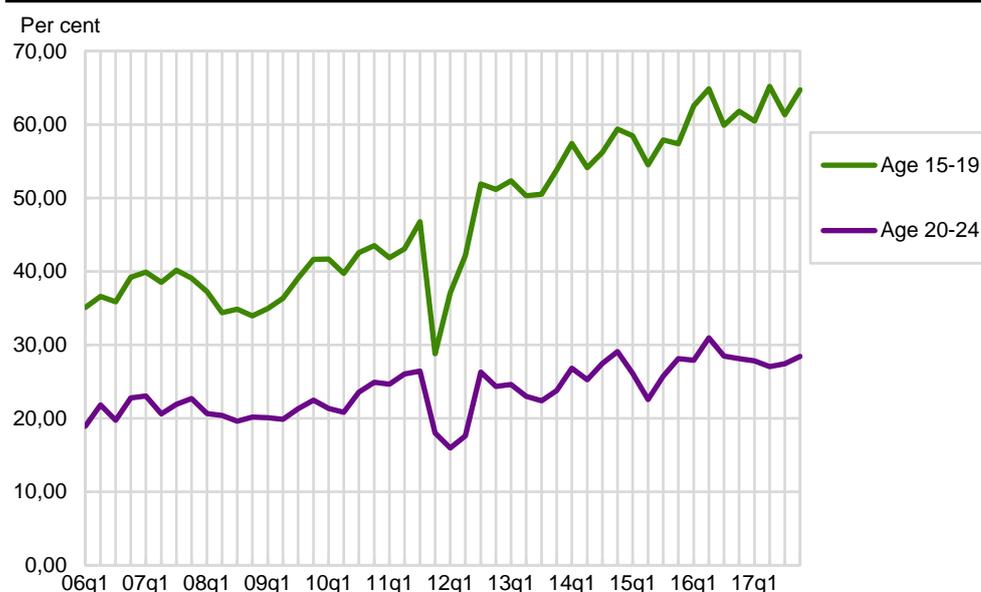
Table 2.2 Distribution of self-respondent and proxy interviews by gender as a per cent

Gender	Interview type	2006	2017
Male	Self-respondent interview	42.35	43.03
	Proxy interview	8.19	7.48
Female	Self-respondent interview	42.74	39.91
	Proxy interview	6.72	9.58

Source: Statistics Norway

2.2. Age groups

Persons who answer the NLFS are between the ages 15 to 74. In this analysis, we divide the respondents into 5-years age groups (12 groups total). For the age groups 15 to 19 and 20 to 24, they don't appear to follow the same trend compared to the rest, 25 to 74 years, whom have a steadier and more similar distribution to Table 2.1. The two youngest age groups, 15 to 19 and 20 to 24 years, have an increasing percentage of proxy interviews. Since around 2012, the youngest age group has more proxy interviews than self-respondent interviews. In Figure 2.2, the relation for self-respondent and proxy interviews are divided into the age groups 15 to 19 and 20 to 24 years. Around 40 per cent respond as proxy interviews in 2006 and around 60 per cent respond as self-respondent interviews, but by 2017 this had swapped. In 2017 the age group 15 to 24-year olds have almost 45 per cent of the proxy interviews (Table 2.3).

Figure 2.2 Comparison of proxy interviews by age groups 15-19 and 20-24 as a per cent

Source: Statistics Norway

Table 2.3 Distribution of self-respondent and proxy interviews by age groups as a per cent

Age group	Interview type	2006	2017
15-24	Self-respondent interview	12.25	9.15
	Proxy interview	5.10	7.62
25-54	Self-respondent interview	50.06	47.38
	Proxy interview	6.16	5.45
55-74	Self-respondent interview	22.77	26.41
	Proxy interview	3.66	3.99

Source: Statistics Norway

2.3. Marital status

Marital status is divided into four groups: single, married, partner and separated or divorced. Those with marital status as partner and separated or divorced don't usually have proxy interviews, while those with marital status as married and

single usually have a parent or guardian, or a spouse who is also sampled and can respond on their behalf. This can be seen in Table 2.4.

Table 2.4 Distribution of self-respondent and proxy interviews by marital status as a per cent

Marital status	Interview type	2006	2017
Single	Self-respondent interview	21.48	20.18
	Proxy interview	5.92	8.19
Married	Self-respondent interview	40.39	36.19
	Proxy interview	8.23	8.29
Partner	Self-respondent interview	14.34	17.80
	Proxy interview	0.44	0.35
Divorced or separated	Self-respondent interview	8.88	8.76
	Proxy interview	0.32	0.23

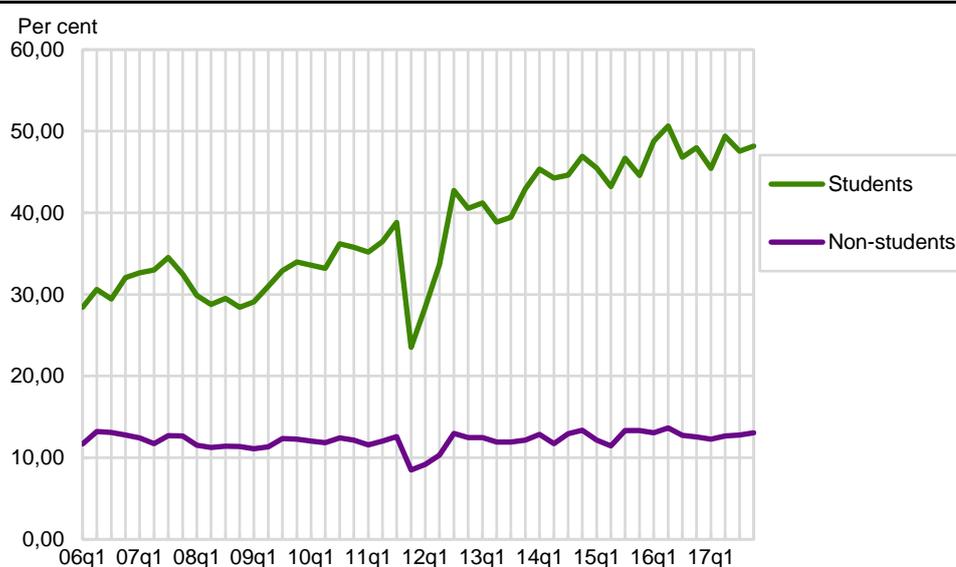
Source: Statistics Norway

In Table 2.4, single persons represent about 20 per cent of the self-respondent interviews, while married persons represent about 40 per cent of the self-respondent interviews in 2006, and married persons have decreased by 4.2 percentage points to 2017 in self-respondent. However, those with a partner increased by 3.46 percentage points from 2006 to 2017, from 14.34 per cent to 17.80 per cent. Single persons with proxy interviews have increased by 2.27 percentage points.

2.4. Student status

We also want to know how students respond, either by self-respondent or proxy. For non-students, they have the same trend as in Table 2.1, and represent the largest proportion of the respondents, see Table 2.5. Students, which already represent a small group, have almost half of the proxies in 2017. This has increased by approximately 20 per cent from 2006. Figure 2.3 shows the distribution for proxy interviews by students and non-students, and Table 2.5 shows the percentage of self-respondent and proxy-respondent for student and non-students in 2006 and 2017.

Figure 2.3 Comparison of proportion of proxy interviews by student status as a per cent



Source: Statistics Norway

Table 2.5 Distribution of self-respondent and proxy interviews by student status as a per cent

Student status	Interview type	2006	2017
Students	Self-respondent interview	8.87	6.54
	Proxy interview	3.83	5.95
Non-students	Self-respondent interview	76.22	76.40
	Proxy interview	11.08	11.11

Source: Statistics Norway

2.5. Other interesting background variables

Other interesting background variables include region in Norway, citizenship and the country of origin.

Norway is divided into 7 geographical regions. We found no major difference between self-respondent and proxy interview percentages among these regional groups, but the Oslo-area has a higher percentage of self-respondents compared to the rest of the country.

Two remaining variables that we have investigated are citizenship and country of origin. These variables were divided into three categories: Norwegian, western countries, and not-western countries. For country of origin, this may be country of birth or country of birth/citizenship for parents or grandparents. Comparing the two variables, citizenship and country of origin, there is no significant differences between their three categories. However, not surprisingly, self-respondent is highest for Norwegians, and less for western-countries and not-western countries. This may be due to language barriers.

2.6. Proportion of proxy interview for student status

The effect in proxy interviews on employment controlling for age and register-employment for students and non-students were investigated separately by Solheim, Håland & Lagerstrøm (2001). In 2000, the participants in the survey were 16 to 74 years compared with 15 to 74 years in more recent years. We would like to make a comparison between years 2000, 2006 and 2017 on student status for self-respondent interviews compared to proxy interviews with the age group 15 to 29.

Table 2.6 shows the proportion of students and non-students who gave proxy interviews for age groups 16 to 29 years in 2000 and 15 to 29 years in 2006 and 2017. In year 2000, there were about twice as many proxy interviews for students compared to non-students, and it was about 30 per cent proxy interviews for students and around half for non-students. From 2000 until 2006, the proportion of students who answer as a proxy did not change. However, from 2006 to 2017 the student proportion increased majorly, and approximately half of the students responded as proxy.

Table 2.6 Proportion of proxy interview for students and non-students in 2000, 2006 and 2017 for age group 15-29 as a per cent

Year	Student status	1 st quarter	2 nd quarter	3 rd quarter	4 th quarter
2000	Students	33.70	29.30	30.90	30.00
	Non-students	19.40	14.90	16.10	14.80
2006	Students	30.13	32.64	31.18	33.53
	Non-students	14.47	16.16	16.66	16.90
2017	Students	47.50	52.08	49.63	50.43
	Non-students	19.17	18.19	19.34	19.59

Source: Statistics Norway

3. Investigating interactions between variables

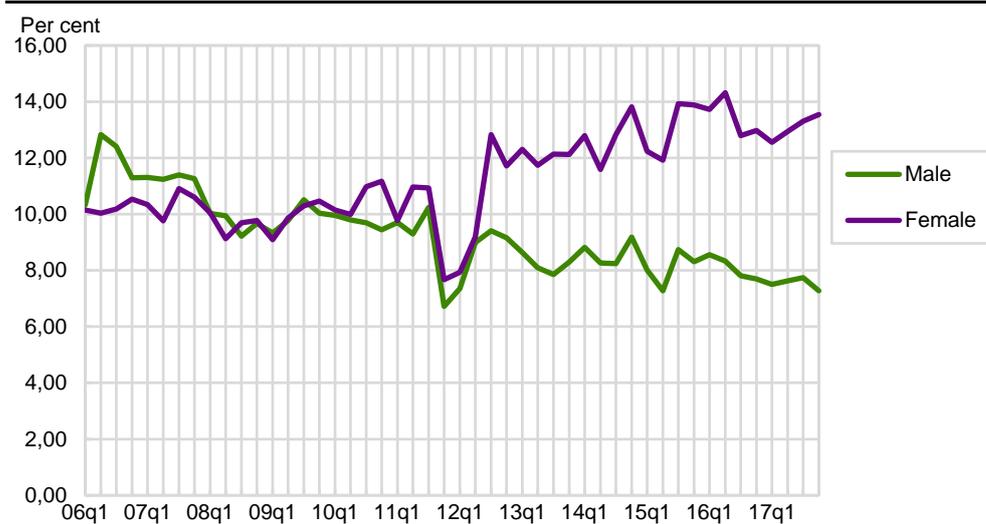
In addition to investigating variables directly, we also look at interactions between age and gender, age and student status, and age and marital status. The cross-variables are divided in broader groups to reduce the number of categories and avoid groups with too small observation numbers. Categories which had similar trends were combined.

3.1. Age and gender

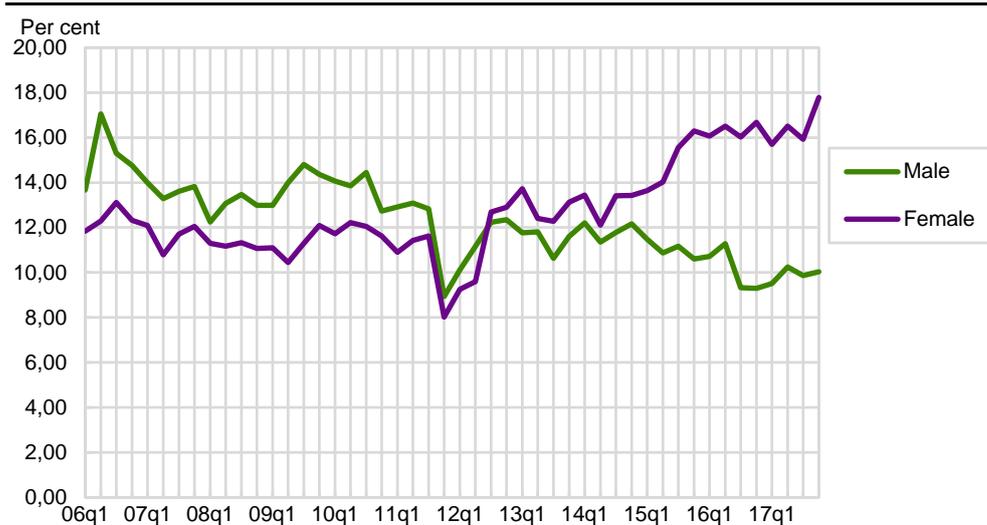
Age and gender were crossed by dividing age into three broader groups, 15 to 24, 25 to 54 and 55 to 74 years, and genders, female and male. Gender does not play a role for the age group 15 to 24 when it comes to responding for themselves or by another person. Persons aged 25 to 54 and 55 to 74 have all similar trends until the year 2012, see Figure 3.1, and show small difference in gender.

For both gender in age group 15 to 24 years, see Table 3.1, the proxy interview rate increased by 1.11 percentage points for men and 1.42 for women. However, for person aged 25 to 54 years, proxy-respondent rates increased by 0.58 percentage points for women, while men decreased by 1.3 percentage points for this same age group. This is reflected in Figure 3.1.

Figure 3.1 Comparison of proportion of proxy interviews by gender for the age group 25-54 as a per cent



Source: Statistics Norway

Figure 3.2 Comparison of proportion of proxy interviews by gender for the age group 55-74 as a per cent

Source: Statistics Norway

Table 3.1 Distribution of self-respondent and proxy interviews by gender and age groups 15-24, 25-54 and 55-74 as a per cent

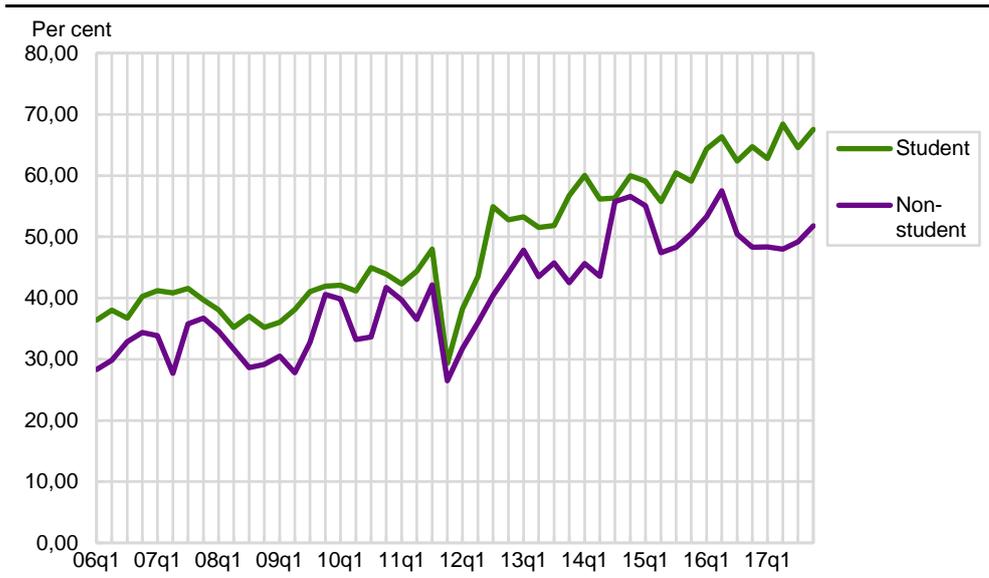
Gender	Age group	Interview type	2006	2017
Male	15-24	Self-respondent interview	6.13	4.69
		Proxy interview	2.85	3.96
	25-54	Self-respondent interview	24.77	24.42
		Proxy interview	3.29	1.99
Female	15-24	Self-respondent interview	11.46	13.91
		Proxy interview	2.05	1.53
	25-54	Self-respondent interview	6.12	4.46
		Proxy interview	2.24	3.66
55-74	Self-respondent interview	25.29	22.95	
	Proxy interview	2.88	3.46	
55-74	Self-respondent interview	11.32	12.49	
	Proxy interview	1.60	2.46	

Source: Statistics Norway

3.2. Age and student status

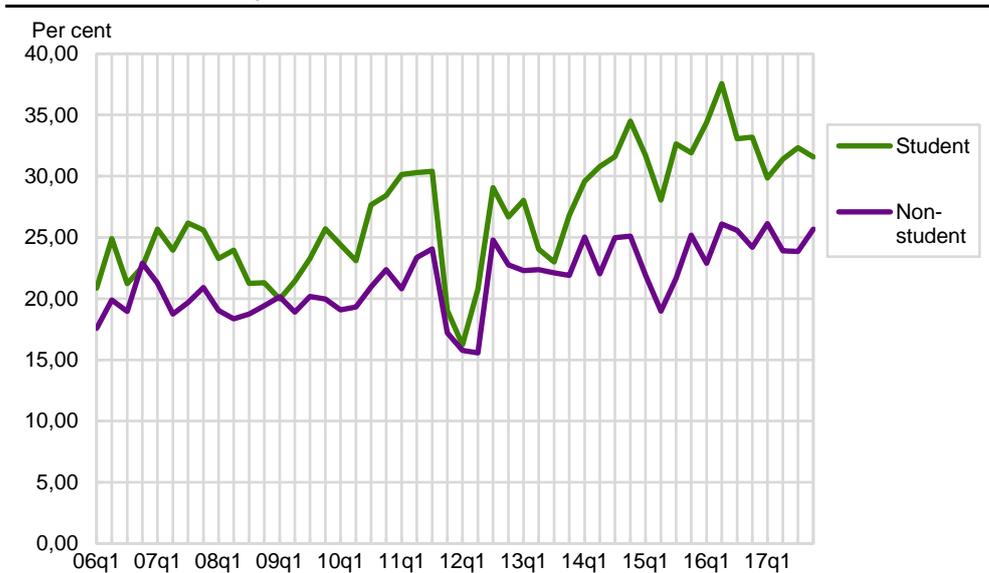
Secondly, students and non-students appear to differ in their percentage of self-respondent and proxy interviews. In Table 3.2 we look at three age groups, 15 to 24 years, 25 to 54 years, and 55 to 74 years. Figure 3.3 and Figure 3.4 show the comparison for self-respondent and proxy interview of students and non-students. For both figures, the proxy interviews for students are higher than non-students. Specially, in Figure 3.4, for students aged 15 to 19-years follows the same increasing trend such as in Figure 2.2 and Figure 2.3.

Figure 3.3 Comparison of proportion of proxy interviews by student status for the age group 15-19 as a per cent



Source: Statistics Norway

Figure 3.4 Comparison of proportion of proxy interviews by student status for the age group 20-24 as a per cent



Source: Statistics Norway

Table 3.2 Distribution of self-respondent and proxy interviews by student status and age groups 15-24, 25-54 and 55-74 as a per cent

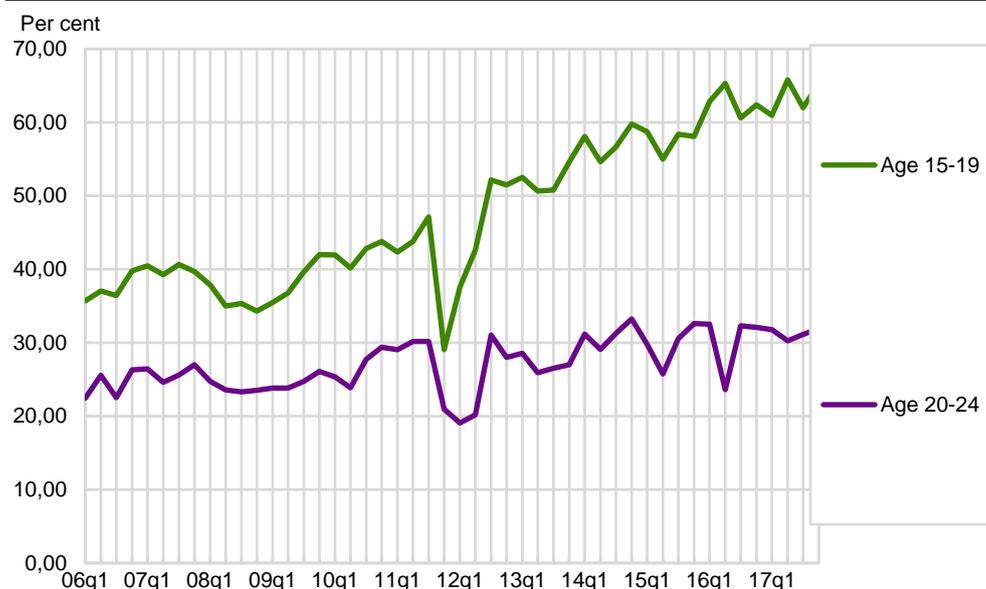
Student status	Age group	Interview type	2006	2017
Students	15-24	Self-respondent interview	7.15	4.90
		Proxy interview	3.59	5.72
	25-54	Self-respondent interview	1.66	1.61
		Proxy interview	0.24	0.22
	55-74	Self-respondent interview	0.06	0.03
		Proxy interview	0.00	0.01
Non-students	15-24	Self-respondent interview	5.10	4.25
		Proxy interview	3.59	5.72
	25-54	Self-respondent interview	48.40	45.77
		Proxy interview	0.24	0.22
	55-74	Self-respondent interview	22.71	26.37
		Proxy interview	0.00	0.01

Source: Statistics Norway

3.3. Age and marital status

Thirdly, for the age group 15 to 24 years, most are single and respond by proxy. This is shown in Table 3.3 for single and married for broad age-groups. Partner and separated or divorced is not shown in this table. Figure 3.5 shows a higher percentage for 15 to 19-year olds in proxy interview and follows the same trend as previous figures. For 15 to 19-year olds the proxy interviews are increasing, while 20 to 24-year olds are steadier.

Figure 3.5 Comparison of proportion of proxy interviews by single for the age groups 15-19 and 20-24 as a per cent



Source: Statistics Norway

Table 3.3 Distribution of self-respondent and proxy interviews by marital status and age groups 15-19 and 20-24 as a per cent

Marital status	Age group	Interview type	2006	2017
Single	15-19	Self-respondent	5.72	3.03
		Proxy	3.40	5.28
	20-24	Self-respondent	4.48	4.54
		Proxy	1.43	2.07
	25-54	Self-respondent	9.86	10.33
		Proxy	0.93	0.74
55-74	Self-respondent	2.86	4.56	
	Proxy	0.32	0.20	
Married	15-19	Self-respondent	0.02	0.00
		Proxy	0.01	0.00
	20-24	Self-respondent	0.28	0.15
		Proxy	0.08	0.06
	25-54	Self-respondent	24.65	20.07
		Proxy	4.91	4.51
55-74	Self-respondent	15.43	15.97	
	Proxy	3.23	3.72	

Source: Statistics Norway

4. Odds ratio for proxy interview

We would like to make a comparison between the years 2000, 2006 and 2017, for both odds ratio on employment for self-respondent interviews compared to proxy interviews as well as other combinations to investigate long-term trends in possible measurement error on the key labour market variables. The results for year 2000 are investigated by Solheim et al. (2001). The age groups for this chapter are 15 to 29 years and all (15 to 74 years).

In the survey, there is more than one variable that describes the employment status, but we use the variable that has a 3 digits status code: employed, unemployed and outside the labour force. This variable is summed into 2 classes: persons who have a job (employed) and persons who do not have a job (unemployed).

We also have access to register employment status from administrative data with classes such as register employed, register self-employed, register unemployed and register outside the labour force etc., but this variable is grouped into two categories register employed and register unemployed as well.

In Appendix A, a general description of odds ratio and how the results can be interpreted is provided. In brief, it provides a measure of the ratio of self-respondent and proxy interviews for employed and can indicate whether there may be measurement errors caused by interview type.

We would like to see what impact the interview types (self-respondent and proxy) have on employment for persons aged 15 to 29 years. The results are compared between 2000, 2006 and 2017 for both students and non-students, see Table 4.1, and divided into gender in Table 4.2. Also, the odds ratios for genders in self-respondent interviews compared to proxy interviews are given in Table 4.3.

In Table 4.1, we look at register employment for students in 2017 for the 1st quarter. The odds ratio is 2.25, which means that the odds for register-employment is 2.25 times higher for self-respondent persons in the NLFS compared to proxy responding-persons in the NLFS. When comparing the odds ratios for register and survey, register values are higher than from the survey. In general, the register-employment for students with proxy interviews is about 55 per cent of the employment estimated by self-respondent interviews in 2000. The effect for non-students is less, the employment ratio sampled by proxy interview is about 65 per cent of the employment ratio estimated by self-respondent interview. However, by 2017, this has changed. The employment for students sampled by self-respondent interviews is more than double of the employment ratio for proxy interview. This varies a lot and is similar for non-students as well.

Table 4.1 Odds ratio for employment for self-respondent interviews compare to proxy interviews for students and non-students in 2000, 2006 and 2017 for persons aged 15-29

Register/Survey	Year	Student status	1 st quarter	2 nd quarter	3 rd quarter	4 th quarter
Register	2000	All	-	-	-	-
		Students	0.54	0.57	0.60	0.58
		Non-students	0.66	0.65	0.72	0.63
	2006	All	2.29	1.93	2.07	2.26
		Students	1.61	1.27	1.40	1.52
		Non-students	1.78	1.46	1.85	1.85
	2017	All	3.49	3.56	5.13	3.70
		Students	2.25	2.06	2.23	2.08
		Non-students	2.30	2.18	1.91	2.85
Survey	2006	All	0.98	1.07	0.84	1.24
		Students	0.72	0.78	0.38	0.74
		Non-students	1.08	1.04	1.19	1.56
	2017	All	2.13	2.17	1.49	1.70
		Students	1.78	1.59	1.41	1.31
		Non-students	1.46	1.11	0.70	1.45

Source: Statistics Norway

Table 4.2 Odds ratio for register employment for self-respondent interviews compare to proxy interviews for female and male in 2006 and 2017

Gender	Year	Age group	Student status	1 st quarter	2 nd quarter	3 rd quarter	4 th quarter
Female	2006	All	All	2.24	1.83	1.74	1.98
			Students	1.68	1.22	1.33	1.43
			Non-students	1.77	1.48	1.44	1.48
		15-29	All	2.53	2.04	2.25	2.27
			Students	1.76	1.25	1.35	1.41
			Non-students	2.12	1.70	2.20	1.88
	2017	All	All	3.06	3.10	2.59	3.00
			Students	2.35	2.13	1.87	2.23
			Non-students	1.78	1.71	1.63	1.60
		15-29	All	3.79	3.99	3.29	3.96
			Students	2.44	2.21	1.98	2.23
			Non-students	2.41	2.04	2.09	2.70
Male	2006	All	All	1.95	1.78	1.89	2.04
			Students	1.42	1.24	1.30	1.49
			Non-students	1.63	1.46	1.66	1.70
		15-29	All	2.06	1.84	1.90	2.30
			Students	1.48	1.24	1.41	1.56
			Non-students	1.46	1.27	1.53	2.01
	2017	All	All	2.16	2.16	1.96	2.15
			Students	2.06	1.98	2.33	1.95
			Non-students	1.59	1.61	1.50	1.65
		15-29	All	3.20	3.17	3.01	3.45
			Students	2.10	1.93	2.43	1.96
			Non-students	2.23	2.37	1.73	3.17

Source: Statistics Norway

Table 4.3 Odds ratio for gender for self-respondent interviews compare to proxy interviews for students and non-students in 2006 and 2017 for age group 15-29

Year	Student status	1 st quarter	2 nd quarter	3 rd quarter	4 th quarter
2006	All	0.84	0.78	0.87	0.83
	Students	0.84	0.80	0.88	0.81
	Non-students	0.64	0.60	0.72	0.70
2017	All	0.97	0.99	0.97	0.96
	Students	0.88	0.88	0.84	0.89
	Non-students	0.83	0.94	0.90	0.82

Source: Statistics Norway

5. Non-response

Non-response can tell us about the representativeness of the observed data, and the effect on the estimates and results. The information on non-respondents comes from auxiliary variables from register data. We want to look at the non-response for population groups such as age groups, gender, and the combination of them.

For each quarter in 2006 and 2017, the information of observed data for respondents and non-respondents can be seen in Table 5.1. The respondents include self-respondent and proxy interviews. Between these two years, the non-respondents increased as well as the sample.

Table 5.1 Information of observations in the survey for respondents and non-respondents in years 2006 and 2017

Year	Quarter	All	Respondents	Non-respondents
2006	1 st quarter	24 556	20 745	3 811
	2 nd quarter	24 474	20 984	3 490
	3 rd quarter	24 453	21 300	3 153
	4 th quarter	24 465	21 461	3 004
2017	1 st quarter	24 654	20 150	4 504
	2 nd quarter	24 669	20 225	4 444
	3 rd quarter	24 648	20 535	4 113
	4 th quarter	24 636	20 512	4 124

Source: Statistics Norway

Table 5.2 includes the distribution of non-response, self-respondents, and proxy-respondents, and the response includes both self-respondent and proxy interviews. The population groups are divided for all (no sub-group), gender, age group, and gender and age group combined. The values for this table are an average of all four quarters in year 2017. The non-response is on average 17.43 per cent in 2017. However, for non-response, they have a relatively small difference between sub-groups.

Table 5.2 Distribution on non-response, self-respondent and proxy for average of quarter in 2017 for gender, age group and gender and age group combined as a per cent

Gender	Age group	Non-response	Self-respondent	Proxy
-	-	17.43	68.43	14.13
Male	-	17.55	70.22	12.22
Female	-	15.56	68.09	16.34
-	15-24	16.35	45.76	37.86
-	25-54	19.27	72.40	8.32
-	55-74	11.59	76.67	11.74
Male	15-24	15.98	45.67	38.31
	25-54	20.88	73.16	5.96
	55-74	12.17	78.99	8.84
Female	15-24	16.74	45.85	37.39
	25-54	17.61	71.60	10.78
	55-74	10.98	74.25	14.77

Source: Statistics Norway

6. Weighting for all respondents and for self-respondents only

All respondents have a weight in the survey, which will give an estimate of the total population. The weight for each participant in this survey shows how many persons they represent in the population. The estimation methodology for calibration and weighting can be further studied in Oguz-Alper (2018). This is important for estimating employment status from the NLFS. The weights are given after the non-response is removed. The two-main focuses are: 1) the number of

observations for the employment status, and 2) the total population given by the weights for the employment status. Here we are looking at employed, unemployed, the labour force and outside the labour force.

Further, the test-run of a multi-mode data collection for the labour force survey with the use of CAWI will not permit proxy interviews. To investigate how this might affect estimates, we use NLFS data and treat proxy interviews as part of the non-response, and using new weights calibrated on the self-respondents only. These new weights will tell us about the bias of proxy interviews as well as the non-response, and how well our calibration is able to adjust for the increased non-response. In the tables below, the weights for self-respondents only, with proxy interviews as part of the non-response are called *new weights*. The weights that include both self-respondents and proxies are called *original weights*.

In Table 6.1, Table 6.2, Table 6.3 and Table 6.4, each table represents the survey employment status: in the labour force, employed, outside the labour force, and unemployed. Each table compare the percentage for original weights, number of observations for all respondents and for self-respondents only, and new weights. Employed, in the labour force and outside the labour force are divided by the total population, while unemployed is divided by the labour force.

In general, the labour force, employed and unemployed increased after removing proxy interviews and giving them new weights, while outside the labour force decreased. For the youngest group 15 to 24-year olds the labour force, employed and unemployed increased by 3.55, 2.43 and 1.26 percentage points respectively, and outside the labour force decreased by 3.55 percentage points.

Table 6.1 The labour force rate – Comparing the percentage using the original weights, unweighted and new weights by annual average of 2017.

Gender	Age group	Original weights	Unweighted	New weights	Unweighted (self-respondents only)
-	-	69.67	70.83	70.71	74.87
Male	-	72.20	73.01	73.19	77.06
Female	-	67.04	68.60	68.12	72.51
-	15-24	54.56	54.84	58.11	67.23
-	25-54	85.62	88.10	86.24	89.41
-	55-74	48.01	49.71	48.36	51.46
Male	15-24	54.67	55.28	57.97	68.20
	25-54	87.71	90.48	88.35	91.58
	55-74	52.39	53.13	52.66	54.59
Female	15-24	54.45	54.37	58.25	66.20
	25-54	83.42	85.72	84.01	87.11
	55-74	43.61	46.17	44.04	47.98

Source: Statistics Norway

Table 6.2 Employment rate – Comparing the percentage using the original weights, unweighted and new weights by annual average of 2017.

Gender	Age group	Original weights	Unweighted	New weights	Unweighted (self-respondents only)
-	-	66.73	68.36	67.55	72.53
Male	-	68.85	70.29	69.58	74.50
Female	-	64.52	66.38	65.42	70.41
-	15-24	48.67	49.37	51.10	61.16
-	25-54	82.46	85.57	83.01	86.88
-	55-74	47.20	49.00	47.59	50.76
Male	15-24	48.06	49.18	50.11	61.37
	25-54	84.21	87.84	84.77	88.91
	55-74	51.28	52.17	51.58	53.66
Female	15-24	49.32	49.58	52.16	60.93
	25-54	80.62	83.30	81.16	84.73
	55-74	43.11	45.72	43.59	47.53

Source: Statistics Norway

Table 6.3 Outside the labour force rate – Comparing the percentage using the original weights, unweighted and new weights by annual average of 2017.

Gender	Age group	Original weights	Unweighted	New weights	Unweighted (self-respondents only)
-	-	30.33	29.17	29.29	25.13
Male	-	27.80	26.99	26.81	22.94
Female	-	32.96	31.40	31.88	27.49
-	15-24	45.44	45.16	41.89	32.77
-	25-54	14.38	11.90	13.76	10.59
-	55-74	51.99	50.29	51.64	48.54
Male	15-24	45.33	44.72	42.03	31.80
	25-54	12.29	9.52	11.65	8.42
	55-74	47.61	46.87	47.34	45.41
Female	15-24	45.55	45.63	41.75	33.80
	25-54	16.58	14.28	15.99	12.89
	55-74	56.39	53.83	55.96	52.02

Source: Statistics Norway

Table 6.4 Unemployment rate – Comparing the percentage using the original weights, unweighted and new weights by annual average of 2017.

Gender	Age group	Original weights	Unweighted	New weights	Unweighted (self-respondents only)
-	-	4.22	3.49	4.47	3.13
Male	-	4.65	3.73	4.93	3.32
Female	-	3.74	3.22	3.95	2.90
-	15-24	10.80	9.97	12.06	9.03
-	25-54	3.69	2.88	3.75	2.83
-	55-74	1.69	1.43	1.59	1.36
Male	15-24	12.09	11.03	13.56	10.01
	25-54	3.98	2.93	4.05	2.93
	55-74	2.12	1.81	2.05	1.70
Female	15-24	9.40	8.81	10.45	7.98
	25-54	3.36	2.82	3.40	2.73
	55-74	1.15	0.97	1.04	0.94

Source: Statistics Norway

7. Discussion

It is important to keep the response rate high for data quality. Self-respondent interviews in the NLFS are likely to give the highest data quality. However, when self-respondents are difficult to contact, a proxy interview can substitute the self-respondent to keep a high response rate. A low response rate can interfere with the data quality, as much as high proxy interview. Therefore, it has been a goal to maintain the proxy interview rates at around 15 per cent.

Two main reasons that can affect the proxy interview rates and give instabilities are: Teenagers or young adults who are living in their parents' home, and whom have a good overview of their schedule and working agreement. This applies for spouses whom have one working agreement as well.

In this analysis, proxy interviews were more prevalent among the younger groups. In 2006, persons aged 15 to 19 years gave more self-respondent interviews than proxy interviews, but in 2017 it was the opposite, and they swapped at around 2012. Young students respond more often by proxy, whom usually are single and haven't settled yet. Many young persons have a mobile phone, but only respond to the persons in their contact list. Solheim et al. (2001) have a hypothesis that 15 to 19-year olds still live at home, and parents can answer on their behalf while having a good overview of their schedule and working agreement, while 20 to 24-year olds and 25 to 29-year olds have higher chance of living by themselves. If they are students and not living at home, parents may not have a good overview of their working agreement, and more difficult to answer correctly. Thomsen & Villund (2011) have an assumption that a considerable number of those interviewed by proxy are less reachable or completely unreachable with reasonable time, and it's better if someone answers on their behalf as a proxy rather than non-response. Non-response and proxy response are associated with young, urban and of foreign origin.

Another interesting finding is that married women in the age of 25 to 54 years and 55 to 74 years respond more by proxy than men. This age group has more married persons compare to the younger generation. One reason could be that the survey is given by telephone, and will first call the eldest respondent in the family. Generally, for married persons this is a man, which means the first person the survey contacts are men, and the men can respond on their behalf.

The employment status was classified into four classes: employed, unemployed, in the labour force and outside the labour force. We made a comparison between statistics based on the original weights and the new weights. In general, after removing proxy interviews, the labour force rate, employment rate and unemployment rates increased, and the proportion outside the labour force decreased. However, for the youngest age group 15 to 24 years for both genders, the labour force rate, employment rate and unemployment rate increased quite a bit, but the proportion outside the labour force decreased.

Thomsen et al. (2011) looked at three different estimates of employment rate under the proxy at random (PAR) model, to see a clearer picture of whether proxy interviews should be included or not. The three estimates are based on self-respondent interviews, both self-respondent and proxy interviews and only self-respondent but adjusted by post-stratification using register-based employment. The results are an overestimation for self-respondent interviews, due to selection effect where parents answering on their children's behalf, and an underestimation for all interviews (self-respondent and proxy interviews). At last, they observed the effect of non-response, only the self-respondent estimate has an even larger bias.

Including proxy interviews have about the same effect as post-stratification of the self-respondents by using register employment as post-stratification variable. Extending the response sample size by including proxy interviews introduces some underreporting, but gives a more representative response sample. The study was based on binominal variables, for instance, register employment and register unemployment.

Zhang, Thomsen & Kleven (2013) give an understanding on what impact auxiliary data (register data) and proxy interviews have on a survey and estimates of employment status. They write about the systematic difference between those who are 'easy-to-reach' respondents versus those who are 'hard-to-reach' respondents. The proxy interviews provide data on some 'hard-to-reach' persons who have a labour-market situation more like those who are not reached at all, which will probably result in a better employment rate estimates even though they introduce some underreporting. These respondents have approximately the same effect as post-stratification of the self-responses, using register-employment status as the auxiliary variable, which will give a good check on the collected values. The register data provide a richer source of relevant auxiliary information, in addition to data collected in previous surveys and censuses. Due to their importance in effective dealings with non-sampling errors, one should make every effort to increase their availability in the statistical system, and at the same time, develop efficient statistical methods that capitalise on the combined data sources. After the data have been collected, re-weighting and imputation values of missing values have to account for non-response are practically unavoidable. Auxiliary data are again necessary, without which one would have to assume that the data are a missing completely at random (MCAR), to carry out the only feasible adjustment.

Kleven, Lagerstrøm & Thomsen (2008) drew the same conclusion. Proxy interviews reduce non-response errors, but this can be lost in term of the 'total survey error' by the fact that proxies give more incorrect answers. Proxy interviews can give lower variance in the estimates and less bias since the persons interviewed by proxy is believed to be systematically different from those interviewed directly. A disadvantage in using proxy interviews can be that the answers given on behalf of another person can be more incorrect than given from the respondent. However, if post-stratification is used in the estimation, it seems that proxy interviewing should be avoided.

Statistics Norway has got better register data for estimation over the years through A-ordningen and can predict good for a larger group of employments in the NLFS. However, for the two important labour force categories such as unemployment and persons outside the labour force, there are less good administrative data for estimation.

Lemaitre (1988) examined a re-interviewed subsample on response errors in the Canadian LFS. The initial interview and re-interview allowed proxy response as well as self-response, and the combined data contains some individual units about the same reference week. Of interest here are the topics that were important to classify a person as employed: 'Had a job, did not work', and 'Worked during reference week', there was more inconsistency when the interview and re-interview were of two different types (one self-respondent and one proxy). Inconsistency between different type interviews can be attributed to proxy error. However, if both were self-respondents or both were proxies, the inconsistency was about half than of two different interview types. Although proxy interviews introduce some measurement errors, the study reveals that two self-respondent interviews also produced up to 5 per cent inconsistency in answers. The study finds more inconsistency when both are proxies than when both are self-respondents. This

indicates lower reliability in proxy interviews than in self-respondent ones. The inconsistency is lowest when both interviews are self-response, and this indicates high reliability.

7.1. Final remarks

In conclusion, proxy interviews have a major influence on the NLFS. Without proxy interviews, and assuming this group is non-respondents, our calibration model is not able to adjust to produce the same estimates for employment and unemployment as those published today. This is especially the case for the youngest age group.

References

- Bø, T. P. & Håland, I. (2015). *Dokumentasjon av Arbeidskraftundersøkelsen (AKU) etter omleggingen i 2006* (Documents 2015/15), Statistics Norway. <https://www.ssb.no/arbeid-og-lonn/artikler-og-publikasjoner/attachment/218011?ts=14bdf041650>
- Kleven, Ø., Lagerstrøm, B. O. & Thomsen, I. (2008). A simple model for studying the effects of proxy interviewing: The Norwegian Labour Force Survey as a Case Study (Documents, 2008/2), Statistics Norway. https://www.ssb.no/a/english/publikasjoner/pdf/doc_200802_en/doc_200802_en.pdf
- Lemaitre, G. (1988). A Look at Response Errors in the Labour Force Survey. *The Canadian Journal of Statistics*, 16, pp.127-141.
- Lohr, S. L. (2010). *Sampling: Design and Analysis* (2nd edition). Brooks/Cole cengage learning. International Edition.
- Oguz-Alper, M. (2018). *New estimation methodology for the Norwegian Labour Force Survey* (Documents 2018/16), Statistics Norway. <https://www.ssb.no/arbeid-og-lonn/artikler-og-publikasjoner/attachment/346996?ts=162d85820c8>
- Solheim, L., Håland, I. & Lagerstrøm, B. O. (2001). Proxy interview and measurement error of employment for the age group 16-29: The Norwegian Labour Force Survey, Statistics Norway. Paper presented at 12th International Workshop of Household Survey Non-response, Oslo.
- Thomsen, I. & Villund, O. (2011). Using Register Data to Evaluate the Effects of Proxy Interviews in the Norwegian Labour Force Survey. *Journal of Official Statistics*, 27(1), pp.87-98.
- Villund, O. (2010). *Evaluating employment classification: A quality study linking survey data and register data* (Documents 2010/9), Statistics Norway. https://www.ssb.no/a/english/publikasjoner/pdf/doc_201009_en/doc_201009_en.pdf
- Zhang, L.-C., Thomsen, I. & Kleven, Ø. (2013). On the Use of Auxiliary and Paradata for Dealing with Non-Sampling Errors in Household Surveys. *International Statistical Review*, 81(2), pp.270-288.

A. General description of odds ratio

An odds ratio measures the relation between an exposure and an outcome. This means that the odds ratio represents the odds that an outcome will occur given a particular exposure, compared to the odds of the outcome occurring in the absence of that exposure. This is most common in case-control studies, but also in cross-sectional and cohort study designs as well.

Odds ratios are used to compare the relative odds of the occurrence of the outcome of interest, given exposure to the variable of interest (in our case it could be students, gender, interview type etc). The odds ratio can also be used to determine whether a particular exposure is a risk factor for a particular outcome, and to compare the magnitude of various risk factor for that outcome.

Definition of three types of result for odds ratio:

- Odds ratio = 1: Exposure does not affect odds of outcome
- Odds ratio > 1: Exposure associated with higher odds of outcome
- Odds ratio < 1: Exposure associated with lower odds of outcome

Example:

Odds ratio can be calculated by a two-by-two frequency table

		Outcome	
		Cases	Non-cases
Exposure	Exposed	a	b
	Non-exposed	c	d

Where

- a – Number of exposed cases
- b – Number of exposed non-cases
- c – Number of unexposed cases
- d – Number of unexposed non-cases

The formula of odds ratio is then:

$$OR = \frac{\text{odds}(\text{cases}|\text{exposed})}{\text{odds}(\text{cases}|\text{non-exposed})} = \frac{\frac{a}{a+b}}{\frac{c}{c+d}} \bigg/ \frac{\frac{b}{a+b}}{\frac{d}{c+d}} = \frac{a/b}{c/d} = \frac{ad}{bc}$$

If $OR > 1$, then the odds of exposed cases is $(OR - 1) \times 100$ per cent higher than the odds of the non-exposed cases. If $OR < 1$, then the odds of exposed cases is $(1 - OR) \times 100$ per cent lower than the odds of the non-exposed cases.

List of figures

Figure 2.1	Comparison of proportion of proxy interviews by gender as a per cent	7
Figure 2.2	Comparison of proxy interviews by age groups 15-19 and 20-24 as a per cent	8
Figure 2.3	Comparison of proportion of proxy interviews by student status as a per cent	9
Figure 3.1	Comparison of proportion of proxy interviews by gender for the age group 25-54 as a per cent	11
Figure 3.2	Comparison of proportion of proxy interviews by gender for the age group 55-74 as a per cent	12
Figure 3.3	Comparison of proportion of proxy interviews by student status for the age group 15-19 as a per cent	13
Figure 3.4	Comparison of proportion of proxy interviews by student status for the age group 20-24 as a per cent	13
Figure 3.5	Comparison of proportion of proxy interviews by single for the age groups 15-19 and 20-24 as a per cent	14

List of tables

Table 2.1	Distribution on average for interview type of approximately 20 000 observations per quarter	7
Table 2.2	Distribution of self-respondent and proxy interviews by gender as a per cent	8
Table 2.3	Distribution of self-respondent and proxy interviews by age groups as a per cent	8
Table 2.4	Distribution of self-respondent and proxy interviews by marital status as a per cent	9
Table 2.5	Distribution of self-respondent and proxy interviews by student status as a per cent	10
Table 2.6	Proportion of proxy interview for students and non-students in 2000, 2006 and 2017 for age group 15-29 as a per cent.....	10
Table 3.1	Distribution of self-respondent and proxy interviews by gender and age groups 15-24, 25-54 and 55-74 as a per cent.....	12
Table 3.2	Distribution of self-respondent and proxy interviews by student status and age groups 15-24, 25-54 and 55-74 as a per cent.....	13
Table 3.3	Distribution of self-respondent and proxy interviews by marital status and age groups 15-19 and 20-24 as a per cent.....	14
Table 4.1	Odds ratio for employment for self-respondent interviews compare to proxy interviews for students and non-students in 2000, 2006 and 2017 for persons aged 15-29.....	16
Table 4.2	Odds ratio for register employment for self-respondent interviews compare to proxy interviews for female and male in 2006 and 2017	16
Table 4.3	Odds ratio for gender for self-respondent interviews compare to proxy interviews for students and non-students in 2006 and 2017 for age group 15-29.....	16
Table 5.1	Information of observations in the survey for respondents and non-respondents in years 2006 and 2017.....	17
Table 5.2	Distribution on non-response, self-respondent and proxy for average of quarter in 2017 for gender, age group and gender and age group combined as a per cent.....	17
Table 6.1	The labour force rate – Comparing the percentage using the original weights, unweighted and new weights by annual average of 2017.	18
Table 6.2	Employment rate – Comparing the percentage using the original weights, unweighted and new weights by annual average of 2017.....	19
Table 6.3	Outside the labour force rate – Comparing the percentage using the original weights, unweighted and new weights by annual average of 2017... ..	19
Table 6.4	Unemployment rate – Comparing the percentage using the original weights, unweighted and new weights by annual average of 2017.	19

Statistics Norway

Postal address:
PO Box 8131 Dept
NO-0033 Oslo

Office address:
Akersveien 26, Oslo
Oterveien 23, Kongsvinger

E-mail: ssb@ssb.no
Internet: www.ssb.no
Telephone: + 47 62 88 50 00

ISBN 978-82-537-9861-5 (electronic)



Statistisk sentralbyrå
Statistics Norway