Statistisk sentralbyrå
Statistics Norway

# Causes and effects of measurement errors in educational attainment

Experiences from The European Social Survey in Norway

Øyvin Kleven and Kristen Ringdal

TALL SOM FORTELLER
NOTATER / DOCUMENTS
2020 / 35

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



Published 20 October 2020

ISBN 978-82-587-1199-2 (electronic)
ISSN 2535-7271 (elektronisk)

| **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| Not for publication | : |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

In most sample surveys conducted by Statistics Norway educational attainment is merged by administrative registers. In some cases, however the respondents are asked in the survey questionnaire about their educational attainment. This gives the opportunity to study the effect off measurement errors on surveys. From 2002 – 2016 Statistics Norway conducted the fieldwork of the Norwegian edition of The European Social Survey (ESS). The final product of the ESS is a public use file that everybody can download. Due to the Statistics Act of Norway the policy of Statistics Norway prohibits to give information from administrative registers without some terms. For quality control or research purposes inside Statistics Norway it is in accordance with the policy to merge administrative information to any given survey. From 2004 the respondents are asked about their educational attainment in ESS. As part of the internal process quality control of the data collection of the N-ESS all the elements of the sample were merged to relevant administrative registers.

In 2006 and 2017 some results from the quality control where presented at international conferences in front of an audience of methodologist. Based on the feedback from the audiences we believe that a further investigation in the causes and effects of measurement errors in educational attainment is useful. The report is written in English because we feel that the results may be of interest to also non-Norwegians.

This publication has been prepared by Mr. Øyvin Kleven, Division for Population Statistics, Statistics Norway and Professor Kristen Ringdal, Department of Sociology and Political Science, Norwegian University of Science and Technology. Professor Ringdal served from 2002 – 2018 as The Norwegian National Coordinator for The European Social Survey. Mr. Kleven was project leader for the survey in Statistics Norway.

Statistics Norway, 08. October 2020

Ann Kristin Brændvang

# Abstract

From 2002 – 2016 Statistics Norway conducted the fieldwork in the Norwegian edition of The European Social Survey (ESS). ESS is a cross-national survey program that has been conducted every two years since 2002, each round covering around 25 European countries. The main aims are to describe stability and change in social structure, social conditions, and attitudes in Europe. The samples are representative of all persons aged 15 and above resident within private households in each country, regardless of their nationality, citizenship or language. Individuals are selected by strict random probability methods at every stage. All the interviews are conducted in face-to-face mode.

The final product of the ESS is a public use file that everybody can download. Due to the Statistics Act of Norway the policy of Statistics Norway prohibits to give information from administrative registers without some terms. The policy is that public use files only contains survey information from the questionnaire (except for gender, age and residence/region). In the Norwegian edition of the European Social Survey (N-ESS) the respondents are asked about their educational attainment instead of merging it from a register. For quality control or research purposes inside Statistics Norway it is in accordance with the policy to merge administrative information to any given survey. As part of the internal process quality control of the data collection of the N-ESS all the elements of the sample were merged to relevant administrative registers. Hence N-ESS offer a unique possibility to study data quality with respect to educational attainment. We have two independent sources of information for all respondents and can study the effects of non-sampling errors in surveys.

In this document we examine the agreement between level of education as measured both in surveys and from register information and to examine whether these two measurements of education would give similar outcomes in regression analyses of a set of dependent variables. The analysis was based on Rounds 5–8 of the Norwegian part of the ESS (N-ESS) combined with register data on education from Statistics Norway. We examined the differences between the two measurements of education both with crosstabulation and with a multinomial regression analysis of a threefold classification of agreement. Finally, seven dependent variables that varied by source, number of levels, in continuous and categorical versions, and a set of control variables were regressed on education. The agreement rate for the total sample aged 15–104 was 62.4%, or 66.3% after excluding cases with missing values, and the rate increased with the aggregation of education. The multinomial analysis showed that agreement varied by age, mostly due to the low agreement rate for young respondents, who also showed the highest propensity to report a higher education than that shown in the register measurement. The validation analysis indicated that the differences in the effects of education were small, and even in a few comparisons with significant outcomes, these differences would not lead to different substantial conclusions. In conclusion, our main expectations were confirmed: the agreement between the survey and the register measurements was relatively high, and the differences were mainly located in the adjacent categories. The agreement was highest for the middle-aged respondents and lowest for the youngest respondents. The tendency to report a higher level of education than that in the register measurement was negatively related to age, with an opposite relationship for reporting a lower level of education than that in the register measurement. Finally, the validation analysis indicated that the measurements based on the survey and the register measurements in most instances may be used interchangeably. However, in projects targeting young people, a survey measurement of education is recommended because of the time lag of the register information.

# Contents

# 1. Introduction

Educational attainment, measured by the highest level of education, is related to many social phenomena of interest and is often used as a proxy for social status. The distribution of educational attainment from primary to tertiary levels of education is in most countries published as official statistics from administrative records or censuses. In social surveys, the level of education is measured by one or more questions. The purpose of this measure in surveys is usually not to estimate the distribution of educational attainment but to use level of education as a predictor or control variable in empirical analyses. However, when the level of education from surveys is compared with official statistics, their distributions will often differ. In survey methodology, it is well established that a survey estimate is not equal to a true value because of random variability and various sources of bias. Register data based on administrative records are often seen as the "gold standard", but register data are also affected by sources of errors and cannot always be regarded as a true criterion for survey estimates. The purpose of the current paper is to examine similarities and differences between the level of education measured in the Norwegian part of the European Social Survey (NO-ESS) and the measurement of education obtained from register data.

Our empirical analysis will be guided by the following research questions:
To what extent do the measurements of the highest level of education measured in the survey and measured from register data agree? Does the agreement between the two measurements vary by demographic characteristics of the respondents, especially age? Do the two ways of measuring levels of education give similar outcomes when used as explanatory variables in regression analyses of a range of dependent variables where education commonly is used as an explanatory variable?

# 2. Background

## 2.1. Total survey error framework

Classic survey methodology textbooks distinguish between sampling errors and nonsampling errors (Deming 1950; Hansen, Hurwitz and Madow 1953). Sampling errors arise as only a subset of the population is selected. Nonsampling errors are due to mistakes and/or system deficiencies and include all errors that can be made during data collection and data processing. All nonsampling error sources may produce random as well as systematic errors. Systematic errors lead to biased estimates, whereas random errors affect the variance of estimators.

A systematic and comprehensive approach to potential error sources in data from sample surveys is the so-called total survey error framework (e.g., Groves 1989, Biemer and Lyberg 2003, Groves et al. 2009), which starts at conception, moves to collection and processing and then to the statistics produced. The point of departure is that sample surveys rely on two types of inference (Groves et al. 2009): (1) that from the questions to construct and (2) that from the sample statistics to the population statistics. The inference process involves two coordinated sets of steps, namely, obtaining answers to questions constructed to mirror the constructs and identifying and measuring sample units that form a microcosmos of the target population. In Table 1, the two sources of error, the measurement process and representation process, are detailed.

**Table 1 Life cycle of statistics from microdata and total error framework**

| Stage | Measurement | *Error Source* | Representation (units) | *Error Source* |
|---|---|---|---|---|
| | Construct | | Target population | |
| Conception | \/ | *Validity* | \/ | *Coverage error* |
| | Measurement | | Sampling frame | |
| | \/ | *Measurement error* | \/ | *Sampling error* |
| Collection | Response | | Sample | |
| | | | \/ | *Nonresponse error* |
| | \/ | | Respondents | |
| | | *Processing error* | \/ | *Adjustment error* |
| Processing | Edited response | | Postsurvey adjustments | |
| | \/ | | \/ | |
| | → | **Statistics produced** | ← | |

Source: Adapted from Groves et al. 2009.

The measurement shall ideally be designed to perfectly reflect the constructs we are trying to measure. When the measuring instrument is used, there will be a "response". There are a variety of means by which to produce a response, e.g., one can search through one's memory or one can access records to report. The responses are collected and stored, and then they will often undergo some form of editing where we look for inconsistency, etc.

The error sources are listed in italics next to the key elements. Starting with measurement, the first step is to transform the abstract construct into a questionnaire or a description of how to measure a phenomenon. By convention, we do not use the word "error" to describe the mismatches between a construct and its associated measurement. Validity is the extent to which the measures reflect the underlying construct. Measurement error is the mismatch of the true value of the measurement as applied to the sample unit and the value provided. Measurement errors occur during data collection and cause the recorded values of variables to be different from the true values. Between data collection and the beginning of statistical analysis, data must undergo processing that comprises data entry, data editing, often coding and imputation. Errors introduced in these stages are called processing errors.

The right column of Table 1 concerns the representational process. In the NO-ESS, the target population is all inhabitants in Norway who are 15 years old or older. A sampling frame is a listing of all units in the target population. In our case, this list is the National Population Register that contains all the registered legal inhabitants in Norway at a certain date. Coverage errors (or frame errors) are due to divergences between the frame population and the target population. Then, a sample is selected from the sampling frame. Sampling error is deliberately introduced into sample survey statistics. Normally, because of cost constraints, only a small share of the units in the sampling frame are measured. This deliberate "error" introduces deviation in the estimates from the sample compared to statistics produced based on the complete sampling frame. There is an important distinction between sampling bias and sampling variance. Sampling bias arises when some members of the sampling frame are given no chance (or a reduced chance) of selection. Sampling variance arises because many different sets of samples could be drawn from the sampling frame. In almost all measurements in surveys or censuses, there are missing data or nonresponses. Nonresponse error is the difference between statistics based on the respondents and the statistics that would be obtained if all respondents in the sample had been measured. Information from the respondents is used to create a "raw" microdataset. To improve the quality of the estimates, we normally perform some adjustments to the "raw" dataset. Because of nonresponses and/or mismatches between the sample frame and the target population, either we reweigh the respondents or missing data are replaced with predicted response through imputation. The last error source, adjustment

error, arises in the construction of statistical estimators to describe the full target population. Postsurvey adjustments involve efforts to improve the sample estimate in the face of coverage, sampling, and nonresponse errors, but they can also increase these errors in some situations.

The sum of all the errors is total survey error (TSE). This is merely the difference between the estimate from the survey and the true value in the population. Total survey error can rarely be reduced to a quantifiable number. Quantifying all the errors in the model is normally very difficult. For example, the TSE of educational attainment from the NO-ESS is not merely the difference between the official estimates from the register and the NO-ESS because the register estimate also has error attached to it. In most cases, the total error framework is a structural methodological approach for assessing and describing the different error sources.

## 2.2. Expanding the total error framework to also include registers

Statisticians within national statistical institutes have argued for a long time that administrative registers do not provide perfect statistical data (e.g., Hoffmann 1995). Bakker (2011) argues that the same "error model" used for sample surveys (TSE) can be adapted to register data: "it is likely that errors that normally emerge in surveys will also occur in registers". Zhang (2012) expanded this idea to a model for a "two-phase life-cycle model of integrated statistical microdata". The use of administrative data for statistical or research purposes is secondary to nature and in contrast to the primary usage of sample survey data that are designed and collected for certain research purposes. Administrative date are owned and maintained by external register owners. The administrative data have already gone through a sequence of conception, collection and processing before they are delivered to the statistical agency, and importantly, they are collected for administrative purposes (not statistical purposes). Hence, they are almost never fit for statistical purposes right away, but most undergo a process in the statistical office. Using an administrative register for statistical purposes is often only possible after it has been combined with data from other sources. To produce statistics regarding the highest attained education level, one could use a register of examinations. To control the data quality, one would also take into account the register of school enrolment and other relevant registers (Zhang 2012:44).

A register-based statistical system is suitable to study the different error sources in survey statistics, as different errors can be separated (Zhang, Thomsen and Kleven 2013). That is, the effect of nonresponse error on level of education can be estimated in a straightforward manner from the register as the difference between the gross sample (sample - non eligible) and the respondents (net sample). In a previous analysis of the NO-ESS, we demonstrated that there is a selection effect leading to nonresponse bias towards highly educated respondents. Respondents with high education tend to be more willing to participate in the survey compared to respondents with low education (Kleven and Ringdal 2006, 2017). Our previous analysis also suggests that in addition to the selection effect, there is a measurement error leading to overreporting. However, measurement cannot be computed straightforward as the net difference between the level of education based on the register and the level of education from the questionnaire in the survey. As already elaborated, administrative registers do not provide perfect statistical data.

## 2.3. Assessment of three different data sources for educational attainment

The measurement process in the NO-ESS consists of face-to-face interviews based on a random sample. The construction of the education register is more complex. Originally, the registers were established by a census. The data collection method used in the census was a paper questionnaire (self-completion). Information in the register regarding persons with no later recorded education relies solely on information provided from the persons themselves. The same is true for information regarding immigrants and some types of education completed abroad. For most inhabitants of Norway, however, the information in the database stems from administrative records. Table 2 shows an assessment of the different error sources in the three different data sources: sample surveys, census records and administrative records. The coverage error will be low for all three data sources. Only the sample survey will have a sampling error. There will be missing data in all three sources, but this error source will be of far more importance in the sample survey. However, in the register, this source can be important for young people and immigrants. This is mostly due to the time lag in the register. A survey will often be more up to date than a register based on administrative records, although the time delay in registers today is a much smaller problem compared to that during the predigital era. The questions and answers in the survey are created to resemble the categories in the current database. However, the wording in the old censuses was created to fit the school system at the time and was subsequently bridged into current educational codes. In Table 2, different errors that can occur during the measurement process are listed. Measurement error can be systematic or random. Their causes are commonly split into the following 3 sources:

1. The *instrument*: the questionnaire or measuring device used for data collection may lead to the recording of wrong values.
2. *Respondent*: respondents may, consciously or unconsciously, give erroneous data;
3. *Interviewer*: interviewers may influence the answers given by respondents.

The instrument in surveys and census is a questionnaire. In administrative records the instrument is a form or a description of meta data. The instrument can have an impact on the quality of the input of the data for surveys, censuses and administrative registers. Both surveys and censuses will have respondents; hence, errors caused by respondents are relevant here but not in data from administrative records.

A response to a survey question involves a cognitive process, including the comprehension of the question, the retrieval of relevant information, the use of that information to make required judgements and selection and the reporting of an answer (Tourangeau, Rips and Rasinski (2000). Studies have shown that there are many pitfalls related to obtaining an accurate response to a question. Groves et al. (2009), e.g., listed seven respondent-related problems that may lead to measurement errors in a survey:

(1) failure to encode the information sought,
(2) misinterpretation of the questions,
(3) forgetting and other memory problems,
(4) flawed judgement or estimation problems,
(5) problems in formatting an answer,
(6) more or less deliberate misreporting, and
(7) failure to follow instructions.

All seven problems are highly relevant for surveys and censuses but are less important in administrative records. Administrative records are not untouched by human hands; in the predigital era, these records relied heavily on the accuracy of office staff, but today, problems related to reporting accurately are considerably lower. Having an interviewer present can both be an advantage and a disadvantage. Interviewers may help the respondent choose the right answer. On the other hand, an interviewer may influence the respondents to give erroneous answers, merely by the presence of the interviewer. Data processing errors can be present in all three data sources. Generally, processing errors can be both systematic and random. Keying errors can occur, but this type of error is normally random and has little impact.

**Table 2. An assessment of error sources in the measurement of education in the integrated file**

| | | The Register for the Population's Level of Education (NUDB) | |
|---|---|---|---|
| | **Educational attainment from the European Social Survey** | **Educational attainment from mandatory census/follow-up** | **Educational attainment reported from educational authorities** |
| Coverage error | | Low | Low |
| Sampling error | Present but variable (sampling variance) | Low | Low |
| Missing data (nonresponse) | Present for both systematic and variable | Low | Generally low, but high for young people and newly arrived immigrants |
| Time lag (update) | Low | Low | Present |
| Validity issues (specification error) | Question and answer category's like the NUDB | Different wordings in questionnaire, later bridged to NUDB | NUDB |
| Measurement error | Present for both systematic and variable | Present for both systematic and variable | Low |
| *Questionnaire or measuring device* | Relevant | Relevant | Relevant |
| *Failure to encode the information sought* | Relevant | Relevant | Low |
| *Misinterpretation of the questions* | Relevant | Relevant | Low |
| *Forgetting and other memory problems* | Relevant | Relevant | Low |
| *Flawed judgement or estimation problems* | Relevant | Relevant | Low |
| *Problems in formatting an answer* | Relevant | Relevant | Low |
| *More or less deliberate misreporting* | Relevant | Relevant | Low |
| *Failure to follow instructions* | Relevant | Relevant | Low |
| *Interviewer influence* | Relevant | No | No |
| Data processing error | | | |
| *Keying error* | Present but variable | Present but variable | Present but variable |

## 2.4. Social desirability bias

Social desirability bias is a significant problem in survey research if the questions deal with socially desirable or undesirable behaviour or attitudes. This bias may also affect factual questions such as educational attainment. In survey methodology textbooks, asking standard demographic questions such as level of education is seldom elaborated. In these textbooks, asking questions about attitudes and behaviour is given more attention (Bradburn et al. 2004), but knowledge of pitfalls related to asking what seem to be straightforward questions about education has long been acknowledged by practitioners inside of national statistical institutes.

The Population Census of 1950 was the first time that questions on level of education were asked by Statistics Norway. In the census report, it is stated that questions on education have been asked in population censuses in other countries but that the quality of the answers has not always been good:

> *"It is often harder to get correct answers to questions about education than on the more "common" questions. The reason may be that people are not used to answering questions about education in the same way as questions concerning date of birth, occupation and marital status and that some people therefore think that it is of no business for the surveyor. Others who think they do not have the education they "should have", because of a feeling of social prestige, report a higher education than they actually have, or completely fail to answer the question. Others may fail to answer the question because "it has been so long since they went to school" or because "their education is so low so there is nothing to talk about". All of these factors have probably affected the answers received by the census …. However, it is unlikely that this has a great impact on the estimates."* (Statistics Norway 1957, p. 7–8)

The social desirability bias is expected to be greater with an interviewer present compared to the use of self-completion questionnaires. The European Social Survey (ESS) is solely based on face-to-face interviewing. All the interviewers, however, receive basic training in which asking threatening questions on behaviour and attitudes is included. In addition, they also receive special briefing sessions on the ESS.

## 2.5. Summary of expectations

Based on the background section, we expect that the measurement of level of education from the survey interview and the measurement from the register will agree for most respondents. We expect that the differences in the two measurements will mostly be in adjacent categories. The survey measurement is expected to be upwardly biased due to the social desirability effect. Young respondents may tend to report a higher education compared to the register simply because of the time lag in the register measurement. Elderly respondents may find it hard to choose the right level of education because the educational system has changed significantly over time, and previously low or middle education has been "upgraded" in the register. Thus, we expect the agreement between the survey and the register measurement to be highest for the middle-aged respondents and lower for the younger and older respondents. Since these differences in the measurement of education will mostly be in adjacent categories, we expect the survey and register measurements of education to function equivalently as independent variables in analyses of a range of dependent variables.

# 3. Data

## 3.1. The European Social Survey (ESS)

The European Social Survey (ESS) is an academically driven cross-national survey programme that has been conducted every two years since 2002, with each round covering approximately 25 European countries. The main aims of the survey are to describe stability and change in social structure, social conditions, and attitudes in Europe; the final products are public-use files (see: https:// www.europeansocialsurvey.org/). Norway has participated in all rounds of the ESS, with the fieldwork being conducted by Statistics Norway (R1–R8). The samples are representative of all persons aged 15 and above who are residents within private households in each country, regardless of their nationality, citizenship or language. Individuals are selected by strict random probability methods at every stage. All interviews are conducted in a face-to-face mode.

In the NO-ESS, the respondents are asked about their educational attainment instead of merging it from a register due to national policy issues regarding public-use files. For quality control or research purposes inside Statistics Norway, it is in accordance with policy to merge administrative information with any given survey. As part of the internal process quality control of the data collection of the NO-ESS, all the elements of the sample were merged with relevant administrative registers. Hence, the NO-ESS offers a unique possibility for studying data quality with respect to educational attainment. We have two independent sources of information for all respondents and can study the effects of nonsampling errors in surveys. Our empirical analyses are based on an integrated file of the Norwegian ESS surveys from 2010–2016 (ESS Rounds 5–8). Earlier rounds are excluded because of changes in the showcard for the question about the highest level of education.

## 3.2. Integrating microdata from the register-based statistical system with the NO-ESS

The National Population Register (NPR) in Norway was established in 1964 based on the 1960 Census. The census in 1970 was used to check and update the NPR and establish a register of education. The core of the NPR is a universal and unique personal identifier (personal number). A universal and unique identifier is given to every resident at birth or upon entering the country for residence. For some decades, administrative registers have been an important data source for official statistics in Norway. Administrative registers provide frames and valuable auxiliary information for sample surveys. Statistics Norway, like other Nordic national statistical institutes, has made systematic efforts to combine data and integrate various administrative data for statistical purposes (Nordbotten 2010; UNECE 2007).

In most surveys conducted by Statistics Norway, educational attainment is not part of the survey questionnaire; this information is merged from the register. The register is viewed by most people to be of higher quality, and omitting questions from the questionnaire lowers the response burden for the respondents. In a situation with a lower response rate for most surveys, keeping the questionnaire as short as possible is viewed as important. The National Education Database (NUDB) contains information on all residents 16 years of age or older (see https://www.ssb.no/a/english/mikrodata/datasamling/nudb/nudb_20130607-en.html).

The register was originally based on information from the 1970 Population and Housing Census and is updated each year with new information on completed degrees in Norway. Education completed abroad and that supported by the State Education Loan Fund is included. Statistics Norway conducted mandatory self-completion data collection among all registered inhabitants with an unknown education level and education completed abroad in 1990, 1999 and 2011 to update the database. Each year, the database is updated with the records of current students and completed education information for the previous academic year. Course codes from the completed education records are used to update the population's level of education where applicable. If an individual has completed two courses, the one with the highest level is chosen. Specific fields of study are chosen over general fields, and newer courses are counted before older courses.

While missing data are minimal for persons who have completed their education in Norway, there is a share of missing data among immigrants who completed their education abroad before immigrating to Norway. Although census surveys have filled in some of the missing data, approximately 25% of immigrants still had an unknown level of education as of 2014. All educational activities in the register are coded and classified according to the revised Norwegian Standard Classification of Education (NUS2000), which classifies educational programmes by level and field of study (see Table 2). Our analysis is based on a microdata set where register information (including education) is merged with responses from the NO-ESS. In the National Population Register, which is the sampling frame of the NO-ESS, several auxiliary variables from the register-based statistical system were available and were used during data collection.

## 3.3. The measurement of education

The measurement of level of education in the ESS is based on the following question: “What is the highest level of education you have achieved? Please use this card”. The Norwegian showcard was designed to resemble the categories in the register, from 0 ‘No education’ to 8 ‘Ph.D./doctoral degree’. In the current version from 2010, however, there are 14 categories due to the horizontal differentiation of higher education. In Table 3, the 14 categories on the ESS showcard are displayed together with the first digit level of the Norwegian Standard for Classification of Education and the ISCED 2011 levels. The ISCED 2011 levels are slightly modified; it was not possible to distinguish ISCED 4 (postsecondary nontertiary) from ISCED 5 (short-cycle tertiary) in the register data we had access to; hence, ISCED 5 was collapsed into ISCED 4.

**Table 3 Levels of education in ESS, Norwegian Standard Classification of Education and ISCED 2011**

| ESS response categories – 2010 2016 | Norwegian Standard for Classification of Education | ISCED 2011 level |
|---|---|---|
| 1. Not completed primary education | 0. No education, preschool education | 0 |
| 2. Primary or first stage of basic | 1. Primary education | 1 |
| 3. Lower secondary | 2. Lower secondary education | 2 |
| 4. Folk high school (folkehøgskole) | 2. Lower secondary education | 2 |
| 5. Upper secondary (up to 2 years) | 3. Upper secondary basic | 3 |
| 6. Upper secondary (more than 2 years) | 4. Completed upper secondary | 3 |
| 7. Noncredit preparatory courses at colleges/universities | 5. Postsecondary nontertiary | 4 |
| 8. Certificate from postsecondary supplementary programmes (vocational education, technical vocational schools) | 5. Postsecondary nontertiary | 4 |
| 9. University and college programmes, 2 to 2.5-year duration | 6. First stage of tertiary education, undergraduate level | 6 |
| 10. College (Bachelor-, cand.mag) 3-4 years | 6. First stage of tertiary education, undergraduate level | 6 |
| 11. University (Bachelor-, cand.mag) 3-4 years | 6. First stage of tertiary education, undergraduate level | 6 |
| 12. College (Master-,magister) 5-6 years | 7. First stage of tertiary education, graduate level | 7 |
| 13. University (Master-,magister) 5-6 years | 7. First stage of tertiary education, graduate level | 7 |
| 14. Doctoral (Ph.D., postgraduate) | 8. Second stage of tertiary education (postgraduate education) | 8 |

# 4. Results

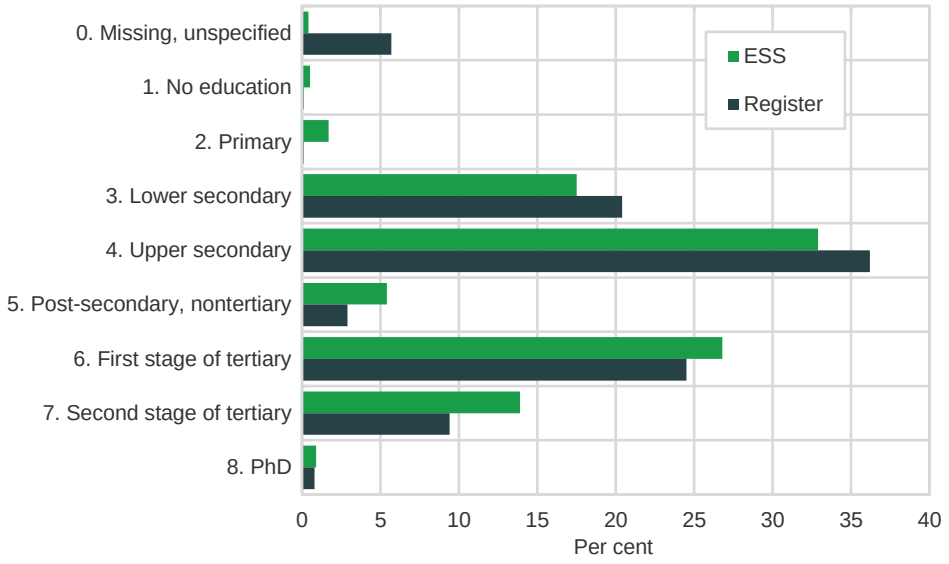
## 4.1. Similarities and differences between survey and register measurements of education

Figure 1 shows the distributions of levels of education from the NO-ESS 2010–2016 and from the register for the net sample aged 15 to 104. This version is the most detailed possible since the integrated file only contains the first digit of the Norwegian Standard Classification of Education from the NUDB database. There were almost no missing data in the NO-ESS surveys but more than 5% missing or unspecified codes in the register data. Very few respondents were classified as having no education or only primary education in both measurements. This is reasonable since lower secondary education is compulsory for all children. There were also relatively few persons with postsecondary and nontertiary education, and persons with the second stage of tertiary education (Ph.D.) are rare.

The NO-ESS had lower percentages for secondary education than the register, and vice versa for tertiary education. The largest difference of 5.5% pertained to the second stage of tertiary education. The second largest difference of 5.3% was found for the first code "Missing, unspecified". More than half of the missing values regarded respondents below 17 years of age.

The distribution displayed in Figure 1 shows the net differences between the register and the survey measurements of education. Some of the misclassifications between the register and the survey will of course cancel each other out. To understand the structure of these misclassifications, we need to examine the gross differences in Table 4.

**Figure 1 Levels of education in Norway, NO-ESS 2010–2016 combined with register data**




Note: Total sample, aged 15–104 at the time of the interviews, n = 6157
Source: European Social Survey, Norway 2010-2016; NUDB, Statistics Norway

The answers to the question on the highest level of education in the survey are displayed in the rows of the table, while the information from the educational register is displayed in the columns. In a situation with perfect agreement between the two measures, all the respondents would be located on the main diagonal. The cells above the main diagonal indicate cases where the survey measurement shows a lower level of education than the register measurement of level of education. The

cells below the main diagonal comprise the cases where the respondent reported a higher level of education than the register measurement.

The agreement rate for the total sample of respondents (2010–2016) is 62.4%. Most of the differences between the register and the survey measurements occur among "nearest neighbours". This is evident from level 3, "Lower secondary", and the levels above it. Out of the 1252 individuals classified in level 3 by the register data, 629 had identical codes from the survey response, while 477 respondents reported the next level of "Upper secondary education". Thus, the main source of difference for this category was reporting a higher level in the survey than in the register. This is also evident for category 6, "First state of tertiary education" (Bachelor level). The survey measurement of level 4, "Upper Secondary", is one level higher than the register measurement for 195 respondents and one level lower for 287 respondents. At the two highest levels, the agreement is 93% for the Master level and 67% for the Ph.D. level. The latter figure is downward biased by the fact that the survey code for Ph.D. was added to code 7 (Master level) for two of the four surveys.

**Table 4 Level of education in the register and the survey among respondents in European Social Survey Norway 2010-2016. Absolute frequencies, n = 6153, age range 15–104**

| | Register | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Survey | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | Total |
| 0 | 5 | 1 | 1 | 8 | 10 | 1 | 0 | 0 | 0 | 26 |
| 1 | 21 | 1 | 0 | 5 | 1 | 0 | 1 | 0 | 0 | 29 |
| 2 | 46 | 0 | 3 | 44 | 10 | 0 | 2 | 0 | 0 | 105 |
| 3 | 147 | 3 | 1 | 629 | 287 | 5 | 5 | 0 | 0 | 1077 |
| 4 | 53 | 0 | 2 | 477 | 1399 | 44 | 46 | 2 | 0 | 2023 |
| 5 | 7 | 0 | 0 | 44 | 195 | 62 | 24 | 0 | 0 | 332 |
| 6 | 42 | 0 | 0 | 41 | 308 | 61 | 1170 | 28 | 1 | 1651 |
| 7 | 28 | 0 | 0 | 2 | 18 | 3 | 255 | 536 | 15 | 857 |
| 8 | 3 | 0 | 0 | 2 | 0 | 0 | 3 | 12 | 33 | 53 |
| Total | 352 | 5 | 7 | 1252 | 2228 | 176 | 1506 | 578 | 49 | 6153 |

Note: Codes for level of education: for register and survey: 0 Missing or unspecified. 1 No education, 2 Primary, 3 Lower secondary, 4 Upper secondary, 5 Postsecondary nontertiary, 6 First stage of tertiary, undergraduate, 7 First stage of tertiary, graduate (Master), 8 Second stage of tertiary education (Ph.D.)
Source: European Social Survey, Norway 2010-2016; NUDB, Statistics Norway

A more systematic summary of the degree of agreement between the survey and the register measurement by age is found in Table 5, where the two youngest age categories are filtered out due to the high percentage of missing values in the register measurement. In total, the survey measurement is lower than the register measurement for 9.1% of the respondents, it is identical for 66.2% of the respondents, and the survey measurement gave the highest level for 24.8% of the respondents. The two measurements of education were more than one level apart for only 8.8% of the respondents. In particular, close agreement showed variation by age. The survey measurement was close to but higher than the register measurement for 30.1% of the respondents aged 17-24, while for 1.1%, the survey measurement was close to but lower than the register measurement. For the oldest age category (70+), the results are in the opposite direction. In the last line, where the identical levels and close agreement are summed, the age difference in agreement disappears.

**Table 5 Agreement between survey and register measurement of education by age, in percent**

| Degree of agreement | 17- 24 | 25-69 | 70+ | Total |
|---|---|---|---|---|
| Survey more than one level below register | 1.9 | 1.0 | 3.3 | 1.4 |
| Survey close but lower | 1.1 | 6.3 | 22.9 | 7.7 |
| Identical levels | 60.4 | 68.6 | 57.7 | 66.2 |
| Survey close but higher | 30.1 | 16.3 | 10.7 | 17.4 |
| Survey more than one level above register | 6.5 | 7.9 | 5.4 | 7.4 |
| Total | 100.0 | 100.1 | 100.0 | 100.1 |
| Sum identical and close agreement | 91.1 | 91.6 | 91.3 | 91.3 |
| n | 734 | 4 329 | 734 | 5 762 |

Note: Close lower/higher: survey measurement one level lower/higher than the register measurement. Total sample with age filer 17–104 at the time of the interviews, listwise deletion. Source: European Social Survey, Norway 2010-2016; NUDB, Statistics Norway

The next step is to examine the agreement between the survey and the register measurements of education when aggregated to fewer levels. Table 6 documents the aggregated levels starting with the nine categories in Table 4. In the 8-level version, the only difference is the "Missing, unspecified" category, which is defined as missing. In the 7-level version, "No education" is joined with "Primary education" as the first level. In the 5-level version, "Lower secondary" is added to the first level, and categories 6 and 7 (Bachelor and Master levels) are joined. In the 4-level version, categories 4 and 5 are joined, and the final aggregation has three levels: primary, secondary and tertiary education.

**Table 6 Aggregation of levels of education**

| The nine levels of education as in Table 4 | 8 | 7 | 5 | 4 | 3 |
|---|---|---|---|---|---|
| 0. Missing, unspecified | . | . | . | . | . |
| 1. No education | 1 | 1 | 1 | 1 | 1 |
| 2. Primary education | 2 | 1 | 1 | 1 | 1 |
| 3. Lower secondary school | 3 | 2 | 1 | 1 | 1 |
| 4. Upper secondary school | 4 | 3 | 2 | 2 | 2 |
| 5. Postsecondary, nontertiary level | 5 | 4 | 3 | 2 | 2 |
| 6. First stage of tertiary, undergraduate level | 6 | 5 | 4 | 3 | 3 |
| 7. First stage of tertiary, graduate (Master) level | 7 | 6 | 5 | 4 | 3 |
| 8. Second stage of tertiary level (Ph.D.) | 8 | 7 | 5 | 4 | 3 |

Note: "." Indicates a missing code and will be excluded from the analyses

In Table 7, the agreement between the survey response and the register data on the level of education is reported for five levels of aggregation. The agreement between the two measurements of level of education with both eight and seven categories was 66.3%, whereas 9% of the survey respondents reported a lower level of education than that obtained from the register, and 24.7% reported a disagreement in the opposite direction. The level of agreement increased with the increased aggregation into fewer levels. For three categories, "Below upper secondary education", "Upper secondary education", and "Tertiary education", the agreement was 76.8%.

**Table 7 Agreement in levels of education obtained from survey and from register data, NO-ESS 2010–2016, in percent**

| Survey vs Register | 8 | 7 | 5 | 4 | 3 |
|---|---|---|---|---|---|
| Survey lowest level | 9.0 | 9.0 | 7.9 | 7.1 | 6.6 |
| Agreement (identical levels) | 66.3 | 66.3 | 67.7 | 71.8 | 76.8 |
| Survey highest level | 24.7 | 24.7 | 24.4 | 21.0 | 16.6 |
| Total | 100.0 | 100.0 | 100.0 | 99.9 | 100.0 |

Note: Total sample, aged 15–104 at the time of the interviews, n = 5780 (listwise deletion).
Source: European Social Survey, Norway 2010-2016; NUDB, Statistics Norway

The four-category version is the one most often used in Norway, with the following categories: 'Below upper secondary education', 'Upper secondary education', 'First stage of tertiary, undergraduate' and 'First stage of tertiary, graduate level and second stage of tertiary postgraduate level'. For this version, the agreement was 71.8%, whereas 7.1% of the respondents reported a lower level, and 21% of the respondents reported a higher level of education in the survey than that obtained from the register.

## 4.2. A multinomial regression analysis of the differences in the two measurements of level of education

The dependent variable was a three-fold classification of the agreement between the levels of education (4-category version) from the survey interviews and those from the register data (NUDB). The first value, "Survey lowest", includes all the respondents for whom the level of education from the NO-ESS was lower than the level obtained from the register data. The second value indicates that the level of education from the NO-ESS and that from the register data were identical. This is the reference category for the multinomial regression analysis. The third category, "Survey highest", includes all the respondents with a level of education from the NO-ESS that was higher than the level obtained from the register data.

The covariates included male gender, age (linear, quadratic and cubic terms), the number of years of full-time education reported from the survey, and the survey year. In Table 8, the regression coefficients are reported together with their probability values and the relative risk ratio (RRR). The results indicate that the regression coefficients for neither gender nor survey year were statistically significant.

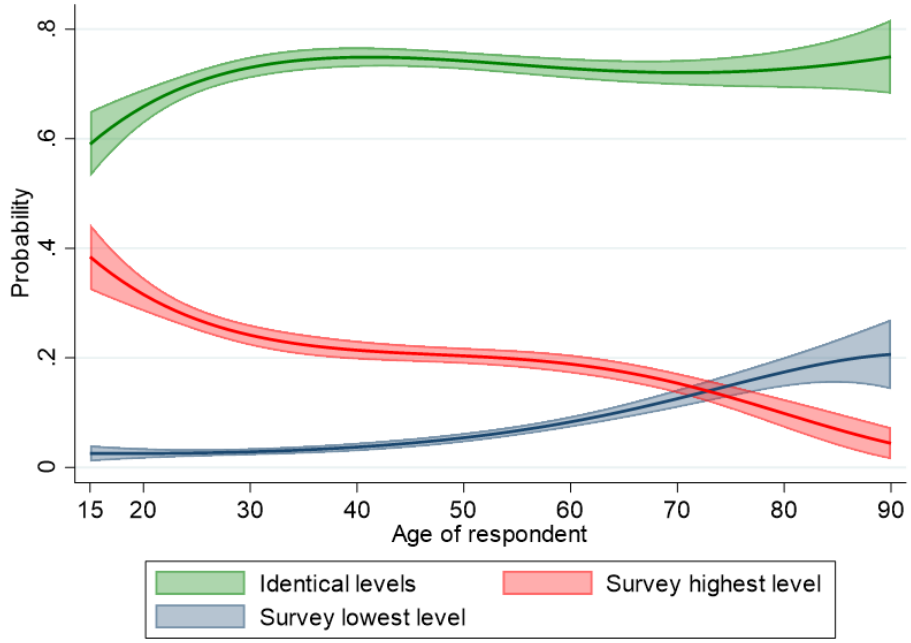
**Table 8 A multinomial regression analysis of differences in survey and register classification of levels of education, NO-ESS 2010 – 2016, n = 5775**

| Survey lowest | Variable descriptions | B | P | RRR |
|---|---|---|---|---|
| Male | Gender, 1 = male | -0.152 | .161 | 0.859 |
| Age | Age in years | -0.092 | .124 | 0.912 |
| Age2 | Age2 (age squared) | 0.002 | .039 | 1.002 |
| Age3 | Age3 (cubic age) | -0.000 | .047 | 0.999 |
| Eduyrs | Years of education | -0.140 | .000 | 0.869 |
| Year | Survey year (base = 2010) | | | |
| 2012 | | -0.016 | .917 | 0.984 |
| 2014 | | 0.130 | .405 | 1.139 |
| 2016 | | 0.097 | .533 | 1.102 |
| Constant | | -0.462 | .627 | |
| *Identical levels –* | reference category | | | |
| *Survey highest* | Variable descriptions | B | P | RRR |
| Male | Gender, 1 = male | 0.093 | .162 | 1.097 |
| Age | Age in years | -0.146 | .000 | 0.864 |
| Age2 | Age2 (age squared) | 0.003 | .001 | 1.003 |
| Age3 | Age3 (cubic age) | -0.000 | .001 | 0.999 |
| Eduyrs | Years of education | 0.030 | .002 | 1.030 |
| Year | Survey year (reference category: 2010) | | | |
| 2012 | | -0.142 | .126 | 0.868 |
| 2014 | | -0.088 | .353 | 0.916 |
| 2016 | | -0.045 | .628 | 0.956 |
| Constant | | 0.778 | .132 | |

Note: B: regression coefficient in logit scale, RRR: (exp(B), relative risk ratio. Total sample, aged 15–104 at the time of the interviews. Categories of the dependent variables: Survey lowest: lower level of education in survey than in register; Survey highest: higher level of education in survey than in register. Identical levels is the reference category.
Source: European Social Survey, Norway 2010-2016; NUDB, Statistics Norway

Age is represented with linear, quadratic and cubic terms, which appeared to be the best way to represent age in the model. Replacing the three parameters of age with a 7-category version of age gave similar results. All three age coefficients showed a statistically significant relationship with the dependent variable. To make the interpretation of the age effect easier, we have included in Figure 2 the predicted probabilities from the model in Table 8. The upper age was set to 90 years because there were few respondents aged 90 or above.

**Figure 2. Similarities and differences in the survey and register measurements of levels of education by age in Norway. Predicted probabilities with 95% CI from estimates in Table 8, n = 5775**




The figure shows the predicted probabilities for the three outcomes by the age of the respondents. The upper curve shows the predicted probability for the survey and the register measurement to be identical. This probability is lowest for the youngest respondents (< 35), it remains high and stable for respondents up to approximately 50 years of age, and then it becomes slightly lower until it increases again for the oldest respondents.

The middle curve shows the predicted probability that the level of education from the survey is higher than the corresponding measurement from the register data. This probability is clearly related to age. For the youngest respondents, the predicted probability is close to .4. The probability decreases rapidly at first and then more slowly until approximately the age of 60, when the probability more rapidly decreases to .04 for the oldest respondents.

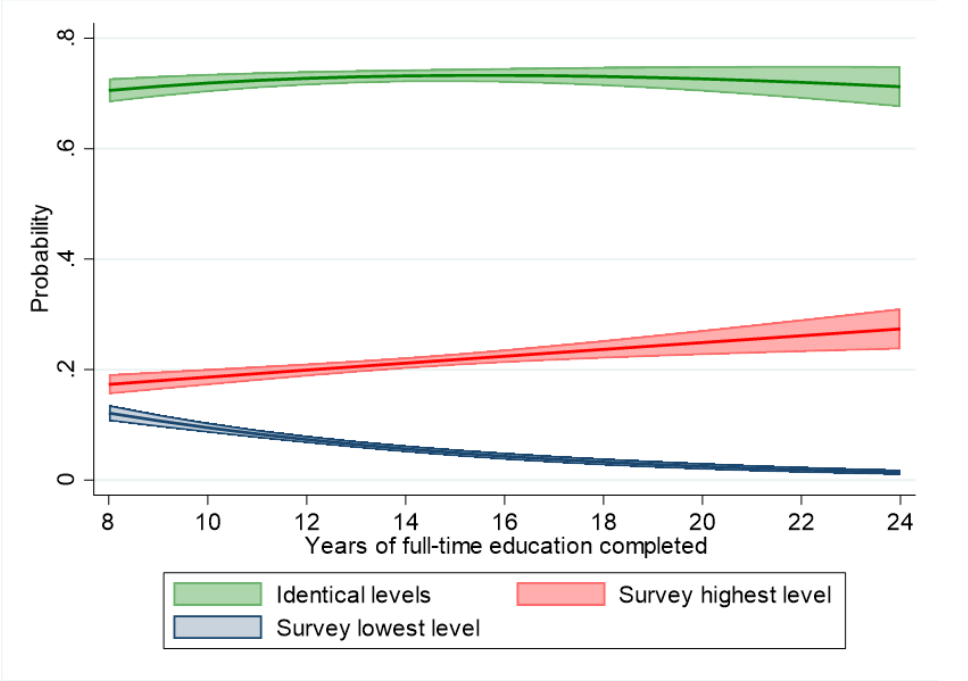
The bottom curve shows the predicted probability that the survey measurement of the highest level of education is lower than the measurement based on register data. This probability is very low for the youngest respondents, but it increases by age, slowly at first, then it increases more rapidly from the age of 50 until it reaches .2 for the oldest respondents.

Years of full-time education reported in the ESS were added to the model as an alternative and more continuous measurement of education than the levels derived from the 14 categories on the showcard. Years of education were statistically significantly related to the three-category outcome variable. The coefficient for years of education was negative and statistically significant for both the “Survey lowest” ($B = -0.14$, $p < .000$) and “Survey highest” outcomes ($B = -0.03$, $p < .01$). We have added Figure 3 for the ease of interpretation of the results, as the figure also shows the predicted probabilities for the reference category of the dependent variable.

The probability of identical measurements is rather stable at around .7 for the range of full-time education from 8 to 24 years of age. The probability for the survey to

show the highest level of education increased linearly from close to .2 for the respondents who reported 8 years of full-time education to just below .3 for those who reported 24 years of education. The probability for the survey to give the lowest level is highest (.12) for the respondents who reported the fewest years of education and decreases evenly until the probability is below .02 for those who reported the highest years of full-time education.

**Figure 3. Difference in survey and register classification of levels of education by reported years of full-time education completed in Norway, predicted probabilities with 95% CI from estimates in Table 8, n = 5775**




## 4.3. The performance of the educational measures in multiple regression analyses

This validation analysis provides answers to the following questions: Does it matter for the estimated coefficients whether we use the levels of education from the surveys or from the register data? Do different levels of aggregation of education make a difference? Is the effect of education best captured by a linear or a nonlinear model?

We have chosen a variety of dependent variables for which education will naturally be among the independent variables. The descriptive statistics for the variables used in the regression analyses are displayed in Table 9. Note that this table and the regression analysis are estimated for the respondents aged 25–79 years of age for two reasons. First, we excluded the respondents who were younger than 25 years of age since they had not yet had the time to complete a high-level education. Second, by also excluding the oldest respondents, we avoided the tails of the age distributions where the agreement between the survey and the register measurement of education is poorest.

The descriptive statistics for the dependent variables in the regression analyses are found in the upper part of Table 9, with a more thorough documentation available in the appendix. The first dependent variable takes the value of one for those who reported having voted in the last national election (“Stortingsvalg”) and zero for those who did not vote. Although this variable is dichotomous, we chose to use OLS regression instead of logistic regression for an easier comparison with the other analyses. The next four dependent variables are composite measurements,

with each constructed as a mean of a set of items. For these, the Cronbach's alpha values are reported. The first scale is based on the three items of engaging in conventional political activities; having contacted a politician, government or local government official; or having worked in a political party or action group or worked in another organisation or association. The second scale is based on four items of unconventional political activities or political action, namely, worn a campaign badge/sticker, signed a petition, taken part in a lawful public demonstration, and boycotted certain products. In terms of Cronbach's alpha (0.51 and 0.55), both scales had lower internal consistency than desired based on a reflective measurement model. This aspect is, however, less relevant if we apply a formative measurement model. Regardless, the low reliability is of less importance since our purpose is only to compare the effects of various representations of level of education.

In the next section follows two three-item scales on attitudes towards immigration. The first is about allowing more immigrants from the same race/ethnicity, those of different race/ethnicity than the majority population, and people from poor countries outside Europe. The second scale is whether immigration is good or bad for the economy, for cultural life, and for making Norway a better or worse place to live. The Cronbach's alpha values for the two scales (0.88 and 0.81) were quite satisfactory.

The last dependent variables were picked from a battery of attitudinal questions. The first is whether "The government should take measures to reduce differences in income levels." The second is whether "Gay men and lesbians should be free to live their own life as they wish." High values on the scales imply high political activity, liberal attitudes towards immigration, homosexuality and endorsing government action to reduce income inequality.

First among the independent variables are level of education in four and seven categories, for both survey or register measurement, treated as both continuous and categorical variables. Note that the descriptive statistics in the table are based on continuous versions. The correlation between the survey and the register measurement is .84 for the 7-level version and .83 for the four-level version. Furthermore, the correlations between the four- and seven-level versions are .98 for both the survey and the register measurement.

Perception of income adequacy was chosen as a measurement of economic coping, as it often functions better than measured income in our experience. High values indicate coping well with one's current income. Next is the variable of gender, with the value of one being assigned to males. The age variables each have five categories with the first serving as the reference category for the following regression analyses. The main activity variable has six categories, with the first (in work) serving as the reference category. The last variable in the table is that of survey year, with 2010 serving as the reference category in the regression analyses.

We do not pretend to develop theoretical-based models for each dependent variable but rather apply a common model to all outcomes. Since the purpose is to examine the differences in the coefficients of education, our description of the findings will not include a substantial interpretation of the effects.

**Table 9 Descriptive statistics for the variables in the validation analysis**

| Dependent variables | Mean | Std. | Min | Max | n |
|---|---|---|---|---|---|
| Voted in last national election | 0.90 | | 0 | 1 | 4,607 |
| Conventional political activities (scale, $\alpha$ = .51) | 0.67 | 0.85 | 0 | 3 | 4,799 |
| Unconventional political activities (scale, $\alpha$ = .55) | 1.04 | 1.11 | 0 | 4 | 4,799 |
| Allow more immigrants (scale, $\alpha$ = .88) | 2.94 | 0.65 | 1 | 4 | 4,779 |
| Immigrants make the country better (scale, $\alpha$ = .81) | 5.80 | 1.83 | 0 | 10 | 4,796 |
| Attitude towards homosexuality | 4.26 | 0.90 | 1 | 5 | 4,785 |
| Government should reduce income differences | 3.56 | 1.02 | 1 | 5 | 4,783 |
| *Independent variables (regressors)* | | | | | |
| Education in 7 levels/categories, survey | 4.08 | 1.40 | 1 | 7 | 4,799 |
| Education in 7 levels/categories, register | 3.85 | 1.36 | 1 | 7 | 4,799 |
| Education in 4 levels/categories, survey | 2.53 | 0.94 | 1 | 4 | 4,799 |
| Education in 4 levels/categories, register | 2.39 | 0.90 | 1 | 4 | 4,799 |
| Coping on current income | 3.55 | 0.65 | 1 | 4 | 4,798 |
| Male gender | 0.53 | | 0 | 1 | 4,799 |
| Age | | | | | 4,799 |
| 25-40 (reference category) | 0.30 | | 0 | 1 | |
| 41-50 | 0.23 | | 0 | 1 | |
| 51-60 | 0.21 | | 0 | 1 | |
| 61-70 | 0.18 | | 0 | 1 | |
| 71-79 | 0.08 | | 0 | 1 | |
| Main activity | | | | | 4,798 |
| In work (reference category) | 0.69 | | 0 | 1 | |
| In education | 0.02 | | 0 | 1 | |
| Unemployed | 0.02 | | 0 | 1 | |
| Pensioned | 0.22 | | 0 | 1 | |
| Housewife | 0.04 | | 0 | 1 | |
| Other activities | 0.01 | | 0 | 1 | |
| Survey year | | | | | 4,799 |
| 2010 (reference category) | 0.25 | | | | |
| 2012 | 0.27 | | 0 | 1 | |
| 2014 | 0.23 | | 0 | 1 | |
| 2016 | 0.25 | | 0 | 1 | |

Note: *Std.*: standard deviation, n: number of observations. Scale indicate composite measurements, $\alpha$: Cronbach's alpha.
Source: European Social Survey, Norway 2010-2016; NUDB, Statistics Norway

For each dependent variable, we estimated OLS regression models for the eight combinations of source of the measurements, the level of aggregation (7 vs 4 categories) and representation as a continuous or categorical regressor. The results are reported in Table 10, where the metric regression coefficient for the level of education is reported together with the partial etas squared as a measure of effect size. All models include (male) gender, age in five categories, main activity for the last seven days in six categories, and survey year in four categories as the controls. The coefficients of the control variables are, however, not reported in the table. In reading the table, the metric coefficients for the survey measurement and the register measurement of level of education may only be compared within the same level of aggregation. Thus, for the first dependent variable (voted), the two measurements gave close to identical regression coefficients for both levels of aggregation (B = 0.029 vs B = 0.028 and B= 0.042 vs B = 0.040). We may not, however, compare the coefficients between the two aggregations since the scale of the educational variables affects the coefficients. For this purpose, the partial etas squared are useful. These values are also close to identical in comparisons across the two levels of aggregation. Thus, it does not seem to matter for the estimates of the effects of education whether we use education at seven or four levels or base the measurement on survey or register data.

This summary is, however, based on the assumption that the relationship between the dependent variable and the level of education is linear. The second line of the partial etas squared is based on a model where the levels of education are represented by a set of dummy variables to accommodate nonlinearity. Thus, we may compare the partial etas squared within the columns to assess the degree to which the linear model is adequate. For the first dependent variable, the partial eta squared is clearly higher for the categorical than for the continuous representation of level of education. For the seven-level version from the register data, the

difference between the etas squared ($\eta^2$ = .014 vs. $\eta^2$ = .023) is statistically significant at the .05 level. This outcome is marked in the table by printing the coefficients in bold italic typeface. For the survey measurement, the difference in partial etas squared is very close to statistical significance. In Figure 4, the predicted probabilities for having voted are displayed for the two models. The linear model tracks the nonlinear model well for all educational categories except the first one, having a far lower probability of having voted than the linear model. The category included, however, only a handful of respondents in the register-based measurement. Thus, as expected, the difference between the linear and the nonlinear representation of level of education is smaller and not statistically significant in the four-level aggregation.

**Figure 4. A comparison of predictions of having voted using register measurement of education treated as metric and categorical, n = 4605. Predictions from the regression analysis in Table 10**

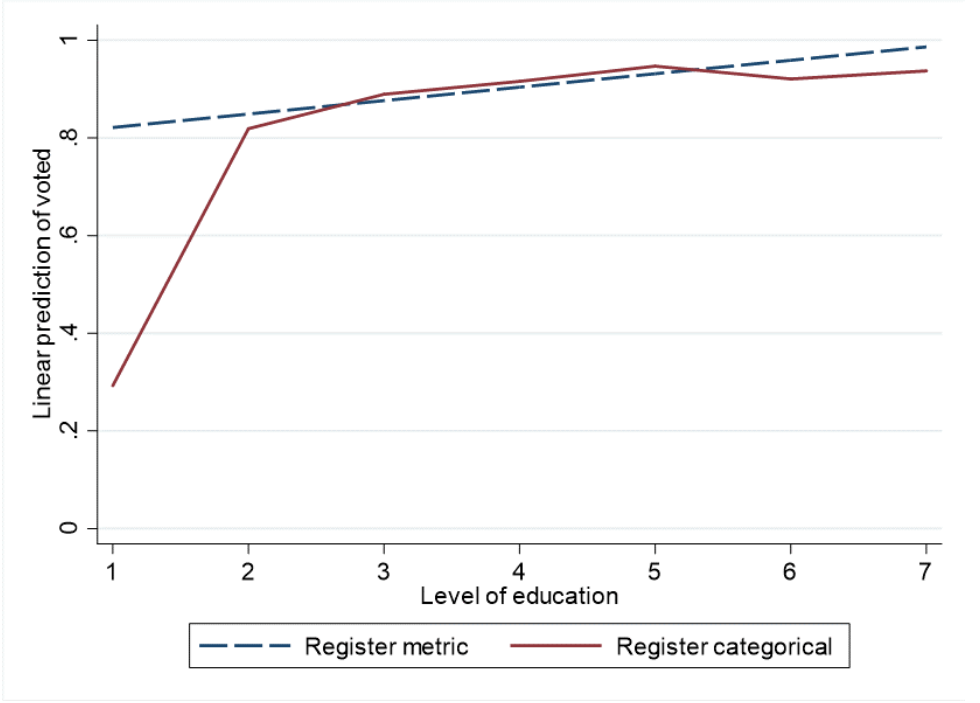


**Table 10. Regression analyses with levels of education from surveys and register in 7 and 4 value-/category versions, age 25–79**

| | 7 values/categories | | 4 values/categories | | |
|---|---|---|---|---|---|
| Dependent variables: | Survey | Register | Survey | Register | n |
| Voted (0 - 1) | | | | | 4605 |
| Metric regression coefficient | 0.029 | 0.028 | 0.042 | 0.040 | |
| Partial eta squared | 0.016 | ***0.014*** | 0.015 | 0.013 | |
| Education in categories, partial eta squared | 0.020 | ***0.023*** | 0.018 | 0.019 | |
| Conventional political participation | | | | | 4797 |
| Metric regression coefficient | 0.107 | 0.083 | **0.160** | **0.125** | |
| Partial eta squared | **0.028** | **0.017** | **0.028** | **0.016** | |
| Education in categories, partial eta squared | **0.029** | **0.019** | **0.028** | **0.019** | |
| Unconventional political participation | | | | | 4797 |
| Metric regression coefficient | **0.139** | **0.099** | **0.201** | **0.142** | |
| Partial eta squared | **0.028** | **0.014** | **0.026** | **0.012** | |
| Education in categories, partial eta squared | **0.029** | **0.016** | **0.027** | **0.016** | |
| Immigration: allow more | | | | | 4778 |
| Metric regression coefficient | 0.100 | 0.108 | 0.146 | 0.158 | |
| Partial eta squared | 0.043 | 0.048 | 0.040 | 0.045 | |
| Education in categories, partial eta squared | 0.044 | 0.052 | 0.043 | 0.050 | |
| Immigration: better | | | | | 4794 |
| Metric regression coefficient | 0.407 | 0.416 | 0.617 | 0.614 | |
| Partial eta squared | 0.089 | 0.090 | 0.090 | 0.086 | |
| Education in categories, partial eta squared | 0.093 | 0.098 | 0.093 | 0.094 | |
| Attitude to homosexuality | | | | | 4783 |
| Metric regression coefficient | 0.101 | 0.088 | 0.151 | 0.125 | |
| Partial eta squared | 0.024 | 0.018 | **0.024** | **0.015** | |
| Education in categories, partial eta squared | 0.025 | 0.020 | 0.024 | 0.018 | |
| Government reduce income differences | | | | | 4781 |
| Metric regression coefficient | -0.055 | -0.042 | -0.079 | -0.069 | |
| Partial eta squared | 0.005 | 0.003 | 0.005 | 0.004 | |
| Education in categories, partial eta squared | 0.008 | 0.007 | 0.006 | 0.004 | |

Note: All models include the following control variables: age in four categories, gender, main activity in six categories, perception of income adequacy, and survey year in four categories. Only the coefficients for level of education are reported in the table. All regression coefficients and etas squared were statistically significant at the 0.01 level. Coefficients in bold indicate within-line significant differences at the .05 level. Coefficients in bold italics indicate within-column significant differences at the .05 level. A statistically significant difference between two etas squared means that the 95% CI for one of the estimates did not include the estimate for the other eta squared.
Source: European Social Survey, Norway 2010-2016; NUDB, Statistics Norway

The results for conventional political activities indicate that the survey measurement of education gives higher estimates of the regression coefficients. The difference in regression coefficients is, however, statistically significant at the .05 level only for the four-level aggregation of education (B = 0.160 vs B = 0.125). This outcome is marked in the table by printing the coefficients involved in bolt typeface. The differences between the partial etas squared are, however, statistically significant in both comparisons. In summary, the effects of level of education differ between the survey-based or register-based measurement of levels of education; however, the effects do not depend upon the level of aggregation. Furthermore, the linear model seems to be adequate, since the etas squared for the nonlinear models are only slightly higher than their linear cousins.

For unconventional political activities, the regression coefficients for the survey measurement of education are significantly lower than those for the register measurement for both the 7-level (B = 0.139 vs. 0.099) and the 4-level aggregation (B = 0.201 vs. 0.142). Additionally, the differences in the partial etas squared are statistically significant at the .05 level. Figures 5-6 show that the distances between the linear predictions are relatively small, indicating that the choice between the two measurements would not matter.

**Figure 5.** **A comparison of linear predictions of unconventional political activities using survey and register measurement of education (seven levels), n = 4797. Predictions computed from the regression analysis in Table 10.**

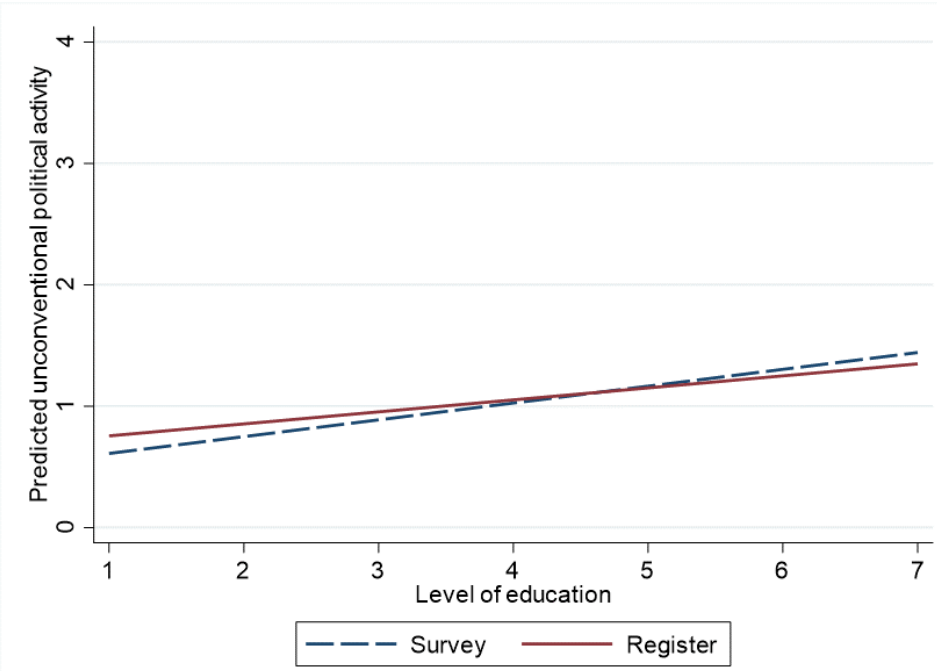
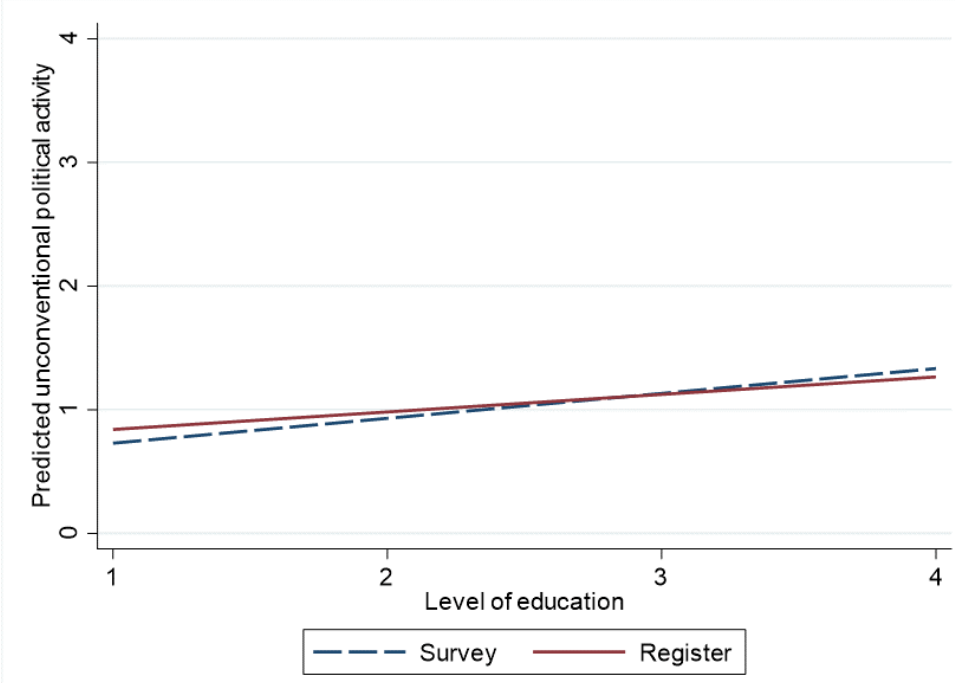



**Figure 6.** **A comparison of linear predictions of unconventional political activities using survey and register measurement of education (four levels), n = 4797. Predictions computed from the regression analysis in Table 10**




Finally, the similarities in the etas squared for the linear and nonlinear models indicate that the linear models are quite adequate.

For the remaining dependent variables (except one), namely, the two scales on attitude towards immigration and attitude towards reducing income differences, the differences are small in all comparisons between the survey-based and register-based measurement of education, between the 7-level and 4-level versions of

education, and between the linear and nonlinear models. However, for attitude towards homosexuality, the partial eta squared is lower for the register-based than for the survey-based measurement of education and for both the 7-level and the 4-level aggregation. Only the latter difference, however, was statistically significant ($\eta^2 = .025$ vs. $\eta^2 = .018$)

Thus, the overall impression is of similarities in all three types of comparisons for four of the seven dependent variables: the two scales of attitude towards immigration, attitude towards homosexuality and attitude towards income differences. The largest differences between the linear models were found for political activities. For voting, the linear models seemed to be inadequate (Figure 6). This finding does, however, mainly rest on the first educational category with only a handful of respondents in the register-based measurement. The most consistent differences were found for unconventional political participation. Although the survey measurement performed best, the differences in the linear prediction appeared to be small.

# 5. Discussion

The purpose of the paper was twofold. First, we examined the agreement between level of education as measured in both surveys and register information and how this measurement varied by respondents' age. Second, we examined whether the two measurements of education would give similar outcomes in the regression analyses of a set of dependent variables.

We found that agreement (identical measurements) between the survey and register measurements of education was most frequently found for the middle-aged respondents and lower for the younger and older respondents. Furthermore, the agreement rate increased as the number of levels in the measurement of education decreased. A mismatch between the survey and the register measurements was often located in the adjacent levels and was termed close agreement. This close agreement also varied by age. The percentage of respondents with close agreement in which the survey gave the lowest level varied from almost zero for the youngest age category to more than 20% for the oldest. The tendency was the opposite for close agreement in which the survey measurement gave the highest level.

The age effect was the most interesting finding in the multinomial regression analysis of the three-fold classification of agreement in the survey and the register measurements of education. The probability of identical levels for the two measurements was lowest for respondents below 35 years of age. The probability of reporting a higher level of education than the register measurement was high for the youngest respondents and decreased evenly to be the lowest for the oldest respondents. The probability for the survey to give the lowest level of education was very low for the youngest and increased to become highest for the oldest respondents.

In the validation analysis, seven dependent variables were regressed on the 7- and 4-level measurements of education from the two sources (survey and register) in continuous and categorical representation, with a set of control variables.

Was the 7-level measurement better than the 4-level measurement of education? In terms of effect size, the differences seem to be small, and the correlation between the two versions was just below 0.98 for both the survey and the register measurements of education. An advantage of the 4-level version is that it does not

have any small categories and is thus not exposed to nonlinearity from small categories at the extremes.

Did the effect of education depend upon the source of the measurement? The survey-based measurement gave higher estimates than the measurement obtained from the register for the effect of education for the two scales of political participation. The linear predictions did, however, indicate that these differences were not large enough to be substantially important. For the two scales for attitudes towards immigrants, the attitude towards homosexuality, and for the question on income equality, the two measurements gave quite similar results.

Did the representation of level of education as a continuous variable mask the nonlinear effects of education? We found only one comparison (i.e., having voted) where the nonlinear model returned a significantly higher eta squared than the linear model. This finding depends, however, on the first level of education, with only a handful of respondents. Thus, the representation of education as a continuous variable should in most instances suffice for both the 7-level and the 4-level versions. It is, however, especially important to check for nonlinearity when ordinal regressors are treated as continuous.

The great advantage of using register information instead of measuring education by asking questions is lowering the response burden for the respondent and avoiding the potential problem of overreporting in surveys due to social desirability. The greatest disadvantage in using register measurement is missing data for some respondents (mainly young people and immigrants). If the study population comprises young people or immigrants, the register measurement of education will clearly be insufficient. Additionally, the register of education is not readily available to field agencies. Only Statistics Norway has a copy of the register for its own use, and other field agencies must apply to the owner of the register to obtain the data.

Can the register measurements be used in combination with the survey question data, for instance, to let the interviewer read out the individual's registered education and let the respondent confirm or change the suggestion from the register? This approach would probably not be useful if only the first digit (level) of the educational code was read to the respondent. However, it could probably function if the highest detailed education were used. This would, however, potentially result in considerable postcoding work.

# 6. Conclusions

The purpose of the current paper was to examine the agreement between the level of education measured in surveys and that from register information and to examine whether the two measurements of education would give similar outcomes in regression analyses of a set of dependent variables.

Among the Norwegian respondents in ESS R5–R8, the agreement between the survey and the register measurements was relatively high, and mismatches were mainly located in the adjacent categories. The agreement was related to age and was highest for the middle-aged respondents and lowest for the youngest respondents.

Based on previous research and relevant literature, we expected the level of education from the survey interviews to be higher than the level obtained from the register due to the social desirability effect. Our findings are consistent with this expectation, as more respondents "overreport" than "underreport" their education compared to the register measurement. The tendency to report a higher level of education than that in the register measurement was negatively related to age, with the opposite relationship shown for reporting a lower level of education than that in the register measurement. This outcome indicates that some of the "overreporting", especially for young people, may be due to lag in the register measurement.

Finally, the validation analysis indicated that the measurement based on the survey and that from the register measurement in most instances may be used interchangeably. Thus, it would be possible to base measures of level of education solely on register data. However, in projects targeting young people or immigrants, the survey measurement of education is recommended because of the time lag and the coverage of the register information.

# References

Bakker, B. F. M. (2011). Micro-integration. State of the art. In ESSnet on Data Integration WP1 State of the Art on Statistical Methodologies for Data Integration (77-107). (ESSnet). Luxembourg: Eurostat.

Biemer, P. P. & Lyberg, L. E. (2003) *Introduction to Survey Quality.* New York: Wiley.

Bradburn, N.M., Wansink, B. & Sudman, S. (2004). *Asking questions: The definitive guide to questionnaire design – for market research, political polls, and social and health questionnaires*. San Francisco, Calif: Jossey-Bass.

Demming, W.E. (1950) *Some theory of sampling.* New York: Wiley.

ESS Round 5: European Social Survey Round 5 Data (2010). Data file edition 3.4. NSD - Norwegian Centre for Research Data, Norway – Data Archive and distributor of ESS data for ESS ERIC. doi:10.21338/NSD-ESS5-2010.

ESS Round 6: European Social Survey Round 6 Data (2012). Data file edition 2.4. NSD - Norwegian Centre for Research Data, Norway – Data Archive and distributor of ESS data for ESS ERIC. doi:10.21338/NSD-ESS6-2012.

ESS Round 7: European Social Survey Round 7 Data (2014). Data file edition 2.2. NSD - Norwegian Centre for Research Data, Norway – Data Archive and distributor of ESS data for ESS ERIC. doi:10.21338/NSD-ESS7-2014.

ESS Round 8: European Social Survey Round 8 Data (2016). Data file edition 2.1. NSD - Norwegian Centre for Research Data, Norway – Data Archive and distributor of ESS data for ESS ERIC. doi:10.21338/NSD-ESS8-2016.

Groves, R.M. (1989) *Survey errors and survey costs.* New York: Wiley

Groves, R.M., Fowler, F.J., Jr., Couper, M.P., Lepkowski, J.M., Singer, E. & Tourangeau, R. (2009). *Survey methodology* 2nd ed. New York: Wiley.

Hansen, M. H., W. N. Hurwitz & Madow, W.G. (1953) *Sample survey methods and theory volume I. Methods and applications.*

Hoffmann, E. (1995) "We must use adminstrative data for official statistics – but how should we use them?" *: Statistical Journal of the United Nations Economic Commission for Europe, 12*( 1), 41-48.

Kleven, Ø. & Ringdal, K. (2006). "A comparison of two ways of measuring level of education in Norway: questions in survey interviews versus register data." *Paper presented at Q2006—European Conference on Quality in Survey Statistics, Cardiff.*

Kleven, Ø. & Ringdal, K. (2017). "Level of education – Comparing survey interviews with administrative records" *Paper presented at The 7th Conference of the European Survey Research Association (ESRA), Lisbon 2017*

Nordbotten, S. (2010): The Use of Administrative Data in Official Statistics – Past, Present, and Future – With Special Reference to the Nordic Countries. In M. Carlson, H. Nyquist & M. Viliani: *Official Statistics – Methodology and Applications in Honour of Daniel Thorburn* (Chapter 17, 205–225 ). Retrieved from: https://officialstatistics.wordpress.com/contents/.

Statistics Norway (1957). *NOS XI 258. Population Census*, 1. December 1950. Oslo: Statistics Norway.

Tourangeau, R., Rips, L. & Rasinski, K. (2000) *The Psychology of Survey Response.* Cambridge: Cambridge Univerity Press.

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), 270–288.

Zhang, L.-C. (2012).Topics of statistical theory for register-based statistics and data integration. *Statistica Neerlandica 66*(1), 41-63.

UNECE (2007) *Register-based statistics in the Nordic countries: review of best practices with focus on population and social statistics.* Geneva: United Nations Economic Commission for Europe, retrieved from http://www.unece.org/index.php?id=17470.

# Appendix A: Documentation of the dependent variables in the OLS regression analysis

The first theme is political activities, which includes voting in the last national election, conventional and unconventional political activities. The variables "voted" takes the value of one for those who reported to have voted and zero for those who reported that they did not vote.

The core questions in the ESS includes a group of seven question on political activities asked in this way:

> *"There are different ways of trying to improve things in [country] or help prevent things from going wrong. During the last 12 months, have you done any of the following? Have you…*
> *...contacted a politician, government or local government official?*
> *...worked in a political party or action group?*
> *...worked in another organisation or association?*
> *...worn or displayed a campaign badge/sticker?*
> *...signed a petition?*
> *...taken part in a lawful public demonstration?*
> *...boycotted certain products?"*

A factor analysis indicated a two-factor solution. The first three items related to organized, conventional political activities loaded strongest on the second factor and the remaining items with ad hoc or unorganized political activities loaded strongest on the first factor. Two summated scales were formed for conventional and unconventional political activities, based on the counts of number of types of political activities for each respondent. In terms of Cronbach's alpha (0.51 and 0.55) both scales showed lower internal consistency than desired. Among the ESS core questions, there are six questions on attitudes to immigration. The first three is about how many immigrants to allow with response categories ranging from *"Allow many to come and live here"* to "*Allow none*":

> *"Now some questions about people from other countries coming to live in Norway.*
>
> *Now, using this card, to what extent do you think Norway should allow people of the same race or ethnic group as most Norway's people to come and live here.*
>
> *How about people of a different race or ethnic group from most Norway people?*
>
> *How about people from the poorer countries outside Europe?"*

The second set of questions asked about the consequences of immigration:

*"Would you say it is generally bad or good for Norway's economy that people come to live here from other countries? Please use this card." The response scale that ranged from 0 "Bad for the economy" to 10 " Good for the economy".*

*"And, using this card, would you say that Norway's cultural life is generally undermined or enriched by people coming to live here from other countries?" The response scale ranged from 0 "Cultural life undermined" to 10 "Cultural life enriched".*

*"Is Norway made a worse or a better place to live by people coming to live here from other countries? Please use this card. The response scale ranged from 0 "Worse place to live" to 10 "Better place to live",*

From the two sets of items, we constructed two scales as the means scores on the three items in each set. In the first set, we reversed the codes so that high values on both scales indicated that the respondents would allow more immigrants and

evaluated the consequences of immigration to be positive for the country. Cronbach's alpha (0.88 and 0.81) for the two scales were quite satisfactory.

We have also included two single questions about income differences and homosexuality among the dependent variables:

*"Using this card, please say to what extent you agree or disagree with each of the following statements:*

*The government should take measures to reduce differences in income levels.*
*Gay men and lesbians should be free to live their own life as they wish."*
*The five response categories tanged from 1 "Agree strongly" to 5 "Disagree strongly".*

In the regression analysis, the scales were reversed, so that high values indicate agreement with the statements.