Documents 9/2010

Ole Villund

Evaluating employment classification

A quality study linking survey data and register data

Statistisk sentralbyrå • Statistics Norway
Oslo–Kongsvinger

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

ISBN 978-82-537-7833-4 Printed version
ISBN 978-82-537-7834-1 Electronic version
ISSN 1891-5906
Subject: 06.90

Print: Statistics Norway

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Not for publication	: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

Individual data on employment are used for a wide range of official statistics and research purposes. Statistics Norway collects data from both sample survey and administrative registers for this purpose. Especially when making detailed tables or measuring labour market flows, it is important that each unit have a correct classification of the employment status. This report provides an assessment of the employment data quality, as well as some discussion on theory and method.

The analyses have been performed with data that are already collected, and represents a cost-effective alternative compared to quality studies that relies on extra data collection.

Abstract

We present a quality assessment of employment data sources that are used for official statistics. Specifically, we want to evaluate the employment classification at the micro level, and link survey- and register-based employment data at the individual level. We do not assume that one of the data sources has all the true answers. Instead, we divide each data source into subgroups with varying quality. The subgroups of survey data with the best quality are used to check the corresponding register data records, and vice versa.

The survey data are collected for the Norwegian Labour Force Survey (LFS), a relatively large sample survey. The register data source is composed of records from the employee register, tax return register (for self-employed) and The Norwegian Tax Administration's "End of the Year Certificate Register". The certificate register consists mainly of wage data, which is used to classify jobs that are not registered in the employee register.

Overall, we find a high agreement rate at the micro level. Over 90 percent of individuals are classified to the same employment status in both sources. On the macro level, we try to differentiate between systematic- and random measurement errors. The register data seem to systematically overestimate employment, by about 1.5 percent, compared to only 0.1 percent estimated random error. In official Norwegian employment statistics the total number of employed in the register-based statistics is adjusted to match the 4th quarter LFS employment estimate. Consequently, the measurement errors estimated for subgroups will not show up as divergence between the two official statistics.

Proxy, or indirect, interviews by family members are allowed in the LFS survey. We find that proxy interviews underestimate employment rate by about 5 percentage point, when controlling for register-employment status. Proxy interviews constitute about 15 percent of the total sample, so the overall effect is noticeable but not alarming for the total employment rate. However, among young people there is both a high incidence of proxy interviews and substantial underestimation of employment due to the proxy interviews. Especially among students, this seems to be pronounced, resulting in nearly 15 percentage point underestimation of employment rate among students.

Overall, we find about 1.4 percentage point overestimation of employment rate, and an agreement rate of over 92 percent. However, there is considerable variation in quality between the source registers. Only 68 percent of those classified as employed from wage-certificate data, are classified as employed in the survey data at the same time. Small jobs are overrepresented among those classified as employed based on the wage-certificate data. Consequently, misclassifications in this group will especially affect analyses of young people and other marginal employed. For the tax return register data, the corresponding figure is 87 percent. That could mean up to 13 percent misclassifications, which of course will affect register-based analyses of self-employment.

Contents

1. Introduction

1.1. Purpose

This report focus on the micro-level quality of employment information in statistical data. Correct classification of employment status is important in order to estimate accurate employment rate. In addition, the micro level quality is important for many research projects that use the same employment data sources. Furthermore, there is a growing interest for estimates of labour market transitions and gross flows. Flow estimates are more vulnerable to classification errors, than level estimates are.

The goal for this report is to assess the classification error of employment status in the data used for official statistics and labour market research. We propose to achieve this by linking data at the micro level, using data sources that we have already have available. With this approach, we can check the data quality without performing additional surveys, thus saving time and money. Furthermore we can monitor the data on a more regularly basis instead of sporadic tests.

The target groups for this report are statistics producers and researchers in Statistics Norway and other statistical offices, as well as other researchers who use administrative registers.

1.2. Data

The results in this report are based on a specially prepared data file, made by micro linking survey data to register data. The survey data are interview responses and technical data from the LFS (Labour Force Survey). The register files are made from a combination of data from several administrative registers. The resulting statistical register is regularly used for official register-based employment statistics. The linking at the individual level uses a universal identification number and reference time variables.

The data sources mentioned each have a variable for labour market status. For simplicity, we define employment status as a dichotomous categorical attribute of an individual. This simplified employment status is represented by a binary variable in each of the data sources. We concentrate on classification errors of the employment status of the individual, meaning whether a person is employed or not. More on the data sources are presented in subsection 2.3.

1.3. Limitations

The target population is limited to residents between 15 and 74 years old, and the data we examine are restricted to the 4th quarter in the years 2005 – 2008. Since the LFS sample is a rotational panel lasting 8 quarters, the effective sample size is about half the number of records in the data file. Most results are averages over multiple years, in order to increase the precision of detailed analyses. If the method were incorporated into a more continuous quality monitoring, some results should be presented in a more aggregated form.

Our approach is to identify conditions where classification errors are more likely, and define these conditions independently in each data source, based on knowledge about the data sources. We propose to utilize existing data sources within limited resources at the cost of not achieving stringent experimental conditions. In this framework, complete information about the true employment status for every individual is out of reach. For practical reasons, we have limited the number of auxiliary variables.

1.4. Focus

The quality of the Norwegian LFS has been studied from several different aspects. We know for instance that nonresponse bias is a major problem, and that non-employed and immigrants are underrepresented in the response sample (Villund 2008).

Here we focus on the classification (measurement) error of one particular variable, the employment status. Additionally, we discuss the data sources in some detail.

For those unfamiliar with the Norwegian official employment statistics, we explain the purpose of making a register-based labour market statistics parallel to the LFS (Labour Force Survey). The LFS is the main resource for continuous monitoring of the labour market and is published monthly. The register-based employment data is used for yearly structural statistics, to make very detailed tables and to deliver micro data for research purposes. The analyses of register data in this report therefore emphasize employment status at the micro level, not total employment rate.

1.5. Outline

Chapter 2 discusses the theoretical framework, while chapter 3 presents definitions and calculations. Chapter 4 examines the results and offers some conclusions.

2. Theory

In classical test theory, measurement error is the difference between measured value and true value. The obvious problem of course is that the true values are seldom readily available. In an experimental design, there are several methods available to estimate measurement errors, for instance parallel tests or re-interviewing. Here, we look for practical methods that can be implemented with the production of official statistics. Comparing existing data sources will also be more cost-effective than performing additional surveys.

Some preliminary studies at Statistics Norway have used data from administrative registers to assess survey-based data (Villund 2009) and vice versa (Foss 2004). There seem to be growing acceptance for the notion that neither data source is the "gold standard" (Kleven et al. 2008).

2.1. Administrative registers

Virtaharju et al. (2003) sums up some main advantages of using administrative registers. In the Nordic Countries, employment statistical data are compiled from population registers and other administrative records. Among the main advantages of register-based statistics are lower production costs, reduction of response burden, and further uses of the available register data. One important feature is the availability of data for smaller geographical areas and small population subgroups. Data quality is also better for items that would otherwise involve long-term memory. The coverage of the data is usually high. There is supportive information used for the detailed coding of occupations, educational data, and employer industry and size. The possibility to link various types of data sets such as demographic and business information is also essential for research purposes.

Longva et al. (1998) points out the need for developing new methods for assessing the quality of administrative register data. They bring up the fundamental problem of determining the “true value” when linking of independent data sources show conflicting values. For instance when we link the employee register with register of unemployed, an algorithm using auxiliary variables decide how to classify the labour market status of people that are registered in both data sources.

Another key issue mentioned in dealing with administrative registers is identification of units. Both the survey- and the register-data used in our study contain unique and universal personal identification numbers. That means linking of records for each person is straightforward. As indicated in chapter 2, linking job-related data requires correct timing of job spells, and the quality of timing-information vary between registers.

2.2. Framework

We use a linked data file that contains two independent employment states for each individual. That means we can compare the two employment classifications for each individual. If one of them were known to be correct, the analysis would be easy.

We first examine briefly the case with three binary employment variables: two measurements and a true value. Table 2-1 lists the eight different combinations that are possible and the resulting evaluation.

Table A: Theoretical combinations of three independent binary variables of employment status.

LFS Employment	Register Employment	True Employment	LFS Evaluation	Register Evaluation
0	0	0	Right	Right
		1	Wrong	Wrong
	1	0	Right	Wrong
		1	Wrong	Right
1	0	0	Wrong	Right
		1	Right	Wrong
	1	0	Wrong	Wrong
		1	Right	Right

However, from theory and earlier observations, it is plausible that both data sources have classification errors. We do not have the true value from some third independent source of employment data; consequently, the evaluation in table A is not feasible.

Even though neither variable represents the true value for every individual, we make some plausible assumptions about the quality in some parts of each data source. Simply put, we wish to find at least some data that are more trustworthy than the rest. From these assumptions, we define a structure of subgroups that ideally should have homogenous classification errors. We emphasize that this group structure is independent from stratification with employment-homogenous strata. Furthermore, the group structures in the LFS and register data are constructed independently, based on theoretical properties and empirical knowledge of each data source. A subgroup where all the classifications were correct in one source could then be used to evaluate the classification in the other source within the same group. Table 2-2 suggests how this can come about.

Table B: The case with two binary measurements of employment status and LFS subgroups.

LFS Employment	Register Employment	*LFS Group*	Register Evaluation
0	0	*Right*	Right
		Wrong	Wrong
	1	*Right*	Wrong
		Wrong	Right
1	0	*Right*	Wrong
		Wrong	Right
	1	*Right*	Right
		Wrong	Wrong

The evaluation in table 2-2 is based on the premise that we are dealing with a binary variable and homogenous groups. If the LFS group value is wrong, then a diverging register value is right. However, we can hardly find groups were every classification is right. Thus, the evaluation itself has measurement errors.

The evaluation is also limited, since it can only be partly achieved by comparing some “good subgroup” in one source with the corresponding individuals in the other. By defining an independent “good subgroup” in the other source, we also assess the quality among the corresponding individuals in the first source. The result will be more a patchwork, than a model for the complete data. Nevertheless, examining large “good subgroups” the analysis covers extensive parts of the labour market population. Assumptions about the data quality in a group should be based on theory and empirical evidence. For instance, both theory and data supports the notion that proxy interviews increase measurement error, although evidence from earlier studies is not overwhelming.

Since the groups are not homogenous with respect to employment, the employment rate difference between corresponding groups serves as an indicator of error. Having defined one group structure, we proceed with the following steps:

1. Identify a group in source A with presumably good quality.
2. Compare the employment rate by subgroups in source B that corresponds to the group in source A.
3. The subgroup in source B that has employment rate closest to the group in source A is deemed the "best".
4. We now use the results from point 1–3 as well as theory to find a group in source B with presumably good quality. This group in B is now defined independently of the data source A.
5. Points 2 and 3 are repeated "in reverse", where source B is used to evaluate some parts of source A. Source B is then divided in a different group structure than source A.

Based on growing knowledge about the sources and empirical studies, these steps can then be repeated with more groups, and thus the quality assessment can be extended. The premises for further development are:

- That we have a sound theory about the data sources, and/or reliable empirical evidence. Since we are comparing employment rate in subgroups, it is desirable to control for employment-related factors, such as age.
- That subgroup structures in the two data sources are not congruent. In order to examine the data independently, and extend the part which is possible to examine at all, the subgroups should preferably not overlap. For instance, young people may be a difficult group to assess in both survey- and register data.

Subsection 2.4 present the group-construction, but first we compare some properties of the data sources, in subsection 2.3.

2.3. Measurement instrument

Since most statistical education deals with sampling and interviewing, the notion of administrative registers as a “measurement instrument” may be unfamiliar. Here we regard the register as a kind of measurement data, and not as the fact sheet. That means administrative registers are similar to survey data in that they have errors. However, administrative registers may have quite different error types than survey data. Several source-specific factors affect the usefulness of a register for statistical purposes. Such factors include the regulations and routines for reporting, and the skills and motivation of those reporting data, and factors concerning those who register the data. For more discussion on this subject, refer to pp. 13–14 in Villund (2009).

Making internationally comparable statistics are rarely the primary purpose of an administrative register. Statistics Norway receives data from many administrative registers intended for other purposes, and the data quality varies accordingly. We take it that any register has both random and systematic errors, resulting in

corresponding classification errors. In addition, the linking and editing of data introduces some errors in the statistical register.

Surveys and administrative registers are obviously different data collection methods, thus having different operational definitions of employment. We recapitulate some of the pertinent differences here, for more details we refer to previous documentation (Bø et al. 2002). Observe that the target population and the theoretical definition of employment are the same in both sources.

Table 2-3 lists the main differences that have a bearing on employment classification errors. "Register" is short for a combined data file made by linking, adding and merging several administrative registers. Note that the *employee register* covers about 80 per cent of the register-based employed. This is fortunately also the most reliable source for register-based employment data.

Figure 2-1: Some main differences between LFS- and register employment data.

	LFS	**Register**
Data collection	Computer assisted telephone interviews with direct questions. Proxy responses from parent or spouse are allowed for most items.	Records filed by employers to various administrative registers, run by several different public agencies.
Size	Quarterly sample of about 24,000 people.	Yearly census with around 3.5 million records.
Units	The unit is technically speaking an interview. In the analysis, we define the unit as person; there is only one interview per subject in each quarterly data file.	In the basic data sources, the unit is usually a job; multiple records about the same person are found at the same time. In the combined file, only one main job is selected for each employed person.
Threshold	The lower limit for being defined as employed is 1 hour of paid work in the reference week. This definition is in accordance with the ILO recommendation. People temporarily absent from work are defined as employed.	The employee register has a lower limit of 4 hours per week and minimum tenure of 6 days. Additional records are added from other registers in order to encompass all employed. These extra registers have no data on working hours.
Reference time	The LFS is a continuous survey, and covers 13 weeks in each quarter. The individuals are allocated specific reference weeks beforehand.	Only one nominal reference week in November are selected each year.
Time variables	The exact dates for each individual's reference week are recorded.	Only the employee register has exact start and stop dates for job spells.
Sub-classification	Direct question are asked to classify employed into: salaried employees, self-employed, and family workers. Part-time workers are asked about their main activity.	Records from the employee register have priority in defining employment. Additional records are added from wage registers in order to encompass more employees. Tax registers are used to identify self-employed. The operational definition is based on entrepreneurial income data.

2.4. Constructing the groups

Survey data

The mode of data collecting in the Norwegian LFS is exclusively CATI (computer-assisted telephone interview). Proxy response by a parent or spouse is allowed for most items, including questions on employment. Theory and empirical evidence indicate that self-response is generally more accurate than proxy response for non-sensitive questions (Groves 2004). On that premise, the response type is an obvious candidate for group construction.

The LFS employment rate based on proxy responses are generally lower than the employment rate based on self-responses. However, the *register-based* employment rate among proxy responses is also lower than the *register-based* employment rate among self-responses. The variation in register-based employment indicates that proxies are not randomly selected. Given the high correlation between the register- and the LFS employment status, it is plausible that the probability of proxy response is associated with true employment status.

Earlier studies have pointed out that the difference in employment rate between proxy- and self-responses is not equal to measurement error (Kleven and Thomsen 2008, Thomsen and Villund 2009). Bearing this in mind, we can better assess measurement quality in the LFS by response type by controlling for register-based employment.

Register data

The subgroups in the administrative register data are constructed independently of the structure in the sample data. The register data production combines several types of administrative registers. Each register has its pros and cons, and the data merging process is aimed at optimizing the employment information. Information of "source-register type" is therefore one of the main candidates for constructing register groups.

We regard the employee register the best data source for employment information: it encompasses the bulk of all employed, and contains many crucial variables including exact start and stop dates. The employment records not found in the employee register are chiefly small jobs and self-employed. The employment information for these t small jobs is collected mainly from wage registers, while the classification of self-employed are based mainly on tax return registers. What we term wage register here is The Norwegian Tax Administration's "End of the Year Certificate Register", which contains reports on wages and other payments from employers and welfare services.

All our analyses hinges on correctly linking data for each person and reference time. One possible source of error is differences in reference time and other time variables. Small jobs not reported to the employee register (with exact dates) but found in the wage register (without exact dates) may be erroneously attributed to 4th quarter. In other words, these cases of divergence between LFS and register are caused by register error and not response error. There may also be some divergence due to fact that the LFS has a wider time-frame than the register-based statistics.

Since we regard the employee register as "high quality data" those records are selected for evaluating the quality in linked LFS-responses.

Control factors

In order to reduce the variation in employment rate in subgroups, we use age as control variable in both types of group structures. Age is an important factor since age is strongly associated with employment status, with proxy probability and with

register source type. It is important then, not to draw the conclusion that response errors are linked directly to age. For instance, both marginal employment and proxy interviews are associated with young age. That mean other substantial factors may be the cause of response error in young age groups.

Special attention has been given to students registered living with their parents, but in reality living on their own. When these students are indirectly interviewed, part-time jobs not known to their parents may be underreported (Solheim et al. 2001). This supports the hypothesis of parents lacking information on the jobs of their offspring. We do not have information in our data about whether the young people are registered students, nor whether they actually live with their parents or not. However, all non-employed and part-time workers are asked about their main activity, and classified as students and several other categories. Data with self-reported main activity are used in a special analysis of students.

Small and short-term jobs are more common among younger people. The employment classification of these jobs often is based on the wage certificate register. This creates an association between age and register source type, and subsequently between age and register classification errors. We do not have evidence that young people *per se*, have poorer response quality in the LFS. Self-employed are generally older than the total employed population (Villund 2005), and the employment information for self-employed are collected mainly from tax registers. These conditions creates another association between age and register source type.

We do not claim at this stage that the groups are homogenous with respect to measurement errors, and anticipate further refinement in future studies.

3. Definitions and Calculations

This chapter lists the definitions and calculations, while chapter 4 presents the results.
As it should be clear from chapter 2, the divergence called "Systematic error" (formula 3-10) can be defined in two ways depending on we are using register to evaluate survey data or vice versa.

Measurements

3-1 $$y_i = \begin{cases} 1 \mid \text{employed in LFS} \\ 0 \mid \text{not employed in LFS} \end{cases}$$ LFS employment status for person *i*

3-2 $$x_i = \begin{cases} 1 \mid \text{employed in register} \\ 0 \mid \text{not employed in register} \end{cases}$$ Register employment status for person *i*

Proportions

3-3 $$p_{00} = \frac{1}{n}\sum_{i=1}^{n}(1-x_i)(1-y_i)$$ Both sources classify *i* as not employed.

3-4 $$p_{01} = \frac{1}{n}\sum_{i=1}^{n}(1-x_i)y_i$$ LFS classify i as employed, but register do not.

3-5 $$p_{10} = \frac{1}{n}\sum_{i=1}^{n}x_i(1-y_i)$$ Register classify i as employed, but the LFS do not.

3-6 $$p_{11} = \frac{1}{n}\sum_{i=1}^{n}x_i y_i$$ Both sources classify i as employed.

3-7 $$\bar{x} = \frac{1}{n}\sum_{i=1}^{n}x_i = p_{10} + p_{11}$$ Register employment rate.

3-8 $$\bar{y} = \frac{1}{n}\sum_{i=1}^{n}y_i = p_{01} + p_{11}$$ LFS employment rate.

Analyses

3-9 $g = p_{00} + p_{11}$ Agreement rate.

3-10 $E(b) = \bar{y} - \bar{x} = p_{01} - p_{10}$ Systematic error.

3-11 $$Var(b) = \frac{p_{01} + p_{10} - (p_{01} - p_{10})^2}{n}$$ Random error.

A note on the calculation of "Random error" (3-11) is included in Annex C page 33.

4. Results

This chapter present and discuss the main findings, as well as some inferences about the overall quality of employment classification. The main results are:

- Overall results
  - Overall agreement at the micro level between survey- and register employment is over 90 percent.
  - At the macro level, register-based employment is about 1.5 percentage point higher than the survey-based. The overall random error is only one tenth of the systematic error.
- Survey quality
  - Proxy interviews lead to an underestimation of employment rate of about 5 percentage points, when controlled for register employment.
  - Proxy interviews seem to give a higher underestimation of employment rate among younger people, but varies with status as a student and with register based employment status.
  - Among students with registered employment status, proxy interviews lead to an underestimation of employment rate of near 15 percentage points.
- Register quality
  - The overall agreement is over 92 percent, but varies with the register source type.
  - Agreement of employment classification varies from 68 percent (Wage register), 87 percent (Tax register) to 97 percent (Employee register)
  - Near 13 percent is classified as employed in the survey, but not in the register data. The figure for the opposite situation varies from 32 percent (Wage register), 13 percent (Tax register) to 3.3 percent (Employee register)

4.1. LFS and the effects of proxy responses

Comparing employment rates

Table 1 shows the LFS- and register employment rate, the overall agreement rate and the systematic- and random error. The four rightmost columns show the combinations of register- and employment status at the micro level.

Table 1: Employment by response type and age. Linked LFS and register data. Average 2005-2008.

		Employment rate			Error		Employment status combinations			
		LFS	REGISTER	Agreement	Systematic	Random	p00	p01	p10	p11
All ages	All responses	0.7147	0.7297	0.9129	-0.0151	0.0010	0.2342	0.0360	0.0511	0.6786
	Proxy response	0.5732	0.6247	0.8763	-0.0515	0.0012	0.3391	0.0361	0.0876	0.5371
	Self response	0.7388	0.7476	0.9191	-0.0089	0.0010	0.2164	0.0360	0.0449	0.7028
15-29	All responses	0.6314	0.6701	0.8430	-0.0387	0.0014	0.2708	0.0591	0.0979	0.5723
	Proxy response	0.4606	0.5408	0.8112	-0.0801	0.0015	0.4049	0.0544	0.1345	0.4063
	Self response	0.6849	0.7107	0.8530	-0.0258	0.0013	0.2287	0.0606	0.0864	0.6243
30-59	All responses	0.8638	0.8568	0.9459	0.0070	0.0008	0.1127	0.0305	0.0235	0.8333
	Proxy response	0.7877	0.8060	0.9254	-0.0184	0.0009	0.1659	0.0281	0.0465	0.7596
	Self response	0.8731	0.8630	0.9485	0.0101	0.0008	0.1061	0.0308	0.0207	0.8423
60-74	All responses	0.3541	0.4072	0.9020	-0.0531	0.0011	0.5704	0.0224	0.0756	0.3317
	Proxy response	0.2904	0.3591	0.9045	-0.0687	0.0010	0.6275	0.0134	0.0821	0.2770
	Self response	0.3641	0.4148	0.9016	-0.0507	0.0011	0.5614	0.0238	0.0745	0.3402

The results imply that the agreement rate is generally very high, over 91 per cent. However, the agreement rate for people under 30 years is 10 percentage points below that of the 30–54 years old. The overall systematic error is 1.5 percentage points in the whole sample. The systematic divergences both in the younger and older age groups are many times that of the middle age group. Random measurement errors are comparatively small, and with no remarkable variation between the subgroups.

Comparing response types

We observe that both employment rate and the difference in employment rate vary with response type. We now define self-responses as a measurement-homogenous group. The proxy effect is defined as the difference in employment rate between a group of proxy responses and the corresponding group of self-responses. It is important to recognize the non-random nature of the proxy selection, mentioned in chapter 2.

Table 2 shows the calculated proxy effect in the total sample and controlled for age and register employment status.

Table 2: Proxy effect by age and register employment. Linked LFS and register data. Average 2005-2008.

	Employment rate			Difference
	All responses	Proxy response	Self response	Proxy effect
Total	0.7147	0.5732	0.7388	-0.1655
15-29	0.6314	0.4606	0.6849	-0.2243
30-59	0.8638	0.7877	0.8731	-0.0855
60-74	0.3541	0.2904	0.3641	-0.0736
Employed	0.9300	0.8598	0.9400	-0.0802
15-29	0.8540	0.7513	0.8785	-0.1272
30-59	0.9725	0.9423	0.9760	-0.0337
60-74	0.8145	0.7714	0.8203	-0.0489
Not employed	0.1333	0.0962	0.1427	-0.0464
15-29	0.1792	0.1184	0.2095	-0.0911
30-59	0.2131	0.1449	0.2250	-0.0801
60-74	0.0378	0.0210	0.0407	-0.0198

The result is that the proxy effect is more than halved when controlling for register employment status. This agrees with earlier findings, and indicates that the difference is the result of a selection effect as well as proxy error. However, the proxy effect is greater among young people, even when controlling for register employment status.

Before we analyze the effect of age, we wish to address some known issues in the heterogeneous register-data file. Younger people tend to have smaller or more short-term jobs. These jobs are less likely to be registered in the employee register, which is the register with best quality on job spells. Consequently, other register-data types are not equally suitable for quality control of short-term jobs, since the timing of job spells is important.

Table 3 shows the proxy effect (again calculated as a percentage points difference in employment rate) by age and register type. The proxy effect for the 15-29 year olds is about 9.6 percentage points for the employee register, compared to 12.7 percentage points for the wage register.

Table 3: Proxy effect by age and register employment type. Linked LFS- and register data. Average 2005-2008.

		Employment rate			Difference
		All	Proxy	Self	Proxy effect
All ages	All registers	0.7147	0.5732	0.7388	-0.1655
	Not register-employed	0.1333	0.0962	0.1427	-0.0464
	Employee register	0.9610	0.9126	0.9674	-0.0548
	Tax return register	0.8639	0.8126	0.8714	-0.0588
	Wage registers	0.6707	0.5709	0.6957	-0.1248
15-29 years	All registers	0.6314	0.4606	0.6849	-0.2243
	Not register-employed	0.1792	0.1184	0.2095	-0.0911
	Employee register	0.9019	0.8224	0.9188	-0.0964
	Tax return register	0.8809	0.7209	0.9103	-0.1893
	Wage registers	0.6606	0.5675	0.6943	-0.1269
30-59 years	All registers	0.8638	0.7877	0.8731	-0.0855
	Not register-employed	0.2131	0.1449	0.2250	-0.0801
	Employee register	0.9834	0.9594	0.9861	-0.0268
	Tax return register	0.9320	0.9268	0.9327	-0.0060
	Wage registers	0.8070	0.6569	0.8274	-0.1706
60-74 years	All registers	0.3541	0.2904	0.3641	-0.0736
	Not register-employed	0.0378	0.0210	0.0407	-0.0198
	Employee register	0.9184	0.9043	0.9201	-0.0159
	Tax return register	0.6589	0.5796	0.6732	-0.0936
	Wage registers	0.3893	0.4200	0.3844	0.0356

For the 30-59 years olds the proxy effect is only 2.7 percentage points for those who are recorded in the employee register. In this age group, nearly 99 per cent of those classified as employed based on the employee register, are also classified as employed based on self-responses in the LFS.

The proxy effect for 30-59 years olds who are classified as employed based on wage register with 17.1 percentage points, compared to 12.7 for the 15-29 years old. It should be remarked that the youngest age group have substantially lower employment rate in both response types. Some variation in the proxy effect should therefore be expected. Nevertheless, the wage-register subgroup shows a higher proxy effect for both age groups. This suggests that the extra proxy effect may in fact be a disturbance in the evaluation method. At least some of the seeming proxy error in the survey may be caused by misclassification in register. To be specific, some register-based employment classification is wrong with respect to reference time, because the source data lack exact timing of job spells. This means the individual was employed sometime during the same year but not in the LFS reference week.

One conclusion is that quality assessment using combined register data should address the variation in quality between the sources. Register data without accurate timing of job spells should be used with care, if at all, for groups where short-time work is predominant, such as young people.

Overall, the proxy response errors probably have limited total impact on the employment rate estimates. For most employees there is very high convergence between the LFS and employee register, which covers most of all employed people. It should also be noted that allowing proxy responses increases the data collection efficiency, and possible includes more young people who would otherwise not respond at all. However, the results should also make it clear that estimates of employment in young age groups may be less accurate. That means detailed labour market figures for young people will have poorer quality, whether we use survey-based or register-based data.

The next subsection deals with a specific group of young people and their activities besides working.

Students

The LFS CATI-questionnaire uses a dependent questioning path to guide the interview. People who during the interview are identified as part-time workers and non-workers are asked about their main activity status. Table 4 shows that there is a higher proxy effect among those who report that they are students. This is also the case among those recorded in the employee register. For young students the proxy effect is higher also compared to young people in general.

Table 4: Proxy effect by age, register source and main activity. Linked LFS- and register data. Average 2005-2008.

		Proxy effect		
		Not register-employed	Employee register	Wage registers
All main activities	All ages .	-0.0464	-0.0548	-0.1248
	15-29	-0.0911	-0.0964	-0.1269
	30-74	-0.0454	-0.0260	-0.1180
Other activities	All ages .	-0.0407	-0.0248	-0.0457
	15-29	-0.1646	-0.0179	-0.0071
	30-74	-0.0471	-0.0258	-0.1218
Students	All ages .	-0.0449	-0.1413	-0.1182
	15-29	-0.0512	-0.1468	-0.1234
	30-74	0.0117	-0.1805	-0.1014

We cannot rule out measurement errors in the "main activity status" itself with this data. In the case of a proxy interview, a parent is asked about whether the son or daughter is working and if not, if he or she is a student. We think that parents have better information about their offspring's student status than about their employment status. Some misclassifications of "main status" due to response errors are possible, but we believe that this constitutes a smaller problem than employment classification errors.
From the results at hand, we conclude that proxy response causes some underreporting of survey-based employment, and that this error is linked more to student status than to young age in general.

As we have observed in all tables, nearly every group exhibits a negative proxy effect, even after controlling for the available factors. Including proxy responses has several advantages, but it could be argued that it should be used with caution for some groups, for instance students.

4.2. Register quality

In this subsection, we turn to the register-based employment data quality. Here we use LFS sample data to evaluate the register data. We must now identify subgroups in the LFS with the best quality and comparability. Based on the differences in

operational definitions and the findings on LFS quality, we do not use the total sample for the register evaluation. As the case was with evaluating the survey data, some subgroups will remain unchartered territory.

Selecting a subsample

To remove the effect of proxy response errors, we use only direct self-responses from the LFS. Furthermore, we address the working hours threshold for the operational definitions of employment. The only register with a working hours variable have a minimum limit of 4 hours/week, compared to 1 hour/week in the survey. In order to remove the effect of different operational definition, we exclude jobs with 1–4 hours/week from the sample.

Therefore, the remaining data is not a random selection of the population, and inference about the total quality is limited. We reiterate that our focus is the quality at the micro level; we do not propose to make alternative macro level (employment rate) estimates.

Some documentation of the total sample and subsample is included in Annex A, table I.

The most pronounced difference is underrepresentation of the youngest age groups, 15-24 years olds constitutes 16.9 per cent of the total sample, compared to 13.4 in the subsample. Over 58.7 per cent of the subsample is between 25 and 54 years old, compared to 55.8 per cent in the total sample. The proportion of men and women in the subsample is about the same as in the total sample.

Comparing register sources

We now use this subsample of the presumably best LFS-data, to evaluate the register employment classification. Table 5 present the overall agreement rates and cross-classifications, for each register-data source-type. Whereas the total overall agreement rate is over 92 per cent, it is under 68 per cent for people defined as employed according to the wage register. In other words, nearly one-third of people classified as employed based on the wage-register, answer in the survey that they did not work at that time.

We draw the conclusion that most of the observed divergence is caused by classification errors in this register-based data source. It is not plausible that so many people in fact work, when they answer themselves that they do not work. The cause of this divergence is more likely invalid linking due to inaccurate timing of job spells in the wage-register based data. This is in accordance with earlier findings (Foss 2004).

The consequence of these errors is low consistency between the classifications at the micro level. Detailed tables of employment situation will have poorer accuracy for the subgroup based on wage-register data. For the tax-register records, the divergence is over 13 per cent. Since the tax-register is used to identify self-employed, detailed analyses of self-employment will be less accurate.

The quality of the employee register is, as expected, much better but the observed divergence is asymmetrical. Only about 1 in 30 people are classified as employed in employee register data but not in the survey data. Of those not classified as employed in register data, about 1 in 8 people are classified as employed in the LFS.

Table 5: LFS and Register Employment states by register source. Subsample of linked LFS- and register data. Average 2005-2008. Percent.

	Employment rate			Employment status combinations			
	LFS Employ-ment	Register Employment	Overall agreement	P_00	P_01	P_10	P_11
Total	73.5	74.9	92.3	21.9	3.2	4.5	70.3
Not register-employed	12.7	0.0	87.3	87.3	12.7	0.0	0.0
Employee register	96.7	100.0	96.7	0.0	0.0	3.3	96.7
Tax return register	86.9	100.0	86.9	0.0	0.0	13.1	86.9
Wage registers	67.8	100.0	67.8	0.0	0.0	32.2	67.8

The divergences can be caused by register time-lag and reference time difference. We therefore look closer at the variation in divergence by LFS reference weeks. We could expect that the quality would be better in the register reference week than in the rest of the quarter. The tax register data (classifying self-employed) is showing a peak in quality around the reference week, although the pattern is not very clear. The employee register data have a more or less constant high quality, while the wage-register data has persistently low quality and considerable variation. More results are included in Annex A.

Register subtype

The shortcomings of the wage-certificate data have been known since the development of the "combined register method" for register-based employment statistics. To improve this part of the data, additional register have be linked to the wage-data. Table 6 present some result for two more register-data subtypes. "Auxiliary registers" are registers with municipal, governmental or private-sector wage data linked to wage-certificate data, while "wage-certificate data only" means unmatched records. Extra information, such as industry and working hours, is collected by linking these additional registers. This method can be considered a kind of deterministic imputation.

Even more important for correct classification of employment status, is accurate start- and stop timing of job spells. The results show close to 80 percent agreement for the Auxiliary registers, compared to around 63 percent for the unmatched records. Unfortunately, for the overall quality, there are almost 3 times more unmatched records than matched.

Table 6: LFS and Register Employment states by register subtype. Subsample of linked LFS- and register data. Average 2005-2008. Percent.

	P_00	P_01	P_10	P_11
Total	21.9	3.2	4.5	70.3
Not register-employed	87.3	12.7	0.0	0.0
Employee register	0.0	0.0	3.3	96.7
Tax return register	0.0	0.0	13.1	86.9
Wage-certificate data only	0.0	0.0	36.9	63.1
Wage registers auxiliary	0.0	0.0	20.7	79.3

4.3. Further development

The approach outlined here could form a starting point for developing quality monitoring of the employment data. We have used only data sources that are readily available in the statistics production, with no extra data collecting.
One line of development is to check up more variables, such as: whether the job is permanent or temporary, different types of working contracts, working hours, and how long the job is held.

The general method of linking independent labour market data sources can of course be applied to other variables, for instance the classification of employed into employees and self-employed. Earlier studies include working hours (Villund 2009) and wages (Braathen 2004). The latter study include a thorough discussion about linking job data from the LFS and register at the micro level.

References

Bråthen, M.: "Collecting data on wages for the Labour Force Survey" Statistics Norway Documents 2004/07

Bø, Tor Petter and I. Håland: "Dokumentasjon av Arbeidskraftundersøkelsen" Statistics Norway 2002/24

Foss, Aslaug Hurlen: "Kvaliteten i arbeidsmarkedsdelen i Folke- og boligtellingen 2001" Statistics Norway 2004/4

Groves, Robert M., F. J. Fowler Jr., M. P. Couper, J. M. Lepkowski, E. Singer and R. Tourangeau: "Survey methodology" Wiley, Hoboken New Jersey, 2004

Kleven, Øyvin, B. O. Lagerstrøm and I. Thomsen: "A simple model for studying the effects of proxy interviewing" Paper presented at the 17th annual International Workshop on Household Survey Nonresponse, Omaha, Nebraska, US, August 2006. Published in Statistics Norway Documents 2008/2

Longva Svein, I. Thomsen and P. I. Severeide: "Reducing Costs of Censuses in Norway Through Use of Administrative Registers" International Statistical Review / Revue Internationale de Statistique, Vol. 66, No. 2 (Aug., 1998), pp. 223-234 International Statistical Institute

Moore, Jeffrey C.: "Self/proxy Response Status and Survey Response Quality" Journal of Official Statistics, 1988, Vol.4 No.2, 155-172

Solheim, Leif, I. Håland and B.O. Lagerstrøm: "Proxy interview and measurement error in the Norwegian Labour Force Survey" Paper presented at the 12th International Workshop of Household Survey Nonresponse, Oslo, 12-14 September 2001

Thomsen, Ib, Ø. Kleven and L. C. Zhang: "Dealing with non-sampling errors using administrative data" Paper presented at the IAOS Conference, Lisboa 2007

Thomsen, Ib and L. C. Zhang: "The Effects of Using Administrative Registers in Economic Short Term Statistics: The Norwegian Labour Force Survey as a Case Study" Journal of Official Statistics, Vol.17, No.2, 2001. pp. 285-294

Villund, Ole: "Measuring working hours in the Norwegian Labour Force Survey. A pilot study of data quality using administrative registers." Statistics Norway Reports 2009/3

Villund, Ole: "Immigrant participation in the Norwegian Labour Force Survey 2006-2007" Statistics Norway Document 2008/7

Virtaharju, Markku, M. Åkerblom, M. Roessingh: "The Measurement of Knowledge Stocks and Knowledge Flows In the New Economy" Report to The New Economy Statistical Information System (NESIS) 2003 ref. D5.6.3

ANNEX A

Table I: sample size by age and gender. Linked LFS and register data and subsample of self-responses and jobs with working hours at least 4 hours/week. Average 2005–2008.

		Sample size		Age percent		Gender percent	
		Total sample	Subsample	Total sample	Subsample	Total sample	Subsample
Total …		84757	71465	100.0	100.0	100.0	100.0
	15-24 years old ….	14307	9605	16.9	13.4	100.0	100.0
	15-19	7599	4496	9.0	6.3	100.0	100.0
	20-24	6708	5109	7.9	7.1	100.0	100.0
	25-54 years old …	47305	41967	55.8	58.7	100.0	100.0
	25-29	6346	5601	7.5	7.8	100.0	100.0
	30-39	15837	14154	18.7	19.8	100.0	100.0
	40-54	25122	22212	29.6	31.1	100.0	100.0
	55-74 :	23145	19893	27.3	27.8	100.0	100.0
	55-66	16778	14503	19.8	20.3	100.0	100.0
	67-74	6367	5390	7.5	7.5	100.0	100.0
Men …		42725	35672	100.0	100.0	50.4	49.9
	15-24 years old ….	7407	4879	17.3	13.7	51.8	50.8
	15-19	3920	2284	9.2	6.4	51.6	50.8
	20-24	3487	2595	8.2	7.3	52.0	50.8
	25-54 years old …	23533	20804	55.1	58.3	49.7	49.6
	25-29	3318	2935	7.8	8.2	52.3	52.4
	30-39	7870	7057	18.4	19.8	49.7	49.9
	40-54	12345	10812	28.9	30.3	49.1	48.7
	55-74 :	11785	9989	27.6	28.0	50.9	50.2
	55-66	8459	7233	19.8	20.3	50.4	49.9
	67-74	3326	2756	7.8	7.7	52.2	51.1
Women		42032	35793	100.0	100.0	49.6	50.1
	15-24 years old ….	6900	4726	16.4	13.2	48.2	49.2
	15-19	3679	2212	8.8	6.2	48.4	49.2
	20-24	3221	2514	7.7	7.0	48.0	49.2
	25-54 years old …	23772	21163	56.6	59.1	50.3	50.4
	25-29	3028	2666	7.2	7.4	47.7	47.6
	30-39	7967	7097	19.0	19.8	50.3	50.1
	40-54	12777	11400	30.4	31.8	50.9	51.3
	55-74 :	11360	9904	27.0	27.7	49.1	49.8
	55-66	8319	7270	19.8	20.3	49.6	50.1
	67-74	3041	2634	7.2	7.4	47.8	48.9

Diagram I: Agreement rate by reference week and register data source type. 2005–2008

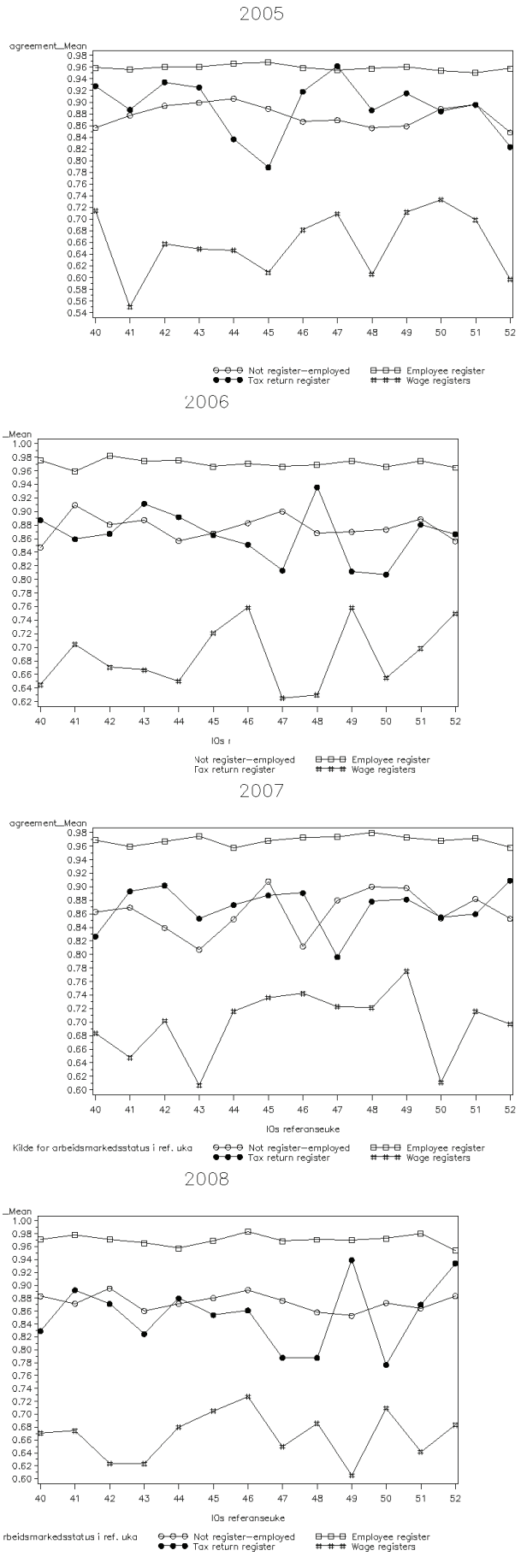
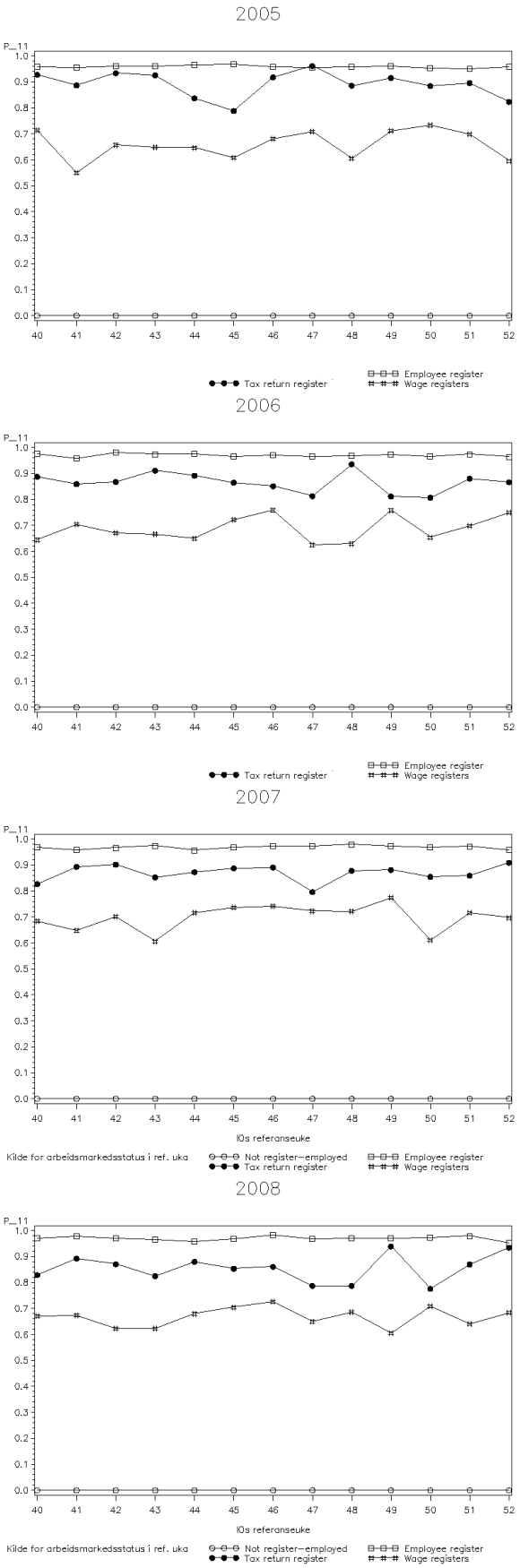


Diagram II: Rate of employed in both LFS and register, by reference week and register data source type. 2005–2008



ANNEX B

Some remarks on employment rate estimates

In the production of official statistics based on the LFS, a somewhat complicated weighting-scheme is used. A key component in the weighting scheme is micro-linking of register-based employment data. Post-stratification by register-based employment reduces some of the nonresponse bias, as well as it substantially reduces variance.

However, in this study we tried to isolate the measurement errors, not involving the ordinary weighting process. All figures are unweighted averages from the special file consisting of survey data linked to register data from 2005–2008. Consequently, the employment rates presented in this report are not entirely consistent with official statistics.

Note also that register-based employment in the regular weighting scheme is not exactly the same as register-based employment in the linked special file used in this report. The register data used for the regular weighting procedure has to have a shorter production time than the statistical register. It therefore encompasses a slightly smaller proportion of employed people. Consequently, the employment rate based on the ordinary variable is lower than from the variable in the specially prepared file.

In the official Norwegian employment statistics the total number of employed in the register-based statistics is adjusted to match the 4th quarter LFS employment estimate. However, the number of survey-employed is not equal to the number of register-employed in the specially linked sample file, due to differences in the data material.

ANNEX C

A note on the formula for random measurement errors

We define the individual measurement error as the difference between true value and measured value. The random error is defined as the variance of these differences. We show here how this is derived from general formulae for variance of a difference in proportions.
Employment status is here defined as a dichotomous categorical variable, so we apply formulae for binominal distribution.

$p_k = \frac{n_k}{n} = \frac{1}{n} n_k$ Definition of a proportion

$\mathrm{Var}(\mathrm{n_k}) = \mathrm{Var}(\mathrm{np_k}) = \frac{1}{\mathrm{n}^2}\mathrm{Var}(\mathrm{p_k})$ Variance of a cell count

$Var(n_k) = np_k(1-p_k)$

The variance of a proportion thus:

$$Var(p_k) = \frac{1}{n^2} np_k(1-p_k) = \frac{p_k(1-p_k)}{n}$$

Proportions of the same set are not independent, and the covariance between two proportions is:

$$Co\,\mathrm{var}(n_k, n_l) = -np_k p_l$$

$$Co\,\mathrm{var}(p_k, p_l) = \frac{1}{n^2}(-np_k p_l) = -\frac{p_k p_l}{n}$$

The variance of a difference of proportions follows thus:

$$Var(p_1 - p_2) = Var(p_1) + Var(p_2) - 2Co\,\mathrm{var}(p_1, p_2) =$$

$$\frac{p_1(1-p_1) + p_2(1-p_2) - 2(-p_1 p_2)}{n} =$$

$$\frac{p_1 - {p_1}^2 + p_2 - {p_2}^2 + 2p_1 p_2}{n} =$$

$$\frac{p_1 + p_2 - ({p_1}^2 - 2p_1 p_2 + {p_2}^2)}{n} =$$

$$\frac{p_1 + p_2 - (p_1 - p_2)^2}{n}$$

Substituting P_{01} and P_{10} give the formula 3-11 in page 17.
To arrive at the figures in the tables we take the square root, in order to have the same unit as e

List of figures

List of tables