

Ole Villund

**Measuring working hours in the Norwegian
Labour Force Survey**

A pilot study of data quality using administrative registers

Reports This series contains statistical analyses and method and model descriptions from the different research and statistics areas. Results of various single surveys are also published here, usually with supplementary comments and analyses.

©Statistics Norway, February 2009 When using material from this publication, Statistics Norway shall be quoted as the source.	Symbols in tables	Symbol
	Category not applicable	.
	Data not available	..
	Data not yet available	...
	Not for publication	:
	Nil	-
ISBN 978-82-537-7529-6 Printed version	Less than 0.5 of unit employed	0
ISBN 978-82-537-7530-2 Electronic version	Less than 0.05 of unit employed	0.0
ISSN 0806-2056	Provisional or preliminary figure	*
Subject: 06.90	Break in the homogeneity of a vertical series	—
Print: Statistics Norway	Break in the homogeneity of a horizontal series	
	Decimal punctuation mark	.

Preface

Statistics Norway has several opportunities to link survey sample data with administrative register data at the micro level. This report presents a method for assessing the quality of contractual working hours in the Norwegian Labour Force Survey (LFS). Responses of contractual hours of work from the employees are compared with the contractual hours of work reported by employers. Most employee jobs are reported to a central governmental register and several variables are used to identify each job. To the extent that the micro linking is correct, we can compare two independent variables for the same attribute. The quantitative assessment is based on the assumption that the register data is correct, and so this assumption is also examined.

The results indicate that LFS survey-based working hours generally are longer in the survey than in the register. Compared with register data reported from employers, some groups of employees seem to over-report contractual working hours. The differences vary with working time length, generally in jobs with few hours per week. Also the type of working time arrangements between industries and occupations affects the working hours data quality. Not surprisingly are the measured errors greatest in jobs without a fixed number of hours per week.

Contents

Preface	3
1. Introduction	5
1.1. Summary	5
1.2. Purpose	6
1.3. Subject matter	6
1.4. Some international findings	7
1.5. Outline	8
2. Data	9
2.1. LFS sample survey data.....	9
2.2. Employee register data	10
2.3. Linking	10
2.4. Limitations	11
2.5. Comparability	12
3. Method	13
3.1. Assumptions	13
3.2. Discussion of quality aspects	13
3.3. Calculations.....	15
3.4. Analysis	16
3.5. Limitations	16
4. Results	17
4.1. Summary	17
4.2. Job and employer characteristics	17
4.3. Employee and interview characteristics	19
References	22
List of Tables	23

1. Introduction

1.1. Summary

We have linked survey sample data with employee data from administrative registers. The combined data sources are used to assess the quality on measured contractual working hours in the Norwegian Labour Force Survey (LFS). The data files are linked at the job-level. Information on contractual hours of work in the LFS is given by the employee himself or by a member of his household. In the register data, the employers report information on contractual hours of work to a central governmental agency. We therefore have two independent sources of contractual working hours. Measurement error is estimated from the differences between the two independent working hours variables. The quantitative assessment is based on the assumption that the register data is correct, but this assumption is discussed at some length.

The results indicate that LFS survey-based mean contractual working hours generally are higher in the survey than in the register. Compared with register data reported from employers, some groups of employees seem to over-report contractual working hours. The differences vary with working time length and type of working time arrangements between industries and occupations. The measurement differences are greatest in jobs without a fixed number of hours per week, and generally in jobs with few hours per week. Proxy interviews¹ result in even larger divergences.

- Overall, the survey based contractual working hours are systematically longer, and vary more, than the register based.
- Systematic measurement error, as we have defined it, is generally much larger than random measurement error.
- For jobs with "normal" working hours and the same number every week, the relative measurement errors are quite small.
- For jobs with shorter hours, the survey seem to overestimate working hours, while it seems to underestimate the working hours for jobs with long hours.
- Jobs without an agreement of working hours have by far the largest systematic measurement errors regardless of the number of hours reported.
- Teachers in primary schools have somewhat larger systematic measurement errors than other occupations at this level. The pattern is the same as the general on: the overestimating shorter working hours, and underestimate the longer working hours.
- Teaching professionals have about the same overall systematic measurement errors as other professional occupations with fixed working hours, but a larger underestimation of the longer working hours.
- There is great variation in systematic measurement errors between industries, also when controlling for contractual working hours length and working hours arrangements.
- Proxy interviews seem to cause more overestimation of contractual working hours, especially in jobs with varying contractual working hours and long hours.

The observed differences can be attributed to measurement error in the survey, given that the suggested evaluation method is valid. We therefore discuss other factors that can offer alternative explanations of some of the divergences. One such factor is errors in the administrative registers, for instance due to errors in reporting to the register or errors in the punching of reported data. Another factor is what we might call evaluation error, for instance due to error in the linking of survey data with register data. In this pilot study we focus on discussing the evaluation method and some descriptive statistical results. This is meant to form a basis for further

¹ Proxy interviews, whereby a spouse or parent are interviewed on behalf of the subject, are used in the Norwegian LFS in order to increase response efficiency. Proxy interviews constitutes about 15 per cent of the responses.

studies, in which more details and some fundamental causes of measurement errors could be investigated.

1.2. Purpose

Statistics for working hours are important both to researchers and policy makers, both from a social and economic perspective. Working hours data are for instance used to measure labour productivity and to examine the social implications of working time arrangements.

Statistics Norway has looked at methods for evaluating the quality of working hours data in the Norwegian Labour Force Survey (LFS). The purpose of this document is to discuss methods for evaluating measurement errors in working hours, and to present some results from recent LFS data.

This document is primarily intended for users and makers of official statistics of working hours, and researchers that uses LFS data. The original purpose was to provide background information for a presentation to the European LFS workshop in Stockholm 2008.

We combine data from administrative registers and survey samples by merging data at the job-level. The document may therefore be of interest in a wider discussion about the benefits and challenges of job-linked labour market data from different sources, and to those interested in measurement errors and survey quality in general.

1.3. Subject matter

One can distinguish between several essentially different types of "working hours", and for instance Evans et al. (2001) distinguish between four concepts. "Actual working hours" is perhaps the most sought-after in economic analysis, for instance in estimating labour productivity. "Usual working hours" are distinguished from "legal working hours" and "normal working hours", by that the former is usually obtained from surveys, and the two latter can be derived from laws and agreements. "Legal working hours" and "normal working hours" might be called "formal working hours" to distinguish these concepts from the empirical working hours.

Many authors have examined one or more of the same concepts, but the words for the same underlying concept vary. We suggest that the many working hours concepts and terms can be functionally divided in two major categories:

- Working hours that vary with irregular events such as overtime, sickness and holidays.
- Working hours that do not vary with irregular events.

The second category can encompass several different concepts, and the variable of interest in our study belongs to them. Perhaps "regular working hours" might be a general term for this group of concepts. Note that even though regular working hours do not vary with irregular events, some types may vary regularly. This variation follows a pattern such as modular working hours arrangements, shift work and seasonal work.

It is possible to include unpaid working time in the working hours concept. It could be a source of error if some employees include unpaid working time, while other report paid hours only.

Our variable of interest is the LFS data contractual working hours. This concept is of the "regular" type, akin to the types called "usual", "basic", "agreed", "contractual" or "scheduled" working hours. It should be noted that "usual" working hours often include regular overtime, which is not the case in our concept.

In previous LFS publications the variable was called "settled working hours", but in this report we have chosen a more universally accepted term.

In the LFS, employees that respond that they have the same number of contractual working hours every week, report this *fixed weekly number* directly. Employees with jobs where the contractual working hours vary from week to week are asked to provide the *average agreed weekly* working hours. Finally some employees respond that they do not have an agreement of working hours, neither a fixed number nor a regularly varying number. They are asked to provide the *average weekly* working hours.

We define the variable of interest as "contractual working hours" which is in fact the three variables mentioned above, folded into one. The LFS includes questions both on the working time arrangement and the working hours per week. We can identify the working hours arrangements by an additional categorical variable and the measurement quality can be analyzed according to the subtypes of contractual working hours.

1.4. Some international findings

Tijdens et al. (2007) have studied 26 different questionnaires from large-scale national surveys. They point out that the heterogeneity of working hours measurements has been debated for over 40 years, and differences in measurements are still among the major methodological problems for this labour market attribute. Their study concludes that at present, there is great variation in both the terminology and the measurements of working hours. We conclude that it is sensible to go to some length in explaining explicitly what we are trying to measure, a conclusion that has shaped this study.

Williams (2002) compares estimates of hours worked in the UK from two independent sources, the Labour Force Survey (LFS) and the New Earnings Survey (NES). The two estimates are independent data because the LFS contains responses from the employees, while the employers are required to fill out the NES questionnaire. When comparing the estimates for basic working hours the LFS figures are generally higher than the corresponding figures in the NES. The differences are especially large in professional occupations, with more than ten per cent longer basic working hours in the LFS estimates. Of occupational subgroups, teachers had the greatest difference with 20 per cent longer basic working hours according to the LFS than the NES.

Evans et al. (2001) report on international trends in working hours for OECD countries. In most countries, the long-term decline in annual hours worked has slowed down in recent years. The OECD report emphasizes that the evidence for widespread flexibility is limited, but that there is trend with increasing use of flexible working arrangements. Employers and employees have different views on which arrangements of working time that are "flexible". Employees and employers can have quite different preferences in part-time work or working hours schedules.

Lemaitre (1988) reports on response errors in the Canadian LFS by examining a re-interviewed subsample. A part of the subsample was both directly interviewed and interviewed by proxy about the same reference week. Of interest to us is the proportion of inconsistent answers when the interview and re-interview were of two different types, compared to both being proxy interviews or both direct self-response.² In cases where both interviews were self-response or both proxies, the inconsistency was about the same, 31.8 and 30.3 per cent respectively. Cases where one interview or re-interview was proxy and the other one direct, gave a considerably higher inconsistency, 42.3 per cent. The overall inconsistency were 35.8 per

² We only refer to the question about usual working hours for main job (Q13).

cent when measured at a "quasi-continuous" scale, but only 23.8 per cent when measured at a collapsed level with seven categories. The result of the categorization was about the same for proxy interviews and self-response.

Klapfer et al. (2007) report on an Austrian pilot study where a small sample of people were selected for cognitive interviews about working hours concepts. This method entails in-depth interviews where the respondents own thoughts around the concept are important. The results for "hours usually worked" are interesting to us, even though it doesn't map our variable of interest exactly. In the Austrian LFS "hours usually worked" include the typical hours per week, including recurrent overtime, not the average working hours including exceptional overtime. When given a concrete example with working hours figures for several weeks, only half of the respondent calculated the correct value. However most respondent seem to have a common understanding of "usual working hours". The divergences arise with the inclusion of some, but not all overtime or additional hours. Furthermore the study indicate that self-employed have more difficulties than employees in specifying usual working hours. Typical reasons given for this were freedom to organize the hours and highly variable working hours.

What do these findings mean to us at this point? The different perspective of employers and employees may have impact on the quality of working hours data. This may be important in a study such as this, where employee survey data and register data reported from the employers are linked. The trends in flexible or irregular working time are of course importance to estimating survey quality, because it is more difficult to measure usual working hours in unusual working time arrangements. Proxy interviews cause further measurement complications, although one study indicates that proxy errors are small compared to the total level of measurement error. The problems concerning working hours for self-employed, should not impact our study since we are concentrating on employees.

1.5. Outline

The general topic of this study is data quality, more specifically errors of observation, and the measurement quality at the job-level. The quality of survey-based official statistics of working hours will also depend upon other aspects, such as sample design and estimation methods. But at this point in time we wish to concentrate on the data quality at the job-level. If the input data is not good enough, post-survey methods can only compensate for this to a limited degree. Furthermore we now have available register data linked with survey data, both data sources containing a contractual working hours variable. This data material gives us an extraordinarily opportunity to study measurement quality at a detailed level.

The quality of working hours data will depend on attributes of the respondent and the data collection method. The quality of the evaluation method itself will depend on two other factors: the job-linking and the comparability of the two working hours concepts. This can be formulated in two questions: Are we comparing the same unit? Are we comparing the same variable? Chapters 2 will deal with the linking challenges and the comparability of the data. Chapter 3 will discuss some theoretical quality aspects, as well as documents the actual calculation method. Finally will chapter 4 give an overview of the results, and draw some conclusions about the quality.

2. Data

Data summary

The analyses are based on sample survey data merged with administrative register data at the micro level. Each of the two independent data sources contains its own variable with contractual working hours value for each job. The statistical unit is the current main job, and the target population are employees. The analyses include quarterly data from the period 2004 – 2007.

In the rest of this subsection we discuss some specifics of the data sources and variables. The sample in this study contains job-linked information from two well-documented data sources, both routinely used in the production of official Norwegian labour market statistics. The two next sections describe some characteristics of the data sources, while section 2.3 deals with the data linking.

2.1. LFS sample survey data

The Norwegian LFS is a relatively large survey with a sample size of 24.000 units per quarter. The sampling design is a one-stage cluster sampling of families, where every family member aged 16-74 years old are selected in the sampled families. The sampling plan is a continuously rotating panel, where every subject is interviewed in a specified week in each quarter, in all eight times over a two-year period. The LFS data are organized in quarterly files that consist only of one interview per person per quarter. The panel data makes it possible to compare information for the same job from different reference times.

The measurement error figures are estimates themselves, and as such have a limited precision. In order to calculate the precision, we need to consider the effective sample size or number of independent units. Quarterly data set have an effective sample size equal to the number of records. This is not the case for annual averages, where the effective sample size is less than the number of records. However, for some subgroups analysis we calculate averages for several years in order to give more precise quality figures. The precision estimate should then be based on the reduced effective sample size.

We have chosen quarterly data from the period 2004 – 2007 in order to use recent data, and to have a large sample size. This reference period includes two years with the former and current LFS questionnaire respectively. As mentioned, the LFS questionnaire was revised in 2006, but the questions about working hours was not changed. Possible breaks in the time series or changes in the quality must therefore be attributed to other changes in the questionnaire, survey protocol and practical routines³.

The sample survey data contains working hours reported by self-response or by proxy interview with a family member. The measurement instrument is a computer-assisted questionnaire. Only telephone interviews are used, and proxy interviews are allowed when the proxy is the spouse or parent of the interview subject. The module concerning working hours consists of several questions, and filters to guide the interview to the appropriate combination of questions. The reference period for a LFS interview is a specified week in each quarter, and the measurement time frame of working hours is the whole week. Thus the LFS measurement refers to a sum of contractual working hours per week, for the job.

³ We have calculated that the measurement error was significantly lower in 1. Quarter of 2006 than 4. Quarter of 2005. The two-year average measurement error for 2004-2005 is lower than the two-year average for 2006-2007, when using all records in the combined data. However simulations to control for effective sample size (unique record pr person) indicate no significant change.

The nominal precision is 1/100 hour or 36 seconds, but due to the imputation method the final data contain more decimals.⁴

2.2. Employee register data

The register-based contractual working hours data used in this study comes from a large administrative register, which we for short call "employee register". This is the Norwegian Central register on Employers and Employees, which is run by the Norwegian Labour and Welfare Service. The employer- and employee data are used by several other public agencies, as well as Statistics Norway for official statistics and research.

Employers are required to report contractual working hours to the employee register when hiring new employees. All changes in working hours in existing jobs should also be reported. Most employers send registrations to the employee register electronically from their payroll systems. The authorities works close with the commercial developers of IT-systems in order to get a smooth delivery of data with a minimal burden on the employer. A growing number of small businesses report to the employee register by an Internet web form, though some still use the paper form. Reporting working hours data to the employee register has been mandatory for several years, but it always takes considerable time before a register of this size is complete with data of good quality.

The working hours in the register data is referring to the sum of hours per week. If a job contract stipulates a percentage of a position and not a number of hours per week, the employer should report the proportion multiplied by 37.50. If a job entails a varying number of hours from week to week, the average number of hours per week should be reported. For sporadic work or very short-term jobs, the actual number of hours is reported, often estimated from the salary paid.

The unit is a job spell, and the contractual working hours for each job should be reported separately regardless if the employee has multiple jobs. This means that one person can have more than one record in the register at a reference time. The formal precision for reporting is 1/100 hour or 36 seconds.

2.3. Linking

To match records at the job-level, we would ideally want to have unique employee identification, employer identification and exact reference time. For a more detailed discussion on job-level linking of LFS data with various administrative registers, see Bråthen (2004). Several other Statistics Norway projects have also involved what we may call a step-wise or gradual linking process. Those familiar with relational databases may compare this to the process of selecting attributes for a *composite primary key*. We discuss the ideal key, the candidate attributes (available variables) and some findings on measurement sensitivity to the linking, before we document the actual key used in the analyses.

Linking data at the job level requires a positive identification of jobs, which in theory would require the same definition of job in each data source. In this study we will concentrate on the ideal and available data attributes, not the definitions of jobs or definition of employees.

Table 2.1 gives an overview of unique job identification, the ideal and the available variables by data source.

⁴ Technical note: in the register data, the numerical value was transformed from a 4-character type. This transformation should in theory give the same precision as the survey data type. In the actual IT-system however, the differences in decimals had to be resolved by using a "fuzz-removing" function (ignoring differences smaller than 10^{-12})

Table 2.1. Job identification, by unit, attribute and data source

Unit	Ideal attribute	Available in register	Available in LFS
Employee	Unique personal identification.	All residents are identified by a unique number code.	All residents are identified by a unique number code.
Employer	Unique identification, preferably of local business unit.	Unique number codes for employers at enterprise- and local business unit level.	Firm name, sector and economic activity.
Time	Exact job spell, start and stop dates	Start dates, and register reference time.	Reference week, and start year/month.

The first steps in the linking procedure are:

- We simplify the reference time unit to quarter + year. The justification for not using actual dates, are for practical reasons and some quality issues.
- Employees are identified by the resident number codes (PID).
- We define exactly one main job for each employee. Main job is the job with the longest hours, or in ties the most recent job. By selecting only the main job in each data sources, we can link the data at the person-level.

This initial data set now contains one record per employee per quarterly data, but with a considerable number of mismatched jobs. We need more control over the precise job-linking, especially the employer unit. The next steps in the linking procedure are to construct three control parameters:

- Is the NACE code for industry equal in the two data sources?
- Is the ISCO code for occupation equal in the two data sources?
- The LFS data also contains a manually coded variable that indicates if the register based business information is the same as in the LFS. We define the same business as equal local unit or equal enterprise unit.

We now perform a sensitivity analysis by calculating the mean differences, by these three linking control questions. Table 2.2 shows the results, which clearly indicates the effect of better job-linking on the seemingly "measurement errors".

Table 2.2. Difference in mean contractual working hours, by job-level data linking parameters. LFS/Register data. Average 2004-2007. Hours

NACE	ISCO	Business		
		Total	Unequal	Equal
Total	Total	1.49	3.05	1.12
	Unequal	1.54	3.53	1.03
	Equal	1.42	2.21	1.26
Unequal	Total	4.24	5.28	2.19
	Unequal	4.67	5.86	2.10
	Equal	2.32	2.21	2.47
Equal	Total	1.24	2.15	1.09
	Unequal	1.14	2.10	0.98
	Equal	1.38	2.21	1.23

From these results, and considering the sample subgroups sizes, we define a final data set from the initial person-level set. In the final data set only records with matching business and matching NACE is selected. We choose to disregard ISCO as linking parameters because of a large proportion of missing occupations in the register data for the early years. The missing values are heavily biased due to administrative reasons, and occupation was not a useful to the linking process.

2.4. Limitations

A considerable number of people have multiple jobs with different contractual working hours. In order to simplify both the linking and the analysis, we select only the main job for each person. This means, for instance, that we do not consider the data quality of the individuals combined working hours. There may still be errors in the job-linking, resulting in overestimation of measurement errors. Linking error may include wrong timing, for instance for persons who have changed jobs within the same company or have changed contractual working hours in the same job. Some employer linking error may also occur, while errors at the employee level are probably negligible.

All types of jobs have a contractual working hours value in the LFS. However, we limit the target population to *employee jobs* because we do not have independent data sources for the working hours of self-employed and family workers. We can only speculate that self-employed have a different working time structure altogether, for instance with no fixed or planned hours, and suggest that the data quality for this groups should be analyzed separately.

2.5. Comparability

The two different contractual working hours data seem to cover much of the same information content, and the variables are the same data type and magnitude. However, in the following paragraphs we point out some known differences in the formal or objective characteristics of the two types of contractual working hours. Subjective or individual factors affecting the quality will be discussed in chapter 3.

An important dissimilarity is the definition of an employee job, specifically the minimum time limits. The administrative register require only jobs with at least 4 hours pr week and a duration of at least 6 days, to be reported. The definition of employment in the LFS has a lower limit of only 1 hour per week. This means that a number of marginal jobs are measured in the survey, but not reported to the register. The difference in minimum limits can have different types of consequences for the evaluation of data quality, such as:

- The target groups are not congruent, which increases the possibility of linking errors.
- The measurement differences will be asymmetrical.
 - A job with 4 hours/week reported to the register can have a LFS response value both below and above 4 hours/week.
 - Jobs with LFS response value below 4 hours/week will have minimum 4 hours/week reported to the register.

The maximum hours per week is also different with a limit of 168 in the LFS and 99 in register, but this affect very few jobs.

The frequency and regularity of the data collection process is different in the survey and the administrative register, which may have some bearing on the quality differences. As we have mentioned the subjects in the LFS are interviewed every quarter. Previously interviewed participants who have been classified as employed, are in subsequent interviews asked if they have the same "position"⁵. If they respond that they have the same position, several questions about the job are not asked, but questions about working hours are asked again. On one hand, we want to minimize interview time and random errors. The procedure should pick up changes in the working hours also for people that report that they have the same job.

The quality of working hours data for both states and transitions, can also be affected by the mode of reporting to the administrative register. A large company will typically report working hours, occupation and other employee data directly from the payroll IT-system, at a monthly basis. Also many small companies report electronically and regularly from similar large IT-systems, because they buy their administrative services from specialized firms. Many small or mid-sized companies use web- or paper form to report employee data to the register. These companies will have to actively initiate its reporting procedure when there are staff changes. There may be more variation as to what kind of routines and what level of compliance these companies have to the mandatory employee reports. This doesn't necessarily mean that large companies always report working hours at a better quality. Especially the data quality for transitions may be affected by how "automatic" the reporting routines are, i.e. how well the data are checked before it is sent to the administrative register.

⁵ The wording is carefully chosen in Norwegian, but all the connotations and possible overlapping between concepts doesn't translate exactly to English. Norwegian "stilling" = position, "yrke"=occupation, "jobb"=job. The latter also goes by other names, depending on the context and operational definition.

3. Method

In this report we have used pretty straightforward statistical methods for a quantitative evaluation. The bulk of chapter 3 is of a theoretical nature, and some readers may want to proceed directly to section 3-3 for documentation of the calculation methods.

3.1. Assumptions

In chapter 2 we mentioned several aspects that can affect the quality of working hours data, and the possibility of errors also in administrative registers. But in order to quantify the measurement errors we must make some basic assumptions. The method we propose is based on the premise that the employer-reported working hours from the administrative register is the correct value, and that only the survey measurement can have errors. That doesn't mean we generally believe that employer-reported or other register-based data are free from errors. It means only that the quantitative part of the evaluation is simplified when we assume that one of the two variables is the correct one. We will therefore address some of the characteristics of administrative register from a non-quantitative perspective.

Our study includes data from administrative registers, a data source often viewed as complete and "certain", at least in the sense that it doesn't have a sample error. However, it is known that administrative registers have their shortcomings for statistical purposes. Factors that can affect register-based working hours data include:

- The definitions and formal regulations for the register.
- The supporting information given to the employers.
- The data type and form of data delivery, such as paper, web or payroll-system.
- The individual company's routines for reporting to the register.
- Subjective individual factor affecting data quality. Administrative data is processed by humans, and subjected to human errors.

An evaluation of the administrative registers themselves could include assessment of the definitions, regulations and routines for reporting working hours from the employers to the authorities, perhaps followed by a business survey aimed at specific groups.

3.2. Discussion of quality aspects

It can be useful to recapitulate some theoretical ideas, even though our method doesn't pick up all the subtleties at this time. The main point is to try to place administrative registers in a more general analytical framework. Survey textbooks, such as Groves et al. (2004) define the response, the measurement variable and the construct (the underlying concept) as three different entities. This division makes it possible to define different quality aspects: the validity and reliability of the measurement instrument and errors in the measurement process.

The validity of the measurement is how well it reflects the construct or the underlying concept, the "true value". There are however a host of different validity "types", i.e. methods for evaluating validity. In the LFS the validity would be whether the questionnaire module really measure contractual working hours as we have defined this concept. The wording of the questions, the filtering and interview sequence can all affect the validity. The validity can be quantified as the correlation between the measurement and the real value over a number of trials.

If we have repeated measurement over a number of trials, we can also define the reliability of the measurement instrument. The reliability is how well it measures the same value from time to time, defined as the true proportion of variance in obtained scores (Guilford 1954). Reliability is usually quantified by some kind of correlation-like coefficient. As mentioned in chapter 2, previously interviewed

participants who have been classified as employed, are asked about their working hours in subsequent interviews, even if they answer that they have the same position. Thus the data make it possible to estimate the reliability of working hours measurement.

The concepts of validity and reliability are more often discussed in psychometrics, and for concept that is not measured directly. Many theoretical concepts may be only measured through several indicator variables, not directly from a manifest variable. An econometric quantity such as working hours might seem like a more clear-cut concept since we in fact have data variable on file. In reality, the survey-based working hours data consists of responses from several questions that are combined according to an operational definition. These responses are subject to human errors, such as random mistakes, memory effects, self-presentation etc. We do not have the actual paper contracts or any physically measurement data for the employee jobs, such as readouts from time clocks. Another matter is the operational definition and treatment of the working hours data.

One alternative way to check working hours quality, could be to link working hours with data on hourly wage and wages-paid. Further studies that link survey- and administrative register data can benefit the quality of both data types, and may reveal errors in both sources.

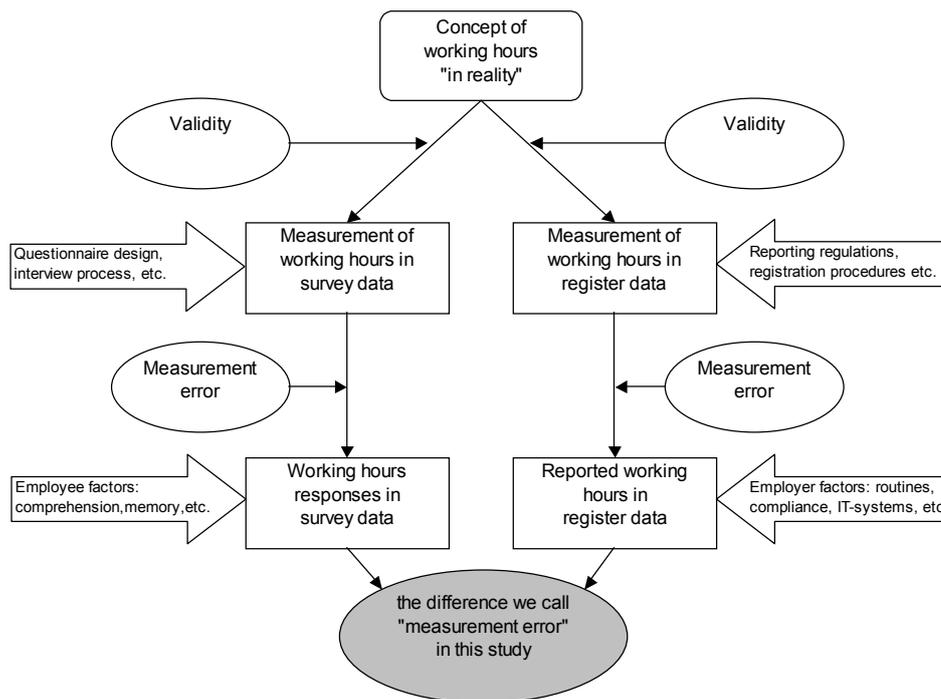
The measurement error can theoretically be defined as the difference between the value that should have been measured and the response value. In this perspective the measurement error is ideally the divergence between a perfect measurement and an actual measurement. The "true value" of the reality concept concerns validity, and not measurement error. Measurement error in this sense has more to do with cognitive processes during an interview, such as memory and response editing, the social setting and so forth. However, the method proposed here doesn't discern between validity as such and measurement errors.

In section 2-5 and 3-1 we pointed out that some objective factors could affect also the quality of register working hours. We suggest that factors such as the definitions, regulations and guidelines of the register affects register data *validity*. The IT-systems and practical routines can in turn affect the register data *reliability*. We propose that also the *measurement errors* the in register data itself should be investigated. The reported and recorded working hours depend on the cognitive abilities and the practical work of several people. That means the registered value of working hours is subjected to random errors such as punching mistakes and accidental omissions. Misunderstandings about the reporting guidelines, short-cuts and other habitual factors can lead to systematic errors.

Diagram 3-1 illustrates how we place our measurement error analysis in a theoretical framework. This framework is partly based on the "survey lifecycle from a quality perspective" from Groves et al. (2004). We have left a number of quality aspects that are not included in our study, but added a parallel "register universe". The left part of the framework is about survey data and the right part is the corresponding entities of the register data source. Register values is considered a kind of measurement in itself, complete with its own errors.

From this framework, it should become clear that what we term "measurement error" in this study is a very indirect estimate of a theoretical measurement error concept. In practical terms it means that components of our results are errors in the evaluation method and not errors in the evaluated data. We propose it as a starting point for examining the quality of working hours data and further development of methods.

Figure 3.1. Theoretical framework



3.3. Calculations

We define:

$$i = \{1, 2, 3, \dots, n\} \quad \text{Units} = \text{the jobs in the sample}$$

$$y_i = W_{i,LFS} \quad \text{Response} = \text{contractual working hours in the LFS survey data}$$

$$Y_i = W_{i,REG} \quad \text{Reported contractual working hours in the employee register data}$$

We define the LFS response as a function of the register value and an error term:

$$y_i = Y_i + \varepsilon_i$$

Then we can define the individual measurement errors as the differences between the register values and the survey measurements:

$$\varepsilon_i = y_i - Y_i \quad \text{Individual measurement error}$$

Taking in all these individual errors ε from all jobs i , we calculate:

$$\eta = E(\varepsilon) \quad \text{Systematic error}$$

$$\zeta = \sqrt{\frac{Var(\varepsilon)}{n}} \quad \text{Random error}$$

Systematic measurement errors can be interpreted as *response bias*: a tendency to answer consistently too high or too low when asked about the working hours. Random measurement errors can be interpreted as *response variance*: answers with mistakes and faults that don't follow any pattern.

The systematic error as we measure it, can be interpreted as the expected error overall. Systematic errors can make the LFS estimate consistently either too high or too low. For instance, if there are systematic over-reporting, with repeated samples the LFS responses will more often be above than below the true values. That means the estimates can be stable and seem certain, but are in fact not accurate. Systematic error has a positive or negative sign indicating over-reporting and under-reporting.

The random error as we measure it, can be interpreted as the typical error per job. Random errors can make a LFS estimate too high or too low, but with repeated samples the estimates will hover around the expected value. If there were only random errors, the risk of a too low estimate equals the risk of a too high estimate. Large random errors make the estimates uncertain, even if the expected mean is correct. Random error figures have only positive values.

In conclusion both random and systematic measurement errors are important in order to evaluate the contractual working hours data quality.

3.4. Analysis

For several reasons, we wish to compare the measurement errors in various groups. The different objectives of study will have consequences for the choice of analytical categories. One important goal is to document the quality for the benefit of users of official working hours statistics. For instance, the quality of working hours for men, women and different age groups will be interesting to social researchers using LFS data. A more long-term goal is to identify problem areas in order to improve quality, such as adjusting the measurement itself, i.e. the survey questions, the interview routines, or improving estimation procedures. To this end, it is for example interesting to evaluate the quality in proxy interviews, or look at effects of different questions.

An overview of some possibilities for comparing contractual working hours data quality:

Unit	Characteristics
Job	Permanent or temporary Full-time / part-time Fixed, varying, or no contractual working hours arrangement Occupation Contractual working hours value
Employee	Gender Age group Level of education
Employer	Type of industry Size and type of company
Survey interview	Proxy interview or direct self-response Number of times participating (wave) Duration Previous interview data, non-response etc.
Register record	Attributes of the employee data reporting Paper form, web form or payroll-system delivery Specific IT-system used Time lag in reporting

3.5. Limitations

In the present study, the variables used are:

- Contractual working hours value
- Fixed, varying, or no contractual working hours arrangement
- Occupation
- Industry and company size
- Gender, age group, level of education
- Proxy interview or direct self-response

Even though we use some statistical parameters such as the mean and variance, we make no assumptions on the distribution function of contractual working hours. Our results show a highly peaked and clearly non-normal distribution.

4. Results

The results are presented with a descriptive summary and tables.

4.1. Summary

- Overall, the survey based contractual working hours are systematically longer, and vary more, than the register based.
- Systematic measurement error, as we have defined it, is generally much larger than random measurement error.
- For jobs with "normal" working hours and the same number every week, the relative measurement errors are quite small.
- For jobs with shorter hours, the survey seem to overestimate working hours, while it seems to underestimate the working hours for jobs with long hours.
- Jobs without an agreement of working hours have by far the largest systematic measurement errors regardless of the number of hours reported.
- Teachers in primary schools have somewhat larger systematic measurement errors than other occupations at this level. The pattern is the same as the general on: the overestimating shorter working hours, and underestimate the longer working hours.
- Teaching professionals have about the same overall systematic measurement errors as other professional occupations with fixed working hours, but a larger underestimation of the longer working hours.
- There is great variation in systematic measurement errors between industries, also when controlling for contractual working hours length and working hours arrangements.
- Proxy interviews seem to cause more overestimation of contractual working hours, especially in jobs with varying contractual working hours and long hours.

4.2. Job and employer characteristics

Contractual working hours and working time arrangements

Table 4.1. Analysis of contractual working hours, by working time arrangement. LFS/register. Average 2004-2007

	Sample size	Mean			Std.dev.		
		Register	Survey	Difference ¹	Register	Survey	Difference ²
Total	153 260	33.29	34.38	1.09	8.06	8.47	6.10
Changing from week to week .	25 089	29.36	30.27	0.91	9.70	10.33	7.22
Same hours every week	118 313	34.25	34.95	0.70	7.07	6.69	4.79
No agreed hours per week	9 858	31.77	37.96	6.19	10.86	15.65	11.91
4 - 37.49 hours/week	55 028	25.46	28.04	2.58	9.05	9.80	7.50
Changing from week to week .	17 149	25.42	26.79	1.37	9.29	9.69	6.99
Same hours every week	34 709	26.09	28.98	2.90	8.53	9.11	6.98
No agreed hours per week	3 170	18.86	24.45	5.59	10.50	14.77	12.69
37.5 hours/week	92 244	37.50	37.87	0.37	0.00	4.61	4.61
Changing from week to week .	7 131	37.50	37.67	0.17	0.00	6.82	6.82
Same hours every week	79 042	37.50	37.40	-0.10	0.00	2.75	2.75
No agreed hours per week	6 071	37.50	44.14	6.64	0.00	11.31	11.31
37.51 - 99 hours/week	5 988	40.43	38.94	-1.49	4.76	6.86	8.43
Changing from week to week .	809	41.17	38.81	-2.36	5.66	9.95	12.41
Same hours every week	4 562	40.13	37.93	-2.20	4.29	3.99	6.02
No agreed hours per week	617	41.71	46.60	4.90	6.26	11.99	13.25

¹ The mean of the differences is equal to the difference of the means.

² This refers to the empirical standard deviation of the differences, which is **not** equal to the difference in the empirical standard deviations. The standard deviation of the differences is used to calculate the random measurement errors. That there are differences in standard deviation is also interesting, because it indicates a different distribution of contractual working hours in the two data sources.

Table 4.2. Measurement errors in contractual working hours, by working time arrangement. LFS/register. Average 2004-2007

	Measurement error Hours		Relative measurement error Per cent	
	Systematic	Random	Systematic	Random
Total	1.09	0.02	3.3	0.05
Changing from week to week	0.91	0.05	3.1	0.16
Same hours every week	0.70	0.01	2.0	0.04
No agreed hours per week	6.19	0.12	19.5	0.38
4 - 37.49 hours/week	2.58	0.03	10.1	0.13
Changing from week to week	1.37	0.05	5.4	0.21
Same hours every week	2.90	0.04	11.1	0.14
No agreed hours per week	5.59	0.23	29.6	1.20
37.5 hours/week	0.37	0.02	1.0	0.04
Changing from week to week	0.17	0.08	0.5	0.22
Same hours every week	-0.10	0.01	-0.3	0.03
No agreed hours per week	6.64	0.15	17.7	0.39
37.51 - 99 hours/week	-1.49	0.11	-3.7	0.27
Changing from week to week	-2.36	0.44	-5.7	1.06
Same hours every week	-2.20	0.09	-5.5	0.22
No agreed hours per week	4.90	0.53	11.7	1.28

Occupation

Table 4.3. Systematic measurement errors in contractual working hours, by working time arrangement and selected occupations¹. LFS/register. Average 2004-2007

	Contractual working hours			
	Total	4 - 37.49	37.5	37.51 - 99
All occupations	1.09	2.58	0.37	-1.49
Changing	0.91	1.37	0.17	-2.36
Fixed number	0.70	2.90	-0.10	-2.20
No settlement	6.19	5.59	6.64	4.90
Teaching professionals	0.67	3.05	0.22	-2.32
Changing				
Fixed number	0.44	3.12	-0.12	-2.35
No settlement				
Other professionals	1.02	2.60	0.51	0.98
Changing	2.14	1.34	1.58	5.80
Fixed number	0.46	2.72	0.02	-0.65
No settlement	7.34	9.34	6.92	7.82
Teachers in primary school	0.88	2.89	0.01	-2.13
Changing				
Fixed number	0.81	2.87	-0.08	-2.18
No settlement				
Other associate professionals	0.88	2.12	0.37	-2.19
Changing	0.56	0.94	0.35	-7.00
Fixed number	0.56	2.45	-0.07	-2.26
No settlement	6.32	6.93	6.47	3.13
All other occupations	1.21	2.66	0.38	-2.22
Changing	0.90	1.48	-0.02	-5.55
Fixed number	0.79	3.05	-0.13	-3.03
No settlement	6.07	5.21	6.74	4.97

¹ The selection is based on teachers as a group with special working hours arrangements. We compare with occupations of the same skill level, and all other occupations.

Industry

Table 4.4. Systematic measurement errors in contractual working hours, by working time arrangements and selected industries¹. LFS/register. Average 2004-2007

	Total	4 - 37.49 hours/week	37.5 hours/week	37.51 – 99 hours/week
All working time arrangements	1.09	2.58	0.37	-1.49
28 Manufacturing of metal products.....	0.45	4.15	0.14	
29 Manufacturing of machinery	0.36	4.00	0.18	
35 Manufacturing of ships and oil rigs ..	0.15	2.96	-0.13	
45 Building and construction	0.73	5.19	0.54	-1.61
61 Sea transportation	3.56	2.06	3.73	
52 Shops and retail sales	2.19	3.38	0.52	0.07
55 Hotels and restaurants	3.25	3.87	1.77	
64 Mail and telecom	-0.30	-0.40	-0.26	
80 Education	0.98	2.98	0.13	-2.09
85 Health- and social services	1.07	1.87	-0.20	-3.44
Other business and services	1.09	2.95	0.55	-0.71
Other manufacturing and supply	0.70	2.59	0.09	-1.03
Fixed number of hours per week	0.70	2.90	-0.10	-2.20
28 Manufacturing of metal products	0.08	4.27	-0.24	
29 Manufacturing of machinery	0.07	4.02	-0.12	
35 Manufacturing of ships and oil rigs...	0.21	4.90	-0.07	
45 Building and construction	0.15	5.01	0.01	-3.27
61 Sea transportation	3.45	2.33	3.72	
52 Shops and retail sales	1.57	3.41	-0.40	-2.21
55 Hotels and restaurants	2.47	4.12	-0.16	
64 Mail and telecom	-0.12	0.06	-0.17	
80 Education	0.87	2.99	-0.06	-2.13
85 Health- and social services	1.27	2.71	-0.22	-3.12
Other business and services	0.47	2.68	-0.09	-1.83
Other manufacturing and supply	0.38	3.00	-0.19	-1.58

¹ The selection consists of five major industry groups with long mean contractual working hours, and five with short hours.

4.3. Employee and interview characteristics

Demographics

Table 4.5. Analysis of contractual working hours, by age and gender. LFS/register. Average 2004-2007

	Sample size	Mean			Std.dev.		
		Register	Survey	Difference	Register	Survey	Difference
Total	153 260	33.29	34.38	1.09	8.06	8.47	6.10
16-19 years	2 042	17.42	18.74	1.32	12.83	13.63	8.92
20-24 years	7 424	28.01	29.93	1.92	12.09	11.99	9.22
25-39 years	47 798	33.96	35.17	1.20	7.40	7.55	6.12
40-54 years	63 383	34.13	35.24	1.11	6.84	7.36	5.53
55-66 years	31 462	33.09	33.83	0.74	7.88	8.67	5.78
67-74 years	1 151	27.12	25.86	-1.26	12.01	12.90	9.45
Men	79 535	35.87	37.04	1.17	5.83	7.02	5.94
16-19 years	1 031	21.80	23.31	1.51	13.71	14.38	8.41
20-24 years	3 948	31.71	33.43	1.72	10.59	10.61	8.06
25-39 years	26 192	36.16	37.34	1.19	5.11	6.11	5.64
40-54 years	31 602	36.66	37.91	1.25	3.97	5.72	5.53
55-66 years	16 130	36.06	36.97	0.91	5.31	7.02	6.12
67-74 years	632	27.90	27.02	-0.88	12.13	13.53	10.54
Women	73 725	30.52	31.52	1.00	9.14	8.95	6.25
16-19 years	1 011	12.96	14.08	1.12	10.07	11.01	9.41
20-24 years	3 476	23.80	25.94	2.15	12.30	12.22	10.37
25-39 years	21 606	31.30	32.53	1.22	8.75	8.26	6.66
40-54 years	31 781	31.61	32.58	0.97	8.06	7.84	5.52
55-66 years	15 332	29.97	30.53	0.56	8.87	9.01	5.40
67-74 years	519	26.17	24.44	-1.73	11.81	11.95	7.89

Table 4.6. Measurement errors in contractual working hours, by age and gender. LFS/register. Average 2004-2007

	Measurement error Hours		Relative measurement error Per cent	
	Systematic	Random	Systematic	Random
Total	1.09	0.02	3.3	0.0
16-19 years	1.32	0.20	7.6	1.1
20-24 years	1.92	0.11	6.9	0.4
25-39 years	1.20	0.03	3.5	0.1
40-54 years	1.11	0.02	3.3	0.1
55-66 years	0.74	0.03	2.2	0.1
67-74 years	-1.26	0.28	-4.7	1.0
Men	1.17	0.02	3.3	0.1
16-19 years	1.51	0.26	6.9	1.2
20-24 years	1.72	0.13	5.4	0.4
25-39 years	1.19	0.03	3.3	0.1
40-54 years	1.25	0.03	3.4	0.1
55-66 years	0.91	0.05	2.5	0.1
67-74 years	-0.88	0.42	-3.1	1.5
Women	1.00	0.02	3.3	0.1
16-19 years	1.12	0.30	8.6	2.3
20-24 years	2.15	0.18	9.0	0.7
25-39 years	1.22	0.05	3.9	0.1
40-54 years	0.97	0.03	3.1	0.1
55-66 years	0.56	0.04	1.9	0.1
67-74 years	-1.73	0.35	-6.6	1.3

Table 4.7. Systematic measurement errors in contractual working hours, by age, gender and working time arrangements. LFS/register. Average 2004-2007

	Total	Working time arrangement		
		Changing	Fixed number	No settlement
Total	1.09	0.91	0.70	6.19
16-24 years	1.79	1.82	1.54	3.12
25-39 years	1.20	1.13	0.85	6.44
40-54 years	1.11	0.77	0.69	7.63
55-74 years	0.67	0.28	0.31	5.06
Men	1.17	1.35	0.42	7.81
16-24 years	1.68	2.42	1.16	4.22
25-39 years	1.19	1.34	0.55	8.20
40-54 years	1.25	1.22	0.41	9.06
55-74 years	0.84	1.00	0.05	6.43
Women	1.00	0.61	1.01	2.66
16-24 years	1.92	1.44	2.14	2.31
25-39 years	1.22	0.95	1.22	2.67
40-54 years	0.97	0.47	0.98	3.78
55-74 years	0.49	-0.12	0.59	1.26

Table 4.8. Systematic measurement errors in contractual working hours, by gender, working hours and working time arrangements. LFS/register. Average 2004-2007

	Total	Men	Women
Total	1.09	1.17	1.00
'0400 - 0799	8.94	10.04	8.50
'0800 - 1499	6.27	7.63	5.74
'1500 - 1799	4.66	5.75	4.42
'1800 - 2799	3.00	5.61	2.54
'2800 - 3749	0.78	1.51	0.43
'3750	0.37	0.74	-0.38
'3751 - 9999	-1.49	-0.35	-3.70
Changing from week to week	0.91	1.35	0.61
'0400 - 0799	7.82	9.33	7.40
'0800 - 1499	4.93	5.75	4.67
'1500 - 1799	3.28	4.69	3.07
'1800 - 2799	1.28	3.87	1.02
'2800 - 3749	-0.31	0.62	-0.86
'3750	0.17	0.82	-2.02
'3751 - 9999	-2.36	1.53	-7.73
Same number every week	0.70	0.42	1.01
'0400 - 0799	9.66	10.31	9.37
'0800 - 1499	6.72	6.98	6.62
'1500 - 1799	5.71	6.40	5.57
'1800 - 2799	3.60	5.26	3.28
'2800 - 3749	1.23	1.71	1.02
'3750	-0.10	0.05	-0.36
'3751 - 9999	-2.20	-1.79	-2.98
No contractual working hours	6.19	7.81	2.66
'0400 - 0799	9.14	10.23	8.54
'0800 - 1499	7.79	11.87	5.48
'1500 - 1799	4.71	5.60	4.08
'1800 - 2799	5.42	10.68	2.34
'2800 - 3749	1.90	5.86	-1.44
'3750	6.64	7.68	1.65
'3751 - 9999	4.90	6.66	-1.85

Proxy interviews

Table 4.9. Systematic measurement errors in contractual working hours, by working time arrangements and proxy or direct interviews. LFS/register. Average 2004-2007

	Total	Direct	Proxy	Proxy minus direct
Total	1.09	1.06	1.32	0.26
Changing from week to week ..	0.91	0.87	1.30	0.44
Fixed number of hours/week	0.70	0.68	0.87	0.19
No settlement of working hours	6.19	6.16	6.46	0.31
4 – 37.49 hours/week	2.58	2.53	2.96	0.43
Changing	1.37	1.33	1.75	0.42
Fixed number	2.90	2.86	3.22	0.36
No settlement	5.59	5.64	5.19	-0.45
37.5 hours/week	0.37	0.35	0.51	0.16
Changing	0.17	0.15	0.40	0.26
Fixed number	-0.10	-0.10	-0.04	0.06
No settlement	6.64	6.56	7.27	0.72
37.51 – 99 hours/week	-1.49	-1.51	-1.32	0.19
Changing	-2.36	-2.68	0.72	3.40
Fixed number	-2.20	-2.17	-2.45	-0.28
No settlement	4.90	4.86	5.24	0.38

References

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

Evans, J. M., D. C. Lippoldt and P. Marianna (2001): "Trends in Working Hours in OECD Countries", OECD Labour Market and Social Policy Occasional Papers, No. 45, OECD

Guilford, J. P. (1954): "Psychometric methods" McGraw-Hill, New York

Klapfer, Karin and Mitterndorfer, B. (2007): "LFS - Quality of Data on Hours Worked. Pilot Study" Bundesanstalt Statistik Österreich, Wien

Lemaitre, Georges (1988): "A look at response errors in the Labour Force Survey" The Canadian Journal of Statistics 1988, Vol.16, 127-141

Thomsen, Ib and Villund, O. (2008): "The effect of proxy interviews on Norwegian Labour Force Survey employment figures", Statistics Norway (to be published)

Tijdens, Kea and A. Dragstra (2007): "How Many Hours Do You Usually Work?" Time & Society VOL. 16 No. 1, Sage Publications, London

Williams, Richard D. (2002): "Hours worked: a comparison of estimates from the Labour Force and New Earnings Survey" Technical report in Labour Market trends, Office for National Statistics, UK

Williams, Richard D. (2004): "Investigating hours worked measurements" Technical report in Labour Market trends, Office for National Statistics, UK

List of Tables

2. Data

2.1. Job identification, by unit, attribute and data source	11
2.2. Difference in mean contractual working hours, by job-level data linking parameters. LFS/Register data. Average 2004-2007. Hours	11

4. Results

4.1. Analysis of contractual working hours, by working time arrangement. LFS/register. Average 2004-2007	17
4.2. Measurement errors in contractual working hours, by working time arrangement. LFS/register. Average 2004-2007	18
4.3. Systematic measurement errors in contractual working hours, by working time arrangement and selected occupations. LFS/register. Average 2004-2007	18
4.4. Systematic measurement errors in contractual working hours, by working time arrangements and selected industries. LFS/register. Average 2004-2007	19
4.5. Analysis of contractual working hours, by age and gender. LFS/register. Average 2004-2007	19
4.6. Measurement errors in contractual working hours, by age and gender. LFS/register. Average 2004-2007	20
4.7. Systematic measurement errors in contractual working hours, by age, gender and working time arrangements. LFS/register. Average 2004-2007	20
4.8. Systematic measurement errors in contractual working hours, by gender, working hours and working time arrangements. LFS/register. Average 2004-2007	21
4.9. Systematic measurement errors in contractual working hours, by working time arrangements and proxy or direct interviews. LFS/register. Average 2004-2007	21