STATISTICAL INFORMATION FROM DIFFERENT SOURCES EXPERIENCES WITH THE DUTCH LABOUR ACCOUNTS

Jan Walschots

Statistics Netherlands

Introduction

Last year the Voorburg Group decided to put employment on the agenda for the 1996 meeting. A first session is intended to tell the Group what we do and don't know about employment. In the report of the closing session of the 1995 meeting different subjects are mentioned that can be touched upon, like skills, education, status or distribution over enterprises. In this paper, that is specially prepared for the 1996 meeting of the Voorburg Group in Newport (South Wales), UK, we will not focus on one or only a few characteristics of employment. Primarily we will discuss the problem of how to get consistent data from all different kinds of sources on different characteristics. We will pose the question how comparable data on employment from different sources are, and how comparability can be improved. The answer comes from the experiences with the Dutch Labour Accounts. After presenting this system and some of its results, we will at the end present some experiences and problems in describing labour market data by educational level of the employee. An extensive description of the development of the Labour Accounts was given by Leunis and Verhage[1].

1. Different sources of information on employment

In the first part we will briefly look at different kinds of sources from which information on employment may be got. Leunis and Verhage distinguish three methods of observation in the Labour Statistics, each having their strong points and their restrictions. These are establishment surveys, personal or household surveys and registrations. The differences, which are summed up in figure 1, may be illustrated by comparing three sources with data on employment.

1. The Annual Survey on Employment and Earnings (ASEE). There are some 71,000 establishments, reporting on their number of employees, and also on the wages of 960,000 individual employees. This annual survey measures the number of jobs, and also for a number of employees individual data on personal, wage and job characteristics. These include sex, age, type of labour contract, economic activity, seniority, hours worked, company size, place of work. Before the ASEE started in 1991 two separate surveys were conducted: the Statistics of Working Persons and the Annual Earnings Survey.
2. The Labour Force Survey (LFS). A household survey conducted on some 110,000 addresses yearly. The most important results are on the labour force. In the Dutch LFS the labour force is from 1992 onwards defined as all persons having a paid job, or actively seeking for a paid job of at least 12 hours per week. Within the labour force are distinguished employees, self-employed and family-workers, and unemployed. Among other variables we measure for every employee characteristics as sex, age, marital status, educational level, profession, hours worked, place of residence and place of work, economic activity.
3. The Statistics on Paid days and Total Wages (SPTW). This statistic is based on a register with data on paid days of employees in social insurance (Sickness, disablement and unemployment benefits).

[1] Leunis, W.P. and C.G. Verhage: Arbeidsrekeningen, kern van het statistisch systeem over arbeid, 1996 (Labour accounts, Core of the statistical system on labour).

In the ASEE sampling errors on the numbers of jobs are smaller than in the LFS: one company reporting its 2367 jobs gives a more precise result than when, say, 25 of its employees are in the sample of the LFS and their response is grossed up to give an estimate of the total. The register used for the SPTW gives even smaller sampling errors, because data on every small enterprise are included, whereas only a sample of the small firms is included in the ASEE. On the other hand the SPTW does not cover all employment, but only paid days of employees, insured for sickness benefits. Also SPTW does not contain much detail on the employment it measures, since that is irrelevant for the purpose of the registration. The ASEE and the LFS have much more detailed information.

Figure 1 Strong points and restrictions of different kinds of surveys

Strong points	Restrictions
Establishment Surveys	
Limited sampling errors because of concentration of employment in large companies that are fully covered.	Measurement error in one unit can have strong influence on results.
Accurate data because questions agree with company administration.	Limitations to questions because of agreement with company administration.
Co-ordination among statistics is possible; characteristics of units are defined in the company register.	Delays in actualising of the company register lead to sampling- and non-sampling errors.
Household Surveys	
Information on an individual level. The population of persons performing paid labour is fully observed.	Inaccurate observation of formal/administrative characteristics. Relatively large sampling errors.
Observation from Registrations	
Integral observation: no sampling errors.	Differences between actual and desired populations.
Accurate data.	Not every registration is updated regularly.
Data are available at low costs.	Definitions are set by the aim of the register. The number of variables is often limited.
Use of registrations limits the statistical burden for companies.	Changes in regulations can lead to changes in available information.

We also note that a variable that must be registered in the companies administration for tax and other purposes, like a persons wage level, can be measured far more accurately from establishment surveys than from a household survey. On the other hand variables that need not be recorded for formal/administrative reasons in the company administration, like

the educational level of the employee and even his profession, are better measured in the

LFS.

In the Business Register of Statistics Netherlands all firms are included and have a coding for their economic activity, and a size class indicator. Among other uses this register is used as a sampling frame for establishment surveys. This means that in all establishment surveys the codes for economic activity and size-class are equal, and also that they need not be asked in every survey. In a household survey the coding of the economic activity of the company where a respondent is working, or its size, is more difficult.

2. Conflicting data on employment

Different sources give different information. We will give some examples. For the years 1987 to 1992 the three statistical sources mentioned published the following results on employment:

Table 1 Employment statistics from three sources

	number of jobs (ASEE)	employees (LFS) [1)]	typical workers (SPTW) [2)]
	* 1000		
1987	4764	5166	3503
1988	4888	5333	3578
1989	5009	5454	3774
1990	5180	5626	3950
1991	5395	5799	4046
1992	5471	5259	
1993		5261	

1) In 1992 break in series as a result of change in definition

2) The number of typical workers is the number of paid days divided by 260

These results have two important defects. Firstly different sources give different results on the same variable. Secondly logical relations between variables do not appear in the statistical results. Not only the absolute levels of the results of these statistics differ, also the developments of employment . Part of the differences are explained by the units of measurement: jobs, persons employed, and typical workers. However, also relations between variables that should hold in reality do not hold in statistical results. An example is the relation between jobs and persons employed . If a person has two (or more) jobs, he is counted as one employed person in the LFS, but in the ASEE jobs are counted, so this person counts twice (or more). In theory the number of employees (5799 thousand in 1991) plus the number of second jobs (which was 208 thousand) should add up to the total number of jobs. This sum (6007 thousand) is clearly higher than the number measured by ASEE (5395 thousand). After a change in definition in the LFS in 1992, the difference seems to be far less important. The results seem to indicate that 5259 thousand persons have 5471 thousand jobs. In section 6 we will see that this is not really the case.

Another example of a discrepancy between different statistics is between the number of

jobs, the average annual earnings, and the wage sum. These are the results for 1991.

Table 2 Jobs, earnings and wage sum from different surveys, 1991

	Unit	result 1991
Number of jobs of employees	*1000	5395
average annual earnings	*1000 Dfl	42.0
Number of jobs * average annual earnings	bln Dfl	226.6

Wage sum	bln Dfl	212.8
Discrepancy	bln Dfl	13.8

One would expect that the wage sum, which is measured in a quarterly establishment survey, agrees with the product of the number of employees and their average annual earnings, also measured from establishment surveys. There appears to be a discrepancy of some 14 billion guilders, more than 6%. In order to deal with these problems methods of statistical integration were developed, resulting in the Labour Accounts.

3. The aims and general principles of Statistical Integration

Three main aims are to be met by Statistical Integration:

- Completeness of the description for the whole population;
- Consistency of definitions of all variables;
- Comparability of data over time.

In this paper we will not deal with the third aim, comparability over time.
The integrated data set has to be a description concerning **all elements of the subject population**. In many cases the original statistical inquiries will be set up for only part of the total population of units. A range of practical and institutional factors may cause that. For example for the smaller jobs there may be no legal obligation to the employer to pay social security contributions. Then in business administration there may be no individual records for the persons involved, and they may be not reported in a statistical inquiry about employees. But in the integrated overview they still have to be present. The integrated statistic will be based on several basic statistical surveys. That may give a first sight on differences in population and a tool to make estimations about specific groups within the total population. But even then one must be aware of the risk of missing units, and permanently look for other sources that may give information about missing elements in order to arrive in the end at an integrated statistic covering all units of the population.
Behind the integrated statistic there will be a **consistent set of definitions** covering all variables in the field. That means that variables are linked to each other by definitional equations. These equations show which variables have to add up to arrive at another variable and how the total population is subdivided into different groups of units. Although in basic statistics one may use a definition for a variable close to the perception of the unit filling in the questionnaires, for the integrated system all definitions for variables have to fit in the theoretic system.
For the process of construction of the integrated statistical data there are two general principles:

- there must be a clear relation between the sources and the integrated system;
- at different stages, there may be different concepts.

In the statistical process of integration the source data are transformed into integrated data. Data according to the definitions in the sources which are close to the perception of the units filling in the questionnaires, are transformed to data according to the definitions of the system. There are several kinds of definitional differences. Some of them are related to the transition to the total of the population. Others are related to the content of the variables. In the system there may be a wider or smaller concept than was used in the original statistical inquiry or in the registers used. For each step in the process of statistical integration, data have to be made and a full bookkeeping of all additions and subtractions has to be kept. By that there will be in the end a clear relation between the original data and the data of the integrated statistic. This is not only important in order to be able to follow the same steps for the next report. It also is a matter of scientific behaviour and may be important for the open relations between the staff working on basic statistics and those who make the integrated overview.
For various variables there may be at least three different concepts or definitions:

- definitions used in questionnaires;
- definitions used in the process of integration;
- definitions for publications of the integrated system.

As set out before the definitions used in the questionnaires have to be close to the general perceptions of the respondents. In case of units with a kind of bookkeeping, the questionnaire can be based on the bookkeeping definitions. In other cases some experimental questioning may make clear what respondents have in mind at some definition or expression. Staying close to the respondents language is the best guarantee to get valid information. However, it is also a cause for the problem of different concepts in different statistical sources.

When integrating information it will not always be efficient to go directly to the definitions of the system. The distance between the definitions of the system and the concepts in basic statistics may be rather large. Then it is more efficient to compare the basic statistics in first instance using a concept that is as close as possible to the concepts of both inquiries. Once the basic data are reconciled on such a concept the transition to the system definitions is made. In the end the publication will usually be in terms of the definitions of the integrated system. But for some users it may be helpful to publish also basic information on other definitions. The bookkeeping of all steps of the process will then be an important source to explain differences in concepts.

4. The process of Statistical Integration

The strength of the accounts approach is essentially derived from the redundancy in the data sources. This makes it possible to compare measurements of the same variable or of related variables from different sources and thus to discover measurement errors and correct for them. First, however, corrections have to be made for differences in definitions and populations. To correct for differences in *definitions*, a set of consistent, multi-purpose definitions is selected for use in the accounts framework. These are compared with the definitions in the sources and estimates are made of the numerical size of the definitional difference. This is generally done at some meso-level of aggregation, sometimes at the micro level. Thus, estimates of the size of definitional differences between wages from an earnings survey and tax records may be made for groups of workers distinguished by age, sex and industry. Naturally, the level of aggregation at which the definitional corrections are made, depends on the properties of the source and therefore differs from source to source. This process of eliminating definitional differences is called **harmonisation**. It yields tables that show at some level of aggregation, firstly, the value of a variable according to the source definition, secondly the value of the correction that has to be made and thirdly, the value of the variable according to the accounts framework definition.

Correcting for differences in population is called **completion**. As a rule, the accounts framework provides the most comprehensive coverage of the population that can be achieved. Consequently, the process of completion amounts to making estimates of the value of the variable concerned - according to the definition of the accounts framework - for the part of the population that is not covered in the original data source. This too is done at a certain level of aggregation, again depending on the properties of the source. The process of completion, piled on top of the process of harmonisation, leads to tables at a certain level of aggregation that show, firstly, the value of a variable according to the accounts framework definition, for the population covered by the source; secondly, the estimate for the population not covered by the source; and thirdly the estimate for the total population covered by the accounts framework.

The third major step in the compilation of the accounts is the **minimisation of measurement errors**. The first two steps lead to estimates from different sources according to harmonised definitions and the complete population. Therefore, all remaining differences must be due to measurement errors. These are analysed and as far as possible attributed to specific error sources. The systematic errors that are discovered in the sources are corrected and a correction is made for the random errors, smoothing them out by e.g. pooling data for several years after correction for trends.

This third step in the compilation process results in tables that show for each variable for each source which corrections have to be made to eliminate the measurement errors. After this third step, a data system is obtained where all variables are measured according to uniform definitions for the complete population and where all measurement errors that could be found are eliminated. Generally speaking, at this stage only one estimate is left for each variable, but the identities and logical relations between variables need not yet be completely satisfied. To achieve this, a **balancing** procedure is used that, by some formal algorithm, eliminates the discrepancies in the macro relations. These balancing corrections may be viewed as estimates of the remaining measurement errors for which no specific causes have been found. Clearly, the smaller these balancing corrections, the higher the quality of the first three steps in the compilation process.

The total compilation process is usually referred to as ‘integration’. The integration process and the relation between the source data and the macro data are summarised in figure 2.

Figure 2 Integration process: from sources to accounts estimates

	Value of variable according to source, source definition, source population
	plus
Harmonisation	**Definitional corrections**
	gives
	Value of variable according to source, accounts definition, source population
	plus

Completion	**Estimate of variable for remaining population**
	gives
	Value of variable according to source, accounts definition, complete population
	plus
Minimisation of measurement errors and balancing	**Correction for measurement errors and balancing**
	gives
	Accounts estimate of variable

5. The Labour Accounts

The Labour Accounts, in which data on the labour market are integrated, were started from 1987 with data on jobs of employees. Later the system was extended with data on self-employed, and on social security benefits. We will restrict ourselves here to the results concerning the jobs of employees. In figure 3 we give an impression of which variables are in the Labour Accounts, and on the relations that exist between these variables.

Figure 3 Relations between variables in the Labour Accounts

		Persons employed		Labour years
		*		*
		Number of jobs per person		Contractual working time per full-time job
		=		=
Agreed hours per job	*	Number of jobs	=	Labour volume in hours (contractual)
+				+
Overtime hours per job	*	Number of jobs	=	Labour volume overtime
=				=
Paid hours per job	*	Number of jobs	=	labour volume in paid hours
*				*
Wage rate per hour (including overtime payments)				Wage rate per hour (including overtime payments)
=				=
Earnings per year excluding bonuses and gratuities	*	Number of jobs	=	Wage sum excluding bonuses and gratuities
+				+
Bonuses and gratuities per job	*	Number of jobs	=	Bonuses and gratuities (sum)
=				=
Earnings per year including bonuses and gratuities	*	Number of jobs	=	Wage sum including bonuses and gratuities
				+
				Social security contributions
				=
				Wage costs sum

All these relations hold in the real world. One of the aims of the integration process is to make them hold in statistical output, and in the results of the Labour Accounts they all hold. It is important to note that in the real world the relations do not only hold for the total economy but also for each group one can imagine, e.g. for males, females, for persons aged 37, or by economic activity. All variables from figure 3 are now computed and published detailed by sex, hours worked (more than 20 hours per week, or less), type of labour contract (full time, part-time or flexible working time job) and economic activity (71 groups)[2]. We call this the core of the system. In the further development of the Labour

Accounts one of the objectives is to expand the system with ever more detail. For the future the core of the Labour Accounts may be extended with age and/or regional information.

[2] With some exceptions these groups correspond to 2-digits groups of NACE. Some NACE-groups with few employed persons were however combined, but some large 2-digits groups could be split to give more detailed results. Other groups were taken separately for pragmatic reasons. Examples are subsidised workers in social workplaces, that are excluded from the labour volume in the National Accounts, and employees at employment agencies, for which other statistical sources are needed.

Since the integration process is performed at some level of aggregation and not at the micro level, not all data from all sources are fully integrated. The core of the system is fully consistent. Other variables, especially variables that are observed only from one source are linked to the system.

When a characteristic is not in the core of the system there are several ways in which the available statistical information can be upgraded, making use of the Labour Accounts. Sometimes it is possible to compute the distribution by some characteristic of part of the variables in figure 3. This applies e.g. the educational level of the employed. Employment by educational level is measured by the LFS. Therewith the distribution by the educational level of persons employed, jobs and labour years in the Labour Accounts can be computed. There are, however, hardly any data on earnings by educational level. Therefore it was not yet possible to enter educational level of employees in the Labour Accounts for the financial data, like earnings, wage sum or labour costs. For the latter only estimates have been made at a very low level of detail.

In other cases, when it is only possible to compute the distribution of some variable for a part of the population, then the results for that part of the population are accompanied by a specification of the difference between that part and the total population.

6. Results

From section 5 we can conclude that publications on the level and characteristics of employment can be made in three different ways:

1. Characteristics that are fully integrated in the Labour Accounts, according to accounts definitions and accounts populations.
2. Characteristics that are computed for the total Labour Accounts population, but not for all variables in the system.
3. Results according to source definitions and/or source populations. These results are at some aggregate level accompanied by a specification of differences with accounts definitions and populations.

We will give examples of all three of them. First in figure 4 we present the major results for 1994 of the Labour Accounts for employees. A large number of statistical sources have been used to build these accounts. These include the Labour Force Survey, an Annual Survey on Employment and Earnings, Quarterly surveys on employment and the wage sum development, and a Quarterly survey on earnings, Statistics on Paid days and Total Wages (based on a registration of employed persons, insured for sickness-benefits), and also more specific statistics on employment agencies, workers in social workplaces etc. One can no longer say that one variable comes from one survey. After the integration process all results match with each other. Figure 4 deals with the total number of employees. The most important results by sector of economic activity are presented in table 3.

Figure 4 Results for 1994, total population of employees

Persons employed	Labour years
5,778,000	4,786,000

		*		*
		Number of jobs per person 1.032		Contractual working time per full-time job 1740
		=		=
Contractual hours per job 1397	*	Number of jobs 5,963,000	=	Labour volume in hours (contractual) 8,329,000,000
+				+
Overtime hours per job 24	*	Number of jobs 5,963,000	=	Labour volume overtime 143,000,000
=				=
Paid hours per job 1421	*	Number of jobs 5,963,000	=	labour volume in paid hours 8,472,000,000
*				*
Wage rate per hour (including overtime payments) 27.76				Wage rate per hour (including overtime payments) 27.76
=				=
Earnings per year excluding bonuses and gratuities 39,400	*	Number of jobs 5,963,000	=	Wage sum excluding bonuses and gratuities 235,170,000,000
+				+
Bonuses and gratuities per job 4400	*	Number of jobs 5,963,000	=	Bonuses and gratuities (sum) 26,290,000,000
=				=
Earnings per year including bonuses and gratuities 43,800	*	Number of jobs 5,963,000	=	Wage sum including bonuses and gratuities 261,460,000,000
				+
				Social security contributions 44,080,000,000
				=
				Wage costs sum 308,540,000,000

Table 3 Major results of the Labour Accounts 1994 by sector of economic activity

	NACE-section	A,B	C	D	E	F	G	H
Number of jobs	* 1000	114	10	1033	45	366	952	175
Contractual hours per job		1325	1706	1533	1651	1611	1390	1055
Overtime hours per job		17	47	33	26	30	19	12
Paid hours per job		1342	1753	1566	1677	1641	1409	1067
Labour volume in hours (contractual)	* 1,000,000	151	17	1584	74	590	1323	185
Labour volume overtime	* 1,000,000	2	0	34	1	11	18	2
Labour volume in hours paid	* 1,000,000	153	17	1619	75	601	1341	187
Labour years	* 1000	86	10	918	43	346	757	106
Wage rate per hour	Dfl	21,11	41,39	28,49	34,85	29,01	23,82	20,11
Earnings per year excluding bonuses and gratuities	* 1000	28,3	72,6	44,6	58,5	47,6	33,6	21,5
Bonuses and gratuities	* 1000	1,9	12,7	5,5	6,7	4,3	3,9	1,6
Earnings per year including bonuses and gratuities	* 1000	30,2	85,2	50,1	65,1	51,9	37,4	23,1
Wage sum including bonuses and gratuities	* 1,000,000,000	3,4	0,8	51,8	2,9	19,0	35,6	4,0
Wage costs sum	* 1,000,000,000	3,9	1,1	62,3	3,5	22,2	41,8	4,7

	NACE-section[1)]	J	K	L	M	N	O	P
Number of jobs	* 1000	209	716	484	389	815	228	40
Contractual hours per job		1582	1385	1586	1202	1142	1252	632
Overtime hours per job		18	17	9	1	5	11	0
Paid hours per job		1600	1402	1594	1203	1148	1263	632
Labour volume in hours (contractual)	* 1,000,000	330	992	768	467	931	285	25
Labour volume overtime	* 1,000,000	4	12	4	0	4	3	0
Labour volume in hours paid	* 1,000,000	334	1004	772	467	935	288	25
Labour years	* 1000	188	564	433	284	538	162	13
Wage rate per hour	Dfl	32,53	27,47	30,20	36,28	26,11	24,72	14,95
Earnings per year excluding bonuses and gratuities	* 1000	52,1	38,5	48,1	43,7	30,0	31,2	9,4
Bonuses and gratuities	* 1000	10,1	4,8	4,5	4,1	2,8	3,2	0,3
Earnings per year including bonuses and gratuities	* 1000	62,1	43,4	52,6	47,7	32,7	34,4	9,8
Wage sum including bonuses and gratuities	* 1,000,000,000	13,0	31,0	25,5	18,5	26,7	7,8	0,4
Wage costs sum	* 1,000,000,000	16,1	36,1	29,2	21,6	31,6	9,5	0,4

1)The names of the NACE-sections are entered in appendix 1

To establish the value of these results it is best to compare some results on employment from the Labour Accounts with results from two major sources, the LFS and ASEE. In table 4 we give some results on employment of employees that are available in all three publications. Looking first at the total results of the LFS and ASEE the number of employees according to the LFS (5222 thousands) seems to fit quite well with the number of jobs with known job characteristics according to ASEE (5495 thousands). This seems to be even more true when realising that according to the LFS 226 thousand persons have a second job. However, when looking to the results in a Labour accounting framework the differences appear to be much more complicated.

For almost half a million jobs no characteristics are measured in the ASEE. Also the differences are not spread evenly over the characteristics. The differences are greatest among small jobs and women and differ per economic activity.

Table 4 Employees by sex, hours worked, type of labour contract and economic activity, from three publications, 1994

	Labour Accounts	LFS	ASEE	percentage of Labour Accounts	
				LFS	ASEE
	* 1000			%	
Total	5963	5222	5495	88%	92%
Male	3555	3260	3313	92%	93%
Female	2408	1963	2182	82%	91%
>20 hours	4743	4874	4579	103%	97%
<20 hours	1221	348	904	29%	74%
Full-time	3698		3535		96%
Part-time	1680		1514		90%
Flexible	585	425	446	73%	76%
NACE rev 2					
A/B	114	89	91	78%	80%
C	10	10	9	100%	90%
D	1033	963	922	93%	89%
E	45	47	43	104%	96%
F	366	348	338	95%	92%
Service sector					
G	952	761	915	80%	96%
H	175	128	171	73%	98%
I	388	367	370	95%	95%
J	209	206	200	99%	96%
K	716	453	690	63%	96%
L	484	526	430	109%	89%
M	389	377	377	97%	97%
N	815	709	746	87%	92%
O	228	165	192	72%	84%
P	40	6	0	15%	0%

When looking at the Service sector it is clear that in several sections there are rather large differences between the statistics. For the LFS this holds for almost all sections, for G,H,K,N and O in particular. In the ASEE there are also differences for all sections, for L and O in particular.

In section 5 we stated that a publication from the LFS or from the ASEE is always accompanied by a specification of the sources of difference from the Labour Accounts results. In table 5 we give a full comparison of the results from the three sources. In this table we start from the published number of employed persons in the labour force, specify the differences with the number of jobs of employees in the Labour Accounts and end with the differences in the published number of jobs in the ASEE. The major differences between the LFS-results and the Labour Accounts come from differences in definitions used. These differences are eliminated in the harmonisation and completion process. Measurement errors and balancing bridge another gap of 139 thousand jobs. Data on frontier workers are used to go from jobs according to a national concept, used in the LFS, to jobs according to a domestic concept, used in establishment surveys. Now all differences between the Labour Accounts and the LFS have been accounted for. The lower part of the table gives the differences between Labour Accounts and the number of jobs, that the ASEE publishes. It appears that 381 thousand jobs are not registered in company administrations, from which the ASEE gets its data. Part of these are recognised as special groups, for which the sickness benefit insurance does not apply (insurance for sickness benefit is one of the operational criteria in defining employees in establishment surveys). Finally there are three more causes for differences in the published results. The first is the difference between a yearly average, as measured in the

Labour Accounts, and ultimo September results as published in ASEE. Secondly there are some groups in ASEE for which no individual characteristics are available. Thirdly the published results of ASEE were based not solely on the direct observations of that statistic, but also on feedback from a preliminary version of the Labour Accounts for 1994. Later more information for the Labour Accounts 1994 became available which lead to slight differences in a second version. It is not very practical to revise the complete ASEE publication, but the difference must be taken into account.

The differences between Labour Accounts results and the results from the source statistics are of course partly a result of national statistical practices, national regulations and national habits of companies administrations. Therefore in other countries other differences may be found. The main point is, however, that under all circumstances one must be aware of the many differences between statistical sources, and the problems of comparability of the results. We think that the Labour Accounts approach provides us with a very useful methodology for getting complete and consistent information about employment.

Development of the employment in the Service sector.

Employment in the Service sector has risen quite substantially in the last decade. In table 6 we present the development of the number of jobs of employees since 1987. The total number of jobs rose by more than 800 thousand (+15%). Most of this rise was realised in the Service sector. Actually the number of jobs outside the Service sector declined yearly since 1992. The Service sector showed a rise in jobs of 22%. Also within the Service sector there are large differences. There was a decline in the number of jobs in section L (Public administration and defence and compulsory social security), and practically a standstill in section M (Education). The largest growth was realised in sections H (Hotels and restaurants: +91%) and K (Real estate, renting and business activities: +49%).

Table 5 Specification of differences between published numbers on employees from LFS, ASEE and Labour Accounts, 1994

	*1000
Employees in labour force according to LFS	
LFS definition: working more than 12 hours, aged 15-64, main jobs	**5222**
Harmonisation	
+ employees working less than 12 hours, aged 15-64, main jobs	646
+ secondary jobs, persons aged 15-64	226
Completion	
+ employees, aged over 65 years	16
Minimisation of measurement errors and balancing	
reporting error paid labour	-120
observation errors second jobs	-17
correction for sampling errors	-2
Total number of jobs based on LFS	**5971**
Frontier workers (Inhabitants of Germany and Belgium working in the Netherlands)	16
Frontier workers (living in the Netherlands working in Belgium or Germany)	-24
Jobs in the Netherlands (Labour accounts result)	**5963**
Jobs not observed in company administrations	381
of which specific groups	
personnel in private households (NACE P)	40
newspaper boys (NACE D)	58
home care personnel (working semi free-lance) (NACE N)	41
part of the clergy (NACE O)	7
outworkers	8
others	228
Jobs observed in company administrations (1994 average)	5582
Difference between 1994 average and September 30 observation	46
Jobs measured by September 30, 1994 by ASEE	5628
Of which jobs with no data available on individual characteristics	
Military personnel (conscripts)	23
Manager-proprietor of LLC	97
Difference between preliminary versions ASEE and Labour accounts	11
Population for job characteristics from ASEE	**5495**

Table 6 Jobs of employees by economic activity 1987-1995

Total	Non-service sectors						Service sector	
	Total	NACE-section					Total	
	A-F	A,B	C	D	E	F	G-Q	G

	x1000								
1987	5271	1555	90	10	1049	48	359	3716	767
1988	5390	1576	95	10	1057	48	366	3814	794
1989	5536	1625	102	10	1086	48	379	3910	826
1990	5714	1650	99	11	1110	46	384	4064	867
1991	5843	1657	103	11	1112	46	385	4186	909
1992	5928	1643	106	11	1101	46	379	4285	944
1993	5934	1608	111	10	1074	45	368	4325	952
1994	5963	1568	114	10	1033	45	366	4395	952
1995	6085	1557	114	10	1024	43	366	4528	983
	% change of jobs								
1995-1987	15%	0%	27%	2%	-2%	-9%	2%	22%	28%

	Service sector (continued)								
	H	I	J	K	L	M	N	O	P
	x1000								
1987	98	348	194	526	508	379	675	192	28
1988	105	351	198	560	507	382	686	197	32
1989	117	364	202	583	506	387	699	190	37
1990	128	374	202	626	503	392	728	209	36
1991	135	383	206	655	499	395	751	218	35
1992	145	394	209	672	491	393	781	222	35
1993	158	394	210	676	495	396	796	212	35
1994	175	388	209	716	484	389	815	228	40
1995	187	391	210	783	470	392	828	244	37
	% change of jobs								
1995-1987	91%	12%	8%	49%	-7%	3%	23%	27%	31%

The Labour Accounts do not only measure employment in numbers of jobs. Also the numbers of paid hours, which may be more relevant from a productivity point of view, are computed. These results are presented in table 7. The number of paid hours rose from 1987 to 1995 by 11%. Outside the Service sector there was a 3% decline in hours; the Service sector rose by 17%. The difference between the growth rates of jobs and hours are a result of the decline in average hours worked per job. This decline is both a result of a shorter working week for full timers and of the creation of small new jobs. This decline is particularly present in the Hotels and restaurants; a rise in the number of jobs of 91% only lead to a rise in hours of 58%. On the other hand the paid hours per job remained unaltered in section K. Therefore the rise in the number of jobs of 49% combines with a rise in paid hours of 50%.

Table 7 Paid hours worked by employees by economic activity 1987-1995

	Total	Non-service sectors						Service sector	
		Total	NACE-section					Total	
		A-F	A,B	C	D	E	F	G-Q	G
	* mln hours								
1987	7803	2520	131	16	1691	81	601	5283	1138

1988	7935	2542	138	17	1694	80	613	5394	1171
1989	8103	2585	145	17	1722	79	621	5518	1216
1990	8322	2634	145	17	1758	79	635	5688	1264
1991	8446	2627	148	18	1756	78	627	5819	1317
1992	8523	2608	150	18	1741	77	622	5915	1344
1993	8478	2538	150	18	1686	77	608	5940	1342
1994	8472	2465	153	17	1619	75	601	6007	1341
1995	8648	2447	155	17	1601	73	602	6201	1385
	% change of paid hours								
1995-1987	11%	-3%	19%	2%	-5%	-10%	0%	17%	22 %

	Service sector (continued)								
	H	I	J	K	L	M	N	O	P
	* mln hours								
1987	127	591	315	741	824	456	829	240	24
1988	134	601	321	780	820	457	837	248	24
1989	143	615	325	828	812	457	845	252	26
1990	152	634	330	888	806	463	867	259	25
1991	157	650	335	926	795	466	882	267	23
1992	164	667	332	944	783	470	910	278	22
1993	172	664	336	945	787	469	924	278	24
1994	187	654	334	1004	772	467	935	288	25
1995	200	662	337	1110	750	470	959	305	19
	% change of paid hours								
1995-1987	58%	12%	7%	50%	-9%	3%	16%	27%	-20%

7. The educational level of the employees

The educational level of a person is not registered very systematically in company records. It is therefore very difficult to measure the distribution by educational level of employees through establishment surveys. The LFS is in The Netherlands the most important source for measuring it. The educational level is defined as the level of the highest educational course, taking at least six months, that a person has successfully completed. If he has succeeded several courses at the same level, the last one followed is chosen.

As we saw from tables 4 and 5 there is a considerable difference between the populations in the publications of the LFS and of the ASEE. When comparing characteristics of employees from these two statistics, one should always keep these population differences in mind. It was not yet possible to give reliable estimates of the distribution of the educational level in the ASEE. We could however estimate the distribution of educational level for the total Labour Accounts population. In table 6 we compare the Labour Accounts results with the originally published LFS-distribution of educational level of the employed labour force.

Table 6 Educational level of employees according to LFS and Labour Accounts

Educational level	Employees according to LFS		Employees according to Labour accounts		Difference	
	*1000					
total	5222	100%	5963	100%	741	100%
basic (ISCED 0,1)	449	8,6%	567	9,5%	118	15,9%
junior secondary (ISCED 2)	376	7,2%	520	8,7%	144	19,4%
junior vocational (ISCED 2)	826	15,8%	947	15,9%	121	16,3%
senior secondary (ISCED 3)	278	5,3%	381	6,4%	103	13,9%
senior vocational (ISCED 3)	1988	38,1%	2142	35,9%	154	20,8%
vocational colleges (ISCED 5 or 6)	909	17,4%	988	16,6%	79	10,7%
university (ISCED 6 or 7)	387	7,4%	409	6,9%	22	3,0%
unknown	8	0,2%	9	0,2%	1	0,1%

It is clear that the sum of all effects of the integration process is distributed quite differently from the educational distribution in the LFS. We will not try to explain all the differences, but a major reason is that many students, after finishing junior or senior secondary school, have a part-time job of less than 12 hours and are therefore not included in the labour force. By the time they finish their education and start to work full-time they reach a higher educational level or finish a vocational education. Senior vocational or higher education is far more often someone's final education, after which a full-time job is taken.

Measuring wage levels by educational level

Up till now in the Netherlands there has been no regular statistic measuring wages by educational level. The last time a large scale survey was conducted, was in 1979. After that only in 1985 and 1989 small scale surveys were conducted. In 1985 and 1989 only part of the responding companies of the ASEE were asked to report on the educational level of a total of 5000 and 21000 employees, respectively. One of the problems with these results is that the reported distribution of the educational level differs substantially from the distribution, measured by the LFS. In table 7 is shown the distribution by educational level of the labour volume of employees. We compare the Labour Accounts results based on the LFS with the results of the ASEE, both for 1989.

Table 7 Labour volume (paid hours) of employees by educational level, 1989

Educational level	Labour volume of employees		
	Labour accounts, based on LFS, 1989	Earnings Survey 1989, ASEE	difference
basic (ISCED 0,1)	11.7%	9.9%	1.8%
junior secondary (ISCED 2)	5.9%	10.9%	-5.0%

junior vocational (ISCED 2)	18.9%	25.6%	-6.7%
senior secondary (ISCED 3)	4.8%	6.5%	-1.7%
senior vocational (ISCED 3)	37.8%	28.6%	9.2%
vocational colleges (ISCED 5 or 6)	15.6%	13.4%	2.2%
university (ISCED 6 or 7)	5.4%	5.1%	0.3%
total	100.0%	100.0%	

Comparable differences were found in the 1979 and 1985 surveys. Even though the populations from the Labour Accounts and the ASEE are different we may read from table 7 two apparent differences in the reported level of education:

- General education (junior and senior secondary, not vocational) is reported as the final education much less frequently in the LFS than in the ASEE
- On average the reported educational level is considerably higher in the LFS than in the ASEE.

Both results point at under reporting of educational level in the ASEE. A number of courses that are reported in the LFS, and that may have been taken after joining the company, are not registered in the business administration. It is not very likely that the differences are caused by the population differences. Neither is it very likely that the difference is caused by a systematic over reporting in the LFS.

For the future Statistics Netherlands expects to be able to get regular results on wages by educational level by integrating information from the LFS and ASEE on micro-data level. The ASEE is being expanded by enlarging the sample of employees, for whom individual wage data are collected. By using electronic data interchange with companies or with agencies that take care of the salary administration of many small firms it is possible to have data on the wages of all instead of a sample of the employees, without raising the response burden of the companies. A nearly integral set of wage data on employees can be matched with the LFS to add the wage level to the employees from the LFS.

APPENDIX

SECTIONS ACCORDING TO NACE, REV 2

A AGRICULTURE, HUNTING AND FORESTRY
B FISHING
C MINING AND QUARRYING
D MANUFACTURING
E ELECTRICITY, GAS AND WATER SUPPLY
F CONSTRUCTION
G WHOLESALE AND RETAIL TRADE; REPAIR OF MOTOR VEHICLES, MOTORCYCLES
H HOTELS AND RESTAURANTS
I TRANSPORT, STORAGE AND COMMUNICATION
J FINANCIAL INTERMEDIATION
K REAL ESTATE, RENTING AND BUSINESS ACTIVITIES

L PUBLIC ADMINISTRATION AND DEFENCE; COMPULSORY SOCIAL SECURITY
M EDUCATION
N HEALTH AND SOCIAL WORK
O OTHER COMMUNITY, SOCIAL AND PERSONAL SERVICES ACTIVITIES
P PRIVATE HOUSEHOLDS WITH EMPLOYED PERSONS
Q EXTRA-TERRITORIALORGANIZATIONS AND BODIES

In the broadest sense the Service sector comprises sections G to Q, inclusive.