



Statistics Netherlands

Division of Social and Spatial Statistics
Department of Labour and Earnings

*P.O.Box 24500
2490 HA Den Haag
The Netherlands*

Quality report of 2012 Labour Costs Survey in the Netherlands

Ben Dankmeyer and Johan Planting

Summary: This report describes the quality of the 2012 Labour Costs Survey. The outline is based on the requirements in Commission Regulation No 698/2006 & 973/2007 as regards quality evaluation on labour costs statistics.

Keywords: labour costs, compensation of employees, wages and salaries

Project number: 209515/04

Date: December 2014

1. Relevance

This optional dimension is not taken into consideration.

2. Accuracy

2.1 Sampling errors

The framework for compiling the 2012 Labour Costs Survey (LCS) is derived from the Statistics on Employment and Earnings (SEE). The SEE is based upon a complete register from the Dutch Social Security Organization, containing information about the declaration of earnings and social contributions of all employees in The Netherlands, which originates from the Dutch Tax Authorities.

As such, the SEE covers all enterprises with employees in all economic activities throughout the year 2012 in The Netherlands.

As the enterprises in the sample are equal to the enterprises in the population, coefficients of variation are 0 (zero), both for annual labour costs (D), as well as for hourly labour costs (D/B1).

2.2 Non-sampling errors

2.2.1 Coverage errors

The population for the 2012 LCS is derived from the register of earnings and social contributions (RES) from the Dutch Social Security Organization. Since this register covers all enterprises with employees in The Netherlands, the whole reference population is covered.

It has to be noted that the reference population from the RES is slightly different from the population of Statistics Netherlands' (SN) General Business Register (GBR). This difference in population is ascribed to administrative effects, since there is a certain time lag between the updating of information in the RES and the GBR. The information in the RES is considered to be more up-to-date and more accurate. Consequently, the RES contains enterprises which are not (yet) known in the GBR. The difference is, however, small: in 2012 the RES covered almost 6 thousand enterprises (representing 0,2 percent of total labour costs) that were not present in the GBR. Since most statistics produced by SN are using the GBR as reference population, it could be argued that the LCS is slightly 'over-covered' in relation to other statistics.

2.2.2 Measurement and processing errors

Main sources

The data for the 2012 LCS were derived from multiple sources. The following sources were used:

1. Statistics on Employment and Earnings (SEE) 2012.
2. Register with employers' information from the Tax Authorities 2012.
3. Statistics Netherlands' General Business Register (GBR) 2012.
4. Survey on the Production Statistics 2012.
5. Labour Force Survey 2012.
6. Survey on Sick Leave 2012.
7. Survey on Strikes 2012.
8. Labour Costs Survey 2008.
9. National Accounts (including Labour Accounts) 2012.
10. Statistics on Contractual wage indexes 2012.
11. Various 2012 data and information from: the Ministry of Social Affairs and Employment, the Ministry of the Interior and Kingdom Relations, the Social Security Organization, the Pension Funds, Insurance Companies, Collective Labour Agreements, and several Annual reports of companies and organizations. Income Panel Research (Inkomens Panel Onderzoek, IPO)
12. Vocational training costs 2010
13. System of social statistic data (Stelsel van Sociaal-statistisch Bestanden, SSB) 2012
14. Commuter Statistics 2012

The SEE is by far the main source for the 2012 LCS. This is illustrated by the following:

Variable A (number of employees): was fully derived from the SEE.

Variable B (hours worked): normal working hours paid, paid overtime hours, leave days and public holidays were derived from the SEE. Hours of sick leave and maternity leave, hours of short-time work, hours due to frost or strikes and unpaid overtime hours were derived from other sources: added up these other sources formed less than 8 per cent of total hours worked.

Variable C (hours paid): normal working hours paid and paid overtime hours were derived from the SEE. Non-paid hours of sick leave and short-time time work as

well as non-paid hours due to strikes were derived from other sources: added up these other sources formed less than 1 percent of total hours paid.

Variable D (total labour costs): more than 85 percent could directly be derived from the SEE. For the other 15 percent a combination of the SEE and at least one other source was used.

Since the SEE is the main source for the 2012 LCS, the description of the methods used to reduce measurement and processing errors will be restricted to the 2012 SEE. In section 2.2.4 (model assumption errors) some shortcomings of the other sources used will be considered.

Measurement and processing errors in SEE

With each salary payment, all employers in The Netherlands supply a declaration of earnings and social contributions of their employees to the Dutch Tax Authorities. The declaration does not only concern current employees, but also ex-employees having social benefits or pension. The declaration is redirected to the Dutch Social Security Organization, which is the controller of the register of earnings and social contributions (RES) of all (ex) employees in The Netherlands. Every month, Statistics Netherlands receives information from the RES. The information is used to make the Statistics on Employment and Earnings (SEE).

The whole process of the SEE, from the RES entering SN, to the making of tables for publication, is accompanied by numerous data checks and corrections.

Many control/correction steps are performed fully automatically (For example, the deletion of identical, non-unique, and information in the RES. Or the imputation of missing information: for some enterprises information of one or more periods in 2008 is missing. By combining these enterprises with an employers' register from the Tax Authorities, it can be decided whether the information is correctly missing or not. If the information should be there, the data of all employees concerned are imputed with an automatic procedure).

Other control/correction steps are performed in a semi-automatic way (i.e. the possible error is detected automatically, but the solution has to be performed by hand). These checks are mostly concerned with value range (outliers, or non-existing codes) and time-relation (unlikely differences with previous or later periods). The semi-automatic control/correction is based on a top-down analysis system, which automatically brings forward the (possible) errors that have the largest effect on the output-level.

In 2012 almost 6.1 per cent of all records from the RES were corrected in some way.

An important issue is the deduction of the exact population of employees from the RES. Although the RES is a complete register, a well-defined definition of employees is not evident. Therefore, the number of employees is deducted by using

a set of criteria based on several tax codes, the social security or tax wage, and paid hours. Since the quality of (especially) the variable ‘paid hours’ in the RES is not optimal, possible errors in the determination of the number of employees (and consequently in the amount of labour costs) may occur.

Another possible error may occur from the fact that enterprise-units in the RES may differ from the statistical units in Statistics Netherlands’ General Business Register (GBR). To ascribe the labour costs results to NACE-groups, the enterprise-units in the RES have to be linked to the statistical units in the GBR. Although most units are linked in a 1:1 ratio, a few hundred enterprises are linked in a 1:n or m:n ratio. The labour costs of those latter enterprises cover almost 3.5 per cent of all labour costs. Consequently, the distribution of those labour costs to NACE-groups is not certain. It is estimated that at the publication level (NACE 2-digit), at most 2 per cent of all labour costs are possibly ascribed to a wrong NACE-group.

Variables

Some variables could not be properly estimated due to major measurement and processing errors. These variables are discussed in this section.

The variable D123 is slightly underestimated.

For the variable “payments to employees’ savings schemes” (D1112) only, the most important arrangement, the life cycle saving scheme (levenslooplegeling) was calculated.

The variable “wages and salaries in kind” (D1114) could not be calculated in 2008, but in 2012 we could calculate this with the IPO-data in combination with information of the National Accounts.

In 2008 we could not distinguish the variables D1223 and D1224, due to a lack of proper sources. For 2012 we used “Register with employers’ information from the Tax Authorities 2012” to calculate D1223. For D1224 we calculated the payment of a death benefit. The estimated amounts of both variables are included in D122.

The variable “region” (broken down to NUTS level 1) in table C was problematic. For the 2008 LCS (as for the 2004 LCS), the region was based on the location of enterprises, whereas it should be based on the location of local units. However, information on local units is absent in the SEE. As a consequence of this, all employees of a certain enterprise were allocated to the location of the enterprise, instead of the location where they actually work. Since many companies in The Netherlands have their head office in West-Nederland (NUTS code NL3), the number of employees (and matching labour costs) in this region would be over-estimated. In the 2012 LCS we correct this by using the commuter statistics.

2.2.3 Non-response errors

The description of non-response errors will be restricted to the 2012 SEE.

Unit response rate

The 2012 SEE contains over 110 million records of individual employees (1 record per employee for each month or 4-week period). Since occasionally some employee-records are missing for one or more periods, these records are imputed. The imputation-rate is rather small: in 2012 about 200 thousand (0.2 percent) employee-records were imputed.

Item imputation rate

Almost 6.1 per cent of all 110 million employee-records in the 2012 SEE were corrected in some way for missing or wrong data in one or more variables.

Imputations of variables took place by using (a) the same variable of the same employee of another period (month before or after), (b) other variables of the same employee, (c) average values from other employees in the same enterprise or section.

Within the individual employee-records, several variables contribute to the total annual labour costs (D). Of those variables, about 0.1 percent of the records was corrected in some way.

The variable hours paid was corrected in about 4.6 per cent of all records. This relatively high rate is explained by the fact that ‘hours paid’ in the RES, is not necessary for (and therefore not controlled by) the Tax Authorities. Consequently, most of the corrections applied to this variable concerned empty records

Overall imputation rate

The information for this section could not be provided.

2.2.4 Model assumption errors

In this section some shortcomings of other sources than SEE used will be considered.

As already mentioned before, 85 percent of all labour costs could directly be derived from the SEE. For calculating the other 15 percent several other sources were used (see section 2.2.2). From these other sources information was collected about the relative volume of missing costs. These missing costs were expressed as a percentage of gross wages and salaries or as costs per employee. With the help of the SEE variables “gross wages and salaries” and “number of employees” the missing cost components in the SEE could be accurately calculated or imputed.

For most of the cost components that could not be directly derived from SEE, assumptions had to be made (see table 1). For instance, because the data related to a higher aggregation level than NACE 2-digit × six size classes. Another reason could

be because the data originated from another reference year than 2012. In the former case it was assumed that the proportions or amounts could also be applied to lower aggregation levels. In the latter case the data were extrapolated under certain assumptions.

Table 1. Assumptions made

Variable	Shortcomings	Assumptions made
D11141, Company products	Amounts per employee are not exactly known	Amounts per employee were estimated on the basis of several sources and assumptions
D11142, Staff housing	Amounts per employee are not exactly known	Amounts per employee were estimated on the basis of several sources and assumptions
D1212, Collectively agreed, contractual and voluntary social security contributions	For part of D1212, % of gross wages and salaries are known per NACE 2-digit x 4 size classes	Cell % (of gross wages and salaries) were applied to lower size class levels
D1221, Guaranteed remuneration in the event of sickness	Sick leave % were used for this variable, which were available per NACE 2-digit x 3 size classes	Cell % (of gross wages and salaries) were applied to lower size class levels
D2, Vocational training costs		Cell % (of gross wages and salaries) were estimated on the basis of Vocational training costs 2010
D3, Other expenditure	% of gross wages and salaries are known per NACE 2-digit x 3 size classes	Cell % (of gross wages and salaries) were applied to lower size class levels
D5, Subsidies	For part of D5, % of gross wages and salaries are known per NACE 1-digit x 3 size classes	Cell % (of gross wages and salaries) were applied to lower size class levels

3. Punctuality and timeliness

3.1 Punctuality

For compiling the 2012 LCS figures, Statistics Netherlands used a methodology in which, by using existing sources, it was not necessary to conduct a separate survey on labour costs (see section 2.2.2). This methodology was developed for the 2000 LCS and also used for the 2008 LCS and 2012LCS.

The first preparations started in June 2013 with a preliminary investigation. This investigation was aimed at examining the use of all possible existing sources, both inside and outside Statistics Netherlands. As compared to the 2008 LCS, several new sources had to be used. Up from September 2013 several sources were collected and processed. This work was time-consuming because many basic data had to undergo further preparations.

The most important source for the 2012 LCS, the Statistics on Employment and Earnings (SEE), was provided in January 2014:

The data tables for Eurostat were prepared, including confidentiality checks and sent to Eurostat on 26 June 2014. Due to a wrong format of the data a second delivery on July 4th was necessary.

3.2 Timeliness

The results of the Dutch 2012 LCS were published on 1 December 2014 on StatLine, the electronic databank of SN (i.e. 23 months after the reference period).

4. Accessibility and clarity

4.1 Accessibility

Detailed results of the Dutch 2012 LCS are accessible via StatLine, the electronic databank of SN. The data are presented in four tables:

1. Arbeidskosten; structuur arbeidskosten, bedrijfsklasse, 2012:
<http://statline.cbs.nl/Statweb/selection/?DM=SLNL&PA=82845NED&VW=T>
2. Arbeidskosten; bedrijfsgrootte, bedrijfstak, 2012:
<http://statline.cbs.nl/Statweb/selection/?DM=SLNL&PA=82871NED&VW=T>
3. Arbeidskosten; structuur arbeidskosten, arbeidsduur, bedrijfstak, 2012:
<http://statline.cbs.nl/Statweb/selection/?DM=SLNL&PA=82853NED&VW=T>
4. Arbeidskosten; kerncijfertabel 2012:

(Forthcoming January 2015)

All the information on this databank may be consulted, printed out and downloaded free of charge.

There were no results sent to reporting units.

4.2 Clarity

Concomitant with the results on Statline, a description of metadata is given.:

<http://www.cbs.nl/nl-NL/menu/themas/arbeid-sociale-zekerheid/methoden/dataverzameling/korte-onderzoeksbeschrijvingen/arbeidskostenonderzoek.htm>

Apart from a press-announcement, which is common for all publications on the Internet site of SN, no specific actions were carried out to inform users about links to the data.

5. Comparability

5.1 Geographical comparability

Statistical units

In first instance the information about units does not refer to local units, but to enterprises. By using the commuter statistics we could correct for this.

Populations

The population is considered as complete, since the register of earnings and social contributions (RES) covers all enterprises with employees in The Netherlands. However, a small difference with the population of Statistics Netherlands' (SN) General Business Register (GBR) is present. This difference represents about 0.2 percent of total labour costs (see 2.2.1).

Reference times

All variables refer to 2012, although in some cases assumptions had to be made, since direct data on 2012 were lacking (see 2.2.4).

Classifications

There are no differences between the national classifications and the Eurostat classifications.

Variables

Variable D1112 (payments to employees' savings schemes) could not be measured. However, amounts concerned are included in D1111 (direct remuneration and bonuses).

Variable D3 (other expenditure paid by the employer) includes expenditure covering indirect benefits from canteens and cultural, sporting and leisure facilities and services, and working conditions services. It is calculated using information from the national Accounts.

Parental leave hours are deducted from the working hours and based on SSB 2012 information. The method has been improved.

5.2 Comparability over time

Definitions

In 2012 parental hours were included in variable B (hours actually worked). The hours not worked because of parental leave amount to about 0.2 percent of the total hours worked in 2012.

Coverage

There are no changes in coverage between 2008 and 2012.

Methods

The basic methodology to compile the LCS did not change between 2008 and 2012: both in 2008 and 2012 a combination of several existing sources was used (see 2.2.2).

6. Coherence

Coherence with the Labour Force Survey (LFS)

The number of hours worked per employee according to the LCS and LFS are given in table 2. The LFS data were calculated according to the international definitions (i.e. employees aged 15 and over who performed 1 or more hours of work in the reference week).

As can be seen in table 2, the number of hours worked in the LCS is less than in the LFS for all sections, except for sections B and H. The largest difference is observed in section N, where employees work about 8 percent fewer hours in LCS as compared to LFS.

Table 2. Hours worked per employee in LCS and LFS.

LCS: number of hours actually worked per employee (B1/A1) per year in 2012.

LFS: average actual hours worked in the main job per employee per year in 2012.

NACE	LCS	LFS	LCS as % of LFS
Rev.2	hours	hours	%
B	1693	1658	102
C	1547	1574	98
D	1591	1641	97
E	1672	1691	99
F	1610	1681	96
G	1190	1216	98
H	1559	1507	103
I	872	933	94
J	1541	1550	99
K	1461	1524	96
L	1336	1407	95
M	1447	1459	99
N	1152	1257	92
O	1416	1472	96
P	1103	1167	95
Q	1006	1068	94
R	959	1011	95

S	1144	1240	92
Total	1257	1303	96

The differences in hours worked may be explained by the following:

1. Further analysis of the LFS and SEE (as the main source of the LCS) has revealed that the number of overtime hours is substantially higher in the LFS. The overall difference is almost 2 percent (of total hours worked). It is not completely clear why the number of overtime hours in LFS is higher. A possible explanation could be that respondents tend to overestimate the number of overtime hours in the reference week.
2. The methods to determine the number of normal hours worked are rather different in LFS and LCS. Whereas in LFS respondents are interviewed and asked about their normal hours worked, in LCS these hours were fully derived from the Dutch register of earnings and social contributions (RES). Again, respondents in LFS may tend to overestimate the number of normal hours worked in the reference week.
3. The relatively large difference in section N may be explained by the fact that in LFS temporary workers (employed by employment agencies) are not included in this section (they are allocated to the section where they actually work). In contrast, temporary workers (who work relatively few hours per week) are included in section N in the LCS.

Coherence with the Structure of Business Statistics (SBS)

The wages and salaries per employee according to LCS and SBS are given in table 3. A comparison between LCS and SBS was only possible for sections B to J and L to N, since SBS data of the other sections are lacking. Moreover, SBS data of section C are excluding NACE 32991 and section M are excluding NACE 70101. For a proper comparison, those NACE-groups were also excluded from the LCS-figures.

As can be seen in table 3, in most sections the wages and salaries per employee are higher in LCS as compared to SBS. For sections E and N the differences are quite large. For sections C, H and L, amounts in LCS are lower as compared to SBS.

For a better understanding of the differences between LCS and SBS, it is important to compare the underlying wages and salaries, as well as the number of employees from both sources (see tables 4 and 5).

As can be seen in table 5, the number of employees in LCS and SBS are comparable for most sections. This indicates that the differences between LCS and SBS are

mainly explained by differences in wages and salaries. Indeed, the wages and salaries are lower in SBS as compared to LCS for most sections (see table 4).

It is not completely clear why the wages and salaries in SBS are different from the LCS. However, it should be mentioned that in the LCS, both the number of employees (A1) and the wages and salaries (D11) are derived from the same source (the SEE), which is considered to be complete (since it is based on a complete register with all enterprise and employees). In contrast, the SBS makes use of different (sample survey based) sources to derive at the number of employees and the wages and salaries.

The biggest difference can be found in section N. This is the section that includes temporary employment agencies. It is always very difficult to get the right wage information of this group if you use sample data.

Table 3. Wages and salaries per employee in LCS and SBS.

LCS: wages and salaries per employee (D11/A1) in 2012.

SBS: wages and salaries per employee (v13320/v16130) in 2012.

NACE	LCS	SBS	LCS as % of SBS
Rev.2	x1000 euro	x1000 euro	%
B	79,6	76,9	103
C (excl 32991)	42,3	43,3	98
D	58,4	55,0	106
E	41,9	36,1	116
F	40,4	40,5	100
G	25,9	25,6	101
H	33,9	36,0	94
I	12,6	12,4	102
J	49,6	47,9	103
L	39,6	40,8	97
M (excl 70101)	46,1	44,4	104
N	20,3	14,4	141
Total	31,7	30,5	104

Table 4. Wages and salaries in LCS and SBS in 2012.

NACE	LCS	SBS	LCS as % of SBS
Rev.2	x mln euro	x mln euro	%
B	684	661	103
C (excl 32991)	27584	28084	98
D	1549	1419	109
E	1583	1361	116
F	14357	14269	101
G	33884	33264	102
H	12811	13494	95
I	4347	4245	102
J	11472	10979	104
L	2749	2758	100
M (excl 70101)	21391	20053	107
N	16834	11938	141
Total	149246	142526	105

Table 5. Number of employees in LCS and SBS in 2012.

NACE Rev.2	LCS	SBS	LCS as % of SBS
	x 1000	x 1000	%
B	8,6	8,6	100
C (excl 32991)	651,4	648,4	100
D	26,5	25,8	103
E	37,8	37,7	100
F	355,0	352,7	101
G	1308,3	1299,3	101
H	378,2	374,9	101
I	344,8	342,3	101
J	231,5	229,2	101
L	69,3	67,5	103
M (excl 70101)	463,7	451,2	103
N	831,0	831,9	100
Total	4706,0	4669,4	101

Coherence with the Labour Cost Index (LCI)

The growth rates of labour costs between 2008 and 2012 according to LCS and LCI are given in table 6.

As can be seen in table 6, the labour costs for the total economy have grown almost at the same rate in the LCS and LCI. In most sections the differences between LCS and LCI are rather small. In sections D and K, the growth rates are somewhat higher in the LCS. The difference in section C, with a negative growth rate in the LCI, is striking.

The main reason for the differences between LCS and LCI is the methodology used. The labour costs and the number of hours worked in the LCI originate from different sources and populations than in the LCS. In fact, the LCI-figures fit within the National Accounts (NA). Therefore, the differences between the LCS and the NA (see next paragraph) also apply to the differences between LCS and LCI.

In the Quality report of 2008 we established a big difference in the growth rate for section C (Mining and quarrying,), this had diminished from 23.9 to 3.8 %-point.

Section K (Financial) gives a bigger difference than in 2008. Possibly the financial crises have had influence on this.

Table 6. Growth rates of labour costs in LCS and LCI.

LCS: growth rate of hourly labour costs (D/B1) between 2008 and 2012.

LCI: growth rate of the unadjusted LCI between 2008 and 2012.

NACE	LCS	LCI	Difference
Rev.2	%	%	%-points
C	10,5	6,7	3,8
D	17,2	10,6	6,6
E	13,8	10,5	3,3
F	10,4	6,6	3,8
G	11,2	8,9	2,3
H	6,0	7	-1,0
I	3,7	3,4	0,3
J	6,6	6,8	-0,2
K	13,8	6,4	7,4
L	9,1	9,6	-0,5
M	10,6	7,8	2,8
N	9,2	7,2	2,0
O	8,5	8,1	0,4
Total	9,9	7,2	2,7

Coherence with National Accounts (NA)

Differences in compensation of employees expressed per employee between NA and LCS are small in most sections (see table 7). For the whole economy (sections B to S) the difference is less than .5 per cent. In a few sections, compensation costs per

employee are lower in NA. In other sections, compensation costs are considerably higher in NA.

A comparison between the results of 2012 LCS and 2012 NA is faced with difficulties because of conceptual differences between both sources. The Labour Statistics (LA) within National Accounts (NA) derive their data not only from Statistics on Employment and Earnings (SEE), but also from registers and household surveys, as the Labour Force Survey (LFS). Moreover, the results from the LA have to fit within the framework of the NA, which may result in deviations from other sources.

A closer analysis of both NA and LCS did not reveal striking differences that might explain the differences observed.

Table 7. Compensation of employees per employee in LCS and NA.

LCS: compensation of employees per employee (D1/A1) in 2012.

NA: compensation of employees per employee in 2012.

NACE	LCS	NA	LCS as % of NA
Rev.2	x1000 euro	x1000 euro	%
B	96,4	104,3	92
C	50,5	50,3	100
D	75,9	73,5	102
E	54,6	54,8	100
F	51,9	52,1	100
G	32,2	32,2	100

H	43,6	44,5	99
I	15,4	15,2	100
J	61,1	63,0	98
K	77,2	76,3	101
L	48,8	54,8	94
M	58,6	63,7	95
N	25,1	24,2	101
O	58,1	58,2	100
P	44,0	44,3	100
Q	31,7	32,7	99
R	27,4	28,5	99
S	33,7	32,9	101
Total	40,8	41,3	100