Topic 1, item 1.5: An Issue of Weighting - A Comparative Analysis on Growth Rates in Labour Costs

By
Uwe Pedersen, Statistics Denmark
&
Michael Eliasen, Confederation of Danish Employers

Workshop on Labour Cost, Rome, 5-6 may 2015

Introduction

One of the main purposes of the Labour Cost Index (LCI) is to give a short-term measure of developments in labour costs, i.e. the total cost on an hourly basis of employing labour. Although the LCI is based on a common inter-European framework, defined by Regulation (EC) no. 450/2003, many member states still produce other more or less similar quarterly labour cost indicators for national purposes, which might differ in definition and methodology compared to the LCI, sometimes probably resulting in apparently incoherencies between various labour cost indicators.

Statistics Denmark (the NSI) compiles besides the Danish LCI also the national Index of Average Earnings. Moreover the Confederation of Danish Employers (DA) compiles statistics on quarterly labour cost developments covering their member enterprises. For some economic activities these different quarterly labour cost indicators are actually based on almost similar data. Despite these data similarities, differences in methodologies might result in quite amazing "incoherent" results when comparing the various labour cost indicators with each other, leading to substantial confusion by users.

Such apparently incoherent behaviour has existed more or less over almost 20 years in the Index of Average Earnings compiled by Statistics Denmark and the statistics on quarterly labour cost developments, compiled by DA.

The differences in annual growth rates compiled by Statistics Denmark and DA respectively have been and still are generally present within all NACE sections. However, especially within NACE section F (Construction) the lack of coherence is very obvious, as shown in chart 1.

Chart 1
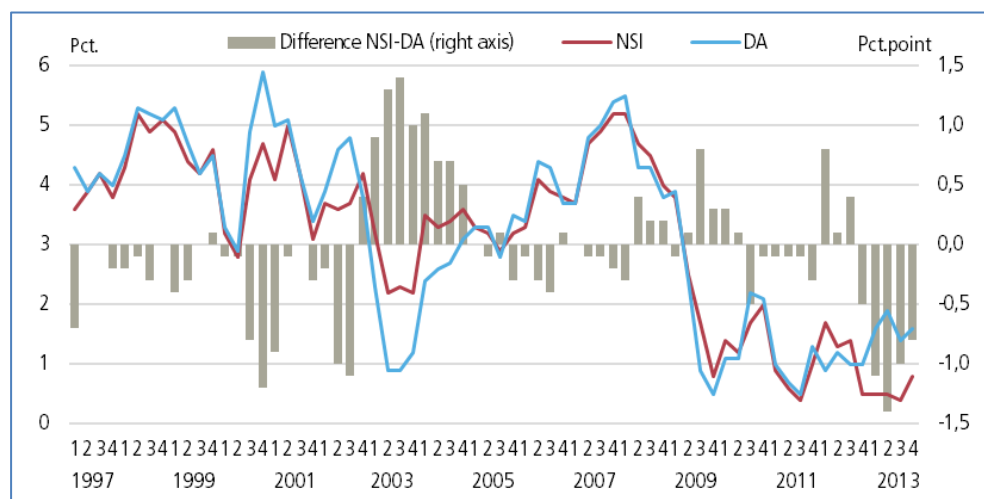
The chart illustrates annual growth rates in percent in labour costs per hour within construction for every quarter since 1997 compiled by DA (blue line) and Statistics Denmark (red line). The right axis in the chart refers to the bars showing the difference between annual growth rates of the two statistics, revealing differences in annual growth up to +/- 1.4 percentage points.

This is in fact surprising, especially because about 90 percent of the labour cost data on construction is initially reported by enterprises, which are organized under DA membership. Thus only a relatively small part of the data from construction enterprises are collected directly by Statistics Denmark, since DA sends a copy of their entire data on labour costs to Statistics Denmark.

Task force set on user demand

The main users of both quarterly labour cost statistics are amongst others primarily the Danish Ministry of Finance, the Danish Ministry of the Interior and Economic Affairs, the Danish National Bank, The Danish Confederation of Trade Unions (LO) and DA. For several years there has been a high demand from those user institutions for a close investigation on the reasons for the different results.

In order to meet the strong user demand and to measure the reasons why these almost similar statistics produce different results, DA and Statistics Denmark in 2013 launched a common study to clarify why the two quarterly labour cost statistics differ. A formal decision to establish a technical task force with representatives from both institutions was set in June 2013 aiming to report back to a committee of main Danish users, which has existed since the beginning of the 1990s and is chaired by Statistics Denmark. It took a year for the taskforce to complete the work and to write a final report, which was presented to the user committee in June 2014.

Working method of the task force

The task force chose to focus its work especially on the construction sector, because - besides the remarkable and obvious incoherencies between the two statistics over the years - there is no other economic activity in the Danish private sector, which is as dominated by DA member enterprises. As mentioned,

approximately 9 out of 10 firms within construction are DA members. By focusing on construction it therefore would be ensured that the population on labour cost data at least initially should be almost identical in the two quarterly statistics.

The task force decided to apply the observed quarterly growth rate 2012 Q2 to 2012 Q3 in hourly labour costs within construction as a case study. It was chosen as a consequence of the fact that Statistics Denmark had published a quarterly growth rate in the Index of Average Earnings amounting to -0.8 percent, i.e. a decrease from 2012 Q2 to 2012 Q3, compared to DA's published increase of 0.4 percent in their quarterly labour cost statistics, resulting in a calculated difference of 1.2 percentage points.

The normal statistical data exchange between DA and Statistics Denmark only regards raw micro data from the payroll systems, i.e. the original data reports sent by the respondent enterprises to DA, which besides applying them in their statistics also sends a similar copy to Statistics Denmark. However these data are only raw and needs to get processed, before they are suitable for statistical purposes, and DA's processing naturally differs from the processing undertaken in the NSI.

The case study clearly had to be carried out on statistical output data after processing, but still on micro level. On the other hand, for data security reasons it is not possible that any micro data ever leaves the "in-house" data servers of Statistics Denmark. So DA sent a copy of their micro output data 2012 Q2 and Q3, which was merged with the output micro data of the NSI for the same period, resulting in a common dataset, comprising the complete original input data from the reporting enterprises and the many statistically processed variables by DA and Statistics Denmark respectively.

After establishment of the common data set, the task force could start the proper analysis of the data, which especially focused on two main topics:
1. an in depth analysis of differences in the data population actually applied in the two labour cost statistics when officially released and the effect on the final results on quarterly growth in hourly labour costs
2. an examination of differences in methodology, when compiling quarterly growth rates and their effect on the final results.

Differences in the data population

The quarterly wage statistics produced by Statistics Denmark and DA builds upon a sample of enterprises engaged in the private sector. Enterprises which have 100 employees or more are automatically obligated to report quarterly pay-roll information to either Statistics Denmark or DA. Enterprises who has less than 100 employees is randomly selected to participate in the survey. In the middle month of each quarter, i.e. February, May, August and November, the enterprises report administrative pay-roll data electronically for each employee. These administrative data consist of wage-cost information (salary, payment in kind etc.) as well as employee specific information (i.e. blue/with collar, ISCO and employment ship information). The information in the pay-roll system has a high degree of precision, since the same information also is delivered to the Danish tax-authorities, and acts as an important part in the income-tax return.

Even though there exists a great deal of overlap between the populations used in production of the two statistics, it is generally the case that Statistics Denmark covers more private firms.

In order to comprehend the contributions of at population differential to the differences in the measured growth rates, the task force tried to eliminate this effect by considering the same observations in both statistics as mentioned above. The results of this analysis are given below.

Differences in data treatment

Even thus the administrative pay-roll data has a high degree of precision some adjustment is needed before this information can be used in the productions of the statistics. When production of the two statistics began in 1993 the methods and principles behind the data treatments was the same. But as time has gone these methods and principles have been drifting apart. These differences have resulted in various differences in the data treatment between Statistics Denmark and DA and can, in some degree, explain why the results differ. Even if we consider the same individual, Statistics Denmark and DA no longer treats the individual pay-role information alike, which causes differences in the way labour costs are calculated and consequently in the measured growth rate.

Differences in the Labour Cost definition

In general Statistics Denmark and DA use the same definition of the labour costs. But at one aspect the unit of measurement in growth rates differ. DA uses the number of actual hours worked as a part of the hourly labour costs, whereas Statistics Denmark uses the number of paid hours. Furthermore Statistics Denmark adjusts/corrects the wage level and the number of hours reported to representing a standard month of 4.33 weeks. This is due to the circumstance that the data report normally only cover 2 weeks of work for blue collar workers paid by the hour and thus have to get adjusted in order to weigh as much as a white collar worker, which gets paid once a month. Adjustment to a standard month is irrelevant concerning DA's labour cost definition because it is based on an hourly basis only.

One would expect that this may result in considerably differences in the unit of measurement (labour cost), but as shown below, differences in the labour cost definitions is not a main explanation of why the two statistics differ.

Compared to the **LCI definition** of labour costs, the labour cost used by Statistics Denmark and DA does neither include irregular payments as well as holiday payment nor a minor part of other labour costs as of payments to mandatory insurance schemes, education funds, taxes, subsidies etc.

Different methodologies in compiling quarterly growth rates

Methodology used by DA

Based on observations for identical enterprises in quarter t-1 and t, the quarterly growth rates are derived by combining growth rates for groups of similar employees in each economic activity (NACE) with a fixed weight. In each industry these similar groups are defined on basis of blue/white collar relationship and their major group of ISCO-08.

In general the quarterly growth rates are derived by:

1) For observations in each economic activity the data is divided between white and blue collar workers. DA is operating with 35 sub-activities (NACE)
2) Every white collard worker is then categorised according to the 9 major groups of ISCO-08 in the starting and ending quarter. This is also done for blue collared workers.
3) Based on this, DA calculates 315 average earnings (hence, each for ever ISCO-08, NACE combination – 9 x 35) as a simple arithmetic mean. This is done for both blue and white collar workers, resulting in 630 average earnings in the starting and ending quarter.
4) These average earnings are then used in calculating an average growth rate. In total DA uses information from 630 growth rates, 315 for blue collar workers and 315 for white collar workers
5) Each of these 315 growth rates are then combined according to a fixed weighting principle, where the fixed weight are changed once every ear. It is the total number of hours worked reported to the yearly SES which acts as weight. This sum is then divided by the same 9 x 35 combination as mentioned above.
6) By combining the 315 growth for blue collar workers and the 315 growth rates for white collar workers, DA ends up with a weighted growth rate for each of the 35 sub-activities and one weighted growth rates for DA as a hole.

Methodology used by Statistics Denmark

The micro data actually applied in the compilations of quarterly growth rates only comprises every single employee employed in enterprises, in which data has passed the statistical validation process for both quarters (i.e. the principal of identical enterprises). This means that missing data for an enterprise in the base quarter t-1 automatically results in discarding the data reported in the current quarter t, even if it has passed all validation processes in the current quarter.

Quarterly movements in the Index of Average Earnings are compiled by weighting average hourly earnings grouped by size classes of the enterprise within every economic activity, according to the following methodology:

1. The micro data regarding the base and the current quarter are grouped by economic activity according to NACE and size classes, corresponding to the enterprise sample design strata.
2. In every NACE*size class stratum an average of hourly earnings is compiled by weighting the hourly earnings from every individual with their number of hours worked respectively, resulting in a weighted average of hourly earnings, which contrary to the DA method take into account if an employee works full time or part time.
3. The stratum averages for quarter t-1 and t are weighted by the corresponding sum of employee according to the Business Register, resulting in a total average of hourly earnings within every NACE section for base and current quarter.
4. For every NACE section the quarterly growth rate is compiled as percentage growth in hourly earnings from quarter t-1 to t.
5. The quarterly growth rates are then finally chain linked to the Index value t-1.

The two methodologies actually represent a fundamental difference in the applied weighting schemes, when compiling the final results on labour cost growth. By weighting growth rates for single groups of employee (DA) with a fixed sum of hours worked, the results are not at all affected by the different average levels of labour costs per hour.

In contrast to this the NSI-method is highly affected by the structural individually diversity of hourly labour costs and moreover quarterly movements in 1) the actual number of employee in the data and 2) movements in the number of hours worked per employee.

This even happens despite the principal of identical enterprises in quarter t-1 and t, since 1) regards the situation of e.g. hiring and firing staff within the enterprises, which might happen between quarter t-1 and t and 2) regards hours worked per employee, which is highly affected by overtime, absence e.g. due to public holidays etc. Even weather conditions might have significant effects on quarterly shifts in hours worked, which basically represents weight shifts, especially within construction enterprises, where days of heavy rain falls might result in a temporary lay-off of construction – especially blue collar - workers, reducing their hours worked substantially.

Main Results of the case study

On data regarding the 10.000 "identical" employees within NACE section F (construction), which have passed the data validation process in the index compilation by Statistics Denmark and the quarterly labour cost statistics by DA as well, it is possible to compile the quarterly growth rate in labour costs per hour based on the two different methodologies and compare the final results.

The actual quite different results compiled are summarized in table 1, evaluating the contribution from 1) applying different labour cost definitions and 2) applying different weighting regime.

Table 1

	NSI labour cost definition		DA labour cost definition	
	(1)	(2)	(3)	(4)
	NSI methodology	DA methodology	NSI methodology	DA methodology
Quarterly growth rate in hourly labour costs	-0,53 pct.	0,41 pct.	-0,79 pct.	0,43 pct.
Difference due to methodology (A)	-0,94 pct.		-1,21 pct.	
Difference due to labour cost definition (B)	0,26 pct.		0,02 pct.	

Applying the hourly labour cost definition of Statistics Denmark and the weighting regime with flexible and current weights, the estimated quarterly growth rate amounts to -0.5 percent, thus below zero. Changing only the weighting regime to DA's methodology, the quarterly growth rate is about 0.4 percent. On the other hand, changing the labour cost definition to DA's, the result based on current weights is about -0.8 percent, whereas a weighting regime as applied by DA also in this case results in about 0.4 percent.

It is obvious, that the difference in weighting regime has a major effect on the final results (A), whereas the difference in definition of hourly labour costs is of minor importance (B).

But could the findings of the common study spawn similar needs for analysis among other LCI compiling member states and lead to a fruitful discussion in the methodology in the index?