The monthly LFS

Content

Summary

This paper gives an introduction to and a description of the monthly publication of the Labour Force Survey (LFS) published for the first time in October 2013. The monthly LFS is based on the already existing quarterly LFS, which has existed since 1994.[1]

The paper will outline the background for the production of monthly results on the LFS and the series that will be published. The main point of focus is on the chosen method for producing the results a *three month moving average with a forecast*. The production of the monthly LFS results will be presented by breaking it down in its different parts; sample size and the data collection, the specific monthly weighting scheme, the forecast method and the formation of the moving average.

The reason for choosing this method is due to the quality issues with producing and disseminating pure monthly estimates. One major issue affecting the outcome was a systematic pattern in the figures, and furthermore, seasonal adjustment of the pure monthly estimates was not possible. However, the chosen method has consequences on how to compare results from the monthly LFS. This will be presented in section 2.

Subsequently, the chosen publication cycle will be presented. The forecast is based on a prognosis that gives a qualified guess on the third month in the moving average that has not yet been collected. Next month the forecasted month will have been collected. The second time the monthly average is published the forecasted month is replaced with collected data. Furthermore, the numbers will also be benchmarked against the quarterly results. This is described in section 3.

The concluding section 4 presents the choices made during the process. Here, the reasons for dropping pure monthly estimates and choosing a three month moving average will be discussed in depth. In addition, the reasons for

[1] LFS in Denmark was a yearly survey from 1983 to 1994. In StatBank Denmark numbers back to 1996 can be found on selected series.

choosing a forecast in the average, and not three collected months, will be described.

Background

The European statistical institute, Eurostat, has published monthly results for the Member States that form part of the Labour Force Survey (LFS). The monthly figures for Denmark have until now not been produced by Statistics Denmark, but have been formed by Eurostat on the basis of the quarterly LFS, which Eurostat has extrapolated with the development in the Danish register-based unemployment. Afterwards, Eurostat has revised the monthly results so they matched the published quarterly figures in the Labour Force Survey. Due to large differences in the population and the seasonal patterns between the register-based unemployment and the LFS, the revisions of the Eurostat produced monthly results that have often been very significant.

As a result it has been a wish from both Eurostat and Statistics Denmark that Statistics Denmark should start producing own monthly results.

In the future, Statistics Denmark will publish two monthly unemployment statistics; the register-based unemployment and the monthly LFS. As a result of the mentioned differences between the two statistics regarding population, definition and seasonal patterns, both the levels as well as the developments are different in the two statistics.[2] Both statistics will be published on the last Thursday of the following month.

Monthly series published

The monthly results in the LFS will be published in the series already published by Eurostat. These can be seen in table 1.

Table 1: Monthly series which will be published.

	Men	Women	Men	Women
	15-24 years		25-74 years	
LFS-unemployment	X	X	X	X
Employed	X	X	X	X

These series are made back to January 2007, and the figures will also be revised in the Eurostat database back to 2007.

[2] For further information about differences between LFS and the register-based unemployment statistics look here

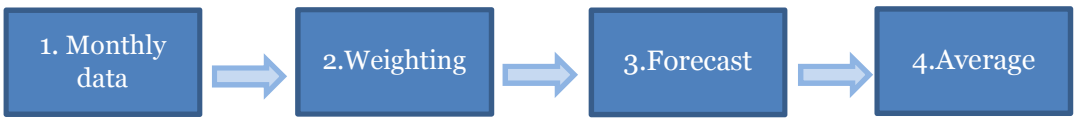
1. The production of monthly LFS results

The Danish LFS is basically a quarterly statistics, implying that the production system will still be optimized to a quarterly survey. The monthly statistics are therefore to be seen as a supplement to the quarterly results and will only be published on a few aggregated series. Respondents participating in the monthly LFS will also be the data basis for the quarterly LFS.
Around 22,000 interviews are carried out in the LFS per quarter. On a monthly basis this results in approximately 7,500 interviews.

In the following section the production of monthly LFS results will be described. It will go through how Statistics Denmark produces unemployment- and employment results from the Labour Force Survey; from the collection of the sample to the point where the figures can be found in StatBank Denmark. As an example, we have chosen the production of results for the month of September 2013.

The monthly LFS results are based on a *three month moving average*, which firstly are published as a result of the four following steps.

Figure 1: Production of monthly results



The monthly data

The sample is collected monthly after the structure 4 weeks – 4 weeks – 5 weeks, implying that the last month in the quarter always contains more interviews than the two previous months. This is balanced in the weighting of the monthly results, so that respondents are equally distributed in the weighting of the three months in the quarter.

Weighting

The LFS operates with a weighting scheme where the incoming survey results are weighted before being published, so they state results for the whole population aged 15-74. As a result of a relatively large non-response the quarterly weighting scheme is pretty advanced in order to make the survey as representative as possible. There is drawn on registers as auxiliary information on, e.g. age, gender, region, educational level and status and socio-economic status. In addition, the register-based unemployment is also a part of the weighting scheme.

The monthly weighting scheme is a simplification of the quarterly weighting scheme, since the amount of auxiliary information as well as the number of subgroups is reduced. It has been necessary to simplify the model due to the monthly sample being much smaller than the quarterly sample. If the model is too complex for a smaller sample size, the model would simply collapse.

Since we are working with a moving average where the publication month is September the collected data from August and September will be weighted to population level. The scheme for the monthly weighting can be seen in table 2.

Table 2 – Monthly LFS weighting scheme

	Variables	Groupings
Crossing information	-age	6 groups
	-gender	2 groups
Crossing information	-age	6 groups
	-education	3 groups
	-socio-economical status	7 groups
	-citizenship	3 groups
	-Register unemployment	4 groups
	-gross income	4 groups

Forecast

The next part of the production concerns the application of a forecast of the third month in the three month moving average – in this example October. When September is collected the whole monthly series back to 2007 is used to give a prognosis of the estimates of the coming month. Four sub-series are forecasted (gender and the age-groups 15-24 and 25-74) for both unemployed and employed. The forecast method is chosen to make the results timely. Alternatively, the last month in the average – October - should be collected. This would have pushed the dissemination of the results a month forward.

Moving average

A three month moving average can be formed on the basis of the two collected months and the last month forecasted and not collected. The publication month is the middle month in the three month moving average. The strength of this model is that the monthly results contain the latest development without this influencing on the timeliness.

Figure 2: The three month moving average with a forecast

Results for:	Average of:		
September	August	September	October
	Collected data	Collected data	Forecast

2. The chosen method: Three month moving average

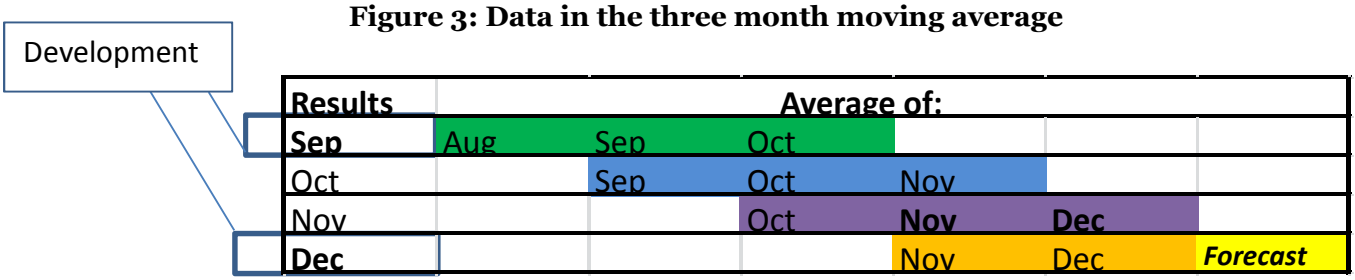
To apply a three month moving average with a forecast has consequences for the interpretation and application of the results.

How to interpret a three month moving average

The three month moving average method results in that comparisons of monthly results from one month to the next are difficult. This is due to the fact that monthly results based on a three month moving average partly shares the same data as the monthly results for the two prior months and the two following months. Developments between following months are thereby hard to interpret. Changes in the unemployment between e.g. August and September will not necessarily be telling of real developments, because it is hard to judge if the change in the unemployment is due to July not being a part of the average any more or due to October now being contained in the three month moving average.[3]

As a result of this it is recommended to go three months back for comparisons and interpretations of developments in the series. This to secure that you compare with the last data point that does not share data with the present month. This is illustrated in figure 3.

Figure 3: Data in the three month moving average



When producing the monthly results for September, August, September and October will be contained in the moving average. As the figure illustrates September month has shared data with the calculated averages for the months October and November. Thus, data for the month October are part of the moving average for all three months, which make comparisons between following months hard to interpret. It is not until the results for December are formed that the series will have a new data point that does not share data with September. Therefore, Statistics Denmark recommends a time span of three months when you are to comment on the results.

Seasonal adjustment

The three month moving average will finally be seasonally adjusted. Regarding the monthly results seasonal adjustment will be indirect. This means that it is the sub-series that are seasonally adjusted and that main series are the sum of sub-series (employed and LFS-unemployed distributed on gender and the two age-groups 15-24 and 25-74). Seasonal adjustment is the last element, so both

[3] Furthermore, estimates on single months will also result in an assessment based on 1/3 of the moving average. You will face the problems concerning pure monthly results discussed in depth in chapter 4, which is subject to great uncertainty and a strange pattern in the development of the unemployment.

the forecast and the moving average are formed when the sub-series are seasonally adjusted.

3. Publication of the data, revisions and uncertainty

Preliminary and collected results

The moving average with a forecast means that there will be an ongoing parallel course in the published numbers on the monthly LFS. The results for September, which are published at the end of October, are *preliminary results*, since the last month contained in the calculated average, October, is not collected, but is a forecast. As a prognosis the forecast can be more or less accurate. When October is collected new results for September are published. These results will first be published at the end of November in StatBank Denmark, where the time series will be revised against the background of the collected results.

Revisions regarding the quarterly results

In order to obtain the largest possible comparability with the monthly results for Denmark published by Eurostat, there will be revisions of the monthly results when the quarters are closed. The monthly results are simply benchmarked up against the quarterly results.

This means, e.g. the monthly results for July, August and September will be benchmarked up against the quarterly results for the 3rd quarter on the relevant series. This is done by correcting the differences between the monthly and quarterly results, so the monthly results will be in accordance with the quarterly results. The results for the 3rd quarter will be published at the end of November. The seasonally adjusted series, however, will change, since one has to go three years back plus the present year in order to find fixed results that will not be affected by seasonal adjustment. For the non-seasonal adjusted series, table 3 shows when the results are published.

Table 3: Publishing of the monthly results

Results for:	Published first time - inclusive forecast	Published second time - data are collected	Published third time – revision against the quarterly results
September	October	November	November
October	November	December	February
November	December	January	February
December	January	February	February

If September is taken as an example, then September will be published for the first time in October. At this point the last month (October) will be a forecast. The result is thus preliminary. When for the last month is collected (October) in the three month moving average forming the results for September, the second publishing will happen (in the end of November).

September is the last month in the 3rd quarter, so in the end of November the quarterly results for the 3rd quarter will be ready. Here, it is possible to find smaller deviations between the collected result from September and the quarterly result. Therefore, the monthly results are revised with the quarterly results as a benchmark. At the end of November the 3rd quarter is published, and the monthly results can be published for the third time based on the revision after the benchmark against the quarterly results.

The timespan between the first and final publication of the results depends on how a given month is placed to the quarterly publication. As an example, the results of October, the first month in the 4th quarter, will first be revised up against the quarterly results in February three months after it was published for the first time.

Uncertainty

Uncertainty on the monthly results is calculated in another manner than on the quarterly results as a result of the chosen method – a three month moving average with forecast. On the quarterly results the uncertainty will be calculated based on a 95 pct. confidence interval directly on the collected results. This is not possible on the monthly results, since the newest results, as mentioned, contain a 1/3 element of a forecast. Here, the uncertainty will be calculated against the background of the two collected months and the forecasted month. They all weight 1/3 of the total uncertainty of the monthly results. Typically, the uncertainty will be larger on the forecasted month, because this month is not collected, but is a prognosis.

The numbers for uncertainty published in 'News from Statistics Denmark', are calculated in the way mentioned above. It is only possible to calculate uncertainty on non-seasonally adjusted results.

The size of revisions

From 1 January 2013 to 1 June 2013 the revisions on the LFS-unemployed between preliminary and collected results have, on average, been -1,000 unemployed persons. The revisions between the collected results and the results that are benchmarked up against the quarterly results have, on average, been, 1.000 LFS-unemployed from January to June 2013. For both revisions the results have been adjusted both upwards and downwards.

In table 4 June 2013 is chosen as an example, since this month is the last data point in the series, where there are final results. This is due to the fact that the whole quarter has to be collected, before it is possible to form a new average against the background of the benchmarking.[4] Here, it is shown that the LFS-unemployment against the background of a preliminary average, including a forecast over July, arrives at 187,000 LFS-unemployed. After the collection of July, the average is revised up with 5,000 LFS-unemployed. Since the quarter is closed and the benchmark is done, the average for June is still 192,000.

Table 4: Revisions in the average of LFS-unemployed aged 15-74

	Preliminary average	Collected average	Benchmarked average

[4] It is not until when the 3rd quarter is published 20 November 2013, that it will be possible to make the same calculations for July, August and September.

June 2013	187.000	192.000	192.000

The revisions made from benchmarking up against the quarterly results are not part of the calculation of uncertainty.

4. Considerations in connection with the monthly publication

Quality and timeliness are central

Two parameters have been decisive for the production of monthly results. The first has been to secure as high a quality as possible. The other has been to secure the timeliness of the results.

This has been the important reasons that made Statistics Denmark choose the above-mentioned model for monthly publication. In this chapter, we will go deeper into the considerations that are the basis for the chosen method, by presenting the methodological choices made as well as the methods not chosen.

Different types of monthly LFS' in Europe

Overall there have on a European level typically been three ways of producing monthly LFS results.

In Sweden the whole production has been rescheduled from being a quarterly statistic to a monthly statistic. This already happened in 1970. Sweden produces all the LFS as a monthly statistic, otherwise Denmark which produces monthly results on a few aggregated series, and has the quarterly statistic as the main survey.

The reason that Sweden has had the opportunity to produce monthly results on the many finely divided groups is due to the size of the Swedish sample. The number of monthly interviews in the Swedish LFS is over 20,000. Compared to this the Danish LFS has about. 7,000-8,000 interviews per month. A reschedule of the Danish LFS to a monthly LFS would require a considerable expansion of the sample, which has not been possible.

This meant that there were two other options for further investigations.

The first option was to publish pure monthly results on the basis of the collected months. A pure monthly result means that the estimate for a given month is formed exclusively on the collected data for the month in question. This is the way the monthly results are published in Finland and Austria.

The other option was to publish results based on a three month moving average. This model is used in Norway and England.

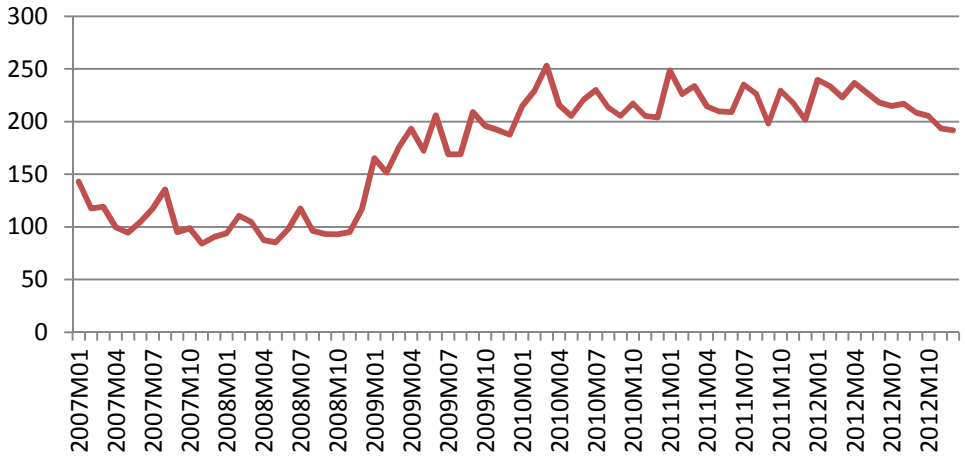
The starting point for the work with the Danish monthly LFS was a wish to establish a monthly statistic based on pure monthly results, since these results are, by far, the most applicable, if you wish to compare following months.

Challenges on producing pure monthly estimates

The pure monthly results soon showed to create a number of challenges. Some of them were not possible to solve.

Looking at the monthly development in the aggregated series on pure monthly results in figure 4, quite large fluctuations appear in the LFS-unemployment. For example you see a decrease from March 2010 to April 2010 in the LFS-unemployed of about 37,000 and an increase in the LFS-unemployed from December 2010 to January 2011 of about 44,000. The series is very volatile due to the relatively small sample size.

Figure 4: Pure monthly results for the LFS-unemployed aged 25-74 (1,000 persons)

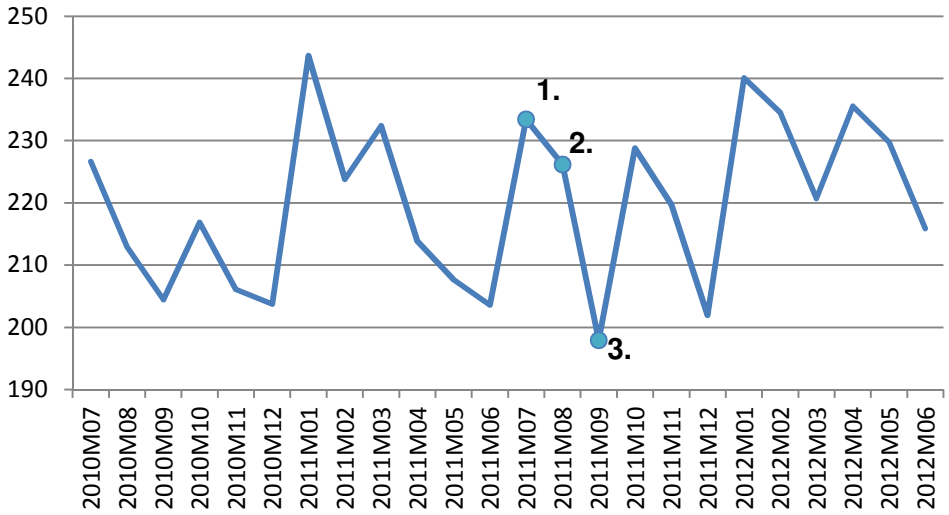


An important element in stabilizing the time series is linked to the seasonal adjustment, since seasonal adjustment will stabilize seasonal fluctuations from month to month.

Here the first major problem arose. It was not possible to seasonally adjust neither the sub-series nor the aggregated series on unemployment and employment.

The most prominent problem was an inexplicable pattern in the monthly results, which is especially pronounced in the years 2010-2011. This can be found in figure 5, where the focus is in on these years, where the pattern is most pronounced.

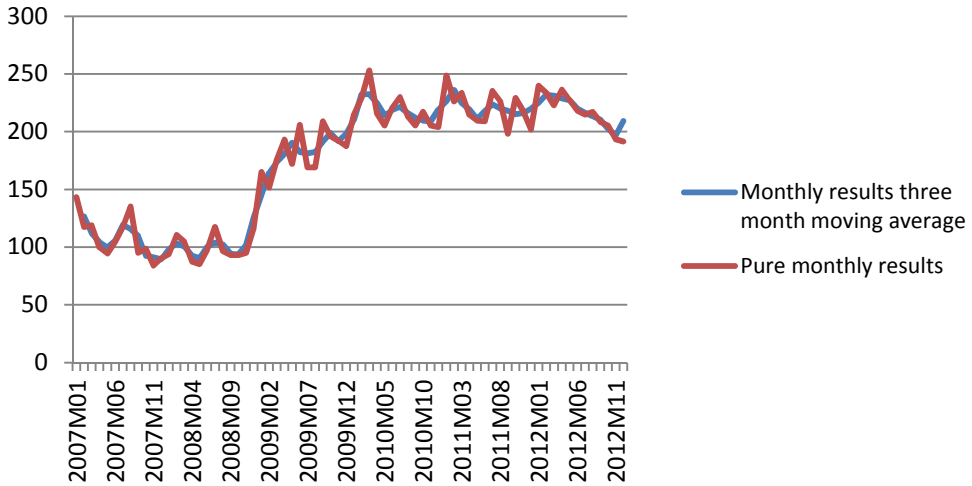
Figure 5: Pattern in the LFS-unemployment on pure monthly results (1,000 persons)



As can be seen, there has been a tendency that the monthly estimates on the three months, which constitutes a quarter, systematically assess the unemployment as high in the first month. Then the unemployment falls in the second month, and finally falls markedly in the third month, which is repeated several quarters in a row. The pattern is not a problem over quarters, since this is compensated, but it posed fundamental questions on the quality of the pure monthly estimates. Thorough investigations were made from the side of Statistics Denmark without finding a convincing explanation to this pattern. This was to be decisive in deciding that the pure monthly estimates were not of a quality for publication.

It turned out that the three month moving average, which method is described in chapter 2, solved more of the problems connected to the pure monthly estimates. First and foremost, the moving average made the time series much more stable. Moreover, it made a seasonal adjustment of all the series possible, and the strange pattern was removed. This is shown in figure 6.

Figure 6: Three month moving average and raw monthly results in the LFS (1,000 persons)



As is shown, the three month moving average is much more stable than the much more volatile series on pure monthly estimates. At the same time, the systematic pattern in the years 2010 -2011 is removed, which can also be seen in the figure.

Choosing the model for the three month average

Also in relation to producing results with a three month moving average choices were to be made regarding, which model should be taken in use. The choice of the model was primarily a balancing of data quality hold against the timeliness of publication. It was central that the results should not be published too long after the publication month, ideally at the end of the following month.

This meant that it collected all three months and publishing one month later was excluded. Furthermore, it was important to publish on the middle month, so the latest development would be contained in the three month moving average. If one instead chose to publish on the last month you would not get the latest development in the average.

So for the chosen method the middle month is published. The last month, October, in this model is a forecast, a prognosis, which is based on the earlier months and tries to give a qualified guess on the development in the last month not yet collected.