

Documentation of real-time unit labor cost and labor share data

Sophocles Mavroeidis Mikkel Plagborg-Møller
James H. Stock

February 16, 2013

1 Overview

The spreadsheet `rt_ulc_p_ls.xls` contains quarterly real-time data on unit labor cost (ULC), the implicit price deflator and the labor share (which is the ratio of ULC and the deflator) for the business and nonfarm business sectors. Data vintages are available for every quarter since 1971Q2. The data is from the Bureau of Labor Statistics (BLS) Preliminary news releases, except for a few vintages that use the Revised release due to the Preliminary one being unavailable. The BLS Productivity and Cost program releases a Preliminary estimate of the previous quarter's ULC and implicit price deflator about a month into each quarter, while the Revised estimate follows about one month later. The data releases are in the form of indexes which change base years approximately every decade.

For most vintages, only index figures for the 5 to 8 most recent quarters are listed in the news releases. When the index is revised significantly, which happens irregularly but often in the third quarter, the news releases include data that goes further back. The spreadsheet includes precisely the data listed in the Preliminary (or when the Preliminary release is unavailable, the Revised) news releases; it does not try to extrapolate. Since most of the data corresponds approximately to what was known around the middle of each quarter, the spreadsheet has the same general format as the real-time dataset available for other macroeconomic series on the website of the Philadelphia Fed.

2 Spreadsheets

In the spreadsheet `rt_ulc_p_ls.xls`, “b” means the business sector, while “nfb” is the nonfarm business sector. The sheet “ulc_b” includes a number of comments (yellow cells) about changes in base years, missing data, etc. The column header lists the data vintage, the precise date at which the news release was disembargoed and the prevailing index year. The sheets starting with “p” are for the price deflator. The sheets starting with “ls” are for the labor share, which until 2001Q3 is calculated as 100 times the ratio of ULC to the deflator (from 2001Q3 on the labor share is directly available in the data source).¹

The spreadsheet `rt_ls_format.xls` extracts five real-time labor share series for use in empirical work. Their labels have the form `RT_dlsmXQ`, with “X” ranging from 1 to 4. For example, `RT_dls2Q` is then-current estimate of the log labor share two quarters ago minus the then-current estimate of the log labor share three quarters ago. The spreadsheet also includes two sheets that calculate vintage *growth rates* of the *log* labor share (these are not influenced by changes in base years).

3 Sources

Vintages from 1971Q2 to 1993Q4 are manually typed in from scanned PDFs of the BLS news releases, available in the St. Louis Fed’s FRASER document database. Vintages from 1994Q1 to 2001Q2 are parsed from electronic news releases available on the BLS website.² Finally, vintages from 2001Q3 and onward are parsed from vintages of the BLS’ internal “edit 60” flat text file, which contains complete series going back to 1947Q1.³ The “edit 60” files for the Preliminary and Revised releases were kindly supplied to me by Shawn Sprague.

An alternative data source is the BLS’ Monthly Labor Review, which normally includes a section called “Current Labor Statistics” that lists ULC and deflator indexes for the most recent 8 or so quarters. Scanned issues of

¹Since the labor share is calculated as a ratio, its average is not precisely equal to 100 in the base year. However, the multiplicative adjustment necessary to correct for this will be miniscule and is therefore omitted for simplicity.

²Parsing is carried out by an Excel macro in the spreadsheet `format_news_release.xls`.

³Parsing is carried out by Excel macros in the spreadsheets `format_edit60_ulc_p.xls` and `format_edit60_ls.xls`. It may appear that the vintage series exhibit too dramatic changes from 2001Q2 to 2001Q3; however, these changes are due to the usual extensive revisions in September of each year and not due to some data incompatibility.

the Monthly Labor Review are available in the EBSCOhost database. However, the scans are sometimes of poor quality, and the published data does not seem to correspond precisely to either the Preliminary or the Revised releases.

4 Data breaks

The base year (during which the indexes average 100) changes a number of times in the dataset.

1. In 1981Q1 it changes from 1967 to 1977.
2. In 1990Q3 it changes to 1982.
3. In 1996Q1 it changes to 1992.
4. In 2010Q3 it changes to 2005.

In 1976Q3 the definitions of the business and nonfarm business sectors changed. Until this quarter, they had been called “total private economy” and “private nonfarm sector” and included output imputed for owner-occupied dwellings and the household and institutional sector, as well as statistical discrepancy. From 1976Q3 to 1982Q2 they were known as “private business sector” and “nonfarm business sector.” In 1982Q3 the word “private” was dropped from the former label (the definition of the series was not changed).

5 Missing data

In the following quarters, only the Revised news release was available to me (often because the Preliminary one was not published), and so the data in the spreadsheet is based on said release: 1979Q1, 1992Q2, 2003Q3, 2004Q4. In 1998Q2 the Preliminary release was published in March, one month later than usual.

In 1976Q3, when the definition of the business and nonfarm business sectors changed, data for the most recent quarter was missing for the nonfarm business sector in the Preliminary news release. It is listed as missing in the the spreadsheet.⁴

⁴Users of the spreadsheet may want to input the Revised data instead in this quarter.