Restoration of consistent business and consumer survey data for Ireland

15/01/2019

Background

The Harmonised EU Programme of Business and Consumer Surveys (BCS) covers all EU Member States and Candidate Countries, allowing for direct comparisons of business cycle developments across countries and the derivation of consistent European aggregates. For many years, Ireland has been the only exception from the EU-wide coverage, after the European Commission's long-standing BCS partner, the Economic and Social Research Institute (ESRI), ended the cooperation in May 2008. Despite considerable efforts, it proved very difficult to find a new partner institute collecting survey data for Ireland, resulting in the exceptional situation of missing BCS data for the country for long periods. Bits and pieces of data were collected in certain sectors for certain periods, being difficult however to reconcile with previously collected data. Only in May 2016 a stable partnership comprising all five sectoral surveys could be reinstalled with the Bank of Ireland (BoI). In May 2019, three years of data will be available, which is the minimum for seasonal adjustment and the minimum length as from which DG ECFIN of the European Commission publishes newly collected data. However, for meaningful business cycle analysis, much longer time series are needed to be able to assess the current situation against historical developments, not least in the country surveillance work of DG ECFIN.

In order to (re-)create a consistent set of survey data for Ireland across the four business sectors (industry, services, retail trade, building) and consumers, an effort was made to link the data collected in 2016-18 with the historical data-sets using econometric techniques. For the business surveys, data is generally missing between May 2008 and April 2011 and May 2012 and April 2016. Consumer survey data is partially missing between May 2008 and April 2009 and May 2015 to April 2016. In all cases, the available partial data sets (from different data providers) appear to feature different long-term averages, thereby requiring level shifts to make the data comparable in time.

The Commission strived to restore almost all monthly survey questions (23 business and 11 consumer survey questions).[1] Moreover, given its importance in gauging the business cycle and, more specifically, in complementing the assessment of the output gap, the quarterly question on capacity utilisation in industry has also been included in the exercise. The work focusses on restoring non-seasonally adjusted data, such that seasonal adjustment can be consistently applied to the reconstructed series subsequently.[2]

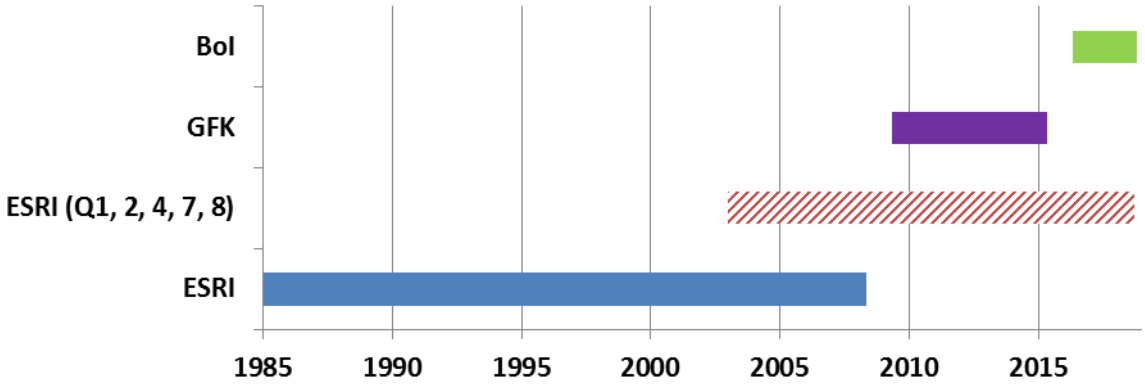
Obviously, it is impossible to generate 'true' data for the missing periods; all recreated data derive from certain assumptions about the co-movement with other data, which can be disputed. With this

[1] Only the monthly questions Q10 of the Consumer survey (good moment to save?) and Q2 of the construction survey (factors limiting production) could not be restored.
[2] See the Appendix for details of how seasonality was dealt with in the reconstruction process.

limitation clearly in mind, the aim of the present note is to describe the employed techniques in detail and series by series, in order to create maximum transparency about the underlying assumptions and techniques.

Consumer survey

For the consumer surveys, three historical BCS data sets are available: ESRI provided all BCS questions as a partner institute from January 1985 to April 2008; moreover, since 2015 it has provided (free of charge, but with a one-month time lag) data for questions 1, 2, 4, 7 and 8 going back to January 2003. Over the 5½-year overlapping period, these five series are practically identical to the original ESRI source. After a marginal adjustment, they can thus be used directly to extend the series. GFK UK provided data for Ireland as a partner institute from May 2009 to April 2015, and the Bank of Ireland (BoI) is the Commission's current partner institute (since May 2016). In order to limit the number of different sources used in the process, GFK data is used only for questions 11 and 12, for which relying on the sole exploitation of the two ESRI and the BoI data sets did not deliver satisfactory results. Not using the GFK data in the reconstruction process of the remaining questions has the advantage that they can be used as an ex-post reasonableness check of the generated data.



The general idea for the restoration of consistent consumer survey series is first to extend the old ESRI data series with the available information on other series up to April 2017.[3] In a second step, if required, the series are adjusted in level so that the average of the first common year with BoI (between May 2016 and April 2017) matches.[4] In a last step, BoI data series are used (without any adjustment or modification) from May 2016. In this way, new observations sent by BoI every month can be added without further adjustment (save seasonal adjustment, which is performed on the complete restored time series).

[3] To this end, this work broadly follows the 'Modelled data approach' described in the UN Handbook on Backcasting (forthcoming).

[4] A level shift can be required to align all series, in order to ensure that their long term developments are consistent. Indeed, survey balances should be interpreted relative to their long term average, and it cannot be assured that series coming from different providers would have the same long term average over comparable samples. For instance in the industry sector, BoI readings for industry question 1 (assessment of past production) are above the highest point ever registered by ESRI between 1985 and 2008. Since there is no reason to think that production growth in 2017 is faster than at any time between 1985 and 2008, a level shift is required to align the series.

For questions 11 and 12, an intermediate step is required, notably the inclusion of GFK data series. After extending the ESRI series, both GFK series and the reconstructed series are aligned so that the averages of the first common year of the two series (between May 2009 and April 2010) match. Then, the extended series values are replaced with GFK values from May 2010 to April 2015 and, finally, both series are again aligned on the last common year (between May 2014 and April 2015), to ensure a smooth transition in May 2015, between the end of the GFK series and the remaining part of the extended ESRI series.

For the other questions, although GFK data was not used directly, it can be used for an ex-post assessment of the reconstruction process. The reconstructed series show high correlations with GFK series over the common sample (see Table 1). Correlation is below 0.8 for questions 6 and 9 only, where GFK series show a range of values which seems too wide to match the volatility of ESRI's series. In addition, correlation of the confidence indicators computed with the reconstructed series[5] and GFK series is very high, at 0.97.

Table 1 - correlation of the reconstructed series and GFK series (May 2009-April 2015)

Question	Question theme	Correlation
COF	Confidence indicator	0.97
1	Past financial situation	0.90
2	Future financial situation	0.89
3	Past general economic situation	0.95
4	Future general economic situation	0.96
5	Past consumer prices	0.91
6	Future consumer prices	0.60
7	Future unemployment	0.95
8	Right moment to make major purchases	0.80
9	Spending on major purchases	0.48

Detailed methodology

The first step (extending ESRI's data series from May 2008 to April 2017) needs some series-specific explanation. There are three different cases, coming with different methods.

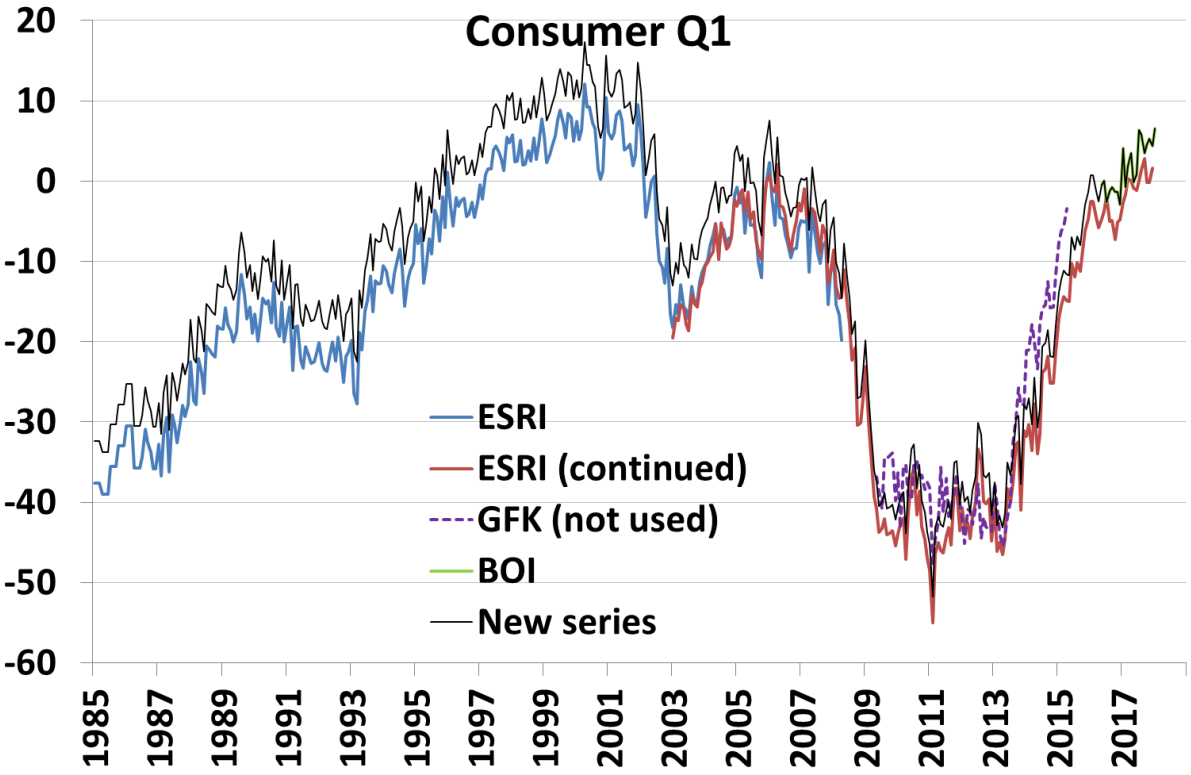
1) The first is to rely as much as possible on ESRI's data for the questions that were not discontinued after April 2008 (questions 1, 2, 4, 7 and 8) and that we received retro-actively in 2015. For these questions, ESRI provided data from January 2003. While the total balances do not exactly match those of the old questions sent by ESRI up to April 2008, the differences are marginal. These series are therefore used directly to extend the old questions up until December 2017, with only a small adjustment in level so that the average of the last common year (between May 2007 and April 2008) matches.

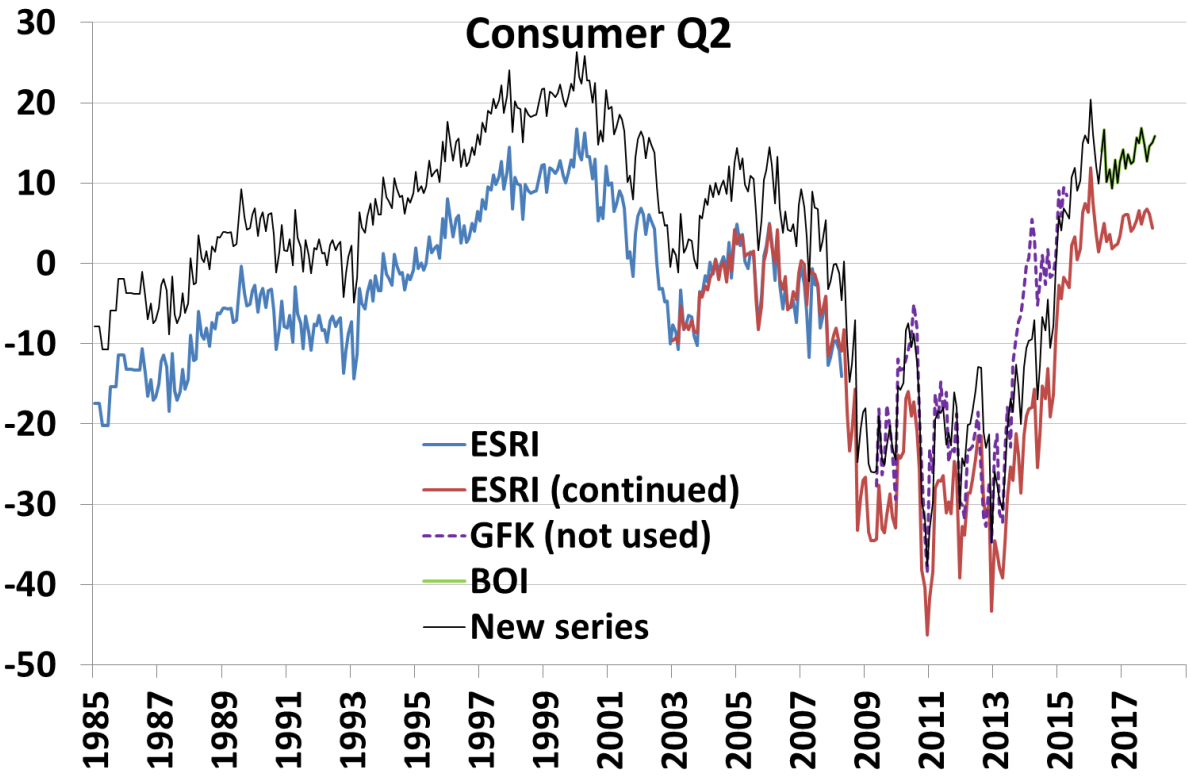
2) The two questions concerning prices (5 and 6), where no post-2008 ESRI data is available, are extended with overall inflation (computed as the year-on-year changes in overall HICP). Question 5 is a backward looking question, so inflation is lagged by 2 months for this question, while an 8-month lead is applied on inflation for the forward looking question 6. The lead or lag is chosen in order to maximise correlation with the ESRI series over the

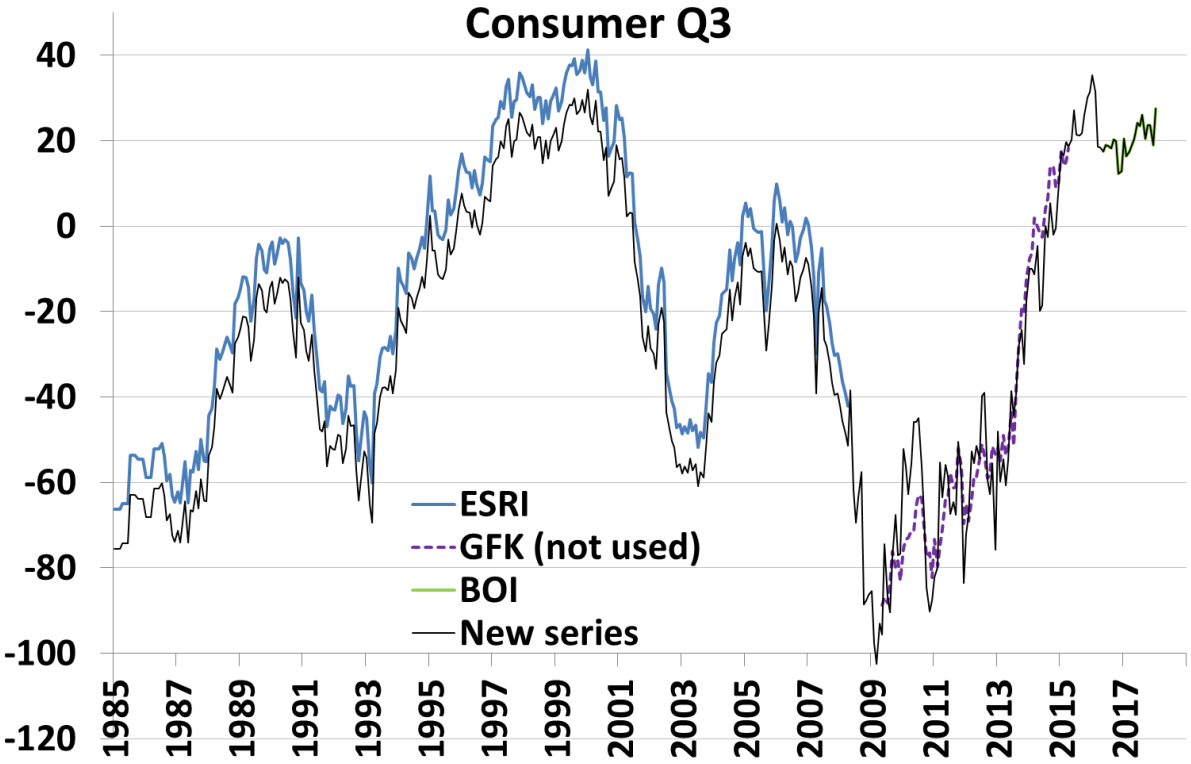
[5] The consumer confidence indicator is based on questions 2, 4, 7 and 11. Among these, only question 11 was reconstructed using GFK series.

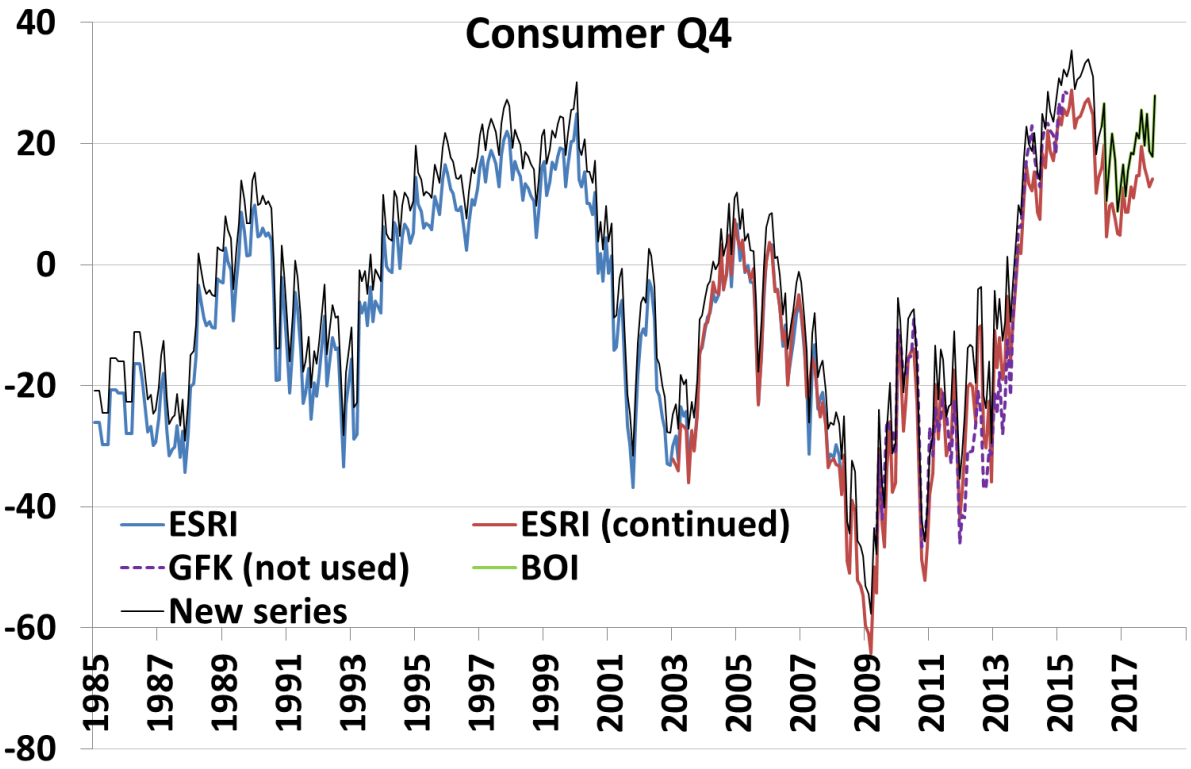
common sample (from January 1997 to April 2008). Moreover, in both cases, the inflation series is rescaled so that its average and standard deviation match that of the ESRI series over the common sample.

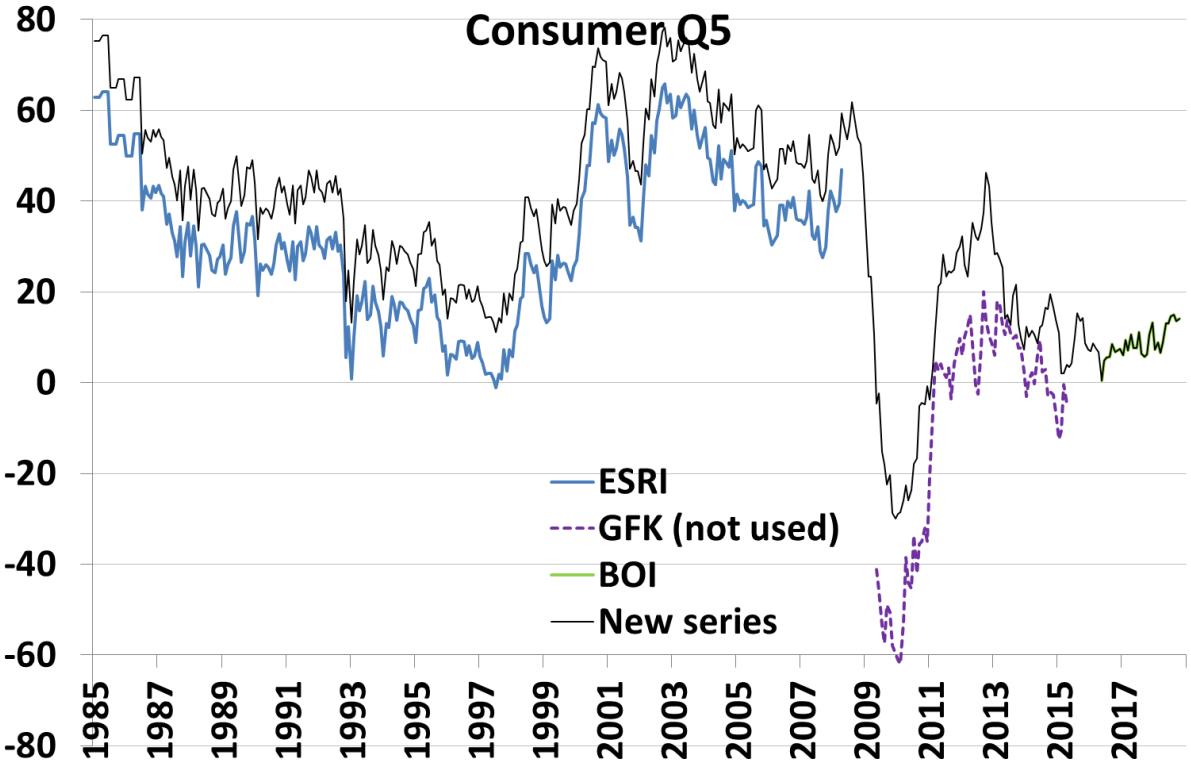
3) Finally, for the remaining questions (3, 9, 11 and 12), where post-2008 ESRI data is also unavailable, the main idea is to use Partial least squares regressions (PLS) in order to reconstruct missing data. PLS is particularly well suited for datasets which are very large and/or feature many collinear time-series and aims at extracting factors from the dataset such that the covariance between the factors and the target series is maximised. In the present cases, the dataset is not that large but includes series that are highly collinear, namely the questions extended on the basis of the second wave of ESRI-data (questions 1, 2, 4, 7 and 8). Practically, the PLS regression first extracts from the dataset two factors that are computed such that the covariance between them and the target variable is maximised. Once the latent factors are computed, ordinary least square regression is used to project the factors on the target variable for the missing data points.

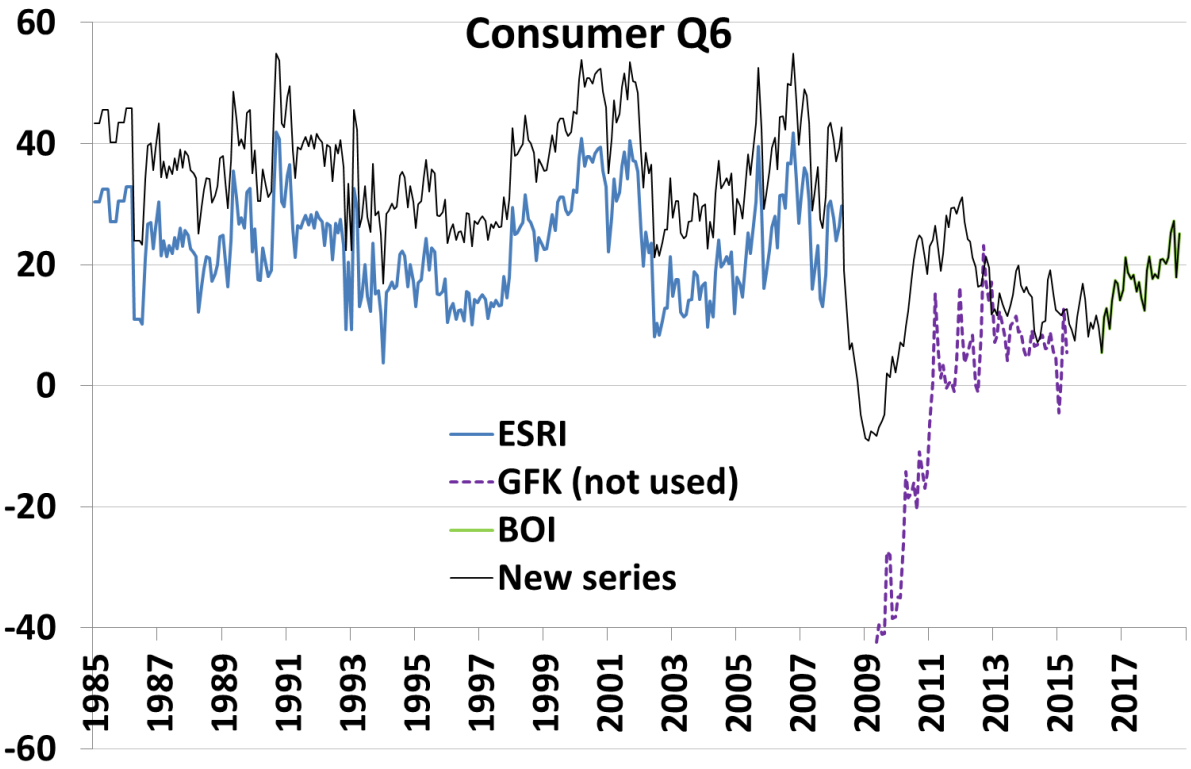


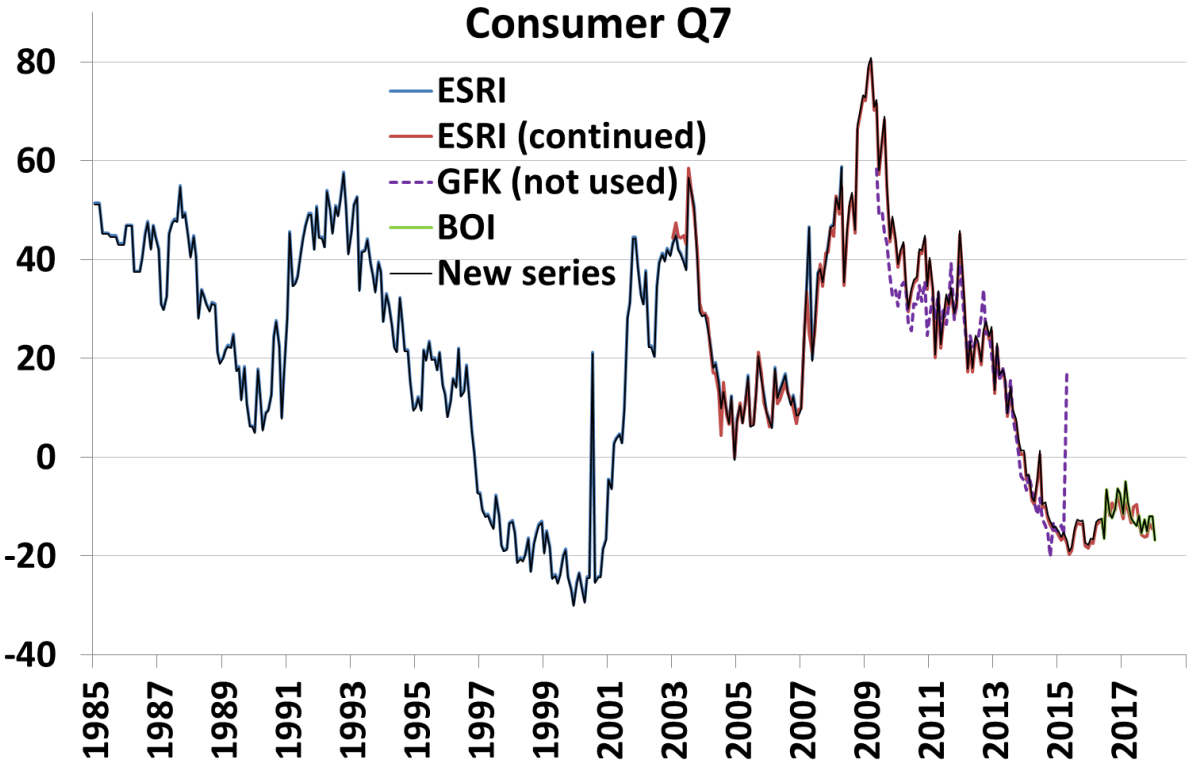

Consumer Q2
30
20
10
0
-10
-20
-30
-40
-50
1985
1987
1989
1991
1993
1995
1997
1999
2001
2003
2005
2007
2009
2011
2013
2015
2017
ESRI
ESRI (continued)
GFK (not used)
BOI
New series

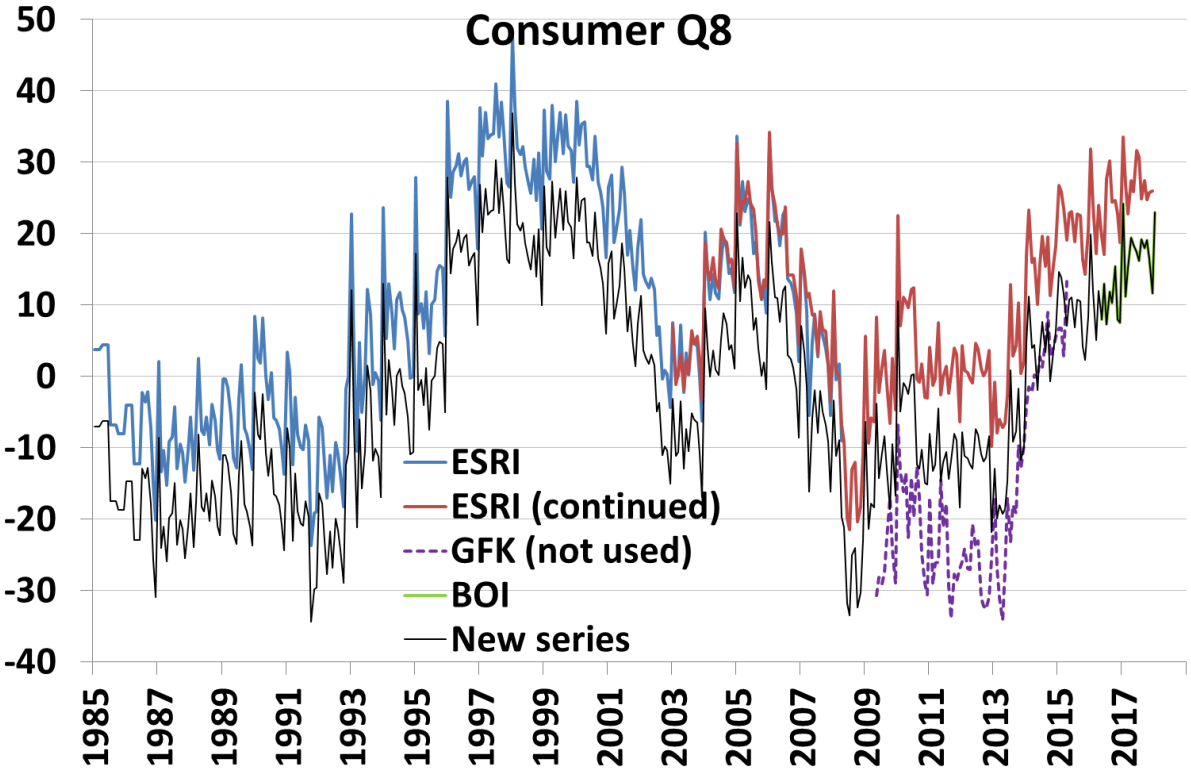

Consumer Q3
40
20
0
-20
-40
-60
-80
-100
-120
1985
1987
1989
1991
1993
1995
1997
1999
2001
2003
2005
2007
2009
2011
2013
2015
2017
ESRI
GFK (not used)
BOI
New series

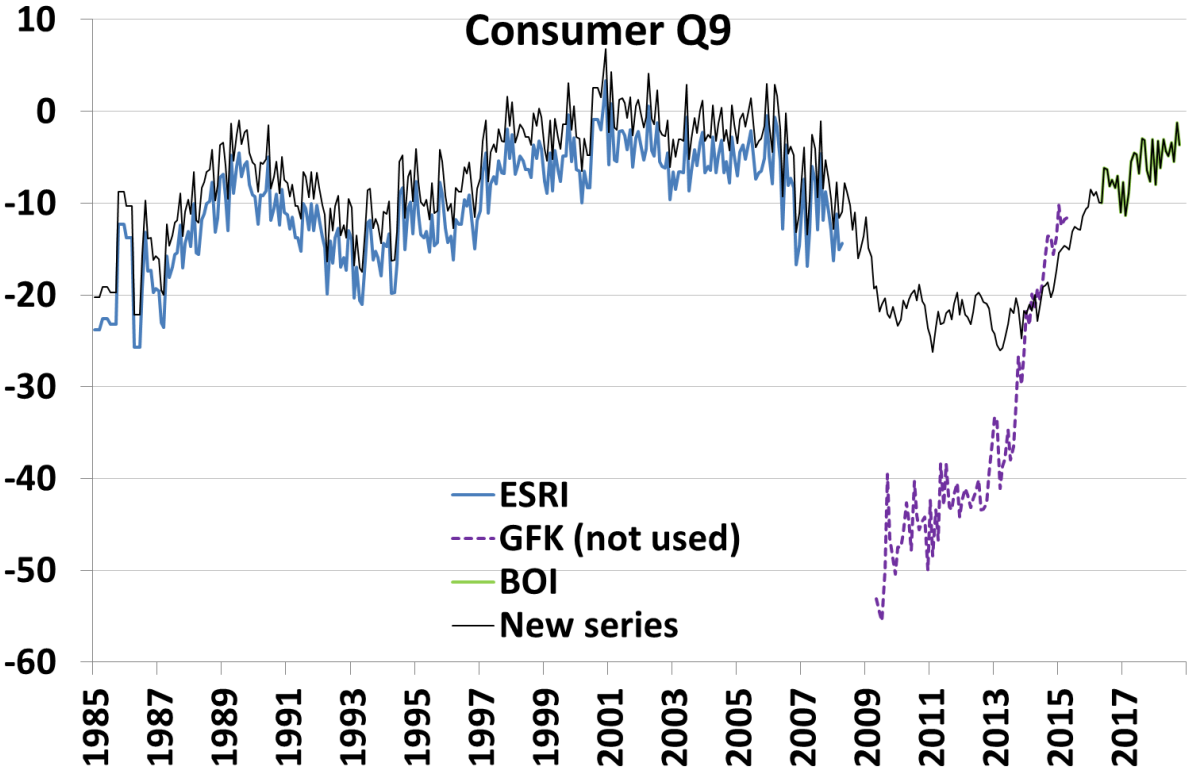

Consumer Q4
40
20
0
-20
-40
-60
ESRI
ESRI (continued)
GFK (not used)
BOI
New series
1985
1987
1989
1991
1993
1995
1997
1999
2001
2003
2005
2007
2009
2011
2013
2015
2017

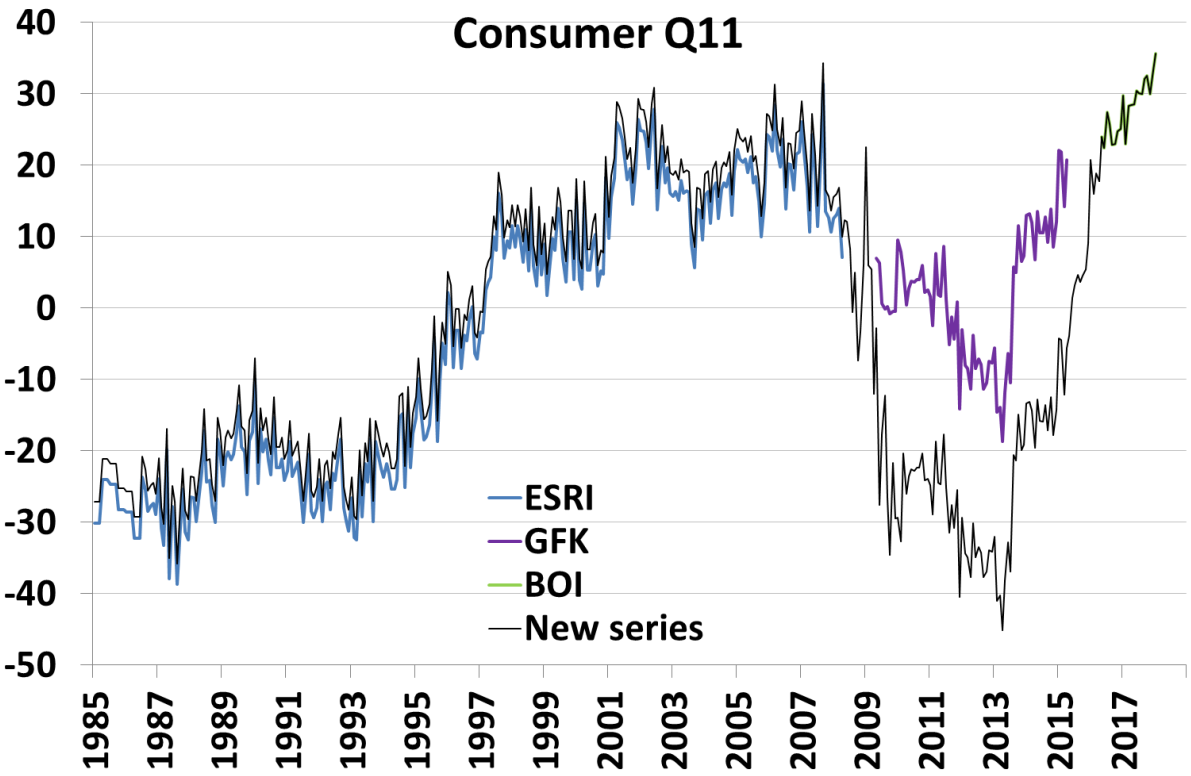

Consumer Q5
80
60
40
20
0
-20
-40
-60
ESRI
GFK (not used)
BOI
New series
1985
1987
1989
1991
1993
1995
1997
1999
2001
2003
2005
2007
2009
2011
2013
2015
2017

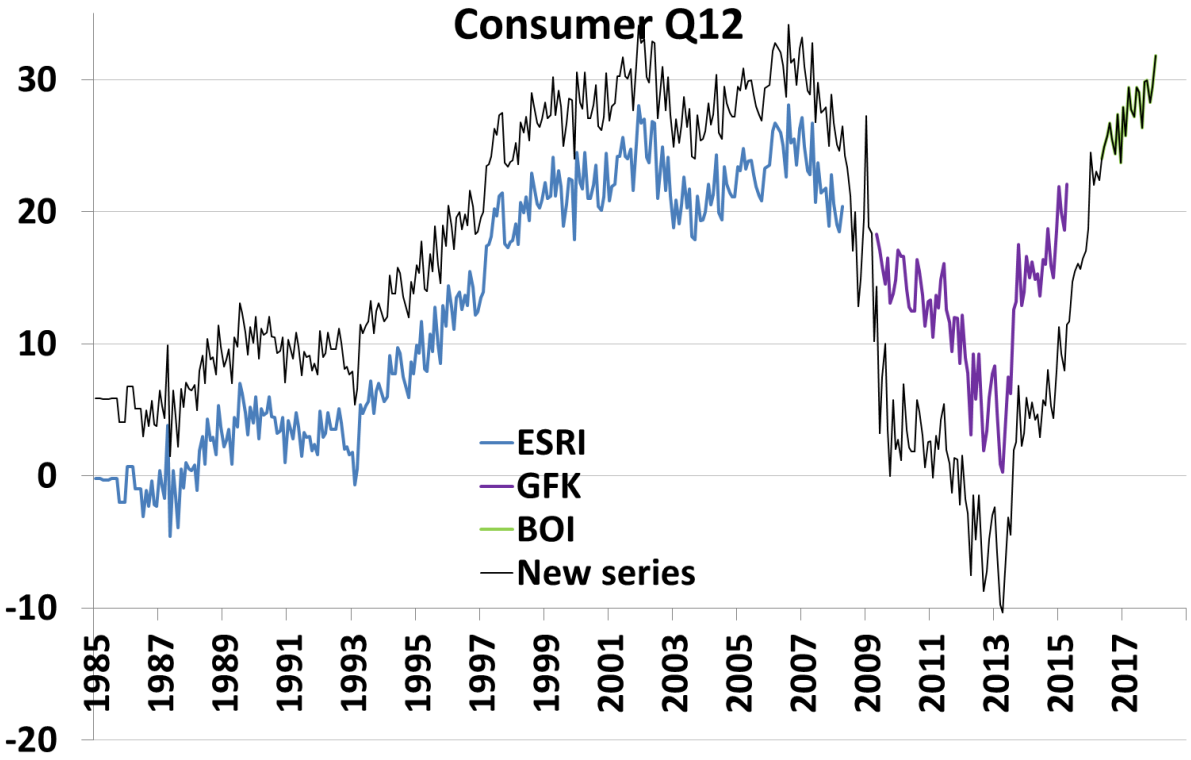

Consumer Q6
60
40
20
0
-20
-40
ESRI
GFK (not used)
BOI
New series
1985
1987
1989
1991
1993
1995
1997
1999
2001
2003
2005
2007
2009
2011
2013
2015
2017

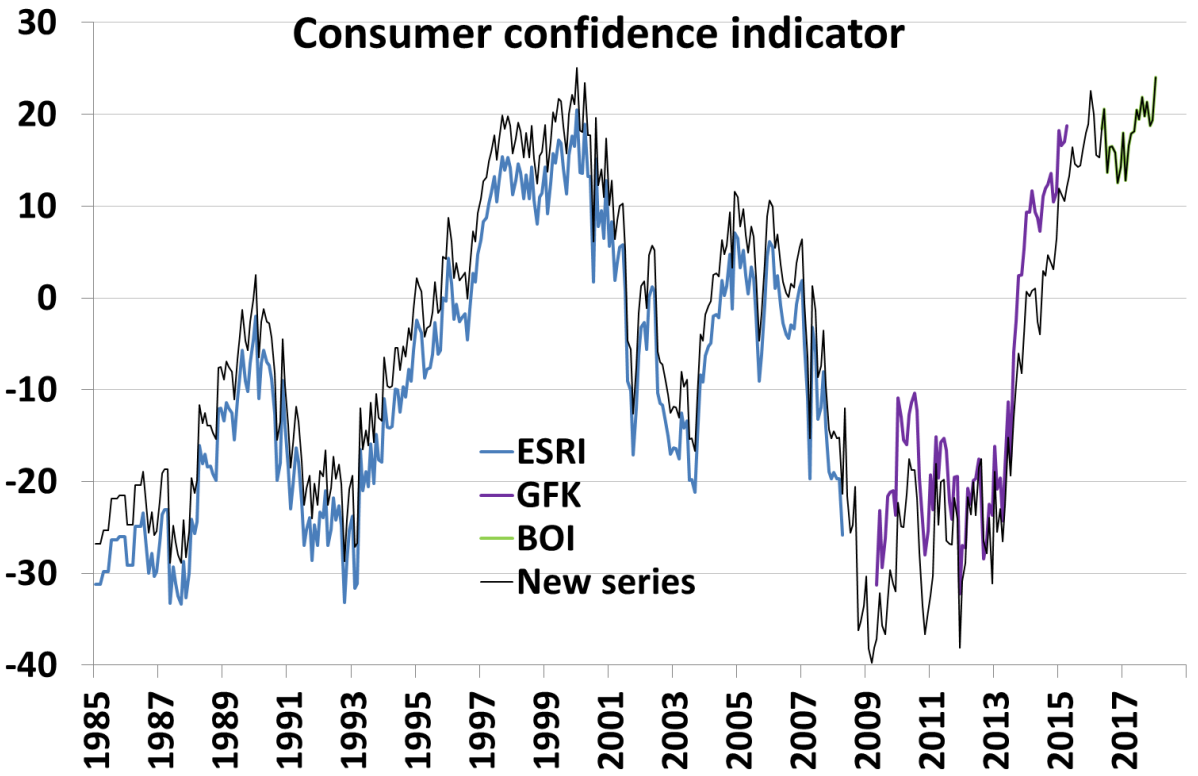

Consumer Q7
80
60
40
20
0
-20
-40
ESRI
ESRI (continued)
GFK (not used)
BOI
New series
1985
1987
1989
1991
1993
1995
1997
1999
2001
2003
2005
2007
2009
2011
2013
2015
2017

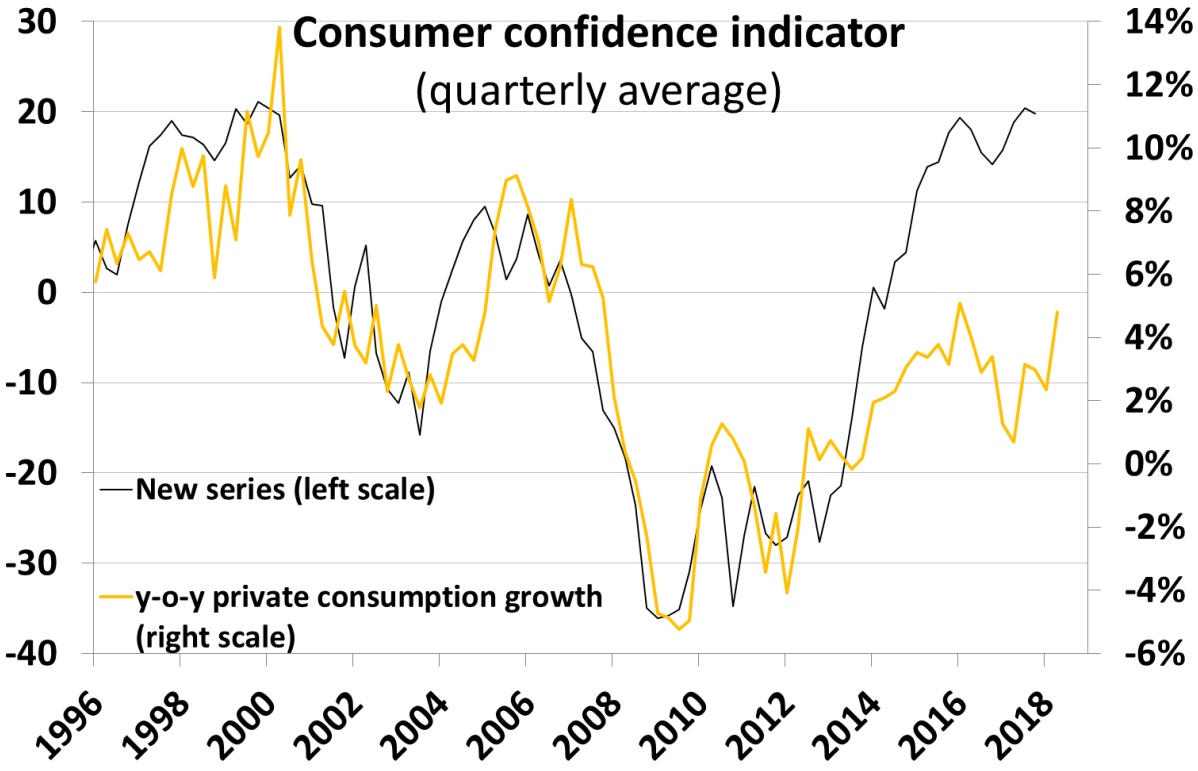

Consumer Q8
50
40
30
20
10
0
-10
-20
-30
-40
ESRI
ESRI (continued)
GFK (not used)
BOI
New series
1985
1987
1989
1991
1993
1995
1997
1999
2001
2003
2005
2007
2009
2011
2013
2015
2017


Consumer Q9
10
0
-10
-20
-30
-40
-50
-60
ESRI
GFK (not used)
BOI
New series
1985
1987
1989
1991
1993
1995
1997
1999
2001
2003
2005
2007
2009
2011
2013
2015
2017


Consumer Q11
40
30
20
10
0
-10
-20
-30
-40
-50
1985
1987
1989
1991
1993
1995
1997
1999
2001
2003
2005
2007
2009
2011
2013
2015
2017
ESRI
GFK
BOI
New series

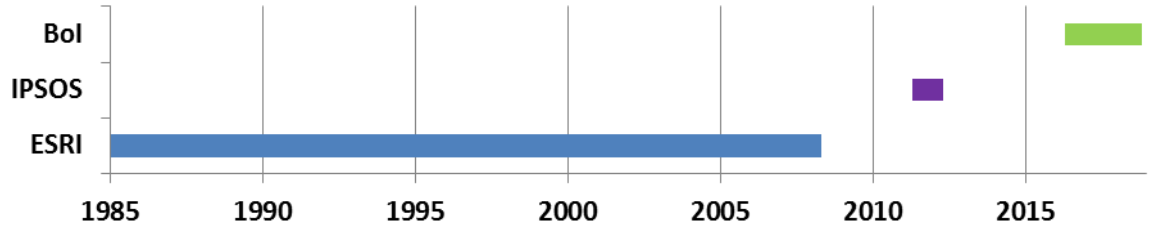

Consumer Q12
30
20
10
0
-10
-20
1985
1987
1989
1991
1993
1995
1997
1999
2001
2003
2005
2007
2009
2011
2013
2015
2017
ESRI
GFK
BOI
New series


Consumer confidence indicator
30
20
10
0
-10
-20
-30
-40
ESRI
GFK
BOI
New series
1985
1987
1989
1991
1993
1995
1997
1999
2001
2003
2005
2007
2009
2011
2013
2015
2017


Consumer confidence indicator
(quarterly average)
30
20
10
0
-10
-20
-30
-40
14%
12%
10%
8%
6%
4%
2%
0%
-2%
-4%
-6%
New series (left scale)
y-o-y private consumption growth
(right scale)
1996
1998
2000
2002
2004
2006
2008
2010
2012
2014
2016
2018

Business surveys

For all business surveys, three BCS data sets are available: ESRI (up to April 2008), IPSOS (from May/June 2011 to April 2012), and BoI (since May 2016). The general idea for business surveys is similar to the consumer survey approach: first ESRI series are extended based on available information from other sources and then the series are adjusted in level. In a last step, BoI data series are used (without any adjustment or modification) from May 2016.



In the first step, the gaps are filled with reconstructed series. As for some of the consumer survey series, the main idea is to use Partial least squares regressions (PLS) in order to reconstruct missing data. First a model is estimated with PLS, with the respective survey question as dependent variable and an explanatory dataset that is tailored to the target variable, over a historical sample including all available data up to April 2008. Then, following the approach described in Chow and Lin (1971),[6] the model is applied to the explanatory data from May 2008 onwards, to simulate the out-of-sample fitted values that are used in the next step. In cases where the estimation sample is too short (because relevant explanatory variables get available rather late - industrial producer prices, for instance, are only available since 2005), or where quarterly series are included in the dataset (which appears to deteriorate the PLS estimation), the PLS-approach had to be discarded. Instead, missing data are generated as the simple average of conceptually close series. These series are rescaled so that their averages and standard deviations match those of the survey question over the historical sample up to 2008.

Once a consistent series is reconstructed (either as the out-of-sample fit of a PLS model or as the average of similar series), it is slightly shifted to make sure that its average between May 2007 and April 2008 matches exactly that of ESRI's data in that period, in order to ensure a smooth transition between the two series.

In a second step, IPSOS data are included for the period 2011/12, provided they do not display excessive volatility.[7] After extending the ESRI series, both IPSOS series and the reconstructed series are aligned so that the average of the common year of the two series (between May 2011 and April 2012) matches. Then, values in the reconstructed series are replaced with all available IPSOS values.

In a third step, the reconstructed series are adjusted in level to ensure a smooth transition to the BoI data from May 2016 onwards. The necessary level shift is computed as the average difference between May 2016 and April 2017 of the BoI series and the reconstructed series. In all cases, it is the reconstructed series which are shifted to the level of the BoI series, rather than the opposite, since

[6] Chow, G., & Lin, A. (1971). Best Linear Unbiased Interpolation, Distribution, and Extrapolation of Time Series by Related Series. The Review of Economics and Statistics, 53(4), 372-375.

[7] Practically, IPSOS series are included for reconstructing series in industry, services and construction, but not in retail trade.

this ensures that no further adjustments of the series will be needed when new data, collected by the BoI, become available.

The following sections detail the settings and data used in the outlined restoration process described as the first step, for each specific survey sector.

Industry

To fill the gaps between the three available survey data sets outlined above, five additional data sets provide useful information for the industry survey data: Series from Markit's PMI data set in the manufacturing sector, the industrial production indices (released by the Central Statistics Office and Eurostat), selected series from the reconstructed consumer survey data set as described previously, producer prices in industry (released by Eurostat) and the modified Gross National Income (GNI, released by the CSO).[8]

For questions 1 to 5, series are reconstructed between May 2008 and April 2017 based on a PLS model. For all these questions, the dataset includes all PMI manufacturing questions, the year-on-year changes of industrial production (in the traditional sector), and the reconstructed consumer questions 3 and 4 (see Table 2).

Table 2 - Dataset used with PLS

BCS question	Theme of the BCS question	Dataset used with PLS		
1	Past production	All PMI manufacturing questions	y-o-y IP (traditional sector[9])	Consumer questions 3 and 4
2	Overall order books			
3	Export order books			
4	Stock of finished products			
5	Production expectations			

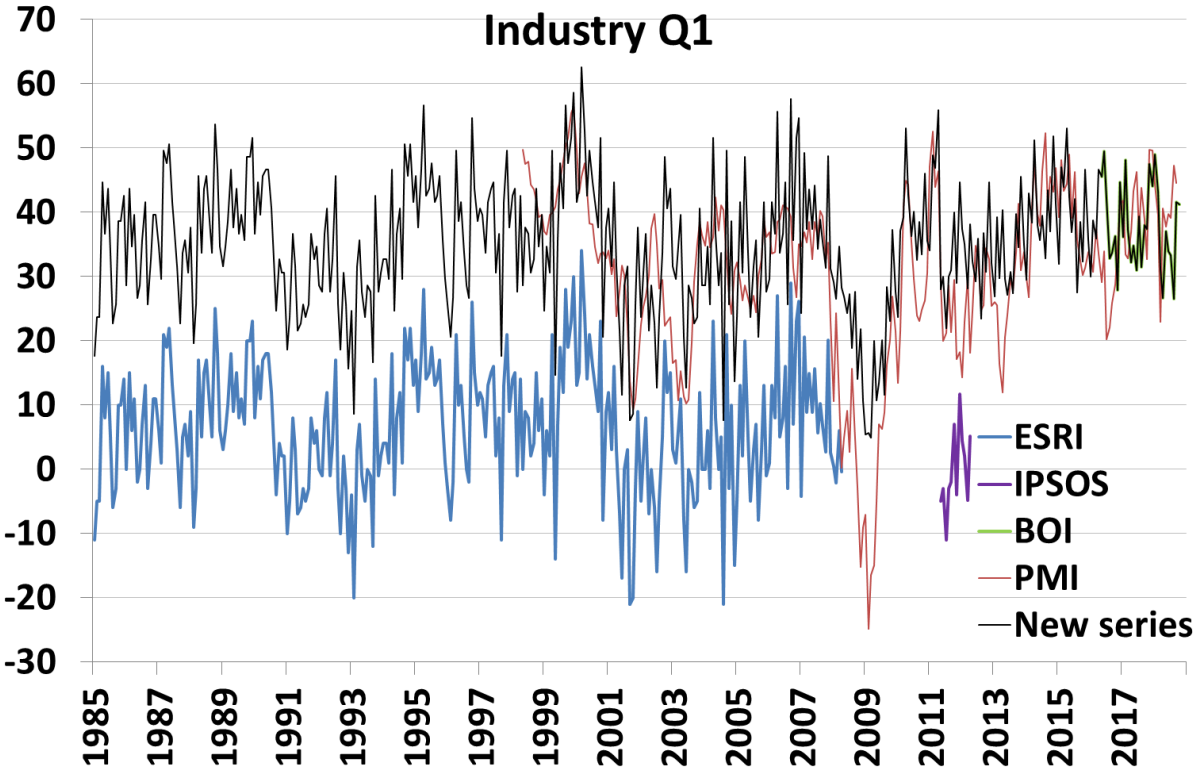
For questions 6, 7 and 13, the reconstructed series are based on a simple arithmetic average (see Table 3), including the closest PMI question and, respectively, industrial producer prices, (interpolated) employment in industry, or the cycle component (extracted with a HP filter) of the modified GNI.

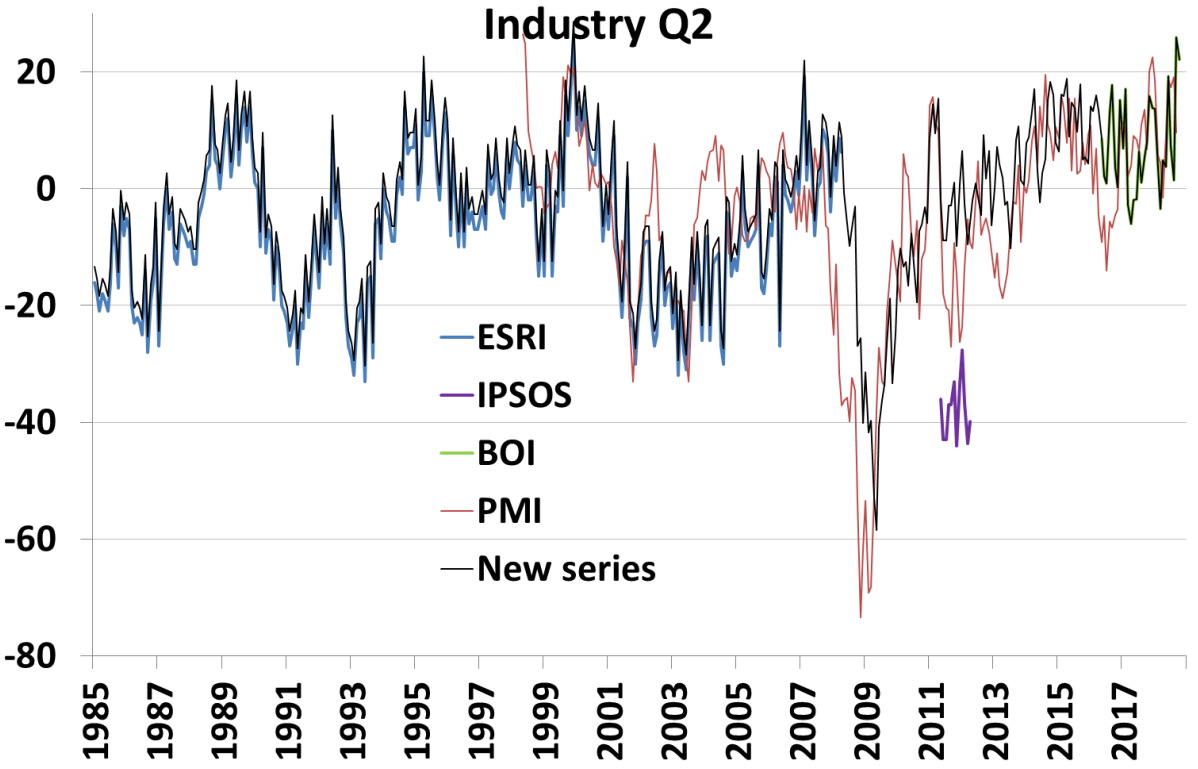
[8] The modified GNI is designed to exclude globalisation effects that are disproportionally impacting the measurement of the size of the Irish economy, see https://www.cso.ie/en/releasesandpublications/ep/p-nie/nie2017/mgni/

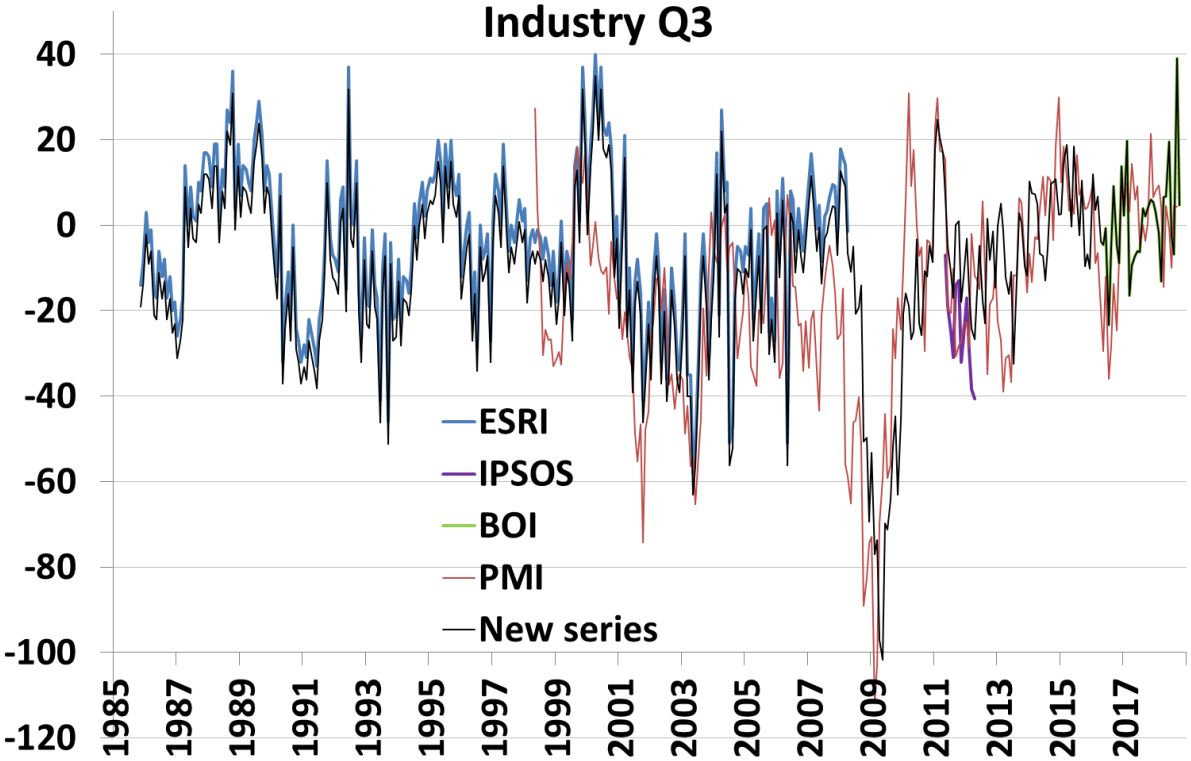
[9] For the period 2012m5-2016m4, industrial production was restricted to the traditional sector. Indeed, due to large multinational corporations having relocated their economic activities, and more specifically their underlying intellectual property, to Ireland, industrial production boomed at the time, with year-on-year growth rates close to +60%. As this boom was not linked to an increased economic activity in every corporation, but caused by the inclusion of new corporations to the sample, this should arguably not be reflected by the surveys. As the traditional sector does not encompass the large multinational corporations, industrial production in this sector doesn't present disproportionally booming growth rates.

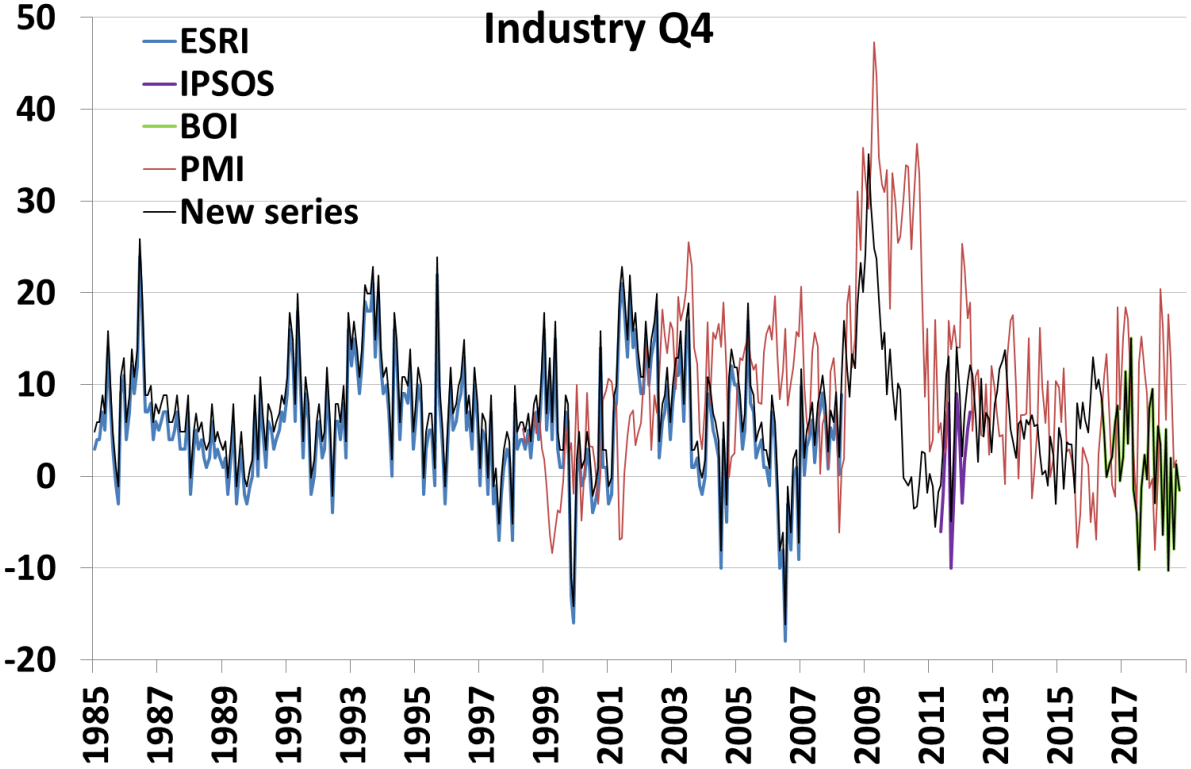
Table 3 - Series included in the average

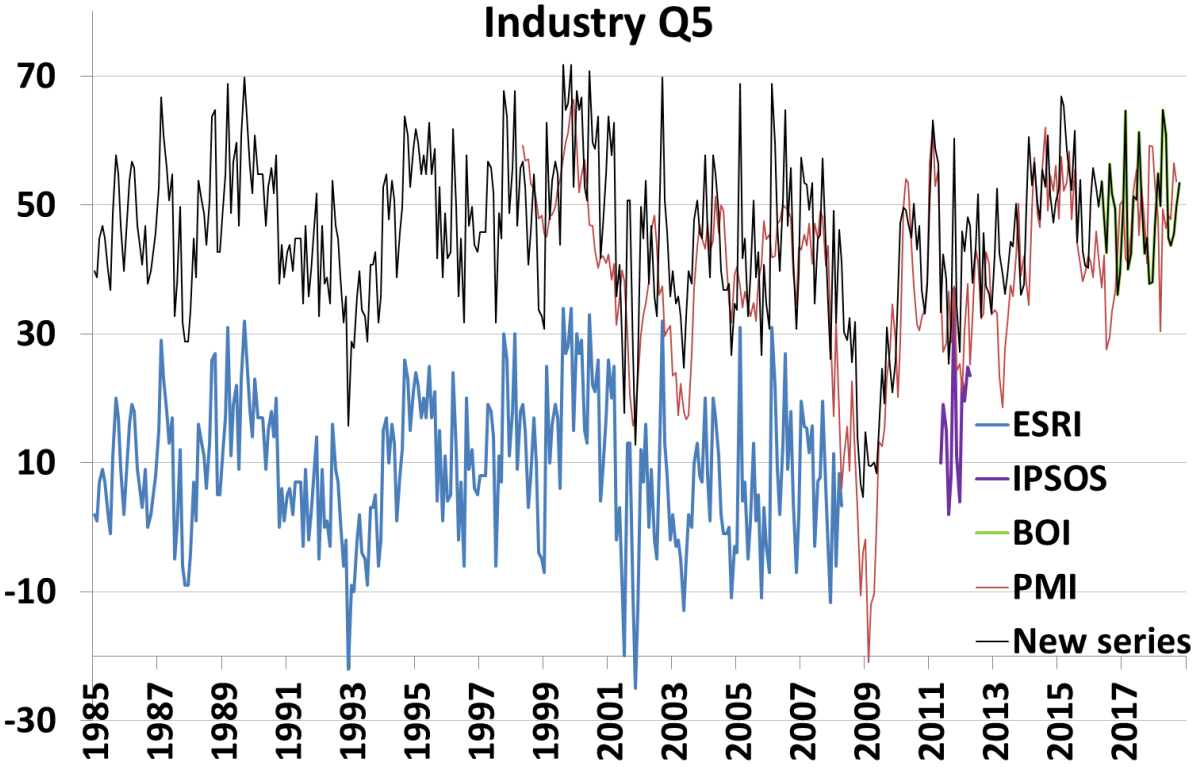
BCS question	Theme of the BCS question	Series included in the average
6	Prices expectations	Manufacturing PMI Input Prices Index + y-o-y industrial producer prices
7	Employment expectations	Manufacturing PMI Employment Index + y-o-y employment in industry
13	Capacity utilisation (quarterly)	Manufacturing PMI Capacity Utilisation Index + cycle component (HP filter) of the modified GNI

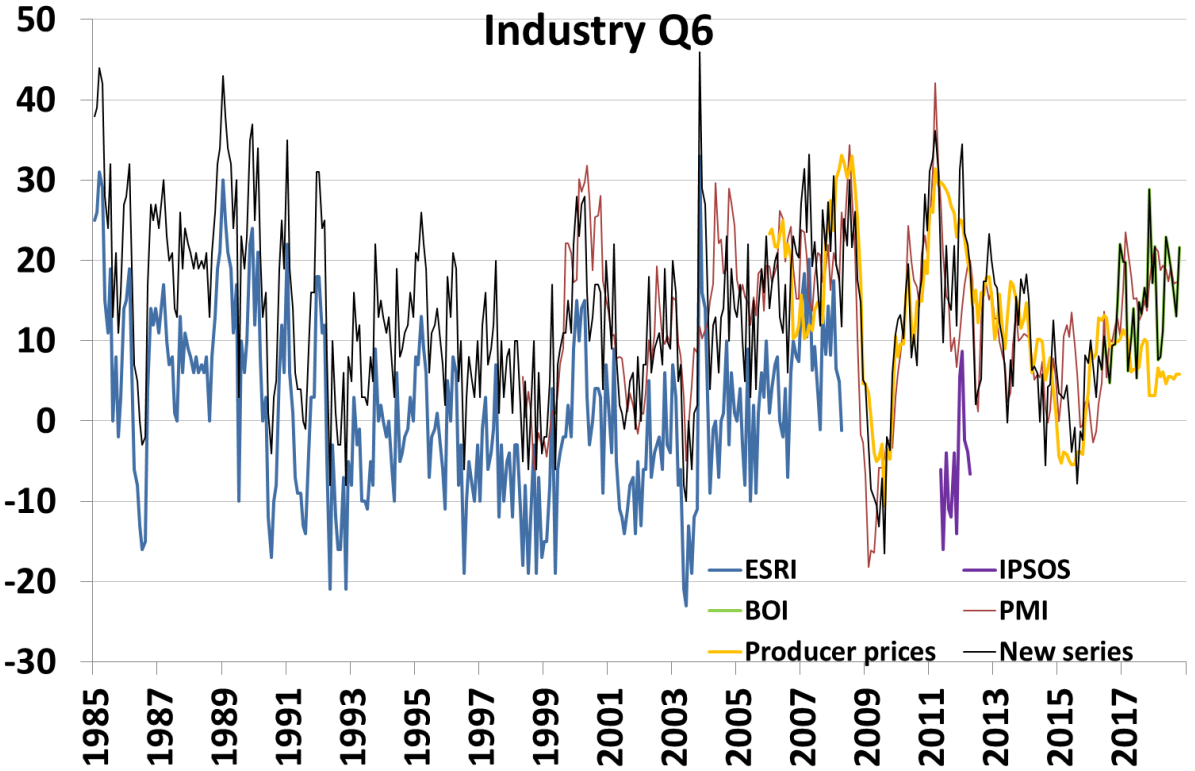

Industry Q1
70
60
50
40
30
20
10
0
-10
-20
-30
1985
1987
1989
1991
1993
1995
1997
1999
2001
2003
2005
2007
2009
2011
2013
2015
2017
ESRI
IPSOS
BOI
PMI
New series

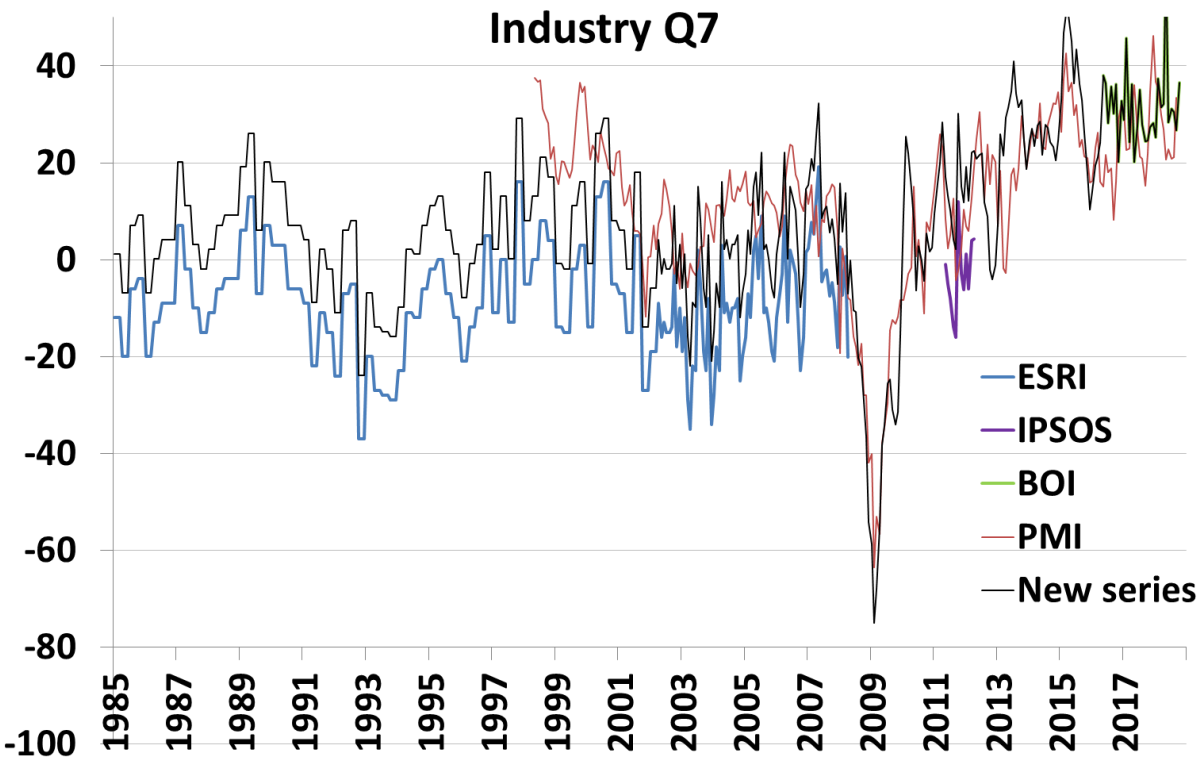

Industry Q2
20
0
-20
-40
-60
-80
1985
1987
1989
1991
1993
1995
1997
1999
2001
2003
2005
2007
2009
2011
2013
2015
2017
ESRI
IPSOS
BOI
PMI
New series

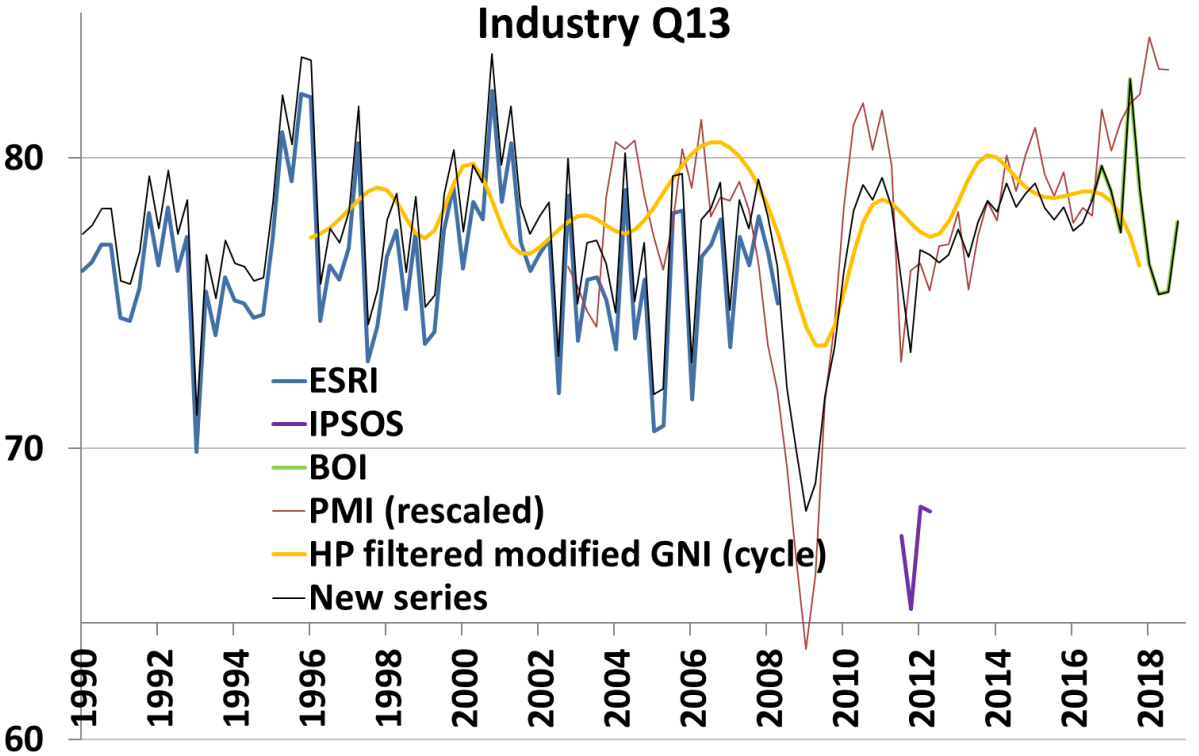

Industry Q3
40
20
0
-20
-40
-60
-80
-100
-120
1985
1987
1989
1991
1993
1995
1997
1999
2001
2003
2005
2007
2009
2011
2013
2015
2017
ESRI
IPSOS
BOI
PMI
New series

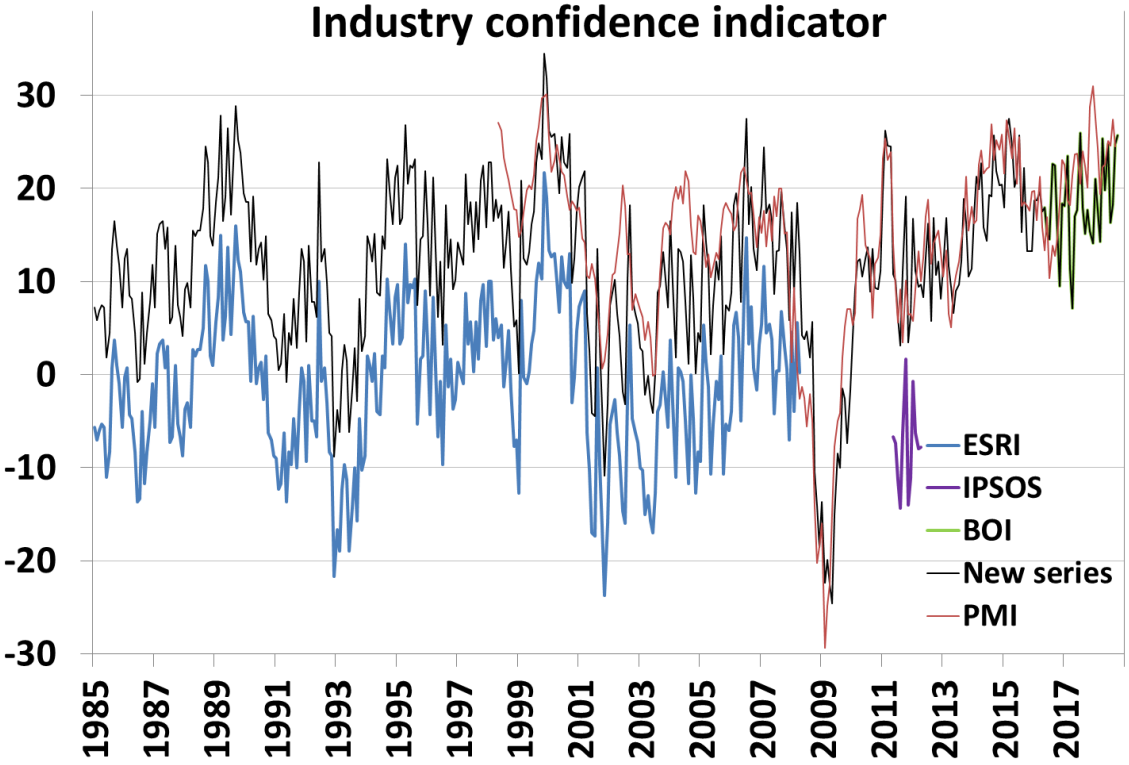

Industry Q4
50
40
30
20
10
0
-10
-20
1985
1987
1989
1991
1993
1995
1997
1999
2001
2003
2005
2007
2009
2011
2013
2015
2017
ESRI
IPSOS
BOI
PMI
New series

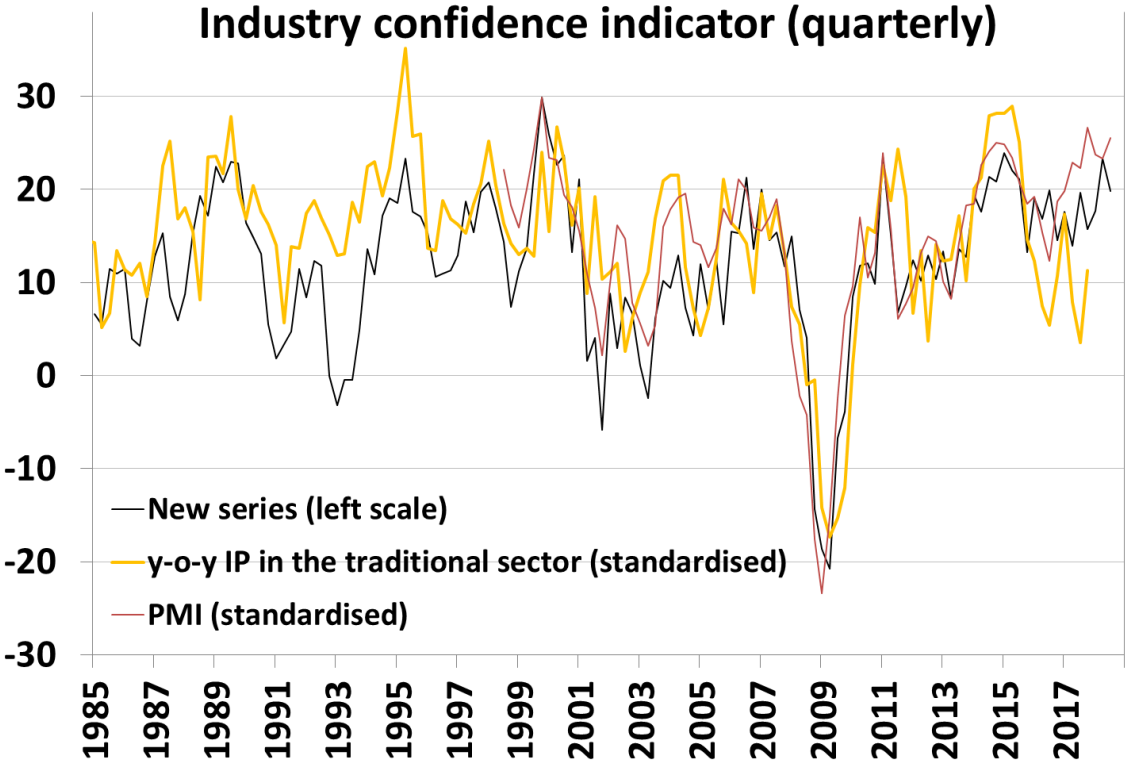

Industry Q5
70
50
30
10
-10
-30
1985
1987
1989
1991
1993
1995
1997
1999
2001
2003
2005
2007
2009
2011
2013
2015
2017
ESRI
IPSOS
BOI
PMI
New series


Industry Q6
50
40
30
20
10
0
-10
-20
-30
1985
1987
1989
1991
1993
1995
1997
1999
2001
2003
2005
2007
2009
2011
2013
2015
2017
ESRI
IPSOS
BOI
PMI
Producer prices
New series


Industry Q7
40
20
0
-20
-40
-60
-80
-100
1985
1987
1989
1991
1993
1995
1997
1999
2001
2003
2005
2007
2009
2011
2013
2015
2017
ESRI
IPSOS
BOI
PMI
New series


Industry Q13
80
70
60
1990
1992
1994
1996
1998
2000
2002
2004
2006
2008
2010
2012
2014
2016
2018
ESRI
IPSOS
BOI
PMI (rescaled)
HP filtered modified GNI (cycle)
New series


Industry confidence indicator
30
20
10
0
-10
-20
-30
1985
1987
1989
1991
1993
1995
1997
1999
2001
2003
2005
2007
2009
2011
2013
2015
2017
ESRI
IPSOS
BOI
New series
PMI


Industry confidence indicator (quarterly)
30
20
10
0
-10
-20
-30
1985
1987
1989
1991
1993
1995
1997
1999
2001
2003
2005
2007
2009
2011
2013
2015
2017
New series (left scale)
y-o-y IP in the traditional sector (standardised)
PMI (standardised)

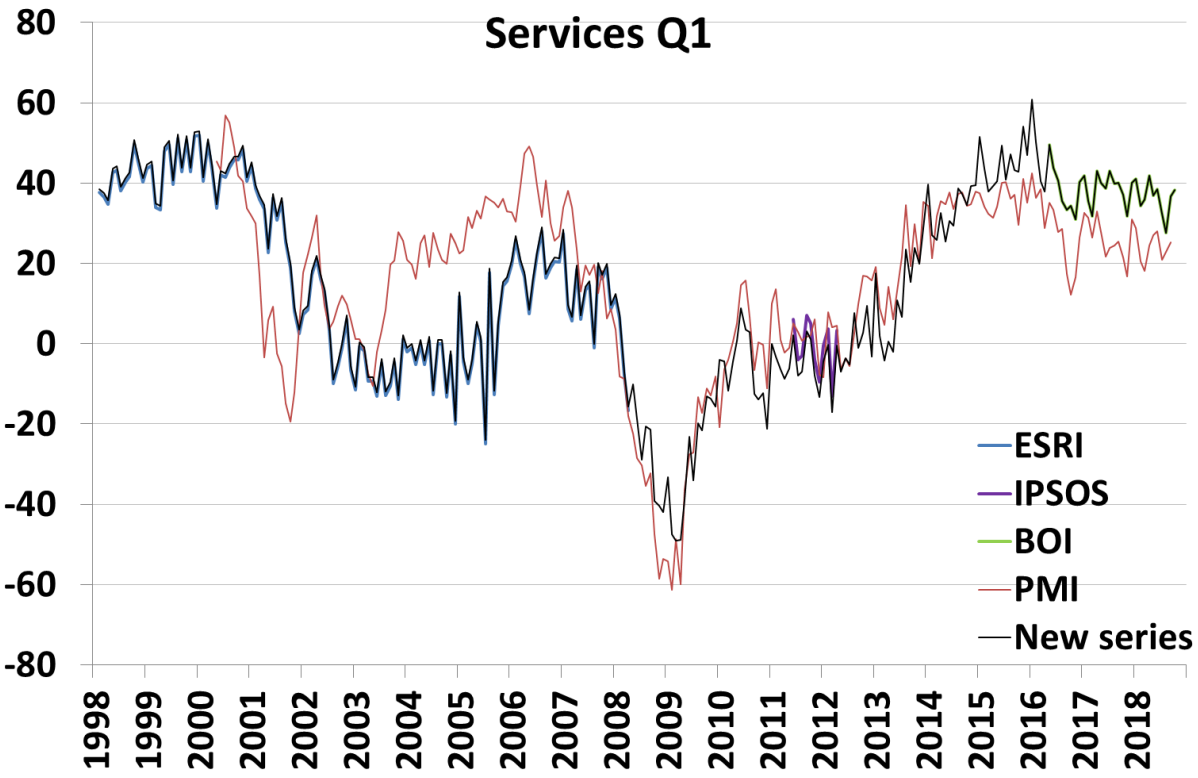
Services

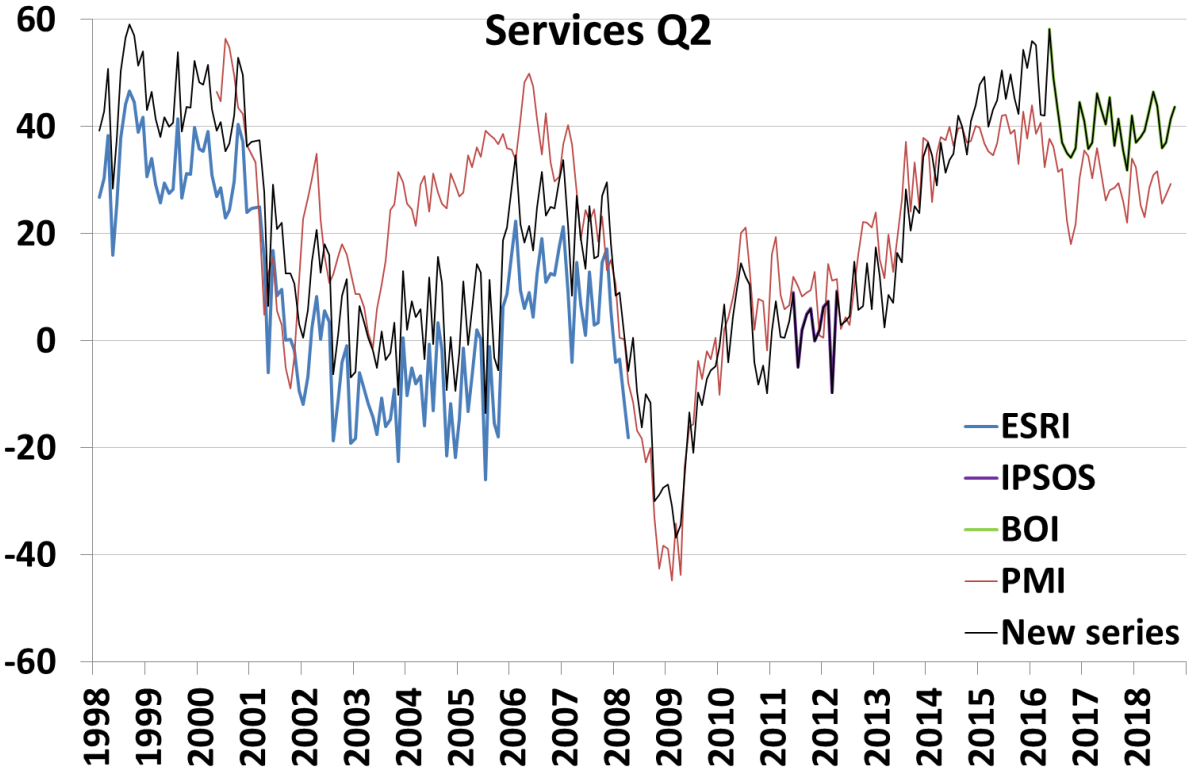
To fill the gaps between the three BCS data sets (from ESRI, IPSOS and BoI), four additional data sets provide useful information for the services survey data: Series from Markit's PMI data set in the services sector, employment in the services sector, except retail trade and repair (released by Eurostat), producer prices in services (released by Eurostat), and selected series from the reconstructed consumer survey data set as described previously.

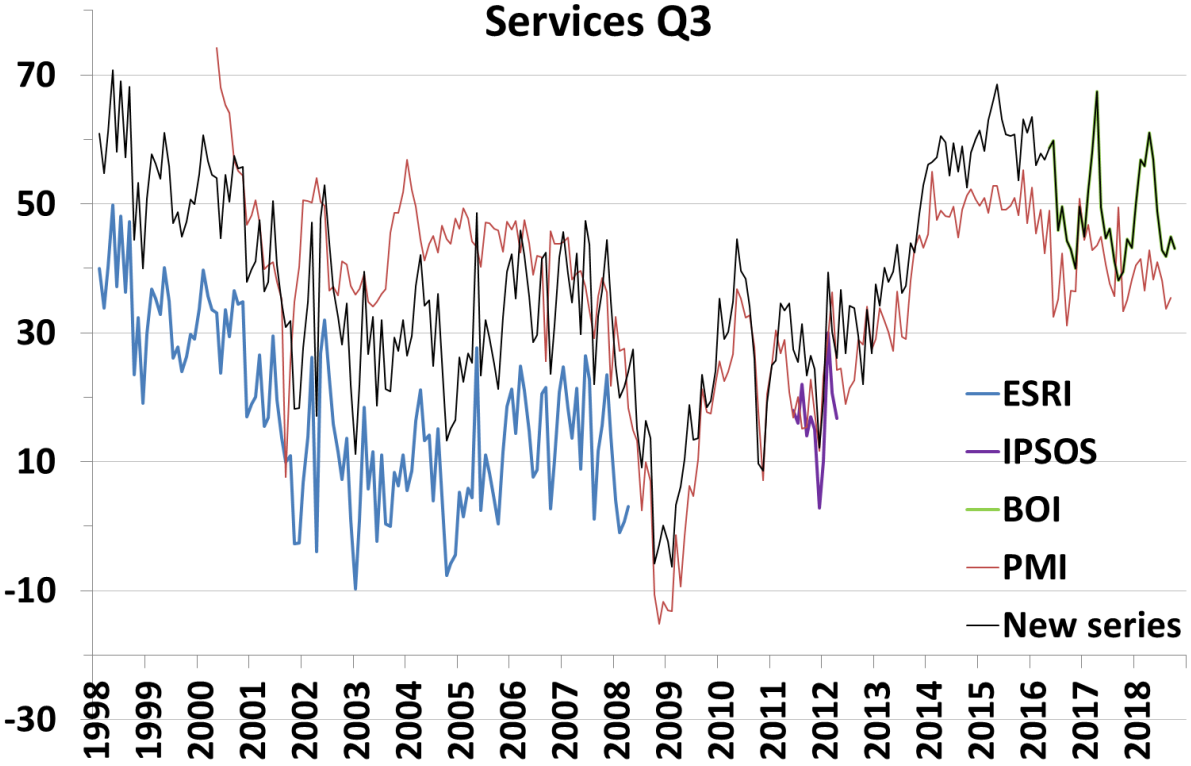
For all questions, the reconstructed series are based on a simple arithmetic average (see Table 4). For questions 1 and 2, the reconstructed series is based on the average of the closest PMI question and the reconstructed consumer question 3. For question 3 (demand expectations), the reconstructed series is based on the average of the Services PMI Future Business Expectations Index and the reconstructed consumer question 4 (future general economic situation). For questions 4, 5 and 6, the average includes the closest PMI question and respectively year-on-year changes in employment in services (for the questions about employment) and year-on-year changes in producer prices in services (for question 6).

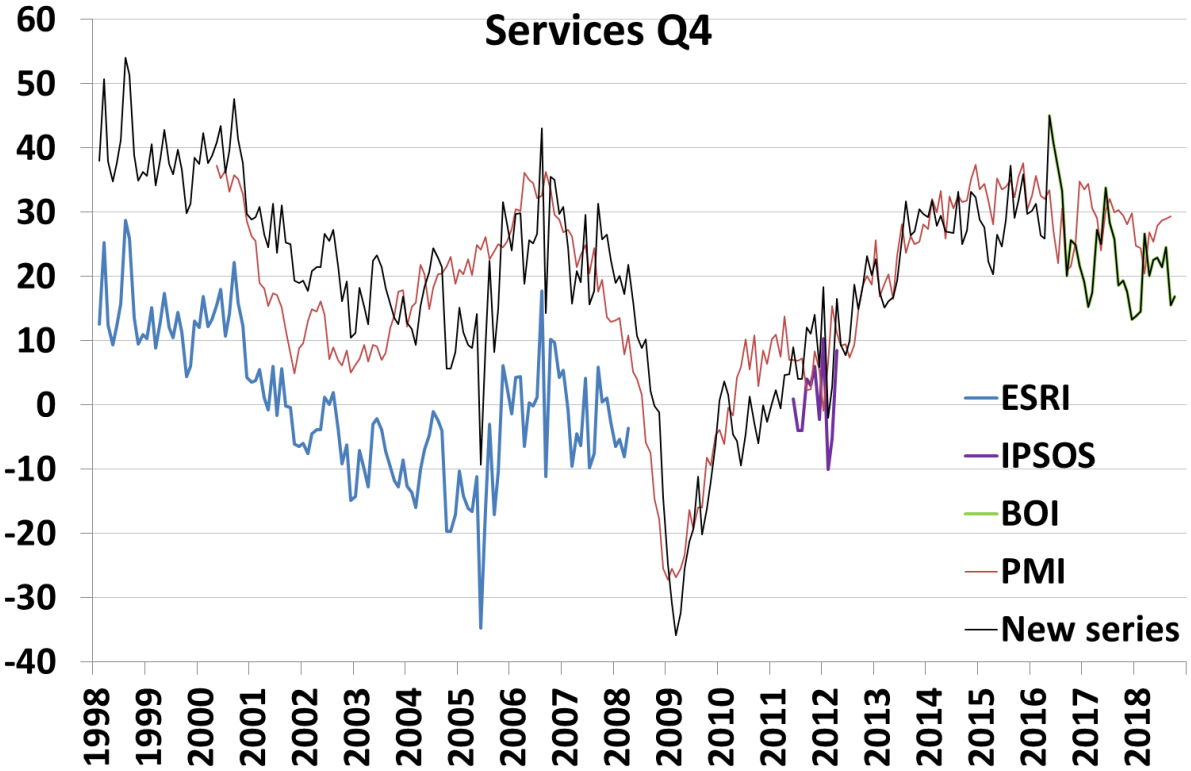
Table 4 - Series included in the average

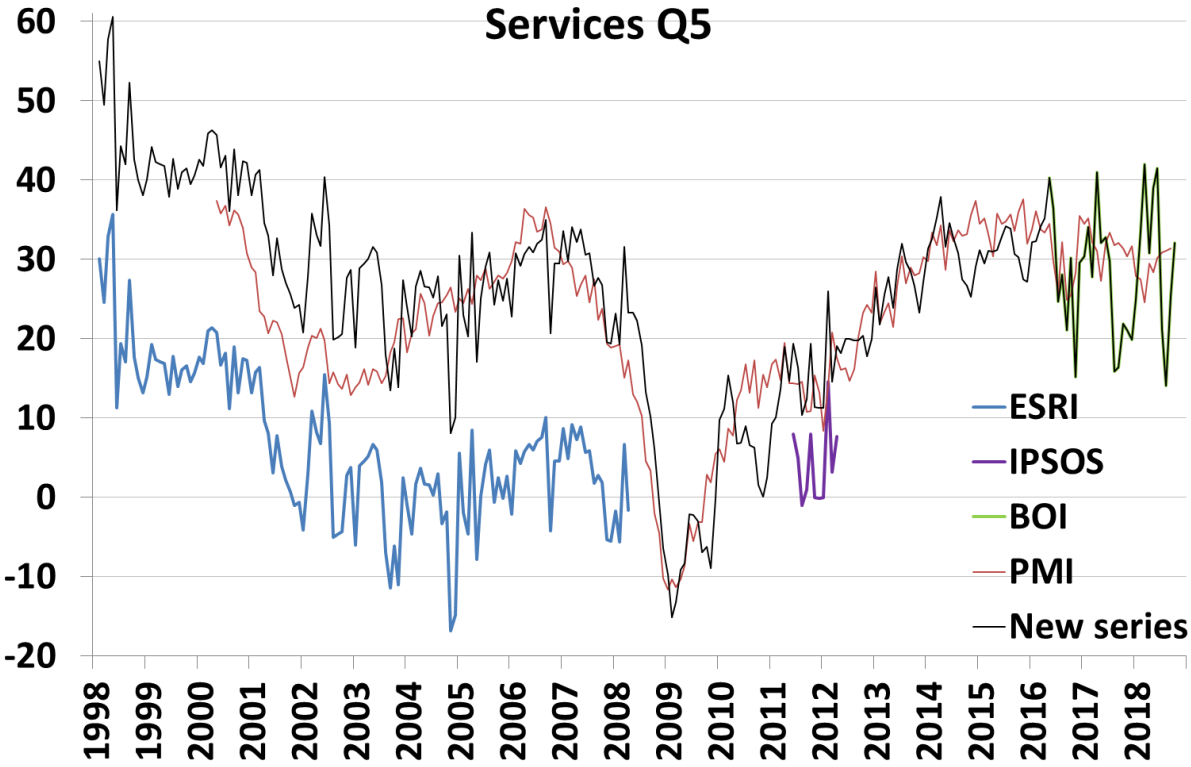
BCS question	Theme of the BCS question	Series included in the average
1	Past business situation	Services PMI Business Activity Index + Consumer question 3
2	Past demand	Services PMI Business Activity Index + Consumer question 3
3	Demand expectations	Services PMI Future Business Expectations Index + Consumer question 4
4	Past employment	Services PMI Employment Index + y-o-y employment in services
5	Employment expectations	Services PMI Employment Index + y-o-y employment in services
6	Prices expectations	Services PMI Output Prices Index + y-o-y producer prices in services

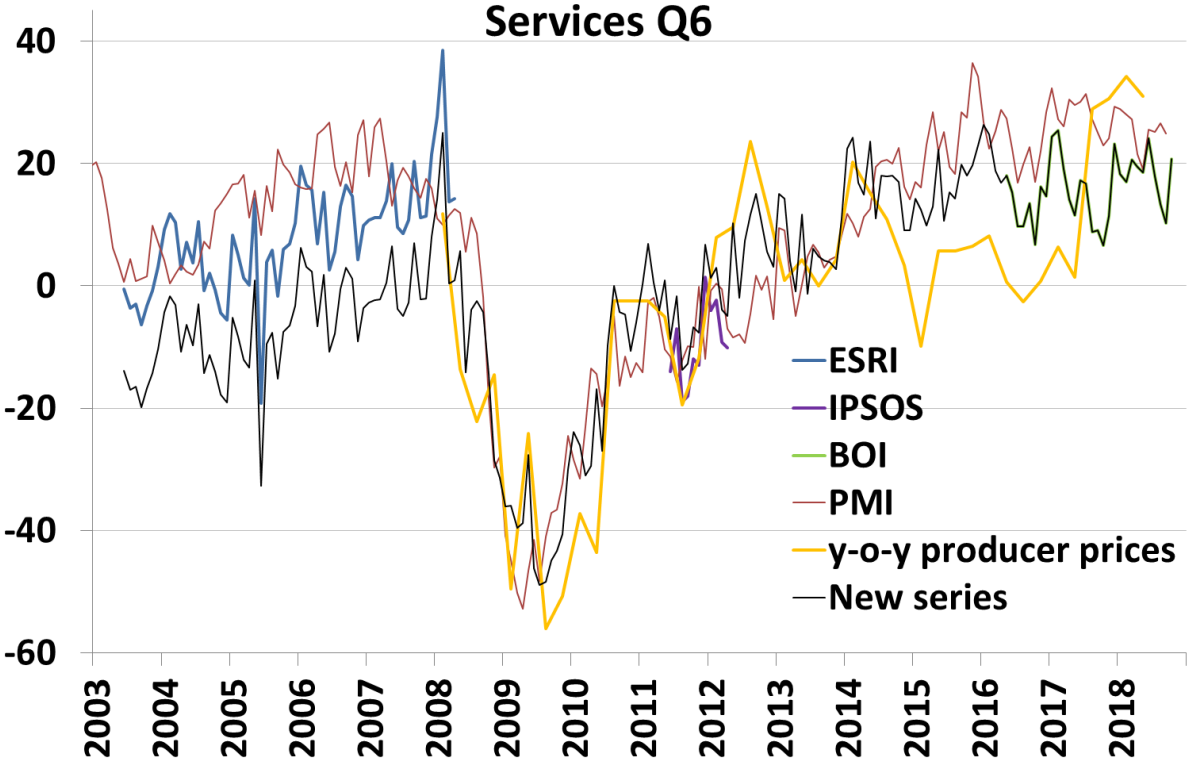

Services Q1
80
60
40
20
0
-20
-40
-60
-80
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
ESRI
IPSOS
BOI
PMI
New series

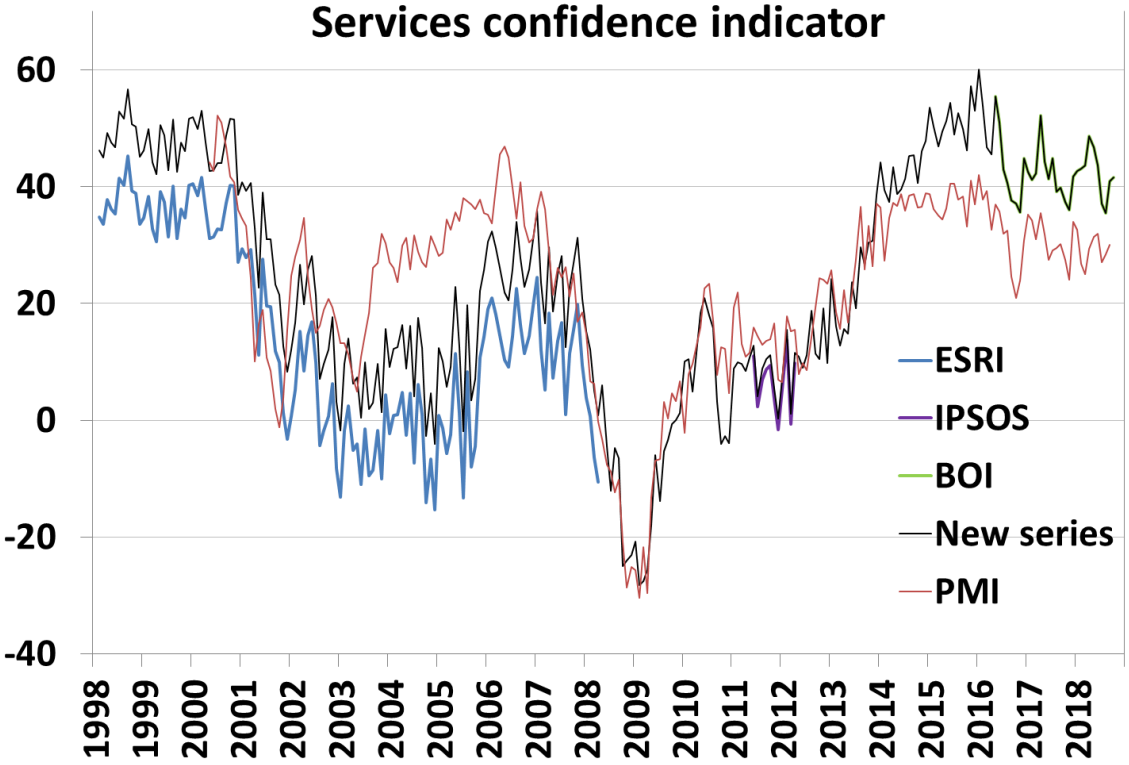

Services Q2
60
40
20
0
-20
-40
-60
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
ESRI
IPSOS
BOI
PMI
New series

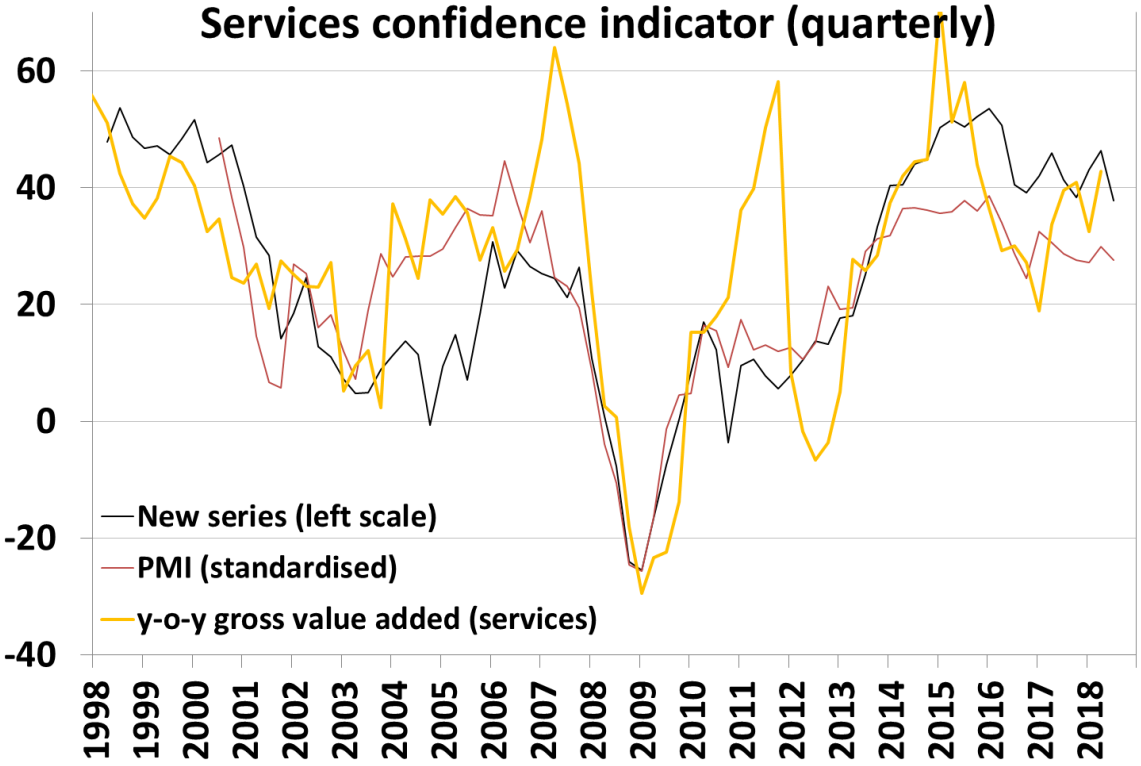

Services Q3
70
50
30
10
-10
-30
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
ESRI
IPSOS
BOI
PMI
New series


Services Q4
60
50
40
30
20
10
0
-10
-20
-30
-40
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
ESRI
IPSOS
BOI
PMI
New series


Services Q5
60
50
40
30
20
10
0
-10
-20
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
ESRI
IPSOS
BOI
PMI
New series


Services Q6
40
20
0
-20
-40
-60
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
ESRI
IPSOS
BOI
PMI
y-o-y producer prices
New series


Services confidence indicator
60
40
20
0
-20
-40
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
ESRI
IPSOS
BOI
New series
PMI


Services confidence indicator (quarterly)
60
40
20
0
-20
-40
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
New series (left scale)
PMI (standardised)
y-o-y gross value added (services)

Retail trade

In the retail trade sector, IPSOS data was discarded due to a too high volatility compared to the other series. To fill the gaps between the other two BCS data sets - ESRI (from November 1997 to April 2008) and Bank of Ireland (from May 2016) – a number of additional data sets provide useful information for the retail trade survey data: Series from Markit's PMI data set, the volume of sales and employment in retail trade (released by Eurostat), inflation for goods (computed as the year-on-year growth rate of HICP for goods), and selected series from the reconstructed industry and consumer survey data sets as described previously.

For questions 1, 3, 4 and 5, series are reconstructed between May 2008 and April 2017 based on a PLS model. For questions 1, 3 and 4, the dataset includes all PMI services questions, the reconstructed consumer questions 3 and 4, and year-on-year changes in the volume of sales in retail trade (see Table 5). For question 5, the dataset includes all PMI services questions and year-on-year changes in employment in retail trade.

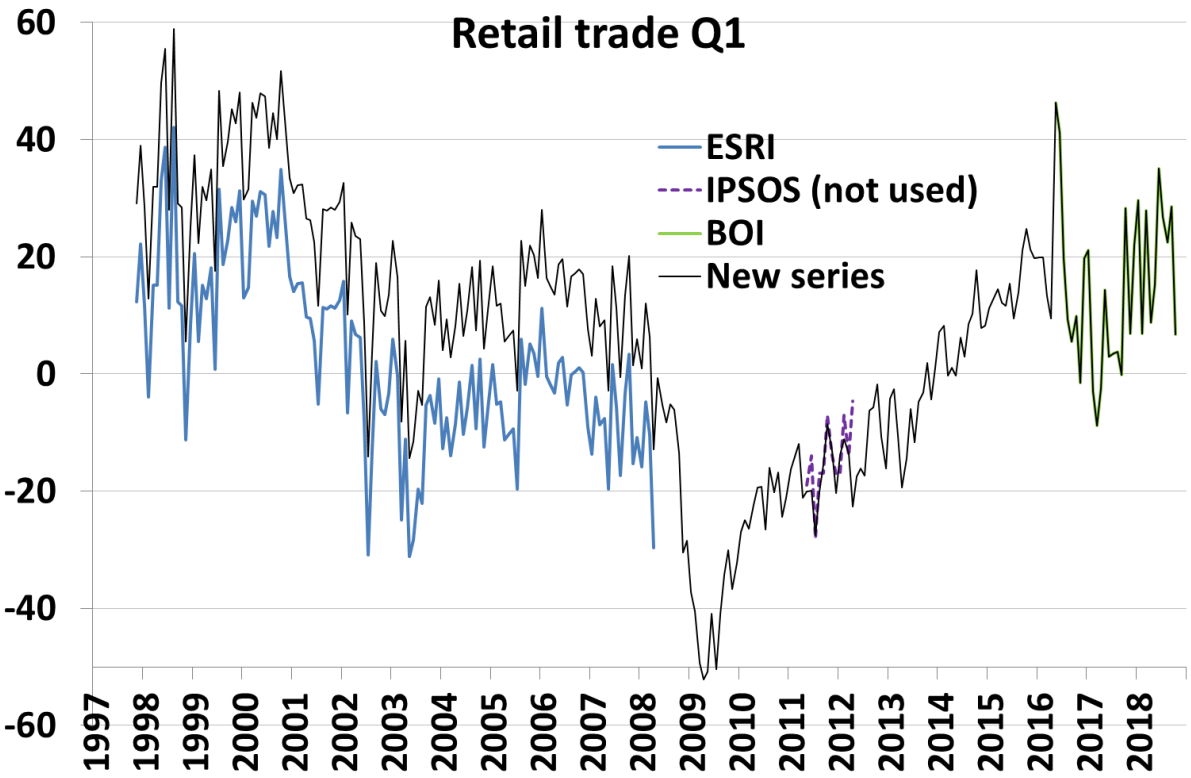
Table 5 - Dataset used with PLS

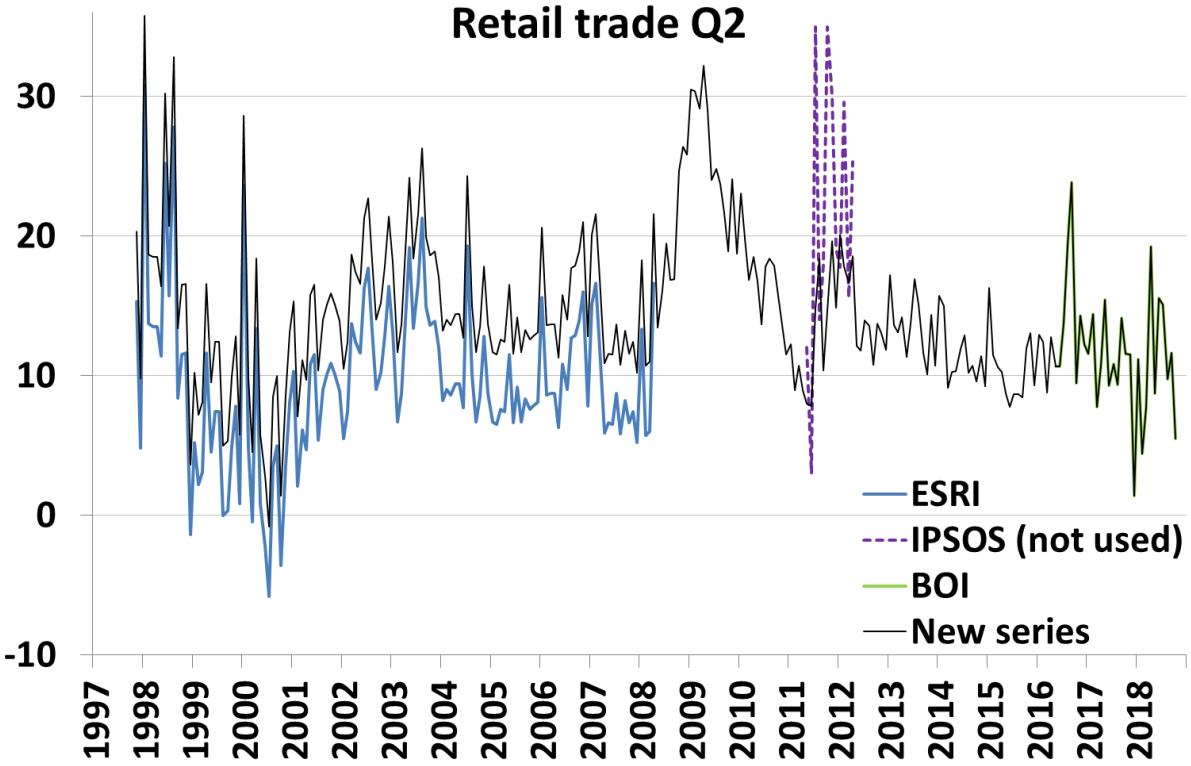
<table>
<tr><th>BCS question</th><th>Theme of the BCS question</th><th colspan="3">Dataset used with PLS</th></tr>
<tr><td>1</td><td>Past business activity</td><td rowspan="3">All PMI services questions</td><td rowspan="2">Consumer questions 3 and 4</td><td rowspan="2">y-o-y retail sales</td></tr>
<tr><td>3</td><td>Expectations on orders placed with suppliers</td></tr>
<tr><td>4</td><td>Expectations on business activity</td><td colspan="2">y-o-y employment</td></tr>
<tr><td>5</td><td>Employment expectations</td></tr>
</table>

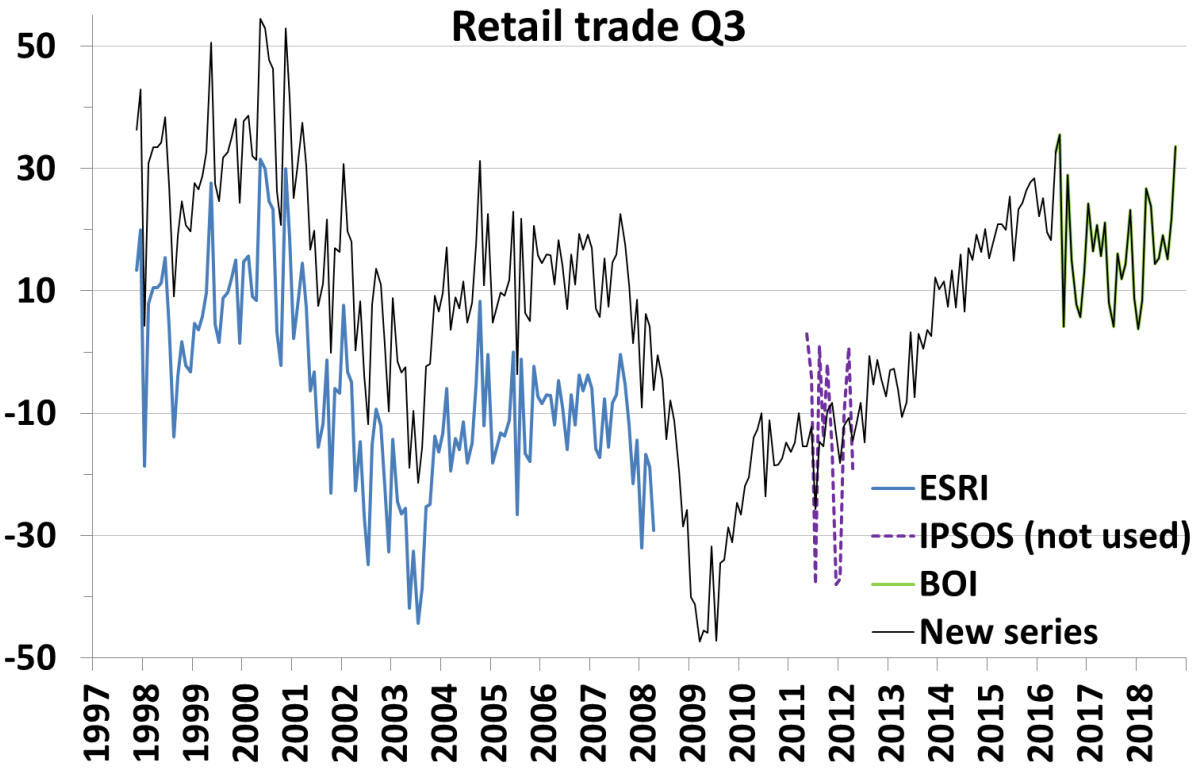
For questions 2 and 6, the reconstructed series are based on a simple arithmetic average (see Table 6). For question 2 (current stock), the series is computed as the average of the Manufacturing PMI question about Stocks of Finished Goods and the reconstructed question 4 (stocks) from the industry survey. For question 6, the series is the average of the Services PMI question about prices and Inflation (HICP) for goods.

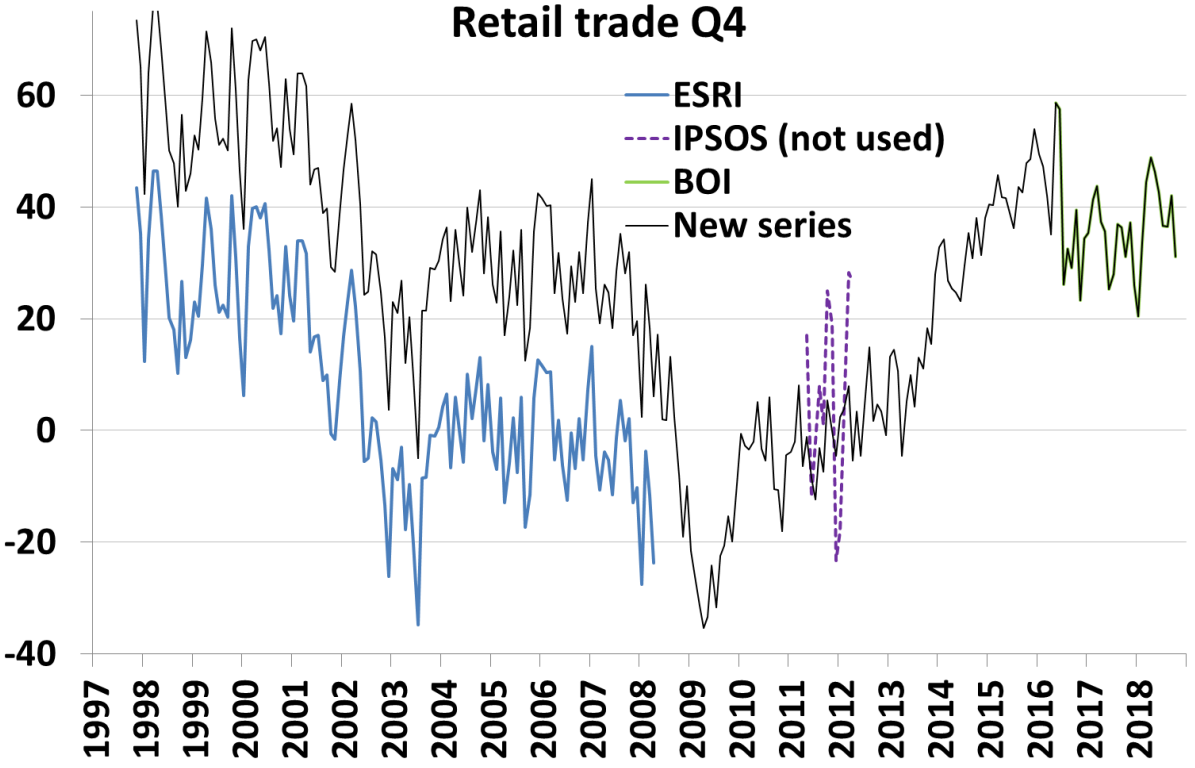
Table 6 – Series included in the average

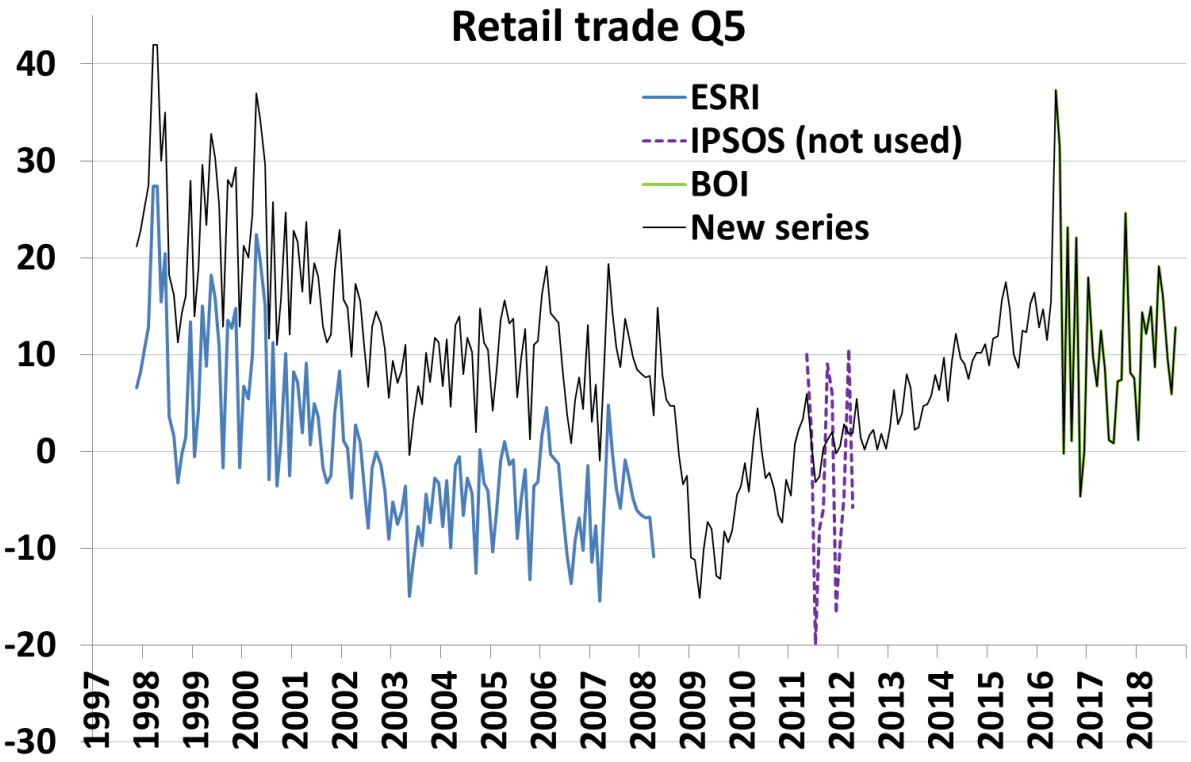
BCS question	Theme of the BCS question	Series included in the average
2	Current stock	Manufacturing PMI Stocks of Finished Goods Index + Industry question 4 (stocks)
6	Prices expectations	Services PMI Output Prices Index + Inflation for goods

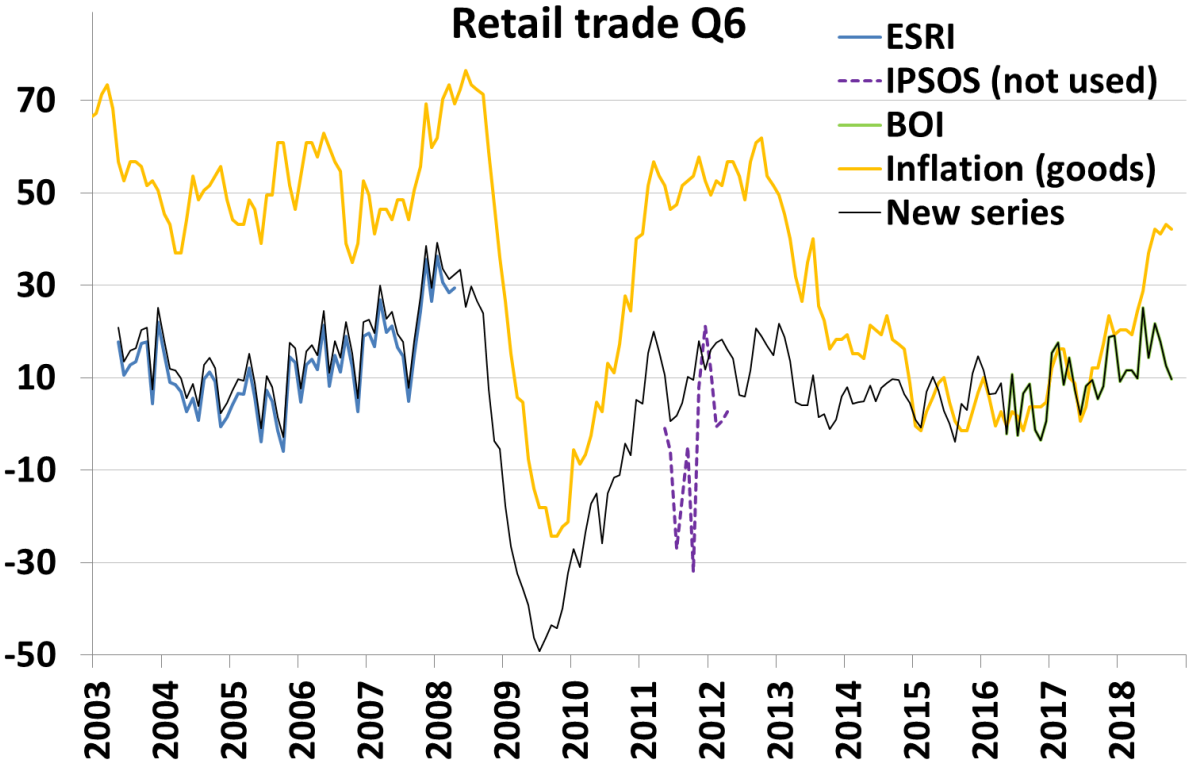

Retail trade Q1
ESRI
IPSOS (not used)
BOI
New series
60
40
20
0
-20
-40
-60
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018

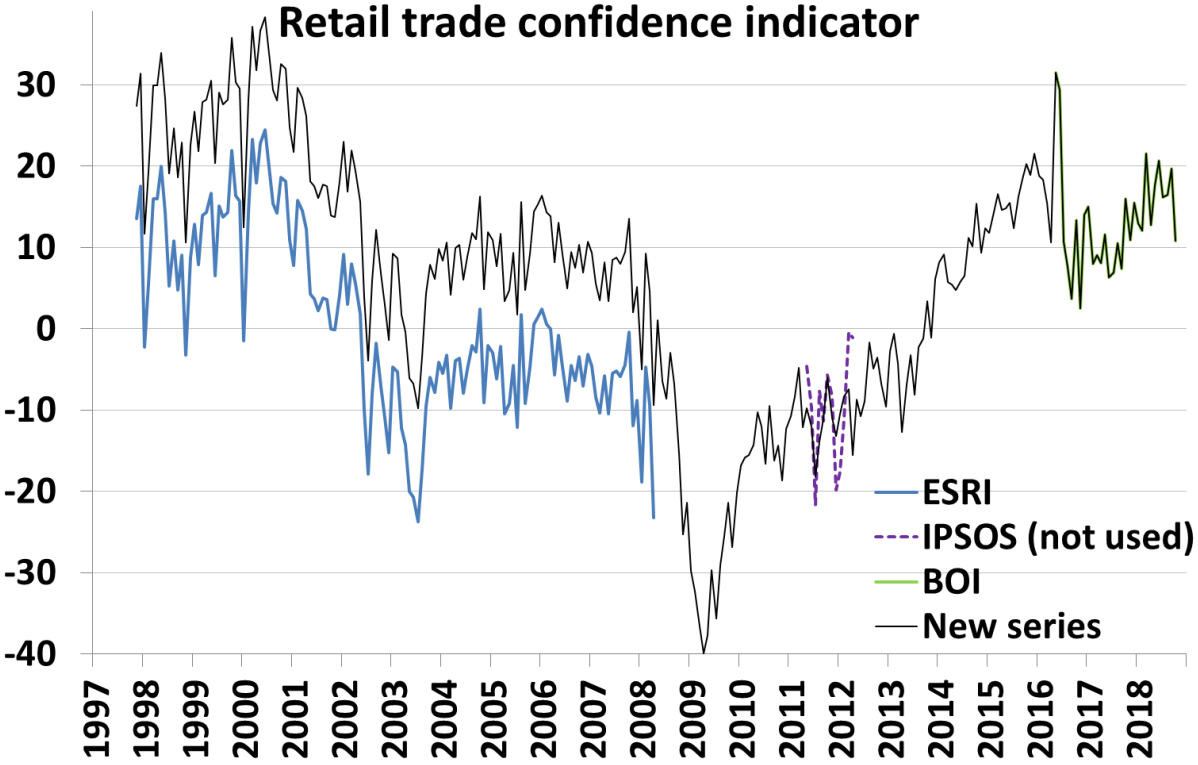

Retail trade Q2
ESRI
IPSOS (not used)
BOI
New series
30
20
10
0
-10
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018

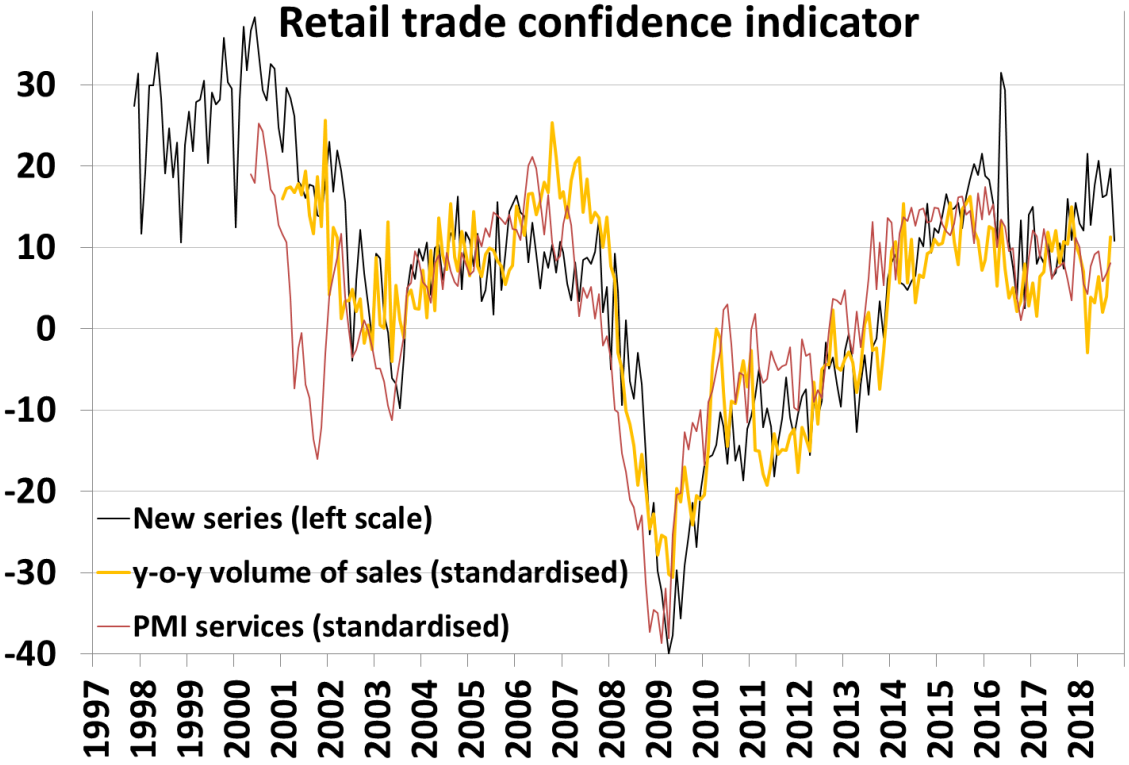

Retail trade Q3
50
30
10
-10
-30
-50
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
ESRI
IPSOS (not used)
BOI
New series


Retail trade Q4
60
40
20
0
-20
-40
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
ESRI
IPSOS (not used)
BOI
New series


Retail trade Q5
ESRI
IPSOS (not used)
BOI
New series
40
30
20
10
0
-10
-20
-30
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018


Retail trade Q6
ESRI
IPSOS (not used)
BOI
Inflation (goods)
New series
70
50
30
10
-10
-30
-50
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018


Retail trade confidence indicator
30
20
10
0
-10
-20
-30
-40
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
ESRI
IPSOS (not used)
BOI
New series


Retail trade confidence indicator
30
20
10
0
-10
-20
-30
-40
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
New series (left scale)
y-o-y volume of sales (standardised)
PMI services (standardised)

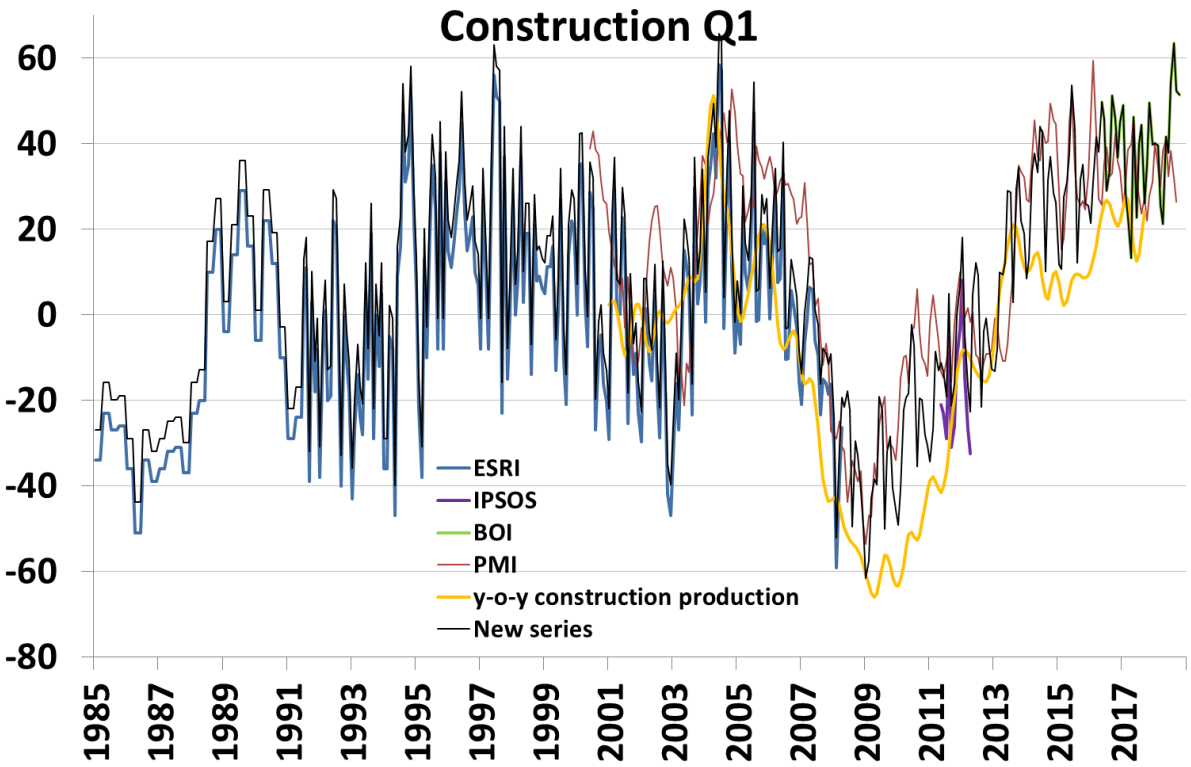
Building

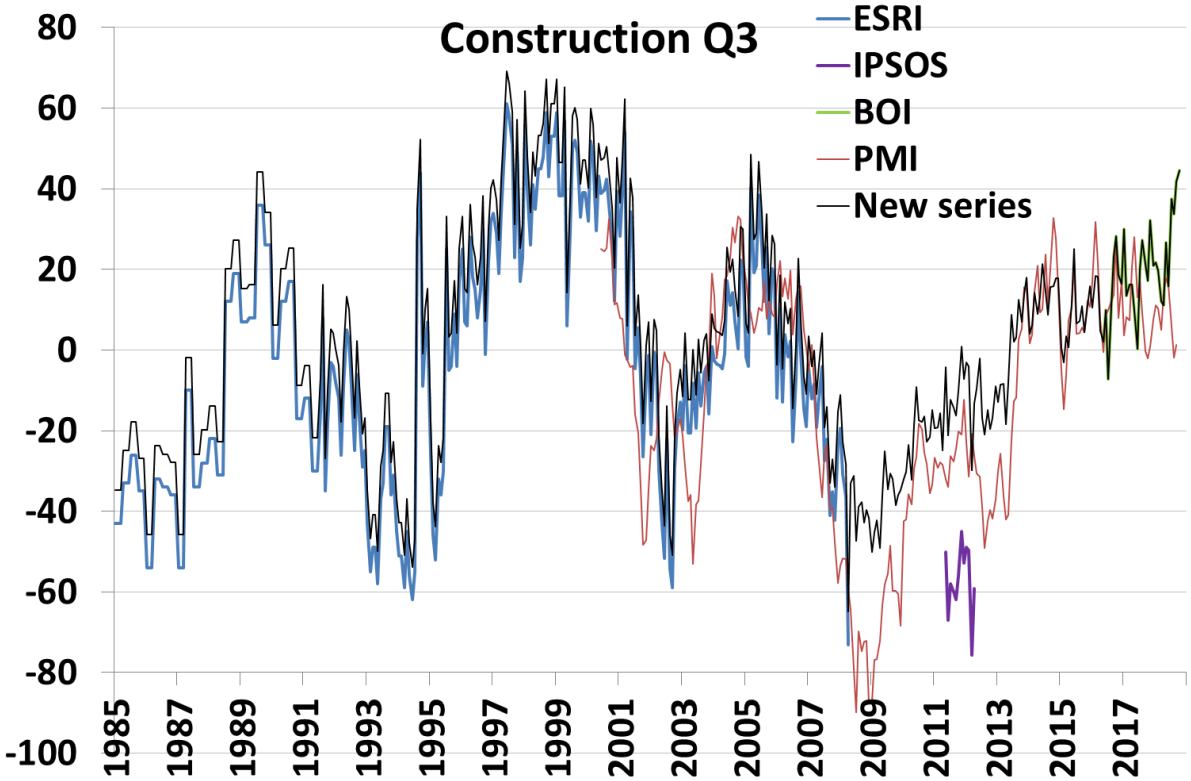
To fill the gaps between the three BCS data sets (namely from ESRI, IPSOS and BoI), four additional data sets provide useful information for the construction survey data: Series from Markit's PMI data set in the construction sector, the construction production indices (released by the Central Statistics Office and Eurostat), construction costs (released by Eurostat), and selected series from the reconstructed consumer survey data set as described previously.

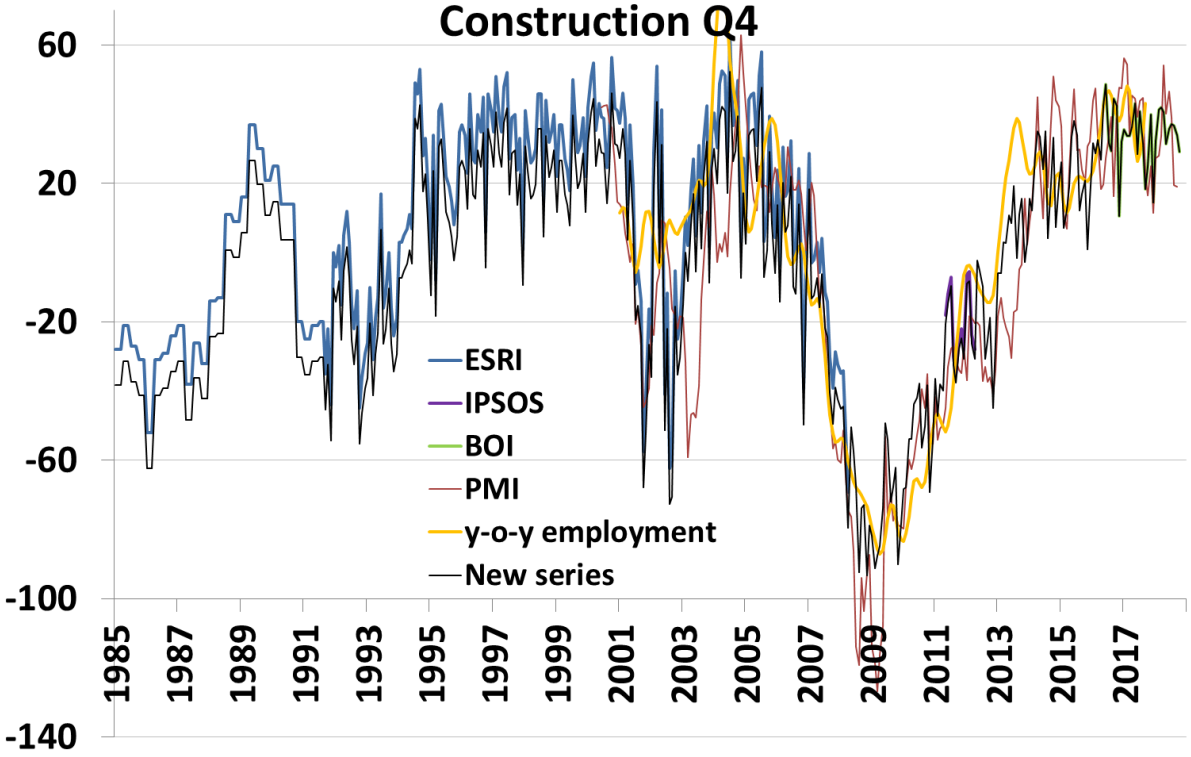
For all questions, series are reconstructed between May 2008 and April 2017 based on a PLS model and the dataset includes the closest PMI survey question. For questions 1, 3 and 4, the dataset includes year-on-year changes in construction production, while for question 5, it encompasses year-on-year changes in construction costs. Finally, the reconstructed consumer question 3 is included in the estimation for questions 1 and 5.

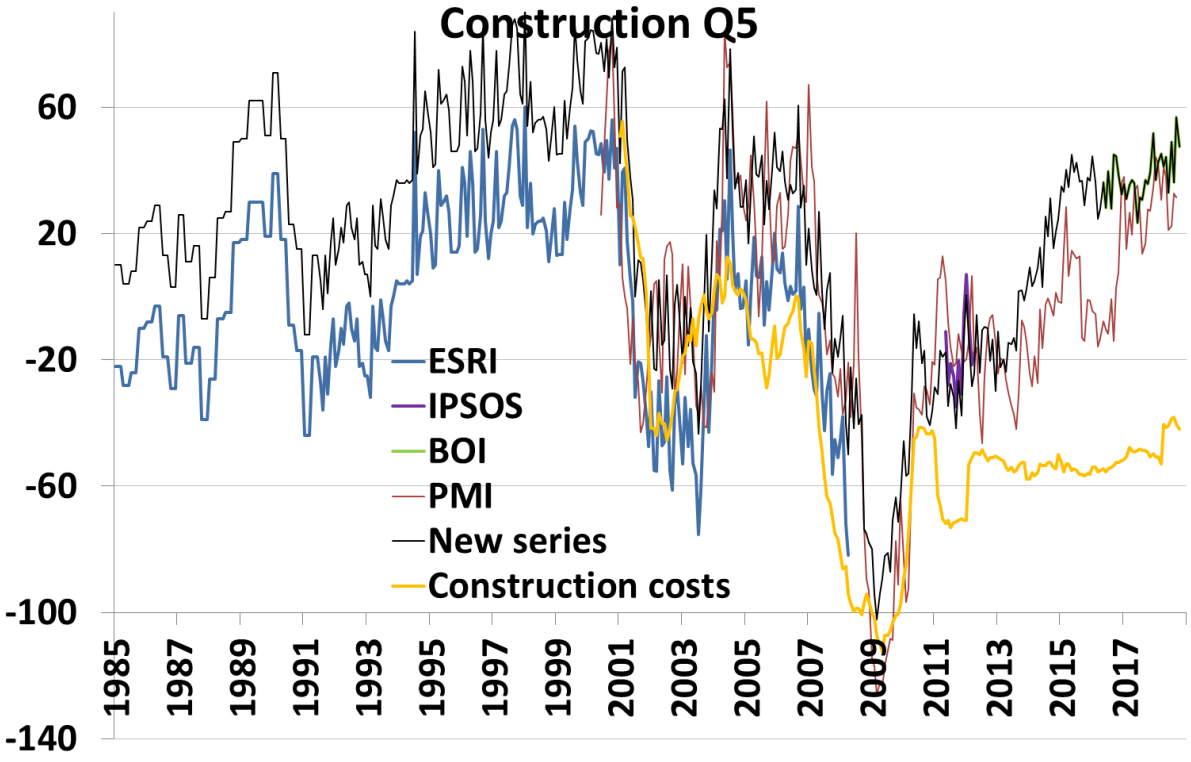
Table 7 - Dataset used with PLS

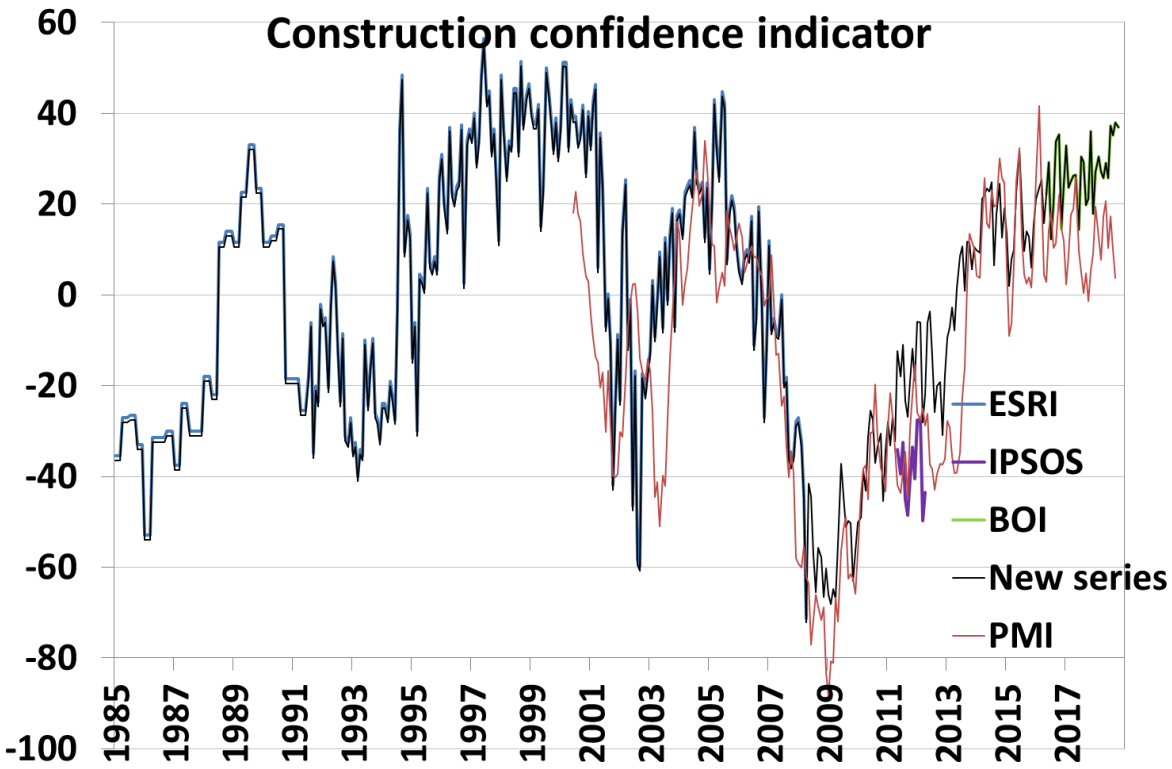
BCS question	Theme of the BCS question	Dataset used with PLS		
1	Past activity	Construction PMI Total Activity	y-o-y construction production	Consumer question 3
3	Overall order books	Construction PMI New Orders		
4	Employment expectations	Construction PMI Employment		
5	Prices expectations	Construction PMI Input Prices	y-o-y construction costs	Consumer question 3

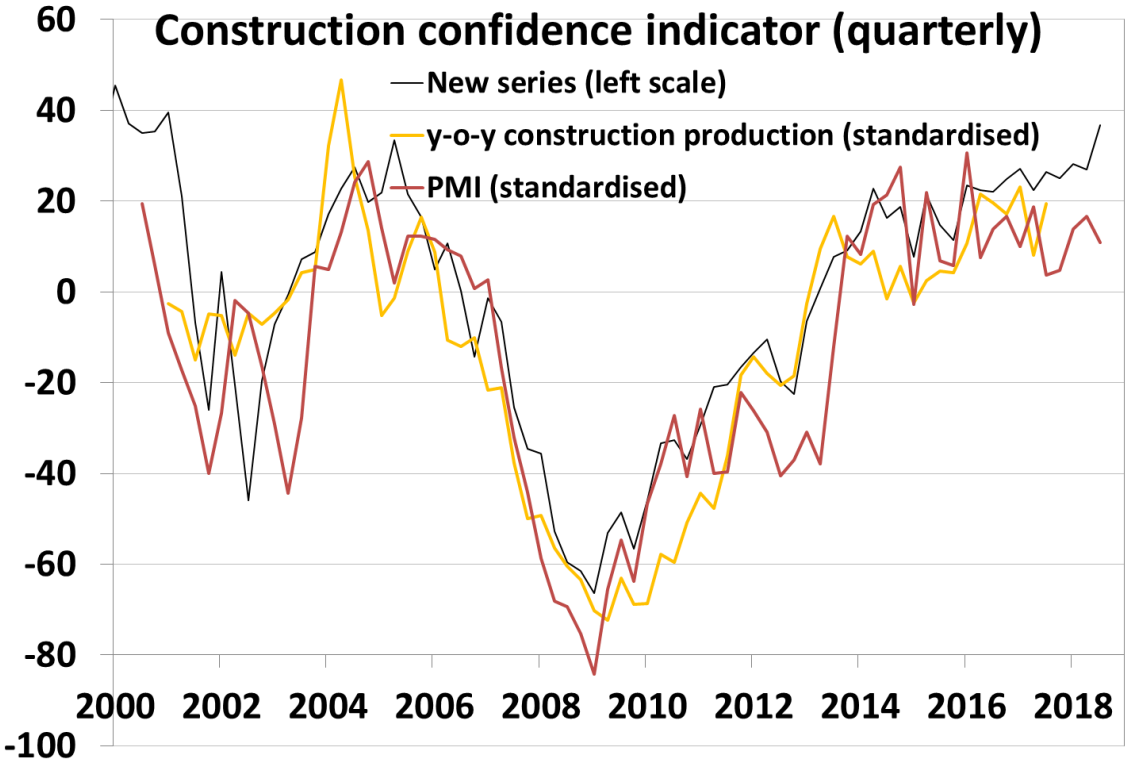

Construction Q1
60
40
20
0
-20
-40
-60
-80
1985
1987
1989
1991
1993
1995
1997
1999
2001
2003
2005
2007
2009
2011
2013
2015
2017
ESRI
IPSOS
BOI
PMI
y-o-y construction production
New series


Construction Q3
80
60
40
20
0
-20
-40
-60
-80
-100
1985
1987
1989
1991
1993
1995
1997
1999
2001
2003
2005
2007
2009
2011
2013
2015
2017
ESRI
IPSOS
BOI
PMI
New series


Construction Q4
60
20
-20
-60
-100
-140
1985
1987
1989
1991
1993
1995
1997
1999
2001
2003
2005
2007
2009
2011
2013
2015
2017
ESRI
IPSOS
BOI
PMI
y-o-y employment
New series

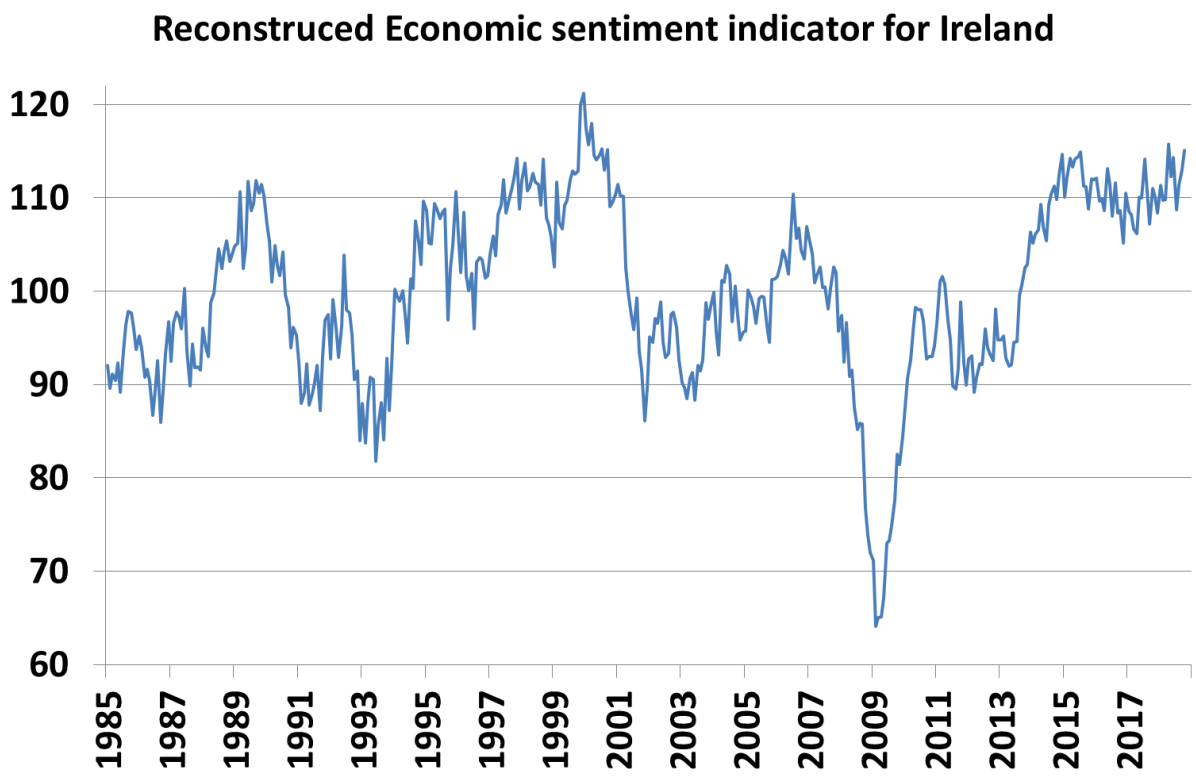

Construction Q5
60
20
-20
-60
-100
-140
1985
1987
1989
1991
1993
1995
1997
1999
2001
2003
2005
2007
2009
2011
2013
2015
2017
ESRI
IPSOS
BOI
PMI
New series
Construction costs


Construction confidence indicator
60
40
20
0
-20
-40
-60
-80
-100
1985
1987
1989
1991
1993
1995
1997
1999
2001
2003
2005
2007
2009
2011
2013
2015
2017
ESRI
IPSOS
BOI
New series
PMI

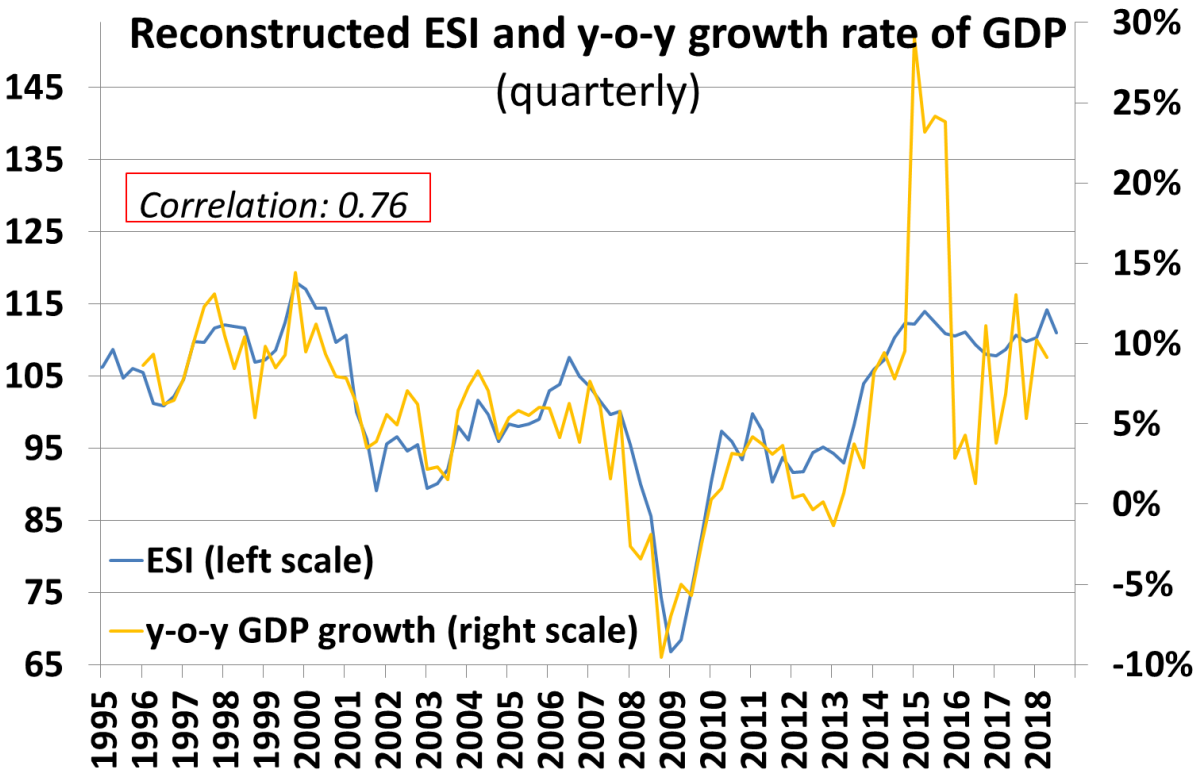

Construction confidence indicator (quarterly)
New series (left scale)
y-o-y construction production (standardised)
PMI (standardised)
60
40
20
0
-20
-40
-60
-80
-100
2000
2002
2004
2006
2008
2010
2012
2014
2016
2018

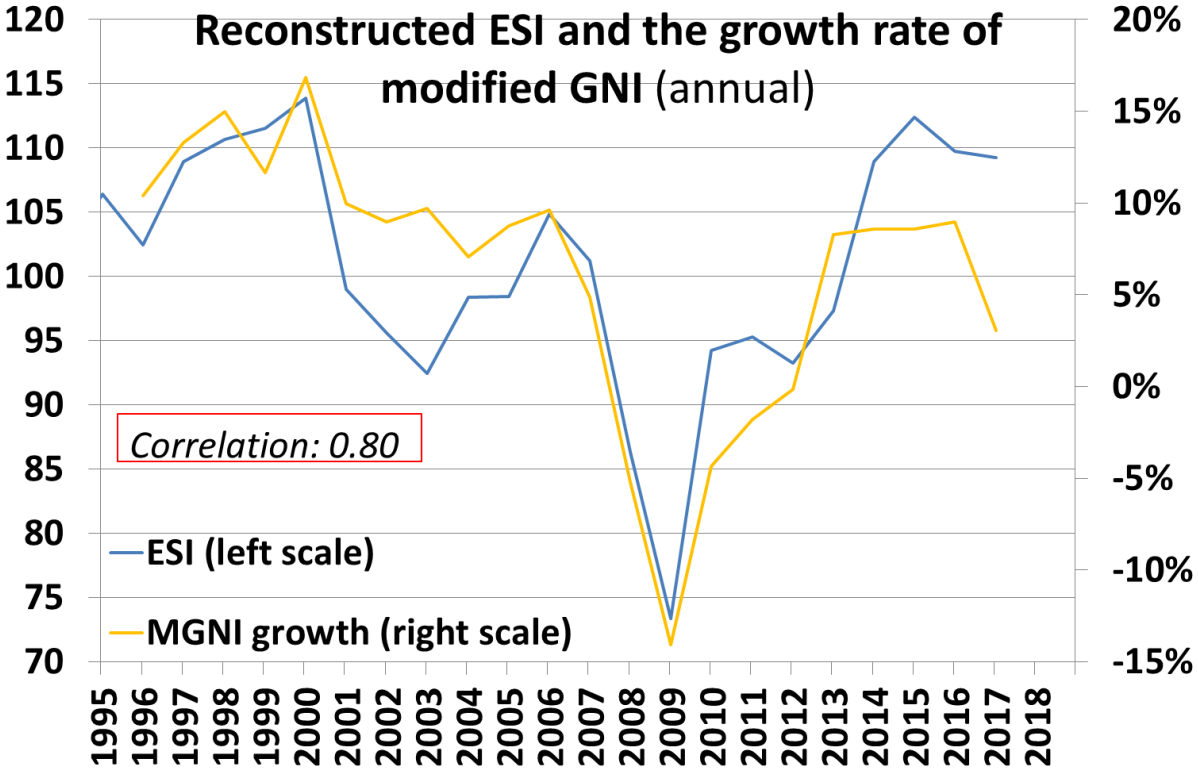
Economic sentiment indicator

After seasonal adjustment of the reconstructed individual questions, they can be aggregated into a consistent Economic sentiment indicator (ESI) for Ireland from 1985 onwards. This monthly series is presented in the graph below.



The following graph compares the quarterly average of the ESI to the year-on-year growth rate of quarterly GDP. Correlation between the two series is quite high at 0.76, a level similar to that for Germany or the United-Kingdom, even though the Irish GDP shows atypically high growth rates in 2015-2016. In annual terms, the correlation of ESI with GDP growth reaches 0.78, and with growth in modified GNI 0.80. Overall, this suggests that the reconstructed series underlying the restored ESI for Ireland are valid indicators of the Irish business cycle.


Reconstructed ESI and y-o-y growth rate of GDP
(quarterly)
Correlation: 0.76
ESI (left scale)
y-o-y GDP growth (right scale)
145
135
125
115
105
95
85
75
65
30%
25%
20%
15%
10%
5%
0%
-5%
-10%
1995 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015 2016 2017 2018


Reconstructed ESI and the growth rate of modified GNI (annual)
Correlation: 0.80
ESI (left scale)
MGNI growth (right scale)
120
115
110
105
100
95
90
85
80
75
70
20%
15%
10%
5%
0%
-5%
-10%
-15%
1995 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015 2016 2017 2018

Appendix: seasonality in the reconstructed series

In May 2019, three years of consistent data provided by BoI will be available. This is the minimum for seasonal adjustment, so the only way to allow seasonal adjustment on BoI data series for now is to provide non-seasonally adjusted (NSA) data going back to more than 3 years.

To this end, the work was carried out directly on NSA data whenever possible. But in many cases, either because reconstructing series was working better with seasonally adjusted (SA) data, or because some of the explanatory variables were only available as seasonally adjusted series, some part of the work had to be carried out on SA series. Practically, only questions from the consumer survey, excluding questions 5 and 6, and the quarterly question about capacity utilisation in Industry were treated directly with NSA data series. For all monthly business survey questions and consumer questions 5 and 6, seasonal adjustment was treated separately. The present appendix details how this was done, and how NSA series were reconstructed afterwards. The idea is to use only SA series in the first step: SA ESRI series are extended based on SA series from PMI, CONS, IP, etc. In a second step, a seasonal component is added back to the reconstructed series.

Technically, this process was handled in the following way. First, all series involved, except those that are too short (BoI and IPSOS) were seasonally adjusted if necessary. Let us consider Industry question 1 as an example:

- As for all other business questions, SA is applied to the original ESRI series.
- For the PMI, Markit provides SA series directly.
- By definition, the year-on-year growth rate of Industrial production is SA.
- And finally, the reconstructed questions 3 and 4 from the consumer surveys do not show seasonality according to X13.

Then, as described in the Industry section, a PLS model is estimated with the SA series from ESRI as dependent variable and an explanatory dataset including all manufacturing PMI questions, the year-on-year growth rate of IP in the traditional sector and the reconstructed questions 3 and 4 from the consumer survey. The model is estimated over a historical sample including all available data up to April 2008. Then, the model is applied to the explanatory data from May 2008, to simulate the out-of-sample fitted values that are used to extend the ESRI series.

In a second step, once a consistent series is reconstructed (based on the out-of-sample fit of the PLS model), a seasonal component is added back to the reconstructed series. Up to April 2008, the seasonal component is computed as the difference of the NSA and SA ESRI series. This seasonal component is then extended from May 2008 to April 2017 by replicating its last year (from May 2007 to April 2008). This way, the reconstructed series up to April 2017 is now NSA. Finally, the third step is then to add NSA data from IPSOS and BoI to this NSA reconstructed series.

This method has 2 main advantages. First, the reconstructed series up to 2008 match exactly ESRI series (NSA), except for the shift that was necessary to align ESRI and BoI series. In addition, this

method ensures that SA will deal perfectly with the reconstructed seasonality up to 2016, as seasonality is constant by design between May 2007 and April 2016.[10]

[10] Yet, it might create a break in seasonality in May 2016, as it is a priori not a given that the seasonal component of BoI series is the same as that of ESRI series in 2007-2008. But as long as SA cannot be applied to BoI data (until May 2019), it is not possible to check whether a break in seasonality actually appears. Only in May 2019 will it be possible to assess the presence of a break in the seasonal component and deal with it properly if needed.