Investigating the Effect of Quarterly Collection of Employee Jobs data on the Estimated Standard Error of Change for Total Turnover on the Monthly Business Survey

B.J. Cotterell (2015)

Executive Summary

Standard errors of change in estimated turnover at lag one have been estimated for the Monthly Business Survey (MBS) for data between September 2012 and March 2014. This work was done in response to a suggestion by the Statistics Authority that there may be a reduced response rate, and hence reduced quality, for the turnover sample in quarter months (i.e. March, June, September and December), since in these months ONS collects employee jobs information from a large number of reporting units along with turnover. This work tests the effects on the standard errors assuming full response and is therefore likely to underestimate the true standard error of change. Note the data used are made up of observed and imputed values.

The data were analysed at the whole survey level and (mostly) the 2 digit SIC level. At the whole survey level, there appears to be an effect caused by the annual IDBR update, with clear peaks in the standard error of change being observed for January 2013 and January 2014. Another peak is caused by the introduction of the new sample allocation in April 2013. This peak is surprisingly small considering the reduction in the sample overlap in that month. Smaller peaks occur at some quarter months but in December 2012 and December 2013 the data are almost flat.

At the industrial level the data are observed to be very variable. For many industries a number of peaks are observed but these are frequently not on quarter months. Charts for two industries are presented in this report, as well as for the whole survey.

It appears that the effect of collecting employee jobs in quarter months on the standard error of change in turnover is at most very weak. It should also be noted that estimated standard errors are intrinsically unstable over time, since standard errors are much more sensitive to outliers than is the case for estimated totals or means.

Introduction

The Monthly Business Survey (MBS) is a survey that collects information on reporting units. It does not cover the whole economy but it does cover a large number of industries. The main variable required is turnover, which is collected on a monthly basis. However, for many businesses, on quarter months MBS also collects employee jobs totals and so-called splits: full time males, part time males, full time females and part time females. It has been suggested that collecting this information may reduce the response rate to the turnover survey on quarter months. Note that work on investigating the effect of the quarterly employee jobs questions on turnover response rates for MBS has been done by Bucknall and Graham (2015). These authors conclude that there is no evidence that the response rate for the turnover sample changes in the quarter months.

It should be noted that estimating standard errors of change in the presence of imputed values is not straightforward. The imputation method used in MBS is deterministic so standard techniques

such as multiple imputation cannot be used. Therefore the estimated standard error estimates do not fully reflect the usual effects of non-response.

A reduced response rate may impact on the estimated standard errors of change. Clearly, ideally one would like these effects to be as small as possible.

Imputation for non response might cause a decrease in the estimated standard errors since the imputations are likely to be close to the centre of the distribution. However the true uncertainty is inflated by non-response.

The assumption that the data for the census strata are known is fallacious, since we know that there are imputed values for non-response. In business surveys, large units are usually extremely important in most industries. Therefore uncertainty in these can have a very large influence on the observed true standard errors. This cannot be treated here, and this is a disadvantage of the work presented.

This work is intended to assess the effect of the suggested reduced response rate on the estimated standard errors of change as they are calculated in the current system.

Method

The estimation of standard errors uses a model assisted approach assuming that the estimator used is a separate ratio estimator. This assumption is unlikely to make much difference as opposed to using some combined ratio estimators. The population data used were from the reporting unit extracts stored on a monthly basis by RD&I Directorate. These were not extracted at the same time as the sample data from MBS used. Therefore the results of this analysis should be considered to be indicative.

A description of the method used in this report can be found in a MSc thesis (Cotterell (2012)). Other information on methods for variances of change is described by Holmes and Skinner (2000) for the Labour Force Survey.

The statistics presented in this report are a ratio of the standard error of change divided by the total in the first period of comparison, expressed as a percentage. For example, for February 2014 the standard error of change between January and February would be divided by the estimated total in January 2014 and then multiplied by 100. Hereinafter this statistic will be referred to as a scaled standard error of change. This method has the effect of showing the relative performance of the method by industry and overall over time. Like the coefficient of variation, it has the disadvantage that it depends on the numerator and denominator of a ratio. However, in general, standard errors of change are more variable and sensitive to outliers than the total.

Note that the two small industries introduced to MBS in April 2013 have been excluded from the analysis, since they did not exist on the survey prior to April 2013.

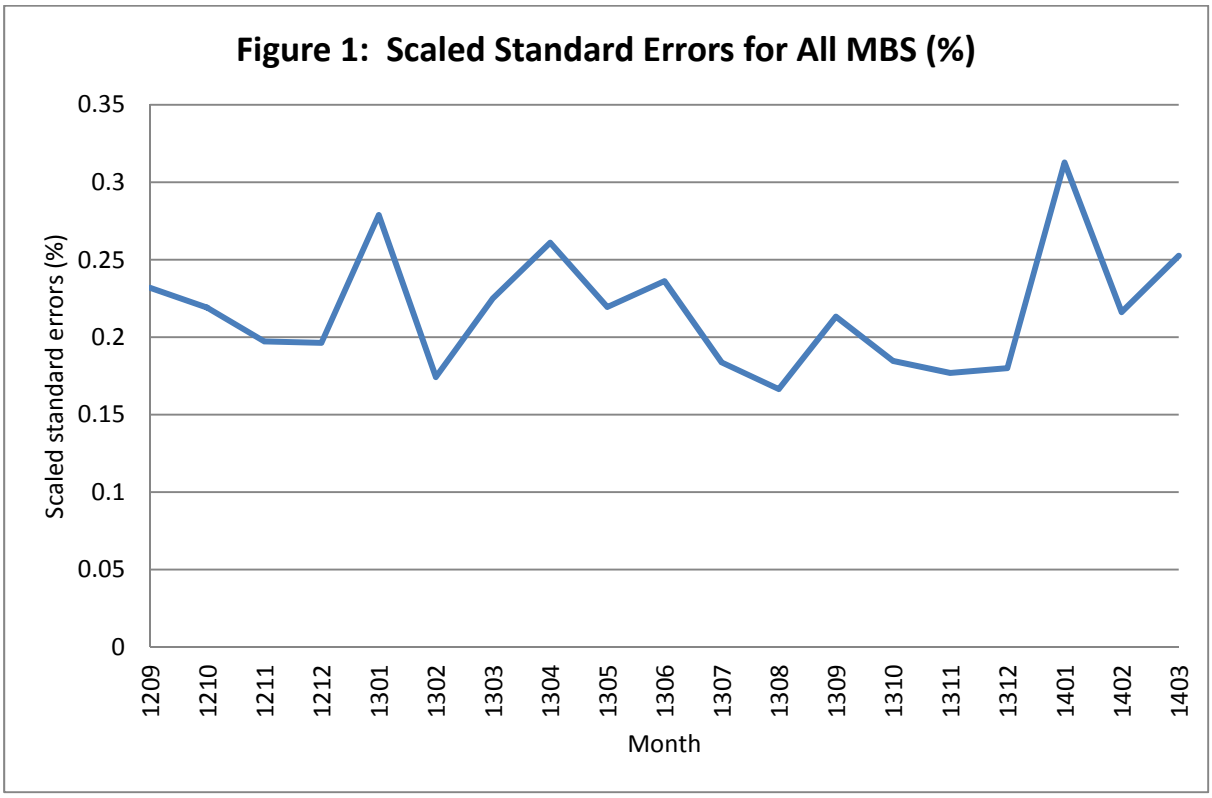
Results

Note that all of the results presented are at lag one; that is, the month on previous month data.

Figure 1 shows estimates of the scaled standard error of change for the whole MBS survey against time for the period September 2012 to March 2014. There are two large and clear peaks at January 2013 and January 2014. These are the periods in which the IDBR is updated using data from the Business Register and Employment Survey (BRES) and other sources. The current to frozen update for turnover also is introduced at this time. At these periods, a number of units change cells (strata) so that the overlap per cell is reduced. This should cause a spike in the chart of the scaled standard errors of change and this is observed, at least for the data used here.

A new sample allocation was introduced in April 2013. In the first period of introduction, the sample overlap will be much lower than in a normal month. This should cause a spike in the scaled standard errors of change in April 2013. A peak is observed at this point; however, it is smaller than the January peaks. This suggests that the introduction of the new allocation showed a smaller effect on the scaled standard errors of change than one might have expected in the first period of implementation.

There are two smaller peaks at June 2013 and September 2013. These occur at quarter months and there is also an increase at March 2013. Note that the latter would not be expected to cause a peak because it is adjacent to the April 2013 figure which shows a larger increase because of the introduction of the new allocation at that period. If the scaled standard error of change was increased on the quarter months, one should see a large increase in the Decembers, although not peaks since they are adjacent to the January peaks. In practice the December values are flat. Thus the evidence of an effect on the quarter periods is weak at best.



An example of an industry is SIC Division 10, as shown in figure 2. This shows clear peaks at the January updates. There are other peaks, but none occur on quarter months. There is no obvious reason for this. This indicates the risk of seeing patterns in data which are possibly coincidental.

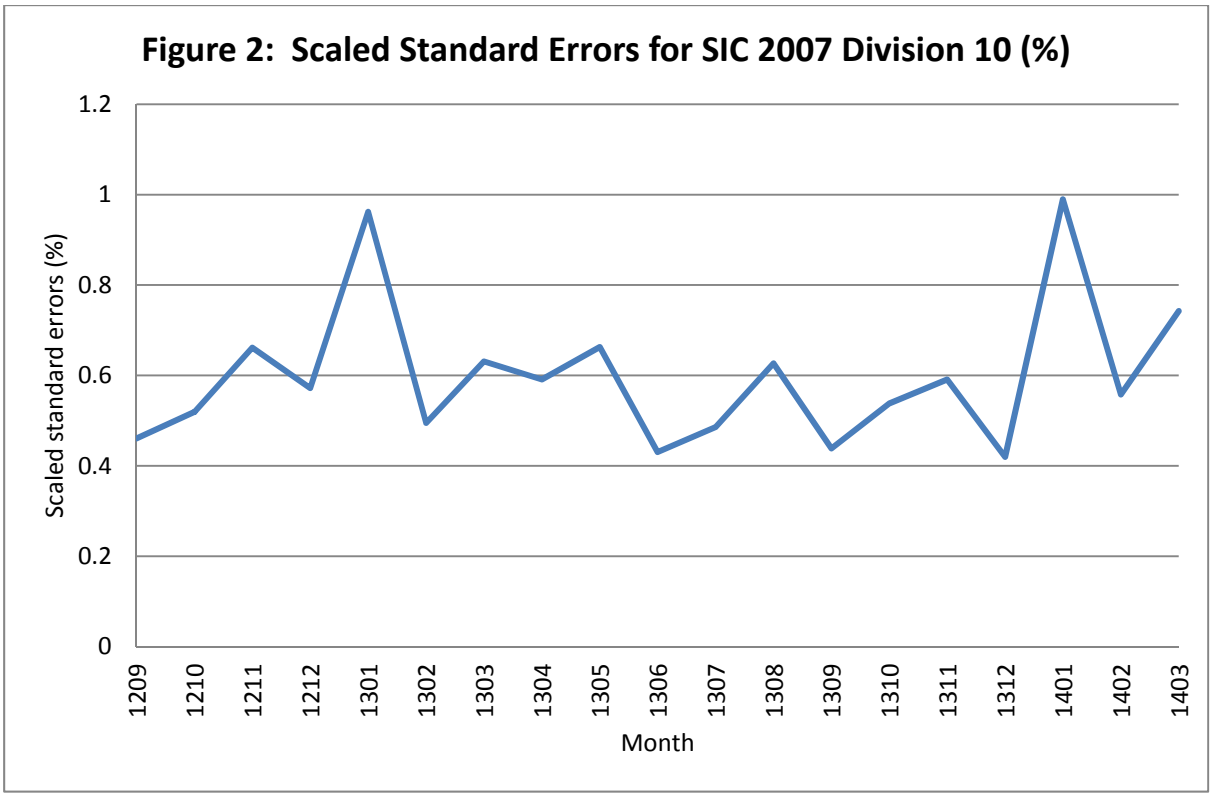
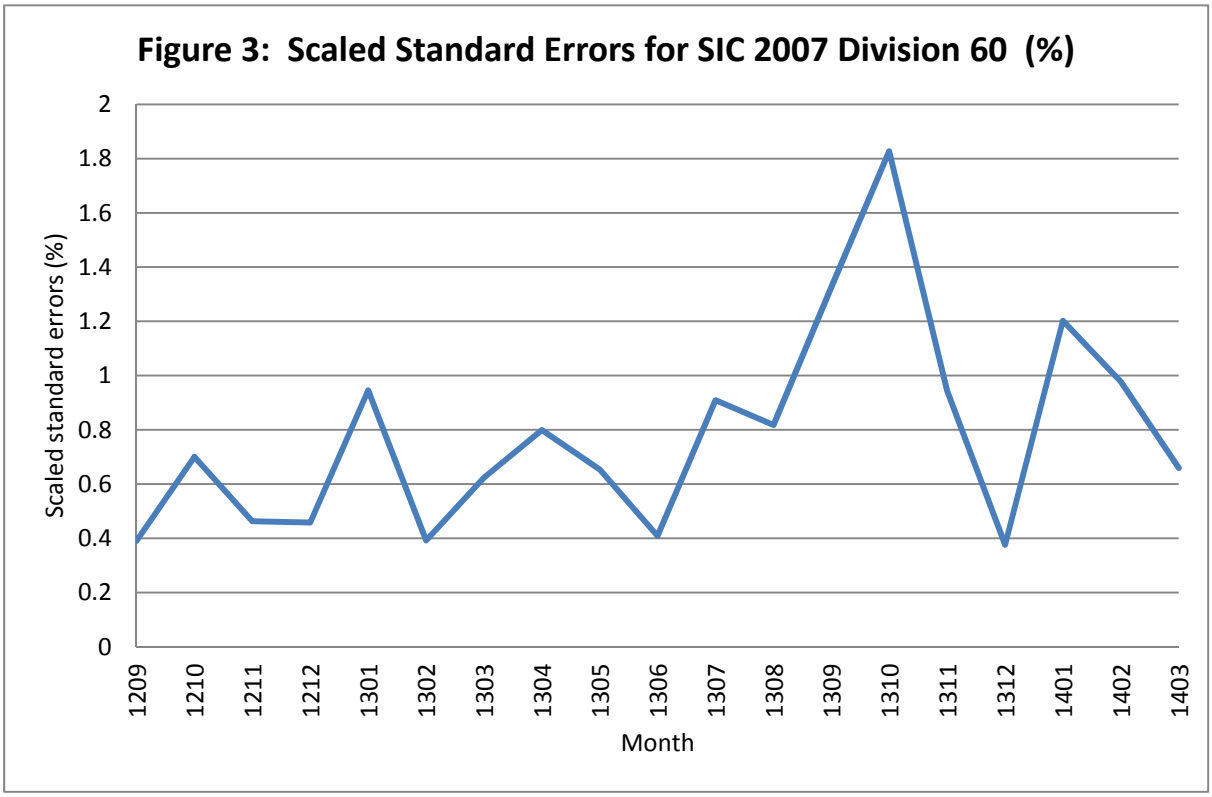


Figure 3 shows the corresponding chart for SIC Division 60. Again none of the peaks occur at quarter months. Note that two outliers in this industry caused a huge rise in the scaled standard error of change in March 2014, the most influential of which was a not cleared error. These values were replaced by imputing one twelfth of the IDBR annual turnover. This brought the standard error into line with the previous values. These were the only changes that were made to the MBS data in March 2014.



Element level data from some previous periods were amended by replacing the not cleared errors with the value found on Common Software (an ONS database) as the latest available data. Note that there are still a lot of errors in the datasets used in this work. However, it is unlikely that the estimates are conservative given the assumption that fully enumerated data are measured exactly in MBS, whereas a number of these are imputed.

Note that changes to the data were made only when there was an extreme change that could not be ignored. In general, the industrial series are very unstable and there was no obvious effect from the introduction of the new sample allocation in April 2013.

Conclusion

There does not appear to be any real evidence that the collection of employee jobs data in quarter months impacts on the scaled standard errors of change for the MBS, especially for the industrial series. The impact of having a higher nonresponse rate on estimated standard errors is not clear, given the imputation method used, and the calculation of standard error estimates, in effect, assumes a fully responding sample. One should also take into account the fact that standard errors can be very unstable over time.

The results of this report should be considered in conjunction with the report on the weighted and not weighted response rates produced by Graham and Bucknall (2015).

References

Cotterell, B.J. (2012): “Estimating Variances of Change for Institutional Surveys”, MSc dissertation, University of Southampton

Graham, B.L. and Bucknall, R.D. (2015): “Impact of Quarterly Employment Question on Monthly Survey Response”, ONS report

Holmes, D.J. and Skinner, C.J. (2000): “Variance Estimation for Labour Force Survey Estimates of Level and Change”, Government Statistical Service Methodology Series No 21.