Retail Sales Index - A Description of the New Methods

Summary

A number of aspects of the methods and systems used to produce the Retail Sales Index (RSI) have been reviewed. The changes to the Retail Sales Inquiry and Retail Sales Index outlined in this article represent the largest investment of methodological development for the RSI in twenty years. They ensure that RSI methods are soundly based and incorporate standard methods used for other ONS business surveys.

The changes are summarised as follows:

- the RSI has been rebased from 1995=100 to 2000=100
- the methods used in the Retail Sales Inquiry have been updated, i.e:
  - survey estimation is now based on ratio estimation, the standard method of estimation for ONS business surveys;
  - the standard business survey method is now being used for dealing with 'small' non responding retailers;
  - a more systematised approach to outlier detection and treatment;
  - a review of trading-day weights.
- a new computer system has been introduced for index construction
- calendar and seasonal adjustment have been updated
- late data, not available at the publication of the current series, have been taken on.

These changes are being introduced to the published series from 24 October 2003. They will affect all published series, and will lead to revisions to levels for all periods and revisions to growth rates from January 2000. The new series shows a different path and more month on month volatility, better reflecting the nature of retail sales than the current series.
Introduction

This article describes in some detail the new methods that have been implemented for the Retail Sales Index. An article 'Retail Sales Index Development: Implementation' [http://www.statistics.gov.uk/cci/article.asp?ID=475] provides an overview of the changes that are being introduced and gives an indication of the impact of the changes on the headline RSI.


This note describes the new methods in each of three main areas where we have introduced improvements and also describes some additional changes that have been introduced. This article is in 4 parts:

- Survey methods used for the Retail Sales Inquiry
- Rebasing and other changes to the deflators
- Calendar and seasonal adjustment
- Other changes in methods and approach

Survey methods

We have reviewed a number of aspects of survey methods and processes used for the Retail Sales Inquiry. In particular we have reviewed

- sample allocation
- survey estimation
- imputation for non response
- outlier detection and treatment

Updating the sample allocation

The total sample size for the retail sales inquiry is 5,000; this sample is allocated to industry by sizeband strata or 'cells'. This breakdown of the population into strata
allows retailers of a similar industry and size to be placed in the same stratum. The main benefit of stratification is that we can produce best estimates of retail turnover with a given sample size. We need to review and update the allocation of retailers to strata from time to time to reflect changes in the retail industry, and to ensure that we are producing best results with a given sample size. We have recently implemented the following changes:

- the allocation has been designed to minimise the standard error of the monthly change in retail sales
- retailers with between 10 and 99 employment will remain in the survey for 27 months
- small retailers with large turnover (annual turnover at least £40million) will always be included in the sample

Changes in the sample allocation were implemented in three stages between September 2002 and January 2003.

Survey estimation

The estimation process takes returned turnover from the sampled retailers and uses it to produce estimates of total retail turnover for all retailers in the population. As part of the current triennial review we have taken forward a recommendation from the last triennial review. The last review recommended that 'the retail sales index should move ... to a system based on grossing to the register each month' - i.e. to use ratio estimation, the standard business survey method of estimation. The Retail Sales Inquiry currently uses a matched-pairs approach for the sampled part of the survey (about 30% of the total retail turnover), rather than 'ratio estimation' that is more commonly used for business surveys. Table 1 compares the main features of matched-pairs and ratio estimation.

**Table 1**

<table>
<thead>
<tr>
<th>Matched pairs</th>
<th>Ratio estimation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Designed to estimate month on month changes</td>
<td>Estimates a level for each period – growth rates derived from level estimates</td>
</tr>
<tr>
<td>Uses data from respondents who have returned for current and previous periods</td>
<td>Uses all returned data</td>
</tr>
<tr>
<td>Therefore, stable over two consecutive periods</td>
<td>Consecutive levels estimates can be</td>
</tr>
</tbody>
</table>
We have investigated the feasibility of introducing ratio estimation to the Retail Sales Inquiry and concluded that we should move to use ratio estimation. The benefits of a move from matched pairs to ratio estimation of retail turnover are:

- we can pick up changes in the real world (via register updates) faster. The matched pairs approach to working means that changes in the real world can only be taken on at periodic rebasing exercises. The move to ratio estimation means we can take on these changes as the register is updated
- ratio estimation allows a more flexible way of working - there are no operational constraints to reviewing and updating estimates for previous months - so we can take on late returns from retailers more quickly. As a result, we are able to introduce the revisions policy used for other economic indicators. We will introduce the national accounts revisions policy when we introduce the ratio estimator
- it is the standard method, as used in other business surveys

**Imputation for non-response**

We routinely response chase retailers who fail to respond to the survey. Even so, some large retailers (those with employment of at least 100) fail to respond in time to be included in provisional survey results. For these large non-responding retailers, figures are manually imputed based on the individual histories of the retailer concerned. Smaller retailers who do not respond are assumed to move in the same way as small responding retailers. We will use the standard ONS business survey method for imputing for 'small' non-respondents for the retail sales inquiry. This standard method calculates, within each industry, the average monthly movements for those retailers who have returned. It then applies that average monthly movement to the previous (returned or imputed) value for non-responding retailers. This produces an imputed turnover figure for the non-responding retailer for the current month. For 'large' non-responding retailers, we will continue to manually impute based on the
individual history of the retailer concerned (which we have shown produces better results for these large retailers).

**Outlier detection and treatment**

Unusually high or low turnover returns from retailers can distort the turnover estimates produced. The current approach completely excludes a retailer’s data if the ratio of the current month's sales to the previous month's sales is less than 0.2 or greater than 5. Retailers with data which is less extreme but atypical may be excluded from the calculation of the month on month movement but their data will be added back to produce the final totals. These less extreme, atypical retailers are currently identified manually. We will introduce a more systematic approach to outlier detection and treatment that will reduce the number of outliers in any particular period. All outliers identified will be treated through a method of post-stratification. The reduction in the number of outliers will increase the volatility of the month on month estimates, but will reduce the chances of misrepresenting the picture of growth over longer time periods.

**Changes to the deflators**

The RSI has been rebased from 1995=100 to 2000=100. As part of this process we have re-weighted the deflators from the current base year (1995) to the new base year (2000). We are taking the opportunity of rebasing to review the way we derive industry level deflators. As a result, we have changed the way detailed deflators are combined together (change from use of an 'arithmetic' to an 'harmonic' mean).

**Introducing a harmonic mean**

The following are the steps of logic that led to us adopting harmonic means for deflators.

1. Because the RSI is a Laspeyres index of volume we need to use Paasche price indices (that is, deflators).
2. The Paasche price index (deflator) between periods 0 and \( t \) is:
\[ d_{it} = \frac{\sum_p p_u q_u}{\sum_i p_{i0} q_{i0}} = \frac{\sum_i p_{i0} q_{i0}}{\sum_p p_u q_{i0}} = \frac{1}{\sum_i w_{it} \frac{p_{i0}}{p_u}} \quad (1) \]

where: \( w_{it} = \frac{p_{i0} q_{i0}}{\sum_j p_j q_j} \) is the current (period \( t \)) weight for commodity \( i \).

Equation (1) shows that the required deflator is a current-weighted harmonic mean of the commodity indices \( \frac{p_{i0}}{p_{it}} \). (The harmonic mean is the reciprocal of the mean of the reciprocals.)

3. We do not know the current weights \( w_{it} \), so we have to use an approximating assumption.

4. If we assume that \( q_{it} = q_{i0} \) (that is, relative volumes are constant), we can replace \( q_{it} \) in equation (1) by \( q_{i0} \) to produce the alternative estimator:

\[ d'_{it} = \frac{\sum_p p_u q_{i0}}{\sum_i p_{i0} q_{i0}} = \frac{\sum_i p_{i0} q_{i0}}{\sum_p p_u q_{i0}} = \sum_i w_{i0} \frac{p_{i0}}{p_i} \quad (2) \]

This is a base-weighted arithmetic mean, as used in the current RSI.

5. However, economic theory suggests that there is an approximately inverse relationship between prices and quantities. This implies that relative values \( \frac{p_{it} q_{it}}{p_{i0} q_{i0}} \) are more likely to be constant than relative volumes. That is \( p_{it} q_{it} = p_{i0} q_{i0} \) or, equivalently, \( w_{it} = w_{i0} \). With this approximating assumption, we have:

\[ d''_{it} = \frac{1}{\sum_i w_{i0} \frac{p_{i0}}{p_u}} \quad (3) \]

This is the base-weighted harmonic mean.

6. To decide between equations (2) and (3), we need to know which of the two assumptions is more plausible. To do so, note that the assumption in paragraph 4
(q_{it\mu q_{i0}}) implies p_{it}q_{it\mu p_{i0}q_{i0}} = (p_{i0}/p_{i0})p_{i0}q_{i0}. So the question is whether current expenditure weights \{p_{it}q_{it}\} are better approximated by base weights \{p_{i0}q_{i0}\}, implying the use of a harmonic mean (equation 3), or by inflated base weights \{(p_{i0}/p_{i0})p_{i0}q_{i0}\}, implying the use of an arithmetic mean (equation 2).

7. To test this we compared expenditure weights for 1999 against 1995 weights and against inflated 1995 weights. The conclusion was that the correspondence between 1999 weights and 1995 weights was stronger than that between 1999 weights and inflated 1995 weights. This implies that the assumption of constant values \{p_{it}q_{it}\mu p_{i0}q_{i0}\} is more plausible than the assumption of constant volumes \{q_{it}\mu q_{i0}\} and supports the use of a harmonic as opposed to arithmetic mean.

**Calendar and seasonal adjustment**

We have taken the opportunity of rebasing and the review of estimation to review RSI calendar and seasonal adjustment, although the basic approach remains unchanged.

**Calendar adjustment**

Calendar adjustment deals with moving bank holidays and the fact that the RSI standard periods move one day a year (two in leap years) relative to the calendar (the phase-shift effect). Calendar adjustment is based on fitting a regression model to estimate the impact of moving bank holidays and the phase-shift effect. The current model was derived using data up to the mid 1990s. We have reviewed the model using data up to the end of 2002, which has given a revised regression model.

The main features of the retail sales index series, which complicate seasonal adjustment, are the calendar effects induced in the series by the 4, 4, 5, recording periods. These can be estimated and removed using the RegARIMA tool in X12ARIMA. This allows an ARIMA model of the time series to be augmented by explanatory variables, specifically designed to identify particular calendar effects. A RegARIMA model of a (log transformed) time series \(y_t\) can be written

\[ y_t = \sum \beta_i x_{it} + z_t, \]
where \( \sum_i \beta_i x_{it} \) is the regression effect, and \( z_t \) is the time series of regression errors fitted by an ARIMA model. In this way, estimation of calendar effects is carried out simultaneously within the framework of a unified statistical model.

The following calendar effects are adjusted for:

- The phase shift effect resulting from the year to year movement of standard recording periods.
- Easter effects, resulting from the Easter holiday moving between the ‘March’ and ‘April’ standard recording periods.
- A May Bank Holiday effect resulting from the holiday moving between ‘May’ and ‘June’ standard recording periods.
- An August Bank holiday effect resulting from the holiday moving between ‘August’ and ‘September’ standard recording periods.

The possibility of other bank holiday effects was considered. The early May Bank holiday moves between April and May, but it occurs in April relatively infrequently, and there was no evidence of any effect in the series. The New Year’s holiday moves between December and January although there is no identifiable New Year’s day effect. This may be because of additional complicating factors such as day of week effects – the New Year’s Day holiday does not always occur on a Monday, as the other Bank Holidays do.

**The Phase shift effect**

The phase shift effect is that variation in the series that is attributable to the movement over successive years of each standard reporting period (SRP) relative to its corresponding calendar month. Figure 1 below illustrates this movement of the SRP for November relative to the calendar month. It is in the context of a strong seasonal pattern, that this movement becomes significant.
A set of twelve phase shift regressors are used to model the phase-shift effect. There is one for each month, and they each reflect the movement of that SRP relative to the corresponding calendar month. Removing the phase shift effect amounts to something close to calendarisation, although falls short of this. Full calendarisation would mean attempting to put the data back onto a calendar month basis. Phase shift adjustment means that SRP’s are effectively centred on their corresponding calendar month, so the series is no longer affected by their movement.

Figure 1 illustrates how the movement of the November SRP relative to the calendar month is used to derive the explanatory variable. The values shown are the number of days between the start of the SRP and the start of the calendar month. A variable is constructed which takes the value zero in all months other than November, while in November it takes the values shown in Figure 1. The variable as a whole is then centred by subtracting from each month the long run monthly mean (calculated over the period 1992-2019, one 28 year cycle) from each month. Monthly means are zero for all months other than November, and turns out to be exactly 1 in November.

The explanatory variables are mean centred in this way so that the regression effects (calculated as \( \sum \beta_i x_{it} \)) which are removed from the series, have no overall mean and
no fixed seasonal component. This ensures that the level of the series is not biased, and that seasonality is confined to the seasonal factors. This centring makes no difference to the model in terms of the forecast, estimated coefficients, log-likelihood etc. This is because the differencing, and seasonal differencing operators of a RegARIMA model, apply also to the regressors. This approach of centring variables by subtraction of monthly means is a standard one used in the calculation of some of the pre-defined regression variables available in X12ARIMA.

Figure 2. Phase shift regressor for November 1986-2019

Bank Holidays

The movement of SRP’s also causes bank holidays to become moving holidays in the sense that they don’t always occur in the same SRP. The Bank Holiday which occurs on the last Monday of May often occurs in the June SRP, but then moves back to the May SRP. This would lead to a lower retail sales figure for May, with higher sales in June. A similar thing occurs when the Bank Holiday which occurs on the last Monday of August switches between the September and August SRP’s. Two explanatory variables are formulated to model the switch in retail activity between these months when this occurs. The variables are defined in the following way:

\[
MBH_i = \begin{cases} 
1 & \text{In May, in years where the bank holiday is in the May SRP} \\
-0.8 & \text{In June, in years where the bank holiday is in the May SRP} \\
0 & \text{otherwise}
\end{cases}
\]
\[
ABH_i = \begin{cases} 
1 & \text{In August, in years where the bank holiday is in the August SRP} \\
-0.8 & \text{In September, in years where the bank holiday is in the August SRP} \\
0 & \text{otherwise}
\end{cases}
\]

The value –0.8 (rather than –1) is used since both June and September have 5 week SRP’s, so a fixed amount of retail activity moving from a 4 week SRP, to a 5 week SRP would be expected to have a slightly different effect on the respective monthly figures.

Again the actual variables used are \( MBH_i - \overline{MBH} \), and \( ABH_i - \overline{ABH} \), where \( \overline{MBH} \) and \( \overline{ABH} \) are the long run monthly means computed over the interval 1992-2019.

**Easter**

Since Easter variation in the series is dependent on the movement of Easter between standard recording periods rather than calendar months, the pre-defined regressors in X12ARIMA are not appropriate.

The date of Easter (Sunday) varies between March 22 and April 25. The start date of the April standard recording period (also a Sunday) moves between 1st April and 6th April. In terms of the position of Easter relative to the boundary between March and April standard recording periods, there are six possible occurrences of Easter. These are:

- **Type 1.** Easter occurs 2 weeks before the start of the April SRP
- **Type 2.** Easter occurs 1 weeks before the start of the April SRP
- **Type 3.** Easter coincides with the start of the April SRP
- **Type 4.** Easter is 1 week after the start of the April SRP
- **Type 5.** Easter is 2 weeks after the start of the April SRP
- **Type 6.** Easter is 3 weeks after the start of the April SRP
Easters of the type 1 and 6 occur rarely, so there is little or no information about their effect. Easters of type 1 are likely to be very similar to Easters of type 2. This is because whether Easter is 1 or 2 weeks before the end of the SRP, in each case the build up period before Easter, and the few days immediately after Easter (when sales are likely to be affected) is entirely within the March SRP. Similarly, Easters of type 6 are likely to be very similar to Easters of type 5. For these reasons no distinction is drawn between type 1 and 2 Easters and type 5 and 6 Easters. Easters are thus classified in the following way:

Type 1 or 2. Easter occurs 1 or 2 weeks before the start of the April SRP
Type 3. Easter coincides with the start of the April SRP
Type 4. Easter is 1 week after the start of the April SRP
Type 5 or 6. Easter is 2 or 3 weeks after the start of the April SRP

The standard approach used in X12ARIMA, assumes the affect of Easter, is an increase (or decrease) in sales, of a fixed amount (to be estimated), for a build up period of $\omega$ days before Easter ($\omega$ to be chosen by the user). This relatively simple formulation allows the Easter effect to be estimated using one parameter.

The Easter effect in retail sales data varies in nature and magnitude across industry groups, and is more complicated than this. However, since it is relatively strong, it is easy to identify with some confidence, and a more sophisticated treatment is possible. Each of the four kinds of Easter is effectively treated separately e.g. only information about Type 3 Easters is used to estimate their effect. No assumptions are made about relationships between the effects of Easters of different kinds. A set of 3 contrast variables are used to model the Easter effect. Defining the four distinct classes of Easter as $i = 1, 2, 3, 4$ – they model the switch in activity that occurs for Easters $i = 1, 2, 3$ compared to when $i = 4$.

$E_\omega$ where $i = 1, 2, 3$ and $t = 1, \ldots, T$

In March in years where Easter is of class $i$
In April in years where Easter is of class $i$
otherwise
\[ E_u = \begin{cases} 
0.8 \\
-1 \\
0 
\end{cases} \]

Again the actual Easter variable used is \( E_u - \bar{E}_u \) where the \( \bar{E}_u \) are the long run monthly means computed over the interval 1986-2019 (non zero only in March and April).

**Model Building**

It was decided that because statistically significant phase shift effects occurred only rarely in months outside the period from October to February inclusive, it would be assumed that phase shift effects were zero in these months - they cannot be reliably estimated.

One advantage of this approach is it means that there are not adjustments for two different things in the same month e.g. adjustments for a phase shift and Easter in April. This avoids problems of multi colinearity between regressors, and also makes the prior adjustments which are estimated, much more interpretable. A single reason can be given for every adjustment:

<table>
<thead>
<tr>
<th>MONTH</th>
<th>REASON FOR ADJUSTMENT</th>
</tr>
</thead>
<tbody>
<tr>
<td>January</td>
<td>Phase shift</td>
</tr>
<tr>
<td>February</td>
<td>Phase shift</td>
</tr>
<tr>
<td>March</td>
<td>Easter</td>
</tr>
<tr>
<td>April</td>
<td>Easter</td>
</tr>
<tr>
<td>May</td>
<td>Late May Bank Holiday</td>
</tr>
<tr>
<td>June</td>
<td>Late May Bank Holiday</td>
</tr>
<tr>
<td>July</td>
<td>No adjustment</td>
</tr>
<tr>
<td>August</td>
<td>August Bank Holiday</td>
</tr>
<tr>
<td>September</td>
<td>August Bank Holiday</td>
</tr>
<tr>
<td>October</td>
<td>Phase shift</td>
</tr>
<tr>
<td>November</td>
<td>Phase shift</td>
</tr>
<tr>
<td>December</td>
<td>Phase shift</td>
</tr>
</tbody>
</table>
Seasonal adjustment

Seasonal adjustment decomposes a series into trend, seasonal and irregular components, and removes the seasonal part. In line with good practice, we have reviewed and updated the seasonal adjustment parameters. This takes account of recent data, including the revisions from January 2000 as a result of rebasing and the move to ratio estimation.

As part of the review of seasonal adjustment we have reviewed the length of the moving average used to identify the seasonal component. The current RSI seasonal adjustment uses a very short moving average. This tends to produce a smoother seasonally adjusted series (as part of the irregular is removed at the same time) but is liable to greater revision in the light of later data. Following advice from ONS methodologists, the new Retail Sales Index now uses a longer moving average in identifying the seasonal component, which will lead to a more volatile seasonally adjusted series - but one that should be revised less when additional monthly data becomes available.

Other changes in methods and approach

Trading day adjustment

Most RSI respondents report data on a standard 4-4-5 week cycle. The 30% or so of respondents who report on a different basis are asked to record the period they are reporting sales for on their questionnaire. Adjustments are made at the beginning of the RSI processing system to bring these non-standard reported sales into line with a standard reporting period. These adjustments depend on day of week and proximity to Easter and bank holidays of the days reported. The current weights have been in place for many years. We have reviewed and re-estimated the weights based on up to date information.

The new weights are based on credit card data. The credit card data comes from two sources: the Association for Payment Clearing Services (APACS) and Streamline. APACS is a trade association of banks and building societies. APACS provided information on average daily spending for each year from 1997 onwards.
Streamline provided daily credit card spend from August 2001. Both APACS and Streamline provided data for 9 types of retailer that closely match the RSI breakdown by type of retailer.

We have used to APACS data to derive daily weights (the average proportion of weekly trading that occurs on each day of the week). We have used Streamline data to adjust the daily weights as appropriate to account for different spending patterns around bank holidays.

**Linking**

We have introduced the improvements to the RSI from January 2000 onwards. We have linked on the existing series prior to 2000 to give a consistent time series. Q1 2000 has been chosen as the link year. The linking method is standard and will maintain the 1995 based RSI growth rates prior to 2000, resulting in a loss in additivity in the historical estimates. The linking method will also revise growth rates from 2000.