# Revisions to the national accounts: nominal, real and price effects[1]

Corné van Walbeek and Evelyne Nyokangi

**ABSTRACT**

Growth rates in the national accounts are published by the South African Reserve Bank's Quarterly Bulletin on a quarterly basis. This paper considers the magnitude of official revisions in quarterly nominal and real GDP and GDE growth rates, as well as the magnitude of the revisions in the growth rates of the implied price deflators, for the period 1986 to 2010. The revisions in the growth rates are substantial. The standard deviation of the difference between the first release growth rate and the growth rate as recorded in the SARB online database is 5.4 percentage points for nominal GDP growth, 1.4 percentage points for real GDP growth, 8.8 percentage points for nominal GDE growth, and 6.8 percentage points for real GDE growth for the period as whole. The revisions were particularly large in the unstable 1986-1995 period. Subsequently the magnitude of the revisions have decreased somewhat.

The paper also considers the magnitude of the revisions between the 1st and 2nd release, between the 2nd and 4th fourth release, between the 4th and 8th release, between the 8th and the 12th release and between the 12th release and the SARB database. Other than real GDP growth, a substantial proportion of the difference between the first release data and the SARB database data is ascribed to revisions made after the 12th release. This suggests that these late changes are due to systemic changes to the data (e.g. changes in the seasonal adjustment factors) rather than to the availability of new data.

The results of the paper suggest that researchers should be careful in the use of time-series data, because the experienced reality (as reflected in the first release data) may be quite different from the data on which time series analysis is based.

## 1. INTRODUCTION

When China published its 2013 first quarter growth rate in April 2013 it was met with disappointment in the market. The year-on-year growth rate of 7.7% was below the market expectation of 8%. As a result, the World Bank revised China's growth forecast for the whole year down by 0.1 percentage points to 8.3% (Reuters, 2013), and the S&P 500 index in the US dropped by 2.3% on the day. This was the largest one-day fall in more than five months, and was led by a decrease of nearly 4% in energy and raw material companies (Hwang, 2013).

This anecdote illustrates the importance and media coverage that is attached to the first release of macroeconomic data. The media writes about these data as though they are cast in stone and

[1] Paper presented at the Biennial Conference of the Economic Society of South Africa, Bloemfontein, 25-27 September 2013.

absolutely accurate. Even small deviations from the "market consensus" result in significant movements on the stock market. The fact of the matter is that first releases of macroeconomic data are based on a very limited view of reality and can change substantially as the statistical authorities collect more data. In fact, the Chinese GDP growth rate was published 15 only days after the end of the quarter to which it refers.

In an ideal world published macroeconomic data are both timely and accurate. In the real world there is a trade-off between timeliness and accuracy. For South Africa's national accounts the first release is typically published within two months after the end of the relevant quarter. In terms of its policy impact, and its impact on the financial markets, it is by far the most important data release. Over time, the data are revised as the statistical authorities gather more information, and the revised data would thus be a more accurate reflection of the real state of affairs in previous years. However, the policy relevance of such data is practically zero, because time has moved on.

When academics use time series data to estimate relationships or reaction functions, or simply to present a descriptive analysis about the economy, they typically use data from an established database, for example from the SARB's online download facility, henceforth called the SARB database (SARB, undated). These data are of different vintages. The latest observation is usually the first release data. Data that refers to older periods have been officially revised (often many times), may have had the base years changed, and may have been seasonally adjusted using different techniques and/or seasonal weights. To the extent that the two sets of data are different, the econometric results, and possible even the policy conclusions, of research based on time series data from the database could be quite different than had the analysis been based on first release data.

Orphanides (2001) illustrated this problem in the context of a simple monetary policy rule in the US. He concluded that "real-time policy recommendations differ considerably from those obtained with ex post revised data" and that "estimated policy reaction functions based on ex post revised data provide misleading descriptions of historical policy and obscure the behaviour suggested by information available to the Federal Reserve in real time" (Orphanides, 2001: 964). Van Walbeek (2006) considers South African national accounting data and finds substantial differences between the first release data and the SARB database. For example, the standard deviation of the difference between the first release real GDP growth rate and the SARB database growth rate is 1.8 percentage *points*; for real GDE the standard deviation of the difference in the two data sources is 5.7 percentage *points* (based on 80 quarters, 1984-2003). What this means in practice is that if the first release of GDE growth is 5%, there is a more than 30% probability that it would be greater than 10.7% or less than -0.7% in the SARB database in a number of years' time. The deviations between the first release and the SARB database for imports, exports, government expenditure and gross fixed capital formation were found to be even larger. Using a simple import function and a simple consumption function as examples, he then illustrated that the choice of data vintage has a very substantial impact on the size of the various coefficients.

The current study aims to expand on the original study in three ways. Firstly, the current study considers a longer period. Whereas the original study considered 80 quarters, the current study considers 100 quarters (1986-2010). Secondly, whereas the original study restricted its focus to the growth in *real* GDP and its expenditure components, the current study also considers the growth in

*nominal* GDP and the growth in the implicit deflators. Thirdly, the current study considers the dynamics of the official revisions. Rather than simply comparing differences in the growth rate between the SARB database and the first release, the current study considers differences between the 1st and 2nd release, between the 2nd and 4th fourth release, between the 4th and 8th release, between the 8th and the 12th release and between the 12th release and the SARB database

The current study is narrower than the previous study in the sense that we focus only on GDP and GDE, whereas the previous study focused on GDP and all its expenditure components. However, the data for other expenditure components are available.

The primary aim of the paper is to raise awareness of the magnitude of the official data revisions, and to show that data extracted from the SARB database differs quite substantially from first release data. Time series econometricians who use the SARB database to estimate relationships may find weaker or stronger relationships if they were to use data of a different vintage. As such the paper is a call to caution about the use of readily available database data, which may bear little resemblance to the first release data. It is hoped that future studies will use different vintages of data to determine the sensitivity of the results to changes in data vintages.

## 2. DATA

In the National Accounts section of the SARB Quarterly Bulletin data on GDP and its expenditure components are presented in a number of formats: (1) annually, in current prices, (2) annually, in constant prices, (3) quarterly, in current prices (not seasonally adjusted) , (4) quarterly, in current prices, in seasonally adjusted annualised rates, (5) quarterly, in constant prices (not seasonally adjusted), and (6) quarterly, in constant prices, in seasonally adjusted annualised rates. In the "Key information" section, the SARB publishes the percentage change in real GDP and its expenditure components in two formats: (1) year-on-year annual changes (i.e. based on (2) above), and (2) quarterly changes, based on seasonally adjusted annualised rates (i.e. based on (6) above). The quarterly percentage changes are annualised using a compound growth formula, i.e. growth rate = $[(Y_t/Y_{t-1})^4-1] \times 100$, where $Y_t$ refers to the current quarter and $Y_{t-1}$ to the previous quarter. Of the eight data series published in the Quarterly Bulletin, three are in current prices, while five are in constant prices.

The SARB does not publish the implicit price deflators but these can easily be calculated by dividing the nominal value by the real value of the magnitude of interest. In line with most market watchers and policy makers, the focus of this paper is not on the levels of national accounting aggregates or the level of the implicit price deflator, but on the percentage change in each of these. The percentage change in the implicit price deflator is approximately the percentage change in the nominal variable less the percentage change in the real variable. For small and even modest changes in the real and nominal variables this approximation is close.[2]

[2] For example, if the nominal variable increases by 10% and the real variable by 4%, then the approximated increase in the price deflator is 6% [= 10% - 4%], whereas the precise increase in the price deflator would be 5.77% [= 1.10/1.04 – 1) x 100].

This study considers quarterly data only. The growth rates of the variables in constant prices are taken directly from the "Key information" section of the SARB Quarterly Bulletin. Since growth rates of the variables in current prices are not available in the SARB Quarterly Bulletin, or in the online database, they were calculated from the seasonally adjusted annualised series (post-scripted "L" in the Quarterly Bulletin), as quarter-on-quarter annualised rates.

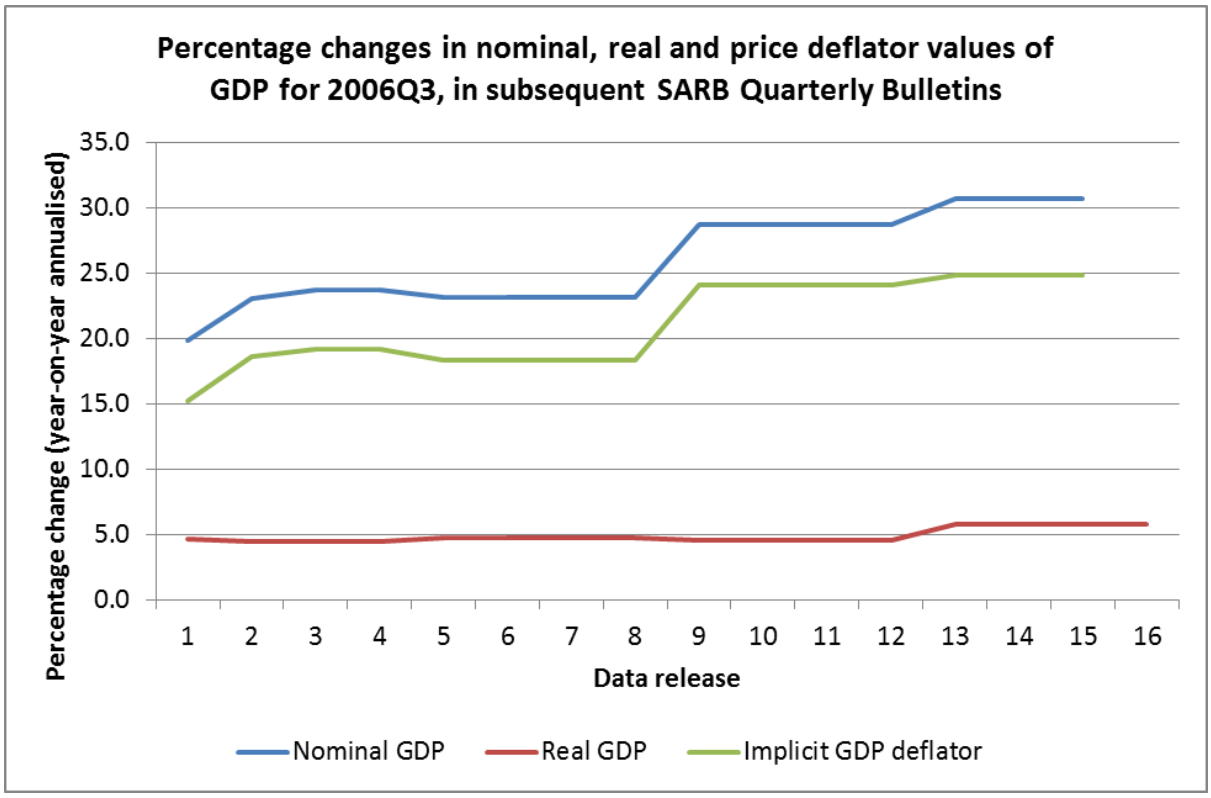
Furthermore, only GDP and GDE are considered in this study. These are the two most comprehensive and well-known quantities that consider aggregate production and aggregate domestic expenditure, respectively. The data for other expenditure components have been captured, but are not presented here mainly for space reasons. The analysis is limited to the period 1986-2010. The reason why we do not consider data subsequent to 2010 is because the data have not yet "settled down". It would create a false impression that the first release data is quite closely correlated with the SARB database data. The data from the SARB database was extracted in August 2013.

Currently the SARB publishes nominal GDP and GDE data (in level terms) for the 16 most recent quarters, and real growth rates for the 20 most recent quarters. Before September 1992, these data were published for only the eight most recent quarters. We captured these data from individual Quarterly Bulletins, and calculated the annualised quarter-on-quarter growth rates in the nominal values for each data release. With 16 data points for each nominal time series, we calculated 15 quarter-on-quarter growth rates.

To illustrate the calculations of the nominal growth rates and how the numbers can change from one release to the next, consider the Quarterly Bulletin of December 2006. Nominal GDP for 2006Q3 (series 6006L) was shown as R1 747 736 million, while nominal GDP for 2006Q2 was R1 670 206 million. Based on these numbers the quarter-on quarter annualised growth rate for 2006Q3 is calculated as 19.9% [$\{(1747736/1670206)^4 - 1\} \times 100$]. The growth rate of real GDP in 2006Q3 was indicated in the "Key information" as 4.7%. The difference between the nominal growth rate and the real growth rate approximates the percentage change in the implicit GDP price deflator (i.e. 19.9% – 4.7% = 15.2%).

In the March 2007 Quarterly Bulletin the nominal numbers for 2006Q3 (and 2006Q2) were revised, such that the second release growth rate for nominal GDP was 23.1% [$\{(1765303/1675959)^4 - 1\} \times 100$], compared to the first release of 19.9%. The real GDP growth rate for 2006Q3 was adjusted down to 4.5% from 4.7% in the first release. By implication, the growth in the implicit GDP deflator increased from 15.2% in the first release to 18.6% in the second. Figure 1 illustrates the published growth rates in real and nominal GDP and the implicit price deflator for 2006Q3 in the various releases.

**Figure 1:** Percentage changes in nominal, real and price deflator values of GDP for 2006Q3, in subsequent SARB Quarterly Bulletins




Source: SARB Quarterly Bulletins, December 2006 to September 2010

## 3. TERMINOLOGY

If $Dn_t$ is the $n^{th}$ release of the data for period t, and $SARB_t$ is the SARB database value for period t, the revisions for period t are defined as follows: $R(x\text{-}y)_t = Dx_t - Dy_t$, with $x > y$. Where it is clear from the text that the subscript t plays no significant role, it is dropped. Section 4 focuses on the differences between the SARB database and the first release, i.e. R(SARB-1).

Section 5 considers the differences between the various releases, in particular R(2-1) (i.e. comparing the second release with the first release), R(4-2), R(8-4), R(12-8) and R(SARB-12). If R(x-y) is positive, the relevant growth rate of the later release is higher than that of the earlier release, and if R(x-y) is negative, the relevant growth rate of the later release is lower than the earlier release. As a convention for this paper, if the expression R(x-y) is used, it refers to one or more of R(2-1), R(4-2), R(8-4), R(12-8) or R(SARB-12), but not to R(SARB-1).

## 4. THE MAGNITUDE OF THE REVISIONS

The two most used data vintages are the first release, because this attracts most attention in the media and has the most policy impact, and the data that appears in the SARB database, because this is the data that is used by academics and researchers in time series analyses. The focus of this section is on the difference between the first release and the SARB database, i.e. R(SARB-1).

The period under investigation (1986-2010) is divided into three sub-periods. The first period (1986-1995) was one of low and volatile GDP growth and even more volatile GDE growth. Based on first release data, 14 of the 40 quarters experienced negative real GDP growth and 16 experienced negative real GDE growth. The second period (1996-2005) was characterised by more stable GDP growth, but volatile GDE growth. Three of the 40 quarters experienced negative real GDP growth (primarily in the 1998 recession), while 11 experienced negative real GDE growth (again based on first release data). The third period (2006-2010) was one of high growth followed by a deep recession in 2009. The recession saw three quarters of negative real GDP and GDE growth.

A number of descriptive statistics are presented below. The average deviation, i.e. R(SARB-1), summarises the bias in the revisions. A positive average implies an upward bias in the revisions, while a negative average implies a downward bias in the revisions. Tests of significance are also performed.

The standard deviation has the usual interpretation as a measure of dispersion.

We also counted the number of deviations (in absolute terms) that were larger than predetermined threshold values. The threshold values that were chosen were 2.5 percentage points and 5 percentage points, respectively.

The results for nominal and real GDP and GDE growth, and the respective deflators, are shown in Table 1.

**Table 1:** SARB growth rates less first release growth rates

| | Gross domestic product | | | Gross domestic expenditure | | |
|---|---|---|---|---|---|---|
| | Nominal growth | Real growth | Price deflator growth | Nominal growth | Real growth | Price deflator growth |
| Average | | | | | | |
| 1986-2010 | 1.97*** | 0.66*** | 1.30* | 1.42 | 0.29 | 1.13 |
| 1986-1995 | 2.12* | 0.16 | 1.97* | 1.11 | -1.00 | 2.11 |
| 1996-2005 | 1.73* | 1.21*** | 0.52 | 1.76 | 1.27 | 0.49 |
| 2006-2010 | 2.12 | 0.57** | 1.55 | 1.39 | 0.92 | 0.47 |
| Standard deviation | | | | | | |
| 1986-2010 | 5.39 | 1.42 | 5.37 | 8.82 | 6.77 | 7.17 |
| 1986-1995 | 6.14 | 1.69 | 6.04 | 12.26 | 9.66 | 9.78 |
| 1996-2005 | 4.36 | 1.17 | 4.29 | 5.92 | 4.16 | 5.03 |
| 2006-2010 | 5.91 | 0.85 | 5.97 | 4.99 | 2.40 | 4.04 |
| Deviation of more than 2.5 percentage points | | | | | | |
| 1986-2010 | 63/100 | 12/100 | 58/100 | 72/100 | 54/100 | 64/00 |
| 1986-1995 | 31/40 | 6/40 | 24/40 | 35/40 | 29/40 | 31/40 |
| 1996-2005 | 21/40 | 6/40 | 21/40 | 25/40 | 17/40 | 24/40 |
| 2006-2010 | 11/20 | 0/20 | 13/20 | 12/20 | 8/20 | 9/20 |
| Deviation of more than 5 percentage points | | | | | | |
| 1986-2010 | 37/100 | 0/100 | 33/100 | 50/100 | 33/100 | 40/100 |
| 1986-1995 | 18/40 | 0/40 | 16/40 | 25/40 | 23/40 | 24/40 |
| 1996-2005 | 12/40 | 0/40 | 10/40 | 19/40 | 9/40 | 11/40 |
| 2006-2010 | 7/20 | 0/20 | 7/20 | 6/20 | 1/20 | 5/20 |

Significance: *** = 1%, ** = 5%, * = 10%

With only one exception (real GDE for 1986-1995), the growth rates published in the SARB database are generally higher than the first releases, i.e. R(SARB-1) > 0. For *nominal* GDP growth, the data has been revised upwards by an average of nearly 2 percentage points for the full period. For the sub-periods, the upward revision is similar to the period as a whole. For *real* GDP growth, average R(SARB-1) was negligible in the 1986-1995 period, but increased sharply to 1.21 percentage points in 1996-2005, after which it decreased to 0.57 percentage points in 2006-2010.

The average R(SARB-1) for implicit GDP deflator growth is about 1.3 percentage points for the period as a whole, and it is positive for the three sub-periods. The first releases thus significantly understate the "inflation rate" of goods and services produced in South Africa.

A broadly similar picture follows for the growth in GDE. For *nominal* GDE growth the average R(SARB-1) is 1.4 percentage points for the period as whole, implying substantial upward revisions. For *real* GDE growth, the average R(SARB-1) is 0.3 percentage points for the whole period, but subject to significant differences between sub-periods. For the 1986-1995 period, the average R(SARB-1) is -1.0 percentage point. Thus the "true" performance of that period was even weaker than was thought at the time. For the subsequent period (1996-2010) the revisions have generally raised the real GDE growth rate.

The growth rate in the implicit GDE deflator has been substantially revised upwards over the period as a whole. The average R(SARB-1) for 1986-1995 is more than 2 percentage points, while for the subsequent period the average R(SARB-1) is just less than 0.5 percentage points.

The standard deviations in Table 1 indicate substantial differences in R(SARB-1) between nominal and real variables, and between GDP and GDE. For *nominal* GDP, the standard deviation of R(SARB-1) is 5.4 percentage points for the full period. For sub-periods, the standard deviation of R(SARB-1) is broadly comparable (between 4.4 and 6.1 percentage points). The standard deviation of R(SARB-1) for *real* GDP is about 1.4 percentage points for the whole period, but is somewhat lower in later sub-periods. The standard deviation of R(SARB-1) of nominal GDP is between 3 and 7 times larger than the standard deviation of R(SARB-1) of real GDP, depending on the sub-period.

The standard deviations of R(SARB-1) of nominal and real GDE growth are substantially higher than the standard deviations of R(SARB-1) of GDP growth. For 1986-2010 the standard deviation of R(SARB-1) of nominal GDE growth is 8.2 percentage points, while it is 6.8 percentage points for real GDE growth. The revisions in 1986-1995 were particularly pronounced, with a standard deviation of R(SARB-1) of 12.3 percentage points for nominal GDE growth and 9.7 percentage points for real GDE growth. For the subsequent periods they were lower, but still much higher than for the standard deviations of GDP growth rates over the same period.

In Figure 2 and Figure 3 the R(SARB-1) values are plotted for nominal growth, real growth and the change in the implicit price deflators, for GDP and GDE respectively. The same vertical scale is chosen, which allows one to compare the relative magnitudes of the revisions. A positive value implies an upward revision and a negative value implies a downward revision. The line graph indicates the R(SARB-1) values of the nominal growth rate. The black bars indicate the R(SARB-1) values of the real growth rates, while the grey bars indicate the R(SARB-1) values of the growth in the implicit price deflator.

**Figure 2:** Decomposition of the difference between the first release and the SARB database: Nominal GDP growth

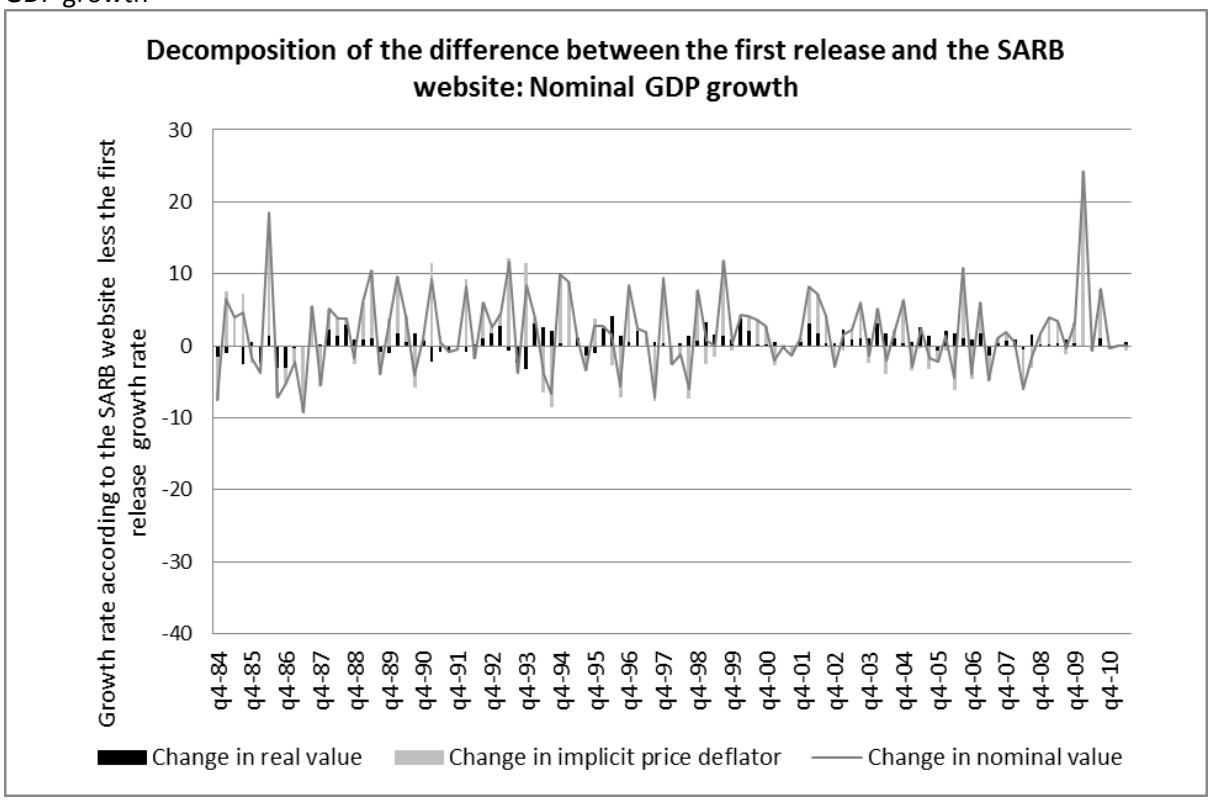
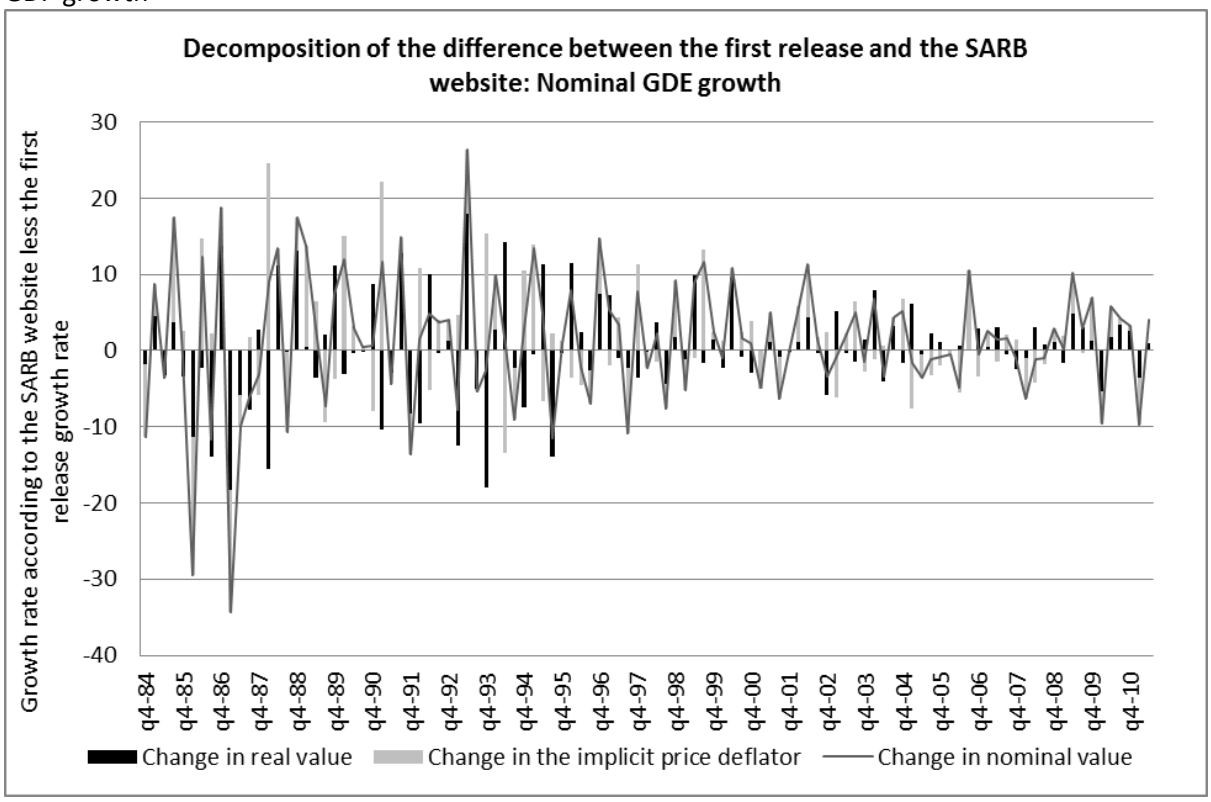



**Figure 3:** Decomposition of the difference between the first release and the SARB database: Nominal GDP growth



Two major issues stand out. Firstly, the revisions are substantially larger for GDE growth than for GDP growth. On the assumption that SARB database data is the most accurate, this implies that the statistical authorities have a better handle on the first release of aggregate production (GDP) data than on aggregate expenditure (GDE) data. The fact that GDP is measured in a number of ways (production, income and expenditure), allowing the statistical authorities to cross-reference the limited data for the first GDP release against other data sources, may explain the relatively "better" first release GDP growth data. The reasons for the large differences in the revisions between GDP and GDE growth is not investigated further in this study, but raises interesting future research opportunities.

Secondly, whereas the revisions in nominal GDP growth are driven primarily by revisions in the implicit GDP price deflator, and to a much lesser extent by revisions in real GDP growth, revisions in nominal GDE growth are attributed to a much larger extent to revisions in real GDE growth.

To further illustrate the magnitude of the differences between the first release and the SARB database, we calculated the number of times the R(SARB-1) values were larger than some arbitrarily chosen number (2.5 and 5 percentage points respectively). The results are shown in the bottom half of Table 1. For each of the sub-periods more than half the values(in absolute terms) of R(SARB-1) values for nominal GDP growth are larger than 2.5 percentage points, and about a third are larger than 5 percentage points. For real GDP the picture is somewhat better. Only 12 of the 100 quarters in the 1986-2010 period have experienced deviations of more than 2.5 percentage points, and none have experienced deviations of more than 5 percentage points.

On the other hand, the R(SARB-1) values for GDE growth are very substantial, both in nominal and real terms. Considering the whole period, 72 of the 100 R(SARB-1) values for *nominal* GDE growth exceed 2.5 percentage points in absolute terms, and 50 of the 100 values exceed 5 percentage points. For *real* GDE growth the pattern is only slightly better: 54 of the 100 deviations exceed 2.5 percentage points in absolute terms, and 33 of the 100 deviations exceed 5 percentage points. Deviations greater than 5 percentage points were concentrated in the unstable 1986-1995 period (23 of 40 quarters), but even in the 1996-2005 period nine of the 40 deviations were larger than 5 percentage points. In the 2006-2010 period only one of the 20 R(SARB-1) values exceed 5 percentage points, but this is likely to increase in years to come as the SARB database numbers are adjusted and revised further.

For the growth in the GDE deflator, 64 of the 100 R(SARB-1) values exceed 2.5 percentage points in absolute value, and 40 exceed 5 percentage points. While there is a somewhat greater concentration of large deviations in the unstable 1986-1995 period, even in the latest sub-period, nearly half (9/20) of R(SARB-1) absolute values are larger than 2.5 percentage points and a quarter (5/20) are larger than 5 percentage points.

## 5. DATA REVISIONS OVER TIME

In the previous section, we compared the SARB database values with the first releases, and it is clear that there are very substantial differences between the two data vintages. In this section we focus on how the data revisions have taken place over time. In other words we ask the following type of questions: Do the changes in the data take place quite quickly after the first release, or is there a

substantial time lag between them? Do the revisions sometimes go in one direction (e.g. a general increase in the growth rates), only to be reversed in a subsequent set of data releases? Et cetera.

In order to perform this analysis we consider the data releases at specific points in time (in fact the first, second, fourth, eighth and twelfth releases in the SARB Quarterly Bulletins, and the SARB database), and then calculate the differences in the appropriate differences in the growth rates between these various releases.

The descriptive statistics, similar to those in Table 1, are provided in Table 2 (for GDP) and Table 3 (for GDE) below. The average R(SARB-1) value is broken down into five subsections, as indicated at the top of the tables, and in principle R(2-1) + R(4-2) + R(8-4) + R(12-8)+R(SARB-12) must equal R(SARB-1). Where this is not the case, it is because of missing values of some of the data releases. This is particularly true for the period 1986-1995. In this 40-quarter period the SARB published data only eight quarters in arrears until 1992, after which the period was extended to 16 quarters (for nominal values) and 20 quarters (for real growth rates). As a result, the average R(SARB-12), R(12-8) and R(8-4) values (the latter only for the nominal values) are available only for a subsample of the 40 quarters. Care is thus required in the arithmetic of these tables, although the general discussion does hold.

If the values of R(2-1), R(4-2),… R(SARB-12) have the same sign, it implies that the adjustment from one vintage to the next is generally in the same direction (for the appropriate period). In most cases, especially after 1996, the values of R(x-y) are positive, indicating that the statistical authorities have generally revised the growth rates upwards. On the other hand, if the values of R(x-y) sub-periods have differing signs, it means that the statistical authorities first adjusted the growth rates in one direction, and subsequently adjusted them in the opposite direction. For example, for real GDP growth for 1986-1995 the average adjustment between the first and the second release (i.e. R(2-1)) was slightly up (by 0.1 percentage points), followed by a slight downward revision between the second and the fourth release (of 0.07 percentage points), followed by a sizeable and statistically significant upward revision between the fourth and the eighth release (of 0.35 percentage points). Later releases again adjusted the real GDP growth rate down.[3]

The relative magnitudes of the average deviations between the releases of interest are shown in the second part of each table. The average R(x-y) value for each sub-period is expressed as a percentage of the sum of the average R(x-y) values for the five sub-periods. Thus, for nominal GDP growth for the period 1986-2010, 35% of the average difference between the first release and the SARB database is accounted for in the revision between the first and second release (i.e. R(2-1)). Subsequent revisions to the data have a relatively minor effect, but the revision from the twelfth release to the SARB database explains 43% of the difference between the first release and the SARB database data.

For the turbulent 1986-1995 period most (73%) of the average R(SARB-1) change in the nominal GDP growth rate is explained by changes in the data between the twelfth release and the SARB database.

[3] As mentioned previously, the sum of the R(x-y) is not equal to R(SARB-1) because R(12-8) and R(SARB-12) are based on 26 observations, while R(2-1), R(4-2), R(8-4) and R(SARB-1) are based on 40 observations.

This probably suggests that the statistical authorities have made systemic changes to the data, e.g. changing the seasonal adjustment factors, and the retrospective inclusion or exclusion of certain sectors in the GDP definition. It seems unlikely that after three years suddenly the statistical authorities found better input data. On the other hand, for the 2006-2010 period 88% of the average R(SARB-1) change in the nominal GDP growth occurs between the first and the second release, i.e. R(2-1), and another 18% occurs between the second and the fourth release, i.e. R(4-2). This is not a completely surprising result either. Given the shortness of the time period between the actual activity and the publication of the data on the SARB database (3-7 years) one would not expect major changes in the growth rates after the twelfth release.

For real GDP growth there is a substantial and significant increase in the published growth rates between the fourth and the eighth release (i.e. R(8-4)), for 1986-2010 and also for the sub-periods. 52% of the average R(SARB-1) value for real GDP growth is explained by revisions between the fourth and the eighth releases (for 1986-2010), while only 12% of average R(SARB-1) is explained by changes after the twelfth release. This suggests that the statistical authorities do revise the real GDP growth rates in response to better data gleaned from the economy, rather than changing the growth rates long after the fact.

For nominal and real GDE growth (Table 3) an inconsistent picture emerges. Other than in 2006-2010, most of the change in the nominal GDE growth rate between the first release data and the SARB database data occurs in the period after the twelfth release. This suggests large systemic changes and thus moves the SARB database data further away from the experienced reality at the time. Descriptive statistics of South Africa's GDE growth experience, or inputs into an econometric model, based on database data, may thus provide a highly distorted view of what was actually happening at the time.

For real GDE growth the data revisions do not follow a consistent upward or downward trend over time. Decreases in the published real GDE growth rates in one set of releases are followed by increases in the published growth rates in subsequent releases, and vice versa. Real GDE growth rates in the second release are often lower than growth rates in the first release (i.e. R(2-1) is negative), but growth rates in the fourth release are typically higher than in the second release.

Also, a large proportion of the R(SARB-1) value is explained by the revisions made after the twelfth release. It thus seems that revisions of real GDE growth are more likely to be systemic than revisions of real GDP growth, since they take place so much longer after the release of the first data.

The standard deviations of the R(x-y) differences and the proportion of R(x-y) differences that are greater than some arbitrary thresholds (1 and 2.5 percentage points respectively) clearly indicate that the GDE growth revisions in the various sub-periods are much larger than the comparable GDP growth revisions.

**Table 2:** Decomposition of the difference in GDP growth rate between the first release and the SARB database

| | Nominal GDP | | | | | | Real GDP | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Second less first R(2-1) | Fourth less second R(4-2) | Eighth less fourth R(8-4) | Twelfth less eighth R(12-8) | SARB less twelfth R(SARB-12) | SARB less first R(SARB-1) | Second less first R(2-1) | Fourth less second R(4-2) | Eighth less fourth R(8-4) | Twelfth less eighth R(12-8) | SARB less twelfth R(SARB-12) | SARB less first R(SARB-1) |
| Average | | | | | | | | | | | | |
| 1986-2010 | 0.69* | 0.01 | 0.25 | 0.18 | 0.86* | 1.97*** | 0.07 | 0.07 | 0.39*** | 0.12* | 0.09 | 0.66*** |
| 1986-1995 | 0.28 | 0.18 | -0.25 | 0.38 | 1.58 | 2.12* | 0.10 | -0.07 | 0.35*** | -0.05 | -0.08 | 0.16 |
| 1996-2005 | 0.45 | -0.35 | 0.64 | 0.24 | 0.74 | 1.73* | 0.08 | 0.21* | 0.50*** | 0.25** | 0.18 | 1.21*** |
| 2006-2010 | 1.99 | 0.40 | 0.02 | -0.21 | 0.07 | 2.12 | 0.02 | 0.07 | 0.26 | 0.08 | 0.14 | 0.57** |
| Average (as a percentage of the total deviation) | | | | | | | | | | | | |
| 1986-2010 | 35 | 1 | 13 | 9 | 43 | 100 | 10 | 10 | 52 | 16 | 12 | 100 |
| 1986-1995 | 13 | 8 | -12 | 18 | 73 | 100 | 39 | -27 | 141 | -19 | -35 | 100 |
| 1996-2005 | 26 | -20 | 37 | 14 | 43 | 100 | 6 | 17 | 41 | 20 | 14 | 100 |
| 2006-2010 | 88 | 18 | 1 | -9 | 3 | 100 | 4 | 12 | 45 | 15 | 24 | 100 |
| Standard deviation | | | | | | | | | | | | |
| 1986-2010 | 3.33 | 2.56 | 2.81 | 2.35 | 3.37 | 5.39 | 0.41 | 0.46 | 0.58 | 0.48 | 0.78 | 1.42 |
| 1986-1995 | 2.99 | 3.07 | 3.10 | 2.12 | 5.07 | 6.14 | 0.56 | 0.43 | 0.50 | 0.27 | 1.00 | 1.69 |
| 1996-2005 | 1.75 | 2.22 | 2.60 | 2.51 | 2.48 | 4.36 | 0.32 | 0.47 | 0.62 | 0.57 | 0.75 | 1.17 |
| 2006-2010 | 5.56 | 2.07 | 2.92 | 2.34 | 1.40 | 5.91 | 0.19 | 0.44 | 0.65 | 0.44 | 0.40 | 0.85 |
| Deviation more than 1 percentage point | | | | | | | | | | | | |
| 1986-2010 | 45/100 | 47/100 | 55/82 | 37/80 | 58/84 | 91/100 | 4/100 | 8/100 | 15/100 | 4/84 | 14/84 | 45/100 |
| 1986-1995 | 21/40 | 18/40 | 13/22 | 9/22 | 21/26 | 37/40 | 3/40 | 3/40 | 3/40 | 0/26 | 6/26 | 20/40 |
| 1996-2005 | 13/40 | 19/40 | 30/40 | 20/40 | 32/40 | 36/40 | 1/40 | 4/40 | 8/40 | 4/40 | 7/40 | 19/40 |
| 2006-2010 | 11/20 | 10/20 | 12/20 | 8/18 | 5/18 | 18/20 | 0/20 | 1/20 | 4/20 | 0/18 | 1/18 | 6/20 |
| Deviation more than 2.5 percentage points | | | | | | | | | | | | |
| 1986-2010 | 25/100 | 26/100 | 26/82 | 29/80 | 29/84 | 63/100 | 0/100 | 0/100 | 0/100 | 0/84 | 0/84 | 12/100 |
| 1986-1995 | 11/40 | 14/40 | 6/22 | 6/22 | 15/26 | 31/40 | 0/40 | 0/40 | 0/40 | 0/26 | 0/26 | 6/40 |
| 1996-2005 | 8/40 | 10/40 | 13/40 | 10/40 | 12/40 | 21/40 | 0/40 | 0/40 | 0/40 | 0/40 | 0/40 | 6/40 |
| 2006-2010 | 6/20 | 2/20 | 7/20 | 3/18 | 2/18 | 11/20 | 0/20 | 0/20 | 0/20 | 0/18 | 0/18 | 0/20 |

**Table 3:** Decomposition of the difference in GDE growth rate between the first release and the SARB database

| | Nominal GDE | | | | | | Real GDE | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Second less first R(2-1) | Fourth less second R(4-2) | Eighth less fourth R(8-4) | Twelfth less eighth R(12-8) | SARB less twelfth R(SARB-12) | SARB less first R(SARB-1) | Second less first R(2-1) | Fourth less second R(4-2) | Eighth less fourth R(8-4) | Twelfth less eighth R(12-8) | SARB less twelfth R(SARB-12) | SARB less first R(SARB-1) |
| Average | | | | | | | | | | | | |
| 1986-2010 | 0.14 | 0.09 | 0.22 | 0.21 | 0.74 | 1.42 | -0.33 | 0.60** | 0.08 | -0.01 | 0.32 | 0.29 |
| 1986-1995 | -0.54 | 0.28 | -0.37 | 0.20 | 1.25 | 1.11 | -0.60 | 0.42 | 0.26 | -0.08 | -0.31 | -1.00 |
| 1996-2005 | 0.36 | -0.33 | 0.65 | 0.29 | 0.79 | 1.76 | -0.24 | 0.70 | 0.32 | 0.09 | 0.41 | 1.27 |
| 2006-2010 | 1.05 | 0.55 | 0.03 | 0.04 | -0.13 | 1.39 | 0.06 | 0.75* | -0.78 | -0.12 | 1.06 | 0.89 |
| Average (as a percentage of the total deviation) | | | | | | | | | | | | |
| 1986-2010 | 10 | 6 | 16 | 15 | 53 | 100 | -49 | 90 | 12 | -1 | 49 | 100 |
| 1986-1995 | -67 | 35 | -46 | 24 | 155 | 100 | 193 | -136 | -84 | 26 | 101 | 100 |
| 1996-2005 | 21 | -19 | 37 | 16 | 45 | 100 | -19 | 55 | 25 | 7 | 32 | 100 |
| 2006-2010 | 68 | 36 | 2 | 3 | -9 | 100 | 6 | 78 | -81 | -13 | 110 | 100 |
| Standard deviation | | | | | | | | | | | | |
| 1986-2010 | 3.93 | 2.89 | 3.74 | 2.69 | 5.71 | 8.82 | 1.75 | 2.16 | 3.03 | 2.48 | 4.96 | 6.95 |
| 1986-1995 | 5.33 | 3.47 | 4.05 | 2.32 | 9.15 | 12.26 | 2.21 | 2.61 | 2.72 | 1.36 | 7.81 | 9.66 |
| 1996-2005 | 2.16 | 2.50 | 3.46 | 2.87 | 3.54 | 5.92 | 1.58 | 1.89 | 2.30 | 1.67 | 2.94 | 4.16 |
| 2006-2010 | 3.23 | 2.33 | 4.00 | 2.85 | 2.43 | 4.99 | 0.71 | 1.66 | 4.57 | 4.58 | 3.07 | 4.29 |
| Deviation more than 1 percentage point | | | | | | | | | | | | |
| 1986-2010 | 46/100 | 59/100 | 63/82 | 46/80 | 65/84 | 88/100 | 26/100 | 36/100 | 58/100 | 30/84 | 54/84 | 76/100 |
| 1986-1995 | 22/40 | 20/40 | 18/22 | 11/22 | 25/26 | 37/40 | 9/40 | 15/40 | 21/40 | 8/26 | 23/26 | 33/40 |
| 1996-2005 | 15/40 | 25/40 | 34/40 | 23/40 | 32/40 | 35/40 | 13/40 | 14/40 | 25/40 | 16/40 | 24/40 | 31/40 |
| 2006-2010 | 9/20 | 14/20 | 11/20 | 12/18 | 8/18 | 16/20 | 4/20 | 7/20 | 12/20 | 6/18 | 7/18 | 12/20 |
| Deviation more than 2.5 percentage points | | | | | | | | | | | | |
| 1986-2010 | 28/100 | 32/100 | 36/82 | 24/80 | 47/84 | 72/100 | 10/100 | 17/100 | 27/100 | 11/84 | 33/84 | 56/100 |
| 1986-1995 | 12/40 | 14/40 | 10/22 | 5/22 | 24/26 | 35/40 | 5/40 | 9/40 | 11/40 | 3/26 | 20/26 | 29/40 |
| 1996-2005 | 11/40 | 12/40 | 19/40 | 12/40 | 18/40 | 25/40 | 5/40 | 7/40 | 8/40 | 3/40 | 10/40 | 17/40 |
| 2006-2010 | 5/20 | 6/20 | 7/20 | 7/18 | 5/18 | 12/20 | 0/20 | 1/20 | 8/20 | 5/18 | 3/18 | 10/20 |

*** = significant at 0.1%, ** = significant at 1% and * = significant at 5%

## 6. IMPLICATIONS AND CONCLUSION

This paper is exploratory and limited in focus. It provides a descriptive overview of official revisions of two important national accounting aggregates, namely GDP and GDE. These are the two most quoted measures of aggregate production and aggregate demand, respectively.

The paper adds to a growing international literature on real-time data analysis. In contrast to the data that is available in significantly revised form in time series databases many months or years after the fact, policy makers have to make decisions based on recently released, incomplete and often quite wrong data. It would be unfair for researchers to criticise policy makers retrospectively using data that has been through various rounds of revision and refinement, and that may be quite different to the data on which the initial policy decision was made.

Within the South African context this study expands on a study that considers the growth in real GDP and the expenditure components for the period 1985-2004 (Van Walbeek, 2006). The present study considers the growth in both real and nominal GDP and GDE.

The main findings are the following:

- Revisions in the growth rate of *nominal* GDP are between three and seven times larger than the revisions of the growth rate of *real* GDP on average, depending on the sub-period. The standard deviation of the difference between the first release of the nominal GDP growth rate and the SARB database is 5.4 percentage points. For the real GDP growth rate the equivalent number is 1.4 percentage points.

- Revisions in the nominal GDE growth rates (s.d. = 8.8 percentage points) are somewhat larger, on average, than revisions in the nominal GDP growth rates (s.d. = 5.4 percentage points). However, the revisions of the real GDE growth rates are substantially larger (s.d. = 6.8 percentage points) than the revisions of the real GDP growth rates (s.d. = 1.4 percentage points).

- The differences between the first release data and the SARB database data are particularly pronounced in the 1986-1995 period, decreasing in the 1996-2005 period, and even more so in the 2006-2010. It seems likely that the large changes in the data in the 1986-1995 period was a function of the general economic instability of the time. However, one cannot compare the 1986-95, 1996-2005 and 2006-2010 periods objectively, because it is likely that the later periods might be subject to still-to-be-affected data changes, which could increase the differences, relative to the first release.

- For nominal and real GDE and nominal GDP (but not real GDP) a significant proportion of the differences in the data between the first release and the SARB database is ascribed to changes in the data after the twelfth release. It seems likely that these are the results of systemic changes to the data (e.g. changes to the seasonal adjustment factors, retrospective definitional changes, and changes to the prices used), rather than new information that was made available.

The bottom line of this discussion is that revisions of the national accounts data are substantial. Revisions to the national accounts are unavoidable, given the desire to get the data into the public domain as quickly as possible. Most lecturers, when explaining the national accounts, indicate that the data gets "better" over time, since the statistical authorities use more comprehensive and more accurate data in calculating the relevant magnitude. Data from the SARB database should thus be most representative of the activity that it purports to measure. If that is true, then the deviations between the first release and the "true" values are worrisome.

First release data are watched very carefully by the financial markets, and even modest changes from the "consensus" view can result in significant moves on the market. What this paper has shown is that the "true" (i.e. SARB database) data can deviate very much from the first release. In fact the subsequent revisions of the data can easily dwarf the deviations from market expectation when the data was first released.

Implicit in the first release data is the caveat that this is the first attempt at measuring the appropriate variable, but this is often forgotten by the market players and commentators. It is hoped that this paper will help to put this back in the foreground.

This paper also has implications for economic policy, especially for monetary policy. During the 1990s, especially, the choice of target received a lot of attention in discussions about monetary policy. In fact, much was made about the choice of a "nominal anchor". The choice of nominal GDP as a target or anchor of monetary policy has never received much serious attention (see, e.g. Fisher, 1995, Taylor, 2000 and Rudebusch, 2002), amongst others, on the grounds that national accounting data are often revised. Had nominal GDP growth been chosen as the target of monetary policy, it would have been a moving target, as the data is revised from one release to the next. This would make it an inappropriate choice for an anchor of monetary policy. This study strongly supports the arguments against the use of nominal GDP as a "nominal anchor" for economic policy.

## 7. BIBLIOGRAPHY


Fisher, S, 1995. Central-bank independence revisited. *American Economic Review* 85(2): 201-206.

Hwang, I, 2013. S&P 500 falls most since November as Chinese growth slows. *Bloomberg Business Week*. Available at http://www.businessweek.com/news/2013-04-15/u-dot-s-dot-stock-index-futures-drop-as-chinese-growth-slows. (Accessed 3 September 2013)

Orphanides, A, 2001. Monetary policy rules based on real-time data. *American Economic Review* 91(4): 964-985.

Reuters, 2013. China growth risks in focus as first quarter data falls short. Available at http://www.reuters.com/article/2013/04/15/us-china-economy-gdp-idUSBRE93E01U20130415. (Accessed 3 September 2013).

Rudebusch, GD, 2002. Assessing nominal income rules for monetary policy with model and data uncertainty. *The Economic Journal* 112(479): 402-432. Available at http://onlinelibrary.wiley.com/doi/10.1111/1468-0297.00036/full

South African Reserve Bank, undated. *Online download facility*. Available at http://www.resbank.co.za/Research/Statistics/Pages/OnlineDownloadFacility.aspx. (Accessed 15 August 2013)

Taylor, J, 2000. Using monetary policy rules in emerging market economies. Unpublished paper. Available at http://www.stanford.edu/~johntayl/Onlinepaperscombinedbyyear/2001/Using_Monetary_Policy_Rules_in_Emerging_Market_Economies.pdf

Van Walbeek, CP, 2006. Official revisions to South African national accounts data: Magnitudes and implications, *South African Journal of Economics* 74(4): 745-765.