MEASURING CORE INFLATION IN SOUTH AFRICA

Stan du Plessis*, Gideon du Rand† & Kevin Kotzé‡

January 2015

Abstract

Measures of core inflation convey critical information about an economy. They have a direct effect on the policy-making process, particularly in inflation-targeting countries, and are utilized in forecasting and modelling exercises. In South Africa the price indices on which inflation is based have been subject to important structural breaks following changes to the underlying basket of goods and the methodology for constructing price indices. This paper seeks to identify a consistent measure of core inflation for South Africa using trimmed-means estimates, measures that exclude changes in food and energy prices, dynamic factor models and wavelet decompositions. After considering the forecasting ability of these measures, which provide an indication of expected second-round inflationary effects, traditional in-sample criteria were used for further comparative purposes. The results suggest that wavelet decompositions provide a useful measure of this critical variable.

JEL Classifications: C43, E31, E52

Keywords: wavelets, trimmed means, dynamic factor models, forecasting, core inflation.

*Department of Economics, University of Stellenbosch

†Department of Economics, University of Stellenbosch

‡*Corresponding author:* School of Economics, University of Cape Town (kevin.kotze@uct.ac.za)

1 Introduction

The inflationary process influences future economic activity through a number of channels. It directly affects the real rate of return and the expected rate of future inflation, which influences the cost of living, increases in nominal wages, grants, agreements, and long-term contracts. In addition, inflationary pressure imposes adjustment costs on the economy, lowering the information content of price changes, while adversely affecting the distribution of income. Bernanke and Gertler (1999) have also noted that inflation contributes towards increased volatility in asset prices, which imposes a direct cost on the economy. Hence, measures of inflationary pressure play a critical role in policy-making, particularly for central banks in inflation-targeting countries that are explicitly responsible for anchoring price levels.

When formulating policy, decision-makers usually focus on the persistent sources of inflationary pressures that are not influenced by temporary short-run price deviations from the underlying trend. This is because changes to policy often have lasting effects that may result in significant transition costs (Blinder, 1997). This measure of the underlying rate of inflation is often termed core inflation, and it is used by the South African Reserve Bank (SARB) to inform policy.[1] Furthermore, as core inflation should be less volatile than changes in the price level, by focusing on the trend in the process, the central bank would not need to make as many changes to interest rates. This would reduce the number of policy shocks, thus providing an environment that would facilitate stable economic growth.

It has also been suggested that measures of core inflation would assist in identifying the economic behaviour that is responsible for generating the underlying inflationary process, after excluding the noise in the process that is largely attributable to relative price shocks (Bryan and Cecchetti, 1994). This would imply that when using core inflation to inform policy, the central bank would not respond to the type of price shocks that are not under its control. For example, relative shortages in the supply of a particular commodity may result in a sudden (relative) price increase of that commodity, which is beyond the control of the central bank. If the measure of core inflation were to exclude the relative increase in the short-term price of this commodity, then the central bank would not respond to it.

In this paper, we consider the relative merits of some of the methods that have been developed to measure core inflation. These include various trimmed means estimates that allow for different forms of asymmetry, which have been used by many central banks (Cecchetti, 2009). An example of such a measure that has been derived for the South African economy is described in Blignaut *et al.* (2009). This measure was constructed from a subsample of the weighted disaggregated consumer prices, after the most volatile constituents have been removed. In addition, we also make use of measures that exclude food and energy prices, such as those discussed in Bryan and Cecchetti (1994) and Silver (2007).

Various modelling techniques have also been employed to provide alternative measures of core inflation. These include the use of dynamic factor models to derive an index that describes the common variation in the disaggregated data. Examples of dynamic factor models that have been used for measures of core inflation include, Cristadoro *et al.* (2005) and Giannone and Matheson (2007), among others. Finally, we also make use of the more recently developed wavelet transformations that are able to distinguish between the trend and short-term noise in the aggregated data. This computationally efficient method is able to preserve the informational content of time-series data despite the influence of nonstationary behaviour. Examples of measures of core inflation that have been derived by wavelet methods include Dowd *et al.* (2010) and Baqaee (2010).

---

[1] The Monetary Policy Review (2014), which is published by the South African Reserve Bank, contains a discussion on various measures of core inflation that exclude volatile items, such as food, non-alcoholic beverages, petrol and energy prices. Ruch and Bester (2013) have noted that several other official statements of the central bank make reference to measures of core inflation and the forecasts of this unobserved variable.

When making use of criteria for comparing these estimates of core inflation, there are largely two aspects of relevance. Firstly, Blinder (1997) emphasised the *out-of-sample* properties of these estimates, where core measures are evaluated on the basis of their ability to predict future inflation (which would imply that they should also provide an accurate measure of the *second-round* inflationary effects).[2] Since these measures of core inflation may also be used in various multivariate models for policy-making and forecasting purposes, this measure would need to provide insight into the evolution of the expected underlying inflationary process. To satisfy this criteria for the evaluation of measures of core inflation, we generate out-of-sample statistics for successive 1 to12 step ahead forecasts over the ten year period 2003M1-2012M12. These forecasts are generated with the aid of a Kalman filter in a state-space model that allows for time-varying parameters.

Then secondly, Bryan and Cecchetti (1994) and Cecchetti (2009) have noted that the *in-sample* statistics of core inflation measures are also of importance when evaluating their potential usefulness. In particular, these measures should provide an accurate estimate of the mean and a lower variance, when compared with the actual reported measure of aggregate inflation. In addition, they also note that the price indices, which may be generated from the measures of core inflation, should share a common stochastic trend with the reported aggregate consumer price index (CPI). To satisfy the aforesaid criteria, we also consider these in-sample statistics for the estimates that are derived from the respective models.

In addition to the computation and evaluation of various measures of core inflation, this paper also considers the properties of the South African price indices, and the effect of changes to the methodology that is used to compile these indices. While changes in the behaviour of economic agents necessitate changes to the way in which we measure economic activity, a change in the technique that is used to measure broad-based consumer prices may result in the imposition of a structural break in the data. When deriving estimates for South African core inflation we need to be cognisant of such an event, as it would affect the way in which certain methods may be applied, particularly when using disaggregated data or when applying these methods to year-on-year, as opposed to month-on-month, inflationary data.

The following section of this paper contains a discussion on the various methods that have been used to measure consumer price movements in South Africa. This section also considers the implications of changes to the methodology that is used to calculate price indices, when deriving various measures of core inflation. Section 3 contains a brief review of the existing measures that have been derived to measure core inflation in South Africa. Thereafter, in section 4 we provide an explanation of the trimmed means estimates, while section 5 contains a discussion relating to the measure that excludes food and energy prices. Section 6 describes the use of dynamic factor models that are used to measure core inflation and section 7 seeks to explain the use of wavelet methods for this purpose. Section 8 contains the comparative results of the out-of-sample and in-sample analyses, before section 9 concludes.

2 Constructing Price Indices for South Africa

Over the period January 1975 to December 2008, consumer price data was reported at a monthly frequency for 33 sub-categories.[3] The basket of goods remained consistent over this

[2]The core inflation estimate that is able to outperform other inflation forecasts would provide a better estimate of expected *second-round* effects, which are of importance to central banks. A recent statement by the Governor of the SARB alludes to this feature of the data: "Whilst the current rate of inflation has increased above the target range, we have not increased interest rates, since we believe that the present rise is transitory and will have little second-round effects." (Marcus, 2011).

[3]Consumer price data is available for earlier periods, but for these earlier periods, the data was reported for fewer sub-categories. For the purposes of this investigation, a starting date of 1975 gives us more than

period, but the weights were amended periodically to reflect changes in consumer preferences. Aron and Muellbauer (2004) suggest that the construction of South African price indices over this period was subject to several methodological inconsistencies and a general lack of transparency, where the construction of seasonally adjusted CPI involved the assimilation of various seasonally unadjusted weighted subcomponents. After attempting to recreate the index from the sub-categories, they suggested that this measure was constructed as a chained Laspeyres index, which compares the old basket of goods, q_0, at the respective previous prices, $P_{n,\,0}$, and current prices $P_{n,\,t}$. This would imply that the aggregate price index, for the N different sub-categories, has been constructed from

$$P_t = \frac{\sum_{n=1}^{N} (P_{n,t} \cdot q_0)}{\sum_{n=1}^{N} (P_{n,0} \cdot q_0)} \ , \qquad (1)$$

where level factors are used to ensure that the data from different base years are comparable.[4]

This measure of consumer prices was complemented by a second measure that was termed CPIX, which was provided over the period 1997M1 to 2008M12 by government's official reporting agency, Statistics South Africa (StatsSA). It effectively sought to exclude the effects of mortgage costs from the consumer price index, in order to avoid the self-referential impact of changes due to the central bank interest rate, which would influence the interest rate on mortgages and inflation. This measure of prices was used to determine a rate of inflation, which provided the initial target in the SARB's inflationary target framework (Van der Merwe, 2004). In certain respects, this measure is similar to those estimates of core inflation that have been derived by other central banks, which often exclude the effects of energy and food (Silver, 2007).

Following a review of the practices that were used to construct these indices by StatsSA, it was decided that a new methodology for calculating consumer inflation would be adopted from January 2008. This revision in methodology ensures that the current calculation is consistent with international practice, and a new basket has been used to reflect consumers' current preferences, which would have been influenced by changes in taste and technology over the last three decades. The new methodology is in accordance with the practice of the International Labour Organisation (ILO) through the application of the United Nations Statistical Division (UNSD) Classification of Individual Consumption by Purpose (COICOP).

This process was conducted in a highly transparent manner, and we now have a document that describes the methodology that is followed to construct this important index.[5] As a result, we can now confirm that the current practice for calculating inflation in South Africa

---

enough data.

[4]Note, that while Aron and Muellbauer (2004) suggests that the construction of the aggregate measure of total inflation prior to 2008 was derived from seasonally adjusted composite series, we found that when using a seasonal filter on the composites, the aggregate measure of total inflation was extremely close to that of the combined seasonally unadjusted weighted inflation rates, to the extent that there are no obvious gains from applying the seasonal filter. It is also worth noting that since the early 1990s, the new approach of using the weighted price indices approximates the reported measure of inflation more closely than the weighted inflation rates of the composites. Aron and Muellbauer (2004) also suggest that the use of a logged seasonally adjusted series may be appropriate; however, we found that the construction of such a series does not replicate that of reported inflation. This may in part be due to the different sample that we have used in this study.

[5]See the website http://www.statssa.gov.za/cpi/index.asp for a list of articles that have been published by Statistics South Africa regarding the change in methodology. Much of what is contained in this discussion relates to the documents: "Shopping for two: The CPI new basket parallel survey - results and comparisons with published CPI data" (Statistics South Africa, 2009) and "The South African CPI Sources and Methods Manual" (Statistics South Africa, 2009).

makes use of a Jevons Index for the prices of disaggregated goods and services,

$$P_{n,\,t} = \prod_{k=1}^{K} \left(\frac{p_{k,\,t}}{p_{k,\,t-1}} \right)^{\frac{1}{K}} \quad , \tag{2}$$

where $p_{k,\,t}$ is the price of a particular good that may be found on the shelf of a particular store, at a particular point in time. The aggregate price of these goods across all stores and regions within a particular sub-category is then given as $P_{n,\,t}$.[6] The price indices for the respective goods are then multiplied by the basket weights to compile a Lowe's index, which may be interpreted as a modified Laysperes Index, for the aggregate CPI,

$$CPI_t = \sum_{n=1}^{N} \left(P_{n,\,t} \cdot q_n^m \right) \ . \tag{3}$$

where the weights of the basket items, q_n^m, are derived from information that is contained in the new Income and Expenditure Survey (IES).[7] This measure of consumer prices may be used to derive the month-on-month rate of headline inflation that is calculated as a simple growth rate,

$$\pi_t = \left(\frac{CPI_t}{CPI_{t-1}} \right) - 1 \ . \tag{4}$$

For the purpose of this analysis, it is worth noting that it is possible to arrive at a similar result for reported *headline inflation* by calculating the rate of inflation for each of the underlying price indices and multiplying them by their respective basket weights,[8]

$$\tilde{\pi}_t = \sum_{n=1}^{N} \left(\pi_{n,\,t} \cdot q_n^m \right) \ . \tag{5}$$

When we compare the measures of aggregate consumer inflation, which are provided by equations (4) and (5), there is a slight difference which reaches a year-on-year maximum of 2.8% when comparing the old measure of inflation, while a maximum discrepancy of 1.9% is found for the new measure of inflation.

In addition to establishing a clear methodological framework, this measure of headline inflation makes use of a larger number of sub-categories price series that are reported by StatsSA. For example, there are 33 reported sub-categories for the older measure of inflation, while there are 44 reported sub-categories for the new measure over the sample that has been used in this paper. From January 2013, the number of sub-categories items increased to 45. The statistics agency is also no longer reliant on retailers to provide information on prices, as it currently employs field teams to collect this information.

Following the introduction of the new measure of consumer price indices, CPIX was no longer calculated and the more encompassing measure, which is termed *headline inflation*, has been used as the target for monetary policy. This may partially be attributed to the fact

[6] For example, $p_{k,t}$ might refer to the price of a 500g package of sugar for a particular brand, while $P_{n,t}$ refers to the price of sub-category for the index of all sugar prices.

[7] The notation q_n^m is used to signify that the new methodology for calculating consumer prices makes use of different information to determine the respective basket weights, when compared to the basket weights that were calculated under the previous methodology (which are denoted q_0).

[8] This is of particular importance for when we start calculating trimmed means, dynamic factor models and the estimates that exclude food and energy prices.

that interest rates are also no longer used as an indicator of housing costs, as this measure has been replaced by a measure of owners' equivalent rent.[9]

During the implementation phase of this project, data was collected according to both methodologies from January 2008 to December 2008.[10] This allows for twelve overlapping data points for the price indices, which enables one to smooth over the month-on-month inflation measures for eleven observations. To construct a combined measure of inflation over the structural break, we smooth the data over this period, using

$$\hat{\pi}_{t+\delta} = \left(\frac{\delta}{12} \cdot \tilde{\pi}_{[\text{New, } t]} \right) + \left(\left[1 - \frac{\delta}{12} \right] \cdot \tilde{\pi}_{[\text{Old, } t]} \right) \; ; \tag{6}$$

where $\hat{\pi}_t$ refers to the approximate measure of inflation in January 2008 and $\delta = 1, \ldots, 11$. In this case, $\tilde{\pi}_{[\text{New, } t]}$ refers to monthly inflation that makes use of the new methodology and $\tilde{\pi}_{[\text{Old, } t]}$ refers to monthly inflation that makes use of the old methodology.

To calculate year-on-year inflation from this measure, we can convert the month-on-month series into an index of prices and recalculate the rate of inflation to allow for the maximum amount of smoothing. Of course, one could have combined the year-on-year estimates from the two price indices, but this would not allow for any overlap, as the last value for the old measure of year-on-year inflation is December 2008 and the first measure that uses actual data for the new measure is January 2009. The result of the smoothing that we implemented is shown in figure 1, which depicts the respective year-on-year measures of inflation over the out-of-sample period, taking into account the period where there were two overlapping measures of inflation.

3 Measures of Core Inflation for South Africa

The recent literature on measures of core inflation for South Africa includes the work of Blignaut *et al.* (2009), Rangasamy (2009) and Ruch and Bester (2013). The first of these derives an asymmetric trimmed means estimate, which seeks to provide an estimate of core-inflation that is similar to that of the Hodrick-Prescott filtered trend. Such a measure would be attractive, as in contrast with the Hodrick-Prescott filter, the trimmed means estimate would not be subject to an endpoint problem when estimated in real time. In addition, Rangasamy (2009) estimates the persistence in each of the sub-categories of the data with a univariate time-varying parameter model, such that larger weights are allocated to those sub-categories that exhibit greater persistence (where the weights may be attributed to the degree of persistence, or a combination of persistence and the original basket weight). The results in this paper suggest that these measures track headline inflation with relatively high precision and may provide reasonable estimates of future inflationary pressure.[11]

[9] In addition, the old alcoholic beverages has been split up into spirits, wine and beer; the old housing has been split into actual rentals, owners equivalent rent, maintenance and repairs, water and other services, and electricity and other fuels. The old line items for household services has been incorporated into domestic workers' wages; the old healthcare has been split into medical products and medical services; petrol is now included as a separate, weighted component; and the old communication has been split into postal services and telecommunication. The old recreation has been slit into recreational equipment and recreational cultural services; the old education has been split into primary and secondary, and tertiary education. A number of other items have been included for restaurants, hotels, insurance, and financial services.

[10] Although new data was collected from 2008, comparable data for the new price indices has been provided back to 2002. However, it is important to note that these values are only approximations and they do not include estimates for some of the new sub-categories, notably, the way in which changes to house prices and rentals are now calculated.

[11] As an extension to this area of research, Balcilar *et al.* (2014) suggest that the degree of persistence in South African data may differ significantly when distinguishing between high and low inflationary episodes.

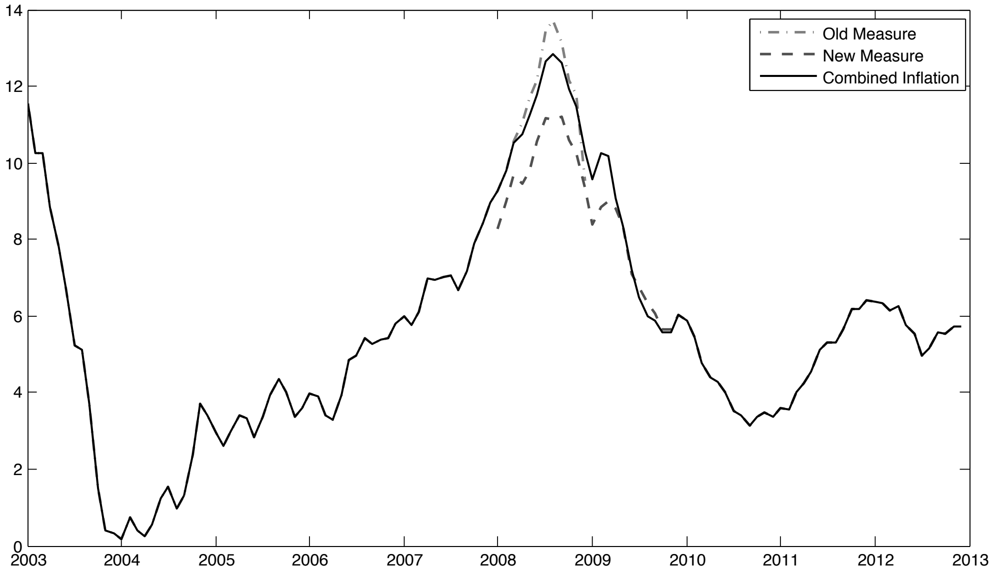

Figure 1: Year-on-year inflation (2003M1-2012M12)

The paper by Ruch and Bester (2013) employs single spectrum analysis to derive a number of measures of core inflation that are unbiased, less volatile than headline inflation, and capable of providing sensible forecasts. This technique involves decomposing the periodicity of the data into various spectra that may be identified by different eigenvalues. The four measures that they report on make use of combinations of the underlying trend, and cyclical components for frequencies over 65, 40 and 24 months.

4 Trimmed Means Estimates of Core Inflation

A trimmed means estimate of core inflation would exclude extreme price movements in the disaggregated data items, as these may be attributed to short-term phenomena that may not influence the persistent rate of inflation. For example, if there is a temporary shortage in the supply of sugar products, we would expect that the relative price of sugar would increase sharply for a short period of time. However, this shortage may be quickly rectified, and as a result, we would not wish for this temporary disturbance to affect long-term monetary policy decisions.

To derive this statistic, we follow Blignaut *et al.* (2009), who make use of an asymmetric trimmed mean estimate, which is calculated after arranging the disaggregated data in ascending order. Thereafter, a number of disaggregated indices may be excluded from the combined measure by setting to zero the corresponding weights for the items that find themselves at the upper and lower ends of the distribution. Hence, this statistic may be calculated as

$$\tilde{\pi}^{T}_{[\gamma 1, \gamma 2;\, t]} = \frac{\sum_{n=1}^{N} \pi_{n,\, t} \cdot \omega_{n,\, t}}{\sum_{n=1}^{N} \omega_{n,\, t}} \,. \tag{7}$$

where γ_1 and γ_2 refer to the amount that is trimmed from each side of the distribution and $\omega_{n,t}$ would refer to the revised weights. In this case an asymmetric trimmed mean estimate would allow for γ_1 to differ from γ_2.

While this statistic may be calculated for both month-on-month and year-on-year inflation, it is worth noting that when performing this calculation on month-on-month inflation, the variation in each disaggregate series is often zero. This may be problematic when one is looking to trim a single series from a number of zero rates of inflation. In this case the choice of which zero to trim would influence the result, depending on whether its respective weight is relatively large or small. There is also no intuition that could be used to guide this decision, and as a result the choice is largely arbitrary. In contrast, the use of year-on-year data has more variation, and as such, the resulting measure would appear to be more reasonable, although this would not allow for one to smooth over the structural break, since the first measure of year-on-year inflation for the new series is January 2009 and the last measure of year-on-year inflation that uses the old method is December 2008.

As an alternative to this trimmed means estimate, we derive an additional trimmed means estimate that excludes those items that experienced the largest deviation, irrespective of whether this change was positive or negative. To derive this measure, we take the absolute value of the respective disaggregated inflation rates to identify the subcategories that are to be trimmed, prior to sorting the observations. Thereafter, we construct the trimmed means estimate using the data that is available prior to taking the absolute value. Doing so allows us to provide more reliable estimates for month-on-month rates of core inflation that are substantially less volatile. However, while this method is intuitively appealing, it is noted that the mean of the process is much lower than the reported measure of aggregate inflation, as large positive price shocks are more common than large negative shocks.

It is also worth noting that the increase in the number of sub-categories complicates the calculation of measures of core inflation that are provided by trimmed means (and dynamic factor models) as the removal of a finite number of sub-categories may have a different impact when the number of underlying sub-categories is either relatively large or relatively small. When reporting on the results, we refer to the *asymmetric trimmed mean* estimate for the traditional trimmed mean estimate that was calculated according to the method employed in Blignaut *et al.* (2009), while the *absolute value trimmed mean* estimate refers to the result from the alternative trimmed means method.

5 Exclusionary Estimates of Core Inflation

Exclusion-based methods seek to exclude sub-categories of CPI that are considered to be particularly volatile. Silver (2007) notes that these measures usually include variants that remove food and energy prices, as these categories may be susceptible to erratic behaviour over a relatively short period of time. While these exclusion-based measures have been used in South Africa as a measure of core inflation (Ruch and Bester, 2013), they may be susceptible to a number of shortcomings, as it has been suggested that changes to food prices are relatively persistent and the contribution of food prices to headline inflation is relatively large (Rangasamy, 2011). In addition, Cogley (2002) suggests that these methods have provided estimates of core inflation that are more volatile than the measures of reported aggregate consumer price inflation in many countries.

For the old measure for CPI, we excluded all the food category items, as well as the subcategories of fuel & power and running costs. Similarly, for the new measure, we exclude all the food category items as well as the subcategories of electricity & other fuels and petrol. These measures were then smoothed over the breakpoint in 2008 using the method that is described in section 2.

6 Dynamic Factor Model Estimates of Core Inflation

Dynamic factor models have been utilised by several researchers to construct measures for core inflation in a number of countries. Recent studies for estimating core inflation with dynamic factor models include: Amstad and Potter (2009) for the United States, Cristadoro *et al.* (2005) for the European Union, Amstad and Fischer (2004) for Switzerland, and Giannone and Matheson (2007) for New Zealand. These models rely on the assumption that the underlying core inflation rate is driven by a small number of common factors, and that the noise in the inflation process is related to localised shocks that affect a limited number of the disaggregated price series. Therefore, the model may be specified as

$$\tilde{\pi}_{n,t} = \mu_n + \lambda_n F_t + \epsilon_{n,t} \ , \tag{8}$$

where, $F_t = (f_{1,t}, \ldots, f_{q,t})'$ represents the q common dynamic factors and λ_n is the matrix of factor loadings that has dimensions $n \times q$. The variable-specific shocks are then contained in the vector $\epsilon_{n,t}$. To give the model a dynamic characterisation, it is assumed that the factors have a vector autoregressive structure, where $\chi_q(L)F_t = \upsilon_t$.

To estimate this model we make use of the procedure of Doz *et al.* (2011), as it allows for the model to be estimated in the presence of missing data. This is of importance in the current setting, since we need to make use of the data that seeks to approximate the new method for calculating consumer prices for the in-sample period 2002M1-2007M12, to derive a consistent estimate for core inflation in 2009M1.

This procedure also facilitates the procedure for generating forecasts within the unified framework, using the Kalman filter with the measurement equation in the state-space representation of the model that treats the vector of factors as an unobserved process. In this case, the autoregressive state equation allows for time-varying parameters in all of the models that are used to generate forecasts.

The number of factors that are included in the optimal dynamic factor model is determined by out-of-sample statistics, allowing various combinations for a maximum of $q = 10$ and a first order lag operator. Note that the decision to utilise a single autoregressive element is due to the need to restrict the dimensionality of the comparative investigation, where all the forecasts for competing models make use of an equivalent state-space framework.[12]

7 Wavelets Estimates of Core Inflation

Most of the commonly used decompositions in economics, such as those that were designed by Hodrick and Prescott (1997), Baxter and King (1999) and Christiano and Fitzgerald (2003), make use of expressions from the time domain. All of these techniques seek to approximate ideal filters, where one is able to identify the trend, cycle and noise components that are located at different periodicities.

Alternatively, these time domain decompositions may be represented in the frequency domain, where the components may be expressed as elements that arise at different Fourier frequencies. Shumway and Stoffer (2010) suggest that the use of the frequency domain is more appropriate for these decompositions of time-series variables. When applying these techniques one would effectively decompose a time series with a number of sine and cosine functions, which define the rate at which the time series oscillates. It is important to note, that this transformation results in the loss of all time-based information, where it is assumed that the periodicity of all the components are consistent throughout the entire sample.

[12] If we made use of additional models that have two autoregressive elements, we would not only need to generate 20 forecasts for the dynamic factor models, but we would also need to generate twice as many forecasts for all the other models, for which we make use of a similar state-space model.

To allow for changes in the periodicity of the respective components, Gabor (1946) developed the Short-Time Fourier Transform (STFT) technique, which involves applying a number of Fourier transforms to different subsamples of the data. Gencay *et al.* (2010) refer to the subsample as a data *window*, where the technique involves sliding the window across the time series and taking a Fourier transform of each subsample. Although this technique would provide potentially useful information on the timing of an event that may have arisen at a particular frequency, it is limited in that the precision of the analysis is affected by the size of the subsample. For instance, one would need a large subsample to identify changes that arise at a low frequency, and small subsamples to identify changes in the higher frequency components.

To overcome the limitations of the above frequency domain techniques, wavelet transformations were developed to capture features of time-series data across a wide range of frequencies that may arise at different points in time. This technique makes use of a number of wavelet functions that are stretched and shifted to describe features that are localised in frequency and time. For example, the wavelet function would be expanded over a relatively long period of time when identifying low-frequency events, and it would be relatively narrow when describing high frequency events.[13] After shifting all of these wavelet functions that have different amplitudes over the entire sample of data, one is able to associate the components with specific time horizons that occur at different locations in time.

Early work with wavelet functions dates back to Haar (1910), who used a number of square-wave functions to decompose time-series data. Unfortunately, the properties of square-wave functions were found to be limited, and as such, a number of alternatives were developed, including those that are discussed in Grossmann and Morlet (1984) and Daubechies (1992).[14] For the computation of these transformations, which make use of various wavelet functions at different scales, most studies currently employ the multiresolution decomposition of Mallat (1989) and Strang and Nguyen (1996).

To describe the use of this technique, one could allow for the case where a variable is composed of a trend and a number of higher-frequency components. In this instance, the trend may be represented by a father wavelet, $\phi(t)$, while the mother wavelets, $\psi(t)$, are used to describe information at lower scales (i.e. higher frequencies). Using an orthogonal wavelet transformation, one could then describe variable x_t as

$$x_t = \sum_k s_{0,k}\phi_{0,k}(t) + \sum_{j=0}^{J}\sum_k d_{j,k}\psi_{j,k}(t) \ , \tag{9}$$

where J refers to the number of scales, and k refers to the location of the wavelet in time. The $s_{0,k}$ coefficients are termed smooth coefficients, since they represent the trend, and the $d_{j,k}$ coefficients are termed the detailed coefficients, since they represent finer details in the data.

The mother wavelet functions, $\psi_{1,k}(t), \ldots, \psi_{J,k}(t)$, are then generated by shifts in the location of the wavelet in time and scale, such that

$$\psi_{j,k}(t) = 2^{-j/2}\psi\left(\frac{t-2^j k}{2^j}\right), \ \ j = 1, \ldots, J \ , \tag{10}$$

where the shift parameter is represented by $2^j k$ and the scale parameter is 2^j. This choice of dyadic scaling factors is arbitrary but efficient (Daubechies, 1992). As depicted in the daublet wavelet functions in figure 2, smaller values of j (which produce a smaller scale

[13]The wavelets literature refers to the use of scales rather than frequency bands, where the highest scale refers to the lowest frequency and vice versa.

[14]See, Hubbard (1998) and Heil and Walnut (2006) for a detailed account of the history of wavelet analysis.

parameter 2^j), would provide the relatively tall and narrow wavelet function on the left. For larger values of j, the wavelet function is more spread out and of lower amplitude. In addition, after shifting this function by one period, we produce the function that is depicted on the right of figure 2.

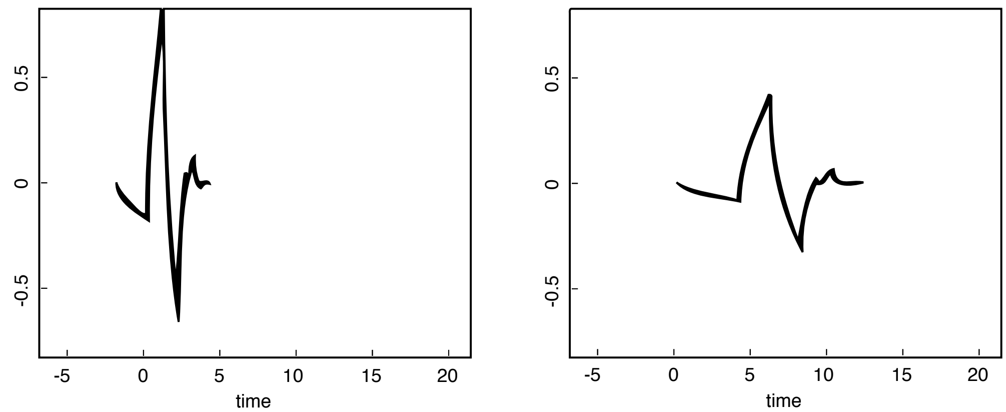

Figure 2: Daublet (4) wavelet functions - $\psi_{1,0}(t)$ and $\psi_{2,1}(t)$

Early applications of wavelet methods in economics include the work of Ramsey and Zhang (1997), which made use of a wavelet decomposition of exchange rate data to describe the distribution of this data at different frequencies. In addition, Ramsey and Lampart (1998*b*) made use of a decomposition of money and income data to describe the relationship between these variables at different frequencies, while Ramsey and Lampart (1998*a*) considered the relationship between income and expenditure (i.e. permanent income hypothesis) at different time scales.[15]

In this study we make use of a maximum overlap discrete wavelet transform (MODWT), which does not restrict the sample size to a multiple of 2^j. In addition, this technique is also able to preserve the phase properties of the data, where it can match the smoothed terms to the underlying price index. As we are primarily interested in removing the noise from the data, we make use of various smoothed wavelet functions that include daublets 2-6, coiflets 1-5, and symlets 2-6, where higher orders refer to functions that focus more intently on higher frequency information. We also make use of both two and three scales, where J is set at 2 and 3, respectively.[16]

8 Results

8.1 Out-of-sample results

To generate forecasts for each of the respective models we make use of a Kalman filter, after the model has been cast in the following state-space framework, with respective measurement

[15] See Ramsey (2002), Schleicher (2002) and Crowley (2007) for a more general overview of the use of these methods in economics.

[16] In this study we perform a simple wavelet analysis that seeks to identify the trend, or father wavelet. Note that these methods could also be used to remove noise from each of the respective scales, should they extend over a particular threshold, before each of the signals is combined to represent the de-noised signal.

and state equations,

$$\begin{aligned}\hat{\pi}_t &= \mu + \alpha_t \hat{\pi}_{t-1} + \varepsilon_t \qquad (11)\\ \alpha_{t+1} &= \alpha_t + \zeta_t \qquad (12)\end{aligned}$$

where α_t is the time-varying parameter. This would ensure that the framework that is used to generate the forecasts for all the models is equivalent to that which is employed by the dynamic factor model. For completeness, we have also included forecasts that were generated from the reported measure of aggregate consumer price inflation. This model is used to generate forecasts that extend from 1-12 steps ahead, over the ten-year period, 2003M1-2012M12. These forecasts are compared to the own forecast values of future inflation, by calculating the root-mean squared errors and the statistics of Diebold and Mariano (1995), which are used to determine whether the differences in the root-mean squared errors of competing models are statistically significant. To calculate the own forecasts, we make use of the reported measure of headline inflation in the above state-space model to generate the respective forecast.

Figure 3 displays the respective root-mean square error statistics for the average of the 1-12 step ahead forecasts. In this case we report only on the optimal specifications for each of the above classes of models, which are those that have the lowest average root-mean squared errors over time. All of these measures report relatively high forecasting errors from 2003-2004 and from the period 2007-2009. These periods relate to those where year-on-year inflation was at a reasonably high level (i.e. well in excess of the target rate, as may be seen in figure 1).

During these periods, the dynamic factor model would appear to do relatively poorly (particularly after a change in the trend of the process), while it does particularly well when inflation is more stable and predicable. The asymmetric trimmed means estimate would appear to do relatively well after the process has remained stable for an extended period of time, but at other times its forecasting error is relatively large. The results are similar for those of the absolute value trimmed mean estimate, although at almost all points in time its forecasting error is larger than the asymmetric trimmed means estimate. The measure that excludes food and energy prices performs relatively poorly, during most periods of time.

The wavelets estimate would appear to provide relatively smaller errors when inflation is high (or trending either upwards or downwards), while its errors are often somewhat larger when inflation is low (where there is less of a positive or negative trend). These estimates would also appear to perform poorly when there is a change in the direction of the trend. These results for the wavelets estimates should not be surprising, since the technique is able to explicitly model any trend in the process, including those that may exhibit stochastic tendencies.

To determine whether the forecasting errors from the core inflation measures are significantly different from those that are provided by the own forecasts that make use of the reported measure of headline (aggregate) inflation, table 1 summarises the sum of the significant Diebold-Mariano statistics. These results suggest that on 46 occasions the dynamic factor model produced forecasts that were significantly superior to those that were generated by the equivalent models for the reported measure of headline inflation, while on 20 occasions it was significantly inferior. The wavelets estimates also generated a large number of forecasts that significantly outperformed those of the own-inflation forecasts, while both trimmed means estimates generated many more estimates that are significantly inferior. These results also suggest that the measure that excludes food and energy prices performs particularly poorly.[17]

[17] Tables 4 to 6 contains the relative root-mean squared errors and Diebold-Mariano statistics for each period of time, where the relative root-mean squared error is calculated as $[(\mathrm{RMSE}_{\mathrm{DFM}}/\mathrm{RMSE}_{\mathrm{actual}}) - 1] \times 100$.

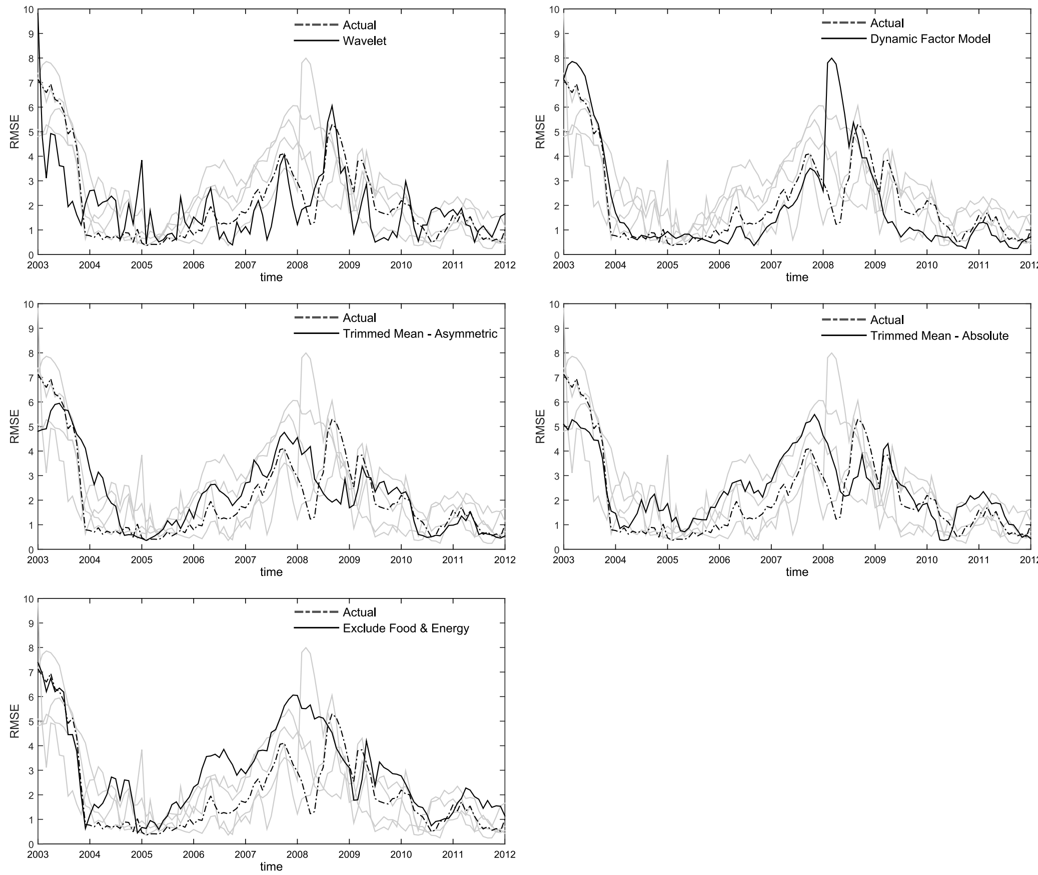

Figure 3: Root-mean squared error (2003M1-2012M12): year-on-year inflation

	Factor Model	Wavelet	Trim - Absol	Trim - Asym	Excl
Superior	46	46	25	30	13
Inferior	20	38	69	52	79

Table 1: Summary of Diebold-Mariano statistics (2003M1-2012M12): year-on-year inflation

We turn our attention to each of the individual step ahead forecasts, where we report on each 1-12 step ahead forecast, after taking the average of these measures for the entire out-of sample period. The results are shown in figure 4, where we note that the wavelets measure of core inflation is consistently below that which is provided by the own inflation forecast at almost all horizons. At short to medium horizons, the dynamic factor model would appear to perform rather poorly; however, as the forecast horizon increases, so too does its relative forecasting potential. The trimmed mean estimates would appear to provide similar results, although the trimmed mean absolute value hardly manages to outperform the forecasts that use aggregated inflation data, while at longer horizons, the errors from the asymmetric trimmed mean estimate are equivalent to those of the dynamic factor model. Once again, the measure that excludes food and energy prices performs particularly poorly, supporting the findings of Rangasamy (2011).

To investigate the statistical significance of these results, we have summarised the Diebold-Mariano statistics for the 1-12 step ahead forecasts in table 2. In this case there are four

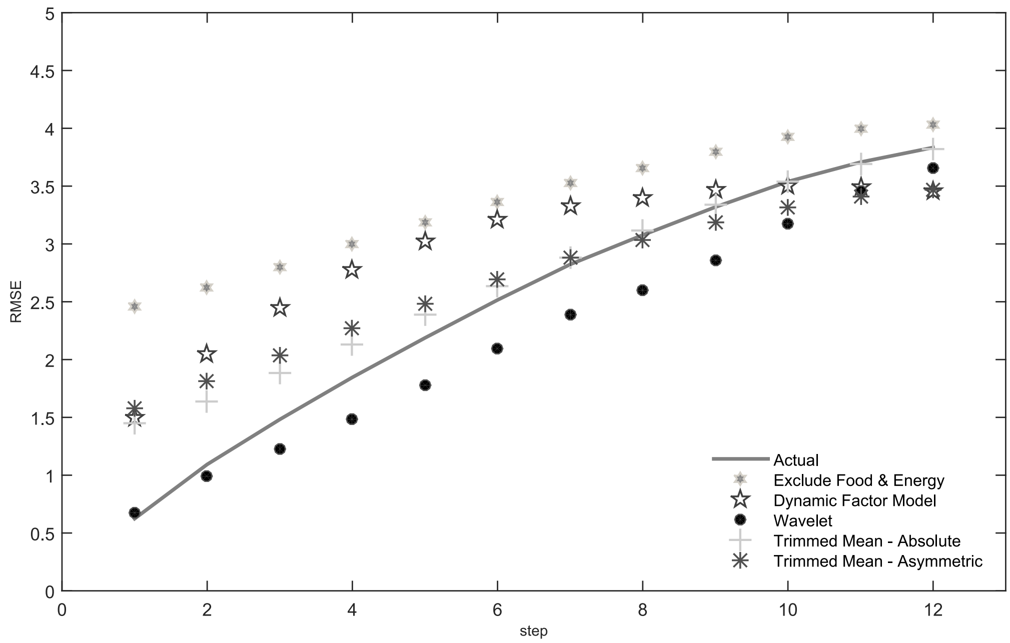

Figure 4: Root-mean squared error (1-12 step ahead): year-on-year inflation

occasions where the wavelets measure is superior, while there are no occasions where it is inferior. This contrast strongly with the other measures, which are never significantly superior, while they are inferior on a number of occasions.

	Factor Model	Wavelet	Trim - Absol	Trim - Asym	Excl
Superior	0	4	0	0	0
Inferior	5	0	3	4	9

Table 2: Summary of Diebold-Mariano statistics (1-12 step ahead): year-on-year inflation

8.2 In-sample results

Bryan and Cecchetti (1994) suggest that estimates of core inflation should provide an accurate estimate of the first moment of actual inflation, while the second moment for core inflation should be lower than actual inflation. In addition, they also suggest that the measure of core inflation should be cointegrated with the actual measure of inflation. When we consider the in-sample statistics for the models that were generated over the full year-on-year sample of 1976M1-2012M12, we note in table 3 that the estimates that provide the most appropriate approximation of the mean and median of actual inflation are derived from the wavelets and dynamic factor models, while the greatest discrepancy is provided by the two trimmed means estimates. The standard deviation for most of the estimates are below actual inflation, with the exception of the measure that excludes food and energy prices.

In addition, we note that with the exception of the absolute-value trimmed mean estimate and the measure that excludes food and energy prices, all of the measures of core inflation

	Actual	Factor Model	Wavelet	Trim - Absol	Trim - Asym	Excl
mean	9.95	9.95	9.95	8.5	9.78	9.5
median	10	9.89	10.25	8.79	9.79	9.73
std	4.5	4.38	4.46	3.84	4.4	4.89

Table 3: In-Sample Descriptive Statistics (1976M1 - 2012M12): Year-on-Year

would appear to be cointegrated when using the Engle-Granger method.[18] The final optimal estimates are then illustrated in figure 5 for the entire sample period; and in figure 6 for the out-of-sample period.

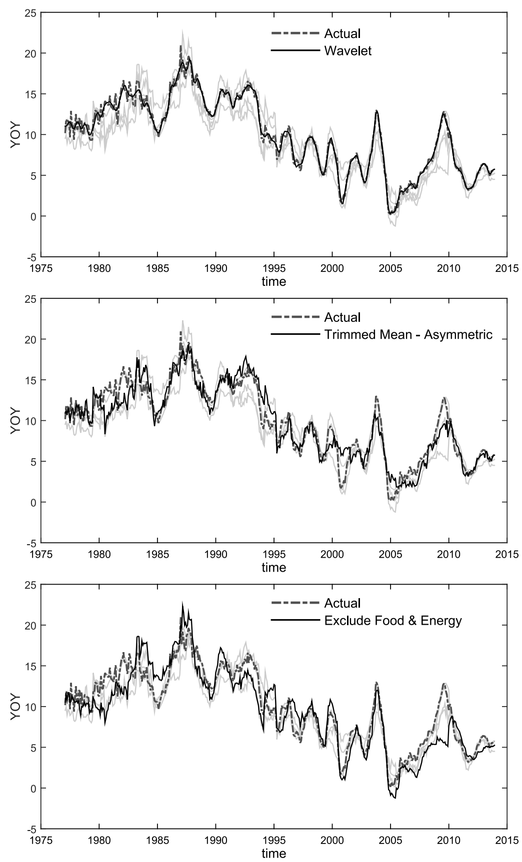

Figure 5: Estimates of year-on-year core inflation (1976M1 - 2012M12)

9 Conclusion

This paper considers the use of various methods that may be used to provide an estimate for core inflation in South Africa. It focuses on the application of trimmed mean, exclusion based estimates, dynamic factor models, and wavelet approaches, where we estimate a total

[18]When testing for cointegration, we first convert the year-on-year measures of inflation (including those for core inflation) into approximate price indices, which are integrated of the first order.

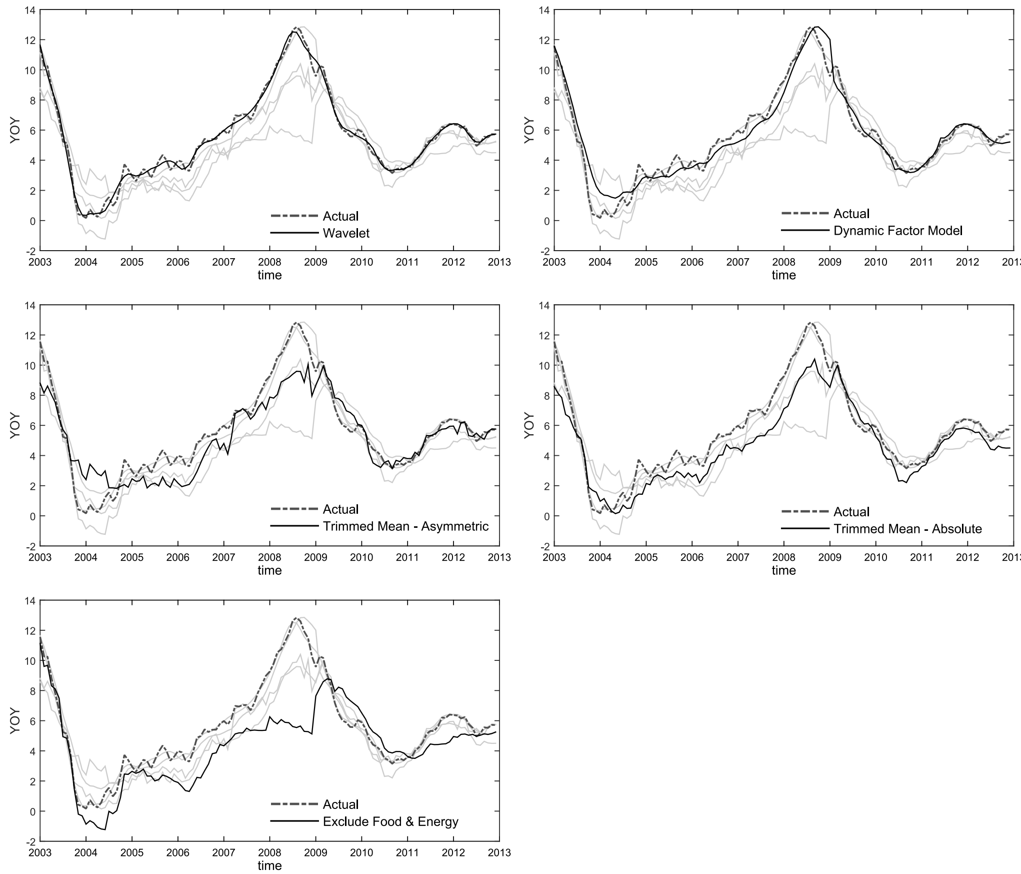

Figure 6: Estimates of year-on-year core inflation (2003M1 - 2012M12)

of 225 different estimates for this key variable. The results are then evaluated using both out-of-sample and in-sample methods.

We find that when comparing the out-of-sample estimates over the ten-year period, 2003M1-2012M12, the only measure of core inflation that is able to improve upon the actual inflation forecasts over all (or for most) horizons, is provided by wavelet estimates. This measure would also appear to provide suitable estimates of future inflation during most periods of time when compared with other proxies.

Hence, policy-makers may wish to include this measure to the range of indicators that are currently used for core inflation. However, they should be aware of the limitations of the wavelets measure, as the results indicate that it could fail to possibly provide superior estimates when there is a change in the trend of the underlying process (which may require a change in policy). Techniques that may be applied to address this concern are beyond the scope of this paper and are left as a potential area for future research.

It is also worth noting that over much longer forecasting horizons (i.e. one year), the dynamic factor model and the asymmetric trimmed means approach improve upon the out-of-sample results of the wavelets methods. In addition, the dynamic factor model also appears to perform reasonably well during the later part of the out-of-sample forecasting period. The results from these estimates would also appear to be most encouraging when the trend in the underlying process is relatively flat. Given these results, policy-makers may wish to incorporate the asymmetric trimmed means and dynamic factor model measures as additional potential indicators of core inflation.

The results from the absolute value trimmed mean estimate are rather poor and do not appear to provide any additional useful information, over that which is provided by the asymmetric trimmed mean estimate. In addition, the results from the measure that excludes food and energy prices are extremely poor, which suggests that the excluded items are important contributors to inflationary pressure. These findings would suggest that both of these measures may be of little use to policy makers as they could provide misleading indicators of core inflation. In addition, they are also of little use when seeking to forecast inflation.

In closing, it is also worth noting that the results suggest that the long-term forecasts of the superior measures of core inflation outperform equivalent long-term forecasts that make use of actual headline inflation (i.e. over a one-year horizon). This should encourage further research into measures of this key unobserved variable.

References

Amstad, M. and Fischer, A. (2004). Sequential information flow and real-time diagnosis of swiss inflation: Intra-monthly dcf estimates for a low-inflation environment. CEPR Discussion Paper.

Amstad, M. and Potter, S. (2009). Real time underlying inflation gauges for monetary policymakers. *FRB of New York Staff Report No. 420*.

Aron, J. and Muellbauer, J. (2004). Construction of cpix data for forecasting and modelling in south africa. *South African Journal of Economics*, vol. 72, no. 5, pp. 884–912.

Balcilar, M., Gupta, R. and Jooste, C. (2014). Analysing south africa's inflation persistence using an arfima model with markov-switching fractional differencing parameter. University of Pretoria, Working Paper Series.

Baqaee, D. (2010). Using wavelets to measure core inflation: The case of new zealand. *The North American Journal of Economics and Finance*, vol. 21, no. 3, pp. 241–255.

Baxter, M. and King, R.G. (1999). Measuring business cycles: Approximate band-pass filters for economic time series. *Review of Economics and Statistics*, pp. 81:575–93.

Bernanke, B.S. and Gertler, M. (1999). Monetary policy and asset price volatility. *Economic Review*, , no. Q IV, pp. 17–51.

Blignaut, Z., Farrell, G., Munyama, V. and Rangasamy, L. (2009). A note on the trimmed mean measure of core inflation in south africa. *South African Journal of Economics*, vol. 77, no. 4.

Blinder, A.S. (1997 May/June). Measuring short run inflation for central bankers - commentary. *Federal Reserve Bank of St. Louis Review*, pp. 157–160.

Bryan, M.F. and Cecchetti, S.G. (1994). Measuring core inflation. In: Mankiw, N.G. (ed.), *Monetary Policy*, pp. 67–123. University of Chicago Press, Chicago.

Cecchetti, S.G. (2009). Monetary policy and the measurement of inflation: Prices, wages and expectations. Tech. Rep. BIS Papers No 49, p1-12, Bank of International Settlements.

Christiano, L. and Fitzgerald, T. (2003). The band pass filter. *International Economic Review*, vol. 44, no. 435-465.

Cogley, T. (2002). A simple adaptive measure of core inflation. *Journal of Money, Credit and Banking*, vol. 34, no. 1, pp. 94–113.

Cristadoro, R., Forni, M., Reichlin, L. and Veronese, G. (2005). A core inflation indicator for the euro area. *Journal of Money, Credit and Banking*, vol. 37, no. 3, pp. 539–560.

Crowley, P.M. (2007). A guide to wavelets for economists. *Journal of Economic Surveys*, vol. 21, no. 2, pp. 207–267.

Daubechies, I. (1992). *Ten Lectures on Wavelets*, vol. 61. SIAM.

Diebold, F.X. and Mariano, R.S. (1995). Predictive accuracy. *Journal of Business and Economic Statistics*, vol. 13, no. 3, pp. 253–263.

Dowd, K., Cotter, J. and Loh, L. (2010). Us core inflation: A wavelet analysis. *Macroeconomic Dynamics*, vol. 15, no. 4, p. 513.

Doz, C., Giannone, D. and Reichlin, L. (2011). A two-step estimator for large approximate dynamic factor models based on kalman filtering. *Journal of Econometrics*, vol. 164, pp. 188–205.

Gabor, D. (1946). Theory of communication. *Journal of the I.E.E*, vol. 93(3), pp. 429–457.

Gencay, R., Selcuk, F. and Whitcher, B. (2010). *An Introdution to Wavelet and other Filtering Methods in Finance and Economics*. Academic Press, San Diego.

Giannone, D. and Matheson, T.D. (2007). A new core inflation indicator for new zealand. *International Journal of Central Banking*, vol. 3, no. 4, pp. 145–180.

Grossmann, A. and Morlet, J. (1984). Decomposition of hardy functions into square integrable wavelets of constant shape. *SIAM journal on mathematical analysis*, vol. 15, no. 4, pp. 723–736.

Haar, A. (1910). Zur theorie der orthogonalen funktionensysteme. *Mathematische Annalen*, vol. 69, no. 3, pp. 331–371.

Heil, C. and Walnut, D.F. (2006). *Fundamental Papers in Wavelet Theory*. Princeton University Press, Princeton, New Jersey.

Hodrick, R.J. and Prescott, E.C. (1997). Postwar u.s. business cycles: an empirical investigation. *Journal of Money, Credit and Banking*, vol. 29, no. 1, pp. 1–16.

Hubbard, B.B. (1998). *The World According to Wavelets: The Story of a Mathematical Technique in the Making*. 2nd edn. A K Peters/CRC Press.

Mallat, S.G. (1989). A theory for multiresolution signal decomposition: The wavelet representation. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 11, no. 7, pp. 674–693.

Marcus, G. (2011). Statement of the monetary policy committee. Tech. Rep., South African Reserve Bank.

Ramsey, J.B. (2002). Wavelets in economics and finance: Past and future. *Studies in Nonlinear Dynamics & Econometrics*, vol. 6, no. 3.

Ramsey, J.B. and Lampart, C. (1998 April*a*). The decomposition of economic relationships by time scale using wavelets: Expenditure and income. *Studies in Nonlinear Dynamics and Econometrics*, vol. 3, no. 1, pp. 23–42.

Ramsey, J.B. and Lampart, C. (1998*b*). Decomposition of economic relationships by timescale using wavelets. *Macroeconomic Dynamics*, vol. 2, no. 1, pp. 49–71.

Ramsey, J.B. and Zhang, Z. (1997). The analysis of foreign exchange data using waveform dictionaries. *Journal of Empirical Finance*, vol. 4, no. 4, pp. 341–372.

Rangasamy, L. (2009). Inflation persistence and core inflation: The case of south africa. *South African Journal of Economics*, vol. 77(3), no. 2, pp. 430–444.

Rangasamy, L. (2011). Food inflation in south africa: Some implications for economic policy. *South African Journal of Economics*, vol. 79, no. 2, pp. 184–201.

Ruch, F. and Bester, D. (2013). Towards a measure of core inflation using singular spectrum analysis. *South African Journal of Economics*, vol. 81(3), pp. 307–329.

Schleicher, C. (2002). An introduction to wavelets for economists. Tech. Rep. Working Paper 2002-3, Bank of Canada.

Shumway, R. and Stoffer, D. (2010). *Time Series Analysis and its Applications*. Springer-Verlag, New York.

Silver, M. (2007). Core inflation: Measurement and statistical issues in choosing among alternative measures. *IMF Staff Papers*, vol. 54(1), no. 54(1).

South African Reserve Bank (2014 June). Monetary policy review. Tech. Rep., South African Reserve Bank.

Statistics South Africa (2009 3 February 2009). Shopping for two: The cpi new basket parallel survey - results and comparisons with published cpi data. Tech. Rep., Statistics South Africa, http://www.statssa.gov.za/cpi/index.asp.

Strang, G. and Nguyen, T. (1996). *Wavelets and Filter Banks*. Wellesley Cambridge Press.

Van der Merwe, E.J. (2004). Inflation targeting in south africa. South African Reserve Bank Occasional Paper No. 19.

10 Appendix - Additional Results

	Actual	Factor Model		Wavelet		Trim - Abs		Trim - Asym		Excl. F & E	
	RMSE	RRMSE	DM	RRMSE	DM	RRMSE	DM	RRMSE	DM	RRMSE	DM
2003M1	7.13	0.31	-0.4	37.6	-3.3	-28.66	3.45	-32.64	3.37	3.52	-2.11
2003M2	6.83	12.92	-2.99	-30.26	3.04	-28.77	3.55	-28.66	3.59	1.38	-1.46
2003M3	6.58	19.27	-3.67	-52.8	3.19	-19.9	3.82	-25.52	3.63	-5.83	4.66
2003M4	6.94	12.14	-3.25	-29.08	3.9	-25.67	4.87	-18.89	5.2	-2.84	10.18
2003M5	6.31	19.85	-4.16	-23.02	3.66	-21.78	5.23	-6.63	6.95	-1.83	4.85
2003M6	6.21	16.87	-5.01	-41.78	4.29	-21.07	5.55	-4.29	6.36	2.16	-9
2003M7	5.82	11.21	-5.48	-38.58	4.47	-18.39	6.11	-2.33	6.92	6.14	-9.46
2003M8	4.9	15.78	-7.91	-60.29	5.55	-9.31	4.55	15.18	-7.37	-9.25	6.51
2003M9	5.12	3.55	-9.32	-57.78	7.31	-13.68	6.2	0.48	-0.77	-13.14	8.29
2003M10	4.33	3.7	-3.51	-62.63	7.41	-14.68	5.08	8.25	-8.25	-12.69	8.92
2003M11	2.36	43.54	-16.68	-49.12	3.87	-4.58	1.12	88.53	-21.37	-16.81	6.86
2003M12	0.79	197.96	-7.19	160.31	-1.98	95.45	-4.43	415.73	-9.79	-21.07	2.15
2004M1	0.76	122.03	-4.36	240	-2.3	91.5	-3.52	327.25	-6.98	53.23	-0.98
2004M2	0.69	88.52	-3.44	277.61	-2.43	21.89	-1.75	279.79	-5.56	132.24	-2.1
2004M3	0.88	45.36	-4.13	124.01	-1.75	9.22	-0.3	256.99	-6.92	66.33	-1.37
2004M4	0.6	71.57	-2.04	248.78	-2.74	44.19	-1.65	376.21	-6.05	171.69	-3.23
2004M5	0.74	-7.93	0.27	195.24	-3.36	57.15	-2.91	192.6	-3.82	173.52	-5.68
2004M6	0.66	-9.59	0.45	258.18	-3.94	119.93	-3.99	182.92	-3.5	310.57	-6.81
2004M7	0.75	4.36	-0.15	192.99	-4.27	168.15	-4.81	134.15	-3.04	253.32	-7.21
2004M8	0.62	16.84	-0.75	-6.35	1.21	139.43	-4.06	53.6	-1.48	186.38	-4.56
2004M9	0.89	-23.89	2.11	14.95	-1.73	129.01	-8.13	-36.57	1.35	191.18	-10.26
2004M10	0.87	-14.71	2.43	136.92	-4.73	157.69	-8.68	-31.53	1.41	195.32	-10.92
2004M11	0.49	36.76	-1.22	152.78	-3.84	219.48	-5.17	19.62	-1.21	230.42	-4.78
2004M12	1.02	-23.7	1.67	138.04	-3.47	55.18	-2.22	-52	3.7	-56.39	2.91
2005M1	0.47	97.47	-1.68	710.49	-2.9	292.44	-5.03	-6.93	0.33	41.5	-1.16
2005M2	0.36	112.57	-1.71	17.21	-1.1	210.2	-3.6	-0.77	0.06	74.05	-1.75
2005M3	0.41	59.07	-1.17	332.08	-4.11	216.17	-5.59	8.61	-0.92	131.17	-3.88
2005M4	0.41	77.5	-1.77	83.31	-2.19	85.07	-2.18	31.43	-1.21	109.11	-2.68
2005M5	0.4	106.13	-1.51	26.9	-1.22	77.59	-2.17	64.35	-1.38	46.64	-1.61
2005M6	0.5	65.18	-1.85	20.04	-1.77	74.71	-1.9	64.99	-1.32	90.15	-2
2005M7	0.71	-3.97	0.19	13.53	-1.3	76.94	-5.8	62.66	-4.18	60.19	-5.2
2005M8	0.54	23.18	-1.09	58.98	-1.26	125.13	-2.52	79.68	-1.66	194.34	-3.78
2005M9	0.64	-2.4	0.11	-6.67	0.45	86.68	-1.96	83.24	-1.89	165.35	-3.27
2005M10	0.77	-25.98	2.42	202.84	-4.49	55.68	-1.25	9.63	-0.33	85.09	-2.03
2005M11	0.67	-25.73	2.71	94.87	-3.62	153.78	-2.33	78.65	-2.26	148.36	-2.75
2005M12	1.02	-55	2.18	3.18	-1.2	65.72	-4.44	59.1	-6.45	85.99	-5.54

Table 4: Relative root-mean squared Error and Diebold and Mariano statistics (2003M1-2012M12)

	Actual	Factor Model		Wavelet		Trim - Abs		Trim - Asym		Excl. F & E	
	RMSE	RRMSE	DM	RRMSE	DM	RRMSE	DM	RRMSE	DM	RRMSE	DM
2006M1	0.79	-25.86	2.04	92.83	-2.61	161.3	-3.55	58.64	-4.04	191.98	-5.2
2006M2	0.99	-49.49	2.84	35.06	-2.62	138.01	-4.14	90.22	-5.29	146.75	-5.68
2006M3	0.95	-54.89	3.29	27.07	-3.1	131.65	-4.45	136.99	-6.54	220.46	-6.41
2006M4	1.59	-54.51	4.06	41.44	-3.95	68.55	-9.14	49.55	-12.64	122.82	-10.33
2006M5	1.95	-43.04	3.71	38.54	-3.93	41.38	-9.58	34.62	-22.57	83.45	-25.8
2006M6	1.55	-26.04	1.98	4.27	-6.42	81.26	-8.61	70.12	-22.01	134.79	-15.08
2006M7	1.21	-42.34	1.74	-48.89	1.88	93.51	-7.73	95.71	-16.61	188.79	-10.38
2006M8	1.27	-51.42	2.13	-27.65	1.5	115.6	-6.22	76.87	-10.3	201.86	-9.81
2006M9	1.22	-70.56	2.47	-57.79	2.07	86.76	-6.48	71.16	-7	185.83	-9.44
2006M10	1.31	-55.34	2.89	-70.52	3.15	82.17	-7.09	35.07	-6.99	145.02	-9.36
2006M11	1.45	-40.35	2.57	-9.53	2.66	46.14	-5.62	37.38	-4.84	92.84	-8.26
2006M12	1.74	-36	2.43	-48.43	3.11	52.58	-6.32	24.68	-4.79	75.53	-8.15
2007M1	1.68	-25.52	1.81	-32.67	2.56	63.81	-5.81	60.08	-6.89	69.45	-5.99
2007M2	1.92	-31.59	2.07	-56.55	2.47	67.26	-5.65	42.48	-6.05	62.88	-6.75
2007M3	2.43	-27.33	2.27	-26.62	3.05	47.18	-7.51	49.13	-8.52	38.33	-6.92
2007M4	2.65	-24.08	2.08	-17.45	2.7	45.45	-5.79	27.95	-6.01	42.64	-5.73
2007M5	2.18	-14.06	1.6	-26.85	2.22	84.54	-4.76	38.47	-2.81	74.53	-4.33
2007M6	2.71	-25.82	2.34	-77.79	2.23	59.31	-5.17	7.89	-2.85	39.94	-5.57
2007M7	2.99	-24.77	2.61	-62.41	2.65	45.36	-5.06	8.85	-2.83	51.36	-4.91
2007M8	3.4	-21.24	2.73	-36.77	2.93	35.5	-5.84	16.99	-3.55	36.72	-5.2
2007M9	4.05	-19.78	3.2	-10.57	4.26	26.68	-6.95	12.54	-4.47	25.85	-6.92
2007M10	4.1	-14.41	2.67	-1.05	1.51	26.83	-8.23	16.03	-5.07	36.93	-7.36
2007M11	3.68	-7.82	1.74	-21.55	2.99	48.92	-7.8	21.92	-5.47	59.61	-8.11
2007M12	3.24	-3.67	0.96	-39.86	3.58	60.51	-8.87	32.11	-5.78	86.92	-8.96
2008M1	2.93	-12	2.14	-58.47	3.34	57.4	-10.37	54.89	-8.4	105.9	-10.64
2008M2	2.53	207.28	-11.13	-28.07	0.89	69.58	-8.34	51.89	-7.31	117.34	-9.2
2008M3	1.9	318.83	-7.74	1.74	-0.04	96.92	-6.1	109.29	-7.1	188.58	-9.07
2008M4	1.22	533	-7.25	61.95	-1.14	163.19	-3.85	241.68	-6.14	362.08	-7.52
2008M5	1.31	417.79	-7.66	103.45	-1.71	78.38	-2.02	117.09	-3.12	286.57	-5.26
2008M6	2.71	110	-4.89	16.99	-2.24	-20.24	0.76	-4.24	0.11	91.13	-2.32
2008M7	3.11	40.67	-1.78	5.85	-2.32	-29.28	1.2	-26.98	0.79	64.19	-1.68
2008M8	4.75	13.01	-0.46	13.14	-2.72	-39.69	2.6	-56.65	1.92	1.53	-0.04
2008M9	5.29	-10.84	0.41	14.28	-2.95	-43.95	3.04	-63.73	2.55	-14.47	0.48
2008M10	5.14	-23.51	0.96	-2.65	2.04	-25.11	4.36	-60.41	2.91	-24.07	0.82
2008M11	4.69	-16.14	0.64	-29.47	2.9	-37.71	3.44	-60.75	2.7	-23.36	0.77
2008M12	4.04	-17.03	0.68	-11.72	3.1	-31.7	3.8	-29.31	3.99	-17.89	0.58

Table 5: Relative root-mean squared error and Diebold and Mariano statistics (2003M1-2012M12)

	Actual	Factor Model		Wavelet		Trim - Abs		Trim - Asym		Excl. F & E	
	RMSE	RRMSE	DM	RRMSE	DM	RRMSE	DM	RRMSE	DM	RRMSE	DM
2009M1	3.2	-9.64	0.36	-33.69	2.93	-23.85	2.97	-47.32	2.19	-3.4	0.09
2009M2	2.54	-2.89	0.08	-66.65	2.81	-2.12	0.55	-28.95	2.6	-29.96	1.68
2009M3	3.79	-56.87	2.77	-63.09	4.42	7.09	-1.67	-36.26	4.8	-52.72	4.2
2009M4	3.83	-66.51	4.15	-42.12	4.88	12.23	-3.37	-11.93	6.05	-31.32	7.57
2009M5	3.29	-59.35	4.3	-13.09	3.24	-2.54	6.92	-10.39	3.51	27.39	-3.15
2009M6	2.74	-62.03	4.16	-61.03	4.11	2.9	-9.69	6.14	-4.39	22.97	-5.93
2009M7	1.84	-61.76	2.53	-72.78	2.9	41.57	-5.67	26.35	-9.43	28.09	-13.52
2009M8	1.73	-47.31	1.56	-65.52	2.1	46.34	-6.57	56	-7.76	92.8	-6.99
2009M9	1.66	-33.77	1.01	-58.8	1.75	34.65	-6.28	65.79	-6.43	86.62	-8.23
2009M10	1.62	-39.59	1.35	-64.37	2.12	21.34	-4.09	37.87	-5.81	85.5	-5.67
2009M11	1.84	-46.41	1.77	-48.82	2.67	-15.98	2.27	30.14	-6.27	55.27	-6.59
2009M12	1.86	-51.86	2.17	-62.55	2.98	-2.91	0.96	35.08	-5.37	51.87	-6.5
2010M1	2.19	-68.84	3.74	-25.27	4.7	-14.8	4.79	3.03	-4.77	25.95	-10.2
2010M2	2.03	-74.12	5.19	46.27	-3.95	-32.72	6.23	14.94	-5.39	19.85	-13.57
2010M3	2	-61.36	5.99	15.16	-7.24	-41.53	6.36	-0.04	0.16	4.94	-1.5
2010M4	1.37	-45.62	5.56	-20.96	3.45	-72.17	5.18	-16.88	6.17	23.97	-5.45
2010M5	1.23	-43.31	4.93	-56.12	3.96	-70.9	4.48	-50.87	5.04	24.22	-3.94
2010M6	1.09	-35.41	3.72	8.96	-0.26	-63.33	4.28	-49.13	5.52	34.72	-4.76
2010M7	0.81	-37.67	2.68	47.05	-0.87	6.64	-0.17	-40.68	3.82	47.6	-3.74
2010M8	0.49	-29.71	1.45	261.61	-2.19	243.07	-2.93	1.7	-0.36	47.67	-3.22
2010M9	0.57	-41.69	1.14	233.69	-2.52	206.17	-3.31	-0.53	0.02	59.35	-2.17
2010M10	0.93	-73.78	2.07	96.16	-2.79	99.12	-4.62	-41.25	2.21	1.09	-0.32
2010M11	0.91	-19.44	1.38	67.7	-2.6	139.47	-3.38	-18.33	1.7	12.32	-2.35
2010M12	1.19	-14.73	1.45	43.93	-2.84	70.41	-3.56	-18.95	2.6	-12.63	2.08
2011M1	1.59	-19.91	2.65	14.71	-3.17	33.6	-4.34	-38.29	3.14	-19.49	3.54
2011M2	1.34	-2.46	0.93	19.18	-3.5	74.97	-3.78	-22.33	3.7	24.4	-3.09
2011M3	1.71	-27.34	5.53	9.5	-3.99	23.69	-4.54	-20.92	4.86	8.03	-4.28
2011M4	1.33	-40.96	3.53	-8.47	3.36	39.07	-4.45	-12.38	5.77	71.23	-5.18
2011M5	1.54	-51.86	4.51	-25.56	3.98	22.52	-6.47	-1.57	0.89	42.12	-7.8
2011M6	1.12	-55.36	3.35	-55.01	3.52	63.88	-5.54	-1.97	4.37	70.96	-11.22
2011M7	0.85	-33.62	1.38	16.73	-0.37	58.74	-6.1	-29.34	4.38	113.49	-8.7
2011M8	0.6	-50.32	2.34	91.36	-1.2	67.16	-3.29	10.73	-1.39	137.98	-4.91
2011M9	0.7	-65.49	3.5	29.42	-0.59	42.02	-3.49	-12.11	1.91	151.67	-6.6
2011M10	0.55	-55.63	3.13	28.11	-1.15	38.45	-1.48	6.66	-0.39	170.24	-3.72
2011M11	0.62	-16.88	0.91	78	-2.43	-21.74	0.7	-21.87	0.69	148.26	-3.04
2011M12	0.51	21.82	-1.48	203.64	-3.01	13.74	-2.03	-10.52	0.39	200.1	-3.66
2012M1	1	-29.01	1.9	64.57	-3.5	-57.2	2.74	-45.77	3.16	13.44	-0.41

Table 6: Relative root-mean squared error and Diebold and Mariano statistics (2003M1-2012M12)

	Actual	Factor Model		Wavelet		Trim - Abs		Trim - Asym		Excl. F & E	
	RMSE	RRMSE	DM	RRMSE	DM	RRMSE	DM	RRMSE	DM	RRMSE	DM
1 step	0.62	139.03	-3.18	7.54	-1.2	132.56	-5.6	153.56	-5.17	293.63	-3.98
2 step	1.09	87.61	-2.87	-9.64	1.17	50.17	-3.72	65.99	-4.21	139.79	-4.1
3 step	1.48	65.5	-2.64	-17.55	2.01	26.72	-2.52	37.78	-3.31	89.1	-4.19
4 step	1.84	50.16	-2.4	-19.79	2.26	15.41	-1.66	22.89	-2.4	62.23	-4.28
5 step	2.18	37.91	-2.16	-18.93	2.19	8.99	-1.07	13.61	-1.64	45.54	-4.31
6 step	2.51	27.44	-1.9	-16.69	2.02	4.62	-0.6	6.89	-0.95	33.79	-4.15
7 step	2.82	18.07	-1.56	-15.32	1.92	2.26	-0.32	2.05	-0.31	25.02	-3.69
8 step	3.07	10.55	-1.1	-15.41	1.95	1.42	-0.21	-1.22	0.19	18.93	-2.96
9 step	3.31	4.34	-0.51	-13.8	1.8	0.61	-0.09	-3.9	0.59	14.4	-2.26
10 step	3.53	-0.93	0.11	-10.45	1.4	-0.15	0.02	-6.16	0.9	10.79	-1.63
11 step	3.7	-5.66	0.69	-6.89	0.93	-0.37	0.06	-7.86	1.11	7.82	-1.11
12 step	3.83	-9.75	1.13	-4.63	0.63	-0.26	0.04	-9.41	1.29	5.23	-0.69

Table 7: Relative root-mean squared error and Diebold and Mariano statistics (1-12 step ahead)