7/RT/09 **August 2009**

Research Technical Paper

A Critical Assessment of Existing Estimates of Core Inflation

Colin Bermingham *

Economic Analysis and Research Department
Central Bank and Financial Services Authority of Ireland
P.O. Box 559, Dame Street
Dublin 2
Ireland
http://www.centralbank.ie

*The views expressed in this paper are the personal responsibility of the author. They are not necessarily held either by the CBFSAI or the ESCB. I would like to thank Kieran McQuinn and Maurice Roche for helpful comments. All errors are my own. Email: colin.bermingham@centralbank.ie

Abstract

Core inflation rates are widely calculated. The perceived benefit of core inflation rates is that they help to inform monetary policy. This is achieved by uncovering the underlying trend in inflation or by helping to forecast inflation. Studies which compare core inflation rates frequently assess candidate core rates on these two criteria. Using U.S. data, the two standard tests of core inflation - the ability to track trend inflation and the ability to forecast inflation - are applied to a more comprehensive set of core inflation rates than has been the case in the literature to date. Furthermore, the tests are applied in a more rigorous fashion. A key difference in this paper is the inclusion of benchmarks to the tests, which is non-standard in the literature. Two problems with core inflation rates emerge. Firstly, it is very difficult to distinguish between different core rates according to these tests, as they tend to perform to a very similar level. Secondly, once the benchmarks are introduced to the tests, the core inflation rates fail to outperform the benchmarks. This failure suggests that core inflation rates are of less practical usefulness than previously thought.

Non Technical Summary

Core inflation is an important economic concept. It provides an indication of the underlying trend in prices. This is in contrast to the actual inflation rate, which is inadequate for this purpose, as it is designed to measure changes in the cost of living. As such, the core rate is of particular importance in a policy context as policymakers need a clear understanding of the true trend in prices in order to set interest rates appropriately. However, the core inflation rate is not an actual economic series and instead must be estimated using statistical measures. The practical usefulness of core inflation depends critically on the accuracy of the estimation method used to construct the core rate.

Using U.S. data, this paper proposes the most rigorous examination of core inflation estimates to date. The two standard tests of the usefulness of core inflation are its ability to track trend inflation and its ability to forecast actual inflation. In this paper, improvements are made to way these tests are implemented. Other papers in the literaure frequently omit a benchmark model from the comparison tests and instead only rank candidate core inflation rates. This is despite the convention in other branches of economics of including a benchmark model. In this paper, benchmark models are included in both the forecasting test and the trend tracking test. Alternative tests not normally considered in the core inflation literature are also included in the paper. In addition, the comparison exercise is the most exhaustive to date in terms of the range of core inflation estimators included. Most papers focus on a specific type of core inflation estimator whereas a number of core inflation estimators are included in this paper.

Two problems with the core inflation rates emerge. Firstly, it is very difficult to distinguish between different core rates according to these tests, as they tend to perform to a very similar level. Secondly, once the benchmarks are introduced to the tests, the core inflation rates fail to outperform the benchmarks. This means that U.S. core inflation rates are not particularly helpful policy tools when utilised in the manner suggested by the literauture. This suggests that alternative estimators of core inflation may be needed or that existing core rates should be utilised differently. The paper considers some alternative uses for core rates but the various estimators still perform relatively poorly according to these alternative criteria. However, the existing core rates might still be helpful in a forecasting context using different forecasting models and methods to the conventional approach.

1 Introduction

The fundamental idea underlying the concept of core inflation is that inflation is ultimately determined by monetary growth, which should affect all prices in the economy equi-proportionately. Core inflation is then defined as the common element in all price changes. The concept is important because it provides a clear picture of the underlying trend in prices. This is in contrast to the actual inflation rate, which is inadequate for this purpose, as it is designed to measure changes in the cost of living. As such, the core rate is of particular importance in a policy context. Information regarding the true trend in prices is critical to policymakers given the long and variable lags between the implementation of monetary policy and its effect on inflation.

Like many other important economic concepts such as potential output or the NAIRU, core inflation is not an actual series and instead must be estimated. Its usefulness as a policy tool depends critically on the accuracy of the estimation method used to construct the core rate. Methods used to calculate core inflation include removing volatile items from the calculation of inflation, statistical filters, SVAR methods, trimmed means and factor models so there is a broad range of core inflation estimators. There is an existing literature that compares the relative merits of core inflation rates based on their policy usefulness. The two tests of policy usefulness most commonly used are the ability to track trend inflation and the ability to forecast actual inflation.

Using U.S. data, this paper proposes the most rigorous examination of core inflation estimates to date. The contribution of the paper is threefold. Firstly, in relation to the trend tracking test, the standard approach to date involves estimating the core rate and the trend over the full sample. This paper implements the test in a pseudo real time environment and so provides a more realistic assessment of the ability of core inflation rates to track trend inflation. Secondly, the forecast tests in the core inflation literature frequently omit a benchmark forecast from the comparison and instead only rank candidate core inflation rates. This is despite the convention in the forecasting literature of including a benchmark model. In this paper, benchmarks are included in both the forecasting tests and the trend tracking test. The introduction of a benchmark forecast to a core inflation paper is not novel. However, studies with benchmark forecasts have typically focused on a small number of core estimators and it has generally not been implemented for the US. In contrast, the introduction of the benchmark to the trend tracking test is novel. The final contribution of the paper is that the comparison exercise is the most exhaustive to date in terms of the range of core inflation estimators included. Most papers focus on a specific type of core inflation estimator whereas a number of core inflation estimators are included in this paper.

The paper finds that core inflation rates are no better at forecasting inflation or tracking trend inflation than the benchmarks included in the tests. In short, the benefit of core inflation rates to policymakers is overestimated. New uses of core inflation rates could exist. This paper suggests two alternative tests of core inflation, less stringent that those currently employed, but the performance of existing core inflation estimators is still relatively poor according to these alternative tests. It may also be possible to use existing estimators more efficiently. For example, although this paper shows that the standard forecasts of inflation based on an inflation gap fail to outperform a benchmark, other specifications or estimation techniques involving existing core rates might be found that could improve on the benchmark. The next section contains a literature review and highlights the contribution of the paper. Section 3 outlines the estimators used in the paper, including any issues in the estimation. Section 4 critically evaluates the performance of the core estimators and section 5 concludes the paper.

2 Literature Review

In tackling the issue of core inflation, the initial focus in the literature was simply to construct new estimates. A number of approaches were taken but these can generally be classified as either structural or statistical. The most basic statistical approaches simply involve excluding certain components, such as the volatile food and energy components. This type of core inflation rate is routinely calculated by national statistical agencies. More sophisticated techniques include statistical filters. The Hodrick-Prescott (HP) filter has been widely applied to economic time series, including inflation and provides one core estimate. The HP filter has been criticised in the past, particularly in relation to the well known end-point problem. Baxter and King (1999) propose an alternative filter, based on the spectral decomposition of a time series. It involves filtering parts of the series that lie between certain frequencies and this can be also used as a measure of core inflation.

Bryan et al (1997) propose the use of trimmed means as estimators of core inflation.

Based on the notion that the headline rate can be significantly affected by large price changes in individual components, the trimmed means exclude these items and are considered robust to these outliers. Subsequent to their paper, trimmed mean estimates were calculated for a large number of countries. In this paper, we calculate trimmed means using two alternative weighting systems.

Persistence measures of core inflation can also be calculated. These measures are based on the persistence of the individual components that constitute the inflation rate. Persistence is estimated using an autoregressive model. Cutler (2001) applied this approach to UK data using only one lag for all series whereas Bilke and Stracca (2007) apply a similar approach to Euro Area data but measure persistence with the lag length determined using traditional lag selection tests. One of the core inflation measures examined in this paper is the Bilke and Stracca (2007) approach. This type of core inflation measure is rarely calculated so its inclusion in the comparison should shed some light on its relative merits.

The structural approach considered is the structural VAR as this is clearly the most prevalent structural approach to estimating core inflation. The methodology used is that proposed by Quah and Vahey (1995) with a standard long-run restriction. According to their approach, "inflation is assumed to be affected by two different types of shock, distinguished by their effect on output. The core inflation shock is output neutral after some fixed horizon whereas the non-core shock is allowed to influence output in the long-run." Following identically the method of Quah and Vahey (1995), a bivariate VAR is estimated using the assumption that the core shock is output neutral.

The final type of core inflation rate included in the paper is a dynamic factor model estimate. Factor models are used when analysing a large volume of data such as the individual price series that make up the overall inflation rate. Following the approach of Stock and Watson (2002), the factor model finds the common element in all these price changes. The benefit of this type of approach is that it takes both time series information, cross-sectional information and frequency domain information into account.

The papers mentioned so far relate to the estimation of core inflation. Other papers in the literature aim to compare and assess various core inflation measures. This paper compares core inflation measures but considers a broader range of core inflation series than other papers in the literature. For example, Clark (2001) compares core inflation measures but concentrates chiefly on exclusion based statistical measures. In a study on German data, Landau (2000) includes the structural VAR but omits a number of important statistical estimators. Smith (2004) examines filters, trims and some exclusion measures, as do Rich and Steindel (2007). The scope of this paper includes all major estimation methods.

Many papers rank core inflation rates based on their ability to forecast actual inflation. Given the well documented difficulties associated with forecasting inflation it is somewhat surprising that this is such a popular yardstick. It is in some part due to the manner in which the forecast comparison exercises have been conducted. Although not an exhaustive list, Cogley (2002), Smith(2004), Clark (2001) and Rich and Steindel (2007) include only core inflation rates in the forecast comparison exercise using US data - there is no benchmark forecast included. The inclusion of a benchmark forecast is considered standard practice in the forecasting literature. Model forecasts are compared to forecasts from naive models, such as a no change forecast, in order to assess their forecasting ability. If the model forecast cannot beat the naive forecast, the model is of little worth for forecasting. This paper includes a benchmark in both the forecast test and the trend tracking test.

An additional improvement is also made to the trend tracking test. The trend is routinely defined as a centred moving average of inflation. The standard approach is to estimate the trend and the core rate using the full sample of data and then compare the two. Instead, we estimate the core and trend recursively as this more closely reflects the situation faced in practice. Although we stop short of conducting a full real time exercise, most core rates are based only on inflation data which are rarely revised.

We also consider two alternative metrics to gauge candidate core inflation series. The first criterion considered is the ability to predict changes in the direction of inflation. Although a poor predictor of the magnitude of inflation, core inflation may still be useful as a predictor of the direction of future changes. The second criterion that we examine is a measure of concordance, which has been used by McDermott and Scott (1999) in the business cycle literature. A key property of core inflation is to indicate whether there is excess inflationary pressure in the economy. If core inflation is above overall inflation, there is a negative "inflation gap". The ability of candidate core rates to measure this gap is captured by concordance, which is the degree to which core inflation series agree on the sign of the inflation gap. The performance of the core rates according to these tests do not suggest an alternative use for core inflation rates.

3 Calculation of Core Inflation

This section describes the construction of the core inflation measures. One issue of concern is the stationarity of the inflation rate. For some core measures, the unit root properties of inflation are irrelevant. This mainly applies to the statistical measures. The HP filter simply smoothes the inflation rate to get a core measure so the unit root issue is irrelevant. Similarly, trimmed means and the PCE excluding food and energy inflation rate both exclude some components of inflation. Once these items are excluded, the inflation rate is re-constructed. The unit root properties of inflation do not matter for this type of core inflation measure.

The paper also considers some time series methods to calculate core inflation, such as the persistence and SVAR measures, and the unit root properties of inflation take on more significance here. There is some doubt regarding the empirical unit root properties of inflation as the results can vary depending on the unit root test employed. Consequently, the SVAR model is estimated twice, first assuming inflation to be stationary and second time assuming a unit root. For reasons explained in the relevant section, the persistence measure is only estimated under the assumption that the component inflation rates are stationary.[1] For the bandpass filter, stationarity is also an issue. In this paper, we only apply the filter to the PCE inflation rate so this implicitly assumes the inflation rate to be stationary. The resulting core series has reasonable properties.

The main dataset used in the calculation of the core inflation rates is the Personal Consumption Expenditure (PCE) dataset from the National Income and Products Accounts (NIPA) tables. For some core measures, only the aggregate PCE inflation rate is needed. For other measures, a detailed breakdown of the PCE based on price indices for 206 separate items is used. This specific breakdown of the PCE together with the associated weights needs to be constructed manually from the data available on the website. Specifically, the series are taken from the underlying data which are available on a quarterly basis at http://www.bea.gov/national/nipaweb/nipa_underlying/SelectTable.asp. A cautionary note from the BEA warns that the underlying data may be of a lower quality than the data normally published. However, when the inflation rates of the 206 items were multiplied by

[1] First differencing non-stationary series and then applying the methodology did not result in a persistence measure that differed systematically from the first measure.

the associated weights, it was possible to recover the aggregate PCE inflation rate with a very high level of precision, indicating that there are no quality issues with this part of the dataset. Quarterly data spanning 1960:1-2008:4 is used. For the structural VAR, data on real GDP over the same time period is also used.

3.1 Hodrick-Prescott Filter

The first estimate of core inflation used in the paper is the Hodrick-Prescott (HP) filter. The filter attributes a certain proportion of each shock hitting the series to a change in the trend of the series while the remainder is regarded as temporary noise. The wide use of HP filters in the profession and their ease of calculation warrants their inclusion in the study. Given the quarterly data used in the study, the standard value of 1600 is chosen for the smoothing parameter. The smoothed series is defined as core inflation. Figure 1 graphs the HP filter and it has the familiar properties.

3.2 PCE excluding Food and Energy

The inflation rate excluding food and energy was included as it is routinely computed by statistical agencies and is one of the most commonly referred to measures of core inflation. The idea is to exclude the items that are normally most volatile. A drawback to this measure is that food and energy are not always the most volatile components in the index. In addition, despite some volatility, they may contain some information regarding the core inflation rate which is lost by total exclusion. Figure 2 shows that this core estimator has been lower than actual inflation at the end of the sample given the high energy and food price inflation experienced in recent years.

3.3 Trimmed Mean

Trimmed Means are commonly constructed as measures of core inflation. The use of trimmed means is motivated by the leptokurtic distribution of individual price changes. This means that price change distributions generally have more extreme values than one would expect from a normal distribution and may be unduly influenced by these extreme values. By trimming the distribution, one removes the influence of these outliers and a more representative measure of the underlying inflation rate is obtained.

The standard approach to calculating a 10% trimmed mean is to order the inflation weights of the individual items from the largest to the smallest. Exclude the largest and smallest price changes (5% in each case for a 10% trim), re-scale the weights of the remaining items so that they again sum to one and calculate the inflation rate again as the weighted average of the remaining items. The problem with this sort of approach is that the weights are based on expenditure shares on a representative basket of goods, devised by statistical agencies to approximate changes in the cost of living. Among others, Wynne (1999) argues that there is no reason to believe that this weighting system should still be used when constructing a core inflation measure, which aims to capture the underlying trend in inflation rather than the cost of living. We argue the weights should be ignored for the following reason.

Consider the case where a 10% trimmed means is calculated using the 13-item breakdown of the PCE inflation rate. Assuming an asymmetric trim, the trimming operation results in just 1 of the 13 items being excluded so, in this case, only the most volatile item is removed. However, in the early part of the sample, the food item had a weight of about 25% and food is often one of the most volatile price indexes. Thus, to trim this inflation rate removes 25% of the index in terms of weights. To rescale the weights of the remaining items and call the resulting series the 10% trimmed mean is misleading. The severity of the problem is lessened when trims are applied to datasets with hundreds of items but the basic criticism still applies. For this reason, we prefer to trim the most volatile inflation rates and take a simple average of the remainder. We refer to the first approach as the weighted trim, as the weights are re-scaled following the trimming operation and refer to this second approach as the simple trim. In this application, the simple trim tends to perform better according to most criteria but only marginally. Figure 3 highlights that the weighted and simple trim have behaved quite similarly over the sample period.

3.4 Band Pass Filter

Following the methodology of Baxter and King (1995), a band-pass filter is applied to the PCE inflation rate to construct another core measure. Band-pass filters are based on a spectral decomposition of the time series and thus operates in the frequency domain of the series rather than the time domain. The spectral representation theorem states that a covariance stationary stochastic process can be expressed as a (infinite) weighted sum of periodic functions. It is the frequency domain analogue of Wold's representation theorem in the time domain. The periodic components are mutually orthogonal and have their own variance. The upshot of this is that we can isolate periodic components at specific frequencies.

The ability to isolate certain frequencies means that new series can be created by filtering out certain periodic components at specific frequencies. The implications in terms of constructing a core inflation measure are obvious. The noise component in the headline rate is defined as the high frequency component. By removing this high frequency component, we are left with an underlying series whose behaviour is driven by long-term trends. It is a more sophisticated approach to removing high frequency noise in comparison with the persistence approach but it also differs to the extent that it is applied directly to the PCE inflation rather than its component parts. The filter is implemented so that components of the series with periodic fluctuations with a frequency of less than one quarter are filtered out. This removes the high frequency component of inflation. Once this high frequency component is removed, the underlying series is defined as core inflation. Figure 4 shows that the high frequency element has contributed significantly to the overall inflation rate, particularly during volatile periods. The filtered series does not have a smooth appearance. A smoother series could have been created by choosing a wider frequency filter but the narrow frequency filter is chosen because it has superior forecasting properties.

3.5 Structural VAR

A bivariate SVAR is also used in the paper to calculate another candidate for core inflation. This is the only structural estimate in the paper; the others are purely statistical in their construction. The variables included in the specification are the inflation rate and real GDP. In order to achieve structural identification, the standard restriction that the core inflation shock is output neutral in the long-run is imposed. This is consistent with the idea of a vertical long-run Philips curve and is a traditional identifying assumption in the application of long-run restrictions. Two core measures are calculated based on the assumption of a stationary and non-stationary inflation series. The two series are found to behave quite differently, as demonstrated in Figure 5.

3.6 Persistence Measure

Persistence measures of core inflation are amongst the least well-known and least widely implemented measures of core inflation. They are somewhat similar in spirit to the exclusion measures, such as inflation excluding food and energy. The exclusion measures exclude high variance components as they are considered to constitute noise. Advocates of the persistence approach prefer to classify noise as high frequency rather than high variance components. The idea is to increase the weight of the persistent components of inflation. This approach has some attractive features empirically. Given the possibility that some price series will exhibit both high variance and persistence, do we really want to exclude items based on variance only? The persistence measures are an attempt to address this short-coming in exclusion measures.

The persistence of a component is measured by estimating an AR model and ranking the magnitude of the autoregressive coefficients. The specific implementation has been approached in two ways in the literature. Cutler (2001) recommends estimating the following AR model using monthly data and annual inflation rates:

$$\pi_{i,t} = \alpha_{i,t} + \rho_{i,t}\pi_{i,t-12} + \epsilon_{i,t} \tag{1}$$

The subscript i is used to index across the various components. The estimated magnitude of the autoregessive coefficient is the persistence estimate. If this coefficient is negative, it is evidence of very fast mean reversion and the item in question is given a zero weight in the persistence measure. For the other components, their weight is proportional to the magnitude of the autoregressive coefficient. This approach is somewhat restrictive in terms of the specification of the autoregressive model. The approach implemented in this paper follows that of Bilke and Stracca (2007), who estimate a model of the form:

$$\pi_{i,t} = \alpha_{i,t} + \sum_{j=1}^{q_i} \rho_{i,j}\pi_{i,t-j} + \epsilon_{i,t} \tag{2}$$

In this case, the lag length of the autoregressive model is chosen according to the Schwartz information criteria. Lag lengths up to twelve lags are considered although in most cases, the lag length chosen was quite short - the average lag length was just over 2. The persistence measure is the sum of the estimated autoregressive components. As in Cutler (2001), items with negative sums are given a zero weight. Following Bilke and Stracca (2007), the inflation rates are re-weighted in proportion to the magnitude of the summed AR weights.

The papers in this area are unclear as to whether each series should be tested for a unit root. When there are a large number of series available, clearly some will have unit roots while others will not. However, as the method involves a re-weighting of inflation rates, it seems logically inconsistent if some items are weighted based on the persistence of inflation rates while others are weighted on the persistence of the first difference of inflation rates. For this reason, we apply the AR model to the inflation rates only, which appears to be the standard approach in the literature. Figure 6 shows that the persistence measure has tracked actual inflation quite closely over the sample.

3.7 Exponential Smoother

Cogley's (2002) exponential smoother also aims to capture persistent movements in inflation. However, this persistence is motivated in terms of the behaviour of central banks. The idea is that shifts in mean inflation arising from changes in policy rules are the main source of inflation persistence and core inflation should be designed to adapt to these changes. The exponential smoother is designed to measure changes in mean inflation, whereby the mean of inflation is updated based on new data according to a constant gain algorithm. This updating rule corresponds to simple exponential smoothing, which is a one-sided geometric distributed lag of past inflation:

$$\pi_t^* = g_0 \sum_j (1 - g_0)^j \, \pi_{t-j} \tag{3}$$

where π^* is the exponential smoother, g_0 is the gain which is calibrated based on the values suggested by Cogley (2002) and π_t is actual inflation. As with other estimators which aim to isolate the persistent elements of inflation, this filter removes the high frequency component. It differs from the persistence estimator in the sense that it is applied directly to the PCE inflation rate and it differs from the HP and bandpass filters to the extent that it is a one-sided filter and so does not suffer from an end-point problem. Figure 7 graphs the exponential smoother and it has the characteristic properties of this type of filter.

3.8 Factor Model

The factor model used follows the approach of Stock and Watson (2002) in that we estimate a static representation of a dynamic factor model. This type of model can be estimated using principal components.[2] Each individual inflation rate is assumed to be driven by a small number of common factors and an idiosyncratic error:

$$\pi_{i,t} = \Lambda_i F_t + \epsilon_{i,t} \tag{4}$$

Each inflation rate is related to the factors with unique factor loadings, Λ_i. This allows the common component of the overall inflation rate to be estimated and is defined as core inflation. Prior to estimation, all inflation rates must be transformed to ensure stationarity. By definition, the resulting core estimator is also stationary. The use of factor models is most common in the pure forecasting literature but it has been fairly widely applied in the core inflation literature also. It represents a hybrid of the statistical approaches in the sense that both time series and cross sectional information is used in its construction. Figure 7 shows the factor estimate of core inflation and it's notable that this estimate was considerably higher than actual inflation during the first oil price crisis.

4 Comparison of Core Measures

Having outlined the core measures included in the paper, we now begin the evaluation process and this section contains the key contributions of the paper. The two standard tests are the ability to track trend inflation and the ability to forecast inflation. Improvements are made to these two standard tests and the tests are applied with the most comprehensive set of core estimators to date. Additional tests not normally found in this literature are also applied. The following results, therefore, provide the most realistic appraisal of the practical usefulness of U.S. core inflation estimators.

4.1 Summary Statistics

To begin the analysis of the various core measures, a couple of basic summary statistics are presented for each core series. Although these are the most basic statistics for any

[2]See Stock and Watson for technical details.

series, it is often argued that they are especially important in the core inflation context. In terms of the mean of the series, one would expect a core inflation rate to have a similar mean to the headline inflation rate when considered over a long time span. If core inflation and actual inflation have significantly different means over a sustained period, the core measure is systematically divergent from the headline rate. Clark (2001) cites similarity of means as one criterion to assess the ability of a core measure to track the trend in inflation as "policymakers and other analysts prefer a measure of core inflation that neither understates nor overstates the long-term trend rate of price change". The importance of the standard deviation lies in the fact some of the core measures are constructed on the basis that volatile components are excluded. Thus, once volatile components are excluded, the resulting series should be less volatile. In this section, we examine the summary statistics of the core measures to see if this is a valid means to discriminate between candidate core series.

Table 1 presents summary statistics for the PCE inflation rate and the core inflation series calculated over the period 1963Q2 - 2008Q4. The PCE inflation rate has a mean of 3.79% over the sample and most of the core measures have a mean which is similar to this. Statistical series generally perform strongly on this criterion. The HP filter posts a mean inflation rate of 3.80%. The I(0) SVAR, the persistence measure and the factor model estimate all have means which differ from the PCE mean by less than 0.03.% The band pass filter has a mean of 3.69%, again quite similar to the PCE inflation rate. As all core inflation rates have very similar means over such a long sample, this is not a suitable statistic to choose the best measure.

A second summary statistic often cited in relation to core inflation rates is their variance or standard deviation. Certain core estimators are designed to remove high frequency noise and according to this criterion, good estimators should have a lower variance or standard deviation than the actual inflation rate. The second column of the table shows the standard deviation of inflation and the core measures. The HP filter, SVAR and exponential smoother all do well on this metric. However, the difference in volatility is not of a sufficient magnitude to meaningfully discriminate between these core measures. In addition, six core series have a standard deviation between 1.98, which is the lowest value, and 2.10. The table also presents the correlation of the core measures with the PCE inflation rate and their correlation with a centred moving average of inflation. Again, nearly all series are highly correlated with the PCE inflation rate. They are also highly correlated with the centred moving average so the summary statistics do not provide a basis for choosing amongst core inflation rates.

4.2 Tracking Trend Inflation

The ability to track trend inflation is often considered a key property of a good core inflation rate. Bryan et al (1997), Cecchetti (1997) and Clark (2001) all define the trend in inflation as a Centred Moving Average (CMA) of the headline inflation rate and this is the standard definition of the trend in inflation in the literature. One shortcoming in the literature is that a core inflation rate constructed using the full sample of data is used as the basis for comparison with the CMA trend. In this paper, we construct core inflation rates recursively to more accurately reflect the situation faced in real time. This has two important benefits. Firstly, a common criticism of econometric estimates of the core rate is that it changes every quarter as the model is re-estimated with additional data. By estimating the core inflation rates recursively, we construct the core measure that would have been available to policymakers at each point in time. The core series are estimated recursively over 1960Q1-1989Q1 in the first step and over 1960:1-2008Q4 in the last step, which represents a period of twenty years. The core inflation estimate for the current quarter of each recursive step is compared with the trend. The second benefit of the recursive estimation strategy is that it takes account of the end-point problem with statistical filters and so provides a more realistic indication of their ability to track trend inflation.

Table 2 presents the results of the recursive trend tracking ability of the core inflation series. The first column shows the correlation of the core series with a nine quarter CMA, which is used as the estimate of the trend. When we compare the correlations in this column to those in the last column of Table 1, which represented correlations with core rates estimated over the full sample, we can see that in most cases, there is a slight decline in the core series' ability to track trend inflation once the exercise is performed recursively.[3] The correlation with the CMA trend declines significantly for the stationary SVAR measure of core inflation although the persistence estimate actually increases its ability to track the trend. The HP filter has a high correlation with the trend even though it is reduced to a one-sided filter at the end of the sample. There are six core series which have a correlation

[3] The correlations from the two tables are not strictly comparable as the sample period is considerably shorter for the recursively estimated series in Table 2.

of 0.90 or higher with the CMA trend so this criterion again leaves little to choose amongst the core series.

In general, trend tracking tests have no benchmark. In this paper, a benchmark defined as a five quarter moving average is introduced. This benchmark is adopted as it is trivial to compute and is a one-sided filter. Thus, if we choose to define the trend as a 9-quarter centred moving average, we can choose a benchmark defined as the currently available part of that centred moving average. The last row of the table shows the correlation of this 5-quarter moving average with the centred moving average trend. This benchmark correlates with the trend just as highly as any core series. Therefore, the core inflation rates perform no better than this benchmark.

A second way to measure the ability to track a CMA trend is to calculate average deviations from the trend. The second column of the table presents the Relative Mean Absolute Error (RMAE), calculated over the recursive sample using the formula:

$$RMAE_i = \frac{\sum \|(\pi_{core,i} - CMA)/CMA\|}{n} \tag{5}$$

This calculates the absolute difference between the core series and the trend as a fraction of the trend and finds the average. The persistence series performs well according with a value 0.09 for this statistic. However, the 5-quarter moving average has an average error of 0.10 and the trims both have errors of 0.11. Thus, the range of values for this statistic is quite narrow and no core series dominates. On the basis of the two trend tracking tests, a number of series core inflation series perform quite well but there is no clear front-runner. In addition, the benchmark does just as good a job of tracking the trend when the standard definition of a centred moving average is adopted. Consequently, the ability to track trend inflation, whether through correlations with the trend or deviations from it, is not a fruitful avenue in terms of ranking core inflation series. More importantly, core inflation rates are no more useful than a lagged moving average in terms of tracking a trend when the trend is defined as a centred moving average.

4.3 Forecasting Headline Inflation

The ability to forecast inflation is cited as a key indicator of the policy usefulness of core inflation rates. In this section, competing measures of core inflation are ranked according to their ability to forecast the headline inflation rate and a benchmark forecast is included in the analysis. Forecasts are constucted using the following regression, which is the standard forecasting equation in the literature:

$$\pi_{t+h} - \pi_t = \alpha + \beta\left(\Pi_t - \pi_t\right) + \upsilon_t \tag{6}$$

where π_t is the inflation rate at time t and Π_t is core inflation. The left hand side of the equation is the difference between headline inflation today and headline inflation h periods in the future. On the right hand side, the term in brackets is the difference between core inflation and headline inflation. The basic premise of this forecasting regression is that the difference between headline inflation and core inflation today has predictive power for headline inflation tomorrow. In particular, if there is a large divergence between headline inflation and core inflation, you would expect headline inflation to move back towards core inflation because core inflation is a measure of the general trend in inflation.

The regression computes a forecast over a fixed horizon. For example, using quarterly data and setting $h = 8$ would yield a forecast of headline inflation eight quarters in the future but would not forecast inflation in the intervening periods. In order to get a continuous forecast to the end of the forecasting horizon, eight quarters in this paper, eight regressions of the type above are estimated setting $h = 1...8$. Each candidate core inflation rate is put in the regression equation above and forecasts of the headline inflation rate are generated. A "no change" benchmark forecast is used to compare the performance of the core series. Under this scenario, if inflation is 4 per cent in 2000Q1, the forecast for year-on-year inflation for each quarter in the forecast horizon 2000Q2-2002q2 is also 4 per cent.

The quarterly forecasts are performed on a recursive basis, with one observation added to the sample each time. In the first recursive step, estimates of core inflation are calculated over the sample 1960Q1-2000Q1 and forecasts are performed up to 2002Q2. The process is repeated adding one observation each time so by the end of the final estimation period, there are 28 sets of forecasts for each core estimation method. The forecasting exercise is repeated using only data from the start of the Great Moderation period. In these short sample estimates, the first recursive series are estimated from 1982Q1-2000Q2. Although 1985 is generally accepted as the beginning of the Great Moderation, we choose a period a few years earlier to begin estimation in order to allow more degrees of freedom in the estimation of the econometric series, particularly the SVAR measures. The forecast periods and number of recursive estimates are identical in both forecast exercises; only the estimation period changes. However, we also conduct a third forecast exercise over the full sample when the number of recursive steps is doubled to 56 as a robustness check. In this instance, the first estimation period runs from 1960Q1-1993Q1.

The results of the full sample forecasting exercise with 28 recursive steps are presented in Table 3. The numbers in the table are the ratios of the RMSE from the regression forecasts to the no change benchmark. A value less than one indicates that it is more accurate to forecast inflation using the regression forecast. From Table 3, we can see that the core inflation rates perform very poorly in terms of forecasting. The SVAR, where inflation is assumed to be I(1), has the best forecasting power. It is much more accurate than any other core measure, particularly at longer forecast horizons. However, it is less accurate than the no change forecast except for quarters five and six. Even then, the improvement in forecast power relative to the benchmark is marginal. The short sample estimates in Table 4 paint a similar picture. Table 5, which presents the full sample estimates with additional recursive steps, again shows that no core rate outperforms the benchmark. As this has the largest number of recursive steps, the results of this exercise are potentially the most robust. For this reason, formal forecast comparison tests are performed on the forecasts in this table. The Diebold-Mariano (1995) test of equal predictive ability is almost universally rejected. This indicates that the core inflation based forecasts are statistically inferior to the no change forecast. The only exceptions are the forecasts of quarter 1 and 2 from the HP filter. The systematic failure of core inflation regressions to beat a naive benchmark indicates that core inflation rates are not a useful tool in terms of forecasting PCE inflation.

4.4 Directional Forecasting

Although core inflation rates do a poor job of forecasting the magnitude of inflation, perhaps they are more suitable to predicting changes in the direction of inflation. Taking the forecasts from the previous section, they are evaluated according to their ability to correctly forecast the direction of the change in inflation four quarters ahead and eight quarters ahead. The forecasts are available over the full and short sample with 28 recursive steps and over the full sample with 56 steps. There is no benchmark per se in this exercise although one would wish that the forecasts would beat a coin toss so that the correct direction is forecast at least 50% of the time. The results presented in Table 6 give the percentage of times that the models correctly forecast the direction of change in inflation.

The core inflation rates do not generally perform well according to this statistic. If we look at the first two columns of the table, which represent the full sample estimates with 28 steps, the I(1) SVAR and the simple trim are the only two series to correctly forecast the direction of the change in inflation more than 50% of the time over both four and eight quarters. Columns three and four show the short sample results. The I(0) VAR correctly predicts the direction of change 64% of the time four quarters ahead while the excluding food and energy series does well for the eight quarter forecast. The factor model beats a coin flip for both horizons. For the full sample results with 56 steps, the I(1) SVAR is the only core rate with forecast accuracy greater than 50% at either four quarters or eight quarters. Taking the results as a whole, the failure of any core rate to systematically (i.e. across forecast exercises) beat a coin flip in terms of directional forecasting highlights major shortcomings in core rates as forecast tools. However, the I(1) SVAR is a front-runner in this exercise as it beats a coin flip in the two full sample exercises.

4.5 Concordance

Concordance is a broad measure of the degree to which the various core inflation rates agree with each other in terms of whether core inflation is above or below actual inflation. For example, if one core measure shows core inflation to be above actual inflation but all the others show it to be below actual inflation, one would conclude that it is below on the balance of evidence. A concordance measure puts this type of logic on a firmer statistical footing. In this context, the concordance statistic is a bivariate statistic that measures the degree to which two core inflation rates agree that core inflation is above/below the headline rate. More specifically, it measures the proportion of the time that two series are in the same state. If we define an inflation gap for each core series as the difference between the candidate core measure and headline inflation, we can define a corresponding series $S_{i,t}$ to be equal to 1 when the gap measure is positive and equal to 0 when the gap measure is negative, where the subscript i is an index over the different core inflation series. The degree of concordance for a pair of gap measures is then calculated as:

$$C_{i,j} = T^{-1} \sum \{(S_{i,t}.S_{j,t}) + (1 - S_{i,t})(1 - S_{j,t})\} \tag{7}$$

By construction, the value of the concordance statistic is bounded between zero and one. A value of 0.5 between two core series means that, 50% of the time, the sign of the inflation gap is the same when calculated using both core inflation rates. The concordance statistics are presented in Table 5. The core inflation rate with the highest average concordance is the exponential smoother. On average, it is in agreement with the other core inflation rates 71% of the time regarding the sign of the inflation gap. The excluding food and energy measure also performs well with average concordance of 70%. The I(1) SVAR has the least satisfactory performance according to this statistic.

Although the results appear reasonable here, there are also difficulties with this statistic in terms of ranking core inflation rates. The range of values for the statistic is again quite tight with five core inflation rate scoring between 0.66 and 0.71. The concordance statistic does not separate the different core inflation rates any clearer than the trend tracking statistic. Also, following the poor results of the directional forecasting exercise, one has to question whether any core rate is consistently measuring excess inflationary pressure in the economy.

5 Summary and Conclusions

The implementation of effective monetary policy requires an accurate assessment of the rate of core inflation in an economy. Like other important concepts such as potential output and the NAIRU, the core inflation rate is not an actual series and instead must be estimated. This paper conducts the most rigorous and comprehensive analysis of existing estimates of core inflation to date. There are other papers of this variety in the literature but these often focus on a specific type of core inflation estimator. This paper compares all major estimation methods. The exercise is conducted for the US and improvements are made to the standard comparison tests. In addition, extra tests not generally used in this literature are also applied to the core inflation rates. Two problems emerge in the comparison exercise.

Firstly, the candidate core inflation rates are very difficult to separate according to the comparison tests as a large number of estimators generally perform to a very similar level.

This makes it very difficult to rank the core inflation rates. The two standard tests of core inflation are its ability to track trend inflation and its ability to forecast future inflation. Comparisons are mostly conducted just amongst the core rates. When simple benchmarks are included, no core rate can outperform the benchmark in either test. This calls into question the usefulness of existing core inflation measures. Additional tests not featured in the literature are also examined but the performance of existing core inflation estimates is still relatively poor.

As the literature has not highlighted these shortcoming of core inflation rates to date, future work is needed to determine if these results are specific to this dataset or perhaps specific to the US. It is difficult to foresee how the trend tracking ability of core inflation rates will compare in other studies. However, the general difficulties in forecasting US inflation in the post Moderation period suggest that the forecasting results are unlikely to be overturned for US data although factor model core estimators have demonstrated good forecasting properties for other countries. Also, given the wide variety of increasingly sophisticated techniques, it seems unlikely that forecasting U.S. inflation using OLS on an inflation gap will prove the best approach.

References

[1] Baxter, M. and R. King, (1999), "Measuring Business Cycle: Approximate Band-pass Filter for Economic Time Series", *The Review of Economics and Statistics*, Vol. 81, pp. 575 - 593.

[2] Bilke, L. and L. Stracca (2007), "A persistence-weighted measure of core inflation in the Euro Area", *Economic Modelling*, Vol. 24, pp. 1032 - 1047.

[3] Bryan, M., S. Cecchetti, and R. Wiggins II (1997), "Efficient Inflation Estimation", *NBER Working Paper*, No. 6183.

[4] Cecchetti, S. (1997), "Measuring Short-Run Inflation for Central Bankers", *Federal Reserve Bank of St. Louis Economic Review*, May/June, pp. 143-155.

[5] Cogley, T. (2002), "A Simple Adaptive Measure of Core Inflation", *Journal of Money, Credit and Banking*, Vol. 34(1), pp. 94-113.

[6] Clark, T. (2001), "Comparing Measures of Core Inflation", *Federal Reserve Bank of Kansas Economic Review*, 2nd Quarter.

[7] Cutler, J. (2001), 'A New Measure of Core Inflation in the UK", *MPC Unit Discussion Paper*, Vol. 3.

[8] Diebold F. and R. Mariano, (1995), "Comparing Predictive Accuracy", *Journal of Business and Economic Statistics*, Vol. 13, pp. 253-263.

[9] Landau, B. (2000), "Core Inflation Rates: A Comparison of Methods based on West German Data",*Discussion Paper Series 1: Economic Studies*, 2000/04, Deutsche Bundesbank.

[10] McDermott, J. and A. Scott (1999), "Concordance in Business Cycles", *Reserve Bank Discussion Paper*, G99/7, Reserve Bank of New Zealand.

[11] Quah, D. and S. Vahey (1995), "Measuring Core Inflation", *The Economic Journal*, Vol. 105, pp. 1130 - 1144.

[12] Rich, R. and C. Steindel (2007), "A Comparison of Measures of Core Inflation", *Economic Policy Review*, Federal Reserve Bank of New York, December, pp. 19-38.

[13] Smith J. (2007), “Weighted Median Inflation: Is This Core Inflation”, *Journal of Money, Credit and Banking*, Vol. 36(2), pp. 253-263.

[14] Stock, J. M. Watson (2002), “Macroeconomic Forecasting using Diffusion Indexes”, *Journal of Business and Economic Statistics*, Vol. 20(2), pp. 147-162.

[15] Wynne, M. (1999), “Core Inflation: A Review of Some Conceptual Issues”, *European Central Bank Working Paper*, No. 5.

Figure 1: HP Filtered Inflation Rate

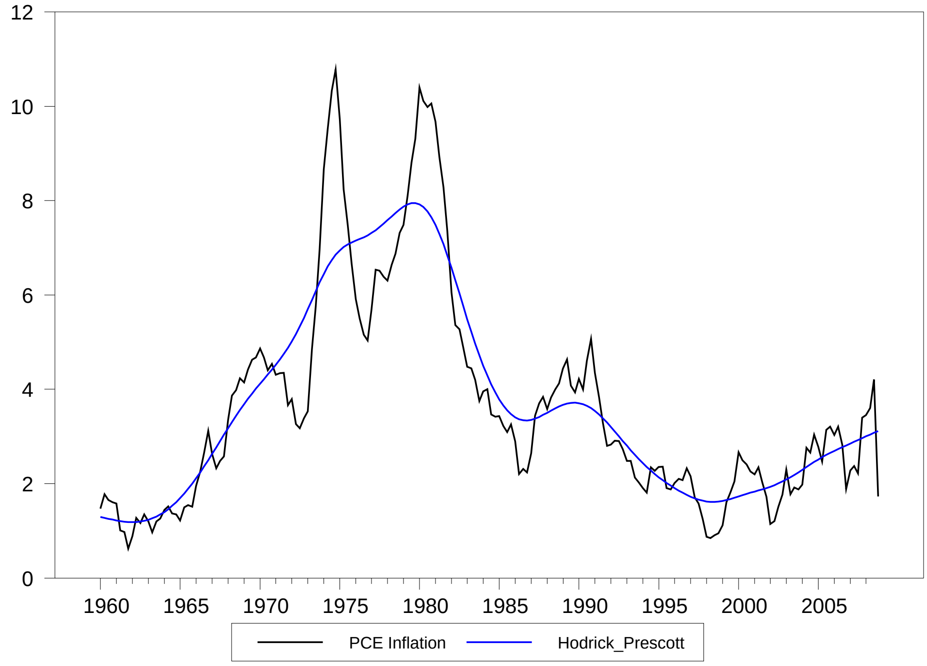

Figure 2: Inflation and Inflation excluding Energy and Food

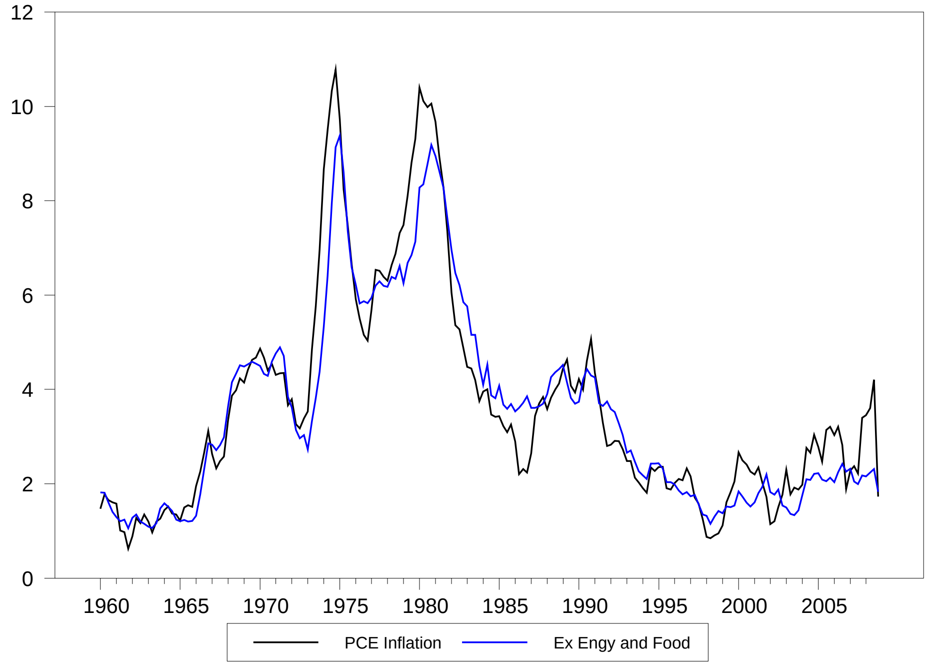

Figure 3: Inflation with Simple and Weighted Trimmed Means

Calculated using all 206 series

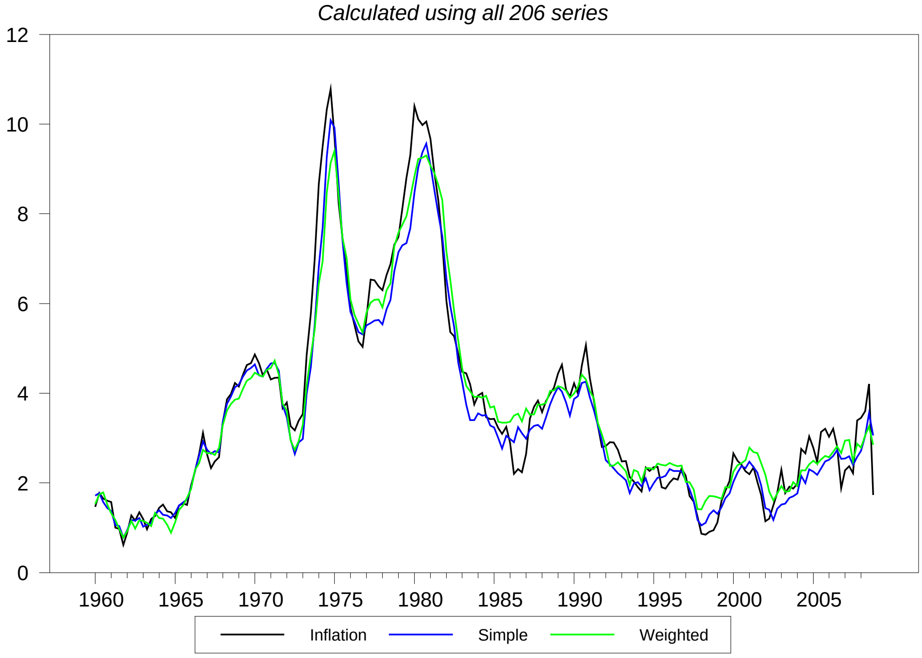

Figure 4: Inflation and Band Pass Filter

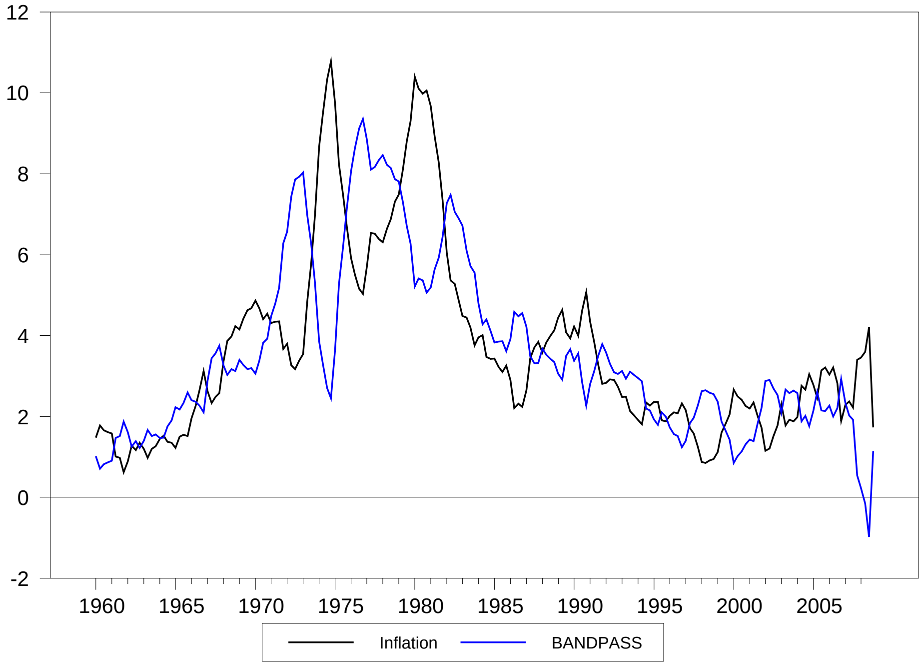

Figure 5: Inflation and SVAR Estimates

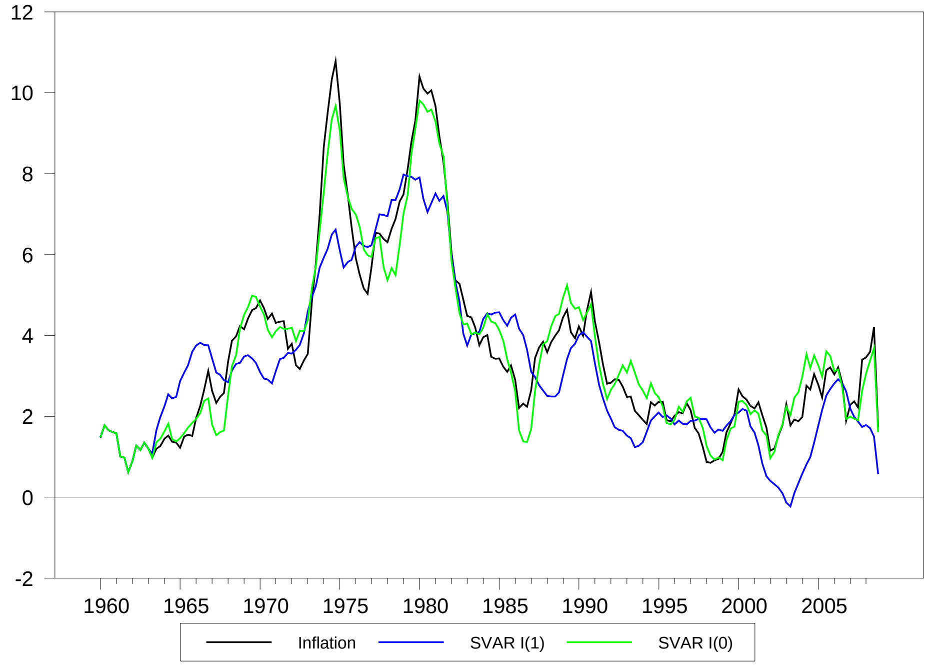

Figure 6: Inflation and Persistence Series

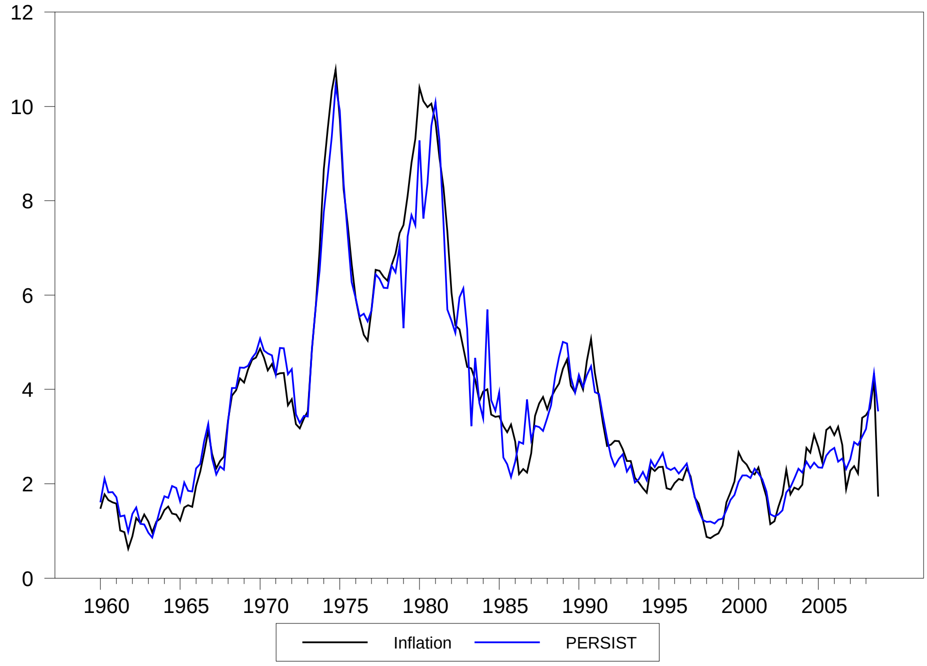

Figure 7: Inflation and Exponential Smoother

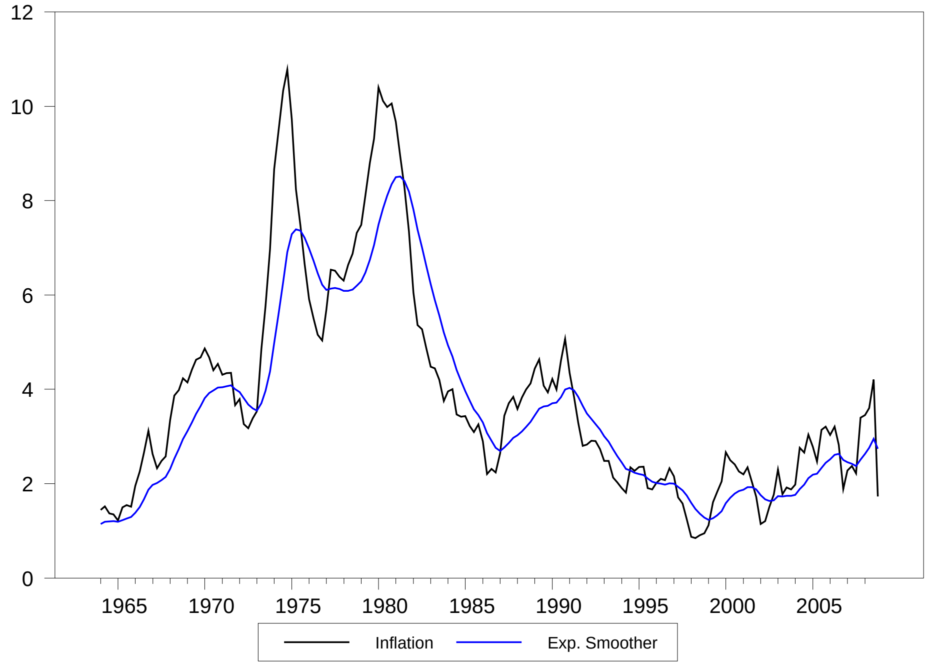

Figure 8: Inflation and Factor Model Core Estimate

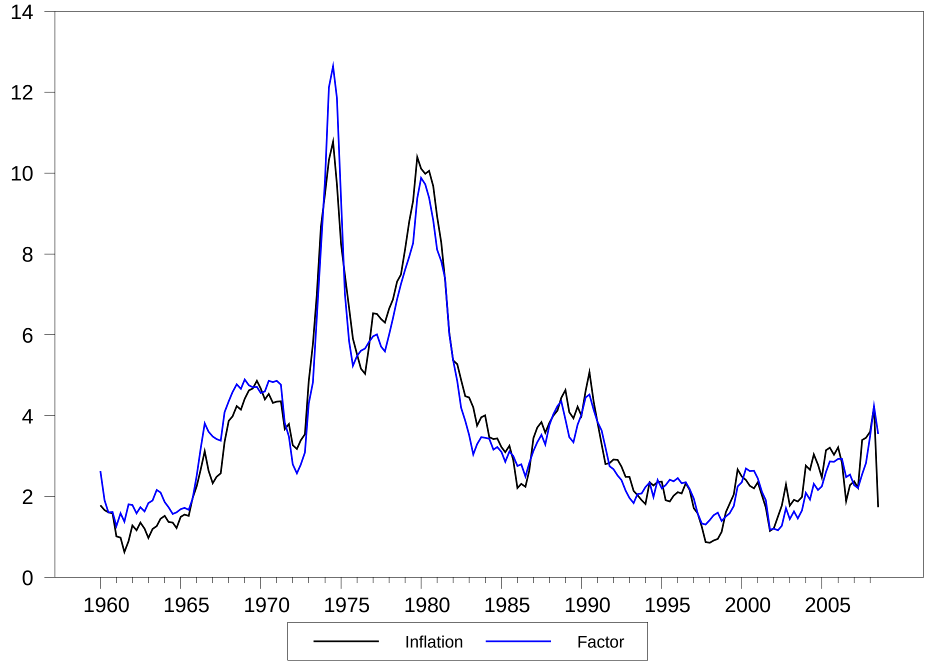

Table 1: **Summary Statistics for Core Measures**

Core Measure	Mean	Std. Dev.	Corr. PCE	Corr. CMA
PCE	3.79	2.31	1.00	NA
HP FILTER	3.80	2.00	0.91	0.96
EXC. FOOD & ENERGY	3.64	2.07	0.94	0.95
TRIM	3.57	2.10	0.98	0.98
WTRIM	3.74	2.07	0.97	0.98
SVAR I(1)	3.37	2.00	0.86	0.89
SVAR I(0)	3.77	2.19	0.98	0.96
PERSIST	3.76	2.11	0.97	0.93
BAND PASS	3.69	2.14	0.57	0.69
EXP. SMOOTH	3.51	1.98	0.90	0.94
FACTOR	3.77	2.27	0.97	0.95

Table 2: **Recursive Trend Tracking Test**

Core Measure	Corr CMA	Deviations
HP FILTER	0.92	0.13
EXC. FOOD & ENERGY	0.90	0.13
TRIM	0.94	0.11
WTRIM	0.93	0.11
SVAR I(1)	0.87	0.14
SVAR I(0)	0.84	0.16
PERSIST	0.95	0.09
BAND PASS	0.75	0.23
EXP. SMOOTH	0.95	0.13
FACTOR	0.90	0.14
MOV5	0.95	0.10

Note: The table shows the ability of each core measure to track the trend based on correlation to or deviations from the trend.

Table 3: RMSE from Full Sample Inflation Forecasts

Horizon	**Forecast Method**				
Quarters	SVAR I(1)	SVAR I(0)	HP Filter	BP Filter	EXP Smooth
1	1.01	1.07	1.04	1.43	1.12
2	1.01	1.15	1.11	1.58	1.26
3	1.00	1.22	1.22	1.51	1.41
4	1.00	1.24	1.35	1.43	1.53
5	0.98	1.30	1.47	1.19	1.64
6	0.99	1.35	1.58	1.09	1.72
7	1.06	1.33	1.62	1.14	1.71
8	1.05	1.33	1.66	1.31	1.72

	Forecast Method				
Quarters	Persistence	Ex. Food Engy	Trim	WTrim	Factor
1	1.06	1.23	1.19	1.20	1.01
2	1.14	1.48	1.37	1.43	1.04
3	1.22	1.70	1.49	1.62	1.08
4	1.28	1.81	1.57	1.76	1.13
5	1.29	1.96	1.60	1.89	1.21
6	1.24	1.99	1.56	1.90	1.28
7	1.21	1.88	1.48	1.82	1.32
8	1.19	1.87	1.38	1.79	1.39

Note: The table shows the ratio of the RMSE from a regression with the named core inflation rate to the "no change" forecast. A value less than one signifies a lower forecast error than the benchmark forecast. First estimation period: 1960:1-2000:1

Table 4: RMSE from Post Moderation Sample

Horizon	Forecast Method				
Quarters	SVAR I(1)	SVAR I(0)	HP Filter	BP Filter	EXP Smooth
1	1.02	1.04	1.12	1.47	1.17
2	1.04	1.05	1.27	1.68	1.34
3	1.04	1.03	1.41	1.64	1.48
4	1.02	1.03	1.56	1.55	1.59
5	1.01	1.02	1.70	1.35	1.69
6	1.02	1.02	1.80	1.27	1.75
7	1.03	1.04	1.85	1.32	1.74
8	1.03	1.04	1.96	1.50	1.81

Horizon	Forecast Method				
Quarters	Persistence	Ex. Food Engy	Trim	WTrim	Factor
1	1.02	1.16	1.10	1.12	1.09
2	1.04	1.38	1.23	1.28	1.21
3	1.04	1.61	1.30	1.37	1.27
4	1.04	1.77	1.38	1.46	1.34
5	1.03	1.92	1.40	1.51	1.28
6	1.04	1.95	1.38	1.48	1.19
7	1.02	1.86	1.38	1.46	1.16
8	1.05	1.91	1.40	1.50	1.10

Note: The table shows the ratio of the RMSE from a regression with the named core inflation rate to the "no change" forecast. A value less than one signifies a lower forecast error than the benchmark forecast. First estimation period: 1982:1-2000:1

Table 5: RMSE from Full Sample with Additional Recursive Steps

Horizon	**Forecast Method**				
Quarters	SVAR I(1)	SVAR I(0)	HP Filter	BP Filter	EXP Smooth
1	1.01*	1.05*	1.02*	1.05	1.10
2	1.03*	1.10	1.06*	1.16	1.24
3	1.06*	1.12	1.14	1.29	1.38
4	1.08*	1.19	1.26	1.42	1.52
5	1.12*	1.26	1.37	1.55	1.64
6	1.16*	1.31	1.47	1.65	1.73
7	1.22	1.32	1.54	1.70	1.75
8	1.25	1.33	1.58	1.76	1.78

	Forecast Method				
Quarters	Persistence	Ex. Food Engy	Trim	WTrim	Factor
1	1.03*	1.18	1.23	1.26	1.02*
2	1.06*	1.37	1.43	1.52	1.06*
3	1.08	1.53	1.55	1.71	1.12
4	1.09	1.63	1.59	1.84	1.20
5	1.08	1.73	1.60	1.93	1.29
6	1.05*	1.76	1.54	1.94	1.39
7	1.01*	1.72	1.46	1.88	1.44
8	1.00*	1.71	1.38	1.83	1.52

Note: The table shows the ratio of the RMSE from a regression with the named core inflation rate to the "no change" forecast. A value less than one signifies a lower forecast error than the benchmark forecast. First estimation period: 1960:1-1993:1.* indicates that the null hypothesis of equal predictive ability is rejected at the 10% level. Rejection of null indicates that core inflation forecasts are not statistically inferior to benchmark. In no case are the core forecasts statistically superior however.

Table 6: Directional Forecasts for Core Measures

	Long, r = 28		Short, r = 28		Long, r = 56	
Core Measure	Q = 4	Q = 8	Q = 4	Q = 8	Q = 4	Q = 8
HP Filter	0.32	0.50	0.25	0.29	0.37	0.41
Persist	0.21	0.46	0.50	0.39	0.41	0.45
SVAR I(1)	0.68	0.57	0.43	0.39	0.59	0.52
SVAR I(0)	0.32	0.43	0.64	0.43	0.39	0.32
Ex. Food Engy	0.32	0.50	0.32	0.54	0.38	0.48
BP Filter	0.11	0.36	0.18	0.36	0.20	0.34
EXP Smooth	0.21	0.32	0.29	0.39	0.38	0.45
WTRIM	0.43	0.46	0.43	0.50	0.45	0.43
TRIM	0.54	0.58	0.36	0.39	0.46	0.46
Factor	0.50	0.39	0.57	0.54	0.48	0.48

Note: The table shows the percentage of the time that the core inflation rate correctly predicts the direction of future price changes four quarters and eight quarters ahead. Like the previous forecast exercise, results are presented for the full sample with 28 recursive steps, the post Moderation sample with 28 recursive steps and the full sample with 56 recursive steps. One would expect a good forecast model to beat a coin flip in the sense that it would forecast the direction of inflation correctly 50% of the time.

Table 7: Concordance of Core Inflation Measures

Core Measure	Persist	SVAR I(1)	SVAR I(0)	HP Filter	Ex. Food Engy
Persist	1.00	0.63	0.59	0.55	0.64
SVAR I(1)	0.63	1.00	0.49	0.53	0.61
SVAR I(0)	0.59	0.49	1.00	0.54	0.58
HP Filter	0.55	0.53	0.54	1.00	0.79
Ex. Food Engy	0.64	0.61	0.58	0.79	1.00
BP Filter	0.63	0.68	0.56	0.70	0.79
Exp Smooth	0.58	0.58	0.51	0.88	0.89
WTrim	0.69	0.59	0.43	0.66	0.65
Trim	0.68	0.63	0.44	0.68	0.69
Factor	0.60	0.63	0.44	0.63	0.66
AVERAGE	0.62	0.59	0.51	0.66	0.70

Core Measure	BP Filter	Exp Smooth	WTrim	Trim	Factor
Persist	0.63	0.58	0.69	0.68	0.60
SVAR I(1)	0.68	0.58	0.59	0.63	0.63
SVAR I(0)	0.56	0.51	0.43	0.44	0.44
HP Filter	0.70	0.88	0.66	0.68	0.63
Ex. Food Engy	0.79	0.89	0.65	0.69	0.66
BP Filter	1.00	0.78	0.64	0.65	0.55
EXP Smooth	0.78	1.00	0.71	0.78	0.68
WTrim	0.64	0.71	1.00	0.84	0.79
Trim	0.65	0.78	0.84	1.00	0.85
Factor	0.55	0.68	0.79	0.85	1.00
AVERAGE	0.66	0.71	0.67	0.69	0.65

Note: The table indicates the degree to which different core measures agree on the sign of the inflation gap. The table needs to be read as a grid reference. For example, the number 0.63 in the second row of the first column indicates that the persistence measure and the SVAR I(1) agree on the sign of the inflation gap 63% of the time. Averages are also provided for each core measure.