Core Inflation: Measurement and Statistical Issues in Choosing Among Alternative Measures

MICK SILVER*

This paper provides an overview of statistical measurement issues relating to alternative measures of core inflation and the criteria for choosing among them. The approaches to measurement considered include exclusion-based methods, limited influence estimators, reweighting, and economic modeling. Criteria for judging which approach to use include credibility, control, deviations from a smoothed reference series, volatility, predictive ability, causality and co-integration tests, and correlation with money supply. The extensive number of measures, criteria by which their appropriateness is assessed, and lack of consistency in the results of country studies argue for a data-driven approach. [JEL C43, E31, E52]

Countries that adopt inflation targeting require a credible, timely measure of inflation to target, and the consumer price index (CPI) is usually adopted for this purpose. Because the price changes of some components of the CPI, including food and vegetables (because of weather conditions) and energy (because of supply shocks), are particularly volatile,

these components are usually excluded from the target. Indirect taxes and interest (mortgage) payments may also be excluded because the former are erratic one-off changes and the latter are a tool of, and therefore should not be a goal of, monetary authorities. Such resulting ''core inflation'' measures are used for inflation targeting, though it is not always clear which components should be excluded.

Inflation targeting benefits from the use of a credible target measure, and the CPI itself may be used as an inflation target if the exclusion of product groups is likely to be perceived as undue manipulation of the target. Core inflation measures may also be used by the monetary authorities as operational guides for analytical and forecasting purposes with respect to achieving the target. In this context, a wider range of core inflation measures can be used for different purposes and their degree of complexity increased. Our concern would no longer be solely with an appropriate target measure, which may be the CPI or a core inflation measure, but with a suite of operational measures of core inflation to be used to better target CPI inflation.

The paper outlines some practical issues regarding CPI measurement to provide a context to the discussions on core inflation measurement. This takes us to the main purpose of the paper, which is to first outline the range of methods available. Although exclusion-based measures are often used as target and/or operational measures of core inflation, there are many alternative approaches that can provide quite different results. Furthermore, the same approach can generally be implemented in different ways and again, this can lead to quite different results. Heath, Roberts, and Bulman (2004), *after* putting aside the less-likely measures, consider 102 measures of core inflation using Australian data.

Because the choice of method matters empirically, there is the second need to choose between the alternative methods. This in turn requires criteria by which different methods can be chosen. The problem with choosing core inflation measures arises not just from the extensive range of possible methods but also from the extensive range of criteria by which measures can be judged to be appropriate. This paper outlines criteria for judging which core inflation measures are the best. There are also a good number of empirical studies to draw on, but they vary according to the country, time period, criteria for selection adopted, and measures considered. Even after taking such variations into account, no unanimity as to the best measures emerges, with conclusions changing even for subperiods of the same country study.

The emerging consensus and, indeed, practice is to use more than one measure for operational purposes (see Roger, 2000; Heath, Roberts, and Bulman, 2004; and Mankikar and Paisley, 2004). If the resulting measures give similar results, then they should give confidence to the monetary authorities in making decisions based on such measures. If they do not, differences in the nature of the measures used should, by construction, allow for insights into the inflationary process.

It is argued in this paper that the evaluation should, where possible, be data driven, that is, be evaluated using country data and objective, transparent criteria. The empirical studies clearly show that methods suitable in one country cannot necessarily be carried over to other countries.

I. The CPI: Practical Considerations

Successful inflation targeting requires confidence that the monetary authorities can achieve the target (range) and confidence that the target is a meaningful measure. To be effective, the index should be one that the public is familiar with. The CPI has the requisite high public profile, and thus its prevalence as a target. Bloem, Armknecht, and Zieschang (2002) provide not only an outline of, but also a case for, alternative indices as the basis for core inflation measure.

Attention should be given to the institutional arrangements regarding the production of core inflation measures. To be effective in anchoring inflation, the target index must be credible, and credibility is derived not just from the quality and suitability of the measures, but also from the transparency of the source data and compilation methods and the credibility of the agency concerned. Generally, an autonomous statistical agency is best placed to fulfill this function, but the central bank has a major role to play in the development of these measures because their use is for monetary policy purposes.

Core inflation measures are derived from CPI data, and errors and bias in CPI data may well pass on to core inflation measures. An understanding of sources of errors and bias in the CPI, outlined in Greenlees and Balk (2004), is important to the successful use of core inflation measures.

Roger (2000) argues a second reason that CPI measurement methods have an impact on the measurement of core inflation. Some of the measures and tests as to which measure is best rely on a characterization of core inflation given in Equation (2) below, whereby shocks are random and normally distributed. Shocks affect relative prices, but in the long run they are held to be accommodated by flexible price setting and consumer substitution toward (away from) products with below (above) average price changes. However, the formulas used for price indices use a fixed basket, and the substitution effects are not accounted for in the measure. Furthermore, the coverage of the CPI may not include all goods and services, and so the full balancing act may not be revealed.

II. The Methods

The methods can be grouped into those suitable for policy assessment, that is, those designed to strip away the noise to identify the signal, and those formulated to predict inflation. There is, of course, something in stripping away noise that makes us better placed to predict, and something in devising a method for prediction that requires the noise to be stripped away. But it

will be apparent from the measures outlined in this section and the criteria/tests for choice of measure in Section III that these needs are quite distinct.

Policy Assessment

- Exclusion-based methods

  - Product groups
  - Indirect taxes
  - One-off shocks
  - Domestically generated inflation
  - Imputation methods

- Trend estimates
- Limited influence estimators

  - Median
  - Trimmed means—symmetric and asymmetric

Prediction

- Reweighting the CPI

  - Persistence weights
  - Volatility weights
  - First principal component

- Economic models

We define the CPI, π_t, as a weighted mean of price changes, $\dot{p}_i^t$, for product group i, from a price reference period 0 to the current period t, and the weights, w_i^b, are normalized, relative expenditures in some period b, prior to period 0:

$$\pi_t = \sum_i w_i^b \frac{p_i^t}{p_i^0} = \sum_i w_i^b \dot{p}_i^t. \tag{1}$$

In practice, the $\dot{p}_i^t$ are generally unweighted elementary aggregate indices for a product group/outlet type/region stratum derived from matched price comparisons for similar items.

We initially characterize CPI inflation, π_t, as core inflation, π_t^*, plus a temporary disturbance, v_t; that is,

$$\pi_t = \pi_t^* + v_t, \tag{2}$$

where v_t is random and *normally* distributed. Implicit in the concept of core inflation is that transitory relative price shocks should not be allowed to influence core inflation. This is because, in theory, prices are expected to be fully anticipated and be "accommodated" by relative price and quantity changes. The shock should not influence mean inflation. The characterization

of v_t in Equation (2), as normally distributed, will be reviewed later in Sections II and III in light of empirical evidence of non-normal price change distributions. Given this concern with this characterization of v_t as normally distributed, it is reasonable to ask why it is phrased as such up front in this paper. This is because there is much in the way that shocks are understood, measures currently used are justified, and criteria are applied for the selection of measures (particularly the test approach in Section III) that rely on this characterization. It is adopted here and relaxed later.

Measures of core inflation are thus required to separate the signal of inflation from the temporary noise or volatility—what Cecchetti (1997) refers to as the "transitory phenomenon" that should not affect policymakers' actions.

Core inflation measures should not be subject to seasonal fluctuations. Twelve-month rates compare the price level in a month with that in the same month in the preceding year, which may be sufficient to remove seasonality. Also, there may well be an interest in month-on-month changes for which core inflation measures may be seasonally adjusted (Fenwick, 2004). For transparency, seasonal factors can be derived from the publicly available X12 program developed by the U.S. Bureau of the Census, although see Diewert (1999) for a moving-year centered index number approach.

Cecchetti (1997) finds that the extent of seasonal variation is quite substantial over subperiods of his study and that core inflation measures cannot be relied on to remove seasonality. He finds a marked *increase* in seasonality when food and energy (F&E) are excluded.

We now consider methods suitable for policy assessment.

Exclusion-Based Methods

Exclusion-based methods exclude component price indices of a CPI that are considered to be particularly volatile. An exclusion-based CPI is easy to understand, timely, and transparent in that the user can replicate the measure. Exclusion-based methods are often used by countries when they first instigate inflation targets. Usual exclusions are F&E, argued on the basis of their undue volatility. Indirect taxes and (mortgage) interest payments are also generally excluded on the grounds that they are erratic and endogenous to monetary policy-making. The adoption of such standard exclusions has the advantage that the authorities are less likely to be perceived as manipulating the target.

Exclusion-based measures benefit from the fact that when more than one is used, say, excluding food and excluding F&E, the difference between the measures provides analytical insights into the inflationary process with regard to identifying the effect on inflation of the excluded sector(s), energy in this case. Volatile products are usually taken to be F&E, but studies have found that the inclusion of certain food items, such as fresh fruit and vegetables, and energy items, such as gasoline, makes the CPI *more* volatile (Cecchetti, 1997). In some countries, all food is excluded when only some

(seasonal) components of food are more volatile (Cutler, 2001). If the exclusion is on the grounds of volatility, empirical work should be undertaken to ensure that such volatility exists, has had some longevity, and preferably has an economic rationale for its continued existence.

The criteria for exclusion should be transparent. For Canada, for example, the eight most volatile components were selected on the basis of historical data and, along with indirect taxes, excluded (Macklem, 2001). These excluded items were fruit, vegetables, gasoline, fuel oil, natural gas, intercity transportation, tobacco, and mortgage interest costs. Their price changes were found to be more than one-and-a-half standard deviations from the mean in at least 25 percent of the 12-month comparisons over a 15-year period. Kearns (1998, Table B1) lists 105 components of the Australian CPI and the number of times out of the past 70 price quarters that a price change fell outside of the left- and right-hand tails of the distribution as defined by one, one-and-a-half, and two standard deviations from the mean. Such analysis provides objective support to the case for excluding particular product groups and preempts any challenge when exclusion is based on the simply stated ground of "excess volatility" (see also Mankikar and Paisley, 2004).

Issues of credibility may play a role in the decision to exclude components. F&E may be very important components of consumer expenditure, and, especially for developing countries, their exclusion may detract from the credibility of the measure among poorer members of the society, as argued for South Africa by Lehohla and Myburgh (2002).

Cutler (2001) notes that in phrasing the problem of separating the signal and the noise, the noise is being defined as sector specific and thus the removal of noise as the removal of sectors. However, as considered by Balke and Wynne (1996), there may well be shocks, owing to unusual weather, oil prices, and exchange rate fluctuations that feed through to a number of sectors to different degrees.

Indirect tax and subsidy changes have a first-round effect in proportion to the degree to which they are immediately passed on. Hogan, Johnson, and Laflèche (2001) show the merits of excluding the effects of both the introduction of value-added tax in Canada in 1991 and the decline of the tobacco tax in 1994. Cutler (2001) shows the importance of removing the effect of a sharp change in local authority taxes in the United Kingdom. *Ad hoc* adjustments to remove the effects of severe indirect tax (and subsidy) changes may, however, diminish the credibility of the index. An alternative and more acceptable procedure is to exclude all indirect taxes and subsidies. However, the resulting "indirect tax/excluded index" is something of a misnomer because it relies on the unlikely assumption that tax changes are passed through immediately and on a one-for-one basis to consumer prices. There may be second-round effects in which they feed through to wages and the prices of other goods and services.

Interest rates are included in the CPIs of some countries as (real) mortgage interest rates with regard to owner-occupied housing or part of

Financial Intermediation Services Indirectly Measured. With their inclusion, an increase in interest rates, for example, designed to lower the rate of inflation, may contribute to inflation. In countries where there are doubts about the credibility of the CPI and concerns over statistical "interference" with the target measure, there may be a case for including interest rate payments in the core inflation measures, but generally they should be excluded (see Rowlatt, 2001, for the United Kingdom).

Other major one-off or erratic shocks may be excluded. These should be known to have a temporary effect on prices, say, because of abandoning tariff barriers or changes in the terms of trade (Roger, 1998) and subsidies (García, 2002). Their *ad hoc* exclusion depends on the perceived confidence the public has in the monetary authorities.

Domestically generated inflation measures are useful when shocks arise from erratic movements in the exchange rate. If traded goods are stripped from the index, the resulting domestically generated inflation may erase the first-round effect of the temporary shocks. Mankikar and Paisley (2004) draw attention to three such measures used by the Bank of England: the GDP deflator excluding export prices, the Retail Prices Index excluding mortgage interest payments (RPIX), and a measure based on unit labor costs. Their trends are found to be quite different. Indeed, the method relies to a large extent on how good the data are to allow such prices to be excluded. Such measures are also useful because they focus on what can be argued to be economic activity that can be more readily managed, being domestically generated.

Imputation-based methods are argued on the basis that exclusion is the same as zero weighting, which is effectively equivalent to allocating mean price changes (with exclusions) to the weights of the excluded products (Roger, 2000). If the purpose of the core inflation measure is to reduce the level of noise that contaminates the signal, it must be recognized that this is at the cost of the loss of information. The volatile product groups contain their own noise and signal and in excluding them some of the latter is lost. It is reasonable to ask whether the weights of excluded items might be better apportioned to product groups likely to experience similar "uncontaminated" price changes, that is, to attempt to recover some of the signal in the excluded product groups rather than implicitly assume that it is the same as the overall mean change.

Trend Estimates

The use for core inflation of trend estimates, T_t, for each period t, from a (say, multiplicative) decomposition of a CPI series, Y_t, has an intuition with regard to its ability to smooth a series. The very essence of the trend component, T_t, is that it abstracts from Y_t the seasonal, S_t, and irregular, I_t, components in the model: $Y_t = T_t \times S_t \times I_t$. The first step in such a decomposition is in fact the estimation of T_t. Moving averages are mainly used to estimate T_t in such decompositions, but reliable estimates are not

available for the more recent time periods, which are of critical importance to inflation targeting. For example, for a 36-month centered moving average (CMA) trend, estimates for the first and last 18 months cannot be computed and would have to be extrapolated based on some model. This is a serious deficiency of the method because it does not provide real-time estimates of core inflation. Diewert (1999) proposes the use of a rolling 12-month average index that, when centered, reduces the delay to 6 months. Regression-based estimates of the trend can, however, be in real time, but they rely on restrictions of parameter stability and functional form. The approach in Section III of this paper is to use the smoothed trend estimates as a retrospective reference series to choose between competing core inflation measures on the basis of their deviations from it.

Limited Influence Estimators

The weighted median requires that price changes of expenditure group i, $\dot{p}_i^t$, are first ranked, as are their associated expenditure weights, w_i^b, and related cumulative normalized, relative expenditure weights, Cw_i^b. The (weighted) median is the value of the middle $\dot{p}_i^t$ such that half of the index's weight is above and half below its value. The median is the middle price change when price changes are ranked according to expenditure share in order of magnitude. It uses all the information in the data set in determining the middle observation, but is unaffected by extreme values at either end of the distribution. The computation of the median is timely and transparent with regard to its replication, and both easy to compile and explain.

Trimmed symmetric means remove specified upper and lower tails of the distribution of $\dot{p}_i^t$. For example, a 20 percent trimmed mean first excludes 10 percent of the weight at the top and bottom of the $\dot{p}_i^t$ ranking. The remaining weights are normalized, and the weighted mean forms the measure. A weighted median is an extreme form of a trimmed symmetric mean suffering from the loss of much of the sector-specific signal information, though gaining from being highly robust to shocks in many product groups. The coverage of trimmed means varies from month to month. Trimming can be used to identify from past data which product groups are most often excluded. A constant product group exclusion-based index can then be compiled.

Trimmed mean estimators are timely, transparent with regard to their replication, and easy to compile. Trimmed mean estimators can be usefully calculated at different levels of trim, there being a trade-off between the ability of the measure to exclude extreme values—the median being most effective in this respect—and the loss of information.

Trimmed means are not without problems. First, the nature of shock or reason behind the extreme values needs to be identified. For example, as Mankikar and Paisley (2004) note, the supply shock outbreak of foot-and-mouth disease in 2001 in the United Kingdom led to large price increases for beef that would be helpfully trimmed out, but the subsequent smaller

readjustments over several months would not be trimmed out. An economy experiencing a series of positive shocks in different sectors, with slow rates of readjustment, would give rise to trimmed core inflation movements that understated inflation. Central banks do not want to react to price fluctuations that will reverse themselves quickly. This implies that they should remove both the first- *and* second-round effects. Further, some product groups have persistent, quite disparate price trends. For example, the prices of electronic goods may be falling rapidly when overall inflation is increasing. Trimming will remove such price trends wrongly identifying them as persistent shocks, and overall trend inflation will be overstated.

Second, it is not just the level of trim that causes trimmed means to differ. Trimmed means will vary according to the level of disaggregation at which the trimming takes place. More disaggregated levels are generally preferred subject to adequate sample sizes.

Third, a problem with trimmed means is that they have been found to be systematically lower than CPI means, implying that they are doing more than trimming random shocks (Kearns, 1998; and Cutler, 2001). Roger (2000) points out that these lower values are in line with the empirical finding of positive skewness of price changes and suggests trimming more of the right-hand tail than that of the left. The skewness of country price changes should be explored before using trimmed means.

Non-normality: Asymmetric and variable trimmed measures are used when the price change distribution is found to be non-normal. There is an extensive literature of findings of non-normal distributions of price changes—leptokurtic and skewed to the right—referenced in Roger (2000, pp. 5–6) and including Bryan and Cecchetti (1996); Silver and Ioannidis (1996); Roger (1997 and 2000); Kearns (1998); Bakhshi and Yates (1999); Cutler (2001); and Heath, Roberts, and Bulman (2004).

There are also reasons to expect a non-normal distribution of price changes. First is the inability of fixed weight CPIs to properly reflect the random nature of the shocks characterized in Equation (2). Second, such a characterization requires flexible price setting. Roger (2000) argues that asymmetries can be expected, especially in developing and transition countries, where there might be more administered prices, trade restrictiveness (which diminishes the elasticity of supply), deregulation and privatization, and productivity differentials between industries (Balke and Wynne, 1996). There is also an argument for price stickiness from search cost theory, which predicts an equilibrium outcome of price dispersion (Burdett and Judd, 1983). Products differentiated by brand and features, irregularly purchased products, and product pricing under high inflation may all be subject to high search costs and inflexible pricing (Sorensen, 2000; and Lach, 2002).

The non-normality of price change distributions has implications for core inflation measurement. The sample mean is the most efficient estimator of the population mean if the distribution is normal, but for even a small departure from normality the sample median is a much more efficient estimator of both

the population mean and median. Roger (2000, pp. 34–5) demonstrates this relative efficiency of the median over the mean when the distribution is leptokurtic and, more specifically, fat tailed. The median is an extreme form of a trimmed mean and it follows that trimmed means are also more efficient estimators than means for non-normal distributions.

For price change distributions skewed to the right, symmetric trimmed means and medians will be biased against a target symmetric arithmetic mean inflation rate. Since the cause of such price changes is considered to be administrative "pent-up" price changes, as opposed to shock, our concern here is to preserve such information because it is part of measured inflation and not to have the trimming throw it out as if it were (normally distributed) noise. This requires that the trimming be less harsh on the right-hand side. One way to correct for this bias, following Roger (1997 and 2000), is to not center the trim on the 50th percentile, but to center it on the *mean* percentile. First, the price changes are ordered (taking account of their weights) into percentiles. Second, the weighted arithmetic mean of all the data is calculated. The mean percentile is the value of the price change of the percentile class in which the mean falls. For a normal distribution, the mean is the value of the observation that corresponds to the 50th percentile. For a positively skewed distribution, the mean will be pulled upward.

Roger (1997) finds for New Zealand data that the 57th percentile is a more appropriate *mean percentile*. Thus, the median (50th percentile) would understate trend inflation, whereas the 57th percentile would correct for the bias and retain the advantages of the median as a measure of core inflation. The appropriate mean percentile should be derived on the basis of country data; Kearns (1998) finds the most appropriate percentile to be the 51st and Mrázek (2006) the 67th for Australian and Czech data.

We have noted that if there is skewness, symmetric trimmed means and medians will be biased against a target symmetric arithmetic mean inflation rate. A further approach is to continue with the biased measures, but to then correct them using a simple rescaling method (Maria, 2004).

An alternative stance is to hold that the non-normality in the price change distribution is the result of noise. The level of trim and central percentile can then be selected as one that best removes the joint effects of skewness and excess kurtosis, say, by recourse to the Jarque-Bera test statistic (see Aucremanne, 2000; and Heath, Roberts, and Bulman, 2004). The least trimming percentage corresponding to each central percentile for which the Jarque-Bera statistic does not reject non-normality is chosen. The central percentile for which the optimal trim rejects the least number of observations is selected as the estimator. Where there are two or more such central percentiles, the one with the lowest Jarque-Bera statistic is selected. The Jarque-Bera $\chi^2(2)$ test is concerned with the rejection of a null hypothesis of normality relating to the (symmetrically combined) differences owing to skewness and kurtosis statistics. However, first, it is not immediately apparent why skewness and kurtosis should be considered equally. Second, the test is one of whether the difference is over and above sampling errors

rather than whether the difference is meaningful. Third, the level of trim from this method will of course vary over time.

Volatility weights, unlike the median and trimmed mean, can include all sectors but give less weight to those most volatile, on the grounds that the concern of a core inflation measure should be to minimize volatility. A volatility weighted index (also referred to as a "neo-Edgeworthian" index) is given by the following equation:

$$\frac{\sum_i^n \frac{1}{vol(\pi_i)} \pi_i^t}{\sum_i^n \frac{1}{vol(\pi_i)}}, \tag{3}$$

where $vol(\pi_i)$ is an indicator of volatility; Diewert (1995) demonstrates that the variance of price relatives is appropriate for this purpose. However, Marques, Neves, and Sarmento (2000) use as volatility weights the standard deviations of the *deviations* in inflation rates, that is,

$$vol(\pi_i) = \sigma_i^t = \sqrt{\sum_t \left[(\pi_i^t - \pi^t) - \overline{(\pi_i^t - \pi^t)} \right]^2}, \tag{4}$$

where $\overline{(\pi_i^t - \pi^t)} = \sum (\pi_i^t - \pi^t)/m$ and π_i^t is the 12-month inflation rate for $i = 1, \dots, n$ product categories in $t = 1, \dots, m$ periods and π^t is the mean of π_i^t. The period over which σ_i^t is calculated will affect the result, and Heath, Roberts, and Bulman (2004) consider a rolling 4-year period as well as the full length of the series.

Such indices can also be double-weighted indices in that the volatility weights are applied to the expenditure weights:

$$\sum_i^n \frac{w_i}{\sigma_i^t} \pi_i^t \Big/ \sum_i^n \frac{w_i}{\sigma_i^t}, \tag{5}$$

or can be unweighted with $w_i = 1/n$, as is implicit in Equation (3), though because the target index will be weighted, there is little sense in excluding such information from this measure. A product group with a very small weight, but smooth price changes, should not be allowed to dominate a core inflation series.

Principal components analysis for core inflation measurement was first proposed by Coimbra and Neves (1997), although this approach has been explored and used by the Banco de Portugal as documented in a number of studies including Machado and others (2001) and Maria (2004). So far, the concern has been with regard to smoothing. Yet monetary authorities responsible for targeting inflation also need to make policy on the basis of expectations as to *future* inflation. Their need is to be led by information that is *durable*, that is, useful in medium- and near-term inflation forecasting (Blinder, 1997, p. 157). This is the "outlook" or "persistence" approach to core inflation measurement.

A number of such measures are considered below, in turn. It is, however, worth pointing out that it is well established that forecast performance can be improved by the combinations of forecasts from different models. Jacobson and Karlsson (2004) use a Bayesian model for forecasting inflation, which combines several methods and provides an indicator of the relative effectiveness of each. However, the concern here is with the individual characteristics of each *prediction-based method.*

Reweighting the CPI

Persistence weights give more weight to product groups considered best able to forecast the target. Cutler (2001) uses a first-order autoregressive (AR) model:

$$\pi_i^t = \alpha_i + \rho_i \pi_i^{t-12} + \varepsilon_i^t, \tag{6}$$

where π_i^t are the normalized, and the estimated ρ_i, given by $\hat{\rho}_i$, are taken as an indicator of the persistence of inflation in each product group i. Product groups with negative $\hat{\rho}_i$ are assigned a weight of zero. The persistence-weighted index is given by the following equation:

$$\frac{\sum_{i \in \rho_i > 0} \hat{\rho}_i \pi_i^t}{\sum_{i \in \rho_i > 0} \hat{\rho}_i}. \tag{7}$$

The index may be double weighted as in Equation (5) for similar reasons. The estimated persistence weights may be changed each year using a rolled-forward monthly data set. The need for a lengthy time series for the estimate means that much of the data on which the estimates are based are neither related to the period required for the predictions nor responsive to changes.

Yet persistence weighting was found by Cutler (2001) to yield sensible results in that it excluded volatile product groups such as seasonal food items, but included nonseasonal food items on the grounds that they were found to have information useful for prediction. Cutler found the persistence-weighted index ranked third over a number of time horizons when compared with seven exclusion-based indices in terms of their predictive ability over and above inflation. When the predictive ability of the persistence-weighted index was tested with the addition of further lags to the one in Equation (6), it, along with two exclusion indices, proved to be superior to the trimmed mean and weighted median indices. Yet the measure has some shortcomings.

First, it is relatively complex. Second, there is much in the construction of the index, such as changes in the type of products included and classification changes over the period in which weights are estimated, that may give rise to unstable or biased coefficients. Third, the definition of the dynamics of persistence given by $\hat{\rho}_i$ in Equation (6) is quite restrictive.

Short-term prediction methods are suitable for one- or two-period-ahead predictions. The time period over which it takes monetary policy to take effect is usually considered as the medium term, say, 18 months. Yet there may also be the need for short-term predictions and the Holt-Winters exponential smoothing and the Box Jenkins approach (see Chatfield and Yar, 1991; and Hyndman and others, 2005) are well-established methods in this regard.

Economic Models

Estimated economic models benefit from being grounded in theory. They incorporate further economic variables, thus realizing a core inflation measure from a multivariate setting, abstracting from (or conditioning on) the effects of these other variables. They also make explicit the economic drivers of core inflation so that the underlying process can be better understood, as can departures from it. However, estimates of core inflation will vary according to the economic model of the determinants of core inflation and econometric issues regarding specification, data, and estimation, which in part require judgment as well as purely statistical considerations.

Following Roger (1998), a simple model owing to Eckstein (1981)—although see Parkin (1984)—provides a short-term aggregate supply curve given by

$$\pi_t = \pi_{t+1}^{LR} + g(x_{t+1}) + \varepsilon_t, \tag{8}$$

where π_{t+1}^{LR} is the long-term trend inflation rate, $g(x_{t+1})$ is a measure of cyclical excess demand pressure, and ε_t are the transient disturbances. Inflation can be simply decomposed into core, long-term trend, cyclical, and residual components. Core inflation, π_t^*, is then

$$\pi_t^* = \pi_t^{LR} = \pi_t - g(x_{t-1}) - \varepsilon_t. \tag{9}$$

Note that under this concept of core inflation, cyclical fluctuations are removed.

Quah and Vahey (1995), on the other hand, distinguish between two types of shocks: those that can influence core inflation and those that have a medium- to long-term effect on real *output*. They use a structural vector autoregressive (SVAR) model for the United Kingdom. The view is that disturbances are benign to output because an economy will adjust to their effects in the medium or long run. However, if there are rigidities, then core inflationary shocks will affect real output in the short run, but not in the medium to long run. The identification restriction in the SVAR estimator allows the data to determine whether the economy quickly adjusts to these core inflationary disturbances. Different identification schemes from those suggested by Quah and Vahey have been used. Folkertsma and Hubrich (2000) compare five schemes using European data finding substantial differences in the range of measurement error, but also express, for policy

purposes, concern about the extent of the errors. Wynne (1999) criticizes the approach on the grounds that each time the index is reestimated, it has to be revised and that such measures are difficult to communicate to the public.

The information set used in economic models need not be confined to the CPI. Bagliano, Golinelli, and Morana (2002), in a study of the euro area (1979–2000), use series on inflation, money, output, and interest rates to estimate a forward-looking measure of core inflation based on the long-term (co-integrating) relationships between these variables. Mankiw and Reis (2003) developed a framework for, and provide estimates for the United States (1957 to 2001) of, a stability price index. The weights for the index are derived as econometric estimates from a model that allows sectoral prices to vary according to (1) their expenditure share, as is appropriate for a CPI, (2) their sensitivity to business cycles, (3) their likelihood of experiencing idiosyncratic shocks, and (4) the flexibility of prices to respond to economic conditions. The estimated weights are those that minimize the volatility in the output gap—the variance of deviations of output from its natural level. The weights used will provide an index that, if kept on target, will lead to the greatest stability in economic activity. Included in the empirical work for the United States is the level of nominal wages, a series that has of course zero weight in the CPI, but is more cyclically sensitive than most other prices, and has a large weight in the stability price index. If the aim is to target a measure that aims at economic stability, the stability price index is shown to be preferable to a CPI.

III. How to Choose Among Methods: Judging Which Is Best

Having outlined a number of measures and their variants and in doing so having said something about their properties, relative merits, and how they might best be implemented, we now turn to the choice among these measures.

With the variety of core inflation methods and their alternative formulations, it is necessary to establish criteria by which countries can choose among measures. A number of empirical country studies have been undertaken involving the use of often different combinations of measures and appraised according to often different criteria. A quite apparent conclusion is that no consensus emerges from the studies. Moreover, even within a country, different criteria suggest different methods and even then, when the same criteria are used for a country, the optimal method chosen often changes over time. Given this lack of consensus, it is proposed that the choice of method should in part be data driven—tailor-made to the empirical realities and needs of the countries. The approach is that each country should examine its own data according to criteria useful to it. How to judge the best method is the subject of this section.

It is recognized, however, that this data-driven approach may not always be practical. The measures outlined above rely on CPI data, and in some cases such data may be deemed to be unreliable by the monetary authorities. The offense of an unreliable CPI will be compounded by then deriving core

inflation measures based on bad CPI data. The concern in this case should be to first improve the CPI.

Credibility and General Considerations

The relative importance of these considerations to a large extent depends on the purpose or needs of the central bank's decision-making process. It is most important, for example, that a core inflation measure used as an inflation target be credible and timely, whereas more attention may be given to the predictive power of a core inflation measure used to operationalize an inflation targeting framework.

Roger (1998) argues that a core inflation measure should be

- timely,
- credible (verifiable by agents independent of the central bank),
- easily understood by the public, and
- not significantly biased with respect to the targeted measure, which would again harm its credibility.

In this regard, the selection of method may in part be based on the perceived confidence in the index. For example, stripping out F&E as volatile may be perceived as fixing the index. South Africa, in choosing between measures for inflation targeting, emphasized the inclusion of items to which poorer households are most sensitive, along with an increased rural coverage (Iahohla and Myburgh, 2002). Transparency in methodology is also an important ingredient in credibility, and the IMF's Dissemination Standards Bulletin Board, which includes the CPI, is important in this respect (San José, Slack, and Sriram, 2002).

The agency responsible for constructing the CPI is also important to its credibility as a target, and Schaechter, Stone, and Zelmer (2000) note the need for the CPI and core inflation to be compiled by an independent agency, typically the country's statistical agency.

Commenting on Roger's considerations, Wynne (1999) notes that a measure of core inflation should also

- be computable in real time;
- be forward looking in some sense;
- be robust and unbiased;
- have a track record of some sort;
- have some theoretical basis, ideally in monetary theory;
- be familiar and understandable to the public; and
- not be subject to revisions.

These are, of course, not absolute criteria, and different policymakers will apply different priorities to each. There may well be, for example, a trade-off between simplicity and bias. As noted in Section II, it is worth differentiating between two broad purposes: measures used for defining an inflation target for policy assessment and measures to help predict and set policy to achieve an

objective. In the first case, credibility, understandability, familiarity, transparency, the ability to compute in real time, and nonrevisability should be heavily prioritized. Wynne (1999) emphasizes that these features are only important to the extent that the central bank seeks to use a measure of underlying inflation as an important part of its routine communications with the public to explain policy decisions. In the second case, it would be important for the series to be forward looking and/or provide analytical insights.

Against all of this, there are statistical criteria relating to how effective a measure is in terms of properties, such as smoothing or prediction, that may beneficially relate to their use. The satisfaction of appropriate statistical criteria can help ground the measures to the extent that objective criteria are used in the selection.

Judging on the Basis of Control

Blinder (1997, p. 160) argues for the automatic exclusion of F&E:

> It all depends on whether recent values of food and energy inflation help forecast future core inflation. As a central banker, I always preferred to view the inflation rate with its food and energy component removed as our basic goal. But not because these components are extremely volatile. The real reason was that the prices of food (really, food at home) and energy are, for the most part, beyond the control of the central bank. The Fed cannot do much about food and energy prices—except, of course, to cause a recession deep enough to ensure that increases in these prices do not lead to overall inflation. But the central bank *can* do something about the rest of the price index—the part that comes out of the industrial *core* of the economy, so to speak [author's emphasis].

Porrado and Velasco (2002) take a similar stance with regard to the CPI not being an appropriate measure because it includes nondomestically produced goods and services. They argue that, because the CPI is affected by exchange rate variations that the central bank has little control over, responding to all CPI fluctuations is an overreaction that destabilizes output. Mankikar and Paisley (2004) discuss the domestically generated core inflation measures used by the Bank of England in relation to this (see also Section II).

Care has to be taken with such stances. On the one hand, they rightly point to a control problem whereby the central bank is attempting to control components of inflation over which it has little control. The argument is to exclude them. On the other hand, a purpose of the inflation targeting framework is to anchor inflation expectations, and such expectations apply to a wider range of components that a central bank will have control over (Hill, 2004).

Judging on the Basis of Deviations from a Reference Series

Countries can evaluate alternative methods, using past data in terms of the deviations of the results from the methods, from a reference, long-term trend

measure of inflation. The criterion used here is not one of predictive ability but policy assessment in terms of adherence to a smoothed reference series. Thus, if the reference index is taken to be a desired (retrospective) measure of core inflation, π_t^*, and the measure of core inflation being assessed is π_t', then the best measure might be one that minimizes its root-mean-squared error (RMSE):

$$\sqrt{\frac{\sum_{t=1}^{T} (\pi_t' - \pi_t^*)^2}{T}}, \tag{10}$$

or its mean absolute deviation (MAD):

$$\frac{\sum_{t=1}^{T} |\pi_t' - \pi_t^*|}{T}. \tag{11}$$

Note the interplay between the outcome and the basis in Equations (10) and (11) used for selecting the measure. If Equation (10) is the criterion for choice, the mean will be a desirable measure of central tendency if there are outliers. The intuition is that the least-squares estimator, in the process of squaring the deviations from the mean, puts a high premium on extreme deviations. As such, a desirable least-squares estimator has to be quite close to the extreme value if it is to satisfy this criterion. Now consider an estimator that seeks to minimize the sum of the *absolute* differences between it and the observations. There is no longer a premium to be put on the measure of central tendency being close to the outlier, and the median can be shown to be the best estimator by this criterion. The mean is only robust to outliers if the distribution is normal, where similar extreme values appear in both tails of the distribution. Because a core inflation measure must function in a way that strips away noise, then it is necessary to identify whether skewness is the result of actual information, as in administered pent-up prices, or is to be characterized as noise, as in Equation (2); such prior decisions will affect the choice of criteria and measure selected. Further, we have to ask whether getting it occasionally badly wrong in a month is something that they must avoid, in which case they should use the RMSE, rather than the MAD, to judge which measures to use. At the very least, Equations (10) and (11) should both be used to see if such considerations matter.

Bryan and Cecchetti (1994 and 1996); Bryan, Cecchetti, and Wiggins (1997); and Cecchetti (1997) use a 36-month CMA of actual inflation as the reference series and the RMSE as a summary measure of the overall deviations, although they comment that, in general, other summary measures lead to similar conclusions. However, Bakhshi and Yates (1999) find that the MAD can provide different results from the RMSE.

The method relies on the suitability of the reference index, the (centered) moving average. It suffers from the inability to provide estimates at the start and end of the series; for example, for a 36-month moving average, trend estimates for the first and last 18 months are not provided. This is a serious deficiency because it does not allow real-time estimates of core inflation,

although Diewert (1999) advocates a moving year centered index average that would lose 6 months' data on either side. However, the CMA is used here as a reference measure—as a means to consider the retrospective performance of difference measures. To its credit, the CMA uses all the CPI data, unlike limited-information estimators. The CMA approach is easy to calculate and has an intuitive justification and plausibility, especially when calculated and graphed over noisy series. However, there have been a number of criticisms of this use of smoothing to generate a reference series.

First, the method treats the first and last 18 months as equally important (Blinder, 1997). If we accept that we are interested in removing noise to better the forecast future inflation rates, then an appropriate procedure is to give more weight to components that better forecast the future.

Second, as Mankikar and Paisley (2004) point out, there is no economic rationale for smoothness to be desirable. Economies can go through periods of sharp fluctuations in demand and supply that have longer-term effects and, as discussed in Balke and Wynne (1996), these may result in *varying* skewness of the distribution over time leading to fluctuations in core inflation (see also Marques, Neves, and Sarmento, 2000).

Third, Bakhshi and Yates (1999) find that the results can depend on the number of periods used in the averaging. Aucremanne (2000); Marques, Neves, and Sarmento (2000); and Heath, Roberts, and Bulman (2004) find that the optimal trim chosen can be very sensitive to the smoothness of the benchmark series chosen as well as the sample periods used in the calculation of the RMSE and MAD statistics. However, if we accept a 36-period CMA as a long-term benchmark, and if alternative moving averages, say, 24 months, give different results, this may be because they are more effective at reflecting shorter-term trend movements, albeit at the cost of some increased sensitivity to noise. Then the monetary authorities must decide on the required balance between smoothness and time horizon.

Justifying the Exclusion of Product Groups on the Basis of Their Volatility

As noted above, exclusion-based measures are often justified on the ground that some product groups are more volatile than others. There may be a priori grounds to expect this, such as for seasonal items and energy. Perceptions of volatility are not sufficient grounds to exclude product groups. There should be objective criteria applied to country data. Product groups found to have experienced relatively high price change volatility, with evidence of some longevity to the volatility—and further, with some reasonable expectation that past patterns will continue into the future—have grounds for their exclusion. It is then appropriate to ask whether the source of the volatility is random shocks or price rigidities, such as irregular administered price changes. The former should be excluded. The latter should be included as part of appropriately measured price changes, otherwise the core measure is likely to understate target inflation. A separate index may, of course, be compiled for analytical purposes that

exclude such changes to allow for the effects of the inclusion to be identified. It is reasonable for exclusions to be data driven, and there are a number of ways of doing this.

One approach is to use the standard deviation of the 12-month rate of inflation for the CPI and then for the CPI excluding F&E. Cecchetti (1997), using U.S. CPI-U data, finds the standard deviation to be 2.33 percent (CPI-U) and 2.58 percent (CPI-U excluding F&E), respectively—F&E were *less* volatile than other components. On the basis of these measures, F&E should not be excluded from a U.S. CPI-U core inflation index.

Another approach is to consider volatility to be deviations from a long-term reference series—a 36-month moving average—as discussed above using Equations (10) and (11). Cecchetti (1997) found the MSE (without the Root) for 12-month (and other) price indices to be *higher* if F&E were included, rather than excluded: 1.29 compared with 1.01, respectively. F&E are *more* volatile by this measure than other components. Cutler (2001), for the United Kingdom, finds on this basis that a series excluding F&E is a little smoother than a series that includes it.

In all of this, it is necessary to consider the appropriate level of disaggregation. Food, for example, if tested for volatility may indeed be less volatile than other groups, but this may be so only because the food products with less volatile price changes have been grouped at too high a level of aggregation with products with nonvolatile price changes. Of course, the level of detail will depend on the availability and reliability of data and the importance of the product to household expenditure.

Blinder (1997, p. 160) is emphatic that if F&E are persistent in the sense that recent values of F&E inflation help forecast future core inflation, then we would "…surely *not* want to take them out of the index. We clearly would want to leave them in" [author's emphasis]. This takes us to judging methods on the basis of predictive ability. Note that the predictive models given in the next section are concerned with predictive models of overall core inflation. Such models may also be applied to data for different product groups and, following Blinder (1997), be used to ascertain which product group contains information to better predict its own, or aggregate, future values to be included or excluded, for an exclusion-based index. However, the context of the next section is the predictive ability of the overall CPI.

Judging on the Basis of Predictive Ability

First is a simple model to estimate regression models of CPI inflation on different core inflation estimates, that is,

$$\hat{\pi}_t = \alpha + \beta \pi^*_{t-12}, \tag{12}$$

and to calculate and compare RMSEs. Bryan and Cecchetti (1993), using U.S. CPI-U 12-month inflation data for February 1967 to December 1979, calculate RMSEs for forecasts beginning in January 1980 over future annual periods of between 1 and 5 years. In all cases, the weighted median has a

lower RMSE than the 15 percent trimmed mean, and the trimmed mean a lower error than the CPI-U. For example, for forecasting over a time horizon of 24 months, the RMSE for the CPI-U to predict itself is 63 percent of the mean price change in that forecast period, whereas that of the median is 57 percent and the trimmed mean 61 percent. The core measures are improvements on the CPI for predicting long-horizon CPI inflation, but not substantially so, especially given the error margins involved.

Second, Laflèche (1997) argues that prediction from an AR model should be used to judge which series are selected, that is,

$$\hat{\pi}_t = \alpha_i + \sum_{j=1}^{n} \beta_j \pi_{t-j}^*, \tag{13}$$

with the core inflation measures with the highest $\bar{R}^2$ selected. Marques, Neves, and Sarmento (2000) note that $\bar{R}^2$ is a measure of their relative standing as opposed to their absolute standing and that the latter is what we are interested in.

For a wider range of acceptable criteria for evaluating forecasts, see Diebold and Mariano (1995). In favor of considering further criteria is the fact that if the same "best" measures are found to arise from a number of criteria, then they give more confidence in their use; and if not, there are pertinent questions to be asked as to why they differ, and insights to be gained.

Third, Jacobson and Karlsson (2004) propose the use of a Bayesian average of forecasts of inflation and, although the work is based on 86 general economic indicators, the principles apply to using just core inflation measures. They find that combining forecasts from the 10 highest ranked indicators produced forecasts with smaller RMSEs than using the forecasts for the individual series.

As will be discussed below, corroboration of the core measures need not just be by a ranking of such measures by their predictive power. The first hurdle should be a (Wald) statistical test to identify, for Equation (13), whether the null hypothesis that the coefficients in β_j are jointly zero can be rejected.

Fourth, an alternative formulation of Equation (13) is the predictive ability of the core inflation measure *over and above* that of the current CPI, using

$$\pi_t = \alpha_0 + \sum_{j=1}^{n} \alpha_{1j} \pi_{t-j} + \sum_{j=1}^{n} \beta_{1j} \pi_{t-j}^* + \varepsilon_{1t}. \tag{14}$$

Positive values of $\sum_{j=1}^{n} \beta_{1j}$, for which the Wald test is statistically significant, demonstrate contributions in predictive power for lagged core inflation over and above that of lagged inflation (Cutler, 2001).

Fifth, as will be discussed below, there is the issue of causality. It is, of course, possible to predict the CPI from core inflation and to predict core inflation from the CPI. We want the prediction of the CPI from core inflation

to dominate. Marques, Neves, and Sarmento (2003) for U.S. data find, in testing Granger-causality, that a series that excludes F&E is a leading indicator of inflation as opposed to a desired lagging one. They note that this is not surprising given that energy and unprocessed food are also intermediate inputs into goods and services, and they are thus likely to affect final prices in a future period.

Sixth, a statistical evaluation of the relationship between a core inflation and headline inflation measure of the types outlined above, and the tests to be outlined in the next section, cannot properly account for the policy responses of the authorities to the core measure. The core measure may appear, for example, to be a bad predictor of inflation only because the authorities respond to, say, increases, by adopting policies to reduce it.

The use of such tests as part of a strategy to evaluate the predictive ability of alternative core inflation measures is one thing. However, a number of writers have used a battery of such tests as the sole basis for choice. Thus, only measures of core inflation that pass all the tests are considered suitable. There is no distinction as to the extent of the volatility or predictive power of the alternative measures. Furthermore, the nature of the tests has implicit in them particular concepts of what core inflation should measure. For example, Marques, Neves, and Sarmento (2003, p. 768) argue that core inflation π_t^* should *not* be evaluated on the grounds that it is a good predictor of inflation, π_t. They note,

> By definition, a good predictor of future inflation must be able to account for short-term movements on the price level, but this is exactly what we cannot or should not expect from a core inflation indicator, as it is just a summary measure of the long run characteristics of inflation.

We turn to the tests they propose.

Judging on the Basis of Tests

There are a number of formal statistical tests that their proponents argue a good core inflation index should satisfy. Heath, Roberts, and Bulman (2004) consider these following two tests, of unbiasedness and causality, of a desirable measure of π_t^* as essential.

Unbiasedness

Consider the following:

$$\pi_t = \pi_t^* + v_t, \tag{15}$$

where π_t is CPI inflation, π_t^* is trend inflation, and v_t is a temporary disturbance in period t. First, core inflation should be unbiased with respect to π_t and jointly $\beta_0 = 0$ and $\beta_1 = 1$ in the estimated equation

$$\pi_t = \beta_0 + \beta_1 \pi_t^* + v_t. \tag{16}$$

Granger-Causality

Second, core inflation, π_t^*, should Granger-cause the CPI, π_t; past values of π_t^* contain useful information to predict π_t beyond past values of π_t; and further, changes in the core precede temporarily changes in π_t, and the converse is not true. The required tests for stationary series would be to reject a joint null hypothesis that $\beta_{1j} = 0$ for all j in Equation (17a), not reject a joint null hypothesis that $\alpha_{2j} = 0$ for all j in Equation (17b).

$$\pi_t = \alpha_{10} + \sum_{j=1}^{n} \alpha_{1j} \pi_{t-j} + \sum_{j=1}^{n} \beta_{1j} \pi_{t-j}^* + \varepsilon_{1t}; \qquad (17a)$$

$$\pi_t^* = \alpha_{20} + \sum_{j=1}^{n} \alpha_{2j} \pi_{t-j} + \sum_{j=1}^{n} \beta_{2j} \pi_{t-j}^* + \varepsilon_{2t}. \qquad (17b)$$

The terms of the n lags can be determined by statistical criteria such as the Schwarz-Bayesian criterion. In Equation (17a), π_t^* is said to behave as an attractor of π_t in the sense that π_t converges to π_t^* in the long term. We learn about the future path of π_t by looking at π_t^*. However, π_t should not be an attractor of π_t^* as in Equation (17b).

An important finding on the choice of method on this basis is that the conclusion may vary with the time period chosen. For example, Heath, Roberts, and Bulman (2004) applied the Granger-causality tests to 102 measures of core inflation for Australia in the period 1987:Q1 to 2003:Q4, finding that 89 of the 102 measures considered passed the test, though for the subperiod 1993:Q1 to 2003:Q4 *none* of the measures passed the tests. The authors note that this period followed the implementation of an inflation-targeting regime (July 2000), and during it inflation had been comparatively stable.

Co-Integration-Based Tests

Co-integrating regressions are widely used in economics to capture long-term equilibrium relationships. There will, of course, be short-term dynamics in a relationship, but these can be captured by an error-correction model. An appropriate measure of core inflation is arguably one that does not have long-term divergences from the CPI. Marques, Neves, and Sarmento (2000 and 2003) and Mankikar and Paisley (2004) use co-integration tests to choose the most appropriate measure using data for Portugal, the United States, and the United Kingdom, respectively; the findings vary across the countries studied. The test approach requires some comment.

First are the causality conditions. Mankikar and Paisley (2004) note that if the monetary authorities adopt policies to target CPI inflation successfully, aside from unforeseeable noise, π_t will behave as an attractor of π_t^* in the sense that π_t^* converges to π_t in the long run. If core inflation is, for example, above the target CPI in one period, mechanisms are put into play to bring

core inflation down toward the target in the next. This is quite different from the conceptualization of the tests in Equations (17a) and (17b).

Second, the information derived in determining co-integrating relationships should, at first sight, also prove useful for predictions rather than just for tests. It might be expected that because error correction models incorporate both short-term fluctuations and deviations from the co-integrated, long-term equilibrium path in the forecasting model, they should prove more accurate than, say, Box-Jenkins autoregressive integrated moving average (ARIMA) models. This gives a sense of validity to the use of such tests. However, Christoffersen and Diebold (1997) challenge the belief that the imposition of co-integrating restrictions produces superior long-horizon forecasts. They show that for the long horizon, using MSE criterion, even *univariate* ARIMA forecasts are equally accurate. They argue that, in part, the problem lies with the use of the MSE to evaluate forecasts. It is the imposition of integration conditions that is helpful, something that ARIMA models achieve through differencing.

Third, Heath, Roberts, and Bulman (2004) point to a problem regarding the periods in which measures are tested against each other. The authors find that in low-inflation periods, the choice of method is less important.

Fourth, it must be borne in mind that what we have here are statistical tests of desirable conditions. The tests are conditioned on the size of the data and their properties and the fit of the models. They are not measures of how close convergence is over specific time horizons or the error margins expected from such an exercise. They are conditions that core inflation measures should satisfy and they are satisfactory in this sense. The tests have in mind particular concepts of inflation. Equation (15) and the test based on Equation (16), for example, implicitly identify the error structure to be normal when empirical studies and theoretical frameworks identified in Section II have argued that part of the signal may be skewed. Furthermore, the concept of equilibrium associated with the co-integration test may be of a time span outside that of use to monetary authorities for policy measures.

Judging on the Basis of Correlation with Money Supply

Bryan and Cecchetti (1993) consider a primary motivation of their study to be finding a measure of core inflation correlated with monetary growth. They test measures of core inflation in terms of the ability of monetary growth to forecast the core inflation measures. Although some of the results are mixed, depending on the time horizons and the measure of money supply used, the weighted median is consistently better correlated than the 15 percent trimmed mean. For measures of money supply M1 (money that can be spent immediately including cash, checking accounts) and M2 (M1 plus most short term savings accounts), the weighted median is better than excluding F&E, though not for the monetary base as a measure of money supply. Granger-causality tests for M1 and M2 both find changes in the money supply to

Granger-cause core inflation as measured by the weighted median and 15 percent trimmed mean.

In models in which inflation is a purely monetary phenomenon and prices are fully flexible, shocks—such as changes in oil prices, tastes, or technology—are instantaneously accommodated and aggregate price inflation is unchanged. But as outlined in Section II, prices are not fully flexible and the CPI is not a satisfactory realization of any such process, owing to its fixed base and inability to properly incorporate all new products and quality changes. The instability of the relationship between monetary aggregates and inflationary pressure argues against judging methods on the basis of correlations with money supply, as does a possible need to model the endogeneity of monetary growth to shocks (Bryan and Cecchetti, 1994).

IV. Concluding Remarks

This paper has critically outlined the many measures of core inflation and, moreover, the many approaches to judging the preferred measure(s). Choice among measures is not a straightforward matter. Empirical research has shown that different measures of core inflation yield different results, that is, that choice of measure matters. Further, different approaches to the choice of measure yield different results, and even for the same approach to choice, the preferred measure may differ across countries and even within a country for different time periods. Choice of measure should thus, in principle, be data driven for each country.

What is apparent, however, is that a consensus has emerged and, for reasons of maintaining credibility, there is a natural starting point for many countries. First is the use of the CPI as the basis for the core inflation measure as the most visible and credible measure to anchor inflation expectations. Second is the widespread adoption of exclusion-based CPIs. There is some commonality in the product groups excluded, and such exclusions can thus be justified as not manipulating the figures. There may, however, be great public sensitivity to the exclusion of items such as F&E. With greater confidence in the ability of the authorities to manipulate the measure without losing credibility, however, the exclusions can be data driven. In all of this, the credibility of the institution producing the CPI and that of the institution issuing the core CPI will be a consideration.

There is a sense in which the above account is much more comforting than the plethora of measures and adoption criteria, with their particular pros and cons, discussed in Sections II and III. Yet these sections come into their own in two respects. One is where the CPI or an exclusion-based CPI is adopted as a credible basis on which to anchor inflation, and it is necessary for the central bank to have further measures to operationalize the targeting framework.

Another is that exclusion-based methods may be found to be not optimal according to the criteria selected by the monetary authorities, and that the credibility trade-off may have little resonance for the circumstances of the

country. In such cases, the monetary authorities are well placed to select from among the methods and the criteria for choice considered in this paper to adapt their core inflation measure to their needs.

All of this should be data driven so that the methods adopted are tailored to the features of the evolution of that country's economy and the choice of measures can be justified on an objective, transparent basis. Research is required to establish an appropriate target measure and operationalizing measures in tandem with the targeting framework to provide a suite of measures to enable the targeting process based on sound statistical criteria.

REFERENCES

Aucremanne, L., 2000, "The Use of Robust Estimators as Measures of Core Inflation," Working Paper No. 2 (Brussels, National Bank of Belgium).

Bagliano, Fabio C., Roberto Golinelli, and Claudio Morana, 2002, "Core Inflation in the Euro Area," *Applied Economic Letters*, Vol. 9 (May 15), pp. 353–357.

Bakhshi, H., and A Yates, 1999, "To Trim or Not to Trim? An Application of a Trimmed Mean Inflation Estimator to the United Kingdom," Working Paper No. 97 (London, Bank of England).

Balke, N.S., and M.A Wynne, 1996, "Supply Shocks and the Distribution of Price Changes," *Federal Reserve Bank of Dallas Economic Review* (First Quarter), pp. 10–18.

Blinder, Alan S., 1997, "Commentary for 'Measuring Short-Run Inflation for Central Bankers'," *Federal Reserve Bank of St. Louis Review*, Vol. 79 (May/June), pp. 157–162.

Bloem, Adriaan, Paul A. Armknecht, and Kimberly D. Zieschang, 2002, "Price Indices for Inflation Targeting," in *Statistical Implications of Inflation Targeting: Getting the Right Numbers and Getting the Numbers Right*, ed. by C. Carson, C. Enoch, and C. Dziobek (Washington, International Monetary Fund), pp. 172–198.

Bryan, Michael F., and Stephen G. Cecchetti, 1993, "The Consumer Price Index as a Measure of Inflation," *Federal Reserve Bank of Cleveland Economic Review*, Vol. 29 (Quarter 4), pp. 15–24.

———, 1994, "Measuring Core Inflation," in *Monetary Policy*, ed by N. Gregory Mankiw (Chicago, University of Chicago Press for National Bureau of Economic Research), pp. 195–215.

———, 1996, "Inflation and the Distribution of Price Changes," NBER Working Paper No. 5793 (Cambridge, Massachusetts, National Bureau of Economic Research).

———, and Rodney L. Wiggins II, 1997, "Efficient Inflation Estimation," NBER Working Paper No. 6183 (Cambridge, Massachusetts, National Bureau of Economic Research).

Burdett, Kenneth, and Kenneth L. Judd, 1983, "Equilibrium Price Dispersion," *Econometrica*, Vol. 51 (July), pp. 955–970.

Cecchetti, Stephen G., 1997, "Measuring Short-Run Inflation for Central Bankers," *Federal Reserve Bank of St. Louis Review*, Vol. 79 (May/June), pp. 143–156.

Chatfield, C., and M. Yar, 1991, "Prediction Intervals for Multiplicative Holt-Winters," *International Journal of Forecasting*, Vol. 7 (May), pp. 31–37.

Christoffersen, Peter F., and Francis X. Diebold, 1997, "Cointegration and Long-Horizon Forecasting," Working Paper No. 97-14 (Philadelphia, Federal Reserve Bank of Philadelphia).

Coimbra, Carlos, and Pedro Duarte Neves, 1997, "Trend Inflation Indicators," *Economic Bulletin, Banco de Portugal*, Vol. 3, No. 1 (March).

Cutler, Joanne, 2001, "A New Measure of Core Inflation in the U.K.," MPC Unit Discussion Paper No. 3 (London, Bank of England).

Diebold, F.X., and R.S. Mariano, 1995, "Comparing Predictive Accuracy," *Journal of Business and Economic Statistics*, Vol. 13, No. 3, pp. 253–263.

Diewert, W. Erwin, 1995, "On the Stochastic Approach to Index Numbers," Department of Economics Discussion Paper No. 95-31 (Vancouver, University of British Columbia).

———, 1999, "Index Number Approaches to Seasonal Adjustment," *Macroeconomic Dynamics*, Vol. 3 (March), pp. 48–68.

Eckstein, O., 1981, *Core Inflation* (Englewood Cliffs, New Jersey: Prentice-Hall).

Fenwick, David, 2004, "Core Inflation, Seasonal Adjustment and Measures of the Underlying Trend," *Statistical Journal of the United Nations Economic Commission for Europe*, Vol. 21, No. 2, pp. 115–124.

Folkertsma, C.K., and K. Hubrich, 2000, "Performance of Core Inflation Measures," WO Research Memorandum No. 639 (Amsterdam, De Nederlandsche Bank).

García, Pablo, 2002, "Design, Measurement, and Communication: Chile's Experience with Inflation Targeting," in *Statistical Implications of Inflation Targeting: Getting the Right Numbers and Getting the Numbers Right*, ed. by C. Carson, C. Enoch, and C. Dziobek (Washington, International Monetary Fund), pp. 157–171.

Greenlees, John, and Bert Balk, 2004, "Errors and Bias," in *Consumer Price Index Manual: Theory and Practice* (Geneva, International Labour Office), Chapter 11, pp. 207–214.

Heath, Alexandra, Ivan Roberts, and Tim Bulman, 2004, "Inflation in Australia: Measurement and Modeling," in *The Future of Inflation Targeting*, ed. by Christopher Kent and Simon Guttmann (Sydney, Reserve Bank of Australia), pp. 167–207, Available via the Internet: http://www.rba.gov.au/PublicationsAndResearch/Conferences/2004/Heath_Roberts_Bulman.pdf.

Hill, Robert J., 2004, "Inflation Measurement for Central Bankers," in *The Future of Inflation Targeting*, ed. by Christopher Kent and Simon Guttmann (Sydney, Reserve Bank of Australia), pp. 140–160, Available via the Internet: http://www.rba.gov.au/PublicationsAndResearch/Conferences/2004/Hill.pdf.

Hogan, Seamus, Marianne Johnson, and Thérèse Laflèche, 2001, "Core Inflation," Technical Report No. 89 (Ottawa, Bank of Canada).

Hyndman, Rob J., Anne B. Koehler, J. Keith Ord, and Ralph D. Snyder, 2005, "Prediction Intervals for Exponential Smoothing Using Two New Classes of State Space Models," *Journal of Forecasting*, Vol. 24 (January), pp. 17–37.

International Labor Organization, International Monetary Fund, Organization for Economic Cooperation and Development, Statistical Office of the European Communities, United Nations, International Bank for Reconstruction and

Development, and World Bank, 2004, *Consumer Price Index Manual: Theory and Practice* (Geneva, International Labour Office), Available via the Internet: http://www.ilo.org/public/english/bureau/stat/guides/cpi/index.htm.

Jacobson, Tor, and Sune Karlsson, 2004, "Finding Good Predictors for Inflation: A Bayesian Model Averaging Approach," *Journal of Forecasting*, Vol. 23 (November), pp. 479–496.

Kearns, J, 1998, "The Distribution and Measurement of Inflation," Research Discussion Paper No. 9810 (Sydney, Reserve Bank of Australia).

Lach, Saul, 2002, "Existence and Persistence of Price Dispersion: An Empirical Analysis," *Review of Economics and Statistics*, Vol. 84, No. 3, pp. 433–444.

Laflèche, T, 1997, "Statistical Measures of the Trend Rate of Inflation," *Bank of Canada Review* (Autumn), pp. 29–47.

Iahohla, Pali J., and Annette Myburgh, 2002, "Statistical Implications of Inflation Targeting in South Africa," in *Statistical Implications of Inflation Targeting: Getting the Right Numbers and Getting the Numbers Right*, ed. by C. Carson, C. Enoch, and C. Dziobek (Washington: International Monetary Fund), pp. 55–75.

Machado, José Ferreira, Carlos Robalo Marques, Pedro Duarte Neves, and Afonso Gonçalves da Silva, 2001, "Using the First Principal Component as a Core Inflation Indicator," Economic Research Department Working Paper No. 9-01 (Lisbon: Banco de Portugal).

Macklem, Tiff, 2001, "A New Measure of Core Inflation," *Bank of Canada Review* (Autumn), pp. 3–12.

Mankikar, Alan, and Jo Paisley, 2004, "Core Inflation: A Critical Guide," Working Paper No. 242 (London: Bank of England). Summary in Bank of England *Quarterly Bulletin*, Vol. 44 (Winter), p. 466.

Mankiw, N. Gregory, and Ricardo Reis, 2003, "What Measures of Inflation Should a Central Bank Target? *Journal of the European Economic Association*, Vol. 1 (September), pp. 1058–1086.

Maria, José Ramos, 2004, "On the Use of the First Principal Component as a Core Inflation Indicator," Economic Research Department Working Paper No. 3-04 (Lisbon: Banco de Portugal).

Marques, Carlos Robalo, Pedro Duarte Neves, and Luís Morais Sarmento, 2000, "Evaluating Core Inflation Indicators," Economic Research Department Working Paper No. 3-00 (Lisbon: Banco de Portugal).

———, 2003, "Evaluating Core Inflation Indicators," *Economic Modelling*, Vol. 20 (July), pp. 765–775.

Mrázek, Jiří, 2006, "Core Inflation Measurement—Czech Experience," paper presented at the United Nations Economic Commission for Europe Conference of European Statisticians, Group of Experts on Consumer Prices Indices, Geneva, May 10–12, Available via the Internet: http://www.unece.org/stats/documents/ece/ces/ge.22/2006/mtg2/sp.3.e.pdf.

Parkin, Michael, 1984, "On 'Core Inflation' by Otto Eckstein: A Review Essay," *Journal of Monetary Economics*, Vol. 14 (September), pp. 251–264.

Porrado, Eric, and Andrés Velasco, 2002, "Alternative Monetary Rules in the Open Economy: A Welfare-Based Approach," in *Inflation Targeting: Design, Performance, Challenges*, Series on Central Banking, Analysis, and Economic Policies, ed. by Norman Loayza and Raimundo Soto (Santiago, Banco Central de Chile), Vol. 5, pp. 295–348.

Quah, Danny, and Shaun P. Vahey, 1995, "Measuring Core Inflation," *Economic Journal*, Vol. 105 (September), pp. 1130–1144.

Roger, Scott, 1997, "A Robust Measure of Core Inflation in New Zealand, 1949–96," Discussion Paper No. G97/7 (Wellington, Reserve Bank of New Zealand).

Roger, Scott, 1998, "Core Inflation: Concepts, Uses and Measurement," Discussion Paper No. G98/9 (Wellington, Reserve Bank of New Zealand).

———, 2000, "Relative Prices, Inflation and Core Inflation," IMF Working Paper 00/58 (Washington, International Monetary Fund).

Rowlatt, Amanda, 2001, "The U.K. Office for National Statistics and the Inflation Target," in *Statistical Implications of Inflation Targeting: Getting the Right Numbers and Getting the Numbers Right*, ed. by C. Carson, C. Enoch, and C. Dziobek (Washington, International Monetary Fund), pp. 125–136.

San José, Armida, Graham L. Slack, and Subramanian S. Sriram, 2002, "Statistical Principles for Inflation Targeting Regimes and the Role of IMF Data Initiatives," in *Statistical Implications of Inflation Targeting: Getting the Right Numbers and Getting the Numbers Right*, ed. by C. Carson, C. Enoch, and C. Dziobek (Washington: International Monetary Fund), pp. 308–339.

Schaechter, Andrea, Mark R. Stone, and Mark Zelmer, 2000, "Adopting Inflation Targeting: Practical Issues for Emerging Market Countries," IMF Occasional Paper No. 202 (Washington, International Monetary Fund).

Silver, Mick, and Christos Ioannidis, 1996, "Inflation, Relative Prices and Their Skewness," *Applied Economics*, Vol. 28 (May), pp. 577–584.

Sorensen, Alan T., 2000, "Equilibrium Price Dispersion in Retail Markets for Prescription Drugs," *Journal of Political Economy*, Vol. 108 (August), pp. 833–850.

Wynne, Mark A., 1999, "Core Inflation: A Review of Some Conceptual Issues," ECB Working Paper No. 5 (Frankfurt, European Central Bank).