UNITED NATIONS **E**

Economic and Social Council

Distr.
GENERAL

ECE/CES/GE.22/2006/6
14 February 2006

ENGLISH ONLY

ECONOMIC COMMISSION FOR EUROPE **STATISTICAL COMMISSION**

CONFERENCE OF EUROPEAN STATISTICIANS

Group of Experts on Consumer Price Indices

Eighth Meeting
Geneva, 10-12 May 2006
Item 6 of the provisional agenda

CORE INFLATION: MEASURES AND THEIR CHOICE[1]

Submitted by International Monetary Fund

The meeting is organised jointly with the International Labour Organization (ILO)

Abstract

This paper provides an overview of statistical measurement issues relating to alternative measures of core inflation, and the criteria for choosing between them. The approaches to measurement considered include exclusion-based methods, imputation 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 cointegration tests, and correlation with money supply. Country practice can differ in how the approaches are implemented and how their veracity is assessed. There is little consistency in the results of country studies to readily suggest guidelines on accepted methods.

JEL Classification Numbers: C43, E31, E52
Keywords: Core Inflation; Core CPI; Inflation Targeting

[1] This paper has been prepared by Mr. Mick Silver, International Monetary Fund, at the invitation of the secretariat.

I. INTRODUCTION

1. 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.[2] Since the price changes of some components of the CPI, including food and vegetables (due to weather conditions) and energy (due to supply shocks) are particularly volatile, these components are usually excluded from the target. So too may be indirect taxes and interest (mortgage) payments, since the former are erratic one-off changes and the latter a tool, 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.

2. The major benefits of inflation targeting will only be realized if the target measure used is credible, 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.

3. The paper outlines in Section II some concepts and practical issues regarding inflation targeting to provide a context to the discussions on measurement and choice of methods. In Section III sources of errors and bias in the CPI are briefly outlined. This is because measures of core inflation are generally derived from the CPI and inadequacies in the latter will generally be passed on to the former. This takes us to the main purpose of the paper, which is quite simple. It is first, to **outline the range of methods available**. While exclusion-based measures are often used as target and/or operational measures of core inflation, there are many alternative approaches and these alternative approaches 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, considered 102 measures of core inflation using Australian data. Section IV provides an outline of the methods.

4. Since 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 and, again, there are a number of such criteria and a number of methods by which each criterion can be formulated and empirically tested. Section V outlines methods for judging which is best by different criteria. There are also a good number of empirical studies to draw on, yet they vary according to the country, time period, criteria for selection adopted, and measures considered. Even when taking into account such variation, no unanimity as to the best measure(s) emerges, with conclusions changing even for sub-periods of the same study.

[2] Bloem, Armknecht, and Zieschang (2002) provides an outline of, and the case for, alternative indexes as the basis for the core inflation measure.

5. 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 this should give confidence to 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.

II. CONCEPTS AND PRACTICAL ISSUES

6. The strongest argument for an inflation targeting framework is to provide monetary authorities with a reliable "nominal anchor." Since paper-money is intrinsically worthless, the nominal anchor ties the price level at a specified time to a specified level. An inflation target acts as an alternative to a nominal anchor, which would fix the value of paper-money to gold or foreign currency, or the amount of money in circulation.

7. The credibility of the targeted measure is of prime importance. The idea is that consumers and producers make decisions as to how much to buy and sell on the basis of what they think inflation will be. If they make a mistake in anticipating inflation, the prices that are charged and the amount produced and sold will be affected. The economy will be operating inefficiently because the market mechanism is distorted. If inflation is kept low then there will be less room for mistakes and less of a welfare loss to the economy due to the unanticipated component of inflation. However, the welfare loss will also be minimized if buyers and sellers have a good idea of what inflation is likely to be—they can anchor their expectations on the basis of a well-anticipated inflationary target. All of this in turn requires confidence that the monetary authorities can achieve the target (range), and confidence that the target is a meaningful measure. If there is little public faith in the target measure, then inflation-targeting will be of little value in this respect. Public expectations of inflation will not be anchored on a target few believe in.

8. To be effective the index should be one the public is familiar with; thus the prevalence of the CPI as a target (though see Bloem, Armknecht, and Zieschang, 2002). The central bank must explain to the public how the core price index is constructed and its relation to the headline rate of inflation, which it may in fact be. The public should not be under the impression that the measure chosen has been selected on the basis that it is likely to guarantee favorable results. The measure should be clearly defined and reproducible, changed as infrequently as possible, and should be produced by, or at least derived from, a CPI produced by an independent statistical authority (Bernanke, and others, 1999, pp. 27–8 and Blejer and others, 2000). The role of inflation targeting in the context of the IMF-supported adjustment programs is considered in Blejer and others (1999 pp. 409–99) and for Brazil as a case study, in Cerisola and Gelos (2005).

9. Different measures of core inflation and criteria for their choice serve different purposes. A timely and credible measure may be required as the target core measure. Measures may also be required which best smooth the data, so that in assessing policy there is an understanding of the extent to which fluctuations in the series can be regarded as having arisen from "noise." Furthermore, forward-looking measures may be required for prediction. It may be that some measures serve more than one function, but this is an empirical matter to be decided for each country.

10. It is argued here that the evaluation should, where possible, be data-driven, that is, be evaluated using recent data and acceptable criteria as outlined below. The empirical studies clearly show that methods suitable in one country cannot be carried over to other countries. However, where data are limited, the methods evaluated for countries with similar patterns of inflation and economies should be used, rather than, for example, using those proven for developed countries in developing countries. Data-driven methods using clear guidelines also lend an element of objectivity to the exercise.

11. In considering measures of core inflation as part of a framework for inflation targeting, attention should be given to the institutional arrangements regarding the production of core inflation measures. To be effective in anchoring inflation they must be credible and such credibility is derived not just from the quality and suitability of the measures, but also the transparency of the source data and compilation methods and the credibility of the agencies concerned. Generally, an autonomous statistical agency is best placed to fulfill this function, but in many cases the central bank might be considered more appropriate. In either case, core inflation should not be considered as an attempt to measure inflation more accurately than the CPI. The central bank has a lead role to play in the development of these measures, because their use is for monetary policy purposes. But as a derived statistic, the statistical agency will have to work closely with the central bank in their development.

III. SOURCES OF ERROR AND BIAS IN A CPI

12. Targeted inflation may be the headline CPI, or a derived core inflation measure. In either case central banks should be aware of the sources of error and bias in their country's CPI. If such errors and bias are serious, then the levels of targeted inflation achieved will not accord with the experiences of the population and, thus, not effectively anchor inflation. This will then cause a loss of confidence in the target which may take many years to restore. Research into sources of errors and bias of the CPI, as well as transparency of the results of the research, are important to the long-run success of an inflation targeting framework.

13. There is a second reason why sources of error and bias impact on the measurement of core inflation (Roger, 2000). Some of the measures and tests in Section V as to which measure is best rely on a characterization given in equation (2) below whereby shocks are random and normally distributed. Shocks affect relative prices, but in the long-run aggregate inflation is unchanged by them. This is because the shocks are held to be accommodated by flexible price-setting—increases are counterbalanced by decreases. However, the formula used for price indexes utilizes a fixed basket, so it does not reveal the substitution effects. Furthermore, the coverage of the CPI may not include all goods and services so that the full balancing act may not be revealed. The practical manner in which CPIs are compiled is that even with annual weight updates, there is, by construction, a period of 12 months during which weights are kept fixed. However, the weights will remain constant for longer than the 12 months since it can take at least 6 months to compile the expenditure surveys. Price updating the weights over this 6 months would include the effects of the shock.

14. Sources of error and bias for a CPI are well outlined by Greenlees and Balk (2004) in a summary chapter of the *CPI Manual* (ILO and others, 2004). They include **sampling error**, both for the price quotes from the monthly price surveys and the weights from the Household Budget

Surveys. Often non-random sampling methods are used which precludes the direct estimation of such errors, though they will exist nonetheless and will be larger with smaller sample sizes and more dispersed price changes. There will be **non-sampling errors** including **non-response errors** from the surveys, **under- and over-coverage** of the desired scope, **inappropriate measures** of the underlying concepts (for example, non-accrual, list prices as opposed to accrual, transaction prices), and **response errors** (such as biased responses for alcohol consumption, processing errors). **Aggregation bias** at the elementary level and higher levels may arise from inappropriate formulas. For example, at the elementary level, the arithmetic average of price ratios (Carli index) and, for non homogeneous goods and services, the arithmetic ratio of averages (Dutot index) have undesirable properties and are biased due to their failure of the time reversal and commensurability tests, respectively (ILO and others, 2004, pp. 363–64). Furthermore these fixed basket elementary aggregate indexes — and at the higher level, the Laspeyres index — give rise to **substitution bias** since they do not properly take account of the change in weights as consumers substitute away from goods and services with above average price increases. Bias may also arise from not fully taking into account the price effect from the introduction of and switches to **new products**, or products from **new outlets**, and the separation of the effect on price of **changes in the quality** of existing goods and services or outlets.

15. The very **concept** of the CPI will dictate its scope and methods; for example, a cost-of-living index (COLI) will exclude expenditure by non-residents, while a "monetary" inflation index will include such expenditures. A COLI's desired aggregation index is a superlative (say Fisher) index that includes substitution effects while a strict fixed-base definition may have a Laspeyres-type index as its goal.

16. It is not the purpose of this paper to examine these sources of errors and bias in any depth and the reader is referred to Greenlees and Balk (2004) and, more generally, the *CPI Manual* (2004) for a summary and details.

IV. THE METHODS[3]

17. The methods can be grouped into those that are suitable for **policy assessment**, that is, they are 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 and the criteria/tests for choice of measure in Section V that there is a substantive difference in measurement and assessment.

Policy assessment
- Exclusion-based methods
    - Product groups
    - Indirect taxes
    - One-off shocks
    - Domestically generated inflation
    - Imputation methods
- Trend estimates

[3] An account of the evolution of different measures of central tendency is given in Roger (2000: Appendix II).

- Limited influence estimators
  - Median
  - Trimmed means—symmetric and asymmetric

Prediction
- Reweighting the CPI
  - Persistence weights
  - Volatility weights
  - First principal component
- Economic models.

18. First, we briefly define the CPI and then consider the use of exclusion-based measures. They are particularly important because of their possible role as inflation targets, as well as measures of core inflation devised to help predict/assess CPI targeting. Section IV continues with an outline of all other methods. The CPI, π_t, is defined as a weighted mean of price changes. Let $\dot{p}_i^t$ be the change in prices of expenditure group *i* from a price reference period 0 to the current period *t*; then:

$$\pi_t = \sum_i w_i^b \frac{p_i^t}{p_i^0} = \sum_i w_i^b \dot{p}_i^t \qquad (1)$$

where the weights, w_i^b, are normalized, relative expenditures for product group *i* available in some period *b,* prior to period 0.[4] In practice, the $\dot{p}_i^t$ are generally unweighted elementary aggregate indexes derived from matched price comparisons for similar items across outlets. There are many advantages to their computation using geometric means, as discussed in Diewert (2004b). Different countries use different, or even a mix of, formulas for such indexes. Common practice is the use of the geometric mean (Jevons index) and ratio of arithmetic means (the Dutot index) of prices—see Diewert (2004b) for details. Equation (1) is a Young index since the period used for the weights, period *b*, differs from the price reference period 0. This is because it takes time to collect and compile data on expenditure weights. If period *b* was the same as period 0 in (1), the index number formula would be Laspeyres. Often statistical offices price-update the expenditure values from period *b* to period 0, the resulting index being a Lowe index. Equation (1) is not theoretically desirable for it does not take account of substitution effects. As consumers substitute expenditure away from goods and services with above average price increases the Young/Lowe/Laspeyres index holds the basket fixed and thus overstates price changes. The preferable target indexes are superlative indexes such as the Fisher and Törnqvist indexes, which would make use, in this context, of period *b* and period *t,* expenditure weights, though data problems limit their application unless calculated retrospectively—see Diewert (2004a) for details.

[4] In practice household budget expenditure surveys (HESs) for the CPIs are conducted less frequently, say annually versus monthly price collection for the monthly index. There will also be a lag between the survey period of the HESs and their compilation before use in the index. Thus the period of weights differs from the price reference period. The resulting indexes are Young or Lowe indexes depending on whether the period *b* weights are price updated or otherwise. Their properties are discussed in Diewert (2004a and 2004b).

We characterize CPI inflation, π_t, as core inflation, π_t^*, plus a temporary disturbance, v_t, i.e.:

$$\pi_t = \pi_t^* + v_t \qquad (2)$$

where v_t is random and *normally* distributed. Core inflation is considered to be a monetary phenomenon. 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 flexible and to adjust for any supply/demand shocks. There would be an instant substitution away from (say) a particularly high price change. The shock would be "accommodated" by relative price and quantity changes. The shock should not influence mean inflation. Only changes in the money supply are held to do so.

19. Measures of core inflation are thus required to separate the signal of inflation from the temporary noise or volatility[5]—what Cecchetti (1997) refers to as "transitory phenomenon" that should not affect policymaker's actions. This should allow a better assessment of the *current* inflationary pressure which is needed for targeting inflation.

Seasonality and the observation interval

20. Before considering alternative measures of core inflation, the periodicity of the price changes has yet to be defined. This may be dictated by the periodicity of the inflation target or, for prediction or analytical purposes, may be decided by the producers/users of the measures. More than one definition may be used to gain further insights into the underlying data.

21. Inflation targets should not be subject to seasonal fluctuations and, thus, should be 12-month rates, comparing the price level in a month with that in the same month in the preceding year. These will be less volatile than month-on-month rates. There may well also be an interest in month-on-month changes for which core inflation measures may be seasonally adjusted to gain insights into price changes.

22. Cecchetti (1997) found the extent of seasonal variation to be quite substantial over sub-periods of his study.6 He also found that core inflation measures cannot be relied upon to remove seasonality. He found that the extent of the seasonality of the overall CPI to be similar for trimmed means and medians, but there was a marked *increase* in seasonality when food and energy were excluded. Fenwick (2004) argues the case for seasonally adjusting an

[5] It is worth distinguishing noise from bias. There is an extensive literature on sources of bias in a CPI or PPI (see Greenlees and Balk, 2004). Such bias includes substitution bias, as the fixed weights of a price index do not reflect a change in the basket of goods away from (toward) goods with above (below) average price changes; bias from an inability to properly incorporate the effects of quality changes and new goods; formula bias; and sample selection bias. Cecchetti (1997) has argued that such bias may be time varying and thus mistaken for noise. As some bias can be removed by the statistical procedures used to extract noise, Blinder (1997, p. 158) cautions against using such procedures as the basis for bias removal since an important issue for central bankers in low inflation settings is the inflation *level*, as well as its trend, and this is best considered by first reducing the bias—by "fixing" the index—and then use statistical signal extraction procedures for the noise.

[6] Seasonality was found to be quite limited for U.S. CPI quarterly data over the period 1967:01 to 1996:04—an R^2 of 0.07 when regressing the series on seasonal dummies. However, on further consideration the extent of the seasonal influence was found to vary dramatically, with an equivalent R^2 of 0.34 for 1982:01 to 1996:04.

exclusion-based index in order to better identify the long-term trend. Irregular changes, such as those due to mortgage interest payments and indirect taxes are first removed, since these irregular components may obscure or confound the seasonal patterns. The resulting series may then be seasonally adjusted using, for transparency, seasonal factors derived from the publicly available X12 program developed by the US Bureau of the Census.[7]

23. If month-on-month series are required or seasonality remains in the 12-month series, there is the question as to whether to adjust for the seasonal influences at the aggregate or component levels. A first step might be to test for seasonality at each of the component levels and, if present, seasonally adjust the series. Bryan and Cecchetti (1996) caution against this and argue that seasonal adjustment should be undertaken at the aggregate level since first, it is the aggregate series that central bankers are most interested in. Second, the pre-test decision not to seasonally adjust a disaggregated component series might be subject to relatively high type I errors—that is, the probability of rejecting the null hypothesis of no seasonality, when there is in fact seasonality—due to the high variances often found in relative price movements.

24. The use of 12-month rates cannot of course deal with irregular, yet periodic, price changes. Mankikar and Paisley (2004) illustrate how a regular change in the price level—say every year in December the government increases postage charges—will lead to stable 12-month inflation rates. However, if the changes are infrequent—for example, a mix of October, November, and December—the 12-month series will be highly volatile. Moving averages over 12-months or, in this example, even 3-months would remove the volatility. Some prices are collected only on a quarterly or annual basis and this would give rise to spikes in these periods, which would be otherwise smoothed by averaging. The averaging of CPI data over a few periods—low frequency data—may in itself provide a good means for reducing noise. Blinder (1997, p. 159), drawing on his central bank experience, notes: "My view, in brief, is that it takes at least three consecutive months numbers before you have any meaningful information."

25. Cecchetti (1997) also found substantial gains from averaging the data over longer periods. The noise from using the 10 percent trimmed mean was cut by over 50 percent when averaged over three months rather than one month, and by over 70 percent when averaged over six months. A long period, say 36-month moving average, thus may be used as a target index that eliminates noise. It should be noted that comparing prices averaged over say three periods, rather than monthly prices, may smooth the data, but up-to-date timely information would be smothered in the averaging.[8]

[7] There are two series which may be derived. The first applies estimated seasonal factors from the X-12 program—derived in turn from moving average trend estimates—to an exclusion-based index to remove its seasonality. The second, is the moving average trend estimates from the X-12 program of the an exclusion-based index. The former series includes irregular fluctuations, while the latter smoothes them. However, a (say) 12 month moving average trend cannot provide trend estimates for the first 6 months and last 6 months of the series (see section III(C). Fenwick (2004) acknowledges that such methods suffer from a loss of information at the start and end of the series.

[8] When calculating moving averages, the resulting average is generally treated as the observation at the center of the time span (for example, month two of a three month span and month seven of a 13-month span). Thus, the start and the end of the moving average series are shortened by length of the average chosen.

A. Exclusion-Based Methods

26. An exclusion-based CPI for use as an inflation target has much to commend it. It is easy to understand, timely, and transparent, in that the user can replicate the measure. The CPI usually has the credibility and exposure required of a measure whose purpose in anchoring inflation is to affect inflation expectations. Exclusion-based methods are often used by countries when they first instigate inflation targets. A common approach is simply to exclude certain product groups. Usual exclusions are food and energy (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 used by a number of countries has the advantage that the authorities are less likely to be perceived to be manipulating the targeting. However, the grounds cited for the exclusion of F&E are their volatility and such components need not be the most volatile. In this section we argue that the decision to exclude specific sectors should, in part, be data driven not only on the grounds of minimizing volatility but also for signaling an objectivity to the choice of target.

27. It is stressed that if the CPI, or some derivative, is used as a target measure, exclusion-based measures may also be used to help facilitate the targeting process and, indeed, more than one exclusion-measure may be used. 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

28. Usual exclusions are food and energy (F&E) argued on the basis of their undue volatility. Some studies have found the inclusion of certain food items, such as fresh fruit and vegetables, and energy items, such as petrol, makes the CPI *more* volatile (Cecchetti, 1997). In some countries all of 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 such volatility exists and, preferably, that there is an economic rationale for its continued existence.

29. Rather than excluding “standard” volatile product groups, a preferred procedure is data-driven, in that each country examines its own past data to determine which components are the most volatile. Such a procedure may suffer from two problems: first, that the components that are found to be volatile may become relatively stable over time, and second, the components established as not being volatile may become volatile. Prior empirical work would be required to ensure that the components selected for exclusion also had some longevity in their volatility. The selection of components to be excluded would also have to be regularly assessed, but on a pre-established, timed basis so as not to give the appearance of interference with the methodology. For Canada, for example, the eight most volatile components were selected on the basis of historical data and excluded, along with indirect taxes (Macklem, 2001). Excluded items were: fruit, vegetables, gasoline, fuel oil, natural gas, inter-city 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—alternative measures of price change volatility are discussed in Section V.

30. Kearns (1998: Table B1) lists 105 components of the Australian CPI and the number of times, out of the last 70 price quarters, 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."

31. Mankikar and Paisley (2004) advocate the identification by the trimmed mean of which product groups are excluded in the majority of cases. For the U.K. they found that of the 21 component product groups (from 81) that were excluded 50 percent of the time—by a 15 percent trimmed mean, between 1975 and 2002—five were seasonal food items, two energy, and four discounted products in monthly sales. The continuity of products being excluded provides some justification for the continued use of the trimming. It might be argued that it also provides justification for the exclusion of such product groups. This would be on the grounds that policymakers would have a better handle on what core inflation consistently measured via its exclusion, rather than some trimming rules.

32. Heath and others (2004) have argued that it should not just be an empirical matter, but the rationale for excess volatility should also be explained. For example, it might be argued that a core inflation measure excludes volatile items, such as seasonal fruit and vegetable since they are subject to large temporary price fluctuations; petrol is affected by fluctuations in the exchange rate and world oil prices and then there are large irregular prices set by the public sector.

33. 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, such exclusions may lose the credibility of the measure amongst poorer members of society, as argued for South Africa by Lehohla and Myburgh (2002).

34. An alternative procedure is not to remove established product groups, but frequently (say monthly, or every three months) establish which product groups are, for example, the eight most volatile components and exclude them. Such a rule suffers from the problem that such components may change. The index is then influenced by the changing mix of the product groups in the basket. There is an analog to chained indexes here whereby the basket is changed regularly to represent updated expenditure patterns. Here it is a basket regularly updated to exclude the most volatile sectors and differences between this index and the sector-specific exclusion index may be of interest as they will at the very least demonstrate a need to reconsider the basket of excluded items.

35. Cutler (2001) notes that in phrasing the problem of separating the signal and the noise, the noise is being defined as being 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, such as weather, oil prices, and exchange rates that feed through to a number of sectors to different degrees.

36. The exclusion of energy price changes is more difficult to justify since such shocks can be substantial, but can have a durable effect on inflation. The calculation of core inflation with and without energy may provide useful information.

Indirect taxes

37. Indirect taxes have a first-round effect of raising prices in proportion to the tax change. They may feed through to wages and the prices of other goods and services. The aim would be to remove the first-round effect. Consider a hike in a sales tax; the core inflation series should have the price series adjusted to remove this one-off change. An increase, for example, would otherwise raise 12-month inflation rates for the duration of 12 months of comparisons, but then the effect would no longer exist as prices, including taxes, are compared with the same tax component 12-months ago. Such changes obscure the long-run trend of the series. For example, Hogan and others (2001) show the merits of excluding both the effects of the introduction of VAT in Canada in 1991 and the decline of the tobacco tax in 1994. Cutler (2001) reports the effect of a sharp change in local authority taxes in the U.K. increasing, between April 1990 and March 1991, by an average 12-month rate of 34 percent and then falling by 29 percent April 1991 to March 1992. She fond that when the tax effect was excluded, the "hump" change was *more* protracted, the increase and fall in the tax change disguising a longer (albeit slightly lower) momentum to an inflationary high in this period.

38. *Ad hoc* adjustments to remove severe indirect tax changes effects may diminish the credibility of the index. An alternative and more acceptable procedure is to exclude all indirect taxes. Hogan and others (2001) note, however, that the resulting indirect tax-excluded index relies on the unlikely assumption that tax changes are passed through immediately and on a one-for-one basis to consumer prices.

39. More than one exclusion-based index provides analytical insights. An index excluding the product groups established to be the most volatile and one excluding such groups and indirect taxes will both be useful and provide complementary insights. In all cases, public confidence permitting, the effects of erratic, one-off shocks should be excluded.

Interest rates

40. In some countries mortgage interest rates are included in their CPI with regard to owner-occupied housing. Interest may also be included as part of FISIM. With their inclusion, for example, an increase in interest rates designed to lower the rate of inflation would contribute to inflation. In countries where there are doubts as to the credibility of the CPI and concerns over statistical "interference" with the target measure, there may be a case for including interest rates in the core inflation measures if they are a small proportion of expenditure, but generally they should be excluded. In the U.K., for example, mortgage interest payments are excluded from the target for inflation (see Rowlatt, 2001).

Other major one-off or erratic shocks

41. Some extreme price changes occur because of one-off or irregular shocks which are known to have a temporary effect on prices, but they are not part of demand-induced inflation. For example, price changes of postage stamps may occur irregularly, and when they occur they may be large. But they form part of the inflationary process being induced by cost/demand pressures and should be included. Excluded should be exogenous shocks which give rise to one-off price change, say due to 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; for if it is low, the public are unlikely to anchor inflation expectations to the targeted measure, thus removing much of the point of the exercise (as argued in Section II). Such procedures are only as good as the judgments they are based on. By definition a "good" call as to what are short-term demand-supply shocks will remove such effects effectively, and without any spillover to other product groups.

Domestically generated inflation

42. External shocks in some economies may be usefully considered as arising from erratic movements in the exchange rate. If such traded goods are stripped from the index, the resulting domestically generated inflation may erase some 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 RPIX (consumer prices) excluding imports, and a measure based on unit labor costs (ULC). 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. Yet in many economies a few traded goods can be responsible for a large proportion of trade, and if the exchange rate is erratic, such domestically generated inflation series may give useful insights into the underlying pressure of inflation, although such series should not be relied upon as measures of core inflation.

Imputation-Based Methods

43. If the purpose of the core inflation measure is to exclude volatile product groups—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. Roger (2000) has noted 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. 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 assuming it is the same as the mean. For example, a crude way to divide them would be into durable goods, non-durable goods and services, though other categorizations may be more plausible. The weight for excluded groups would be assigned to product groups likely to experience similar, core, smoother price changes. In such a case the method of exclusion might suffer less from the loss of information.

B. Trend Estimates

44. The use of trend estimates, 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 moving average, trend-moving

average estimates for the first and last 18 months are not provided and would have to be extrapolated based on some model. This is a serious deficiency of the method as it does not provide real-time estimates of core inflation. 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 V(C) of this paper is to make use of the smoothed nature of the trend estimates as retrospective reference series in that we choose between competing core inflation measures on the basis of their past deviations from the smooth trend estimates.

C. Limited Influence Estimators

The weighted median

45. The calculation of the weighted median requires that price changes of expenditure group *i* from period *t* to *t*-12, defined now as $\dot{p}_i^t$, are first ranked, as are their associated expenditures weights, w_i^b, and related cumulative normalized, relative expenditure weights, Cw_i^b. For example, if we have 105 product groups with weights given as a ratio of 1,000, "telephone services" may have the lowest price change, say $\dot{p}_i^t$=0.87 and a weight w_i^b of 30 and Cw_i^b=30; "milk and cheese" may be next with a weight 40 and Cw_i^b=70 and so forth as given in Table 1 below.

Table 1. Ordering of price relatives to determine weighted median value

Product group	$\dot{p}_i^t$	w_i^b	Cw_i^b
Telephone services	0.8700	30	30
Milk and cheese	0.8800	40	70
↓	↓	↓	↓
Pet foods	1.0220	10	490
Men's clothing	1.0240	30	520
Men's footwear	1.0245	8	528
↓	↓	↓	↓

46. 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. In our example, the median price change would be the price change of the product group corresponding to the $n/2^{th} = 500^{th}$ cumulative weight, i.e., that of men's clothing which is 1.024, a 2.4 percent change. Of course the price relative of 1.024 is itself an average of values which may only just fall into the "men's clothing" price interval or just before "men's footwear." The median is better calculated treating the price changes as a continuum rather than a set of discrete steps. A standard textbook adjustment[9] is to use:

[9] This differs a little from the one in Kearns (1998).

$$\dot{p}_{m-1} + (\dot{p}_{m+1} - \dot{p}_{m-1}) \times \left(\frac{n/2 - Cw_{m-1}}{Cw_m - Cw_{m-1}} \right) \tag{3}$$

$$= 1.022 + (1.0245 - 1.022) \times \left(\frac{500 - 490}{520 - 490} \right)$$

$$= 1.02283$$

where the subscripts m, $m - 1$ and $m + 1$ refer to the median observation, observation before and after it respectively.

47. The median, as the middle price change, when price changes are ranked in order of magnitude, benefits from the fact that it can be easily explained. It, of course, makes use of all the information in the data set in determining the middle observation, but is unaffected by extreme values at either end of the distribution. In the above example telephone services may have had a price change of 0.002, a fall of 99.8 percent, as opposed to the fall of 13 percent given, but the median would remain the same. In this manner it strips out extreme price changes. Trimmed means are a less extreme version of the median. As will be seen below, the median is a more efficient and robust estimator of the population mean when the distribution of price changes is not normal. The computation of the median is timely and transparent with regard to its replication, and easy to compile and explain. It is an extreme form of a trimmed symmetric mean suffering from the loss of much of sector-specific signal information, though gaining from being highly robust to shocks in many product groups.

Trimmed symmetric means

48. A trimmed mean removes 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 of the $\dot{p}_i^t$ ranking and 10 percent of the weight at the bottom of the $\dot{p}_i^t$ ranking. The remaining weights are normalized and the weighted mean of the remaining 80 percent of price changes form the measure.

49. The calculation could be undertaken to exclude the whole product groups, if some or all of their cumulative weight intrudes into the top and bottom 10 percent, or to give the borderline product group a weight appropriate to how much it intrudes into the cumulative distribution. The former identifies trimming as a method of identifying outlier product groups and excluding them. It has the advantage of being well defined since, for a particular price comparison, it can be described in terms of which product groups are excluded. However, the exclusion/inclusion of product groups can change over time so such definitions have little merit. A smoother index would be one which incorporated the proportion of weight that strictly lay inside the middle 80 percent. For example, consider a 10 percent trim which should exclude the bottom 0.05(1000)= 50 units of the weights. In Table 1 we might exclude just ‘telephone services’ and ‘milk and cheese’, but it would be more appropriate to set the weight for telephone services to zero and change the weight of ‘milk and cheese’ to 20, and then normalize the weights after similar adjustments to the top tail.

50. Trimmed mean estimators are timely, transparent with regard to their replication, and easy to compile once a decision has been made on the nature and size of the trim. Such a decision is the only judgmental intervention involved, though more than one trim may be used so that any perceived arbitrariness of the trim can be countered and analytical insights gained through using more than one comparable measure. For such purposes trimmed mean estimators cannot be defined in terms of the product-groups excluded since this may well change.

51. Trimmed mean estimators can be 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. Aucremanne (2000) and Heath and others (2004) consider the purpose of trimming to reduce the price distribution to one that is normally distributed and use the non-rejection of a Jarque-Bera normality test as an indicator of having achieved an appropriate trim. The Jarque-Bera $\chi^2(2)$ test is concerned with the rejection of a null hypothesis of normality relating to the (symmetrically combined) differences due to skewness and kurtosis statistics. There are some concerns with this approach. 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, and even as a necessary condition, the conclusion depends on the power of the test. Third, the level of trim from this method would of course vary over time. This and alternative criteria for the selection of formulas are discussed in more detail in Section V.

52. Trimmed means are not without problems. First, the users of such measures should be aware of the nature of shocks taking place. As Mankikar and Paisley (2004) note, the supply shock outbreak of foot-and–mouth disease in 2001 in the U.K. led to large price rises for beef which would be helpfully trimmed out, but the subsequent smaller readjustments back over several months would not be trimmed out. Core inflation would appear to be falling when it would not be, it would be readjusting. Similar effects may arise in the tourism sector following a natural or terrorist disaster. An economy experiencing a series of positive shocks in different sectors with slow rates of adjustment back, would give rise to trimmed core inflation movements that understated inflation. In a similar vein Mankikar and Paisley (2004) note Bakhshi and Yates' (1999) advice that if only a few price setters initially respond to an aggregate demand increase, then trimming out the price changes removes the valuable information in the tails. "...knowing the source of the shock is crucial in determining whether it is wise to trim." Mankikar and Paisley (2004: p. 17).

53. However, applying such wisdom is problematic with trimmed means because once defined, they are not subject to manipulation. For this reason trimmed means should be used at more than one level of trim, including the median as an extreme, so that disparities can be revealed and perhaps explained in terms of the factors underlying them. Compilers can easily experiment with such levels of aggregation and trim to determine their effect. Such experimentation can be repeated periodically, or as and when the economy goes through major changes.

54. 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.

55. Third, a problem with trimmed means is that they have been found to be systematically lower than CPI means inferring that they are doing more than trimming random shocks (Cutler, 2001 and Kearns, 1998). Roger (2000) points out that these lower values are in line with the empirical finding of skewness of price changes and suggests trimming more of the right-hand tail than the left. The skewness of price changes should be explored by countries using their own data. Indeed if means and medians are being calculated, a measure of skewness advocated by Pearson based on their difference is given by:

$$Skew_{Pearson} = \frac{3\left(\pi^{t}_{w\Delta p} - Median^{t}_{w\Delta p}\right)}{\sigma^{t}_{w\Delta p}}, \qquad (4)$$

where $\sigma^{t}_{w\Delta p}$ is the weighted standard deviation of price changes and statistics of plus and minus 3 denote very high positive and very high negative skewness respectively. Where the distribution is found to be non-normal, symmetric trimming should not, at least alone, be relied upon.

Asymmetric and variable trimmed measures

The non-normality of price-change distribution

56. There is an extensive literature on findings of, and theoretical reasons to expect, non-normal distributions of price changes. The consistent body of empirical evidence—referenced in Roger (2000: pp. 5–6)[10]—finds the distribution of price changes to be skewed to the right and leptokurtic (fat-tailed). For example, an early study was Bryan and Cecchetti (1996) who examined 36 components of the U.S. CPI between January 1967 and April 1996. The price changes were considered over month-on-month (k=1), month-on-three-month (k=3), annual (k=12), and k=24 periods. Over k=1 and k=3 the weighted kurtosis figures for the CPI was found to be about 8, the unweighted statistics being even higher; these are very high.[11] The kurtosis decreased as k increased to 24 periods, though not to normality. The skewness was always positive over all values of k though decreased as k increased. Again it was substantial with values of 0.322 and 0.233 for the CPI for k=1 and k=3 respectively.

57. The non-normality of the distribution of price changes has implications for the measurement of core inflation. First, the sample median is a *less* efficient estimator of the population mean than the median when the distribution is leptokurtic. Second, the mean would be pulled up by extreme observations in a positively skewed distribution, and if skewness is positively related to mean inflation, the mean would increase over time as a result of what may be perceived to be noise rather than signal. However, if the distribution of price changes is normal, the mean, trimmed mean and median will give the same result. It is necessary to consider whether the non-normality can be considered to be part of the signal, and thus should be left to influence the core measure, or part of the noise and removed by exclusion or trimming.

[10] Studies include Roger (1997, 2000), Bryan and Cecchetti (1996), Silver and Ioannidis (1996), Bakshi and Yates (1997), Kearns (1998), Heath and others (2004), and Cutler (2004).

[11] In a Monte Carlo experiment of samples of 36, 95 percent of kurtosis statistics from an empirical normal distribution lie between 1.67 and 5.57 (Bryan and Cecchetti, 1996).

58. Explanations of non-normality of price changes lie (i) with imperfections in the CPI measure, and (ii) with prices that are not fully flexible. The former was discussed in Section II and relate to the inability of fixed weight CPIs to properly reflect the random nature of the shocks characterized in equation (2). Such a characterization further requires flexible prices. In practice prices are not fully flexible, at least in the short-to-medium term. Furthermore, shocks may have a long duration even if prices are fully flexible, if time is necessary for the consumers/producers to react to the shock, possibly as a result of institutional factors—see, for example, the effect of the use of subsidies in Chile to cushion the 1999 oil price increase, García, (2002). There are more formal theoretical frameworks to explain inflexible prices.

59. Menu cost models are based on the premise that there are costs to undertaking price changes; price changes are not fully flexible. Thus, price setters will only change their prices if, say for an increase, their desired price is over and above (outside of the bounds of) the menu costs. This results in (asymmetric) staggered price changes which give rise to a positive relationship between skewness and inflation (Ball and Mankiw, 1995). The asymmetry in the skewness arises from the belief that price setters wishing to increase their nominal prices will do so more often than those wishing to decrease their prices. The menu cost argument is that firms let inflation do the work for them; relative price decreases can be achieved by holding nominal prices constant and allowing inflation to achieve the real, relative price decreases.[12]

60. Roger (2000) considered the prime sources of the skewness to be infrequently adjusted prices due to government-set or regulated prices, seasonal goods, and goods sampled less frequently, rather than the price stickiness of menu costs. Flexible prices will have a normal distribution. He argues that when prices are changed there will be a 'blip' of x times the trend rate where x is the periodicity of the pent-up price change. The higher the rate of inflation the higher the skewness of the price change distribution. When price changes do not occur there will be spikes. He demonstrates how the skewness and kurtosis of the distribution of log-normal price changes will be high even if only a relatively small percent, such as 4 percent, of price changes are infrequent.

61. Balke and Wynne (1996) consider the effects of supply shocks in a multi-sector model in which the effect on prices of the shocks varies with the (fixed) productivity of the sector resulting in skewed price changes.[13]

62. There is also an argument for price stickiness from search cost theory in which optimizing consumers with imperfect information search for additional information such that their (rising) marginal search cost equals their (falling) marginal search benefits (Stigler, 1961). This would result in imperfect information that the firm can exploit by charging a higher price (Sorensen, 2000). It is therefore in the interests of firms to adopt strategies which increase search costs. 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 (Van Hoomissen, 1988, Sorensen, 2000 and Lach, 2002).

[12] Though Roger (2000) shows positive skewness is a phenomena in high, medium and low inflation times.
[13] However, Bryan and Cecchetti (1996) find that the correlation between skewness and the mean suggested by the menu cost and multi-sector model can be explained as being due to small sample bias.

63. The next section presents the argument that first, mean percentiles should be used as an alternative to the median when there is skewness; second, variable trims be used to maintain normality; and finally, the mean be not used as an estimator when the distribution is leptokurtic. The latter is raised because exclusion-based methods, for all their merits, may also suffer from non-normal price distribution *after* exclusion. If this is the case, the resulting estimator may be neither efficient nor robust.

Skewness and the mean percentile estimator

64. There is a lot of empirical evidence to suggest that the distribution of price changes may be non-normal, or more particularly, leptokurtic and skewed to the right. Roger (2000) argued 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, as noted above, much empirical evidence of such asymmetries in developed countries (for example, Australia: Kearns (1998), Heath and others (2004); New Zealand: Roger (1997, 1998); USA: Bryan and Ceccecheti (1996)). Our concern here is to preserve such information as it is part of measured inflation, and not have the trimming throw out such information with the 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 not to centre the trim on the 50th percentile, but to centre it on the percentile that ensures the average of price changes in the underlying variable lines up with that corresponding to the target variable. Otherwise symmetric trimmed means, including the median, may result in a series that consistently understates inflation. Mean percentiles are one way of dealing with asymmetric data. We would start by first ascertaining whether there was a skewed price relative distribution using, for example, equation (4). If so, the next step is to order the observations into percentiles. Mean percentiles are calculated by first taking the weighted arithmetic mean of all of the data. For a normal distribution the mean is the value of the observation that corresponds to the 50^{th} percentile. For a positively skewed distribution the mean will be pulled upwards, say to the 53^{rd} percentile. The mean percentile is the value of the price change of the percentile class in which the mean falls.

65. This issue has proved particularly problematic in New Zealand, where strong and persistent right-hand skewness in the distribution of price changes resulted in a large difference between the weighted mean and weighted median, and it has been shown that the 57th percentile was a more appropriate centre (Roger, 1997). In the Australian context, Kearns (1998) found that centers around the 51^{st} percentile were most appropriate.

Variable trimmed measures

66. Aucremanne (2000) takes as a criterion for the choice of the appropriate percentile the one that minimizes the average absolute difference between the target and expected inflation. Heath and others (2004) extend Aucremanne's (2000) use of the Jarque-Bera test statistic as a basis for selecting the level of trim that jointly removes skewness and excess kurtosis. The level of trim required to do this would vary each period and Heath and others (2004) reports some favorable results from its use. The Jarque-Bera statistic is used to decide between measures with varying percentages of trim and with varying central percentiles (between 40 and 60 percent).

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. Some concerns were raised above about the use of the Jarque-Bera statistic, though we return to this in Section V.

67. We have noted that if there is skewness, symmetric means and medians will be biased against a target arithmetic mean. A further approach is to continue with the biased measures, but then correct them. A simple rescaling method was successfully applied for similar reasons by the Banco de Portugal for use with its principal components estimator (see below). They defined the rescaled indicator, the core inflation *level*, as the one corresponding to the fitted values of a regression of the target CPI on the core measure. In order to get an estimator computable in real time, successive regressions needed to be estimated each time, each including an additional observation (José, 2004).

Kurtosis and the efficiency and robustness of a sample estimator

68. The concern here is with utilizing the most **efficient** estimator of core inflation. Roger (2000, pp. 34–35) based on Yule (1911) provides an excellent, detailed account of the relationship between the efficiency of the estimators and the underlying parameters of the two sub-distributions. The efficiency of the estimator depends on the distribution of price relatives that the data are drawn from. The sample mean is the most efficient estimator of the population mean if the distribution is normal. Departures from normality can arise from either kurtosis, skewness, or a combination of both. It can be shown that the sample median is a *more* efficient estimator of the population mean than the median when the distribution is leptokurtic.[14] The median is an extreme form of a trimmed mean and it follows that trimmed means are also more efficient estimators than means for leptokurtic distributions.[15] The question is the degree of kurtosis that should exist in the data before there should be a swap from the mean to a median or trimmed mean for the sake of efficiency. Roger (2000, pp. 34–35) provides an excellent account of the parameters that govern the relationship.

69. Huber (1964) discusses robust estimates based on their sensitivity to outliers. If there are no outliers and the distribution is normal, then the mean and median are the same. However, consider a single outlier value, say an extremely high price change. Its effect is to pull up the mean. Now compare alternative estimators of central tendency. A criterion for choice might be that a good estimator is one whose sum of squared differences between it and the individual values is a minimum. The mean would be such a desirable least squares measure of central tendency. 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. Of course, the mean is pulled up by extreme values. Now consider an estimator that seeks to minimize the sum of the *absolute* differences between it and the observations. Then there is no

[14] It is easier to draw an observation in one tail of the distribution that is not counterbalanced by the draw of an observation on the other side, even if the population distribution is symmetric, when the population is leptokurtic.
[15] Experiments of the quite marked increases in efficiency over sample means are described in Bryan and Cecchetti(1996).

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. As a least squares estimator, the mean is sensitive to outliers and, thus, is only robust if the distribution is normal where similar extreme values appear in both tails of the distribution. Since a core inflation measure must function in a way that strips away noise, then a measure sensitive to noise cannot be advocated, unless the noise is symmetrically distributed. A less sensitive criterion is desired and the mean absolute deviation is more relevant in this context . Note that Section III considers the criteria for choice between alternative methods in terms of prediction or correspondence to smoothed values, and there is choice between using a robust measure (mean absolute deviation) or efficient estimator (root mean squared error).

70. Yet the empirical evidence is of leptokurtic distributions. In this case the mean is not as efficient as the median. Robust estimators are efficient estimators which are not sensitive to non-normalities in the distribution. Roger (2000) cites Hogg's (1967) and Harter's (1974, 1975) advice to avoid the sample mean if the kurtosis exceeds 4 and 3.7 respectively. He also, as discussed in Section IV, notes that the reproducibility and general comprehensibility of the measure are also considered:

> "In weighing up these criteria and the potential trade-off involved, the recommendations of experts in the field, based on their close examination of the properties of different estimators, are pretty straightforward. First, relatively simple estimators, such as trimmed means, tend to be recommended over more complicated estimators on the grounds that they are easier to understand. Second, the higher the kurtosis of the distribution, the less weight the estimator should place on observations in the tail of the distribution." Roger (2000: 42).

Volatility weights

71. Trimmed means and exclusion-based indices lose information in the sectors experiencing extreme price changes. Volatility weights 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:

$$\frac{\sum_{i}^{n} \frac{1}{\mathrm{v}ol(\pi_i)} \pi_i^t}{\sum_{i}^{n} \frac{1}{\mathrm{v}ol(\pi_i)}} \qquad (5)$$

where $\mathrm{v}ol(\pi_i)$ is an indicator of volatility; Diewert (1995) demonstrates that the variance of price relatives is appropriate.[16] However, in Marques and others (2000) the volatility weights used were the standard deviations of the *deviations* in inflation rates, i.e.

[16] There is a practical problem in calculating this index since a measure of the mean price change in period *t* is required to calculate the variance for *i* in *t*, but the mean in turn requires a measure of the variance. Vega and Wynne (2003) uses a simple arithmetic mean as an initial estimate of the mean to calculate the variance, and iterates further.

$$vol(\pi_i) = \sigma_i^t = \sqrt{\sum_t \left[\left(\pi_i^t - \pi^t \right) - \left(\overline{\pi_i^t - \pi^t} \right) \right]^2} \qquad (6)$$

where $\left(\overline{\pi_i^t - \pi^t} \right) = \dfrac{\sum_t \left(\pi_i^t - \pi^t \right)}{m}$ and π_i^t is the 12-month inflation rate in period *t* for *i*=1....*n* product categories in *t*=1, ...*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 and others (2004) considers a rolling 4-year period as well as the full length of the series.

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

$$\frac{\sum_i^n \dfrac{w_i \, p_i^t}{\sigma_i^t}}{\sum_i^n \dfrac{w_i \, p_i^{t-12}}{\sigma_i^t}} \qquad (7)$$

or can also be unweighted with $w_i = 1/n$, as is implicit in equation (5), though since 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.

First principle component

73. The use of principle 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 José (2004). Its use for Portugal was supported by cointegration tests (see Section IV below) by Marques, Neves, and Sarmento (2000). Principal components analysis is a data reduction technique in the sense that given 12-month price changes $\dot{p}_k$ of 1,...,*k* product groups it seeks to establish a smaller number of variables (*j*<*k*) that are linear combinations, z_j of $\dot{p}_k$ such that:

$$\begin{aligned} z_1 &= a_1\dot{p}_1 + a_2\dot{p}_2 + ...a_k\dot{p}_k \\ z_2 &= b_1\dot{p}_1 + b_2\dot{p}_2 + ...b_k\dot{p}_k \\ &\vdots \\ z_j &= c_1\dot{p}_1 + c_2\dot{p}_2 + ...c_k\dot{p}_k \end{aligned} \qquad (8)$$

account for most of the variance of the original set. Each principal component is a weighted basket of price changes, the weights being the coefficients. The first principal component z_1 accounts for the largest proportion of the variance, z_2 accounts for the second largest, etc. Each principal component (PC) is orthogonal to subsequent components. Since the PCs are not scale

invariant, it is first necessary to standardize the price changes. Standard statistical software contain routines for deriving the coefficients and, thus, PCs. Only the first principal component, z_1 is used for this analysis. The higher the proportion of variance explained by z_1 and the greater the stability of the explanatory power of z_1 over time, the more reliable is the method.

74. A very small estimated coefficient, say a_2, attached to product group 2 denotes that the first PC gives a commensurately small weight to the prices changes $\dot{p}_2$ of this product group because it has a relatively small signal (variance) to noise ratio. In studies by Banco de Portugal the weight of the volatile aggregates "unprocessed food" and "energy" were much smaller in the first principal component than in the consumer price index. The concern of the method is to derive a smooth series by way of minimizing the influence of variables with a high noise to signal ratio. A potential problem, however, arises with integrated series of order one. As the indexes change, so too will the sampling variance, even if the change is smooth. Thus, series with relatively high rates of change will appear more volatile if only the variance is looked at. The average level of the first principal component can be shown not to be the same as that of the CPI. Therefore, it needs to be rescaled, and there are several alternative routines for this. A simple one adopted successfully in the Banco de Portugal studies was to run a regression equation between the inflation rate and the first principal component. Then they defined the rescaled indicator, the core inflation *level*, as the one corresponding to the fitted values of the regression. In order to get an estimator calculated in real time, successive regressions needed to be estimated, each including an additional observation (see Machado and others (2001), José (2004), and other related papers on the Banco de Portugal's web site).

D. Prediction

75. 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 not misled by price changes that are believed to be the result of one-off economic shocks, and have no information on future inflation. Policies should be made on the basis of information that shows the more durable phenomena. Blinder (1995, p. 157) defines *durable* as the part useful in medium- and near-term inflation forecasting. This is the "outlook" or "persistence" approach to inflation, which may be over a one- or two-year horizon for relatively stable inflation, but may (also) be much shorter depending on the needs for intervention.

76. 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 (Clemen, 1989). Jacobson and Karlsson (2004) use a Bayesian model for forecasting inflation which can also provide an indicator of the relative effectiveness of each forecasting method. However, the immediate concern is with their individual characteristics.

Reweighting the CPI: persistence weights

77. The concern here is to give more weight to product groups considered best able to forecast. Cutler (2001) estimated the weights for the i=1,...,80 components of the U.K. RPIX from a first-order autoregressive (AR) model:

$$\pi_i^t = \alpha_i + \rho_i \pi_i^{t-12} + \varepsilon_i^t \qquad (9)$$

where π_i^t are the price changes[17] for category *i* as the normalized, and positive $\hat{\rho}_i$ taken as an indicator of the persistence of inflation in each category *i*. Categories with negative $\hat{\rho}_i$ were assigned a weight of zero justified on the grounds of their rapid mean reversion. The persistence weighted index is given by:

$$\frac{\sum_{i \in \rho_i > 0} \hat{\rho}_i \pi_i^t}{\sum_{i \in \rho_i > 0} \hat{\rho}_i} \qquad (10)$$

78. The persistence weights were changed each year using a rolled-forward monthly data set, commencing in January 1976; for example, the 1999 weights were based on estimates over the period January 1976 to December 1998 and, for 2000, on estimates over the period January 1976 to December 1999. The need for a lengthy time series for the estimate means that much of the data on which the estimates are based are quite unrelated to the period of the price comparison. This is especially problematic since the data are treated symmetrically in the estimator, with just as much influence given to 1976 as to 1998 for the 1999 weight. Second, the data are overlapping so that any changes in the estimated coefficients will be smoothed. This is not to negate the usefulness or indeed the concept behind the application, but to draw attention to limitations in operationalizing the procedure.

79. Persistence weights may differ from budget expenditure weights because the former may exclude some components—those with negative estimated parameters—and because of differences in the magnitude of the remaining weights. 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 non-seasonal food items on the grounds that they were found to have information useful in prediction. Petrol and oil had a very much lower weight, higher persistence weights for coal, electricity, gas oil and fuel compensated for this. The exclusion of all food and energy product groups in CPIs for the U.K. was argued to be not justified by Cutler (2001) since components with valuable predictive ability were removed. The case for persistence weights lies with the inclusion of statistical information in the time series properties of the disaggregated components which are useful for prediction.

80. Cutler (2001) compared the predictive ability of a persistence-weighted core index with other exclusion-based measures in terms of their predictive ability over and above inflation. The persistence-weighted index ranked third over a number of time horizons when compared with seven other core indexes. However, when the predictive ability of the persistence-weighted index was tested with the addition of further lags it, along with two exclusion indexes, proved to be superior to the trimmed mean and weighted median indexes, the result being statistically significant at a 5 percent level. Yet the measure has some shortcomings.

[17] Cutler (2001) advocates the use of month-on-month price changes for the components which are then combined by successive multiplication to form an index. The annual core inflation series is derived from this series as the current value of the index compared with the previous 12-month index value. This procedure is held to smooth the effect of step changes in the weights, but for volatile items, could generate chain drift.

81. First, it is relatively complex . Second, as will be discussed in Section IV, tests of predictive ability suffer from the Lucas critique. This is that if policy makers react to the measures, they should have less predictive power. Lucas (1976) argues that the parameters of traditional macroeconometric models depend implicitly on agents' expectations of the policy process and are unlikely to remain stable as policymakers change their behavior. Given historical policy changes and a plausible empirical forward-looking autoregressive model, the estimated parameters of the model should be unstable. The reweighting of indexes using information which naturally includes the responses of policy makers, as is the case with the persistence weighted index, may suffer particularly from this problem, though the econometric evidence for the Lucas critique is not strong (see, for example, Rudebusch, 2005) and, as Cutler (2001) points out, actual decisions by policy makers are based on more than one economic indicator.

82. Third, there is much in the construction of the index, such as changes in the type of products included and classification changes over the periods in which weights are estimated, that may give rise to unstable or biased coefficients. Fourth, the definition of the dynamics of persistence given by $\hat{\rho}_i$ in equation (9) above is quite restrictive with regard to being constrained to each *i*.

83. Finally, as with any method that effectively removes the weighting of a component on the grounds of volatility, it is quite possible that such components, while noisier, will have different long-run trends to other components on the whole. If this is so, this is valuable information which is part of a CPI estimate. It would be ignored on the grounds of its noise to signal ratio, not on the grounds of whether the signal is useful or not.

Short-run prediction

84. An interest in a prediction approach begs a question as to the time period over which predictions are required. This in turn, and in part, relates to the time period over which it takes monetary policy to take effect and is usually considered as the medium term, say 18 months. Yet there may also be the need to respond quickly (or at least be perceived as responding quickly) to short-run changes. As such it is necessary to ask whether there are appropriate methods for short-run, say one period-ahead, forecasts as opposed to the medium- to long-run persistence measures considered above. Section IV provides details of methodologies for choosing between core inflation measures based on predictive criteria. Such criteria must be specific to the time frame required, of which there may be more than one. If both short- and medium-term measures are required, then suitable measures are required for each to be accordingly appraised.

85. Short-term univariate forecasts can be derived using the target series. Two widely used methods are the Holt-Winters exponential smoothing and the Box Jenkins approach. They may be applied to month-on-month or 12-month inflation. There are advantages to using the Holt-Winters exponential smoothing since it has an intuitive explanation: the method decomposes the series into smoothed, trend and seasonal estimates with older data having an (exponentially) decreasing effect in determining the constituent parameters. It can also be shown that the predicted results have a self-correcting characteristic in that if the prediction is, say, above the actual data in one period, it will correct downwards in the next. The seasonal components will allocate (exponentially) more weight to more recent data. The nature of these methods is well documented and exact expressions for the expected values of multi-step-ahead

forecasts and their prediction intervals are given in Chatfield and Yar (1991), and Hyndman and others. (2005). These statistical models are distinguished from the economic models below which have a rationale in economic theory, and whose concern is to predict over much longer time horizons than considered here.

Economic models

86. Core inflation is characterized as a series arising from an estimated econometric model as opposed to being measured from combining, in different ways, price relative and weight information. The grounding of the models in theory is an obvious advantage since it can benefit from the incorporation of further economic variables, thus realizing a core inflation measure from a multivariate setting, abstracting from (or conditioning on) the effects of these other variables. It also makes explicit the economic drivers of core inflation so that the underlying process can be better understood, as can departures from it. However, it is sensitive to the Lucas critique whereby if policy were based on a relationship between core inflation and the CPI, that very relationship would change as a result of the realization of the policy. Furthermore, 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 judgments as well as purely statistical considerations.
Following Roger (1998), consider first a model due to Eckstein (1981) in which the short-run aggregate supply curve is given by:

$$\pi_t = \pi_{t+1}^{LR} + g(x_{t+1}) + \varepsilon_t \qquad (11)$$

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

$$\pi_t^* = \pi_t^{LR} = \pi_t - g(x_{t-1}) - \varepsilon_t . \qquad (12)$$

87. Note that under this concept of core inflation cyclical fluctuations are removed. This may be compared with Ceccechetti (1997) who used a 36-month moving average as a target core inflation, the concern being with removing transitory (in a 36-month sense) noise only. If the purpose is also to remove cycles, then equation (12) is more suitable. A practical issue is with the derivation of suitable measures of excess demand.

88. The methods do not discard information and, as with smoothing, make no assumptions as to the time dimension of the smoothing. It also has the strength of being grounded in an economic model, though this has as a possible weakness the validity of the model itself—see Parkin (1984) for a critique of Eckstein (1981).

89. 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 U.K. The view is that disturbances are benign to output since an economy will adjust to their effects. Core inflation is output-neutral in the long run. If there are rigidities such as menu or search costs or expectations

errors, in that the economic agents do not properly anticipate inflation and make wrong decisions affecting real output, then core inflationary shocks will affect real output in the short-term, but not medium- to long-term. The identification restriction in the SVAR estimator allows the data to determine whether or not the economy quickly adjusts to these core inflationary disturbances. Different identification schemes to those suggested by Quah and Vahey (1995) have been used. Folkertsma and Hubrich (2000) compared five schemes using European data finding remarkable differences[18] in the range of measurement error, but also expressed, for policy purposes, concern about the extent of the errors:

> "The probability of a measurement error exceeding 1 percent when estimating the level of broad core inflation varies with the identification scheme between 11.3 percent and 30.5 percent and for a measurement error exceeding 0.5 percent between 42.3 percent and 60.6 percent." Folkertsma and Hubrich (2000, p. 496).

90. Wynne(1999) criticizes the approach on the grounds that each time the index is re-estimated, it has to be revised and that such measures are difficult to communicate to the public. Roger (1998) considers an important difference between the use of the Eckstein (1981) and Quah and Vahey (1995) models to be the time horizon; with the former the core inflation is not considered to be cyclical over the time horizon of the policy maker.

91. Economic models can aid in the analysis of core inflation. More particularly equation (12) fits in well with simple frameworks in equation (1) with an additional concern of abstracting long-run price movements due to excess-demand pressure. Tests of the effect, and identification of the extent, of parameters on $g(x_{t+1})$ are of interest in this respect.

92. The information set used in economic models need not be confined to the CPI. Bagliano and others (2002), in a study of the Euro area (1979 to 2000), use series on inflation, money, output and interest rates to estimate a forward-looking measure of core inflation based on the long-run (cointegrating) relations among these variables. Mankiw and Reis (2003) developed a framework for, and provides estimates for the US (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 (i) their expenditure share, as is appropriate for a CPI, (ii) their sensitivity to business cycles, (iii) their likelihood of experiencing idiosyncratic shocks, and (iv) 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 would provide an index that, if kept on target, would lead to the greatest stability in economic activity. Included in the empirical work for the US was 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 had a large weight in the stability price index. The energy sector was also found to be pro-cyclical, but, unlike nominal wages, had a much higher likelihood of experiencing idiosyncratic shocks, and as a result of this, a lower weight. If the aim is to target a

[18] The differences are only remarkable in terms of their magnitude with regard to policy use. They are not remarkable in terms of expectations of such results. Folkertsma and Hubrich (2000) draw attention to Cooley and Dwyer (1998) who show how applications of SVAR models often rely on long-run identifying restrictions or on time series properties which are either impossible or inherently difficult to test.

measure that aimed at economic stability, the stability price index was shown to be preferred to a CPI.

V. HOW TO CHOOSE AMONG METHODS: JUDGING WHICH IS BEST

93. 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, the concern now turns to the choice among these measures.

94. Given 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 method that is best is the subject of this section.

95. 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 would 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.

96. A second and related problem is that changes in CPI methodology may have been recently undertaken to bring it up to standard. However, the past CPI series is deemed to be unreliable. It is advisable that data-driven methods, at the very minimum, be based on 36 months of data to allow seasonal components to be estimated to avoid seasonality having an undue influence. As noted above, the basis for core inflation measures should be 12-month rather than month-on-month price relatives, and for practical purposes this requires a bare minimum of 24 months of data. Patterns in economies take time to emerge, and the results of two or three years may be particular to the events of that period. While data-driven exercises are still advised in such cases, the results of (more reliable) studies of countries with similar economies should be borne in mind when making a choice among measures.

97. We now turn to the methods for judging which measures are best. This to a large extent depends on the purpose or needs of the central bank's decision-making process. There are also more general considerations which we consider first.

A. Credibility and General Considerations

98. Roger (1998) argues that a 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 with the public.

99. Credibility can be of prime importance. In part this is because one of the purposes of inflation targeting is to anchor inflation expectations to the target. The idea is that consumers and producers make decisions as to how much to buy and sell on the basis of what they think inflation will be. If they make a mistake in anticipating inflation, the prices that are charged and the amount produced and sold will be wrong. The economy will be operating inefficiently. If inflation is kept low, then there will be less room for mistakes and less of a welfare loss to the economy due to the unanticipated component of inflation. However, the welfare loss will also be minimized if buyers and sellers have a good idea of what inflation is likely to be—they can anchor their expectations on the basis of a well-anticipated inflationary target. All of this in turn requires confidence that the monetary authorities can achieve the target (range), and confidence that the target is a measure meaningful to them. If there is little public faith in the CPI, then inflation-targeting will be of little value in this respect. Public expectations of inflation will not be anchored on a target few believe in.

100. For example, a producer wanting to increase real prices over some period by 5 percent who believes in a CPI inflation target of 2 percent may increase prices by 7 percent. The resource allocation via the price mechanism for the economy is deemed to be working well. But if there is no faith in the CPI measure, whether rightly or otherwise, the producer will have to make a judgment as to what inflation will be, say 4 percent. If inflation really is 2 percent there will be a misallocation of resources by the actual price increase of 9 percent. Of central importance to monetary authorities is that the CPI used for targeting is plausible, and this may take resources to do so as well as measures to improve its image. One reason why the CPI, as opposed to say the producer price index (PPI), is used for inflation targeting is its high profile, in the belief that inflation expectations are more likely to be anchored to it. For these reasons the considerations listed above by Roger (2000) are crucial to the selection of an appropriate "headline" inflation target.

101. In this regard, the selection of method may in part be based on the perceived confidence in the index . For example, stripping out food and energy 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 (Lehola and others, 2002). Transparency in methodology is also an important ingredient in credibility, and the IMF's data dissemination standards for the CPI—the basis of the vast majority of measures—are important in this respect (José and others, 2002).

102. The agency responsible for constructing the CPI is also important to its credibility as a target. As Schaechter and others (2000, p. 9) note:

> "Compilation of the CPI and core inflation by an independent agency, typically the country's statistical agency, can improve credibility by avoiding the perception that the central bank manipulates the data."

103. While this generally holds, it is also the case that central banks may have their own measures used for targeting inflation for the very reason that there is little credibility in the statistical agency's figures. In the long run, however, the aim should be toward the development

of credible CPI statistics from an independent statistical agency for this is likely to better anchor inflation expectations.

104. Wynne(1999), in commenting on Roger's considerations, 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.

105. These are of course not absolute criteria and different policy makers will apply different weights to each. They may also be seen to be prerequisites of good measures, but again this must be weighed against needs. There may well, for example, be a trade-off between simplicity and bias. As noted in Section I, 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, computable in real time, and non-revisable should be heavily weighted. In the second, it would be important that the series were forward-looking and/or provided analytical insights. The weights given to these criteria may well vary among countries. For example, countries undergoing some political/economic transition in which there was a lack of confidence in data preceding the transition, might put more weight on deriving a credible measure rather than other statistical criteria.

106. 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. Marques and others (2002) comment that, while many of the above criteria are sensible, they are somewhat vague and do little to clarify exactly what statistical conditions a suitable underlying inflation indicator should satisfy.

107. 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. Yet the complexity of the statistical criteria used may harm the transparency of the selection and, thus, acceptability of the measures. Again, it is a matter for individual countries to consider. It is reiterated that more than one measure may be used and the CPI, or simple exclusion-based derivatives of it, may be used as the target with core inflation measures of different complexities used to help operationalize the targeting.

108. The question considered next, is what can statistical offices do, using past data, to determine which measure(s) are appropriate? More credibility can be assigned to measures that have been, at least in part, selected on the basis of objective statistical criteria. Furthermore, especially for operational purposes, methods chosen by appropriate statistical criteria will, by

definition, do the job better. We look at the types of data analysis available to judge which methods are best.

B. Judging on the Basis of Control

109. Blinder (1997, p. 160) argued 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].

110. Porrado and Velasco (1999) take a similar stance with regard to the CPI not being an appropriate measure since it includes non-domestically produced goods and services. They argue that since the CPI is affected by exchange rate variations that the central bank has no control of, 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 IV(A)).

111. 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 no control. The argument is to exclude them. On the other, as considered in Section II, 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).

C. Judging on the Basis of Deviations from a Reference Series

112. 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. First is the need to calculate a reference index. Then the results of alternative methods can be graphed alongside the reference one. Summary averages of the deviations of the results from each method from the reference series can also be calculated so that methods can be more readily compared against each other. Thus, if the reference index is taken to be a measure of core inflation, π_t^*, and the measure of core inflation being assessed is $\pi_t^{'}$, then the best measure might be one that minimizes its root mean square error (RMSE):

$$\sqrt{\frac{\sum_{t=1}^{T}\left(\pi_t^{'}-\pi_t^{*}\right)^2}{T}} \qquad (13)$$

or its mean absolute deviation (MAD):

$$\frac{\sum_{t=1}^{T}\left|\pi_t^{'}-\pi_t^{*}\right|}{T} \qquad (14)$$

113. Bryan and Cecchetti (1994), Bryan, Cecchetti and Wiggins (1997) and Cecchetti (1996, 1997) used a 36-month centered moving average of actual inflation as the reference series and the root mean squared error (RMSE) as a summary measure of the overall deviations,[19] although they commented that in general other summary measures led to similar conclusions. However, Bakshi and Yates (1999) found that the mean absolute deviation (MAD) can provide different results from the RMSE. If the measure departs in a single month by a large margin from the reference series, then the squaring of the difference between the two series puts a much larger penalty on the deviation than the mean absolute deviation would. Central bankers have to ask themselves 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. A strategy might be to compute both; if they give similar answers, the choice of minimization criterion is not contentious. If differences are substantial, then this may well arise from a few abnormal deviations, and their nature and their importance should be evaluated for inflation targeting.

114. It is of course possible to consider the distribution of the deviations and again Cecchetti (1997) is helpful. As well as the average deviation in equations (13) and (14) above, he calculates the 12.5 and 87.5 percentiles of the deviations from the reference index, i.e. he calculates the range of deviations within which the middle 75 percent of deviations lie. For example, for 12-month rates, he finds the MSE (without the Root) of the US CPI-U to have a 12.5 percent percentile *range* (which excludes 12.5 percent on either side) of 0.42 to -0.51. This method can be seen to be useful in assessing how robust the decisions as to the best measures are to outlier deviations. Similar conclusions should result when comparing the RMSE and the MAD. Such measures are also useful in phrasing the magnitude of the expected deviations of a measure from the reference series. Bear in mind the above 12.5 percent percentile range for 12-month inflation rates was about 1 percent wide around a 36-month moving average. If the reference series, the centered moving average, corresponded to the target and a target band that was 1 percentage point wide was used, historical experience implies that the measure would be outside of the band one-quarter of the time.

115. It is also worth noting that both summary measures in equations (13) and (14) assume a symmetric aversion to over- and under-estimating the reference rate. It may be that central bankers are more worried about getting a large increase wrong than a similarly large decrease.

[19] This analysis was a little more complicated being interested in the efficiency of the estimator. The summary measures were bootstrap estimators.

The RMSE and MAD would not choose methods that take this into account. It can be seen, however, that the use of asymmetric percentile ranges may serve us well in this respect—say to choose the measure that has the widest range between the 10^{th} and 95^{th} percentile.

116. The method relies on the suitability of the reference index, the (centered) moving average. This is a smoothing technique whose purpose is to estimate the trend of data in a decomposition of a time series. It suffers from its 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 as it does not allow real-time estimates of core inflation. However, it is used here as a reference measure—as a means to consider the retrospective performance of difference measures. The identification of this reference trend is by smoothing the data through moving averages. A trend estimate for a month, say July 1995, is the centered moving average (CMA) of 36 successive price change observations, from January 1994 to December 1996. Any seasonal or other irregularities are smoothed in the averaging. The next estimate for August 1995 is derived by dropping the first observation, January 1994, and moving the average along to include January 1997, and so forth. If there is a regular monthly pattern, say high in January but low in December, the CMA on a long series of monthly data would estimate the trend through such highs and lows. If the data in 1997 are progressively higher, for example, than in previous years, the CMA would follow the trend; it would progressively increase as it dropped lower values and included new higher ones. There may be erratic large price changes and this would affect the index in a particular month as the average moves to include them. However, if the smoothing takes place over a lengthy time period, the expectation is that there will also be countervailing large price decreases, and a CMA over a period of 36 months should serve to smooth these out. To its credit, the CMA uses all of the CPI data, unlike limited-information estimators. The CMA approach easy to calculate and has an intuitive justification and plausibility, especially when calculated and graphed over noisy series. Such smoothed series can be calculated for 12-month price comparisons or on a month-on-month basis; the CMA itself is an estimate of the trend component as distinct from seasonal, cyclical and residual factors. However, there have been a number of criticisms over this use of smoothing to generate a reference series.

117. 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 forecast future inflation rates, then an appropriate procedure would be to give more weight to components that better forecast the future, rather that to weight in accordance with its level of noise in the past.

118. Second, as Mankikar and Paisley (2004) have pointed out, there is no economic rationale for smoothness to be desirable. Economies can go through periods of sharp fluctuations in demand and supply which 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. Marques and others (2000) note that centered moving averages are known to preserve linearity, i.e. they are optimal estimators when the trend of a series is a linear function of time. They also show that if the series is integrated of order 1, I(1)—see Section V(F) —then the CMA has nice properties. Their very nature makes them a poor device for replicating a core series that is anything but smooth.

119. Third, Bakshi and Yates(1999) have found the results can depend on the number of periods used in the averaging. Aucremanne (2000) and Heath and others (2004), using Australian data, found 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. Marques and others (2000) evaluated the properties of three moving averages over time horizons of 13-, 25-, and 37-months. They found only the 37-month (36-centered) moving average met their time series tests (outlined below). In view of this ambiguity, Heath and others (2004) argued that rather than focus on a single optimal measure of underlying inflation, a central bank should consider a collection of underlying measures. An alternative stance is to 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. It is then for the monetary authorities to decide on the required balance between smoothness and time horizon.

120. Yet a core inflation measure that best follows a 36-month CMA has an interesting property. An estimate of the moving average for, say June 1995 is based, as noted above, upon data smoothed over the period January 1994 to December 1996. A core inflation measure that has the smallest root mean squared error would compare the core inflation measure with this smoothed value for June 1995, and other months. It would retrospectively compare a core measure that uses data for the current period or the preceding periods only, with a reference series that smoothes data 18 months into the future and past of each value. It answers a hypothetical question as to what would be the optimal measure in terms of a reference index that is deemed to be correct because it has extended and averaged over a very long period into the past and future to smooth out its noise, but maintains enough information to minimize bias. This is no small task and a CMA measure serves us well in this respect.

121. Cecchetti (1997) evaluated alternative methods in terms of their RMSE from a long-run trend in inflation. He used U.S. monthly data between January 1982 and April 1996 for 36 components and found the 10 percent trimmed mean to be the most appropriate. It was also considered better to include F & E than exclude it. There is no uniformity in findings from this approach. For example, Kearns (1998), using Australia data and deviations from a smoothed mean as the criterion of choice, following Bryan and Cecchetti (1994), found a weighted median optimal. Countries must be encouraged to undertake their own analysis to find what is best for them.

D. Justifying the Exclusion of Product Groups on the Basis of their Volatility

122. As noted above, exclusion-based measures are often justified on the grounds 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. Product groups may also be excluded since they are conceptually undesirable, such as interest (mortgage) payments and one-off price changes. The concern here is with the first case, product groups excluded on the grounds of their perceived volatility. Perceptions of volatility are not sufficient grounds for the exclusion of product groups. The relative volatility of price changes should also be examined empirically. If they are found to be more volatile and if it is considered reasonable that past patterns will continue into the future, there is then grounds for their exclusion. Often food and energy product groups are excluded on *a*

priori reasons alone. Is it reasonable for a country to first consider, from their own data, whether food and energy are more volatile? There are a number of ways of doing this.

123. First, is to use the standard deviation of the 12-month rate of inflation for the CPI and then, for the CPI excluding F&E. Here no reference index is used, just a measure of the extent to which the series fluctuates over time. Cecchetti (1997) using U.S. CPI-U data found the standard deviation to be 2.33 percent (CPI-U) and 2.58 percent (CPI-U excluding F & E) percent 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. He found this also held for sub-periods of the data and also for when the price changes were averaged over 3-months, though not 6-months. There is also the question of how much less should the standard deviation be to justify exclusion? While F-tests of statistical significance may be used to ascertain whether a null hypothesis of no difference between variances can be rejected at a specified level, this is only a necessary condition for the exclusion of a volatile product group. The difference must also be substantial in magnitude to sufficiently smooth the index. Graphs should also be used with such summary measures.

124. The second approach is to consider volatility to be deviations from a long-term reference series—a 36-month moving average—as discussed above using equations (13) and (14), or symmetric /asymmetric percentile bounds. Cecchetti (1997) found the MSE (without the Root) for a 12-month (and other) price indexes 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 U.K., found on this basis that a series excluding F&E was a little smoother than a series which included it. Cecchetti (1997) also considered the use of 12.5 percent percentile bands outlined above for the CPI-U series with and without F & E. For 12-month rates, he found the MSE of the CPI-U to have a 12.5 percent percentile range of 0.42 to -0.51 while excluding F & E to have a range of 0.90 to -0.18. The latter was larger, again demonstrating that F & E are more volatile, even when the extreme 25 percent of price deviations are ignored.

125. 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 only because the food products with less volatile price changes have been grouped at too high a level of aggregation with products with non-volatile price changes. Cutler (2001) undertook her analysis of measures of core inflation for the U.K. using 77 product groups. She used 25 separate food groups and identified that the level of persistence varied substantially within these food group products. Potatoes, vegetables, and fish proved to lack any persistence (signal), though even in this finely wrought study, there was no further decomposition as in the manner suggested above. Of course the level of detail will depend on the availability and reliability of data and the importance of the product to household expenditure.

126. A starting operational point to considering the level of disaggregation would be elementary aggregate indexes—the level at which weights are first applied. Note how even the 58 COICOP product groups can obscure quite different patterns. At this level food is treated as a single group. This needs to be further disaggregated into its 9 classes, but even this is unsatisfactory. For example, fish is given by division 01.1.3 and includes:

01.1.3 Fish (ND)

- Fresh, chilled or frozen fish;
- Fresh, chilled or frozen seafood (crustaceans including land crabs, molluscs and other shellfish, land and sea snails, frogs);
- Dried, smoked or salted fish and seafood;
- Other preserved or processed fish and seafood and fish and seafood-based preparations (canned fish and seafood, caviar and other hard roes, fish pies, etc.).

Includes: fish and seafood purchased live for consumption as food.

Excludes: soups, broths and stocks containing fish (01.1.9). Frozen, dried, salted, 'other preserved or processed (including canned) fish will be less susceptible to seasonal fluctuations than fresh or chilled fish. Similarly with vegetables.

127. Frozen, smoked, and salted fish may be far less volatile than other groups. Blinder (1997) 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 below in the next section provide forecasts, and different measures of core inflation can be used in them to identify which provides the best forecast. They may also be used for different product groups and, following Blinder (1997), used to ascertain which product group contains information to better predict its own, or aggregate, future values. However, it is the former that is the context of this section. We now turn to the later.

E. Judging on the Basis of Predictive Ability

128. First, is a simple model, to estimate regression models of the CPI inflation figures on different core inflation estimates and calculated RMSEs for forecasts, i.e., the model is

$$\hat{\pi}_t = \alpha + \beta \pi^*_{t-12} \qquad (15)$$

and our interest is in the RMSE deviation between $\hat{\pi}_{i,t}$ and $\pi_{i,t}$. Bryan and Cecchetti (1993) using US CPI-U 12-month inflation data for February 1967 to December 1979 calculated RMSEs for forecasts beginning in January 1980 over future annual periods between one and five years. In all cases the weighted median had 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 was 63 percent of the mean price change in that forecast period, while that of the median was 57 percent and the trimmed mean 61 percent. The core measures were improvements on the CPI for predicting long-horizon CPI inflation, but not substantially so, especially given the error margins involved. As Cecchetti (1997) demonstrated, if the purpose of core inflation was to provide reliable predictors of inflation, they may provide more reliable ones than a CPI-U, but the extent of this left a lot to be desired.

129. Second, Laflèche (1997) argued that prediction from an auto-regressive (AR) model be used to judge which series are selected, i.e.:

$$\hat{\pi}_t = \alpha_i + \sum_{j=1}^{n} \beta_j \pi^*_{t-j} \qquad (16)$$

with the core inflation measures with the highest $\overline{R}^2$ selected. Marques and others (2000) note that $\overline{R}^2$ is a measure of their relative standing as opposed to their absolute standing and that it is the latter in which we are interested. Indeed $\overline{R}^2$ is only a measure of the residual sum of squares relative to a naïve model, where only the mean is eliminated, and this is unsuitable for time series data. An alternative might be relative to a random walk with drift and seasonal components (see Maddala, 1992, p. 550).

130. While the RMSE and MAD are well-used and acceptable criteria for evaluating forecasts, further insights into the suitability of alternative measures can be gleaned from the use of a wider range of criteria (see Clements and Hendry, 1993; and Diebold and Mariano, 1995). In favor of considering further criteria is that if the same "best" measures are found to arise from a number of criteria, then this gives 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. Against all of this needs to be a perception of transparency and objectivity in choice. Again, for core inflation measures used to operationalize targeting, this is less of a concern. But for the choice of, for example, which sector to exclude to derive an inflation target on the basis of it being one that predicts well, simple measures and rules of selection should be used for reasons of credibility.

131. Third, we noted that Marques and others (2000) considered it best to evaluate the predictions in terms of *absolute* rather than relative values. This can be undertaken by the construction of prediction intervals for the forecasts. Such intervals will depend on the fit (standard error) of the regression, the distance the future period is away from the data, and the variability in the series. Thus for a prediction from equation (15), the standard error, which can be used to generate, say, 95 percent confidence intervals for $\hat{\pi}_t$, is $\pm 1.96 SE(\hat{\pi}_t)$ where $SE(\hat{\pi}_t)$ is given by:

$$SE(\hat{\pi}_t) = \sqrt{\sigma^2 \left(1 + \frac{1}{n} + \frac{\left(\pi^*_{t-12} - \overline{\pi}^*_{t-12}\right)}{\sigma^2_{\pi^*_{t-12}}} \right)} \qquad (17)$$

and σ^2 is the standard error of the regression, $\left(\pi^*_{t-12} - \overline{\pi}_{t-12}\right)$ is the distance π^*_{t-12} is from the mean of variable π_{t-12} and $\sigma^2_{\pi^*_{t-12}}$ the variance of π^*_{t-12}. The prediction interval bounds on the prediction is determined by the fit of the regression model (better fit—smaller bounds), the sample size (larger sample size—smaller bounds), the distance the out-of-sample observation on the past inflation rate is away from its means (closer to the mean—smaller bounds) and the variability of price changes (larger dispersion—smaller bounds). For more than one lag, as in equation (16) the standard error of the prediction is:

$$SE(\hat{\pi}^*_{t-j}) = \sqrt{\sigma^2 \left(1 + \frac{1}{n}\right) + \sum_{j=1}^{k} \sum_{l=1}^{k} \left(z^t_j - \overline{z}^t_j\right)\left(z^t_l - \overline{z}^t_l\right) \mathrm{cov}\left(\hat{\beta}_j, \hat{\beta}_l\right)} \qquad (18)$$

132. Fourth, Jacobson and Karlsson (2004) proposed the use of a Bayesian average of forecasts of inflation and, while the work is based on 86 general economic indicators, the principles apply to using just core inflation measures. They used quarterly data for 1983:Q1 to 2000:Q3 for Sweden on 80 predictor variables and regressed inflation in period $t + j$ on each of the series in period t. They adopted a Bayesian treatment of model uncertainty and ranked the indicators in terms of their posterior probabilities. They found combining forecasts from the 10 highest ranked indicators produced forecasts with smaller RMSEs than the forecasts for the individual series. The method, in this context, also provides a basis for ranking the effectiveness of different core measures in terms of their forecasting accuracy. However, it should be noted that the exercise has to be repeated for different values of j. This may confirm the suitability of a measure over a range of forecast periods, or it may alternatively suggest different methods for different forecast horizons.

133. As will be discussed below, corroboration of the core measures need not just be by a ranking of such measures by their predictive power. A first hurdle should be a (Wald) statistical test to identify, for equation (16), whether the null hypothesis that the coefficients in $\beta_{i,j}$, are jointly zero, can be rejected. This might be considered as a necessary condition for evaluating the predictive power of different core measures using equation (18).

134. Fifth, an alternative formulation of the test is to consider it in terms of the predictive ability of the core inflation measure *over and above* that of the current *CP*I, using:

$$\pi_t = \alpha_0 + \sum_{j=1}^{n} \alpha_{1j} \pi_{t-j} + \sum_{j=1}^{n} \beta_{1j} \pi^*_{t-j} + \varepsilon_{1t} \qquad (19)$$

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) adopted this approach for the U.K. finding that her persistence-weighted core inflation measure, along with some exclusion measures, outperformed trimmed and weighted medians.

135. Sixth, 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 would want the prediction of the CPI from core inflation to dominate. Marques *et al*. (2003) for US data found, in testing Granger-causality, that a series that excluded F&E was 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, thus, are likely to affect final prices in a future period.

136. 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 of 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 *et al*. (2003)

argues that core inflation π_t^* should *not* be evaluated on the grounds that it is a good predictor of inflation, π_t . They note that:

> "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." Marques and others (2003: p. 768).

We turn to the tests they propose.

F. Judging on the Basis of Tests

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

Unbiasedness

Consider the following:

$$\pi_t = \pi_t^* + v_t \qquad (20)$$

where π_t is CPI inflation π_t^* is trend inflation and v_t a temporary disturbance in period *t.* First, that core inflation should be unbiased with respect to π_t, since when there are no shocks, $\pi_t = \pi_t^*$ in this model. A test of unbiasedness would be that jointly $\beta_0 = 0$ and $\beta_1 = 1$ in the estimated equation:

$$\pi_t = \beta_0 + \beta_1 \pi_t^* + v_t \qquad (21)$$

Granger-causality

138. Secondly, that π_t^* should Granger-cause π_t , that the measure of core inflation, π_t^* should better predict the CPI, π_t , than the CPI, would itself. The test requires ordinary least squares (OLS) (assuming stationary) estimates of :

$$\pi_t = \alpha_{10} + \sum_{j=1}^{n} \alpha_{1j} \pi_{t-j} + \sum_{j=1}^{n} \beta_{1j} \pi_{t-j}^* + \varepsilon_{1t} \qquad \text{and} \qquad (22a)$$

$$\pi_t^* = \alpha_{20} + \sum_{j=1}^{n} \alpha_{2j} \pi_{t-j} + \sum_{j=1}^{n} \beta_{2j} \pi_{t-j}^* + \varepsilon_{2t} \qquad (22b)$$

and Wald tests for the joint hypothesis that $\beta_{1j} = 0$ and $\alpha_{2j} = 0$ for all *j*. The terms of the *n* lags can be determined by statistical criteria such as the Schwarz-Bayesian criterion. As noted above,

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.

139. An important finding on the choice of method on this basis is that it may vary with the time period chosen. For example, Heath and others (2004) applied the Granger-causality tests to 102 measure 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 sub-period 1993:Q1 to 2003:Q4 *none o*f the measures passed the tests. The authors note that this period followed the implementation of an inflation-targeting regime (July 2000) and inflation had been comparatively stable.

Cointegration-based tests

140. Cointegrating regressions are widely used in economics to capture long-run equilibrium relationships. There will of course be short-run dynamics in a relationship, but this can be captured by what is called an error-correction model. An appropriate measure of core inflation is argued to be one that does not have long-run divergences from the CPI. To understand this approach it is necessary to consider a feature of time series. A time series π_t is said to be integrated of order 1, I(1), if its first-difference, $\Delta\pi_t$, is a stationary series, I(0). If *both* the CPI and core inflation are found to be I(1)[20], then they are said to be cointegrated if there exists a β such that for the regression $\pi_t = \beta\pi_t^* + \nu$, it is found that $\left(\pi_t - \beta\pi_t^*\right)$ is I(0)—that is, they will not drift far apart over time. If they are not cointegrated, they can drift far apart; any regression relationship between the CPI and core inflation might be spurious. Marques and others(2000 and 2003), have proposed that an appropriate test would be that the CPI should respond to deviations from the cointegrating relationship.[21] While the details of cointegration tests are outside the scope of this paper, they are readily available in introductory econometrics texts such as Maddala (1988) and can be implemented in standard econometric software.

141. Briefly, when inflation π_t is I(1), then π_t^* is a measure of core inflation if:

(i) π_t^* is I(1) and π_t and π_t^* are cointegrated with unitary coefficient, β , i.e. $\left(\pi_t - \pi_t^*\right)$ is a stationary variable with mean zero, i.e.: $\delta_t = 0$ in $\left(\pi_t - \pi_t^*\right) = \delta_t + \nu_t$;

(ii) there is an error correction mechanism (ECM) given by $z_{t-1} = \left(\pi_{t-1} - \pi_{t-1}^*\right)$ for $\Delta\pi_t$, i.e. $\Delta\pi_t$ may be written as:

$$\Delta\pi_t = \sum_{j=1}^{m}\delta_j\Delta\pi_{t-j} + \sum_{k=1}^{n}\gamma_j\Delta\pi_{t-k}^* - \lambda\left(\pi_{t-1} - \pi_{t-1}^*\right) + \nu_t \qquad (23)$$

[20] This is an empirical matter resolved through unit root tests, though in much of the literature inflation has been found to be I(1). If the inflation rate is found to be stationary, then Marques and others (2000) Appendix A considers alternative tests.

[21] They acknowledge Freeman (1998) as first proposing the basic idea.

(iii) π_t^* is strongly exogenous for the parameters in the above error-correction equation.

142. The conditions require that core inflation and measured inflation cannot exhibit systematically diverging long-run trends. Note that the error correction applies to $\Delta\pi_t$ on the left-hand-side of the above equation and not to $\Delta\pi_t^*$, the argument being that in the long run $\Delta\pi_t$ must converge to $\Delta\pi_t^*$—short-run adjustments are to inflation to allow it to converge to core inflation, not visa versa. Consider the case where π_t is above π_t^* in a period. Because of the error correction, π_t will converge to π_t^*. This includes the requirement of equation (22a) above in which π_t^* Granger-causes π_t.

143. The phrasing of equation (23) with $\Delta\pi_t$ on the left-hand-side, as opposed to $\Delta\pi_t^*$, is important. If $\Delta\pi_t^*$ were used, it would imply that in the long run π_t^* must converge to π_t. The exogeneity condition in (iii) above requires that π_t does not Granger-cause π_t^*—there is no error correction equation to determine $\Delta\pi_t^*$ since this is only determined by its own past values, i.e., the error correction model for π_t^* is:

$$\Delta\pi_t^* = \sum_{j=1}^{m}\theta_j\Delta\pi_{t-j}^* + \sum_{k=1}^{n}\vartheta_j\Delta\pi_{t-k} - \varsigma\left(\pi_{t-1} - \pi_{t-1}^*\right) + \eta_t \qquad (24)$$

where $\varsigma = \vartheta_1 = \vartheta_2 = ... = \vartheta_n = 0$, so that:

$$\Delta\pi_t^* = \sum_{j=1}^{m}\theta_j\Delta\pi_{t-j}^* + \eta_t\,. \qquad (25)$$

144. Estimation and testing are relatively simple in modern econometric software. The approach considered by Marques and others (2003) is to: (i) test for unit root tests on the series $\left(\pi_t - \pi_t^*\right)$ and that $\delta_t = 0$; (ii) test the null hypothesis that $\lambda = 0$ using a *t*-test; and (iii) test the null hypothesis that $\varsigma = \vartheta_1 = \vartheta_2 = ... = \vartheta_n = 0$.

145. Marques and others (2000), for Portugal, compared eight measures: a 37-month, 25-month and 13-month moving average, 10 percent and 25 percent trimmed means, underlying inflation (excluding F&E), first principal component, and a standard deviation-weighted CPI. All passed a test to establish whether they were AR(1), as was inflation, however, the 10 percent and 25 percent trimmed means failed the test of $\beta_0 = 0$ (given a unit root, $\beta_1 = 1$). Condition (ii), the short-run error correction was passed by just about all measures. For two measures, the trimmed mean and the underlying inflation index, the core measure led inflation, as opposed to the requirement that it is lagged indicator. They found only the 37-month moving average, the first principle component, and the standard deviation weighted CPI passed all tests.

146. Trimmed means met all conditions except the first. Therefore, Marques and others (2000) advocated asymmetric trimming. They argued against moving averages because the 13-month and 25-month averages did not satisfy the test conditions postulated. However, the 36-month did

pass these tests at the standard 5 percent level and was therefore not to be dismissed on these grounds.

147. Yet the results differed in a subsequent study for the U.S. Marques and others (2003) used U.S. data for the period of January 1983 to December 2000, on three measures of core inflation—the CPI excluding F&E, a trimmed mean excluding eight percent on either side, and the weighted median. They employed cointegration tests and found that the measure which excluded F&E did not meet the test criteria. They found that in the long run π_t^* converged to π_t —the causality ran the wrong way in that core inflation led rather than lagged. Again, this may be because food and energy to a large extent enter as intermediate inputs into the production process. The trimmed mean and weighted median met the test requirements.

148. Mankikar and Paisley (2004) adopted cointegration tests for a range of U.K. core inflation measures. They tested six exclusion indexes, a persistence weighted index, three domestically generated inflation indexes, the Quah and Vahey index, a trimmed mean and weighted median. Only three indexes met all conditions: two exclusion ones and a DGI (using ULC) index. While all series proved to be I(1), only these three had a zero mean, though the persistence weighted, Quah and Vahey, and one exclusion index satisfied the causality conditions. The test approach requires some comment.

149. First are the causality conditions. Mankikar and Paisley (2004) note that if the monetary authorities target CPI inflation successfully, aside from unforeseeable noise, then their use of core inflation in say $t - j$ determines the CPI in t. Core inflation is attracted to CPI inflation since if core inflation is, for example, above the target CPI in one period, mechanisms are put into play to bring it down towards the target in the next. This is quite different from the conceptualization of the tests (ii) and (iii) which require CPI to be attracted to core inflation. Such institutional influences are over and above the Lucas critique, whereby if policy were based on a relationship between core inflation and the CPI, that very relationship would change as a result of the realization of the policy (see Section IV-D above).

150. Second, the information derived in determining cointegrating relationships should, at first sight, also prove useful for predictions, rather than just tests. It might be expected that since ECM models incorporate both short-term fluctuations and deviations from the cointegrated, long-run equilibrium path in the forecasting model, they should prove more accurate than say Box-Jenkins ARIMA models. This gives a sense of validity to the use of such tests. However, Christoffersen and Diebold (1997) challenged the belief that the imposition of cointegrating restrictions produced superior long-horizon forecasts. They showed that for the long-horizon, using MSE criterion, even *univariate* ARIMA forecasts are equally accurate. They argued that, in part, the problem lies with the use of the MSE to evaluate forecasts.[22] It is the imposition of integration conditions that is helpful, something that ARIMA models achieve through differencing.

151. Third, Heath and others (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

[22] They consider the trace MSE *difference*, as opposed, to the trace MSE *ratio*—see also McCrae (2002).

of method is less important. Indeed over a period when there was a fall in volatility and increased stability of the CPI, they found a constant inflation rate was a better predictor of future CPI inflation than measures of core inflation, though they recognized the usefulness of core inflation measures for analytical purposes. It follows that lessons learnt from empirical studies using high-inflation volatile data, should not be carried over to choice of method for low inflation periods (for similar findings see de Brouwer, 2004.)

152. 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 a core inflation measures should satisfy and are satisfactory in this sense. Yet they also suffer from type I error. The non-rejection of the hypotheses is taken to be a failure of the test. However, it may arise from inadequacies of the model or data, leading to poor power efficiency. The tests have in mind particular concepts of inflation. Equations (20) and (21), for example, implicitly identify the error structure to be normal when empirical studies and theoretical frameworks identified in Section IV(C) above (under "Asymmetric and variable trimmed measures") have argued that part of the signal may be skewed. Furthermore, the concept of equilibrium associated with the cointegration test may be of a time span outside of that of use to monetary authorities for policy measures.

G. Judging on the Basis of Correlation with Money Supply

153. Bryan and Cecchetti (1993) considered a primary motivation of their study to be to find a measure of core inflation correlated with monetary growth. They tested measures of core inflation in terms of the ability of monetary growth to forecast the core inflation measures. While some of the results were mixed, depending on the time horizons and the measure of money supply used, the weighted median was consistently better correlated than the 15 percent trimmed mean, and for measures of money supply M1 and M2, the weighted median was 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 found changes in the money supply to Granger-cause core inflation as measured by the weighted median and 15 percent trimmed mean.

154. In models in which inflation is a purely monetary phenomena and prices are fully flexible, shocks, such as changes in oil prices, tastes, technology, are instantaneously accommodated and aggregate price inflation is unchanged. But, as outlined in Section IV(C) above, prices are not fully flexible and the CPI is not a satisfactory realization of any such process, due 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).

VI. Concluding remarks

155. The paper has critically outlined the many approaches and methods to the measurement of core inflation and, moreover, the many approaches to judging the preferred measure(s). It is not a straightforward matter and the empirical research shows that different measures of core

inflation yield different results, that is, that choice of measure matters. Further, that 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 county for different time periods. Choice of measure should thus, in principle, be data-driven for each country based on appropriate criteria selected from Section V. An understanding of the features of the methods and the alternative criteria for choice is the necessary groundwork for the choice of core inflation measure(s) and contributing to this groundwork is the purpose of this paper.

156. What is apparent, however, is that a consensus has emerged and, for reasons of maintaining credibility, there is for many countries a natural starting point. 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 products groups excluded and such exclusions can thus be justified as not manipulating the figures, since their use is widespread. There may, however, be great public sensitivity to the exclusion of items, such as food and energy, noted, for example, for South Africa in Section IV(A). With greater confidence in the ability of the authorities to manipulate the measure without losing credibility, the exclusions can be data-driven depending on the features of the country's data, or alternative methods adopted. The focus on credibility derives from the primary purpose of targeting as one of anchoring inflation expectations as discussed in Section II and IV(A). 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.

157. 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 IV and V above. Yet these Sections come into their own in two respects. The first is where the CPI or an exclusion-based CPI is adopted as a credible basis upon which to anchor inflation, and it is necessary for the central bank to have further measures to operationalize the targeting framework. An examination of the country's past data, using methods discussed in Section V(A), may show other measures to better reflect the smooth pattern of the economy, forward information, possibly conditioned on other economic variables using economic models, and so forth. The selection of such measures should be based on criteria chosen from Section V that are meaningful to the monetary authority and its inflation targeting framework.

158. Second, it may be that exclusion-based methods are found to be not optimal according to the criteria selected by the monetary authorities, and that the credibility trade-off has little resonance for the circumstances of the country. In such case the monetary authorities are well placed to select among the methods and the criteria for choice considered in this paper to adapt their core inflation measure to their needs.

159. 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 so that the choice of measures can be justified on an objective, transparent basis. Research would be 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, *National Bank of Belgium Working Paper*, No. 2.

Bagliano, Fabio C., Golinelli, Roberto, and Claudio Morana, 2002, Core Inflation in the Euro Area, *Applied Economic Letters*, 9, 353-357.

Bakhsi, H., Yates, T., 1999, To Trim or Not to Trim, An Application of a Trimmed Mean Inflation Estimator to the United Kingdom, *Bank of England, Working Paper Series* No. 97. London, Bank of England.

Balke N.S. and Wynne, M.A., 1996, Supply shocks and the Distribution of Price Changes, *Federal Reserve Bank of Dallas, Economic Review,* First Quarter, 10-18.

Ball, L. and Mankiw, N.G., 1995, Relative Price Changes as Aggregate Supply Shocks, *The Quarterly Journal of Economics*, February, 161-193.

Bernanke, Ben S., Laubach, Thomas, Mishkin, Frederic, S., and Adam S. Posen, 1999, *Inflation Targeting: Lessons from International Experience*, N.J.: Princeton University Press.

Blejer, Mario I., Leone, Alfredo M., Ruband, Pau and Gerd Schwartz, 1999, Inflation Targeting in the Context of IMF-Supported Adjustment Program, pp. 409-499. In Loayza, Norman and Raimundo Sato (eds.), *op. cit*.

Blejer, Mario I., Ize, Alain, Leone, Alfredo M., and Sergio Werdang, 2000, I*nflation Targeting in Practice: Strategic and Operational Issues and Applications to Emerging Market Economies*, Washington D.C.: IMF.

Blinder, Alan S. 1997. Commentary on 'Stephen G. Cecchetti (1997)' *op. cit., Federal Reserve Bank of St. Louis Review*, May/June. 157-160.

Bloem, Adriaan, Armknecht, Paul, A. and Kimberly D. Zieschang, 2002, Price Indexes for Inflation Targeting. 172-198. In Carson and others (2002) *op. cit.*

Brouwer, de, Gordon., 2004, Discussant on Heath and others (2004), *op. cit.*

Bryan, Michael F. and Cecchetti, Stephen G., 1993, The Consumer Price Index as a Measure of Inflation', *Federal Reserve Bank of Cleveland Economic Review*, No. 4, pages 15–24.

Bryan, Michael F. and Cecchetti, Stephen G., 1994, Measuring Core Inflation. In Gregory Mankiw, N. (ed.), *Monetary Policy*. University of Chicago Press for NBER, pp. 195-215.

Bryan, Michael F. and Cecchetti, Stephen G. (1996), Inflation and the Distribution of Price Changes, *National Bureau of Economic Research Paper 5793,* Cambridge, Mass.: NBER.

Bryan, Michael F., Cecchetti, Stephen G. and Rodney L. Wiggins II, 1997, Efficient Inflation Estimation, *NBER Working Paper* No. 6183.

Carson, Carol, S., Enoch, Charles and Dziobek, Claudia (eds.)., 2002, S*tatistical Implications of Inflation Targeting: Getting the Right Numbers and Getting the Numbers Right,* Washington, D.C.: Statistics Department, International Monetary Fund.

Cecchetti, Stephen G., 1997, Measuring Short-Run Inflation for Central Bankers, *Review of the Federal Reserve Bank of St. Louis*, May/June, 79, 3, 143-155.

Cerisola, Martin and R. Gaston Gelos, 2005, What Drives Inflation Expectations in Brazil? An Empirical Analysis, *IMF Working Paper Series*, WP/05/109, June 2005.

Chatfield, C., Yar, M., 1991, Prediction Intervals for Multiplicative Holt–Winters. *International Journal of Forecasting,* 7, 31–37.

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

Clemen R.T., 1989, Combining Forecasts: A Review and an Annotated Bibliography, *International Journal of Forecasting* 5: 559–583.

Clements, M.P. and Hendry, D.F., 1993, On the Limitations of Comparing Mean Square Forecast Errors, *Journal of Forecasting*, 12, 617-637.

Coimbra, Carlos and Pedro Duarte Neves, 1997, Trend Inflation Indicators. Banco de Portugal, *Economic Bulletin*, 3, 1, March.

Cutler, Joanne, 2001, A New Measure of Core Inflation in the U.K., External MPC Unit Discussion Paper, 3, March.

Diebold, F.X. and Mariano, R.S., 1995, Comparing Predictive Accuracy, Journal of Business and Economic Statistics, 13, 253-265.

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

Diewert, W. Erwin, 2004a, Basic Index Number Theory, Chapter 15, 263-287. In ILO and others (2004) op cit.

Diewert, W. Erwin, 2004b, Elementary Indices, Chapter 20, 355-370. In ILO and others (2004) *op cit.*

Eckstein, O., 1981, *Core Inflation*, NJ: Prentice Hall.

Fenwick, David, 2004, Core Inflation, Seasonal Adjustment and Measures of the Underlying Trend, Statistical Journal of the United Nations ECE, 21, 115-24.

Folkertsma, C. K. and Hubrich, K. S. E. M., 2000, Performance of Core Inflation Measures, De Nederlandsche Bank NV, Research Memorandum WO&E No. 639/0034.

Freeman, D.G., 1998, Do Core Inflation Measures Help Forecast Inflation? Economics Letters, 58, 143-147.

García, Pablo, 2002, Design, Measurement, Communication: Chile's Experience with Inflation Targeting, 157-171. In Carson and others (2002) op. cit.

Gaspar, Vítor, 2002, Eurostat's HICP and the European Central Banks' Definition of Price Stability. 137-156. In Carson and others (2002) *op. cit.*

Greenlees, John and Bert Balk (2004) Errors and Bias, Chapter 11, 207-214. In ILO and others (2004) op cit.

Heath, Alexandra, Roberts, Ivan and Tim Bulman. 2004. Inflation in Australia: Measurement and Modeling, In The Future of Inflation Targeting ed. Christopher Kent and Simon Guttmann,167-207 Reserve Bank of Australia. Available at: 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* Christopher Kent and Simon Guttmann, (ed.), 140-60, August 2004 Reserve Bank of Australia. Available at: http://www.rba.gov.au/PublicationsAndResearch/Conferences/2004/Hill.pdf.

Harter, H.L., 1974 and 1975, The Method of Least Squares and Some Alternatives, *International Statistical Review*, 1974: Part I, 42, 2, 147-74; Part II, 42, 3, 235-64; 1975: Part III, 43, 1, 1-44; Part IV, 43, 2, 125-190; Part V, 43, 3, 269-78.

Hogan, Seamus, Johnson, Marianne and Thérèse Laflèche, 2001, Core Inflation, *Bank of Canada, Technical Report* No. 89, January.

Hogg, R., 1967, Some Observations on Robust Estimation, *Journal of the American Statistical Association*, 62, December, 1179-86.

Huber, P., 1964, Robust Estimation of a Location Parameter, *Annals of Mathematical Sciences*, 35, 73-101.

Hyndman, Rob J., Koehler, Anne B., Ord, J. Keith and Ralph D. Snyder, 2005, Prediction Intervals for Exponential Smoothing Using Two New Classes of State Space Models, Journal of Forecasting. 24, 17-37.

ILO/IMF/OECD/Eurostat/UN/World Bank, ILO/IMF/OECD/Eurostat/UN/World Bank. 2004. Consumer Price Index Manual: Theory and Practice, Geneva: International Labour Office, 2004. Available at: www.ilo.org/public/english/bureau/stat/guides/cpi/index.htm.

Jacobson, Tor and Karlsson, Sune, 2004, Finding Good Predictors for Inflation: a Bayesian Model Averaging Approach, Journal of Forecasting, 23, 479-496.

José, Ramos Maria, 2004, On the Use of the First Principal Component as a Core Inflation Indicator, Banco de Portugal Working Paper, WP 3-04 January 2004 .

José, Armids san, Slack, Graham L. and Subramanian S Sriram, 2002, Statistical Principles for Inflation Targeting Regimes and the Role of IMF Data Initiatives. 308-339. In Carson and others (2002) op. cit.

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

Lach, Saul., 2002, Existence and Persistence of Price Dispersion, Review of Economics and Statistics, 84, 3, August, 433-444.

Laflèche, T., 1997, Statistical Measures of the Trend Rate of Inflation. Bank of Canada Review, Autumn.

Lehohla, Pali J. and Annette Myburgh, 2002, Statistical Implications of Inflation Targeting in South Africa, 55-75. In Carson and others (2002) op. cit.

Loayza, Norman and Raimundo Sato (eds.), Inflation Targeting: Design, Performance, Challenges, Santiago: Central Bank of Chile.

Lucas, Robert E. (1976). Econometric Policy Evaluation: A Critique. Carnegie-Rochester Conference Series on Public Policy 1, 19-46.

Machado, José Ferreira, Marques, Carlos Robalo, Neves, Pedro Duarte and Afonso Gonçalves da Silva,2001, Using the First Principal Component as a Core Inflation Indicator, Banco de Portugal Working Paper, WP 9-01 September 2001.

Macklem, Tiff, 2001, A New Measure of Core Inflation, Bank of Canada Review, Autumn, 3-12.

Maddala, G. S. Introduction to Econometrics, New York: Macmillan Publishing Co.

Mankikar, Alan and Jo Paisley, 2004, Core Inflation: A Critical Guide. Bank of England Working Paper no. 242, 1-36; Summary in Bank of England. Quarterly Bulletin, Winter, 44,4, 466.

Mankiw, N. Gregory and Ricardo Reis, 2003, What Measures of Inflation Should a Central Bank Target? Journal of the European Economic Association, 1, 5, September, 1058-86.

Marques, C. R., Neves, P D and Sarmento, L. M., 2000, Evaluating Core Inflation Measures, Banco de Portugal Working Paper, No. 3-00.

McCrae, Michael, Lin, Chael, Lin Yan-Xia, Pavlik, Daniel and Chandra M. Gulati, 2002, Can Cointegration-Based Forecasting Outperform Uunivariate Models? An Application to Asian Exchange Rates, Journal of Forecasting 21, 355-380.

Parkin, Michael, 1984, On Core Inflation by Otto Eckstein: A Review Essay, Journal of Monetary Economics, 14, 2, 251-64.

Porrado, Eric and Andrés Velasco, 1999, Alternative Monetary Rules in an Open Economy: A Welfare-Based Approach. In Norman Loayza and Raimundo Sato (eds.) op. cit., 294-348.

Quah, Danny and Shaun P. Vahey, 1995, Measuring Core Inflation, Economic Journal, 105, September, 1130-44.

Roger, Scott, 1997. A Robust Measure of Core Inflation in New Zealand, Reserve Bank of New Zealand. Discussion Paper G97/7 Wellington: New Zealand.

Roger, Scott, 1998, Core Inflation: Concepts,Uses and Measurement. Reserve Bank of New Zealand. Discussion paper G98/9 Wellington: New Zealand.

Roger, Scott, 2000, Relative Prices, Inflation and Core Inflation, International Monetary Fund (IMF) Working Paper WP/00/58, Washington D.C.: IMF.

Rowlatt, Amanda, 2001, The U.K. Office for National Statistics and the Inflation Target, I, 577, December, reprinted in 125-136, Carson and others (2002) op. cit.

Rudebusch, Glenn D., 2005, Assessing the Lucas Critique in Monetary Policy Models Journal of Money, Credit, and Banking, 37, 2, April, 245-72.

Schaechter, Andrea, Stone, Mark, R. and Mark Zelmer, 2000, Adopting Inflation Targeting: Practical Issues for Emerging Market Countries, International Monetary Fund (IMF), Occasional Paper, 202, Washington D.C.: IMF.

Silver, Mick and Christos Ioannidis, 1996, Inflation, Relative Prices and their Skewness, Applied Economics, 28, 577-584.

Stigler, G.J., 1961, The Economics of Information, Journal of Political Economy, 69, June, 213-225.

Sorensen, Alan T., 2000, Equilibrium Price Dispersion in Retail Markets for Prescription Drugs, Journal of Political Economy, 108, 4, 833-850.

Van Hoomisson, T., 1988, Price Dispersion and Inflation: Evidence from Israel, Journal of Political Economy, 96,6, 1303-1314.

Vega, Juan-Luis and Mark A. Wynne, 2003, A First Assessment of Some Measures of Core Inflation for the Euro Area, German Economic Review, 4, 3, 269-306.

Wynne, Mark A., 1999, Core Inflation: A Review of Some Conceptual Issues', ECB Working Paper Series, No. 5.

Yule, U., 1911, An Introduction to the Theory of Statistics, (11th Edition 1937, with Maurice Kendall), London: Charles Griffen.
