StaffPAPERS *is published by the Federal Reserve Bank of Dallas. The views expressed are those of the authors and should not be attributed to the Federal Reserve Bank of Dallas or the Federal Reserve System. Articles may be reprinted if the source is credited and the Federal Reserve Bank of Dallas is provided a copy of the publication or a URL of the website with the material. For permission to reprint or post an article, e-mail the Public Affairs Department at dal.webmaster@dal.frb.org.*

Staff Papers *is available free of charge by writing the Public Affairs Department, Federal Reserve Bank of Dallas, P.O. Box 655906, Dallas, TX 75265-5906; by fax at 214-922-5268; or by phone at 214-922-5254. This publication is available on the Dallas Fed website, www.dallasfed.org.*

No. 4, May 2008

StaffPAPERS Federal Reserve Bank of Dallas

Measuring Core Inflation: Notes from a 2007 Dallas Fed Conference

Jim Dolmas
Senior Economist and Policy Advisor

and

Mark A. Wynne
Vice President and Senior Economist
Director, Globalization and Monetary Policy Institute
Federal Reserve Bank of Dallas

Abstract

In May 2007, the Federal Reserve Bank of Dallas hosted a conference, organized with the Federal Reserve Bank of Cleveland, titled "Price Measurement for Monetary Policy." The conference broadly focused on two issues—the measurement of core inflation and the measurement of inflation expectations. This paper summarizes the conference papers on core inflation.

JEL codes: C43, E31
Keywords: core inflation, price index

The measurement of inflation is critically important to central banks concerned with achieving price stability. While policymakers invariably look at a broad range of inflation indicators, many central banks have assigned a prominent role to core inflation measures. The Federal Reserve, for example, includes a core measure—the price index for personal consumption expenditures (PCE) excluding food and energy—in the inflation forecasts it presents twice a year to Congress.[1] The inflation-targeting Bank of Canada expresses its target in terms of a headline measure but gives a core measure—the consumer price index (CPI) excluding the eight most volatile items—a prominent role as an "operational guide." Similarly, the Sveriges Riksbank expresses its inflation goals in terms of a headline CPI inflation rate while referring also to its "CPIX"—CPI excluding mortgage interest costs and the effects of indirect taxes—or to CPIX excluding energy in explaining the stance of monetary policy in public communications.

"Inflation ex…" measures—like those used by the Federal Reserve, the Bank of Canada, and the Riksbank—are, however, just a small subset of the many core inflation gauges being produced, studied, and debated by economists at central banks and in academia. The Reserve Bank of Australia, for example, expresses its target inflation rate in terms of a headline index but presents several core measures—which it describes as underlying measures of inflation—in its quarterly monetary policy statements. Since November 2004, these underlying measures have consisted of a CPI excluding volatile items, a weighted median CPI, and a trimmed mean CPI. The Reserve Bank of New Zealand follows a similar tack. Its most recent Monetary Policy Statement presents data on CPI inflation—the reserve bank's target measure—plus seven alternative measures, including a trimmed mean and median, a dynamic factor model estimate, an exponentially smoothed CPI, and two "inflation ex…" measures.

Issues of inflation measurement—including the measurement of core inflation—were the focus of a research conference held May 24–25, 2007, at the Federal Reserve Bank of Dallas. The conference, organized by the Federal Reserve Banks of Cleveland and Dallas, brought together economists from several central banks and academia. In this paper, we summarize the conference presentations on core inflation.[2] These papers, most of which either propose new core measures or evaluate existing ones, represent a small but, we believe, informative sample of work being done at the frontier of core inflation measurement.

1. CORE INFLATION: SOME BACKGROUND

Before discussing the research presented at the conference, we think it's useful to have in hand some background on the measurement of core

---

[1] From 2004 through 2007, the inflation forecasts in the semiannual Monetary Policy Report to the Congress were cast solely in terms of PCE ex food and energy. Since February 2008, the Fed's forecasts have included headline PCE as well.

[2] Another theme of the conference was the measurement of inflation expectations.

inflation, in terms of both theory and practice. We are not attempting an exhaustive survey; this literature has been ably surveyed elsewhere—Silver (2007) and Wynne (1999), for example. What we do hope to offer is some context for the work presented at the conference.

The term *core inflation* was coined by Otto Eckstein in 1980, though the concept to which Eckstein affixed that name differs considerably from current usage.[3] Eckstein posited a tripartite decomposition of actual inflation as the sum of a "core" component, relating to steady-state factor price growth; a "shock" component, capturing the effects of exogenous changes in food and energy prices, taxes, the minimum wage, and so on; and a "demand" component, capturing the effects of deviations of unemployment from its natural rate.

Blinder (1982) was one of the first academic papers to consider the now-conventional notion of core inflation as an "inflation ex…" measure—in Blinder's case, ex food, energy, and mortgage interest costs.[4] The practice of reporting "inflation ex…" measures, though, long predates academic research on such indexes. By the time of these early analyses by Eckstein and Blinder, the U.S. Bureau of Labor Statistics had been reporting a CPI ex food and a CPI ex shelter for at least thirty years and had recently begun reporting a CPI ex food and energy.[5]

From the early 1980s to the early 1990s, there was little in the way of scholarly research on core inflation. Beginning in the early 1990s, however, the field saw something of a renaissance, resulting in the development of several important competing measures of core inflation. One of the earliest papers in this new wave of research was Bryan and Pike (1991), which proposed using median consumer price changes as a measure of underlying inflation. Bryan and Cecchetti (1994) elaborated on this approach, providing a more complete theoretical and statistical justification for using limited-influence estimators—like the median or trimmed mean—as measures of the central tendency of the cross-sectional distribution of monthly price changes in the components of the CPI.

The statistical underpinnings of the median and trimmed mean approaches to measuring core inflation derive from the much earlier literature in statistics on robust measures of location. It is well known from this literature that compared with the median or a trimmed mean, the sample mean is a relatively inefficient measure of a distribution's location when the sample is drawn from a distribution with fat tails. In U.S. data, the distributions of monthly price changes for the components of both the CPI

---

[3] Eckstein's concept of core inflation was originally laid out in a 1980 study prepared for the Joint Economic Committee of the U.S. Congress. He further elaborated on the concept in his monograph *Core Inflation* (Eckstein 1981).

[4] At the time of Blinder's analysis, the BLS still used a "payments" approach to measuring the cost of owner-occupied housing in the CPI—i.e., the cost of owner-occupied housing was measured by summing all out-of-pocket expenses associated with owning and maintaining a home, including mortgage interest payments.

[5] We base this claim on the first appearances of these series in the statistical tables of the *Economic Report of the President*. CPI ex food and CPI ex shelter first appear in the 1958 *ERP*, while CPI ex food and energy first appears in the 1980 *ERP*.

and PCE are characterized by large degrees of positive kurtosis.[6]

The median and trimmed mean are examples of core measures that make exclusions based on the cross-sectional distribution of monthly price changes. In contrast, one can think of "inflation ex…" measures like CPI ex food and energy as making exclusions based on the time-series properties of individual components. To be concrete, and establish concepts to which we refer below, we can imagine (to an approximation) that the rate of inflation in a given month (π_t), for an index like the CPI or PCE, is a weighted average of one-month price changes ($\pi_{i,t}$) of a set of N components:

$$\pi_t = \sum_{i=1}^{N} w_{i,t} \pi_{i,t} . \tag{1}$$

Typically, the $w_{i,t}$ weights in commonly used price indexes like the CPI and PCE reflect the importance of each component i in household expenditure.[7] An "inflation ex…" measure excludes the same subset of components each period—all $i \in X$, say, where X might be the set of indexes corresponding to food and energy items. While many criteria might be used to exclude certain items, the justification typically given for excluding items such as food and energy is based on their time-series volatility. The Bank of Canada and the Reserve Bank of Australia, for example, describe the construction of their "inflation ex…" measures in these terms. This gives

$$\pi_t^X = \sum_{i \in I} w_{i,t}^X \pi_{i,t} , \tag{2}$$

where $I = \{1,2,\ldots,N\} \backslash X$ and $w_{i,t}^X = w_{i,t} / \Sigma_{j \in I} w_{j,t}$.

In contrast, the components excluded from the median and trimmed mean differ from period to period, with the exclusions in a given period based on properties of that period's cross-sectional distribution. To describe a trimmed mean inflation rate, assume that a given month's component price changes (and their corresponding weights) are ordered so that

$$\pi_{1,t} \leq \pi_{2,t} \leq \cdots \leq \pi_{N,t} .$$

Then, a trimmed mean inflation rate can be written as

---

[6] See Bryan, Cecchetti, and Wiggins (1997) for facts on the distribution of component price changes for the CPI. Dolmas (2005) discusses evidence of excess kurtosis in PCE data.

[7] The weights may vary each period or less frequently, depending on the underlying price index. Note that for the PCE price index, which is a chain aggregate of its component price indexes, an accurate approximation of π_t as a weighted sum of the component $\pi_{i,t}$'s requires weights that reflect the components' importance in expenditure in both t and $t-1$. The Tornqvist formula is one example of such an approximation; another is given in Dolmas (2005).

$$(3) \qquad \pi_t^{TM} = \sum_{i=l_t}^{u_t} w_{i,t}^{TM} \pi_{i,t} ,$$

where l_t and u_t are such that the total weights of the components excluded from the lower and upper tails, $\sum_{i=1}^{l_t - 1} w_{i,t}$ and $\sum_{i=u_{t+1}}^{N} w_{i,t}$, are particular fractions, α and β, say.[8] The weights $w_{i,t}^{TM}$, of course, are such that $w_{i,t}^{TM} = w_{i,t} / \sum_{j=l_t}^{u_t} w_{j,t}$. The median inflation rate is simply an extreme version of a trimmed mean, in which $\alpha = \beta = ½$.[9]

Another notable approach to measuring core inflation developed in the early 1990s is the dynamic factor model of Bryan and Cecchetti (1993). This approach—like the "inflation ex…" and trimming approaches—takes the component price changes $\pi_{i,t}$ as inputs but draws on information concerning both the time-series and cross-sectional properties of these price changes. In this approach to measuring core inflation, each $\pi_{i,t}$ is viewed as the sum of a common inflation factor and an idiosyncratic relative price movement:

$$(4) \qquad \pi_{i,t} = \pi_t^* + x_{i,t} .$$

Given that $\Sigma_i w_{i,t} = 1$ in equation 1, under the assumption embodied in equation 4 a fixed-weight index such as the CPI—the focus of Bryan and Cecchetti's analysis—gives measured inflation as

$$\pi_t = \pi_t^* + \sum_{i=1}^{N} w_{i,0} x_{i,t} ,$$

where $w_{i,0}$ are base-period expenditure shares.[10] Drawing on basic consumer theory, Bryan and Cecchetti argue that one should not expect $E(\Sigma w_{i,0} x_{i,t}) = 0$, implying that measured inflation π_t will be a biased estimator of the common factor π_t^*. Positing linear laws of motion for π_t^* and the $x_{i,t}$'s, Bryan and Cecchetti apply Kalman filter methods to obtain an estimate of the common inflation factor, π_t^*.

Like the trimmed mean, median, and "inflation ex…" measures, the dynamic factor model reweights and aggregates the individual component price changes $\pi_{i,t}$, though with two major differences: The weights incorporate both time-series and cross-sectional information on the $\pi_{i,t}$'s, and the aggregate includes current and past values of the $\pi_{i,t}$'s. The resulting core measure—call it π_t^{DF}—has the form

---

[8] In practice, these fractions are typically constant over time. The trimmed means produced by the Reserve Bank of Australia and the Federal Reserve Banks of Cleveland and Dallas all have this feature. However, if we follow the statistical justification for trimming to its logical conclusion—viewing trimming as a robust estimation method, given a fat-tailed distribution—then the trimming proportions ought to depend on the degree of excess kurtosis in the monthly distributions of price changes.

[9] To keep the notation simple and not lose the broad concepts in a thicket of details, we're ignoring certain technical details arising from the fact that the distributions of component price changes are lumpy. In practice, for example, there is no item that precisely splits the distribution into halves.

[10] At the time Bryan and Cecchetti (1993) was written, CPI weights were updated every ten years; since then, the frequency has risen to every two years.

$$(5) \qquad \pi_t^{DF} = \sum_i \sum_j w_{i,t-j}^{DF} \pi_{i,t-j} .$$

Despite significant differences, the “inflation ex…” measures, robust measures like the trimmed mean and median, and the dynamic factor model share a common set of basic building blocks—the component price changes underlying the headline index—and produce core measures with a common structure, clear from equations 2, 3, and 5—core measures that are, in effect, reweightings of the component price changes $\pi_{i,t}$ (or $\pi_{i,t}$, $\pi_{i,t-1}, \ldots$, in the dynamic factor model). Other approaches developed in the early 1990s boomlet of core inflation research focus on the aggregate headline indexes themselves, rather than their disaggregated components, taking the history of headline inflation—and perhaps other macroeconomic variables as well—as building blocks. These approaches posit core inflation measures that are either transformations of current and past headline inflation—

$$\pi_t^C = \phi\,(\pi_t,\ \pi_{i,t-1}, \ldots)$$

—or transformations of current and past headline inflation and other macroeconomic variables—

$$\pi_t^C = \phi\,(\pi_t,\ \pi_{t-1}, \ldots;\ X_t,\ X_{t-1}, \ldots).$$

The former would include the exponential smoothing approach due to Cogley (2002), while the latter would include the structural VAR approach of Quah and Vahey (1995).

Cogley uses a simple learning model in the spirit of Sargent (1999) to motivate a core measure that is a geometrically declining distributed lag of current and past headline inflation rates—that is,

$$\phi\left(\pi_t, \pi_{t-1}, \pi_{t-2}, \ldots\right) = g_0 \sum_{j=0}^{\infty} \left(1 - g_0\right)^j \pi_{t-j} ,$$

where the single parameter g_0 can be calibrated based on considerations about the speed with which agents are assumed to detect shifts in the inflation process. Quah and Vahey, in contrast, estimate a bivariate VAR in headline inflation and industrial output. They identify the VAR’s structural shocks by assuming that one of the two shocks—which they term the core inflation shock—has no long-run effect on output. They then obtain a core inflation series by simulating the estimated VAR using only the core inflation shocks as inputs.

As we noted in the introduction, many central banks assign a prominent role to core inflation measures in either formulating monetary policy or communicating policy decisions to the public. Several of those central banks—Australia’s, Sweden’s, New Zealand’s, and Canada’s—are explicit inflation targeters. While we cannot prove causality here, we nonetheless find it striking that the early-1990s resurgence of core inflation research broadly coincides with the increasing popularity of inflation targeting as a framework for monetary policy. More generally—thinking about both the theory and practice of monetary policy—growth in the volume of research on core inflation has occurred against a backdrop of increasing emphasis on rules-based formulations of policy. To us, this seems quite natural: A practical concern for price stability and a theoretical focus on the performance of monetary policy rules both lead logically to questions about the appropriate inflation measure (or measures) to inform, implement, and evaluate monetary policy.

Finally, the reader will note that we've now used the term *core inflation* roughly two dozen times without offering a definition. We will continue to refrain from offering a definition in what follows, which is a choice that is itself something of a characterization of the literature on measuring core inflation. From the research surveyed above, and the papers discussed below, one can compile a list of desiderata that researchers have suggested a core inflation measure should possess—that it should be smooth; that it should track the trend in headline inflation in real time; that it should forecast future headline inflation; and so on. Different authors emphasize different criteria. To the extent that precise definitions have been given, those definitions are not necessarily consistent. Some authors view core inflation as a purely statistical construct. For example, core inflation is some measure of the trend in headline inflation, a trend that cannot be observed in real time but can be estimated. A good measure of core inflation is then a good real-time estimate of the trend in headline inflation.[11] This is clearly a different definition from, say, Quah and Vahey (1995), in which core inflation is the part of measured headline inflation that has no long-run impact on real output. These definitions and the core indexes that follow from them are, in turn, different from the indexes that, according to some variants of the sticky-price New Keynesian model, should be the focus of the monetary authority's attention. Aoki (2001) and Goodfriend and King (1997), for example, conclude that policymakers should focus on an index of the economy's sticky prices and ignore price movements in the economy's flexible-price sectors—arguments that give some theoretical justification for the "inflation ex..." approach.

The lack of consensus on a definition of core inflation was touched on at the conference by Stephen Cecchetti, who urged researchers to think hard about the purpose for which core inflation measures are constructed. The final bullet point of Cecchetti's presentation—"What is it for?"—echoes a thought expressed by the great Irish-born economist Francis Ysidro Edgeworth ninety years ago in one of his many papers on the construction of index numbers: "It is a peculiarity of the problem that much thought must be expended in order to find the meaning of the question before you begin to answer the question." (Edgeworth 1918)

We now discuss the papers presented at the conference.

---

[11] This formulation is made explicitly in Bryan, Ceccheti, and Wiggins (1997) and Dolmas (2005), for example.

2. THE PAPERS

Inflationary Pressure in Retrospect

Several of the conference papers propose new core inflation measures; others evaluate existing measures. One paper, however, by Diana Weymark and Mototsugu Shintani (2006), examines from a historical perspective the inflationary pressures faced by the Federal Reserve over the past four decades.[12] While Weymark and Shintani do not explicitly deal with core inflation, their analysis can inform our thinking on the subject, since core inflation is sometimes defined as that portion of overall inflation that can be controlled through monetary policy.

Weymark and Shintani begin by positing non-model-specific definitions of the inflationary pressure faced by a central bank, the extent to which the central bank's actions alleviated that pressure, and the impact of the central bank's actions on inflation expectations.

They define what they call "ex ante inflationary pressure" through a counterfactual thought experiment, as the change in inflation that would have occurred from one quarter to the next had the central bank held its policy interest rate constant. The extent to which the central bank's interest rate action alleviated inflationary pressure—what the authors call the "effective price stabilization index"—is then defined as the difference between the change in inflation that actually occurred and ex ante inflationary pressure, as a fraction of ex ante inflationary pressure.

Formally, they assume that inflation in period t, π_t, can be written as a function of the central bank's policy rate i_t, a vector of expectational variables x_t^e, and random disturbances e_t and u_t:

$$\pi_t = h(i_t, x_t^e, e_t, u_t),$$

where h is a decreasing function of i_t. Such an expression for π_t can, as they note, be derived from any model in which there is a negative trade-off between the central bank's policy rate and the inflation rate, conditional on inflation expectations and shocks to the economic environment. Examples include models consisting of a New Keynesian Phillips curve and an IS equation, of the sort Clarida, Galí, and Gertler (1999) describe.

To clarify the counterfactuals that define ex ante inflationary pressure and the effective price stabilization index, suppose that the central bank sets the interest rate according to an inertial policy rule of the form $i_t = f(i_{t-1}, e_t, u_t)$ and that the expectational variables x_t^e are consistent with rational expectations. Weymark and Shintani's ex ante inflationary pressure would assume the central bank makes a fully anticipated one-period-only deviation from its rule, setting $i_t = i_{t-1}$ rather than $f(i_{t-1}, e_t, u_t)$. Let $\hat{x}_t^e$ denote the values of the expectational variables, assuming the public understands the one-period deviation the central bank is engaging in—and believes it is credible. The inflation that results from this counterfactual is $\pi_t^0 = h(i_{t-1}, \hat{x}_t^e, e_t, u_t)$, and ex ante inflationary pressure is then

---

[12] Here and below, references to the papers presented at the conference refer to the versions—typically in working paper form—at the time of the conference. Conference papers, as well as presentations by discussants, are available online at http://dallasfed.org/news/research/2007/07price.cfm.

defined as $\pi_t^0 - \pi_{t-1}$. The effective price stabilization index is then defined by $[(\pi_t - \pi_{t-1}) - (\pi_t^0 - \pi_{t-1})]/(\pi_t^0 - \pi_{t-1}) = (\pi_t - \pi_t^0)/(\pi_t^0 - \pi_{t-1})$.

Of course, in most macroeconomic models, central bank actions affect inflation outcomes through two channels—one running from interest rates to resource utilization rates to inflation and one working through inflation expectations. In general, the change in inflation that follows a change in policy will reflect the impact of policy through both channels. To isolate the impact of policy actions on expectations, Weymark and Shintani posit the concept of "ex post inflationary pressure," which they define using another, more complex, counterfactual: the change in the inflation rate that would have occurred if the central bank could have held its policy interest rate constant but still realized the change in inflation expectations that its actual policy response brought about. They then define an "index of policy effectiveness" as the difference between ex ante and ex post inflationary pressure—in essence, the extent to which the central bank's impact on expectations either mitigated or exacerbated inflationary pressures. For example, if x_t^e denotes the value of the expectational variables in period t in a rational expectations equilibrium in which the central bank is following a rule of the form $i_t = f(i_{t-1}, e_t, u_t)$, then ex post inflationary pressure is measured as $h(i_{t-1}, x_t^e, e_t, u_t) - \pi_{t-1}$.

While their definitions are not highly model-dependent, Weymark and Shintani must posit a model to make their definitions operational. The model they use is a variant of the sorts of simple New Keynesian macroeconomic models now in common use.[13] They estimate the parameters of their model using quarterly U.S. data from 1966:Q1 through 2001:Q4, calculate rational expectations equilibria, and then derive values for their indexes using the estimated parameters in conjunction with the historical data on inflation, interest rates, and resource utilization. They then use their indexes to evaluate the conduct of monetary policy over the sample period.

Not surprisingly, Weymark and Shintani find that the period from 1966:Q1 through 1979:Q2 is one in which positive inflationary pressure was quite pronounced, while 1979:Q3 through 2001:Q4 was generally characterized by negative inflationary pressure.

One surprising feature of Weymark and Shintani's results is the similarity they find in the Fed's policy responses to inflationary pressures over the entire period, which spans the chairmanships of William McChesney Martin, Arthur Burns, G. William Miller, Paul Volcker, and Alan Greenspan. Weymark and Shintani find that, by and large, all five chairmen acted to mitigate positive inflationary pressures (or magnify negative inflationary pressures) to a similar degree.

Weymark and Shintani's indexes also provide evidence suggesting that the Fed, especially during Greenspan's tenure, took advantage of negative inflationary pressure to bring down the economy's inflation rate—a strategy that has sometimes been referred to as "opportunistic disinflation."

---

[13] The model they use is a variant on the models Rudebusch (2002) and Clarida, Galí, and Gertler (1999) employ.

Alternative Weights, Alternative Inflation Measures

As we noted in Section 1, the price indexes commonly used to measure inflation—headline measures like the CPI and PCE—aggregate the prices of individual goods and services using weights that reflect some notion of each item's importance in spending or output. This sort of weighting is often based on the theory of cost of living indexes (as in the case of the CPI) or derives from principles governing national income accounting (as in the case of the PCE and GDP price indexes). Price indexes with this form have become sufficiently standard as inflation measures that introductory economics texts frequently discuss their construction as a concomitant to the treatment of inflation.

As we also discussed above, many of the common core versions of these headline indexes can be thought of as applying different weighting schemes to aggregate component prices, with some weighting schemes emphasizing cross-sectional properties of the component price changes (the trimmed mean and median), others emphasizing times-series properties (the "inflation ex…" measures), and at least one taking a hybrid approach (the dynamic factor model).

A paper by Julie K. Smith (2006) examines core PCE measures constructed using several less-conventional approaches to weighting, in addition to the more standard "inflation ex…" measures, trimmed means, and medians. Like several of the core measures discussed in Section 1, the basic building blocks in Smith's framework are the component price changes that are aggregated into the headline index, though in Smith's case, these changes are twelve-month inflation rates for fifty-one broad components of the PCE. The primary criterion Smith uses to evaluate the performance of the various core measures is the ability of each to forecast overall PCE inflation over the subsequent twelve months.

Let $\pi_{t,t+12}$ denote the headline inflation rate from period t to $t+12$, and let $\pi^i_{t-12,t}$ denote the percentage change in component i from $t-12$ to t. The first weighting scheme Smith considers is one that suggests itself immediately, given her criterion for evaluating core inflation measures: If the goal is a best forecast of $\pi_{t,t+12}$ from a linear combination of the components $\pi^i_{t-12,t}$, then eschew any reference to a component's importance in expenditure and simply consider the weights one obtains from estimating a forecasting equation of the form

$$\pi_{t,t+12} = \alpha + \sum_i \beta_i \pi^i_{t-12,t} + \varepsilon_t .$$

Of course, given its construction, the core measure resulting from this exercise performs better, in sample, than the other alternatives Smith considers; surprisingly, though, it bests the alternatives out of sample as well.

Among the alternatives Smith considers is a scheme in which components are weighted according to the persistence of their price changes—i.e., component weights are proportional to the coefficients obtained from estimating AR(1) processes for each component. Hence, the components that receive the greatest weight are those with the property that unusually high or low rates of price increase tend to persist over time; components where unusually high or low rates of price increase are quickly reversed are given less weight. Smith also considers a weighting scheme that combines persistence weights and expenditure weights and one that combines persistence weighting with trimming. Like the dynamic factor model discussed in Section 1, the weights in the latter scheme combine both time-series and cross-sectional information.

A Neo-Edgeworthian Index

In several papers between 1887 and 1925, Edgeworth suggested—vaguely, as Diewert (1995) notes—weighting the components of a price index inversely to their variability.[14] This sort of weighting scheme—eventually formalized in the modern "stochastic approach" to index number construction[15]—was the focus of a paper presented at the conference by Richard Anderson. The paper (Anderson et al. 2007)—coauthored with Jane Binner, Fredrik Andersson, and Thomas Elger—brings sophisticated mathematical machinery to bear on a tricky statistical problem: extracting a common inflation trend from component price increases whose volatilities vary over time.

The jumping-off point for their approach is the assumption embodied in equation 4, which decomposes the price changes for individual items into two parts, an inflation term, common to all items, and an idiosyncratic element:

$$\pi_{i,t} = \pi_t^* + x_{i,t}.$$

If the idiosyncratic elements are, on average, zero—that is, if the $x_{i,t}$'s have $E(x_{i,t}) = 0$—any convex combination of the $\pi_{i,t}$'s, including convex combinations based on expenditure weights or simple averages, provides an unbiased estimate of the inflation component π_t^*.

Basic principles of statistics, though, suggest that one can achieve a more efficient estimate of π_t^* by giving less weight to items whose idiosyncratic components are more volatile. This is the key insight behind Edgeworth's suggestion and the principle behind the classical approach to signal extraction. If, for example, the $x_{i,t}$'s are distributed independently across components, each with mean zero and constant variance σ_i^2, then an efficient estimate of π_t^* would be a weighted average of the form

$$\hat{\pi}_t^* = \sum_i w_i \pi_{i,t},$$

where each w_i has the form

---

[14] See, for example, Edgeworth (1925, 383), where Edgeworth notes that by this principle, "the price of pepper might deserve more weight than the price of cotton." Diewert (1995) gives a complete set of references for the incidence of this idea in Edgeworth's writings.

[15] See Diewert (1995).

$$w_i = \frac{1/\hat{\sigma}_i^2}{\sum_j 1/\hat{\sigma}_j^2},$$

with $\hat{\sigma}_i^2$ an estimate of the variance of $x_{i,t}$.[16] But a cursory look at the data on prices of items that go into the CPI or PCE suggests that the volatility of price changes for those items varies considerably over time. A set of weights based on relative component volatilities calculated from data through, say 1990, might look very different from the weights one would calculate with data after 1990. Also, as Anderson et al. note, echoing an argument in Keynes' ***Treatise on Money***, if the rate of price increase differs across components—if, in fact, there is no π_t^* term common to all the $\pi_{i,t}$'s—then we must take some account of the economic importance of the various component prices in constructing an appropriate notion of an "average price change."[17] The model Anderson et al. elaborate allows for time variation in the mean and variance of each of the $\pi_{i,t}$'s, and the weighting scheme that results combines each component's (time-varying) volatility with its expenditure weight:

$$w_{i,t} = \frac{\left(1/\hat{\sigma}_{i,t}^2\right)\varpi_{i,t}}{\sum_j \left(1/\hat{\sigma}_{j,t}^2\right)\varpi_{j,t}},$$

where $\varpi_{i,t}$ is component i's expenditure weight.

Of course, allowing for variation across time and across components in the means and variances of the $\pi_{i,t}$'s poses tricky issues of identification, which Anderson et al. handle by applying the concept of wavelets, a tool used extensively in engineering and physics. They then apply their techniques to construct core price indexes for the U.S., U.K., and euro area. While they consider their results preliminary, their paper demonstrates the tractability of a promising technique for incorporating time-varying component-level volatility into measures of core inflation.

Core Inflation from an Economic Model

As might be gathered from the preceding discussion, many core inflation measures are designed to perform well according to some statistical criterion. For example, the trimming proportions used in the Dallas Fed's trimmed mean PCE were chosen to minimize the discrepancy in histori-

---

[16] To be precise, the estimates of $\hat{\pi}_t^*$ and the $\hat{\sigma}_i^2$'s are obtained by solving a system of simultaneous equations, since an estimate of the mean of $\pi_{i,t}$ (which is just π_t^*) is necessary to calculate $\hat{\sigma}_i^2$.

[17] Keynes—who was arguing contra the "unweighted" index numbers championed by Jevons, Edgeworth, and others—takes issue with the idea there is a common inflation term (our π_t^*) around which the component price changes are distributed: "There is no bull's-eye. There is no moving but unique centre, to be called the general price level or the objective mean variation of general prices, round which are scattered the moving price levels of individual things. There are all the various, quite definite, conceptions of price-levels of composite commodities appropriate for various purposes and inquiries which have been scheduled above, and many others too. There is nothing else. Jevons was pursuing a mirage." (Keynes 1930, chapter 6)

cal data between the trimmed mean inflation rate and a measure of trend PCE inflation. A criticism of such purely statistical approaches to core inflation is that they lack an economic justification.

Stefano Siviero and Giovanni Veronese (2007) propose a core inflation measure that addresses this criticism. The idea behind their approach, which Siviero presented at the conference, is straightforward: Imagine a central bank that adjusts interest rates not in response to an aggregate index of inflation but in response to a set of components that make up the aggregate index. In particular, suppose the central bank follows a Taylor-type rule that puts separate weights on the component inflation rates, as well as on a measure of economic slack, such as the output gap. The central bank's optimal choice of weights on the component inflation rates determines a measure of core inflation—one that is, in Siviero and Veronese's terminology, "policy-sensible."

For example, consider disaggregating the CPI into four broad components: food ($\pi_{F,t}$), energy ($\pi_{E,t}$), goods other than food and energy ($\pi_{G,t}$), and services other than energy services ($\pi_{S,t}$). Now suppose the Fed adjusts the federal funds rate according to a Taylor-type rule that depends on slack in the economy (measured by, say, the unemployment rate or output gap) and the inflation rates in those four CPI components. That is, imagine that the Fed follows a rule of the form

$$i_t = a + \theta_x x_t + \theta_F \pi_{F,t} + \theta_E \pi_{E,t} + \theta_G \pi_{G,t} + \theta_S \pi_{S,t},$$

where x_t is a measure of slack in period t, the θ_i's are coefficients on the rule's five inputs, and a is an intercept term. One hypothetical set of coefficients on the inflation components might be $(\theta_F, \theta_E, \theta_G, \theta_S) = (0,0,0.3,0.7)$. A rule with these coefficients would call for the Fed to raise the funds rate 0.7 percentage point for every 1 percentage point increase in core services inflation, 0.3 percentage point for every 1 percentage point increase in core goods inflation, and zero percentage points in response to any changes in the inflation rates for food and energy. In the conventional ex food and energy CPI, core services have a weight of roughly 70 percent, while core goods have a weight of roughly 30 percent. Thus, an interest rate rule with these coefficients would correspond, roughly, to a Taylor rule that adjusts the interest rate one-for-one with changes in ex food and energy CPI inflation.

This choice of weights—0.7 on core services and 0.3 on core goods, with 0 on food and 0 on energy—is clearly ad hoc. What Siviero and Veronese are interested in are the coefficients for an optimal rule. To that end, they estimate the parameters of a small-scale New Keynesian macroeconomic model using data for the U.S. and the euro area. The model has sectors corresponding to food, energy, core goods, and core services, with Phillips curves for each sector and a single IS equation describing the evolution of the slack variable. They then use their estimated models for the U.S. and euro areas to calculate optimal Taylor rule coefficients and the policy-sensible core inflation measures those coefficients imply. The rules are optimal, of course, relative to the central bank's preferences, which are assumed to take the form of a quadratic loss function defined over headline inflation, the output gap, and the policy interest rate.[18] Note that the policymaking framework Siviero and Veronese describe is similar in spirit to the framework adopted by the Bank of Canada: The inflation rate the central bank cares about is a headline rate, but the bank uses a core measure as an operational guide—in Siviero and Veronese's model, as an input to the bank's policy rule.[19]

What do they find? Surprisingly, for their U.S. model, the optimal choice of coefficients implies a measure of core CPI inflation close to CPI ex food and energy—that is, the resulting coefficients are roughly proportional to the 7/10, 3/10, 0, 0 pattern described above. For the euro area, the results are quite different. The optimally chosen coefficients put most of the weight on core services inflation; small, but non-negligible weights on food and core goods (with food weighted slightly *more* than core goods); and zero weight on energy. This weighting corresponds to no core measure that we know of.

While Siviero and Veronese consider their results to be tentative, the results suggest that for some economies, the usual ex food and energy core gauges may be close to optimal, while for other economies, an optimal core index may look like none we've seen to date.

A Core Inflation Horse Race

In Section 1, we summarized several of the more prominent approaches to measuring core inflation that have been developed over the past twenty-five years or so. If we think in terms of basic forms rather than particular variants—at the level of what we may consider "phyla," perhaps, rather than "species"—then we have at least five basic forms summarized in Section 1: (1) the "inflation ex…" approach; (2) trimming—including the median as well as the trimmed mean; (3) the dynamic factor model; (4) models like Cogley's, based on the history of headline inflation; and (5) models like Quah and Vahey's, which use the history of headline inflation and other macro variables. The species within each of these categories are also diverse. Within the trimming approach, for example, we have the median, symmetrically trimmed means (for example, the Cleveland Fed's trimmed mean CPI), and asymmetrically trimmed means (for example, the Dallas Fed's trimmed mean PCE). A similar variety exists among "inflation ex…" measures. At the risk of pushing the biological analogy too far, how do all these variants survive?

This is, in a sense, part of the question posed by Robert Rich and Charles Steindel (2007). Rich and Steindel begin by noting that the literature on alternative measures of core inflation has produced little consensus about which alternative is the best. They attribute this lack of consensus

---

[18] Their expected loss can be written in terms of variances as $L = \mathrm{var}(\pi) + \lambda \mathrm{var}(x) + \mu \mathrm{var}(i)$. They consider the case where the central bank cares only about inflation ($\lambda = \mu = 0$), cares equally about inflation and the output gap ($\lambda = 1$, $\mu = 0$), cares equally about inflation and interest rate volatility ($\lambda = 0$, $\mu = 1$), or cares equally about all three ($\lambda = \mu = 1$).

[19] By this comparison, we don't mean to imply that the Bank of Canada sets policy according to a Taylor rule.

to a lack of uniformity in the methods used to compare alternatives: Different authors often apply different statistical criteria (such as different forecast horizons), examine different sample periods, and compare core measures based on different underlying price indexes (PCE for some, CPI for others). As alternatives are compared along different dimensions, different authors tend to find support for different preferred core measures, and no overall winner emerges.

Rich and Steindel attempt to remedy this lack of uniformity by evaluating several prominent core candidates against each other within a common testing framework—a sort of horse race among core measures, run under uniform conditions. They examine core alternatives based on both the CPI and PCE and apply a common set of statistical criteria: the ability to (1) track the trend rate of inflation and match the long-run average rate of inflation, (2) explain within-sample movements in inflation, and (3) forecast inflation out of sample. They also make several robustness checks, including varying the forecast horizon and sample period.

The main conclusion of the horse race—or perhaps better, “triathlon,” given the three criteria they examine—is something of a negative result. Among the candidates they consider—ex food and energy, ex energy, median, and Cogley’s exponentially smoothed inflation—none emerges as the clear winner. They find that the performance of the various alternatives differs markedly across inflation measures, statistical criteria, and sample period. The median CPI, for example, does a relatively poor job tracking the trend in CPI inflation but performs relatively well in the within-sample tests of explanatory ability. Likewise, in out-of-sample forecasts, the exponentially smoothed CPI measure does very well over the period 1990–2004 but rather poorly over the period 1995–2004.

Rich and Steindel interpret their results as suggesting that “the features of transitory price movements are better described by the core inflation measures on a collective, rather than individual, basis. That is, there is too much variability in the nature and sources of transitory price movements to be effectively captured through the design of any one of the core inflation measures.”

The Performance of Trimmed Means

The Rich and Steindel paper does not consider trimmed mean measures of core inflation, except for the extreme case of the median. Trimmed means are the main focus of two other conference papers, one by Andrea Brischetto and Anthony Richards and another by Michael Bryan.

Brischetto and Richards (2007) evaluate the trimmed mean alongside headline and “inflation ex…” measures of consumer price inflation using data from the U.S., Australia, Japan, and the euro area. The underlying price indexes are the CPIs of the U.S., Japan, and Australia and the Harmonized Index of Consumer Prices (HICP) of the euro area.

Like Rich and Steindel, Brischetto and Richards compare the performance of the various inflation measures along several dimensions, including their smoothness, their ability to track trend inflation, and their ability to forecast inflation at short horizons of three to six months. They find that trimmed means outperform the headline and "inflation ex…" measures on a range of criteria.

A novel feature of Brischetto and Richards' approach to trimming is their treatment of the owners' equivalent rent (OER) component of the U.S. CPI. OER is by far the largest component of the CPI, with an expenditure weight of roughly 23 percent. Because of its size and the fact that it is less volatile than other CPI components, OER can tend to dominate the behavior of trimmed means, the more so the greater the fraction of items trimmed. OER invariably amounts to a large fraction of "what's left in the middle" after trimming. In the extreme case of the median—which amounts to trimming out everything but the component whose price increase is at the midpoint of the distribution—the median CPI inflation rate is frequently just the rate of increase in OER. This was the case, for example, with the Cleveland Fed's median CPI for eight straight months, from December 2006 to July 2007.

On the surface, it would seem that OER, as a basic component of expenditure, cannot be disaggregated further. Brischetto and Richards' innovation is to view OER for the economy as a whole as an aggregate of regional OERs. They disaggregate OER, with its 23 percent expenditure weight, into four regional components—Northeast, South, Midwest, and West—with much smaller weights that range from 4 to 9 percent. They apportion the overall weight accorded to OER in the CPI amongst their four regional OERs using U.S. census data on regional housing stocks. Brischetto and Richards show that the median and trimmed means calculated using disaggregated OER outperform the standard versions of these indexes along many of their criteria, though the improvement is much more pronounced with the median than with the trimmed means they consider.[20]

The question Bryan (2007) asks is, Among several alternative CPI measures—the overall CPI, the CPI ex food and energy, the median CPI, and the trimmed mean CPI—which gives the earliest indication that the underlying trend in CPI inflation has changed?

Applying a standard econometric test for structural breaks—specifically, the Bai and Perron (1998) test—Bryan identifies three inflation regimes in monthly CPI data over 1973–2007: a high inflation period from January 1973 to August 1981, with average CPI inflation of 9.4 percent; a moderate inflation period from September 1981 to December 1990, with average inflation of 4.1 percent; and a low inflation period from January 1991 to the end of the sample, with average inflation of 2.7 percent. Importantly, these three periods are determined ex post, by looking at data

---

[20] Since August 2007, the Cleveland Fed has used Brischetto and Richards's approach to disaggregating OER in the calculation of the Bank's median and trimmed mean CPI inflation indexes.

over the full sample, as it exists today.

Now, imagine an analyst operating in real time who—at each date between January 1973 and the present—has only the data that were available at that date. With the benefit of hindsight, we know that a break occurred in August 1981 and that it was sizeable. According to Bryan's results, average inflation fell by 5.4 percentage points from the high inflation period to the moderate inflation period. It's too much to hope that our real-time analyst would detect the break at the moment it occurred, but how soon after August 1981 would the analyst have detected the break in trend? Not surprisingly, the answer depends on whether the analyst was monitoring overall CPI inflation, CPI ex food and energy, the median CPI, or the trimmed mean. Bryan shows that for this large break in trend, an analyst monitoring the overall CPI would have been the first to conclude that a break had occurred, 18 months after the fact. An analyst using any of the three core measures would not have detected the break until much later—at best 32 months after it had occurred (using the trimmed mean) and at worst 39 months after it had occurred (using CPI ex food and energy).

These results clearly don't look good for the trimmed mean and median, but what about the break in January 1991? This break in trend was much more subtle, with average inflation differing across the moderate and low inflation periods by only 1.4 percentage points. Here, the trimmed mean and median prove superior, detecting the break 18 months after the fact (the median) or 20 months after the fact (the trimmed mean). The CPI ex food and energy comes in third, at 26 months after the break, while the overall CPI, in this case, lags far behind at 78 months after the break.

Bryan interprets the trimmed mean's and median's superiority at detecting the more subtle break in trend as indicating that trimming—which reduces the noise in monthly data—is particularly useful in low inflation environments, when the underlying signal is weak relative to the noise in the data. This point is evident to anyone who has listened to an older classical music recording on vinyl LP, then listened to the same recording, remastered and de-hissed, on CD. The value of noise reduction is most pronounced not in the fortissimo sections of the work, but in the pianissimo passages.

3. CONCLUSION

One of the goals of the conference—and this paper—was to capture some of the wide variety of work being done, primarily at central banks but also in academia, on the measurement of core inflation. We think the papers we have summarized above capture that variety well.

We've made no mention thus far of any of the discussion that took place at the conference, but perhaps this is an appropriate point to do so. Much of the discussion reflected the sentiment expressed by Edgeworth in the passage we quoted in Section 1—that is, the difficulty in figuring out the question to which measures of core inflation are the answer. To cite one example of a point raised a few times during the conference, why is the ability to forecast future inflation typically a desideratum for a core inflation measure? Or to put it differently: If one wanted to forecast future inflation, why limit the information set to the past behavior of inflation or its components, rather than make the best forecast one can given all information available?

Rich and Steindel attribute the lack of consensus on the best core inflation measure to, at least partly, differences in the performance criteria authors use to evaluate core measures and the failure of any one measure to dominate across all criteria. That different authors focus on different criteria, however, is not a matter of happenstance and may, we suspect, indicate that they have different assumptions about what a core measure ought to do. One response to Rich and Steindel's results—and they discuss several possible responses, each of which they consider problematic—is to "acknowledge that different core inflation measures seem better suited at performing different tasks, and then adopt the appropriate core inflation measure as the guide for a particular stated purpose." In weighing the desirability of, or difficulty with, this possible response, as well as the others they suggest, Rich and Steindel are explicitly evaluating these possibilities from a policymaking perspective. From a long-term research perspective, however, the direction seems, at least to us, clear: to figure out what the correct criterion is—which may or may not be among those already considered. In other words, now, as in Edgeworth's time, "much thought must be expended in order to find the meaning of the question."

If the number of core inflation measures being produced and evaluated seems large, it's worth remembering that in the early days of price measurement, the number of indexes was almost certainly quite a bit larger. The early history of price index construction—particularly during the late nineteenth and early twentieth centuries—contains numerous debates over the merits of different index forms: weighted versus unweighted, simple averages versus geometric or harmonic means, fixed weights versus time-varying weights, and so forth. The variety of possible forms was captured nicely in Irving Fisher's 1922 magnum opus, *The Making of Index Numbers*, in which he analyzed and ranked 134 different index numbers. Of course, few of those indexes survive today. What caused the extinction of so many of these forms? Most likely it was progress in theory—in particular the growth of modern consumer theory. No doubt something similar will play out with respect to core inflation measurement. As models of monetary policy and its transmission are refined and consensus established, theory should increasingly provide guidance—perhaps along the lines suggested by Siviero and Veronese's paper or in the spirit of Aoki's work (2001)—on the construction of core inflation measures.

REFERENCES

Anderson Richard G., Fredrik N. G. Andersson, Jane Binner, and Thomas Elger (2007), "Core Inflation as Idiosyncratic Persistence: A Wavelet Approach to Measuring Core Inflation" (Paper presented at "Price Measurement for Monetary Policy" conference, Dallas, May 24–25, 2007).

Aoki, Kosuke (2001), "Optimal Monetary Policy Responses to Relative-Price Changes," *Journal of Monetary Economics* 48 (August): 55–80.

Bai, Jushan, and Pierre Perron (1998), "Estimating and Testing Linear Models with Multiple Structural Changes," *Econometrica* 66 (January): 47–78.

Blinder, Alan (1982), "The Anatomy of Double-Digit Inflation in the 1970s," in *Inflation: Causes and Effects*, ed. Robert E. Hall (Chicago: University of Chicago Press), 261–82.

Brischetto, Andrea, and Anthony Richards (2007), "The Performance of Trimmed Mean Measures of Underlying Inflation" (Reserve Bank of Australia paper presented at "Price Measurement for Monetary Policy" conference, Dallas, May 24–25, 2007).

Bryan, Michael F. (2007), "Monitoring Inflation in a Low-Inflation Environment" (Federal Reserve Bank of Cleveland paper presented at "Price Measurement for Monetary Policy" conference, Dallas, May 24–25, 2007).

Bryan, Michael F., and Stephen G. Cecchetti (1993), "The Consumer Price Index as a Measure of Inflation," NBER Working Paper Series, no. 4505 (Cambridge, Mass., National Bureau of Economic Research, October).

Bryan, Michael F., and Stephen G. Cecchetti (1994), "Measuring Core Inflation," in National Bureau of Economic Research Studies in Business Cycles, vol. 29, *Monetary Policy*, ed. N. Gregory Mankiw (Chicago: University of Chicago Press), 195–215.

Bryan, Michael F., Stephen G. Cecchetti, and Rodney L. Wiggins II (1997), "Efficient Inflation Estimation," NBER Working Paper Series, no. 6183 (Cambridge, Mass., National Bureau of Economic Research, September).

Bryan, Michael F., and Christopher J. Pike (1991), "Median Price Changes: An Alternative Approach to Measuring Current Monetary Inflation," Federal Reserve Bank of Cleveland *Economic Commentary*, December 1.

Clarida, Richard, Jordi Galí, and Mark Gertler (1999), "The Science of Monetary Policy: A New Keynesian Perspective," *Journal of Economic Literature* 37 (December): 1661–707.

Cogley, Timothy (2002), "A Simple Adaptive Measure of Core Inflation," *Journal of Money, Credit and Banking* 34 (February): 94–113.

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

Dolmas, Jim (2005), "Trimmed Mean PCE Inflation," Federal Reserve Bank of Dallas Working Paper no. 0506 (Dallas, July).

Eckstein, Otto (1981), *Core Inflation* (Englewood Cliffs, N.J.: Prentice-Hall).

Edgeworth, F. Y. (1918), "The Doctrine of Index-Numbers According to Professor Wesley Mitchell," *Economic Journal* 28 (June): 176–97.

———(1925), "The Plurality of Index-Numbers," *Economic Journal* 35 (September): 379–88.

Fisher, Irving (1922), *The Making of Index Numbers* (Boston: Houghton Mifflin Co.).

Goodfriend, Marvin, and Robert G. King (1997), "The New Neoclassical Synthesis and the Role of Monetary Policy," in *NBER Macroeconomics Annual 1997*, ed. Ben S. Bernanke and Julio J. Rotemberg (Cambridge, Mass.: MIT Press), 231–83.

Keynes, John Maynard (1930), *A Treatise on Money* (London: Macmillan and Co.).

Quah, Danny, and Shaun P. Vahey (1995), "Measuring Core Inflation," *Economic Journal* 105 (September): 1130–44.

Rich, Robert, and Charles Steindel (2007), "A Comparison of Measures of Core Inflation" (Federal Reserve Bank of New York paper presented at "Price Measurement for Monetary Policy" conference, Dallas, May 24–25, 2007).

Rudebusch, Glenn D. (2002), "Assessing Nominal Income Rules for Monetary Policy with Model and Data Uncertainty," *Economic Journal* 112 (April): 402–32.

Sargent, Thomas J. (1999), *The Conquest of American Inflation* (Princeton N.J.: Princeton University Press).

Silver, Mick (2007), "Core Inflation: Measurement and Statistical Issues in Choosing Among Alternative Measures," *IMF Staff Papers* 54 (1): 163–90.

Siviero, Stefano, and Giovanni Veronese (2007), "A Policy-Sensible Benchmark Core Inflation Measure. An Application to US and Euro Area Data" (Bank of Italy Research Department paper presented at "Price Measurement for Monetary Policy" conference, Dallas, May 24–25, 2007).

Smith, Julie K. (2006), "Better Measures of Core Inflation?" (Lafayette College paper presented at "Price Measurement for Monetary Policy" conference, Dallas, May 24–25, 2007).

Weymark, Diana N., and Mototsugu Shintani (2006), "Measuring Inflation Pressure and Monetary Policy Response: A General Approach Applied to U.S. Data, 1966–2001" (Vanderbilt University paper presented at "Price Measurement for Monetary Policy" conference, Dallas, May 24–25, 2007).

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