Inflation Measurement

David E. Lebow and Jeremy B. Rudd
2006-43

NOTE: Staff working papers in the Finance and Economics Discussion Series (FEDS) are preliminary materials circulated to stimulate discussion and critical comment. The analysis and conclusions set forth are those of the authors and do not indicate concurrence by other members of the research staff or the Board of Governors. References in publications to the Finance and Economics Discussion Series (other than acknowledgement) should be cleared with the author(s) to protect the tentative character of these papers.
Abstract

Inflation measurement is the process through which changes in the prices of individual goods and services are combined to yield a measure of general price change. This paper discusses the conceptual framework for thinking about inflation measurement and considers practical issues associated with determining an inflation measure’s scope; with measuring individual prices; and with combining these individual prices into a measure of aggregate inflation. We also discuss the concept of “core inflation,” and summarize the implications of inflation measurement for economic theory and policy.

*JEL classification codes: C43, E31*
Inflation measurement is the process through which changes in the prices of individual goods and services are combined to yield a measure of general price change. In formal terms, we may specify the time-\(t\) rate of aggregate inflation \(P_t\) as

\[
P_t = F(p^1_t, p^2_t, \ldots, p^I_t),
\]

where \(F(.)\) is a function that aggregates a set of \(I\) individual time-\(t\) price changes \(p^i_t\).

Writing the problem in this manner highlights three basic issues associated with inflation measurement: First, we must decide what collection of price changes we wish to include (or, more generally, what should be the measure’s scope); second, we must ensure that the individual price changes are correctly measured; and finally, we must choose a method for combining those changes into a measure of aggregate inflation.

While the problem of inflation measurement can be broadly described in these terms, dealing with the numerous complications that emerge in practice requires some explicit conceptual framework. Probably the simplest way to construct a measure of overall inflation involves defining the aggregate price level in terms of the cost of a fixed basket of goods and services. Such a measure—sometimes labeled a cost-of-goods index, or COGI—has several practical advantages; in particular, for a broad enough basket of goods, the change in a COGI comes very close to what most people intuitively mean by an inflation rate, and a COGI-type measure can easily be defined for any subcomponent of expenditure or production (such as consumption, investment, or the output of intermediate goods). However, this simple measure of inflation faces an important practical difficulty. In a dynamic economy, the composition and nature of output will evolve as existing goods are consumed or produced in different quantities, as the characteristics of existing products change, or as entirely new goods are introduced; these changes make the COGI’s fixed bundle of goods become less representative over time. The COGI approach provides no guidance as to how to address this problem, suggesting that a more comprehensive conceptual framework is needed.
If we confine our attention to consumption prices, then a natural guiding principle is provided by the concept of a *cost-of-living index* (COLI), which measures the expenditure needed for an optimizing consumer to maintain a specified level of utility as prices change. The strength of the COLI framework derives from its grounding in the theory of consumer behavior, which can provide clear-cut suggestions (at least in principle) as to how to deal with such problems as changes in expenditure patterns or the introduction of new goods. That said, this feature of the COLI approach can also be a weakness to the extent that consumer theory provides an incorrect characterization of actual behavior (NRC 2002, pp. 53-58) or is insufficiently well developed to handle a particular practical situation. In addition, because the COLI concept only pertains to consumption, it provides little or no guidance about the construction of broader measures of inflation that include prices for other components of output (these might be of interest, for instance, to a monetary policymaker). The COLI framework therefore provides a natural guide to the construction of a consumer price index (CPI), which attempts to measure the prices of goods and services consumed by households; it will not, however, be able to inform the construction of a price index for overall GDP, which is defined to include the prices of *all* domestically produced final output—whether purchased by consumers, businesses, governments, or the rest of the world. (While a literature does exist on the measurement of price change from a *producer* perspective—see Diewert, 1983, for an overview—it has generally received much less attention than the corresponding consumer-based approach.)

Despite these potential shortcomings, a COLI-based approach is commonly employed as a framework for informing inflation measurement—in the United States, for example, the CPI uses the COLI concept both as its explicit measurement objective as well as a reference for making practical decisions about index construction (U.S. Bureau of Labor Statistics, 2005). In much of what follows, therefore, we follow common practice in using the concept of a cost-of-living index to guide our discussion of the three basic issues—scope, individual price measurement, and aggregation—that are associated with the measurement of inflation. In addition, we discuss the concept of *core inflation*, which can be motivated and interpreted in terms of an alternative approach to inflation
Finally, we conclude by considering the implications of these measurement issues for economic research and policy.

1. What items should be included in an inflation measure?

The scope or domain of a cost-of-goods index—whether it is defined for consumption goods or more broadly—is defined to include all items that are purchased and sold in market transactions, and, hence, that have well-defined prices. (In reality, of course, any inflation measure will only include a subset of goods consumed or produced in the economy, so sampling in order to provide a representative characterization of aggregate price change represents an important practical concern.) By contrast, the scope of a cost-of-living index is much broader than that of a corresponding COGI for consumption goods inasmuch as a COLI needs to account for anything that affects utility, including changes over time in “background” or “environmental” factors such as weather, pollution, crime, or the provision of public goods.

For a COLI-based measure of consumption price inflation, therefore, the relevant set of price changes $p_1, p_2, ..., p_t$ should in principle include changes in both market prices and the “shadow prices” of environmental factors (with the latter defined in the sense of Pollak, 1989). In practice, however, it is almost impossible to correctly measure the effect on utility of these sorts of changes (even if we could do so, inclusion of such factors strays beyond what most people understand by the term “inflation”). These considerations lead to the concept of a “conditional” COLI, which (again following Pollack, 1989) is defined as the smallest change in expenditure that is required in order to maintain a reference utility level following a change in prices, holding the state of the environment fixed. Although intuitive, the concept of a conditional COLI has its own conceptual difficulties. In particular, since preferences over market goods will likely depend on the environment (for example, demand for medical care depends on the incidence of disease), the rate of inflation implied by a conditional COLI will depend on the particular state of the environment that we condition on.
While the concept of a conditional COLI provides useful guidance regarding the relevant domain of a measure of consumption price inflation, it cannot unambiguously solve all questions regarding scope. For example, many households receive an implicit flow of services from owner-occupied housing. Assuming that the “price” of these services could be measured, it is unclear whether they should enter a COLI given that they are not generated by a market transaction or explicit expenditure (and are not closely related to a conventional notion of a price); of course, the initial home purchase does meet these criteria. A similar problem extends to a number of other goods that are consumed by households but not directly purchased by them (one example is banking services furnished without explicit charge, which are included in most national accounts definitions of consumption). And, again, once one moves outside of the realm of private consumption, the conditional COLI framework provides no practical guidance regarding the construction of inflation measures for other components of production or spending (such as investment) or for broader measures of inflation (such as the GDP price index).

Finally, a particularly difficult and controversial issue concerns the proper role of asset prices in inflation measures. If we extend the theory of a cost-of-living index to an intertemporal or multiperiod context (see Pollak, 1975), then expected changes in the price of future consumption streams can affect current inflation through their impact on lifetime utility. We can therefore consider a cost-of-living index that is defined to include current and future prices of consumption goods; furthermore, to the extent that information about future consumption prices is contained in current asset prices, an argument can be made for including these prices in a COLI-based inflation measure (Alchian and Klein, 1973). In practice, however, the volatility of asset prices—as well as the related fact that observed movements in asset prices can stem from sources unrelated to expected future consumption-price changes—typically precludes their inclusion in conventional inflation measures. (The current purchase prices of durable goods—which are often included—provide a partial exception.)
2. **How should the individual price changes be measured?**

A number of practical problems complicate the measurement of individual price changes. First, in a modern economy the characteristics of existing goods can change over time; likewise, new goods and services will constantly be entering—and old goods leaving—production and consumption. Left unaddressed, these problems will render it impossible to track the price changes for an identical set of goods and will cause the set of goods being priced to become increasingly less representative of actual consumption and production. This will obviously affect COGI-based measures of inflation, and it will also affect COLI-based measures to the extent that changes in the characteristics or variety of available goods have an effect on the utility that is realized from their consumption.

Several techniques exist for dealing with nontrivial changes in the characteristics (loosely speaking, the “quality”) of existing products; all of these involve some procedure for dividing the observed price change into a component that reflects changes in the good’s characteristics and a component that reflects “pure” price change—where only this latter component is appropriate for inclusion in an inflation measure. (Moulton and Moses, 1997 and NRC, 2002 provide a detailed description and assessment of these various methods of quality adjustment in the context of the U.S. CPI; see also ILO, et al., 2004.) For example, when the original and modified products exist in the same period, any difference in their prices can be attributed to differences in the goods’ characteristics. Alternatively, in the more common case where a good exists in one form in period $t$ and in another in period $t+1$, the “pure” price change over the intervening period can be imputed from the observed average price change for a similar group of goods. (A “matched model” index, which only includes price changes for goods that remain in the sample without change—and so implicitly assigns that average price change to other items—is a common example.) Finally, additional information may be brought to bear on the problem: Under certain assumptions, for example, data on the cost to manufacturers of modifying the characteristics of a product can be used to compute the effect of these modifications on the good’s price.
When detailed information about a product’s characteristics is available, so-called “hedonic” methods may be used. The hedonic approach relates the observed price of a good to its characteristics; any change in characteristics can then be explicitly controlled for and removed from the good’s total price change. Specifically, when the individual effects of a good’s characteristics on its price are stable over time, a measure of pure price change can be obtained by permitting the level of the price-characteristics relation to shift in each period. In the more realistic case where different hedonic functions exist for each period, a measure of pure price change between periods \( t \) and \( t+1 \) can be defined as \( h^{i+1}(z) / h'(z) \) or \( h^{i+1}(z_{t+1}) / h'(z_{t+1}) \), where \( h'(z) \) denotes the hedonic function in period \( i \) relating the good’s price to its set of characteristics \( z \). (Here, the first expression yields a “Laspeyres-like” price measure as the hedonic function is evaluated with the set of characteristics from the variety that is purchased in the base period; similarly, the second expression yields a “Paasche-like” measure.)

An important advantage of the hedonic approach to dealing with quality change is that it can be explicitly grounded in cost-of-living theory: Under relatively weak conditions, \( h^{i+1}(z) / h'(z) \) provides an upper bound for the compensating variation associated with a given price change; likewise, \( h^{i+1}(z_{t+1}) / h'(z_{t+1}) \) gives a lower bound for the equivalent variation (NRC, 2002, pp.153-154). It is unknown, however, whether these bounds are particularly tight. In addition, statistical agencies typically find real-time production of measures like these too difficult, and instead produce quality-adjusted price changes by scaling the observed price change for a good by the ratio \( h^{t-j}(z) / h^{t-j}(z_{t+1}) \), where the \( t-j \) superscript makes apparent the dependence of the estimated hedonic function on an earlier period’s data. Such a procedure cannot, in general, be justified in terms of a COLI-based approach (Pakes, 2002).

The “new goods” problem can be thought of as a more difficult variant of the quality adjustment problem in which the new good contains features or characteristics that have never existed before (in a sense, the dimension of the “characteristics space” has increased); examples include the introduction of the video cassette recorder or
cellular telephone. In this case, one needs a method for imputing the price of a newly introduced good in the period prior to its first appearance in the economy; as was suggested by Hicks (1940, pp. 114-15), one logical imputation involves setting this pre-introduction price equal to the price at which the demand for the good is just equal to zero. While such an approach can be explicitly motivated in terms of a COLI-based framework, its implementation requires a degree of information about consumer preferences that is unlikely to be realized in practice (see Hausman, 1997, for a representative example). It is therefore common for statistical agencies to attempt to mitigate the new goods problem through the more rapid addition of new items into the set of price changes being tracked over time; while intuitive, this approach may not always ameliorate the effects of new-good introduction (Pakes, 2002).

Another problem that arises in measuring individual price changes relates to the fact that even identical goods can sell for different prices across different sellers. These differentials could reflect true price differences—a particular outlet might simply be able to charge a lower price—but they could also reflect characteristics of the outlet itself, such as customer service or convenience. In the latter case, two otherwise-identical goods should be treated as different products if they are sold at different outlets; similarly, when the outlet used to price a particular good changes, some adjustment—akin to the sorts of quality adjustments discussed above—must be made to the good’s price.

One final issue relating to the measurement of individual price changes is that a goods’s purchase price need not be related to its effect on current-period utility if it provides consumption services in more than one period (as is the case for a durable good) or if it can be stored for later consumption. For a durable good, the conceptually relevant measure of the change in the good’s price in a given period is the change in its user cost. In practice, the user cost turns out to be difficult to estimate and often implies erratic price movements. In the presence of a well-functioning rental market, the cost of hiring a good can serve as a proxy for its user cost; this “rental equivalence” procedure is used in the U.S. CPI for owner-occupied housing. However, the absence of rental markets for...
most durable goods limits the usefulness of this technique, and in practice, the purchase prices of many durable goods are directly included in most inflation measures.

3. How should the individual price changes be combined?

The combination of individual price changes into an aggregate measure of inflation falls into the domain of the theory of index numbers, a full discussion of which is beyond the scope of this survey. We therefore focus on some of the practical issues that arise in choosing and implementing an aggregation formula.

A natural way to construct a cost-of-goods index involves weighting the individual price changes for the components of the fixed market basket by their shares in overall expenditures. When the initial period of the index is the same as the period used to specify the expenditure weights, the resulting measure corresponds to a Laspeyres index. As is well known, however, a Laspeyres index overstates changes in the cost of living when consumer substitution occurs in response to changes in relative prices; hence, alternative formulas that do capture substitution behavior can provide a more accurate approximation to a COLI. Examples include the Törnqvist and Fisher ideal indexes (both members of the “superlative” class of index numbers defined by Diewert, 1976), which employ aggregation weights derived from quantities purchased in both the initial and final periods of the comparison. Although the theory is not as well developed as that for consumer expenditures, similar justifications for commonly employed superlative aggregation formulas may exist for broader measures of output prices as well (for example, see Diewert, 1983, for a production-based interpretation of a Törnqvist index).

In addition, statistical agencies often make use of “chaining” (Fisher, 1911, ch. 10; Forsyth and Fowler, 1981) when constructing long time series of inflation rates; with this procedure, the price changes implied by a sequence of indexes defined over various subperiods are “chained” or cumulated together. In the COGI

---

1 In common parlance, the terms “chain price index” and “superlative index” are often used interchangeably. Here, we are referring to “chaining” in its strictly defined sense.
context, chaining carries an intuitive or pragmatic appeal inasmuch as it ensures that the basket being priced will remain reasonably representative of actual consumption patterns over time. However, chaining by itself cannot correctly capture consumer substitution. (Feenstra and Shapiro, 2003 and Szulc, 1983 consider other problems that can arise with chained indexes.)

In many circumstances, price indexes must be constructed in the absence of timely data on expenditures. A superlative aggregation formula cannot be used in real time in these cases (indeed, the fact that the Laspeyres index only requires expenditures from an earlier base period accounts for much of its appeal). A compromise procedure, which only requires base-period expenditure data, involves using a weighted constant elasticity of substitution (CES) aggregator (this includes the weighted geometric mean—which measures the cost of living when utility takes a Cobb-Douglas form—as a special case). Based on historical evidence, one could form a judgment about the likely degree of substitutability across items and then use an appropriately calibrated CES formula (Shapiro and Wilcox, 1997). Such a procedure is now employed by the U.S. CPI to aggregate individual prices (that is, prices within item-area strata), with a geometric means formula used for the majority of cases and a Laspeyres formula reserved for strata where substitution is deemed unlikely a priori.

Accurately capturing substitution behavior is not the only relevant issue for choosing an aggregation formula. Statistical agencies typically measure a sample of prices (where the number of price quotes for a given subindex may be quite small), and commonly used formulas can differ in their susceptibility to small-sample biases. Indeed, Bradley (2001) has argued that small-sample bias in Laspeyres indexes—not a failure to capture substitution across categories of goods—accounts for most of the observed difference between the published (Laspeyres) version of the U.S. CPI and a superlative (Törnqvist) variant.

At least two other issues arise in choosing how to combine individual price changes into a measure of overall inflation. First, the weights selected for use in
aggregation can reflect explicit or implicit judgments as to which agents are to be represented in the index. By employing aggregate expenditure weights, the typical consumer price measure in effect gives a larger weight to the inflation rates faced by richer households—a so-called “plutocratic” weighting scheme. (Alternatively, we could compute the simple average of each household-specific inflation rate; this “democratic” weighting scheme might be more representative of a “typical” household’s experience.) For some purposes, one might also explicitly choose to measure the inflation rate faced by a particular segment of the population, such as wage earners, the poor, or the elderly.

Second, correct measurement of the quantities used in aggregation is critical. To the extent that these are subject to measurement error (as might occur if they are estimated from survey data), and to the extent that mismeasured weights are systematically associated with items that display above- or below-average price changes, the resulting aggregate inflation rate will be mismeasured. (Lebow and Rudd, 2003, present evidence of this in the U.S. CPI.)

4. The concept of core inflation
Core inflation was originally defined as “the trend rate of increase” of either “the price of aggregate supply” or “the cost of the factors of production” (Eckstein, 1981). More commonly, however, core inflation is understood in a statistical sense as corresponding either to “underlying inflation” (the portion of overall inflation that is free from transitory influences) or to a measure of the common trend in all prices. In line with its various definitions, core inflation can be measured in a variety of ways.

The most prevalent core inflation measures are “exclusion” measures that omit certain items—such as food and energy—from the calculation of overall inflation. The popularity of excluding food and energy derives in part from the experience of the 1970s and early 1980s, which saw sizable supply-driven price hikes for these items. Many prices other than food and energy may move erratically as well, however (indeed, some countries publish exclusion-based measures of core consumer price inflation that omit
housing, the effects of changes in indirect taxes, or other items). Thus, a variant on the exclusion approach involves adjusting the weight of items in inverse proportion to their variability (sometimes termed a “neo-Edgeworthian” index), so that items with erratic prices are downweighted rather than omitted entirely.

A second category of core inflation measures includes limited-influence measures such as medians or trimmed means (Bryan and Cecchetti, 1994). These measures exclude a certain proportion of the largest and smallest price changes each period (in the extreme case of the median, all items but one are excluded each period). In contrast to standard exclusion measures, however, the omitted items will vary period by period. Limited-influence measures sometimes do well in statistical exercises aimed at finding measures that are well correlated with long moving averages of headline inflation, or measures that can serve as good univariate predictors of headline inflation. However, for these limited-influence measures to capture underlying inflation well, true relative price changes must be small compared with transitory fluctuations (which will not always be the case). In addition, construction of these measures is often sensitive to the degree of disaggregation employed and to the length of time over which the individual price changes are measured.

A third set of approaches uses econometric techniques to estimate core inflation (variously defined). For example, in an econometric reduced-form Phillips curve (as was employed in Eckstein’s original study), lagged inflation terms can proxy for the persistent component of inflation once one controls for supply shocks and aggregate demand (the univariate analogue would involve taking simple or weighted averages of past inflation as the core inflation measure). Another approach is to use a dynamic factor model to extract a common component or “signal” from a set of disaggregated inflation rates (Bryan and Cecchetti, 1993). Other econometric approaches have been proposed as well (often invoking economic theory to provide their rationale)—for example, core inflation may be defined as the component of inflation that is uncorrelated with long-run economic activity (Quah and Vahey, 1995), or best correlated with money growth. Of course, these theory-
based underpinnings might be controversial; more generally, econometric approaches might be difficult to understand or communicate.

The neo-Edgeworthian, limited-influence, and dynamic factor approaches to measuring core inflation provide examples of an alternative “statistical” or “stochastic” approach to inflation measurement that has garnered increased interest in recent years (Wynne, 1997). Wynne contends that the economic basis for these inflation concepts is “some concept of ‘monetary’ inflation that…is not necessarily the same thing as changes in the cost of living.” If so, these alternative approaches will in principle imply different decisions about scope and aggregation relative to those implied by a COLI-based framework. In particular, to the extent that these measures seek to capture the portion of aggregate price movement that is attributable to changes in the supply of money, their relevant scope could be the price of any transaction that involves an exchange of money (including prices for financial assets and the purchase prices—not the user costs—for durable goods). In addition, the aggregation weights employed by these stochastic approaches are typically informed by purely statistical considerations, and so need not bear any resemblance to the weights implied by cost-of-living theory.

5. Implications for research and policy

Inflation measurement matters for at least three reasons. First, and most obviously, economic decisions often depend directly—even automatically—on published inflation measures. In the public sphere, many government programs are indexed to inflation measures such as changes in a consumer price index; in the U.S., for example, Social Security benefits, income tax schedules, and coupon payments on inflation-indexed government debt are all directly tied to changes in the CPI. Private contracts, including wage arrangements, are also indexed to changes in the CPI (although such indexation provisions are less common today than they were when inflation was higher and more uncertain).
The use of inflation measures in indexation arrangements, in principle, should help inform the details of inflation measurement. If indexation of a payment is intended to maintain its real purchasing power for a recipient, then this goal is best served by using an inflation measure tailored to that recipient. On this principle, indexation of pension payments would utilize an inflation measure that reflects the consumption patterns of pensioners; income-support payments would use a measure reflecting the consumption of the poor; and so on. Such specialized price indexes can differ from an aggregate price index in both the choice of priced items and in the weights assigned to them.

Inflation measurement is also important because inflation affects economic welfare and therefore serves as a goal of public policy in its own right—in particular, a central objective of monetary policymakers is the maintenance of low and stable inflation. Problems measuring the average level of inflation will therefore affect a central bank’s choice of inflation target (whether explicit or implicit). Indeed, many argue that the Federal Reserve should seek to stabilize measured inflation at some level higher than zero, in part because the U.S. CPI tends to overstate changes in the cost of living (for example, Bernanke, et al., 1999). More problematically, if measurement errors in inflation vary over time in unknown ways, central banks could respond inappropriately to movements in observed inflation rates.

Finally, because real quantities are typically estimated by deflating nominal values with a price index, inflation measurement directly affects the construction of other economic statistics (including real GDP and productivity). Thus, our ability to correctly assess the effects of technological progress, the sources of economic growth, and changes in living standards over time hinges in an obvious way on the accurate measurement of individual and aggregate price movements. Furthermore, if the extent of measurement error in inflation varies over time and across items or places, then growth comparisons could be affected; examples include measuring changes in living standards over long periods (Gordon, 2005) and comparing growth and productivity performance in the U.S. and Europe (Ahmad, et al., 2003).
References


