

IMF Working Paper

ARCHIVES
ROOM HQ C-532

0441

Relative Prices, Inflation and Core Inflation

Scott Roger

IMF Working Paper

Asia and Pacific Department

Relative Prices, Inflation and Core Inflation

Prepared by Scott Roger¹

Authorized for distribution by Kanitta Meesook

March 2000

Abstract

The views expressed in this Working Paper are those of the author(s) and do not necessarily represent those of the IMF or IMF policy. Working Papers describe research in progress by the author(s) and are published to elicit comments and to further debate.

Empirical evidence on the distribution of relative price changes almost invariably reveals high kurtosis and a tendency toward right-skewness. Simple mixed distribution models including volatile and infrequently adjusted prices can account for these and other common features, such as correlation between the mean and variance of relative prices. In such circumstances, robust measures of central tendency are likely to outperform the mean or standard measures of “core” inflation as indicators of generalized inflation. The analysis also supports the use of geometric averaging in CPI construction and the targeting of the geometric mean inflation rate rather than the Laspeyres mean.

JEL Classification Numbers: C43, E31, E52

Keywords: Inflation, relative prices, central tendency

Author's E-Mail Address: SROGER@IMF.ORG

¹The views expressed in the paper are my own and do not necessarily reflect the views of the IMF. I would like to thank, without incriminating, Francisco Nadal De Simone for helpful comments on an earlier draft.

	Contents	Page
I.	Introduction	3
II.	Empirical Evidence on the Distribution of Relative Prices	4
III.	Models of the Distribution of Price Changes	7
	A. A Model with "Volatile" Prices: Mixed Log-Normal Price Changes	7
	B. A Model of Infrequently Adjusted Prices	12
IV.	Implications for Core Inflation Measurement	14
	A. The Choice of the Target Measure of Central Tendency	14
	B. Estimation of the Core Inflation Rate	18
V.	Concluding Comments	24
	References	26
	Figures	
1.	Kurtosis of the Mixed Log-Normal Distribution	11
2.	Skewness of the Mixed Log-Normal Distribution	11
3.	Bias of the Arithmetic Mean of the Mixed Log-Normal Distribution Relative to the Geometric Mean (Measured in Standard Deviations)	12
4.	Skewness of the Distribution of Log-Normal Price Changes Mixed with Infrequently Adjusted Prices	15
5.	Kurtosis of the Distribution of Log-Normal Price Changes Mixed with Infrequently Adjusted Prices	15
6.	Bias of the Arithmetic Mean with respect to the Geometric Mean Inflation Rate	16
7.	Bias of the Median with respect to the Geometric Mean Inflation Rate	16
8.	Mean Percentile of the Mixed Log-Normal Distribution	23
9.	Efficiency of the Sample Mean Percentile Relative to the Sample Mean of the Mixed Log-Normal Distribution	23
	Appendices	
I.	The Mixed Normal Distribution, Kurtosis and the Efficiency of the Median Relative to the Mean	30
II.	Evolution of the Estimation of Central Tendency	37

I. INTRODUCTION

Central banks concerned with inflation performance almost invariably seek to distinguish between a generalized or persistent component of the measured aggregate rate of inflation—driven by demand pressures and expectations—and supply-driven, essentially transitory inflation developments associated with movements in relative prices. The more generalized, persistent element – usually described as “core” or “underlying” inflation - is regarded as a legitimate monetary policy concern, while the influence of specific relative price developments are generally regarded as transient “distortions” that should be largely ignored in forward-looking policy setting, as well as in backward-looking assessment of policy performance.

Typically, “distortions” to the aggregate inflation rate are attributed to one of two sources: (a) the influence of particularly “volatile” prices—such as petroleum or fresh food prices—largely determined by supply factors; and (b) the effects of infrequent price adjustments, most often involving prices set directly by the public sector or, at least, heavily influenced by government regulation or infrequently adjusted excise taxes and levies. In either case, the relative price movements are seen as distorting the aggregate inflation rate insofar as they do not reflect persistent or generalized demand pressures in the economy.

A review of empirical evidence indicates that, over the entire period in which price indices have been constructed and across many countries, regular features of the distribution of price changes have included chronic excess kurtosis, right-skewness (even in periods of low or negative inflation), and positive correlation between the mean and variance of relative price movements. Simple statistical models of price movements are then used to examine the impact of volatile and infrequently adjusted prices on the aggregate price index. These models generate results consistent with the empirical evidence.

A consequence of the high kurtosis associated with high volatility of some prices is that robust estimators—such as the median or trimmed mean measures—are likely to give a much more reliable indication of the general trend of inflation than the CPI mean. Such measures are also likely to be more reliable than measures of core inflation based on excluding or down-weighting prices of particular items.

Chronic skewness in the distribution of price movements—associated with arithmetic measurement of price changes and with infrequent price changes—leads different measures of central tendency to differ from one another. A simple method is suggested for designing a robust measure of core inflation that is unbiased with respect to the target measure.

The analysis also suggests that, instead of targeting a CPI inflation target with an adjustment for the bias of the Laspeyres mean with respect to a superlative index, it would be more appropriate to set the inflation target in terms of the geometric mean. This is in part because the bias of the Laspeyres mean is likely to be more strongly correlated with the variance of relative prices and, therefore, is also likely to be more variable.

II. EMPIRICAL EVIDENCE ON THE DISTRIBUTION OF RELATIVE PRICES

Ever since the first systematic efforts to construct aggregate or composite price indices in the 1860s, it has been observed that the cross-sectional distribution of price movements has not conformed to the Gaussian or Normal distribution.² Instead, the distribution has typically been found to be markedly leptokurtic (i.e., “high-peaked” or “long-tailed”) and more often positively than negatively skewed.

Jevons (1863) found that, over the period from 1782 to 1861, the distribution of annual price relatives for commodity prices in the United Kingdom was not typically symmetric but right-skewed.³ Consequently, Jevons advocated the use of geometric rather than arithmetic averaging in construction of aggregate price indices, since the distribution of logarithmic price changes was closer to Normal. Edgeworth (1887) also drew attention to an apparently systematic tendency towards right-skewness in the distribution of price movements. In addition to citing Jevons’ research, he referred to reports on prices in Massachusetts and Illinois in the 1880s, as well as to Laspeyres’ price indices for Germany.

Mitchell (1915) analyzed the distribution of pooled annual percentage changes in prices for over 200 goods making up the U.S. wholesale price index over the period from 1891 to 1913. Mitchell found that the distribution displayed high kurtosis and right-skewness. Fisher (1922, pp. 230–31, 408–10) also noted the right-skewness of the distribution of U.S. commodity price changes between 1913 and 1918. Mills’ (1927) extension of the analysis of U.S. wholesale prices to 1926 corroborated Mitchell’s findings. Mills summarized his findings with the observations that: “(a) The distributions of price relatives are...erratic...although Type IV of Pearson’s classification predominates; (b) The conditions which give rise to the Normal distribution are seldom realized in the distribution of price relatives...even when generous allowance is made for errors of sampling; and (c) There is some improvement...when logarithms of price relatives are used in place of relatives in natural form...” (p.368).⁴

²The term “Normal” is capitalized when the reference is to the statistical concept.

³The term “price relative” refers to the ratio of the price of a good in one period relative to the price in the previous period and is equal, therefore, to 1 plus the arithmetic rate of change of the price.

⁴Mills’ analysis included calculations of the coefficients of skewness and kurtosis, based on arithmetic and geometric, linked and unlinked, weighted and unweighted annual price relatives. For the (fixed base) arithmetic price relatives, the median coefficients of skewness and excess kurtosis were approximately 0.3 and 2.9, respectively. For chain-linked geometric price relatives, Mills found little skewness but persistent excess kurtosis: the median values of the coefficients of skewness and excess kurtosis were 0.0 and 1.9, respectively.

Keynes (1930) may not have been aware of the accumulation of evidence against Normality in the distribution of price changes but, in any event, was dismissive of the evidence he had come across. In Keynes' words, "If it were to be shown that the curve of dispersion [of price relatives] is of the same type in a great number of different contexts, then one would take notice; but the investigations, so far as they have gone, show nothing of the kind. It is worth mentioning, however, that M. Olivier and Prof. Bowley both conclude that, as a matter of curve-fitting, the geometric curve fits better, in the cases they have examined, than the arithmetic" (p. 85).⁵

Analysis of the distribution of price changes was revived by Vining and Elertowski (1976). Although their investigation focused primarily on the relationship between the mean and cross-sectional variance of price changes, they also found that the distribution of U.S. wholesale and consumer annual price changes, over the 1948–74 period, was typically more kurtotic than a Normal distribution, though not typically right-skewed.⁶

Over the past twenty years, analyses of the distribution of consumer and producer price changes have proliferated. The revival of this area of research has been spurred by the increasing emphasis placed by central banks on the measurement and control of inflation. Studies have included:⁷

- Parks (1978) on Netherlands consumer prices, 1921–63, and U.S. consumption price deflators, 1930–75;
- Balk (1978, 1983) on Netherlands consumer and wholesale prices, 1952–75, and consumption deflators at different levels of aggregation, 1951–77;
- Fischer (1981) on U.S. consumption deflators, 1930–80; and Fischer (1982) on consumption deflators in the United States, 1950–79, and in West Germany, 1970–79;
- Buck and Gahlen (1983) on West German producer prices, 1952–77;
- Blejer (1983) on Argentine consumer prices, 1977–81;
- Assarsson (1984) on Swedish consumer prices, 1951–79;
- Mizon, Safford and Thomas (1989) on U.K. retail prices, 1962–83;
- Ball and Mankiw (1992) on U.S. producer prices, 1948–89;

⁵Maurice Olivier's analysis of French wholesale prices between 1920 and 1924, in *Les Nombres Indices de la Variation des Prix*, is also noted by Mills (1927, p. 338). According to Mills, "Dr. Olivier finds that the distributions of these [wholesale] price relatives in natural form are far removed from the Normal type. Distributions of logarithms of these relatives are closer to the Gaussian form, but remain distinctly more peaked than the Normal curve."

⁶In the case of wholesale prices, the median annual (unweighted) values of skewness and kurtosis were -0.3 and 11.8, respectively, while for consumer prices the median values were -0.2 and 5.2, respectively.

⁷See also Marquez and Vining (1984), pp. 10–13.

- Lourenco and Gruen (1995) on Australian producer prices, 1970–92;
- Roger (1995, 1997) on New Zealand consumer prices, 1981–95 (at a disaggregated level) and 1949–96 (at a fairly aggregated level);
- Bryan and Cecchetti (1996) on U.S. producer prices, 1947–95, and consumer prices, 1967–96;
- Shiratsuka (1997) on Japanese consumer prices, 1971–96;
- Taillon (1997) on Canadian consumer prices, 1985–96;
- Amano and Macklem (1997) on Canadian producer prices, 1962–94;
- Bakhshi and Yates (1997) on U.K. retail prices, 1974–97;
- Kearns (1999) on Australian consumer prices, 1980–98.

The investigations, spanning roughly two centuries of data from a range of small and large economies, with quite different structural characteristics, different exchange rate regimes, and in the context of marked differences in average inflation, show a number of regular features:

- The cross-sectional distribution of price changes almost always shows excess kurtosis.
- Distributions of price changes are often asymmetric, with arithmetic price changes tending to show a bias towards right-skewness, while logarithmic price changes show greater symmetry.
- Sample skewness and kurtosis tends to be lower when prices are aggregated either cross-sectionally or over time.
- Skewness and kurtosis tend to be more pronounced across prices of quite different items than within more homogeneous groupings.
- Measured skewness and kurtosis can differ substantially according to whether the price changes are equally or unequally weighted.
- The mean, variance, skewness and kurtosis of cross-sectional price changes tend to be positively correlated.

Alternative theoretical explanations for some of these features, particularly the relations between the mean rate of inflation and the variance and skewness of price relatives, are surveyed in Fischer (1981, 1982), Cuckierman (1983), and Assarsson (1984). Fischer distinguishes between three broad categories of explanations. First are models in which asymmetry in the distribution of price relatives arises from the existence of some form of “menu” cost of price adjustment that generates “stickiness” in otherwise symmetrically flexible nominal prices. Second are models in which, even in the absence of menu costs, prices are asymmetrically sticky for other reasons such as, for example, asymmetric macroeconomic policy rules. Third are models in which imperfect information or misperceptions can lead to asymmetric price changes, even if prices are otherwise symmetrically flexible and menu costs are inconsequential.

Menu cost explanations for asymmetries in the distribution of price changes have received particular emphasis in recent years, following the work of Ball and Mankiw (1992, 1994).⁸ A particularly important implication of this type of model is that the chronic tendency towards right-skewness in the distribution of relative price movements is a consequence of positive trend or generalized inflation. Firms wishing to raise their prices relative to the general price level must raise their nominal prices, whereas firms wishing to lower their relative prices can hold nominal prices unchanged and let positive generalized inflation erode their relative prices. This implies that the degree of chronic skewness in relative price changes will tend to fall as trend or generalized inflation is reduced. Indeed, skewness should turn negative if there is trend deflation.

Unfortunately, this implication of the menu cost approach does not appear to be consistent with the historical record. Right-skewness appears to be a chronic phenomenon, occurring in periods of high inflation, low inflation and even deflation.

The issue of whether right-skewness in the distribution of relative prices is truly chronic, or simply a product of positive trend inflation, is important for monetary policy and particularly for the measurement of core inflation. If right-skewness is chronic, it implies that relative price variability and the mean rate of inflation will be positively correlated rather than independent.

III. MODELS OF THE DISTRIBUTION OF PRICE CHANGES

In this section, two basic models of the distribution of price changes are outlined to examine the implications of volatile prices and infrequently adjusted prices for the moments of the distribution of price changes and measurement of the aggregate rate of inflation.

A. A Model with "Volatile" Prices: Mixed Log-Normal Price Changes

This model focuses on the presence of some prices characterized by systematically higher volatility than other prices in the aggregate distribution. For analytical convenience, it is assumed that the aggregate *population* distribution of price movements is composed of two sub-distributions, each of which is log-Normal with identical means (See Appendix I). The only difference between the two distributions is that one is characterized by a higher variance than the other. This may arise from (a) one group of price movements exhibiting systematically greater stochastic variation relative to the common mean; and/or (b) one group exhibiting systematically greater dispersion in trend price movements relative to the common mean. In either case, the aggregate distribution of price changes will be a mixed log-Normal distribution.

⁸For a survey of the menu cost literature, see Cassino (1995).

The use of the mixed Normal distribution has a long history. Some of the earliest formal uses of the mixed Normal distribution to approximate leptokurtic distributions (i.e., distributions displaying higher kurtosis than the Normal distribution) appear to have been by Newcomb (1886) and Pearson (1894).⁹ Yule (1911) used the mixed Normal distribution to illustrate analytically the impact of excess kurtosis on the efficiency of the sample mean relative to the sample median. In the 1940s and 1950s Tukey et al used the model (described as a “contaminated” Normal distribution) to examine the issue of robust statistical inference, as did Andrews et al (1972).¹⁰ More recently, Bryan, Cecchetti, and Wiggins (1997) have used the model specifically in the context of the distribution of price changes.

The model in this paper differs from the Bryan, Cecchetti and Wiggins model in three respects. First, it generalizes their model by allowing explicitly for variations in the proportions of the two distributions. Second, it assumes that log price changes, rather than arithmetic price changes, are Normally distributed. Third, it allows explicitly for differences in relative price variances between the two groups to arise from differences in relative price trend rather than solely from stochastic relative price disturbances.¹¹

By construction, the composite population distribution of log price changes is symmetric, since both sub-distributions are symmetric. Beyond this, the shape of the aggregate distribution depends on two key parameters:

- The relative proportions or “mix” of the two distributions, described by the term α , representing the total weight in the aggregate price index of prices characterized by high variance or dispersion around the mean.
- The variability of the high variance group of price changes relative to the lower variance group, described by the term λ , defined as the ratio of the standard deviations of the two sub-distributions.

Loosely speaking, the parameter α governs the “peakedness” of the distribution, while the parameter λ governs the length of the tails.

The symmetry of the sub-distributions, of course, implies that the composite or mixed distribution of log price changes will have zero-skewness. But, as shown in Appendix I, the

⁹See Stigler (1973) for a discussion of Newcomb’s work. In Pearson’s use of the mixed normal model, he did not assume equal means for the two component distributions, so that the mixed distribution displayed asymmetry as well as excess kurtosis.

¹⁰See Harter (1975).

¹¹The inclusion of specific trends in relative prices is not an essential feature of the model, but is included to emphasize that standard measures of variations in relative prices include both persistent and transient elements.

kurtosis of the composite distribution will depend on the values of both α and λ , and will always equal or exceed the kurtosis of a Normal distribution.

Standard aggregate price measures, however, are based on arithmetic, not geometric, indices.¹² If the population distribution of price changes is mixed log-Normal, this will have important consequences for the arithmetic mean and the correlation between the mean and variance of price changes. Fundamentally, this arises from the simple fact that the arithmetic price changes corresponding to offsetting positive and negative log price changes are not offsetting, and the discrepancy increases with the magnitudes of the log price changes. As a result, contamination of a Normal distribution, with even a few large log-Normal price changes, will lead to significant skewness in the arithmetic price change distribution.

To illustrate the impact of the mixed log-Normal distribution on the distribution of arithmetic price changes, a numerical approach was used. This involved generating a sample of 15,000 log-Normally distributed price changes, then splitting the sample and varying the values of α and λ . The impact of variations in the values of α and λ on the kurtosis and skewness of the distribution of arithmetic price changes are shown in Figures 1 and 2.

Figures 1 and 2 illustrate a number of important points:

- In the special case in which the composite distribution is log-Normal (i.e. when $\lambda = 1$ or $\alpha = 0$), kurtosis will be almost equal to that of the Normal distribution and skewness will be very close to zero. This is the distribution implied by the Lucas (1973) "islands" model of price changes, and is clearly inconsistent with the weight of empirical evidence, suggesting that the Lucas model should not be thought of as a literal description of the distribution of price movements.
- The addition of even a very small proportion of prices characterized by high volatility or high trend dispersion dramatically increases the kurtosis and skewness of the composite distribution of arithmetic price changes.¹³
- Positive skewness is a chronic characteristic of the distribution of arithmetic price changes, independent of the prevailing inflation rate. This result contrasts with the result in menu cost models, in which right-skewness is only chronic with positive trend

¹²In recent years, it should be noted that the statistical agencies in an increasing number of countries, including Canada, France, the United States, and Australia, have begun (or indicated an intention to begin) using geometric averages for a significant proportion of prices below the basic level of aggregation. In Sweden a close approximation to geometric averaging is also being used.

¹³The distribution of log price changes will also be highly kurtotic, as shown in Appendix I, but symmetric.

inflation; indeed, in the Ball and Mankiw (1992) model, left-skewness should be associated with a falling trend of prices.

The presence of positive skewness in the distribution of arithmetic price changes (or “natural” price relatives) has two particularly important implications:

- In the symmetric distribution of log price changes, the population mean and median coincide. In the right-skewed distribution of arithmetic price changes, the population mean will exceed the median. The median will continue to closely approximate the geometric mean.¹⁴
- In contrast with the Normal distribution, in which the mean and variance of the distribution are independent, right-skewness in the distribution of arithmetic price changes leads to chronic positive correlation between the mean and the cross-sectional variance of price changes, with the correlation increasing with the degree of skewness.¹⁵

Figure 3 shows that the displacement or bias of the Laspeyres arithmetic mean, relative to the geometric mean, may be substantial depending on the values of α and λ . It should also be noted that the displacement of the mean is measured in terms of standard deviations since, as noted above, the value of the mean and the variance of the distribution are not independent.

¹⁴In a current weighted index, the arithmetic median would coincide exactly with the geometric mean, because the order of price changes is unaffected by log transformation of the prices. As a result, the geometric median and arithmetic median will be identical. With a Laspeyres price index, however, the evolution of the price weights differs between the arithmetic and geometric measures. This can lead to some difference between the arithmetic and geometric medians if the implicit weight shifts in the arithmetic index do not cancel out above and below the median.

¹⁵See Cramér (1946), pp. 348–349. It may be noted that in this model, an increase in the variance of the high variance prices will increase the variance of the aggregate distribution as well as its skewness, unambiguously raising the arithmetic mean relative to the geometric mean and the median. A rise in the variance of the low variance prices will have a more ambiguous effect, since skewness will tend to decline.

Figure 1. Kurtosis of the Mixed Log-Normal Distribution

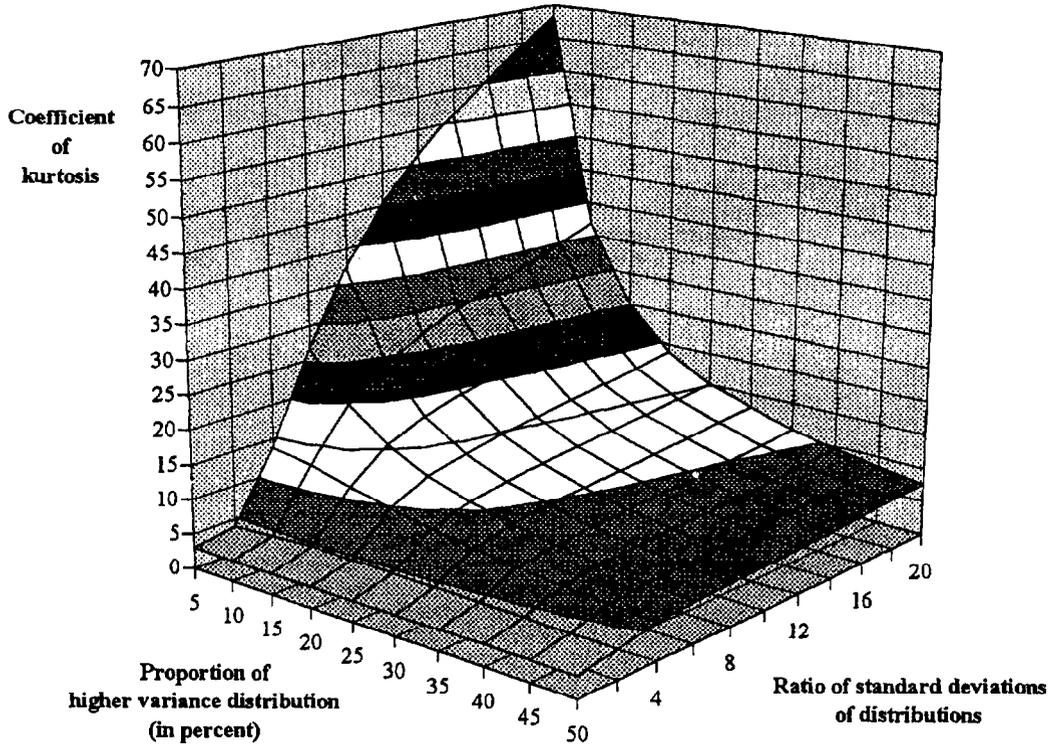
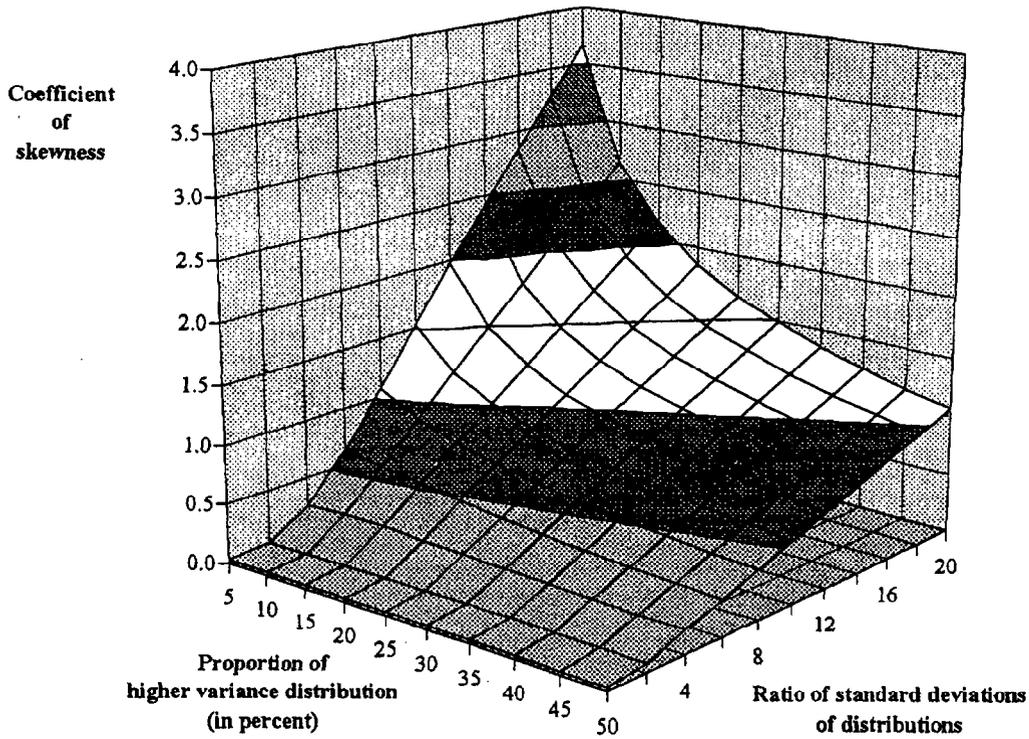
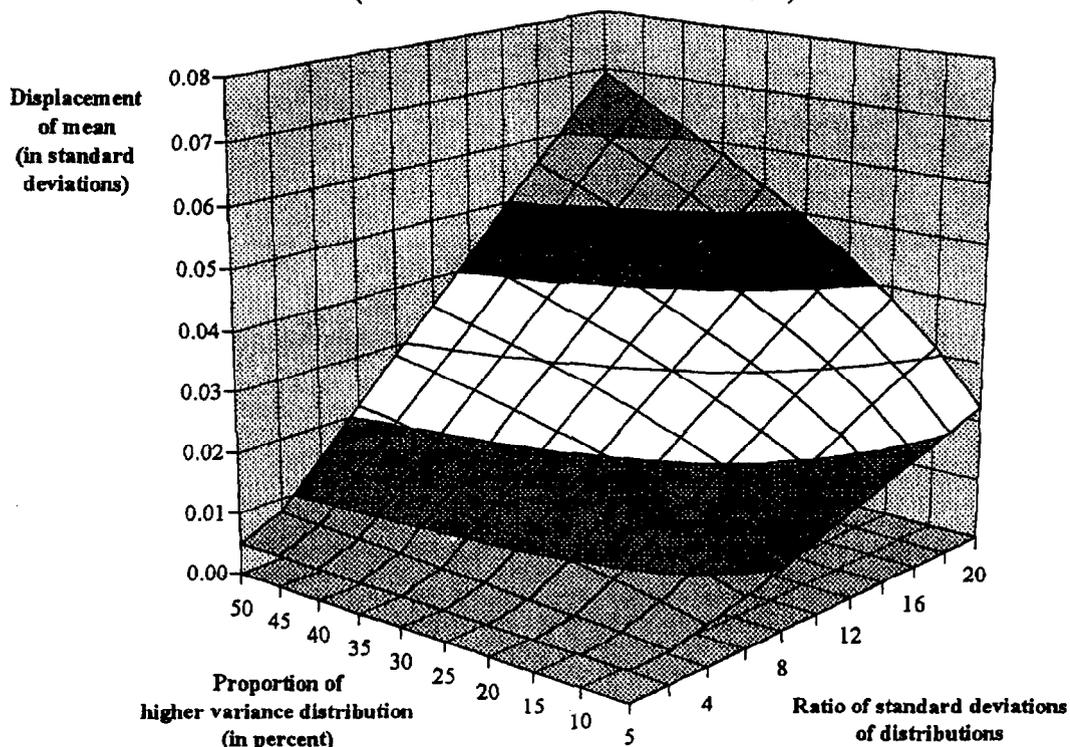


Figure 2. Skewness of the Mixed Log-Normal Distribution



**Figure 3. Bias of the Arithmetic Mean of the Mixed Log-Normal Distribution
Relative to the Geometric Mean
(Measured in Standard Deviations)**



B. A Model of Infrequently Adjusted Prices

In this model the focus of attention is on prices that, for a variety of reasons, may be adjusted infrequently. Infrequent price adjustment appears, empirically, to be common for:

- Government-set or regulated prices. A variety of government excise taxes, as well as public sector fees and levies, are commonly adjusted only once or, at most, a few times per year.
- Seasonal goods. Prices for articles such as skis, bathing suits and some fresh produce may not be available at some times of the year. In this case, the compilers of the CPI may show unchanged prices for such goods in those months.
- Infrequently sampled prices. Even if the CPI is published at the monthly or quarterly frequency, some prices may be sampled less frequently if they are costly to collect. In the CPI the prices of such items may show no change in the non-sampling months.

To allow for such price adjustments, we again employ a composite distribution. A proportion $(1-\alpha)$ of price changes are log-normally distributed, as in the previous model. The remaining

α prices are assumed to have the same trend inflation rate, but to adjust only once every four CPI measurement periods.¹⁶ On average, therefore, three-quarters of the prices will show no change in a given CPI measurement period, while one quarter will rise by four times the trend rate, plus a log-normally distributed error term. That is:

$$\pi_i = \begin{cases} 0 & \text{for } i = 1 \dots 3j/4 \\ 4\Pi + \varepsilon & \text{for } i = 3j/4 \dots j \\ \Pi + \varepsilon & \text{for } i = j+1 \dots n \end{cases}$$

where:

$$\pi_i = \ln(p_{it}/p_{it-1})$$

$$\Pi = \sum_{i=1}^n w_i \pi_i$$

$$\varepsilon \sim IN(0, \sigma^2)$$

$$\sum_{i=1}^j w_i = \alpha \quad \text{and} \quad \sum_{i=1}^n w_i = 1$$

It is worth emphasizing that the nature of price “stickiness” implicit in this model is not the same as the sort of price stickiness that is involved in standard menu cost models à la Ball and Mankiw. This is not meant to suggest that such stickiness does not occur. Nonetheless, even at the basic level of CPI aggregation, there is already a significant degree of aggregation across goods, producers, wholesalers and retailers so that any such price inertia at the individual producer level tends to be lost from view. The kinds of infrequent price adjustment of concern in this section, however, are of concern precisely because they are commonly evident in CPI data at even a relatively aggregated level.

With a positive trend rate of inflation, the mixed distribution of log price changes will include a Normal distribution of flexible prices together with a substantial spike at the zero-price change frequency and a smaller, Normally distributed “blip” centered at four times the periodic trend inflation rate.¹⁷ The higher the trend rate of inflation, the farther out in the tail of the distribution the “blip” will occur, magnifying the skewness of the distribution. Symmetrically, for a negative trend inflation rate, skewness will be negative. For the special case of a zero-trend inflation rate, skewness in the distribution of log price changes will be

¹⁶For Australia and New Zealand, with quarterly CPIs, the implicit assumption is that such prices are typically adjusted once per year. For countries with monthly CPIs, it might be more reasonable to assume that such prices are adjusted only once every 12 periods.

¹⁷In the standard menu cost model, there will be a spike at zero-price change, but no “blip” towards the other end of the distribution, since the “sticky” prices do not adjust periodically to restore relative prices.

zero, but the distribution will show a large spike at the zero-price change frequency, generating high kurtosis.

The skewness and kurtosis of the distribution of arithmetic price changes will be essentially similar to those of log price changes, though slightly asymmetric around the zero-trend inflation point. The skewness and kurtosis of arithmetic price changes, for different positive trend inflation rates, and different proportions (α) of infrequently adjusted prices, are shown below in Figures 4 and 5.

The asymmetry of the distribution of log price changes in the infrequently adjusted price model has important implications for the mean and median:

- As in the previous model, the arithmetic mean will be upward biased with respect to the geometric mean for positive inflation rates. In addition, however, the degree of bias will increase with the geometric inflation rate, as shown in Figure 6.
- In contrast with the previous model, the median will no longer coincide with the geometric inflation rate; instead it will be downward biased (for positive trend inflation) and the magnitude of the bias will increase with the level of the geometric inflation, as shown in Figure 7.

IV. IMPLICATIONS FOR CORE INFLATION MEASUREMENT

In the stochastic approach to core inflation measurement, the observed distribution of price movements in any given period is thought of as a *sample* drawn from an underlying *population* distribution.¹⁸ Some measure of central tendency of the population distribution is thought of as the “core” inflation rate. The same measure of central tendency in the observed distribution of price movements in any given period, however, may be distorted by the effects of relative price disturbances in that period. The challenge in the stochastic approach is to obtain a good estimate of the measure of central tendency in the population distribution from the observed sample distribution. Two basic issues arise. The first is to choose the target measure of central tendency in the population distribution. The second is how best to infer that value from the observed data.

A. The Choice of the Target Measure of Central Tendency

The Fisher Ideal and Törnqvist-Theil indices are today widely regarded as the best price index formulae. Indeed, Diewert (1976) characterized these indices as “superlative”. One of the properties of these indices is that they pass the time reversal test; that is, if the index is calculated using price relatives moving forward in time, this should be the reciprocal of the

¹⁸See, e.g. Clements and Izan (1987), Diewert (1995), or Roger (1998).

Figure 4. Skewness of the Distribution of Log-Normal Price Changes Mixed with Infrequently Adjusted Prices

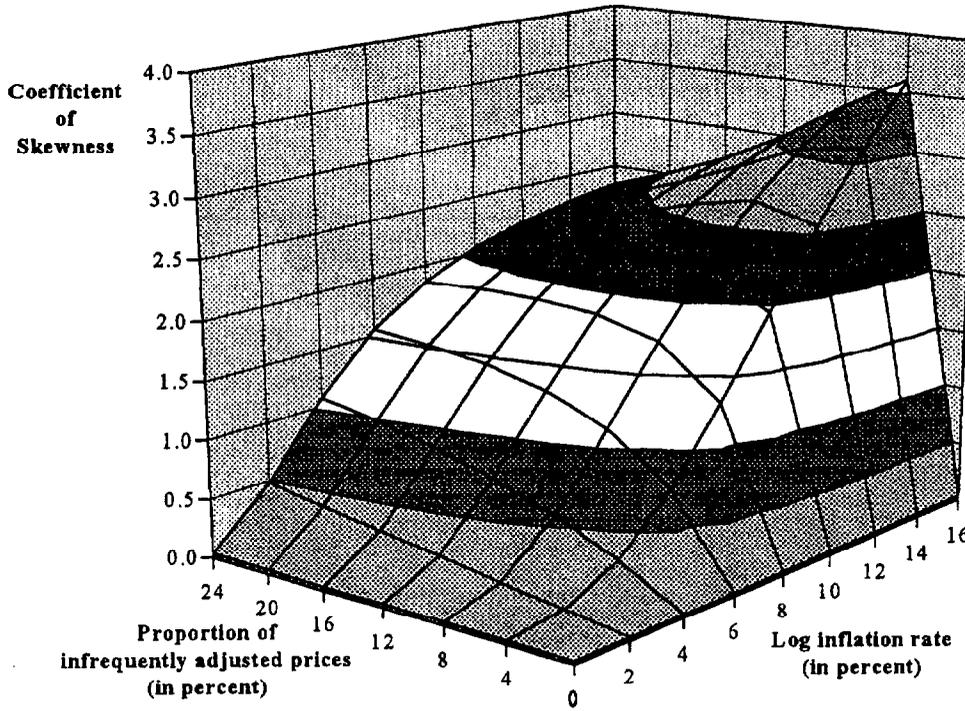


Figure 5. Kurtosis of the Distribution of Log-Normal Price Changes Mixed with Infrequently Adjusted Prices

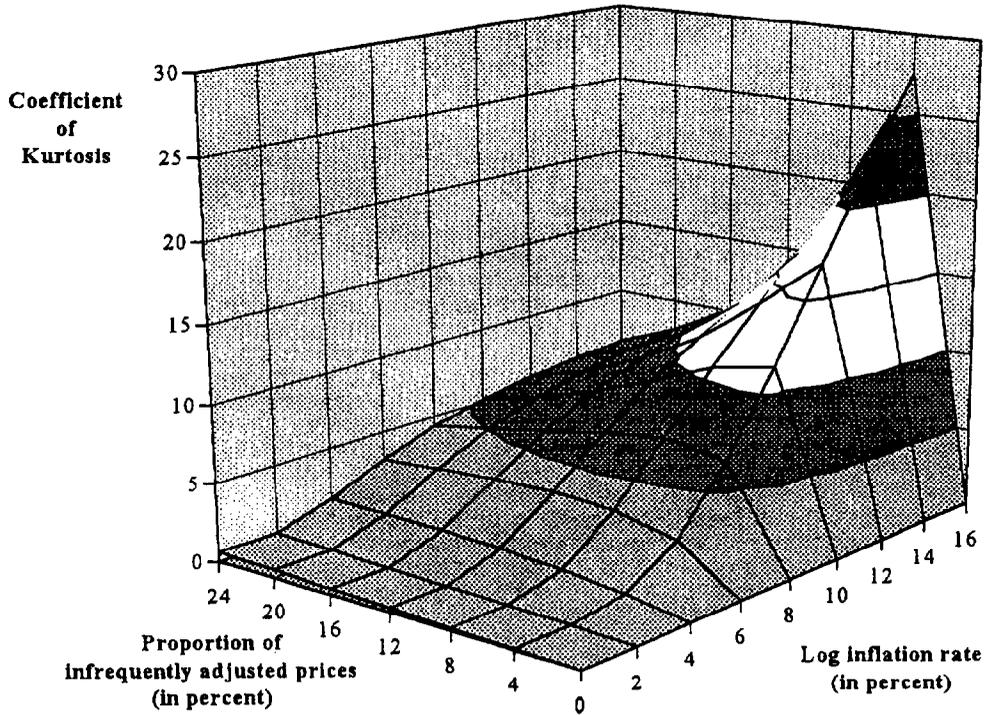


Figure 6. Bias of the Arithmetic Mean with Respect to the Geometric Mean Inflation Rate

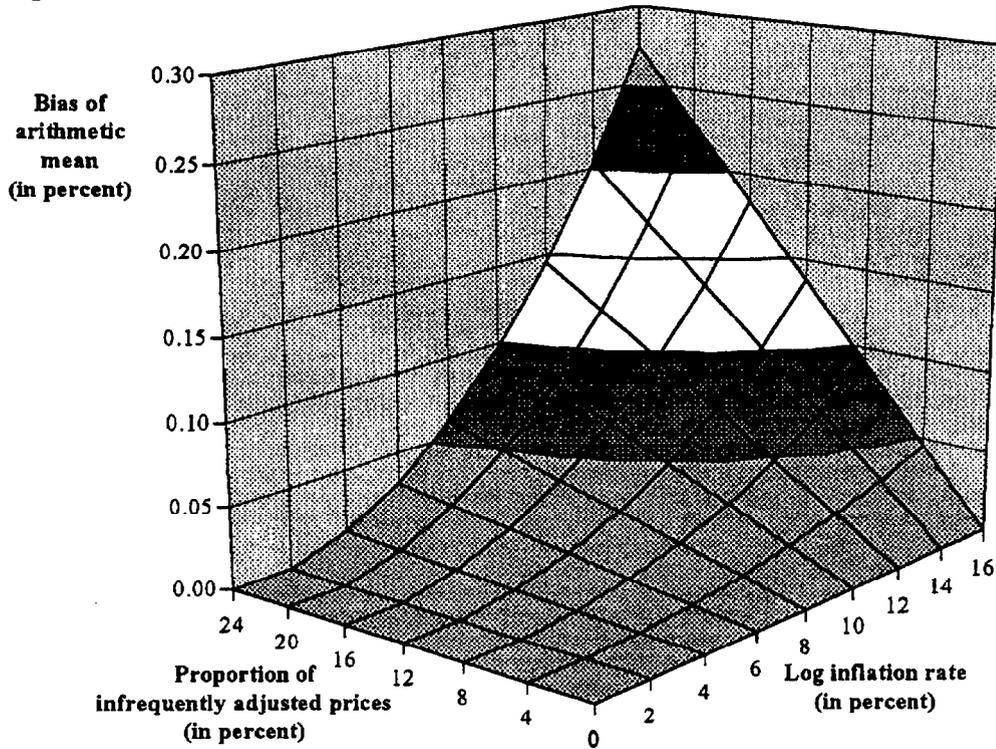
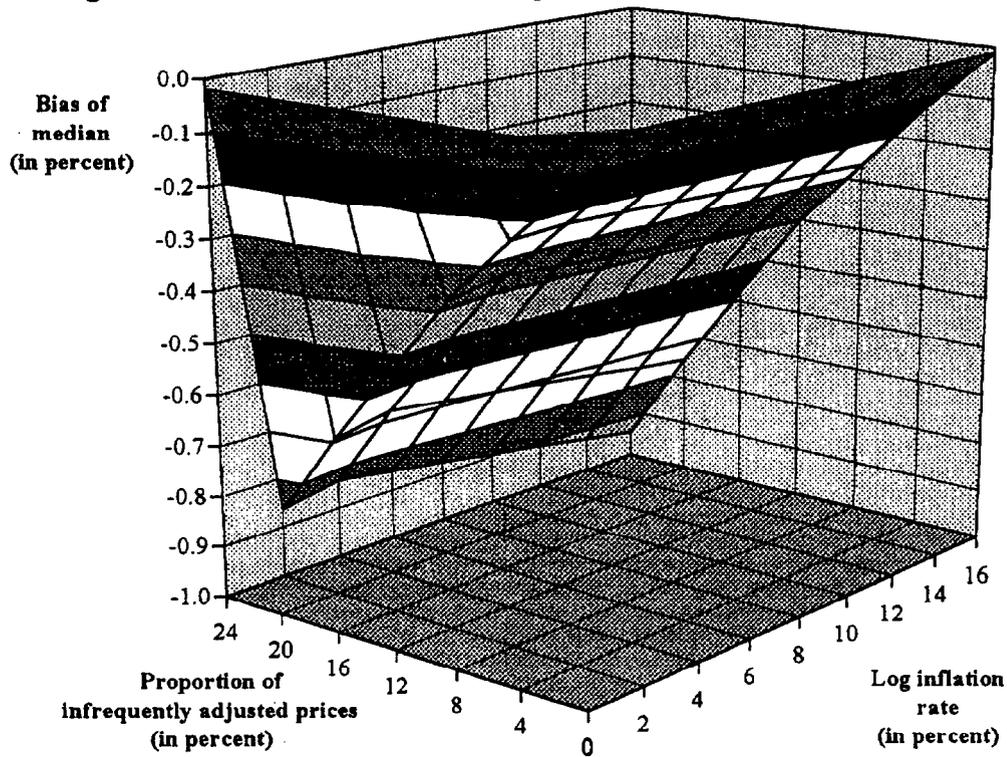


Figure 7. Bias of the Median with Respect to the Geometric Mean Inflation Rate



index using price relatives moving backward in time.¹⁹ An important implication, stressed by Fisher (1922), is that such indices will be unaffected by asymmetry in the distribution of prices.²⁰

Fisher's logic is unassailable, but ignores the reality that the practical choices to be made are not between the Fisher Ideal or Törnqvist-Theil indices, which require both base- and end-period weights to construct, but between different base-period weighted indices which all fail the time reversal test in some degree. In this context, Fisher's own analysis indicates that the Laspeyres (base-period weighted) arithmetic index fails the time reversal test miserably. Indeed, Fisher's graphical presentation (p. 409) shows quite clearly that, for the right-skewed distribution, the Laspeyres arithmetic mean is strongly affected by the skewness of the distribution, while the median and geometric mean are much less affected.

Shapiro and Wilcox (1997) shed useful empirical light on the issue. Using U.S. CPI data for the 1987-95 period, the authors calculate Laspeyres, Fisher Ideal, Törnqvist-Theil and geometric indices. It is found that the geometric index tends to be downward biased relative to the Fisher Ideal and Törnqvist-Theil indices (which virtually coincide), while the Laspeyres index tends to be upward biased, with the bias in the Laspeyres index being about 50 percent larger than that of the geometric index. The authors also construct a CES type index so as to allow for elasticities of substitution intermediate between the value of zero (implicit in the Laspeyres index) and one (implicit in the geometric index). It is found that using a value of 0.7 yields an index that is unbiased with respect to the superlative indices, which again points to the geometric index being somewhat closer to the superlative indices than is the Laspeyres index.

The above discussion suggests that, for practical purposes, there are three main options regarding which measure of central tendency of the population distribution to focus on:

- The Laspeyres mean, essentially because that is the standard formula for CPI calculation, despite its upward bias which, as shown earlier, may be substantial.
- The (weighted) geometric mean, either on the basis that it is a natural measure of central tendency if price changes are approximately log-Normally distributed, or because it is a closer approximation of the superlative indices than is the Laspeyres mean.
- Some point in between the Laspeyres and geometric means that may more closely approximate the superlative indices.

From a monetary policy perspective, each of these alternatives has pros and cons:

¹⁹See Diewert (1987) for an overview of properties of various price indices.

²⁰Fisher thus rejected Jevons' and Edgeworth's arguments, based on their concerns regarding skewness in the distribution of price relatives, for the geometric mean and median, respectively.

- Targeting the Laspeyres arithmetic mean has the advantage of consistency with the official CPI, but also has obvious drawbacks. The recognition that the Laspeyres index is upward biased leads central banks to define “price stability” as a positive measured inflation rate, partly in order to accommodate this bias. Quite apart from the public relations challenge involved, this approach forces the central bank to make some estimate of the magnitude of the bias. Yet, as shown earlier, the magnitude of the bias is not a fixed number, but depends on structural economic factors affecting the shape and spread of the distribution of price changes. As a result, if an unbiased measure of inflation is the ultimate target of policy, the Laspeyres measure of inflation may not be a very reliable intermediate objective.
- If the inflation target is set in terms of the (weighted) geometric mean, this should be a closer approximation to an ideal index. The main drawback is that the CPI inflation rate, based on the Laspeyres index, would tend to consistently exceed the geometric mean. Once again the central bank would face the challenge of explaining to the public why price “stability” implies a positive Laspeyres measure of inflation. Undoubtedly, the credibility of a target defined in terms of the weighted geometric mean would benefit if the national statistics agency were to calculate and publish such an index alongside the Laspeyres index.
- If an approximation to a superlative price index were to be the target, the central bank would face similar public relations challenges as with the adoption of a target for the geometric mean. In addition, however, it would need to find a suitable approximation, perhaps using a weighted average of the Laspeyres and geometric means.²¹ It would likely be even more important for the national statistics agency to calculate the true superlative index, even on an *ex post* basis.

B. Estimation of the Core Inflation Rate

Whichever measure of central tendency in the population distribution of price movements is selected as the policy target, the next issue that arises is how best to estimate this from the sample distribution of price movements observed in each month or quarter. If the underlying population distribution of price movements is symmetric, then the choice of estimator mainly involves considering the robustness and efficiency of alternative estimators. If the distribution is also asymmetric, then attention must also be paid to bias in the estimator.

²¹The appropriate weights will vary from country to country and over time. Since most countries appear to have price distributions more skewed than that of the United States, it is likely that most countries would need to place greater weight on the geometric mean than Shapiro and Wilcox (1997) suggest for the United States.

Estimation with a Symmetric Population

The most appropriate estimator of the mean of a symmetric population distribution depends largely, though not totally, on the kurtosis of the distribution. Carl Gauss long ago showed that the sample mean is the best (lowest variance) linear unbiased estimator of the population mean *if* the population distribution is Normal.²² However, if the distribution is symmetric but is characterized by higher kurtosis than the Normal distribution, then, as shown in Appendix I, the sample mean will be a much less efficient estimator of the population mean than is the sample median. Essentially this stems from the fact that the median is much less affected by extreme price movements than the mean, and such outliers are far more common when the population distribution is highly kurtotic distribution than when it is Normal.

The median and the mean can both be thought of as extremes in the class of trimmed mean estimators, with the mean being the 0 percent trimmed mean and the median being the 50 percent trimmed mean. Analytically it is easier to compare the relative efficiency of these extreme cases than intermediate trimmed means. But that does not necessarily imply that the sample median will be the most efficient estimator of the population mean for every symmetric distribution showing excess kurtosis. Indeed, Bryan, Cecchetti and Wiggins (1997) find, using a model very similar to model 1 above (differing insofar as λ is fixed in their model), that the optimal trim may be as little as 5 to 10 percent. Basically, this result arises from the fact that, for the particular class of distributions considered, after trimming around 10 percent from each end of the distribution, the residual distribution is approximately Normal, so that the optimal estimator is the mean of the remainder.

When the precise shape of the distribution is not known, however, prudence suggests that the central bank may want to place less emphasis on the efficiency of a particular estimator for a particular population distribution and place more emphasis on the robustness of the estimator for a range of distributions. In other words, some sacrifice in terms of the expected average efficiency of the estimate of core inflation may be an acceptable price for reducing the risk of occasional bad errors. In such circumstances, the central bank may be advised to adopt a higher trim than might appear optimal in the case where the population distribution is known.²³ Bryan, Cecchetti and Wiggins' analysis, however, suggests that the efficiency cost of this extra robustness is likely to be quite small – the efficiency of even the median is only marginally lower than for the optimal trimmed mean; a result that seems likely to carry through for other classes of leptokurtic distributions.

²²See, e.g., Harter (1974) or Stigler (1986).

²³Monte Carlo testing of alternative estimators suggests that the median is more robust than smaller trims. It should be emphasized, however, that trimmed means (including the median) form only one class of robust estimators; some others, such as the Gastwirth mean, may be even more robust. See Harter (1975).

Although an increasing number of central banks monitor robust estimators of core inflation, the most common measures used remain based on either exclusion or down-weighting of particularly volatile items in the CPI. The volatile price model presented above suggests that such estimators are likely to be more efficient than the CPI mean. In the volatile price model, for example, if the volatile group of prices was excluded, the remaining distribution would be log-Normal, so that the sample geometric mean would be the most efficient estimator of the population mean. A more sophisticated approach involves re-weighting of prices in inverse proportion to their volatility. The main difficulty with these approaches is that they are not robust.²⁴ As shown in Appendix I, even a very slight departure from Normality in the distribution of price changes can dramatically reduce the efficiency of such re-weighted means relative to more robust estimators. The times when a central bank is most likely to want to have a measure of inflation that is robust to distortions of the distribution are precisely the times when the mean, with or without exclusions, is most likely to break down as a reliable estimator.²⁵

Dealing with Asymmetry

If the population distribution of price changes is asymmetric, then measures of central tendency in prices will no longer coincide. As a consequence, robust estimators based on the implicit assumption of symmetry in the population distribution will be chronically biased. For example, in the case of the mixed log-Normal distribution, the sample median will be an unbiased estimator of the population geometric mean, but a downward biased estimator of the arithmetic mean. If there are infrequently adjusted prices, then even the distribution of log price changes will be asymmetric, with the direction and magnitude of bias depending on the geometric mean inflation rate. The analysis in this paper also indicates that these biases may be large. This matters for two particular reasons. Obviously, if the estimator of the target measure of inflation is biased, it will be important for the central bank to take this into account in determining its policy reactions to developments in the measure or in the use of such a measure in policy accountability. In addition, differences in the bias of alternative estimators is also likely to distort evaluation of their comparative efficiency as measures of core inflation.²⁶

²⁴In addition, such measures may also be biased, since the prices excluded or re-weighted may have different trend rates of increase than the mean of other prices.

²⁵In practice, the breakdown in the efficiency of the mean would tend to be reflected in the central bank needing to explain why certain price movements were “distorting” even its official measure of core inflation.

²⁶To illustrate, consider two alternative estimators of P , the target measure of inflation:

$$x \sim N(P, \sigma_x^2) \text{ and } z \sim N(P + a, \sigma_z^2).$$

Calculating the variances, $V(x)$ and $V(z)$ of the two estimators around the target P we obtain:

(continued...)

The analysis in this paper suggests that the inherent asymmetry of infrequently adjusted prices is particularly problematic, insofar as it will generate asymmetry in any otherwise symmetric distribution, leading to bias in all symmetrically trimmed estimators. In many cases, however, the inertia in measured prices is not so much a real phenomenon, but a reflection of the way in which the statistics agency imputes missing or infrequently measured prices. Instead of assuming no change in the absence of a measured price, the statistics agency could assume no change in the price relative to the price of some like good or group of goods.

Inertia in the prices of many publicly-set or regulated prices is somewhat different. Such prices can be thought of as composed of a flexible "shadow" market price together with a system of implicit subsidies and excise taxes that generate the stepwise pattern of adjustment in the measured price. From this perspective, replacement of the zero rates of change in such prices with the rates of change in prices of comparable items should be regarded as replacing the posted prices with their shadow prices. Apart from reducing asymmetry in the distribution of prices, the use of shadow prices in place of regulated prices would have the important side-benefit of greatly reducing the temptation to resort to explicit or implicit indirect taxes and subsidies in order to manipulate the measure of inflation of particular concern to the central bank.²⁷

Adjustments of these kinds should substantially reduce asymmetry in the distribution of log price changes, but may not eliminate it altogether. Consequently, the median may remain a biased estimator of a geometric mean inflation target. If the arithmetic mean or one of the superlative measures is the policy target, then even if the distribution of log price changes is completely symmetric, the median (or any other trimmed mean) will be a biased estimator of the target.

For the trimmed mean class of estimators, including the median, the bias problem can be minimized fairly readily by trimming the distribution asymmetrically. In a symmetric

$$V(x) = \sigma_x^2 \text{ and } V(z) = (\alpha^2 + \sigma_z^2).$$

Now, even if $\sigma_x^2 > \sigma_z^2$, x will be a lower variance estimator than z of P if $\alpha^2 > (\sigma_x^2 - \sigma_z^2)$.

Yet, if the bias in z were eliminated, it would be the lower variance estimator. This problem may lead, for example, to lightly trimmed means showing lower variance (or RMSE) than more heavily trimmed means when using the arithmetic mean (or a moving average of CPI inflation) as the benchmark.

²⁷For example, in order to forestall a firming of monetary conditions in response to a pickup in market price inflation, the fiscal authority may be tempted (or even pressured by the monetary authority) to delay periodic adjustment of administered prices, with ultimately adverse consequences.

distribution, the population mean will coincide with the median or 50th percentile price change. In an asymmetric distribution the mean (whether of the Laspeyres or superlative kind) will correspond to a different percentile. If the distribution is right-skewed, the mean will correspond to a percentile of the distribution somewhere between the 50th and 100th percentiles. If this percentile can be determined, then the sample value of that percentile can be considered as an estimator for the population mean, just as the sample median is used as an estimator of the population mean in a symmetric distribution.

For convenience, I shall describe the value of the percentile corresponding to the population mean as the *mean percentile*. But it is probably more appropriate to describe the measure as the *Boscovich-Laplace median*. As discussed in Appendix II, the modern definition of the median, as the least absolute errors measure of central tendency, was established by Edgeworth about 1887.²⁸ Edgeworth's median, however, differed slightly from the definition originally proposed by Boscovich in 1757 and more thoroughly investigated by Laplace, beginning around 1774.²⁹ Both definitions of the median set as the primary condition the minimization of the sum of absolute errors. In the Boscovich-Laplace formulation, however, there is the additional criterion that the (probability weighted) sum of errors equals zero. The arithmetic mean possesses the zero sum of errors property, but the Edgeworth median does not. That difference, fundamentally, is why the Edgeworth median and the arithmetic mean of the population distribution do not necessarily coincide. The mean percentile is defined precisely to ensure that, on average, the sum of (arithmetic) errors sums to zero. By imposing this constraint, the mean percentile can, therefore, be regarded as analogous to the Boscovich-Laplace median.

The mean percentile of the distribution, however, depends on the shape of the distribution. As shown earlier, the displacement of the arithmetic mean from the geometric mean depends on the values of α and λ in the case of the mixed log-Normal distribution. Consequently, the value of the mean percentile will also vary with α and λ , as shown in Figure 8.

Moreover, just as the efficiency of the sample median relative to the sample mean of a symmetric population distribution varies with the shape of the distribution, so too will the relative efficiency of the sample mean percentile. Figure 9 shows how the relative efficiency of the sample mean percentile varies with the values of α and λ for the mixed log-Normal distribution. Because the values were generated from a sample distribution (albeit a large one), rather than analytically, the values evolve less smoothly than the data in Appendix I, Figure 2. Nonetheless, the basic message that comes through is that the mean percentile is a much more robust estimator of the population mean than is the sample mean. Although the

²⁸See Harter (1974, Part II), pp. 235–238, for a review of Edgeworth's development of the modern definition of the median.

²⁹See Harter (1974, Part I), pp. 147–52 for a review of the development of the Boscovich-Laplace median.

Figure 8. Mean Percentile of the Mixed Log-Normal Distribution

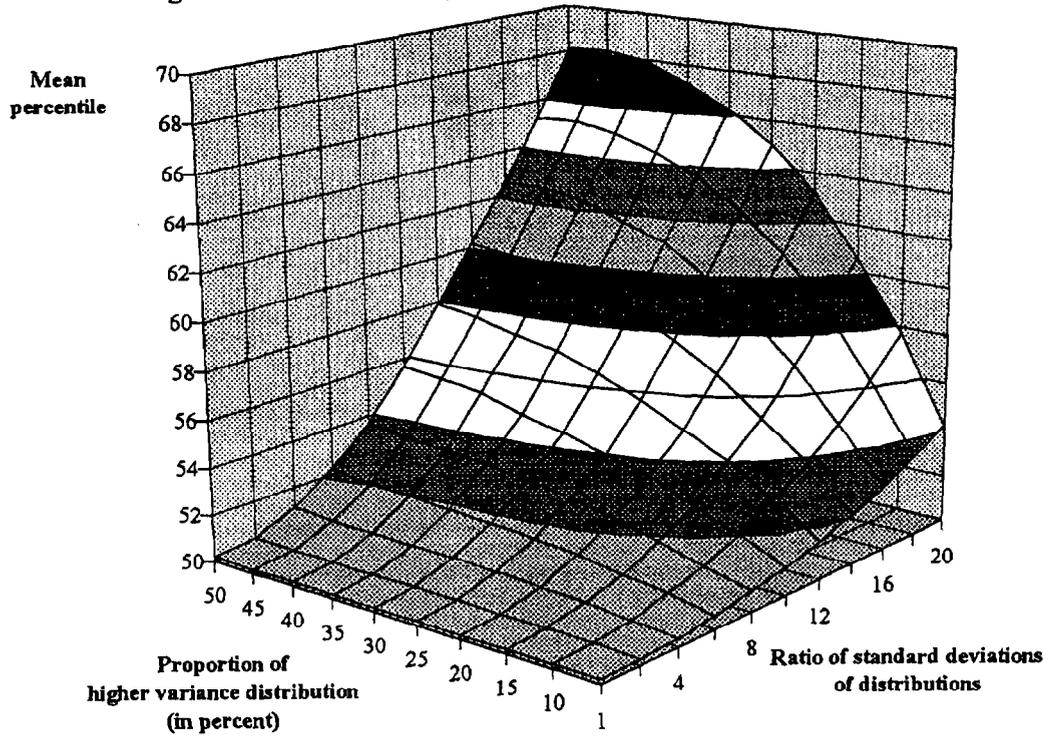
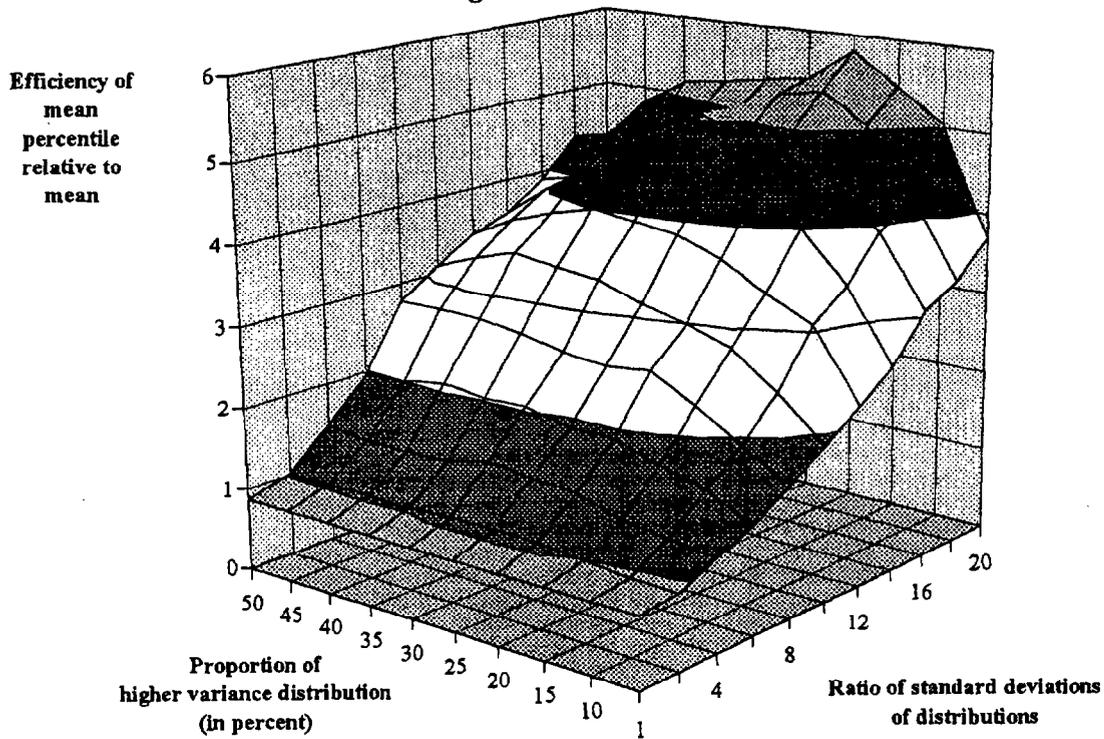


Figure 9. Efficiency of the Sample Mean Percentile Relative to the Sample Mean of the Mixed Log-Normal Distribution



relative efficiency of the mean percentile is not typically as high as in the case of a symmetric distribution, it is still almost always greater than one, and usually by a substantial margin.

The fact that the mean percentile depends on the shape of the distribution of price changes means that its value should be expected to vary over time, as structural factors influencing the parameters of the distribution change, and internationally, reflecting differences in structural factors and price measurement methodologies between countries.³⁰ Within the framework of the mixed Normal distribution, a number of particularly important influences on the values of α and λ may be noted:

- Increasing openness of the economy to trade should increase the elasticity of supply of some goods, reducing their volatility relative to other goods, lowering α .
- Changes in productivity differentials in different industries will tend to affect both α and λ . Increased economic openness could increase λ as productivity and inflation differentials between the traded and non-traded goods sectors of the economy widen. Deregulation and privatization of various industries would also affect α and λ , partly through the impact of reforms on trend productivity in such industries and, partly, through associated changes in price regulation and production subsidies and taxes.
- In federal states the degree of synchronization of adjustments in publicly set or regulated prices will tend to be less than in unitary states, resulting in smaller but more frequent movements in these prices in the CPI. As a consequence, the value of α will tend to be lower in federal states.

V. CONCLUDING COMMENTS

Non-Normality in the distribution of price movements has important implications both for the measurement of inflation in general and for the measurement of core inflation for use in monetary policy. If the characteristic distribution of price movements—whether measured in logarithmic terms or in arithmetic terms—was approximately Normal, the analysis in this paper suggests that both the Laspeyres mean and the geometric mean would tend to be only slightly biased approximations of superlative measures of aggregate price change. Moreover, as estimators of the underlying general trend of inflation of particular concern for monetary policy, the sample arithmetic or geometric mean would be essentially unbiased and more precise, on average, than alternative estimators.

³⁰For example, Roger (1997) finds the mean percentile for New Zealand data at around the 57th percentile; Kearns (1998) obtains the 52nd percentile for Australia; research at the Reserve Bank of Peru obtains the 61st percentile; and Bank Indonesia obtains the 64th percentile.

However, the evidence, historically and internationally, points to the distribution of price changes typically being characterized by high kurtosis and also by right-skewness. The analysis in this paper indicates that these characteristics may lead the Laspeyres mean to be a much more biased approximation to a superlative measure of inflation than is the geometric mean. On this basis, the increasing use by statistical agencies of geometric averaging below the basic level of CPI aggregation is strongly endorsed.³¹ The analysis also suggests that, for central banks setting inflation targets, it would be more appropriate to target the geometric mean than to target the Laspeyres mean, not simply because the Laspeyres mean is likely to be more biased with respect to a superlative measure, but also because the bias will tend to vary over time.

In addition, high kurtosis in the distribution of price changes suggests that robust estimators such as the median or trimmed mean may be far more dependable indicators of the general trend of inflation than either the arithmetic or geometric mean. If the distribution of price changes is asymmetric, perhaps as a result of infrequent adjustment in particular prices, the robust estimators themselves may need to be asymmetric to avoid bias. For statistical agencies, the analysis in this paper suggests that use of robust estimators below the basic level of CPI aggregation is warranted on traditional grounds of statistical inference, as protection against bad samples or errors in measurement and recording of prices. For central banks, the robust measurement methods offer a simple and readily verifiable means of screening out the impact of relative price shocks on the aggregate inflation rate. It is argued that such techniques will be more dependable than the more traditional approach of defining core inflation as the CPI mean excluding a particular group of prices.

Finally, the paper implies that the problems of bias in the Laspeyres CPI and the potential benefits from the use of robust inflation measurement techniques are likely to be even greater in developing and transition economies than in developed industrial countries. Developing and transition economies are likely to be particularly exposed to relative price shocks, both genuine and as a consequence of price measurement difficulties, magnifying kurtosis and skewness in the distribution of arithmetic price movements.

³¹See footnote 12. In this context it can be noted that Dalton et al (1998) estimate that the adoption of geometric averaging for about 60 percent of the U.S. CPI (below the basic level of aggregation) will likely lower the average annual inflation rate by about 0.2 percentage points relative to Laspeyres arithmetic index.

REFERENCES

- Amano, R. and T. Macklem, 1997, "Menu costs, relative prices, and inflation: evidence for Canada", *Bank of Canada Working Paper 97-14*.
- Andrews, D., P. Bickel, F. Hampel, P. Huber, W. Rogers, and J. Tukey, 1972, *Robust estimates of location*, Princeton University Press (Princeton, N.J.).
- Assarsson, B., 1984, "Inflation and relative prices in an open economy", *Lund Economic Studies 31*.
- Bakhshi, H. and A. Yates, 1997, "To trim or not to trim", *mimeo*, Bank of England.
- Ball, L. and N.G. Mankiw, 1992, "Relative-price changes as aggregate supply shocks", *National Bureau of Economic Research Working Paper 4168*.
- Ball, L. and N.G. Mankiw, 1994, "Asymmetric price adjustment and economic fluctuations", *The Economic Journal*, vol. 104, pp. 247-61.
- Balk, B., 1978, "Inflation and its variability", *Economics Letters*, vol. 1, pp. 357-60.
- Balk, B., 1983, "Does there exist a relation between inflation and relative price-change variability?", *Economics Letters*, vol. 13, pp. 173-80.
- Balk, B. (ed.), 1998, *Proceedings of the Third Meeting of the International Working Group on Price Indices*, Statistics Netherlands, (Voorburg).
- Blejer, M., 1983, "On the anatomy of inflation", *Journal of Money, Credit and Banking*, vol. 15 (4), pp. 469-82.
- Bowley, A., 1928, *F. Y. Edgeworth's contributions to mathematical statistics*, (Royal Statistical Society, London)
- Bryan, M. and S. Cecchetti, 1996, "Inflation and the distribution of price changes", *National Bureau of Economic Research Working Paper 5793*.
- Bryan, M., S. Cecchetti and R. Wiggins, 1997, "Efficient inflation estimation", *National Bureau of Economic Research Working Paper 6183*.
- Buck, A. and B. Gahlen, 1983, "On the normality of relative price changes", *Economics Letters*, vol. 11, pp. 231-36.
- Cassino, V., 1995, "Menu costs - a review of the literature", *Reserve Bank of New Zealand Discussion Paper G95/1*.
- Clements, K., and H. Izan, 1987, "The measurement of inflation: a stochastic approach", *Journal of Business and Economic Statistics*, vol. 5(3), pp. 339-50.
- Cramér, H., 1946, *Mathematical methods of statistics*, (Princeton University Press, Princeton).

- Cuckierman, A., 1983, "Relative price variability and inflation: a survey and further results", *Carnegie-Rochester Conference Series on Public Policy*, vol. 19, pp. 103-58.
- Dalton, K., J. Greenlees, and K. Stewart, 1998, "Incorporating a Geometric Mean Formula into the CPI", *Monthly Labor Review*, vol. 121(10), pp.3-7.
- Diewert, W. E., 1976, "Exact and superlative index numbers", *Journal of Econometrics*, vol. 4, 115-45.
- Diewert, W. E., 1987, "Index numbers", in J. Eatwell, M. Millgate and P. Newman, eds., 1987, *The new Palgrave. A dictionary of economics*, MacMillan, (London), pp. 767-80.
- Diewert, W. E., 1995, "On the stochastic approach to index numbers", *Department of Economics Discussion Paper 95-31*, University of British Columbia.
- Edgeworth, F., 1887, "Measurement of change in the value of money", *Memorandum presented to the British Association for the Advancement of Science*, reprinted in F. Edgeworth (1925) *Papers relating to political economy*, vol. 1, (Burt Franklin, New York), pp. 198-297.
- Edgeworth, F., 1888, "Some new methods of measuring variation in general prices", *Journal of the Royal Statistical Society*, vol. 51, pp. 346-68.
- Fischer, S., 1981, "Relative price shocks, relative price variability, and inflation", *Brookings Papers on Economic Activity*, vol. 2, pp. 381-441.
- Fischer, S., 1982, "Relative price variability and inflation in the United States and Germany", *European Economic Review*, vol. 18, pp. 171-196.
- Fisher, I., 1922, *The making of index numbers*, (Houghton Mifflin, Boston).
- Harter, H. L., 1974, 1975, "The method of least squares and some alternatives", *International Statistical Review: Part 1* ("Pre-least-squares era (1632-1804)") in vol. 42(2) (1974), pp. 147-74; *Part 2* ("Eighty years of least squares (1805-1884)") in vol. 42(3) (1974), pp. 235-64; *Part 3* ("The awakening (1885-1945)") in vol. 43(1) (1975), pp. 1-44; *Part 4* ("The modern era I (1946-64)" and "The modern era II (1965-74)") in vol. 43(2) (1975), pp. 125-190; *Part 5* ("Conclusions and recommendations") in vol. 43(3), pp. 269-78.
- Hodges, J. Jr., and Lehmann, E., 1963, "Estimates of location based on rank tests", *Annals of Mathematical Statistics*, vol. 34, pp. 598-611.
- Hogg, R., 1967, "Some observations on robust estimation", *Journal of the American Statistical Association*, vol. 62 (Dec.), pp. 1179-86.
- Huber, P., 1964, "Robust estimation of a location parameter", *Annals of Mathematical Statistics*, vol. 35, pp. 73-101.
- Huber, P., 1972, "Robust statistics: a review", 1972 Wald Lecture, *Annals of Mathematical Statistics*, vol. 43(4), pp. 1041-67.

- Jevons, W. S., 1863, *A serious fall in the value of gold ascertained and its social effects set forth*, reprinted in W. S. Jevons, 1884, *Investigations in currency and finance*, (Macmillan, London), pp. 13–150.
- Judge, G., R. C. Hill, W. Griffiths, H. Luetkepol, and T.-C. Lee, 1988, *Introduction to the theory and practice of econometrics*, 2nd edition, Wiley (New York).
- Kearns, J., 1998, “The distribution and measurement of inflation”, *Reserve Bank of Australia Research Discussion Paper 9810*.
- Keynes, J. M., 1930, *A treatise on money*, (Macmillan, London).
- Lourenco, R. and D. Gruen, 1995, “Price stickiness and inflation”, *Reserve Bank of Australia Research Discussion Paper 9502*.
- Lucas, R., 1973, “Some international evidence on output inflation tradeoffs”, *American Economic Review*, vol. 63, pp. 326–35.
- Marquez, J. and D. Vining, 1984, “Inflation and relative price behavior: a survey of the literature”, in M. Ballabon, ed., *Economic perspectives: an annual survey of Economics*, (Harwood, New York).
- Mills, F., 1927, *The behavior of prices*, (National Bureau of Economic Research, New York).
- Mitchell, W., 1915, “The making and using of index numbers”, in “Introduction to index numbers and wholesale prices in the United States and foreign countries”, *Bureau of Labor Statistics Bulletin 173*, U.S. Department of Labor. This article was updated in *Bulletin 284* (1921) and subsequently reprinted as *Bulletin 656* (1938).
- Mizon, G., J. C. Safford and S. Thomas, 1990, “The distribution of consumer price changes in the United Kingdom”, *Economica*, vol. 57, pp. 249–62.
- Parks, R., 1978, “Inflation and relative price variability”, *Journal of Political Economy*, vol. 86 (1), pp. 79–96.
- Pearson, K., 1894, “Contributions to the mathematical theory of evolution”, *Philosophical Transactions of the Royal Society of London*, Series A, vol. 185 (part 1), pp. 71–111.
- Roger, S., 1995, “Measures of underlying inflation in New Zealand, 1981–95”, *Reserve Bank of New Zealand Discussion Paper G95/5*.
- Roger, S., 1997, “A robust measure of core inflation in New Zealand, 1949–96”, *Reserve Bank of New Zealand Discussion Paper G97/7*.
- Roger, S., 1998, “Core inflation: concepts, uses and measurement”, *Reserve Bank of New Zealand Discussion Paper G98/10*.
- Shapiro, M. and D. Wilcox, 1997, “Alternative strategies for aggregating prices in the CPI”, *Federal Reserve Bank of St. Louis Review*, vol. 70(3), pp. 113–125.

- Shiratsuka, S., 1997, "Inflation measures for monetary policy: measuring the underlying inflation and its implication for monetary policy implementation", *Bank of Japan Monetary and Economic Studies*, vol. 15 (2), pp. 1-26.
- Stigler, S., 1973, "Simon Newcomb, Percy Daniell, and the history of robust estimation 1885-1920", *Journal of the American Statistical Association*, vol. 68, pp. 872-79.
- Stigler, S., 1986, *The history of statistics: the measurement of uncertainty before 1900*, (Harvard University Press, Cambridge, Mass.).
- Taillon, J., 1997, "L'inflation sous-jacente: un indice à médiane pondérée", Division des prix, Statistique Canada.
- Vining, D. and T. Elertowski, 1976, "The relationship between relative prices and the general price level", *American Economic Review*, vol. 66 (4), pp. 699-708.
- Yule, U., 1911, *An introduction to the theory of statistics*, 11th (1937) edition (with M. Kendall), (Charles Griffen, London).

**THE MIXED NORMAL DISTRIBUTION, KURTOSIS AND THE EFFICIENCY OF THE MEDIAN
RELATIVE TO THE MEAN**

In this appendix a simple model of the distribution of price movements is developed. In order to capture the notion of some prices showing greater variance about the mean than others, the aggregate distribution of price changes is modeled as mixture of two Normal distributions, having the same means, but different variances. The attraction of the model is that it allows us to determine analytically the kurtosis of the mixed distribution and the efficiency of the sample median relative to the sample mean as estimators of the population mean and median.

A. Sources of Relative Price Variation

To keep things simple, we begin by focusing on the sources of differences in relative price variance before moving on to consider the mixture of two distributions.

Define the logarithm of the aggregate inflation rate (over the domain of goods $i = 1 \dots n$) as:

$$\ln \Pi_t \equiv \ln(P_t/P_{t-1}) = \sum_{i=1}^n w_i \ln(p_{it} / p_{it-1}) = \sum_{i=1}^n w_i \pi_{it} \quad (1)$$

where:

P_t is the aggregate price index;

p_{it} is the price index for good i in period t ;

w_i is the weight of good i in the aggregate index; and $\sum_{i=1}^n w_i = 1$.

The aggregate index, therefore, is a geometric index. If the constituent price indices are equally weighted (i.e. $w_i = 1/n$, for $i = 1 \dots n$), then the price index becomes the Jevons index. If base period and end period expenditure weights or probabilities are averaged, then the index becomes the Törnqvist-Theil index.

Next, decompose the log inflation rate for each good into general and relative price inflation elements:

$$\pi_{it} = \Pi_t + v_{it} \quad (2)$$

where:

v_{it} is the change in the relative price of good i in period t .³²

Changes in relative prices are assumed to be made up of two components:

³²Note that the *relative price* of good i in period t is defined as p_{it}/P_t , whereas the *price relative* for good i in period t is commonly used to describe p_{it}/p_{it-1} .

$$v_{it} \equiv \beta_i + \varepsilon_{it} \quad (3)$$

where:

β_i is a constant log trend in the relative price of good i ; and $\sum_{i=1}^n w_i \beta_i = 0$;

ε_{it} is a stochastic disturbance term to the relative price of good i .

It is further assumed that:

$$\beta \sim N(0, \sigma_\beta^2) \quad (4)$$

$$\varepsilon \sim N(0, \sigma_\varepsilon^2) \quad (5)$$

$$\text{Covariance}(\varepsilon_{it}, \varepsilon_{jt-s}) = 0 \text{ for all } i \neq j \text{ and for all } t \neq s \quad (6)$$

It follows from (4) and (5) that:

$$v \sim N(0, \sigma^2) \quad (7)$$

where:

$$\sigma^2 = (\sigma_\beta^2 + \sigma_\varepsilon^2) \quad (8)$$

Differences in the variances of price movements around the aggregate mean, therefore, can arise from either transient disturbances or from persistent trends in relative prices.

B. Mixed Normal Price Changes and Kurtosis of the Aggregate Distribution

We now relax the assumption that the variances of prices are identical. Instead, following Yule (1911), we assume that the aggregate distribution of log price changes is composed of two Normal sub-distributions, having identical means but different population variances. The first group, with a total weight of α in the index, is characterized by relative price variance of σ_1^2 , while the second group is characterized by relative price variance of σ_2^2 . That is:

$$\pi_1 \sim N(0, \sigma_1^2) \text{ for } i = 1 \dots j \quad (9)$$

$$\pi_2 \sim N(0, \sigma_2^2) \text{ for } i = j+1 \dots n \quad (10)$$

For convenience, define:

$$\alpha \equiv \sum_{i=1}^j w_i \quad (11)$$

and

$$\lambda \equiv \sigma_1 / \sigma_2; \lambda \geq 0 \quad (12)$$

It can be noted that, as in equation (8), differences in the variances between the two groups can arise from differences in the variances of either the stochastic or trend components. Consequently, the model can be thought of as allowing for a group of particularly "noisy"

prices or, alternatively allowing for a group of prices with trends that are significantly different from most prices.³³ For the purposes of statistical inference, however, it is probably best to not think of specific prices falling into one or the other group but, rather, to think of all prices as being exposed high and low variance shocks with probabilities α and $(1 - \alpha)$, respectively.

The composite distribution of log price changes remains symmetric, because both sub-distributions are Normal about the same mean. The kurtosis of the composite distribution, however, depends on the values of λ (the ratio of σ_1 relative to σ_2) and α , as shown below. The derivation closely follows Bryan, Cecchetti and Wiggins (1997), but generalizes their analysis slightly to allow for variation in both α and λ .

The kurtosis of the distribution is defined as:

$$\beta_2 \equiv \mu_4 / (\mu_2)^2 \quad (13)$$

where:

$$\mu_n = E[x_i - E(x_i)]^n$$

For the mixed distribution described in equations (9) to (12), (and since $E(x_i) = 0$),

$$\mu_n = E(x_i)^n = \sum_{i=1}^j (x_i)^n + \sum_{j+1}^n (x_i)^n = [\alpha \sigma_1^n + (1-\alpha) \sigma_2^n] = \sigma_2^n [1 + \alpha (\lambda^n - 1)] \quad (14)$$

Hence, the kurtosis of the distribution can be re-written as:

$$\beta_2 = \frac{\sigma_2^4 [1 + \alpha (\lambda^4 - 1)]}{\{\sigma_2^2 [1 + \alpha (\lambda^2 - 1)]\}^2} = \frac{\sigma_2^4 [1 + \alpha (\lambda^4 - 1)]}{(\sigma_2^2)^2 [1 + \alpha (\lambda^2 - 1)]^2} \quad (15)$$

Lastly, since π_2 is Normally distributed, it follows that $\sigma_2^4 / (\sigma_2^2)^2 = 3$, and equation (15) simplifies to:

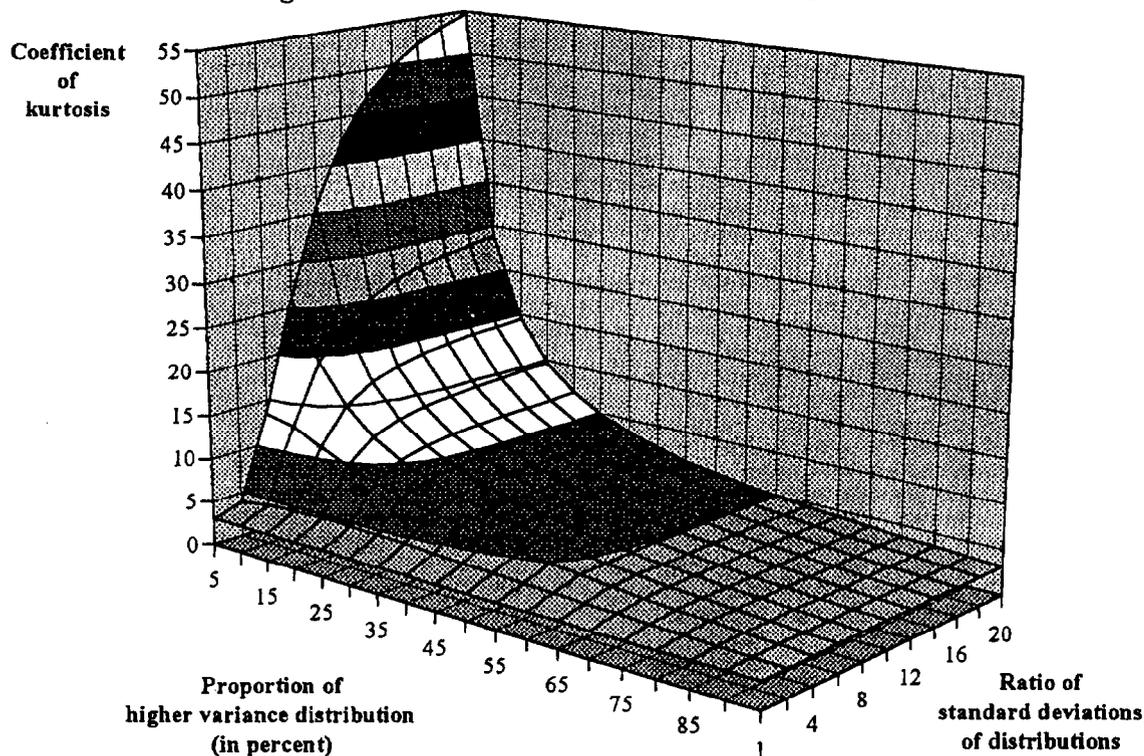
$$\beta_2 = \frac{3[1 + \alpha (\lambda^4 - 1)]}{[1 + \alpha (\lambda^2 - 1)]^2} \quad (16)$$

Now, since $\lambda \geq 0$, the mixed Normal distribution will always show kurtosis equal to or greater than 3, as shown in Figure A1. Figure A1 indicates that:

³³Since relative price variances depend on the *squares* of relative price trends, it makes no difference for these purposes whether the relative price trends are positive or negative.

- Even a very small degree of “contamination” of a Normal distribution by high variance prices will dramatically increase the kurtosis of the mixed distribution.
- Although leptokurtic distributions are often described as either “long-tailed” or “high-peaked”, with the terms used interchangeably, Figure A1 shows that the two are not the same. The classic “long-tailed” distribution occurs with low values of α , so that we can think of an essentially Normal distribution “contaminated” with a small proportion of extreme price changes. A “high-peaked” distribution occurs with a high value for α , so that we can think of an essentially Normal distribution mixed with a small proportion of low variance price changes.³⁴ Since empirical evidence on the distribution of price changes typically shows kurtosis above 6, it is appropriate to think of these distributions as long-tailed; that is, as distributions “contaminated” with a fairly small proportion of exceptionally large price changes.

Figure A1. Kurtosis of the Mixed Normal Distribution



³⁴It can also be noted that for the case of $\alpha = 0.5$, the limiting value of the kurtosis is 6, the same as for the double-exponential or Laplace distribution. The Laplace distribution can, therefore, be thought of as occupying the middle ground between “long-tailed” and “high-peaked” leptokurtic distributions.

C. The Efficiency of the Sample Median Relative to the Sample Mean

It is often assumed that the *sample* mean of a distribution is the most efficient estimator of the *population* mean. In general, however, that is not true. The sample mean should be an unbiased estimator of the population mean, but it will only be the most efficient estimator if the underlying population distribution from which the sample is drawn is Normal. For even quite small departures of the population distribution from the Normal, the sample mean becomes a very inefficient estimator. By contrast, the efficiency of the sample median is much less sensitive to the shape of the underlying population distribution, and this is why it is classified as a *robust* estimator.

The efficiency of the sample median relative to the sample mean, as an estimator of the population mean, is defined as the ratio of the standard errors of the two estimators. For the mixed Normal model, this ratio can be found analytically, in contrast with many other non-Normal distributions.³⁵

Begin with the standard deviation of the sample p^{th} percentile of the frequency distribution, $\sigma_{x(p)}$:

$$\sigma_{x(p)} \equiv \frac{\sigma}{f(p)} \sqrt{\frac{p(1-p)}{n}} = \frac{\sigma}{\sqrt{n}} \frac{\sqrt{p(1-p)}}{f(p)} \quad (17)$$

where:

$f(p)$ is the ordinate of the p^{th} percentile of the cumulative frequency distribution.

The variance of the sample mean is defined as:

$$\sigma_{\bar{x}}^2 \equiv \frac{\sigma^2}{n},$$

So that the sample standard deviation is: $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$ (18)

Combining equations (17) and (18) and rearranging yields an expression for the efficiency of the sample p^{th} percentile relative to the sample mean:

$$\frac{\sigma_{\bar{x}}}{\sigma_{x(p)}} = \frac{f(p)}{\sqrt{p(1-p)}} \quad (19)$$

If we now focus on the 50th percentile, or median, equation (19) can be re-written as:

³⁵This exposition is based closely on Yule (1911), pp. 383–84.

$$\frac{\sigma_{\bar{x}}}{\sigma_m} = \frac{f(0.5)}{\sqrt{0.5(1-0.5)}} = 2f(0.5) \quad (20)$$

In the case of the Normal distribution, where $f(0.5) = 1/\sqrt{2\pi} = 0.399$, the efficiency of the sample median is only about 80 percent of the efficiency of the sample mean or, equivalently, the sample mean is about 25 percent more efficient than the sample median.

In the case of the mixed Normal distribution the relative efficiency of the median is given by:

$$\frac{\sigma_{\bar{x}}}{\sigma_m} = 2f(0.5) = 2 \left[\frac{\alpha}{\sigma_1 \sqrt{2\pi}} + \frac{(1-\alpha)}{\sigma_2 \sqrt{2\pi}} \right] \sigma = 2 \left[\frac{\alpha\sigma_2 + (1-\alpha)\sigma_1}{\sigma_1\sigma_2 \sqrt{2\pi}} \right] \sigma \quad (21)$$

And, since $\sigma_1 = \lambda\sigma_2$, equation (21) can be simplified to:

$$\frac{\sigma_{\bar{x}}}{\sigma_m} = 2 \left[\frac{[\alpha + (1-\alpha)\lambda]}{\lambda\sqrt{2\pi}} \right] \frac{\sigma}{\sigma_2} \quad (22)$$

Now, recall from equation (14) that:

$$\mu_n = \sigma_2^n [1 + \alpha(\lambda^n - 1)]$$

The population variance, σ^2 , of the mixed Normal distribution is, therefore:

$$\sigma^2 = \mu_2 = \sigma_2^2 [1 + \alpha(\lambda^2 - 1)]$$

And the standard deviation is:

$$\sigma = \sigma_2 \sqrt{1 + \alpha(\lambda^2 - 1)} \quad (23)$$

Finally, substituting equations (17) and (21) into (22) yields:

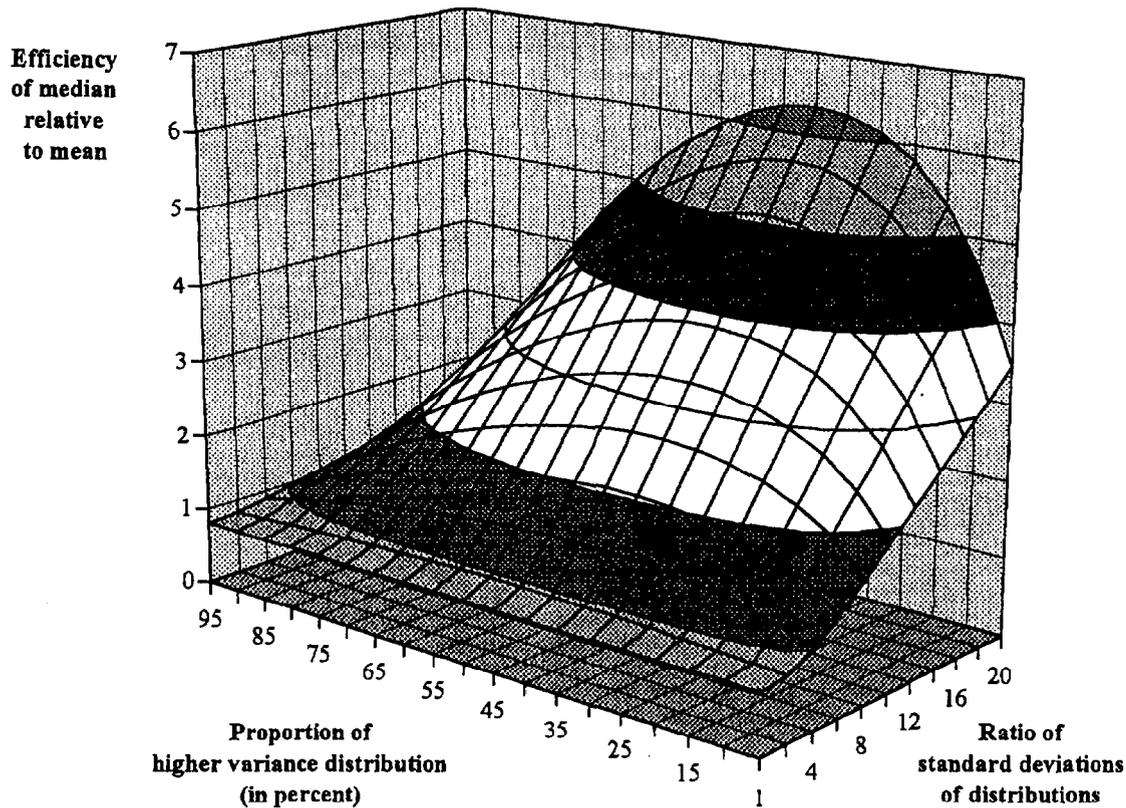
$$\frac{\sigma_{\bar{x}}}{\sigma_m} = \frac{2[\alpha + (1-\alpha)\lambda]}{\lambda\sqrt{2\pi}} \sqrt{1 + \alpha(\lambda^2 - 1)} \quad (24)$$

In contrast to equation (20), equation (24) indicates that the relative efficiency of the sample median will depend on the values of both α and λ , as shown in Figure A2. Figure A2 also indicates that:

- For a population distribution even only slightly different from the Normal, the sample median is a much more efficient estimator of the population mean (and median) than is the sample mean.

- The relative efficiency of the median is not a simple function of the kurtosis of the distribution – otherwise Figures A1 and A2 would look very similar. The intuitive explanation is that kurtosis depends somewhat more heavily than the efficiency of the median relative to the mean on the *length* of the tails of the distribution, as opposed to the *thickness* of the tails. As a consequence, the relative efficiency of the median tends to be greatest roughly midway between very long-tailed, kurtotic distributions (characterized by a very low value of α), and highly peaked, relatively thick-tailed distributions (characterized by a very high value of α).

Figure A2. Efficiency of the Sample Median Relative to the Sample Mean of the Mixed Normal Distribution



EVOLUTION OF THE ESTIMATION OF CENTRAL TENDENCY

In a comprehensive and fascinating survey of the historical development of the theory of statistical inference, Harter (1974, 1975) dates the beginnings of the development of the modern concepts of statistical inference to about 1632 with Galileo Galilei's work on measurement in the field of astronomy. Harter observes that prior to the 17th century, the most common measures of central tendency appeared to be the mode or mid-range (the simple mean of the highest and lowest observations).³⁶

The modern theory of statistical inference was mainly developed in the late 18th and early 19th centuries, most notably by Pierre Laplace and Carl Gauss, and continued to be driven mainly by measurement issues in astronomy. Laplace (1774, 1781, 1793, 1799. See Harter 1974 for details) followed Roger Boscovich (1757) in proposing as criteria for the best measure of central tendency that (a) positive and negative "errors" be equally likely (i.e. sum to zero); and (b) the sum of products of absolute errors and their respective probabilities be minimized.³⁷ Laplace's analysis showed that the measure of central tendency satisfying these criteria depended on the shape of assumed population distribution. If the distribution approached the symmetric or double exponential (i.e. following Laplace's first law of error), then the best measure of central tendency approached the (weighted) median. However, if the distribution approach the Normal (i.e. following Laplace's second law of error), the best measure of central tendency approached the (weighted) arithmetic mean.

Carl Gauss (1806) is commonly credited with the development of the least squares method and the rationale for the use of the arithmetic mean as the standard measure of central tendency.³⁸ Gauss, however, did not claim that the arithmetic mean was *the* best measure of central tendency. What he did show was that the arithmetic mean was the minimum least squares estimator of central tendency *if* the population distribution was Normal.³⁹

³⁶Harter (1975, Part 3, p. 17) notes that Plackett (1958) dates the use of the arithmetic mean only as far back as its use by Tycho Brahe in the late 16th century, and that it only ceased being controversial following the work of de Moivre, Simpson and Lagrange in the 18th century. Interestingly, Huber (1972, p. 1043) quotes an 1821 paper (attributed to Jons Svanberg) indicating that trimmed means were in fairly common use in calculating agricultural yields in France.

³⁷It may be noted that Laplace's formulation differed slightly from Boscovich's insofar as it replaced the simple median with the weighted median.

³⁸Harter (1974, Part 1, pp. 152-54) notes, however, that the method was first published by Adrien Legendre in 1805, and was also developed independently by Robert Adrain (1808).

³⁹See Huber (1972), pp. 1042-43.

Clearly, however, the method of least squares took the statistical world by storm, even leading Laplace (1810, 1811, 1812) to argue in favor of the least squares method and the arithmetic mean as the “most advantageous mean” for large samples. An important dissenting opinion noted by Harter (1974, Part 1, pp. 160–61) was expressed by Peter Lejeune Dirichelet (1836). Dirichelet observed correctly that, regardless of the sample size, the superiority of the arithmetic mean over the median depends on the ratio of the standard errors of the two measures.

The accepted superiority of the arithmetic mean over other estimators of central tendency began to be challenged seriously by Francis Edgeworth in a series of papers published between 1885 and 1888.⁴⁰ Edgeworth argued that:

- For samples from a Normal distribution, the arithmetic mean was most accurate (i.e., had the lowest variance of the sample mean) but that the median was little less accurate. For samples from distributions more kurtotic (i.e., more peaked or long-tailed) than the Normal, the mean was less accurate than the median (as had been shown earlier by Laplace).⁴¹
- The choice of which mean to use as the measure of central tendency of a distribution should be the measure with the smallest standard error (as had been argued earlier by Dirichelet).

Using these principles, Edgeworth (1888) argued in favor of the use of the median as the best measure of central tendency for prices on the basis of evidence presented by Jevons and others that the distribution of prices was typically markedly more kurtotic than Normal.

The distribution of prices was also typically found to be right-skewed, leading Jevons to recommend use of the geometric mean. Edgeworth’s solution was to modify the Boscovich-Laplace formulation of the median (dropping the constraint that the sum of deviations from the measure of central tendency be zero) so that it became simply the least absolute errors measure of central tendency. In the case of a symmetric distribution, the constraint was not a binding one, but for asymmetric distributions, as the distribution of prices appeared to be, dropping the constraint did matter.

To this point, the development of the stochastic approach had focused primarily on the efficiency of alternative estimators of central tendency for particular frequency distributions. In the 1940s the focus shifted towards the development of so-called “robust” estimators. Loosely speaking, “robust” estimators could be characterized as estimators whose efficiency is high for

⁴⁰Harter (1974, Part 2), pp. 235–37. See also Edgeworth (1887, 1888) and Bowley (1928).

⁴¹It may be noted that Edgeworth particularly recommended use of the median in the case of a distribution characterized by “discordant observations”. In modern terminology, this would be described as a “contaminated” distribution and will be more kurtotic than the Normal distribution.

the sorts of distributions of interest, but not extremely sensitive to variations in the distribution.⁴² Eisenhart (1971, reported in Harter (1975, Part 3, p. 148)) points to the wartime needs of precision bombing as a spur to development of robust estimators. Huber (1972, p. 1044) also notes, however, that E.S. Pearson (1931) was concerned even before the war with the poor performance of the arithmetic mean, and measures employing it, for even slightly non-Normal distributions.

Harter and Huber emphasize the contributions by John Tukey and the Statistical Research Group at Princeton University through the late 1940s and 1950s to development of robust methods of statistical inference. Huber (1972) summarizes the insights gained as follows:

- “one never has a very accurate knowledge of the true underlying distribution”;
- “the performance of some of the classical tests or estimates is very unstable under small changes in the underlying distribution”;
- “some alternative tests or estimates (like the Wilcoxon instead of the t -test, or the α -trimmed mean instead of the [arithmetic] mean) lose very little efficiency for an exactly [N]ormal law, but show a much better and more stable performance under deviations from it.” (p. 1045).

In the 1950s and 1960s research led to a proliferation of alternative “robust” estimators of central tendency or location for distributions.⁴³ The most prominent of these were the so-called L -estimators consisting of linear (hence the “ L ”) combinations of order statistics, and include the α -trimmed mean. As with all estimators based on order statistics, L -estimators involve, first, ordering all observations (e.g., price changes, in ascending order). The observations are then re-weighted (or double-weighted) according to their order or relative position in order to arrive at a recalculated mean. In the case of the α -trimmed mean, for example, a fraction α of the observations (by number or by initial weight in the distribution) at each end of the ordered distribution are given a zero-weight, while the weights of the middle $(1-2\alpha)$ observations are scaled up proportionately to sum to unity. The zero-weighting of the extreme observations is equivalent to replacing the observed values with the mean of the remaining observations. In some respects this may be a more helpful way of thinking about L -estimators if the reason for using a robust measure is based on doubts about the accuracy of the extreme observations.

The mean and the median and mid-range can be thought of as very particular L -estimators. The mean can be thought of as the α -trimmed mean, with $\alpha = 0$. In this case, the re-weighted distribution is equal to the original weighting scheme, with the important consequence that the ordering of the observations ceases to have any effect on the value of the statistic. By contrast,

⁴²Huber (1972), pp. 1046–47 offers a more detailed discussion of the meaning of “robustness”.

⁴³See, e.g., Huber (1972) or Judge et al (1988).

with the median, all but the middle observation are zero-weighted. Finally, the mid-range involves zero-weighting all but the observations at either end of the ordered distribution.

L-estimators also include an infinite number of other, more complex, re-weighting possibilities.⁴⁴ These include, notably, linear combinations in which the re-weighting is not binary (i.e. include or exclude) but decreasing or increasing from the center, either linearly or non-linearly; linear combinations of selected order statistics (e.g. the Gastwirth mean, which assigns weights of 0.3, 0.4 and 0.3 to the 33.33rd, 50th and 66.67th percentiles, respectively); or Winsorized means.⁴⁵

In addition to the class of *L*-estimators, two other broad classes of estimators of central tendency were developed: *M*- and *R*-estimators.

M-estimators were first proposed by Peter Huber (1964). These estimators are based, like the mean, median and mid-range, on minimizing some function of the sum of "errors" (i.e., the difference between the observations and the estimator of central tendency). In other words, the estimator *T* is such that:

$$T = T_n(x_1, \dots, x_n) \text{ minimizes } \sum_i \varphi(t) = \sum_i \varphi(x_i - T); \varphi \text{ non-constant}$$

- If $\varphi(t) = |t|^\infty$, then *T* is the sample mid-range - the least maximum errors estimator;
- If $\varphi(t) = |t|^2$, then *T* is the sample mean - the least squared errors estimator;
- If $\varphi(t) = |t|$, then *T* is the sample median - the least absolute errors estimator;
- If $\varphi(t) = -\log f(t)$, (where *f* is the assumed density function), then *T* includes all maximum likelihood estimators.

Huber then defines a robust estimator as the minimum asymptotic variance estimator when there is a Normal distribution contaminated by an unknown symmetric distribution. If there is no contamination, the minimum variance estimator is the sample mean. If there is extreme contamination, the sample median is most robust. For intermediate cases, Huber finds that the form of the most robust estimator is related to a Winsorized mean.⁴⁶

R-estimators were first proposed by Joseph Hodges and Erich Lehmann (1963). These estimators, based on non-parametric rank tests, involve non-linear combinations of order

⁴⁴See, e.g., Andrews et al (1972).

⁴⁵In the α -trimmed mean, the weight of the trimmed observations is used to scale up proportionately the weights of all the remaining observations. In the Winsorized mean, the weight α is added to the initial weights of the remaining highest and lowest observations.

⁴⁶See Huber (1972) and Andrews et al (1972) for details of important modifications to *M*-estimators by Frank Hampel (1968).

statistics. Harter (1975, Part 3, p. 28) notes in particular the "Hodges-Lehmann" estimator, based on the median of pairwise averages of different order statistics.

In 1972, Andrews et al examined the performances of some 68 alternative robust estimators of location or central tendency, over a large number of distributions and several different sample sizes, using Monte Carlo methods. As noted by Harter (1975, part 4, p.155), the results are not readily summarized. Nor do they point to the general superiority of any one estimator or class of estimators. Nonetheless, the authors identify important dimensions in which the estimators differ:

- Estimators differ with respect to the length of tail of the distribution for which they perform best. The authors identify this dimension as "gross-error-sensitivity" of the estimator.
- Estimators also vary significantly in their sensitivity to outliers at different distances. Some are quite strongly influenced by moderate outliers, but very insensitive to distant outliers. In other cases, sensitivity does not vary as significantly between moderate and distant outliers.
- The performances of different estimators also vary to some extent with sample sizes - some perform quite well for small sample sizes, while others do better for larger samples.
- The pros and cons of the various robust estimators should be kept in perspective: they are all far more robust than the sample mean which is a very poor estimator of central tendency for any other distribution than the Normal.

On one level, the lack of a conclusive recommendation in favor of a particular robust estimator or group of estimators is frustrating. But it should not be surprising. Two hundred years ago, Laplace was making the same point: the "best" estimator will differ according to the distribution involved; there is, therefore, no one estimator that is best for all situations.

To be sure, some estimators appear, on balance, to be generally better than some others. But within the group of generally better estimators, the choices to be made appear to depend substantially on one's knowledge or priors regarding the underlying population distribution and on the kinds of risks one particularly wishes to minimize. In this context, Huber (1972, p. 1047) observes that "we often have quite a good idea of the approximate shape of the true underlying distribution...so that it should suffice to consider a neighborhood of only one shape".

Consequently, the range of estimators to consider seriously can be narrowed quite quickly - at least once the empirical distribution has been examined.

Having narrowed the range of estimators to choose from, the basic statistical trade-off is between the efficiency of an estimator for a particular distribution and the robustness of the estimator to variation in the distribution. Huber, following Anscombe (1960), regards the issue as one of insurance; in the choice of estimator he is willing to pay a "premium" in terms of slightly lower efficiency of the estimator if it will provide protection against distortions to, or erroneous assumptions about, the distribution. How much of a "premium" one is willing to pay should depend on how certain one is of the true underlying distribution or how susceptible it may be to change or distortion.

An additional, non-statistical, criterion to be taken into consideration is the reproducibility and general comprehensibility of the measure. This is bound to weigh more heavily in the use of robust estimators in the very public context of accounting for inflation outcomes than in a more purely research context.

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.⁴⁷ Second, the higher the kurtosis of the distribution, the less weight the estimator should place on observations in the tail of the distribution.

Robert Hogg (1967), for example, offered the following simple scheme for the use of different estimators:

- the sample mean if the kurtosis of the distribution is between 2 and 4 (note that the Normal distribution has a coefficient of kurtosis of 3);
- the 25 percent trimmed mean (also known as the mid-mean) if kurtosis is between 4 and 5.5;
- the median if kurtosis is above 5.5 (note that the double exponential distribution has a coefficient of kurtosis of 6.0).

For a moderate sample size, Harter (1974; reported in Harter 1975, Part 4, pp. 174) suggested:

- the mid-range if kurtosis is less than 2.3;
- the mean if kurtosis is between 2.3 and 3.7;
- the median if kurtosis is over 3.7.

Finally, Huber (1972, pp. 1063–64) suggested:

- a 10 or 15 percent trimmed mean for approximately Normal distributions;
- the Gastwirth mean or a simple Hampel estimator for “poorly specified and presumably long-tailed” distributions.⁴⁸

Clearly, the theory of robust estimation of central tendency or location has come a long way since Edgeworth’s day, and the range of estimators has expanded enormously. Nonetheless, Edgeworth’s recommendation (1888) of the median in situations in which the distribution is

⁴⁷Hogg (1967) also notes in favor of the trimmed mean that it is approximately normally distributed with a calculable variance and that the variance appears to be approximately *t*-distributed.

⁴⁸Hampel’s estimates are closely related to Huber’s *M*-estimates.

typically much longer-tailed than for the Normal distribution, together with his observation that "...even where the arithmetic mean is better, it is not likely to be very much better" (p. 363), is fully consonant with the results of subsequent research.

