

# Inflation and the Distribution of Price Changes

Michael F. Bryan and Stephen G. Cecchetti <sup>1</sup>

Revised April 1998

Forthcoming *Review of Economics and Statistics*

<sup>1</sup>Federal Reserve Bank of Cleveland; and Federal Reserve Bank of New York, Ohio State University and Research Associate, National Bureau of Economic Research, respectively. We gratefully acknowledge comments from Stephen Cosslett, Ben Craig, Pok-sang Lam, Nelson Mark, Margaret Mary McConnell, George Pennacchi, Gabriel Perez Quiros, James Stock (the editor), the referees, and seminar participants at the Reserve Bank of Australia; research assistance from Rodney L. Wiggins; editorial assistance from Valerie LaPorte; and data supplied by Laurence Ball and N. Gregory Mankiw. The views stated herein are those of the authors and not necessarily those of the Federal Reserve Bank of Cleveland, the Federal Reserve Bank of New York or of the Board of Governors of the Federal Reserve System.

# INFLATION AND THE DISTRIBUTION OF PRICE CHANGES

## Abstract

This paper reconsiders the empirical evidence connecting inflation to its higher order moments. In particular we examine the statistical properties of the observed positive correlation between the sample mean and the sample cross-sectional skewness of price changes. This correlation has attracted substantial attention over the years and has recently been the focal point of a debate among macroeconomists. We show that the sample mean-skewness correlation suffers from a small-sample bias that accounts for the entirety of the observed correlation. In other words, we establish that one of the stylized facts in the literature on aggregate price behavior need not be a fact at all.

(JEL E31, C12)

# I. Introduction

Views of the sources, costs, and consequences of inflation are often tied to evidence on the relationships among the cross-sectional moments of inflation. For example, one often hears the argument that higher inflation is more variable inflation (cross-sectionally) and that this is costly.<sup>1</sup> Evidence on the statistical relationship between mean inflation and its higher moments (variance and skewness) has been around for some time. Vining and Elwertowski (1976) is the standard reference for the United States, although the authors cite a much earlier literature. They find a large positive correlation between mean inflation and its higher-order moments.<sup>2</sup>

More recently, Ball and Mankiw (1995) have revived interest in the relationship between the mean and skewness of inflation. It is their view that the positive correlation is evidence for a particular sticky-price model of inflation. Balke and Wynne (1996a, 1996b) take issue with this conclusion and propose an alternative theory in which the same positive correlation between the mean and the skewness of price changes arises in a multisectoral flexible-price model.

The purpose of this paper is to examine the properties of the sample statistics on which this work is based. Our primary interest is determining whether the observed sample (time-series) correlation between the mean and skewness of the cross-sectional distribution of price changes is significantly different from zero. We are able to show that this correlation is biased in small samples, and that the bias arises because price-change distributions have tails fatter than those of the normal distribution. For samples of the size we typically see, drawn from a symmetrical distribution, the correlation of the sample mean and sample skewness will be positive, so long as the data-generating process exhibits excess kurtosis.

The intuition for the existence of the bias is straightforward. Consider a sample drawn from a unimodal, zero-mean, symmetric distribution. Take a case in which the initial observations have a sample mean of zero, and imagine taking one additional draw from the extreme positive tail of the distribution. The sample mean and the sample

---

<sup>1</sup>See, for example, the work of Fischer (1981) and Taylor (1981).

<sup>2</sup>Ball and Mankiw (1995) reference a more recent literature that establishes similar properties of data from other countries.

skewness are now both positive. And the further in the tail the added observation, the more positive both the sample mean and the sample skewness, producing a small-sample bias in the covariance between the estimates of the first and third moments of the distribution. So long as the mean of inflation has nonzero time-series variance, the mean-skewness correlation does converge to zero as the sample size (in both the cross-sectional and time-series dimensions) goes to infinity. But in small samples, there is a bias that depends on the second, fourth, and sixth moments of the cross-sectional distribution generating the data, as well as the variance of the time-series mean of inflation and the sample size. In particular, for given variances, sixth moment, and sample size, the greater the kurtosis of the distribution, the more positive the bias.

Our derivation of the small-sample bias points to a problem with the standard interpretation of the sample correlation between cross-sectional mean inflation and its higher-order moments. But the analytical results are necessarily approximate, as we are unable to compute the small-sample properties of the correlation estimator under realistic distributional assumptions. To address this difficulty, we conduct a series of Monte Carlo experiments that allow us to compute the exact small-sample distribution of the mean-skewness correlation. Using consumer and producer price data at various frequencies, we estimate a large positive bias in this statistic that accounts for virtually all of the observed sample correlation. This finding leads us to conclude that such statistics are not useful in distinguishing sticky from flexible price-setting behavior in macroeconomic models.

The remainder of the paper is divided into six sections. Section II summarizes the two theoretical models of inflation that are the focus of the debate surrounding the mean-skewness correlation. We follow with a preliminary examination of the data in Section III. In Section IV, we examine the statistical theory of the bias in the mean-skewness correlation and present preliminary empirical results. Section V reports the application of Monte Carlo techniques to the consumer and producer price data in order to construct exact small-sample distributions for the correlation. In Section VI, we consider some simple robustness checks using consumer price data. We examine the consequences of splitting the data into two subsamples, and of excluding food and energy. The final section summarizes our findings.

## II. Recent Models of Asymmetric Price Change

We proceed with an overview of two recent theoretical models used to interpret the correlation between mean inflation and the cross-sectional skewness of price changes. The first is a simple sticky-price model developed by Ball and Mankiw (1995), and the second is a flexible-price model created by Balke and Wynne (1996a, 1996b).

### A. The Ball and Mankiw Model

Ball and Mankiw (1995) assume that there is a single period at the beginning of which each firm in the economy fixes its price. Following this initial adjustment, each firm is subjected to a mean-zero (real) shock and can pay a menu cost to change its price a second time. Not all firms will experience shocks that are large enough to make the second adjustment worthwhile. As a result, the observed change in the aggregate price level will depend on the shape of the distribution of the idiosyncratic shocks.

In particular, Ball and Mankiw point out that if the distribution of real shocks is skewed, this will move the aggregate price level in the same direction. A positively skewed distribution, for example, will result in an aggregate price level increase. In general, asymmetrical distributions of relative price shocks will cause transitory fluctuations in the mean of prices, producing a positive correlation between the mean and the skewness of the price-change distribution.<sup>3</sup>

The Ball and Mankiw model has additional implications that the authors do not explore. For example, De Abreu Lourenco and Gruen (1995) note an implication for the relationship between the level and cross-sectional variance of inflation and the level of expected inflation. In addition, the sticky-price model implies that the mean-skewness correlation should die out in the long run. In the Ball and Mankiw model, the correlation arises purely from short-run considerations. Once all the price setters in the economy have adjusted to their relative-price shocks, the correlation should disappear. The implication is that as the length of time over which prices changes are measured increases from months to quarters to years, the correlation should weaken and eventually disappear.

---

<sup>3</sup>In Bryan and Cecchetti (1994), we noted how we could remove these transitory fluctuations in aggregate prices by using limited-influence estimators such as trimmed means and the median. Cecchetti (1997) and Bryan, Cecchetti, and Wiggins (1997) extend this earlier analysis.

## B. The Balke and Wynne Model

Balke and Wynne (1996a, 1996b) build a multisector dynamic general equilibrium model with money and flexible prices. In this approach, a productivity (or supply) shock arising in one sector has two implications. First, it changes aggregate output. With the quantity of money fixed and a cash-in-advance constraint, the change in output induces a change in the aggregate price level. Beyond this, the aggregate supply shock feeds through the input-output structure of the economy to create differential impacts on different sectors. Prices in each sector respond to the extent that the supply shock affects the sector's average productivity. Since the input-output matrix is not symmetrical, the resulting price-change distribution is skewed.

Balke and Wynne show that their model implies a positive correlation between the mean and skewness of inflation. The reason is straightforward. The larger the initial supply shock, the larger the change in the aggregate price level. But since the coefficients of the input-output matrix are fixed, the larger the initial sectoral productivity disturbance, the more skewed its impact will be on the distribution of price changes.

While both the sticky- and the flexible-price models imply a positive correlation between the mean and skewness of price changes, they differ in their predictions of the effects at short and long horizons. In the Balke and Wynne model, the effects need not die out in the long run. In fact, if the effect is predominantly a real one, it might become more pronounced over time.

## III. Preliminary Analysis of the Data

Our analysis of data begins with the simple calculation of sample moments. We study two data sets in detail. The first is composed of 36 components of the consumer price index (CPI), available monthly from January 1967 to April 1997. The second uses 29 components of the producer price index (PPI) for finished goods, monthly from January 1967 to April 1997.<sup>4</sup> In both cases, the number of components chosen is the maximum

---

<sup>4</sup>The BLS seasonally adjusts only those components of the CPI that contain statistically significant seasonality. We use the resulting full set of CPI subaggregates that have no seasonality. For the PPI, only seasonally unadjusted series are available, and so we seasonally adjust all of the components using

that enables us to maintain a balanced panel over the entire sample period.

We begin by computing each component's price change over horizon  $k$ :

$$\dot{p}_{it}^k = \frac{1}{k} \ln(p_{it}/p_{it-k}) , \quad (1)$$

where  $p_{it}$  is the index level for component  $i$  at time  $t$ . From this, we define the mean inflation in each time period as

$$\pi_t = \sum_i w_i \dot{p}_{it} , \quad (2)$$

where  $w_i$  is the weight attached to each of the components. (For future reference, we also define  $m_{1t}$  as the sample mean.) The higher-order central moments are then just

$$m_{rt} = \sum_i w_i (\dot{p}_{it} - \pi_t)^r . \quad (3)$$

Skewness and kurtosis are the scaled third and fourth moments, respectively:

$$S_t = \frac{m_{3t}}{[m_{2t}]^{(3/2)}} \quad (4)$$

and

$$\mathcal{K}_t = \frac{m_{4t}}{[m_{2t}]^2} .$$

In computing the moments, we use the actual 1985 expenditure weights to aggregate the 36 components of the CPI and the 1982 output value weights to aggregate the 29 components of the PPI.

Table 1 reports the various moments of both indices computed over horizons of one to 24 months. The most important thing to notice is the kurtosis of these distributions. On average, monthly CPI inflation has kurtosis in excess of twelve. The average kurtosis declines as the horizon increases, but remains relatively large even at low frequencies (kurtosis equals 5.25 for 24-month averages of monthly inflation). The PPI data show the same general pattern; kurtosis is high for one-month changes, falling only gradually as the horizon increases. Furthermore, while the data are often skewed over an individual month or year, the average skewness is nearly zero at all horizons for both data sets.

---

the ARIMA-X11 program in SAS.

Table 1: Sample Moments of Price-Change Distributions

36 Components of the CPI				
Sample 1967:01 – 1997:04				
	$k = 1$	$k = 3$	$k = 12$	$k = 24$
Mean	5.28	5.30	5.37	5.43
Std. dev.	7.88	5.57	3.41	2.76
Skewness	0.32	0.26	0.19	0.14
Kurtosis	12.80	10.91	6.62	5.25
29 Components of the PPI				
Sample 1967:01 – 1997:04				
	$k = 1$	$k = 3$	$k = 12$	$k = 24$
Mean	4.56	4.61	4.73	4.81
Std. dev.	13.41	9.34	5.50	4.00
Skewness	0.07	0.25	0.26	0.18
Kurtosis	11.33	9.55	8.12	6.00

Notes: The table reports time-series averages of the cross-sectional moments of price-change distributions at annual rates.  $k$  is the horizon, so  $k = 12$  refers to 12-month overlapping changes.

## IV. Bias in the Mean-Skewness Correlation

All evidence linking inflation and skewness comes from sample statistics. For example, the sample correlation between the weighted mean and weighted skewness in the 36 components of the CPI, measured at a monthly frequency over the 1967–97 sample, is 0.35. A standard hypothesis test for whether this is different from zero yields an asymptotic t-ratio of 6.74, which implies rejection at any meaningful level of statistical significance. But is such a conclusion warranted? What are the small-sample properties of this correlation, and how might they affect such an inference?

Some simple intuition suggests why time-series estimates might be misleading. Consider a sample drawn from an arbitrary, but symmetrical, distribution. The important insight is that draws deviating from the population mean of the distribution can move the sample first and third moments systematically away from their population values, inducing a correlation.

In fact, we can show analytically how such a correlation arises in small samples.

Suppose that we define  $x_{it}$  as a single observation in a panel of data with  $T$  rows and  $N$  columns. Assume that the  $x_{it}$ 's are i.i.d. across  $i$  and are stationary and ergodic over  $t$ , but with a mean that varies over time. That is, taking expectations over  $i$ ,  $E_i(x_{it}) = Z_t$ , where the notation " $E_i$ " is taken to mean "expectation over  $i$ ," and  $Z_t$ , in the context of the inflation data, represents trend inflation. The time-series behavior of this trend is characterized by  $E_t(Z_t) = 0$ ,  $E_t(Z_t^2) = \sigma_Z^2$  and  $E_t(Z_t Z_s) = 0$  for all  $t \neq s$ . Further, assume that, for all  $i$  and  $t$ ,  $E(x_{it} - Z_t)^2 = \mu_2$ ,  $E(x_{it} - Z_t)^3 = 0$ ,  $E(x_{it} - Z_t)^4 = \mu_4$ , and  $E(x_{it} - Z_t)^6 = \mu_6$ . Independence implies that  $E[(x_{it} - Z_t)^k (x_{js} - Z_s)] = 0$ , for all  $i \neq j$  and  $t \neq s$ .<sup>5</sup>

Following the notation established in Section III, the central sample moments computed over the cross section are written as  $m_{rt}$  but with equal weighting. That is, the  $w_i$ 's of equations (2) and (3), are all set to  $\frac{1}{N}$ , while the component inflation measures are replaced by the draws  $x_{it}$ , and so

$$m_{rt} \equiv \frac{1}{N} \sum_{i=1}^N (x_{it} - m_{1t})^r, \quad (5)$$

where  $m_{1t}$  has the natural definition as the sample mean, corresponding to average inflation ( $\pi_t$ ) in equation (2). Equation (4) defines the sample skewness. We define the time-series averages of the sample moments as

$$\bar{m}_r = \frac{1}{T} \sum_{t=1}^T m_{rt}. \quad (6)$$

Equivalently, the time-series average of the skewness is  $\bar{S}$ .

The statistic of interest is the correlation over  $t$  of the sample mean and skewness,

---

<sup>5</sup>In this case, the i.i.d. assumption is rather innocuous. To see why, simply note that if the  $x_{it}$ 's are i.n.i.d., we would obtain the same results as in the i.i.d. case, but the sample would be drawn from a mixture of all of the distributions assumed in the heteroskedastic case. That is to say, we could consider a sample  $\{x_1, \dots, x_N\}$  drawn from a distribution with second and fourth moments  $\mu_{2i}$  and  $\mu_{4i}$ , or a sample of  $N$  i.i.d. draws from a distribution that is a mixture of distributions in which we draw with probability  $1/N$  from each of the original  $N$  distributions.

each computed over  $i$ . In the notation we have just established, this correlation is

$$\rho_{T,N} = \frac{\frac{1}{T} \sum_{t=1}^T (m_{1t} - \bar{m}_1)(S_t - \bar{S})}{\left[ \frac{1}{T} \sum_{t=1}^T (m_{1t} - \bar{m}_1)^2 \frac{1}{T} \sum_{t=1}^T (S_t - \bar{S})^2 \right]^{1/2}} \quad (7)$$

Assuming that  $\mu_2$  is known, we evaluate  $\rho_{T,N}$  taking the probability limit as  $T$  goes to infinity, which is<sup>6</sup>

$$\text{plim}_{T \rightarrow \infty} \rho_{T,N} = \frac{E(m_{1t}m_{3t})}{[E(m_{1t}^2)E(m_{3t}^2)]^{1/2}}. \quad (8)$$

For the i.i.d. case, these three terms are

$$\begin{aligned} E(m_{1t}m_{3t}) &= \left( \frac{1}{N} - \frac{3}{N^2} + \frac{2}{N^3} \right) (\mu_4 - 3\mu_2^2), \\ E(m_{1t}^2) &= \frac{\mu_2}{N} + \sigma_Z^2, \\ \text{and,} \\ E(m_{3t}^2) &= \frac{\mu_6 + 9\mu_2^3 - 6\mu_4\mu_2}{N}. \end{aligned} \quad (9)$$

Making the substitutions, we find that

$$\bar{\rho}_N = \text{plim}_{T \rightarrow \infty} \rho_{T,N} = \frac{\left(1 - \frac{3}{N} + \frac{2}{N^2}\right) (\mu_4 - 3\mu_2^2)}{\left\{ \mu_2 \left(1 + N \frac{\mu_2}{\sigma_Z^2}\right) (\mu_6 + 9\mu_2^3 - 6\mu_4\mu_2) \right\}^{1/2}}. \quad (10)$$

This expression has several interesting properties. First, note that in the limit, as  $N$  goes to infinity, so long as  $\sigma_Z^2$  is nonzero, the correlation will go to zero. That is to say, the true value of the time-series correlation between the cross-sectional sample mean and skewness is zero in population so long as the cross-sectional mean varies over time. There is, however, a small-sample bias that is a complex function of the higher-order moments of the data-generating process and the sample size.

We pause for a moment to consider the case in which the time-series mean has no variance in population — that is, the case in which  $\sigma_Z^2$  is zero. Taking the limit of

---

<sup>6</sup>If we were to parameterize the temporal dependence of the  $x_{it}$ 's, then we would be able to derive the implied value of  $\rho_{T,N}$ . The resulting expression depends on terms that are of order  $\frac{1}{T}$  and are a function of the dependence in the first moments. Dependence of higher-order moments, that is, ARCH and the like, are never relevant here.

the expression in equation (10) as  $N$  goes to infinity, we no longer get zero. Note that the numerator and the denominator of this expression both have terms of order  $\frac{1}{N}$ . All three of the expressions are going to zero and they do so at the same rate. For this case,  $\text{plim}_{T \rightarrow \infty} \rho_{T,N}$  will *not* go to zero as  $N$  goes to infinity. As a result, the time-series correlation of the sample mean and skewness of data drawn from a process with no time-series variance converges to a constant that depends on the second, fourth, and sixth moments of the data-generating process. The *sign* of this correlation depends on the kurtosis of the distribution, while the size depends the kurtosis and other properties. In the case of price data that always exhibit excess kurtosis (that is,  $\mathcal{K} > 3$ ), the expected value of the correlation in population is positive. This is not a “bias” in the normal sense of the term, because the statistic converges to a value that is not zero. Nevertheless, ignoring this property of the correlation will surely lead to inferences that are incorrect. For example, a test of the hypothesis that an estimate of the correlation is significantly different from zero is unlikely to have the meaning intended.<sup>7</sup>

The special case in which the time-series mean of  $x$  has no variance is interesting because of its analytical similarity to the result that one obtains when considering the properties of a time-series regression of the cross-section mean on the cross-section skewness. Typically, researchers have tried to establish a relationship between these quantities by studying the properties of the regression

$$m_{1t} = \alpha + \beta S_t + \epsilon_t, \quad (11)$$

where  $\alpha$  and  $\beta$  are coefficients to be estimated and  $\epsilon$  a regression error. Again, assuming that  $\mu_2$  is known, it is straightforward to establish that

$$\bar{\beta} = \lim_{N \rightarrow \infty} \left( \text{plim}_{T \rightarrow \infty} \beta_{T,N} \right) = \frac{(\mu_4 - 3\mu_2^2)}{(\mu_6 + 9\mu_2^3 - 6\mu_4\mu_2)}, \quad (12)$$

which is zero only when the data-generating process has a kurtosis of 3. In a study of inflation data, a regression such as equation (11) would yield an estimated coefficient

---

<sup>7</sup>This caution should serve as a general warning about the use of sample correlations in cases in which the covariances and variances of the sample statistics go to zero at the same rate. Even in cases in which the covariances and variance go to zero asymptotically, their ratio need not.

Table 2: Mean-Skewness Correlation

	$k = 1$	$k = 3$	$k = 12$	$k = 24$
36 Components of the CPI Sample 1967:01 to 1997:04				
Sample value	0.35	0.35	0.41	0.47
Approximate Small-sample Bias	0.25	0.23	0.19	0.14
29 Components of the PPI Sample 1967:01 to 1997:04				
Sample value	0.37	0.29	0.15	0.14
Approximate Small-sample Bias	0.22	0.21	0.12	0.08

Notes: The sample value is the time-series correlation of the cross-sectional sample mean and sample skewness of the components of inflation as measured by the CPI or PPI. The approximate small-sample bias is the expected value of the correlation computed in text equation (10), evaluated at the time-series means of the second, fourth, and sixth moments of the data.

$\beta$  that is not zero in population. As a result, interpreting the simple finding that  $\beta$  is different from zero as evidence for or against any economic theory would be incorrect.

Returning to the issue at hand, we ask the following empirical question: How big is the bias in the mean-skewness correlation likely to be? We examine this question in two steps. First, using the actual data described in Section III, we compute the sample values for  $\bar{\rho}_N$ . We refer to this quantity as the “approximate small-sample bias” because it ignores covariance across  $i$ , as well as time-series dependence. This leads us to perform a series of Monte Carlo experiments based on a more comprehensive accounting of the stochastic properties of the data (see Section V). The Monte Carlo experiments allow us to fully characterize the small-sample distribution of  $\rho_{T,N}$  and thus to conduct accurate inference.

Evaluation of the expression for  $\bar{\rho}_N$  in equation (10) requires that we estimate the following quantities:  $\mu_2$ ,  $\mu_4$ ,  $\mu_6$ ,  $\sigma_Z$ , and  $N$ . These are the average second, fourth, and sixth moments of the data, as well as the variance of the time-series mean, and the sample size. For the first three quantities, we follow the procedure employed in the construction of Table 1 and use the time-series averages of the cross-sectional sample moments. We estimate  $\sigma_Z^2$ , the variance in trend inflation, as the variance in the thirty-six-month

centered moving average of sample mean inflation.

Determining the sample size is an interesting and subtle problem. In the case of unequal weighting, which is applicable to the inflation data,  $N$  is analogous to the inverse of the sum of the squared weights. That is,  $N \equiv \frac{1}{\sum_i w_i^2}$ . So long as any single component of the index retains finite weight, the effective sample size will be finite as well. For example, the weight on the home owner's rental-equivalence component of shelter in the CPI can never be reduced much below 0.2. As a result, the number of effective data points will never be more than about 25, and small-sample bias will be a problem regardless of how many prices the government samples.

Table 2 reports the sample value of the mean-skewness correlation, together with the approximate small-sample bias, for two data sets: the 36 components of the CPI and the 29 components of the PPI, both over the 1967:01 – 1997:04 sample.<sup>8</sup> Looking at the case of one-month changes ( $k = 1$ ), we find that the bias and sample values are of comparable magnitude. As we allow the horizon  $k$  to increase, the sample values are increasingly higher than the estimates from our analytical approximation. But this is the case in which the independence assumption is furthest from the truth. Overall, these results reveal a clear potential for substantial biases. Consequently, in the next section we proceed to the Monte Carlo experiments, which allow us to take explicit account of covariance across components and temporal dependence.

Before turning to those experiments, however, we should comment briefly on the correlation between the sample mean and the sample variance. This correlation has attracted even more attention than the mean-skewness correlation. The small-sample covariance of the mean and the variance,  $E(m_{1t}m_{2t})$ , depends on the third moment of the distribution from which the data are drawn and equals  $\left(\frac{1}{N} - \frac{3}{N^2} + \frac{2}{N^3}\right)\mu_3$ . The variance of the sample variance is  $E[(m_{2t} - \mu_2)^2] = \frac{(\mu_4 - \mu_2^2)}{N}$ , while the variance of the sample mean is given in (9). The approximate small-sample bias follows as  $\frac{(1 - \frac{3}{N} + \frac{2}{N^2})\mu_3}{\left[\mu_2 \left(1 + N \frac{\sigma_Z^2}{\mu_2}\right) (\mu_4 - \mu_2^2)\right]^{\frac{1}{2}}}$ . Clearly, the mean-variance correlation will be biased in small samples when the data are drawn from skewed distributions,  $\mu_3 \neq 0$ . When the distribution is symmetrical, which is the case in the price data we study here, there is virtually no bias. Calculations analogous

---

<sup>8</sup>We postpone discussion of the CPI excluding food and energy to Section VI below.

to those in Table 2, using time-series averages of the cross-sectional point estimates of  $\mu_2, \mu_3$ , and  $\mu_4$ , yield estimates of the approximate small-sample bias for the mean-variance correlation that are virtually zero. (See the appendix tables for a complete set of results.) For example, in the case of the CPI, with  $k = 1$ , the sample mean-variance correlation is 0.20 and the approximate small-sample bias is 0.01. Exact calculations based on Monte Carlo experiments, described below and reported in the appendix tables, show a bias that is somewhat larger, but still small enough to allow us to conclude that for the U.S. inflation data, the mean-variance correlation is in fact positive and significantly different from zero.<sup>9</sup>

## V. Exact Small-Sample Bias

Inflation indices exhibit both dependence over time and correlation across the components. The approximate small-sample bias derived analytically in Section IV ignores both of these potentially important characteristics of the data. For this reason, we use Monte Carlo techniques to compute the exact small-sample distribution of the mean-skewness correlation — a strategy that allows us to estimate the bias more accurately and to conduct inference.

We perform a series of three experiments. In each, we begin by specifying a data-generating process from which to draw simulated price change data and to compute mean inflation and cross-sectional skewness. Each random draw has the same size data matrix as the actual panel of inflation data,  $\{p_{it}^j\}$ , where  $i$  is the component,  $t$  is the time period, and  $j$  is the number of the replication.

Using these draws, we compute the sample moments of interest. These are the time-series mean inflation rates, given by

$$\pi_t^j = \sum_{i=1}^N w_i p_{it}^j, \quad (13)$$

---

<sup>9</sup>Roger (1997) notes that the New Zealand price data are substantially skewed, suggesting that the mean-variance correlation may be biased in that case.

and the higher-order moments needed to compute the skewness,

$$m_{rt}^j = \sum_{i=1}^N w_i (p_{it}^j - \pi_t^j)^r . \quad (14)$$

The weights,  $w_i$ , are equal to those employed in the construction of the price indices. Using these, we compute the correlation, over time, of the sample mean and the sample skewness.

The three experiments are outlined below:

1. The first assumes that the component inflation series are all normally distributed, **serially independent**, but correlated cross-sectionally. For each time period, we draw a vector of length  $N$  from a multivariate normal distribution  $N(0, \Sigma)$ , where  $\Sigma$  is the covariance matrix of unconditional distribution of the actual price-change data. For this case, the actual time-series draws will each be mean zero.
2. The second assumes that component inflation series are normally distributed, **serially correlated**, and correlated cross-sectionally. Each panel draw is built up in the following way: First, the components are modeled as AR(12) processes with innovations. These autoregressions are estimated from the actual data, yielding a set of parameter vectors and a matrix of residuals. The panel of residuals is used to compute an estimated variance-covariance matrix for the system,  $\hat{\Sigma}$ . A sequence of draws is then taken from a multivariate  $N(0, \hat{\Sigma})$ . The draws are used recursively to construct the component data series, with the first 12 actual values of each series being used to construct the first artificial time-series observation. Beginning with the actual values from the data implies that we do not restrict average inflation to be zero.
3. The third uses a **bootstrap** methodology. This approach is the same as the previous one, except that the draws are taken from the actual innovation series (with replacement), not from a multivariate normal distribution.

Finally, in addition to monthly changes, we consider overlapping 3-, 12-, and 24-month changes. Throughout, we construct the actual p-values using robust methods. For the serially independent cases, White's (1980) method is used. For the serially correlated

and bootstrap cases, Newey and West’s (1987) procedure with bandwidth parameter set equal to  $1.4k$  is used.

We present results for three data sets: (1) 36 components of the CPI, monthly, from 1967 to 1997; (2) 29 components of the PPI, monthly, from 1967 to 1997; and (3) the Ball and Mankiw annual PPI data, which have a maximum of 301 components and a sample period from 1947 to 1989. For each of these, we present a table that includes the following results: the sample value in the data of the correlation between the mean and the skewness of inflation; the asymptotic p-value computed from the standard statistic (that is, the t-ratio) under the assumption that the asymptotic distribution is accurate; the approximate small-sample bias computed using equation (10); the exact small-sample bias estimated as the median of the empirical distribution; the bias-adjusted correlation (the actual minus the median of the empirical distribution); and finally, an empirical p-value for the test that the correlation observed in the data is equal to the median of the empirical distribution. All results are based on experiments with 2500 replications, and all p-values correspond to two-sided hypothesis tests.

Table 3 reports the results for the monthly consumer price data. In the base case, assuming serial independence, we estimate the exact small-sample bias of the mean-skewness correlation to be about 0.34 at all frequencies — always larger than the value computed using the analytic approximation. These yield estimates of the “bias-adjusted” correlations that range from 0.01 for monthly changes ( $k = 1$ ) to 0.12 for 24-month changes ( $k = 24$ ). The empirical p-values are uniformly large, implying that we cannot reject the null hypothesis that the bias-adjusted correlation differs from zero. The exact small-sample bias diminishes for the more complicated covariance structures — but not by much. The strongest small-sample bias-adjusted correlations are found for 12- and 24-month changes, although we can still easily reject that these correlations are statistically meaningful.

One might ask if there is a simple rule-of-thumb that would allow us to avoid running these complex and time-consuming Monte Carlo experiments. One possibility would be to compute the p-value for the test that the observed correlation equaled the approximate small-sample bias, using the sample standard error. For the  $k = 1$  case, this procedure would involve subtracting 0.22, the approximate small-sample bias, from 0.35, the sample

Table 3: Monte Carlo Results for Correlation between Mean and Skewness  
Consumer Prices, 36 Components, Monthly, 1967–97

Experiment	Statistic	$k = 1$	$k = 3$	$k = 12$	$k = 24$
Sample values	Correlation	0.35	0.35	0.41	0.47
	Asymptotic p-value	0.00	0.00	0.00	0.00
	Approximate small-sample bias	0.25	0.23	0.19	0.14
1. Serially independent	Exact small-sample bias	0.34	0.34	0.34	0.35
	Bias-adjusted correlation	0.01	0.01	0.06	0.12
	Empirical p-value	0.78	0.88	0.62	0.46
2. Serially correlated	Exact small-sample bias	0.38	0.36	0.21	0.10
	Bias-adjusted correlation	-0.03	-0.01	0.20	0.38
	Empirical p-value	0.64	0.92	0.23	0.10
3. Bootstrap	Exact small-sample bias	0.42	0.39	0.25	0.15
	Bias-adjusted correlation	-0.07	-0.04	0.16	0.33
	Empirical p-value	0.25	0.67	0.31	0.11

Notes: Sample correlation and approximate small-sample bias are defined as in Table 2. The asymptotic p-value is computed from a standardized regression of the sample skewness on the sample mean, using the Newey-West procedure with bandwidth parameter equal to  $1.4k$ . Monte Carlo experiments differ according to the assumed data-generating process as described in the text. The exact small-sample bias is the median of the 2500 Monte Carlo replications for each case. The bias-adjusted correlation is the sample value minus the exact small-sample bias. The empirical p-value is the percentage of the Monte Carlo replications greater (in absolute value) than the observed correlation.

Table 4: Monte Carlo Results for Correlation between Mean and Skewness  
 Producer Prices, 29 Components, Monthly, 1967–97

Experiment	Statistic	$k = 1$	$k = 3$	$k = 12$	$k = 24$
Sample values	Correlation	0.37	0.29	0.15	0.14
	Asymptotic p-value	0.00	0.00	0.00	0.00
	Approximate small-sample bias	0.22	0.21	0.12	0.08
1. Serially Independent	Exact small-sample bias	0.37	0.37	0.37	0.38
	Bias-adjusted correlation	0.00	-0.08	-0.23	-0.24
	Empirical p-value	0.95	0.10	0.02	0.11
2. Serially Correlated	Exact small-sample bias	0.41	0.41	0.34	0.38
	Bias-adjusted correlation	-0.04	-0.12	-0.20	-0.24
	Empirical p-value	0.33	0.07	0.15	0.20
3. Bootstrap	Exact small-sample bias	0.42	0.40	0.35	0.39
	Bias-Adjusted correlation	-0.05	-0.11	-0.20	-0.24
	Empirical p-value	0.30	0.11	0.15	0.23

Notes: See Table 3 for a description of the methods.

value, and dividing by 0.052, the sample standard error (from a standardized regression using the Newey and West procedure with bandwidth parameter set equal to 2). The resulting “bias-adjusted” estimate is 0.13, with a t-ratio of 1.92 and a p-value of 0.05. This value can be compared with the result for experiment 3, the bootstrap, reported in the first column of Table 3. There we find an exact small-sample bias-adjusted correlation of  $-0.07$ , with a p-value of 0.25. Thus, the estimate obtained using the simplified procedure is sufficiently inaccurate to justify the use of the exact procedure.

The experiments with the PPI data, reported in Table 4, prove even less conclusive. All of the estimated exact small-sample bias-adjusted correlation coefficients are now negative, and virtually all of the empirical p-values are in excess of 0.10.

Table 5 reports the results using the more detailed PPI data set of Ball and Mankiw. In the case of the unbalanced 301 component data set (in which the number of price components changes over the sample period), the exact small-sample bias-adjusted correlation between the mean and skewness of the price change distribution is essentially zero when serial independence is assumed, but statistically significantly negative (correlation coefficient of  $-0.28$ ) when serial correlation is introduced. Because of the unbalanced

Table 5: Monte Carlo Results for Correlation between Mean and Skewness  
Ball-Mankiw Producer Prices

Experiment	Statistic	Unbalanced Panel	Balanced Panel
		301 Components 1947–89	201 Components 1967–89
Sample values	Correlation	0.43	0.35
	Asymptotic p-value	0.00	0.07
	Approximate small-sample bias	0.14	0.11
1. Serially independent	Exact small-sample bias	0.42	0.58
	Bias-adjusted correlation	0.01	-0.23
	Empirical p-value	0.96	0.19
2. Serially correlated	Exact small-sample bias	0.71	0.56
	Bias-adjusted correlation	-0.28	-0.21
	Empirical p-value	0.00	0.06
3. Bootstrap	Exact small-sample bias	—	0.63
	Bias-adjusted correlation	—	-0.28
	Empirical p-value	—	0.13

Notes: All data are from Ball and Mankiw (1995), restricted to include only those series with nonzero weight in 1982 and with at least six years of data. For all experiments,  $k$  equals one year and each series is assumed to follow a first-order autoregression. See Table 3 for a description of the methods.

nature of the data set, bootstrap experiments are extremely difficult to construct from these data. However, if we consider only a balanced data set, in which the number of components is constant over the sample period (201 components from 1967 to 1989), a full set of experiments can be conducted. The results of this exercise, reported in the second column of the table, are essentially the same as those obtained from using the more comprehensive data set. The exact small-sample bias-adjusted correlation coefficients are found to be insignificantly different from zero, and in the cases that allow for serial correlation, the point estimates are negative.

To the extent that we are able to identify any relationship between the mean and skewness of the price-change distribution, it would appear to be a negative one, although the evidence on this point is quite weak. This result is intriguing because it is consistent with the existence of a nominal rigidity at zero. If such a nominal rigidity existed at zero inflation, it would have implications for the relationship of the mean to the skewness of price-change distributions. Specifically, one would expect that as inflation falls toward zero, the cross-sectional distribution of price changes would have increased mass near zero and, consequently increased skewness. The result is a *negative* correlation between the mean and skewness of inflation.<sup>10</sup>

## VI. Robustness

To confirm the robustness of the conclusions reported in section V, we undertake two tests. First, by examining subsample estimates, we demonstrate that the results are not sensitive to the possible existence of changes in the inflation process over time. And second, by excluding food and energy prices, we show that the correlations are unlikely to be artifacts of “real shocks.”

In Table 6 we report the consequences of separating the data sample into two subsamples — one covering the 1967-81 period of high and volatile inflation, and the other covering 1982-97 period of low and stable inflation. First, note that the observed corre-

---

<sup>10</sup>See, for example, the work of Crawford and Dupasquier (1994) on Canada. Because they find no relationship between the sample mean and sample skewness of wage inflation, from which they conclude that there is no downward wage rigidity. Akerlof, Dickens, and Perry (1996) is a more recent contribution to the literature on the importance of inflation in facilitating relative price and wage changes.

Table 6: Bootstrap Monte Carlo Results for Correlation between Mean and Skewness

CPI, All Items, Subsamples

Statistic	$k = 1$	$k = 3$	$k = 12$	$k = 24$
	<u>Full Sample</u>			
Actual	0.35	0.35	0.41	0.47
Asymptotic p-value	0.00	0.00	0.00	0.00
Approximate small-sample bias	0.25	0.23	0.19	0.14
Exact small-sample bias	0.34	0.34	0.34	0.35
Bias adjusted correlation	0.01	0.01	0.06	0.12
Empirical p-value	0.78	0.88	0.62	0.46
	<u>1967-81</u>			
Actual	0.24	0.20	0.36	0.45
Asymptotic p-value	0.00	0.03	0.02	0.01
Approximate small-sample bias	0.24	0.27	0.22	0.18
Exact small-sample bias	0.28	0.29	0.19	0.12
Bias adjusted correlation	-0.03	-0.09	0.17	0.33
Empirical p-value	0.65	0.50	0.43	0.31
	<u>1982-97</u>			
Actual	0.48	0.52	0.49	0.33
Asymptotic p-value	0.00	0.00	0.00	0.10
Approximate small-sample bias	0.34	0.38	0.33	0.26
Exact small-sample bias	0.44	0.42	0.45	0.41
Bias adjusted correlation	0.04	0.10	0.04	-0.09
Empirical p-value	0.42	0.08	0.67	0.68

Notes: See Table 3 for a description of the methods.

lations are generally larger in the later period. Indeed, for monthly changes ( $k = 1$ ), the correlation in the post-1981 period is twice what it is in the pre-1982 sample. Even so, the results from the bootstrap experiments using consumer price data confirm our earlier findings that the bias-adjusted correlation is never statistically significantly different from zero.

For our second test, we exclude food and energy prices from the data set in bootstrap experiments over the full sample period and two subsamples. Comparing the results reported in Table 7 with those in Table 6, we find that the sample values of the correlation are a bit smaller when food and energy are omitted. For example, when  $k = 1$ , the correlation for the full sample period falls from 0.35 to 0.25. But the p-value using the standard (unadjusted) procedure is still 0.00.

Excluding food and energy does reduce the bias somewhat. In the case of the full-sample estimates, the exact small-sample bias when  $k = 1$  falls from 0.34 to 0.29, yielding bias-adjusted estimates of the correlation that are 0.01 and  $-0.04$ , respectively. Needless to say, these estimates are not significantly different from zero. In fact, none of the bias-adjusted estimates of the mean-skewness correlation are both positive and significantly different from zero.

In sum, the magnitude of the estimated bias in the inflation-skewness correlation is approximately that observed in the data, and this result appears to hold for different time periods and for different data sets, including ones that exclude food and energy. Collectively, these results allow the following conclusions. The common belief that mean inflation and the cross-sectional skewness of price changes are positively related rests on a misinterpretation of the statistics. A consideration of the small-sample properties of the mean-skewness correlation shows that the observed levels in the data are wholly explained by the bias inherent in their calculation. Furthermore, the observed mean-skewness correlation provides no support for either the Ball-Mankiw sticky-price model or the Balke-Wynne flexible price model. In order to distinguish these models, we will need to look elsewhere.

Table 7: Bootstrap Monte Carlo Results for Correlation between Mean and Skewness

CPI, Excluding Food and Energy

Statistic	$k = 1$	$k = 3$	$k = 12$	$k = 24$
	<u>Full Sample</u>			
Actual	0.25	0.24	0.30	0.29
Asymptotic p-value	0.00	0.00	0.02	0.12
Approximate small-sample bias	0.17	0.15	0.09	0.07
Exact small-sample bias	0.29	0.24	0.23	0.20
Bias adjusted correlation	-0.04	0.01	0.07	0.09
Empirical p-value	0.39	0.95	0.58	0.64
	<u>1967-81</u>			
Actual	0.15	0.17	0.34	0.43
Asymptotic p-value	0.02	0.05	0.02	0.05
Approximate small-sample bias	0.19	0.18	0.11	0.08
Exact small-sample bias	0.28	0.27	0.29	0.24
Bias adjusted correlation	-0.13	-0.10	0.05	0.19
Empirical p-value	0.03	0.30	0.77	0.55
	<u>1982-97</u>			
Actual	0.38	0.35	0.36	0.36
Asymptotic p-value	0.00	0.02	0.02	0.15
Approximate small-sample bias	0.22	0.22	0.17	0.14
Exact small-sample bias	0.36	0.30	0.47	0.47
Bias adjusted correlation	0.02	0.04	-0.10	-0.10
Empirical p-value	0.69	0.42	0.39	0.60

Notes: See Table 3 for a description of the methods.

## VII. Conclusions

This paper identifies a small-sample bias in the correlation between inflation and its higher-order moments. This bias is of particular importance in the interpretation of the correlation between mean inflation and the cross-sectional skewness of price changes. Using a variety of Monte Carlo experiments, we show that the correlation observed in the data is fully accounted for by small-sample bias. In other words, we establish that one of the most accepted stylized facts in the literature on aggregate price behavior, that inflation and the skewness of the price-change distribution are positively correlated, need not be a fact at all.

We close by noting that the small-sample bias we have identified does not require us to reinterpret the evidence on the relationship between inflation and the cross-sectional variance of price changes, a correlation that has been the basis for many of the early studies in this area. Only when price-change distributions are asymmetrical on average will there be a small-sample bias in the mean-variance correlation. While may be true for some countries and in some periods, for the U.S. data we study here, it does not. If the correlations between inflation and its higher order moments can yield interesting insights that improve our understanding of the price-setting environment, then the inflation-variance relationship offers the most potential for future study.

## Appendix: Bias in the Mean-Variance Correlation

The tables in this appendix, focussing on the correlation between the sample mean and the sample cross-sectional variance of inflation, replicat the results in Tables 3 and 4 of Section V.

Table A1: Monte Carlo Results for Correlation between Mean and Variance

Consumer Prices, 36 Components, Monthly, 1967–97

Experiment	Statistic	$k = 1$	$k = 3$	$k = 12$	$k = 24$
Sample values	Correlation	0.20	0.27	0.56	0.70
	Asymptotic p-value	0.07	0.05	0.00	0.00
	Approximate small-sample bias	0.01	0.01	0.02	0.03
1. Serially independent	Exact small-sample bias	0.00	0.00	0.00	0.00
	Bias-adjusted correlation	0.20	0.27	0.56	0.70
	Empirical p-value	0.00	0.00	0.00	0.00
2. Serially correlated	Exact small-sample bias	-0.06	-0.08	-0.08	-0.04
	Bias-adjusted correlation	0.26	0.35	0.64	0.74
	Empirical p-value	0.00	0.00	0.00	0.00
3. Bootstrap	Exact small-sample bias	0.10	0.03	-0.04	-0.03
	Bias-adjusted correlation	0.11	0.24	0.60	0.73
	Empirical p-value	0.43	0.09	0.00	0.00

See Table 3 for a description of the methods.

Table A2: Monte Carlo Results for Correlation between Mean and Variance

Producer, 29 Components, Monthly, 1967–97

Experiment	Statistic	$k = 1$	$k = 3$	$k = 12$	$k = 24$
Sample values	Correlation	0.36	0.36	0.59	0.67
	Asymptotic p-value	0.01	0.00	0.00	0.00
	Approximate small-sample bias	0.00	0.01	0.01	0.01
1. Serially independent	Exact small-sample bias	0.00	0.00	0.01	0.01
	Bias-adjusted correlation	0.36	0.36	0.58	0.66
	Empirical p-value	0.00	0.00	0.00	0.00
2. Serially correlated	Exact small-sample bias	-0.01	-0.01	-0.02	-0.01
	Bias-adjusted correlation	0.37	0.37	0.61	0.68
	Empirical p-value	0.00	0.02	0.00	0.00
3. Bootstrap	Exact small-sample bias	0.13	0.02	0.03	0.06
	Bias-adjusted correlation	0.24	0.34	0.56	0.61
	Empirical p-value	0.11	0.10	0.02	0.03

See Table 3 for a description of the methods.

## References

- Akerlof, George A., William T. Dickens and George L. Perry, 'The Macroeconomics of Low Inflation,' *Brookings Papers on Economic Activity* (1996:1) 1-59.
- Ball, Laurence and N. Gregory Mankiw, 'Relative-Price Changes as Aggregate Supply Shocks,' *Quarterly Journal of Economics* 110 (February 1995) 161-193.
- Balke, Nathan S. and Mark A. Wynne, 'An Equilibrium Analysis of Relative Price Changes and Aggregate Inflation,' Research Department Working Paper 96-09, Federal Reserve Bank of Dallas, 1996 a.
- \_\_\_\_\_, and \_\_\_\_\_, 'Supply Shocks and the Distribution of Price Changes,' *Economic Review* of the Federal Reserve Bank of Dallas (First Quarter 1996b) 10-18.
- Bryan, Michael F. and Stephen G. Cecchetti, 'Measuring Core Inflation,' in *Monetary Policy*, N. Gregory Mankiw, ed., Chicago: University of Chicago Press for NBER, 1994, 195-215.
- Bryan, Michael F., Stephen G. Cecchetti and Rodney L. Wiggins II, 'Efficient Inflation Estimation,' N.B.E.R. Working Paper No. 6183, September 1997.
- Cecchetti, Stephen G., 'Measuring Short-Term Inflation for Central Bankers,' *Economic Review of the Federal Reserve Bank of St. Louis*, 79 (May/June 1997) 143-156.
- Crawford, Allan and Chantal Dupasquier, 'Can Inflation Serve as a Lubricant for Market Equilibrium?' in Bank of Canada, *Economic Behavior and Policy Choice Under Price Stability*, Ottawa, Ontario: Bank of Canada, June 1994.
- De Abreu Lourenco, Richard and David Gruen, 'Price Stickiness and Inflation,' Research Discussion Paper #9502 Reserve Bank of Australia, March 1995.
- Fischer, Stanley, 'Towards and Understanding of the Costs of Inflation: II,' *Carnegie-Rochester Conference Series on Public Policy*, 15 (1981): 5-42.
- Newey, Whitney K. and Kenneth D. West, 'A Simple, Positive Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix,' *Econometrica* 55 (May 1987): 703-708.
- Roger, Scott, 'A Robust Measure of Core Inflation in New Zealand, 1949-96,' mimeo., Reserve Bank of New Zealand, March 1997.
- Taylor, John B., 'On the Relationship between the Variability of Inflation and the Average Inflation Rate,' *Carnegie-Rochester Conference Series on Public Policy*, 15 (1981): 57-82.
- Vining, Daniel R., Jr. and Thomas C. Elwertowski, 'The Relationship between Relative Prices and the General Price Level,' *American Economic Review*, 66 (September 1976) 699-708.
- White, Halbert, 'A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity,' *Econometrica* 48 (May 1980) 817-38.