CROSS-SECTIONAL INFLATION ASYMMETRIES AND CORE INFLATION: A COMMENT ON BRYAN AND CECCHETTI

Randal J. Verbrugge*

Abstract—This paper reexamines the evidence relating core inflation to cross-sectional inflation asymmetry using statistical measures that are robust to the criticism of Bryan and Cecchetti. The results here suggest that there does exist significant positive correlation between core inflation and cross-sectional inflation asymmetry, but only at the monthly frequency. Furthermore, a sampling problem is highlighted which underscores the importance of careful Monte Carlo analysis when exact small-sample distributions are unknown.

I. Introduction

ONE GOAL of Bryan and Cecchetti's work (1999) is to shed light on the following question: Is core inflation positively correlated with cross-sectional inflation asymmetry? They show that the observed positive mean-skewness correlation could result entirely from a small-sample property that plagues moments: a single outlier will tend to significantly shift both the sample mean and skewness in the same direction. However, not all measures of central tendency and asymmetry suffer from this property. Hence, whether or not core inflation is positively correlated with cross-sectional inflation asymmetry remains an open question. The purpose of this note is to explore this question further. For brevity, I focus only on the CPI data.

For measures of "central tendency" of cross-sectional price changes (or core inflation), I utilize the (unweighted) median and the (unweighted) trimmed mean;[1] as the measure of cross-sectional asymmetry, I utilize the nonparametric triples U-statistic of Randles et al. (1980). The median and the triples U-statistic are both robust to, and the trimmed mean is partially robust to, outliers. Applying these to monthly CPI data, conventional analysis indicates a positive and statistically significant relationship between cross-sectional inflation asymmetry and core inflation (as measured by the median). Monte Carlo investigation indicates that neither the Bryan and Cecchetti small-sample problem, nor several other potentially troubling data characteristics (heteroskedasticity, kurtosis, asymmetry,[2] serial correlation, and conditional heteroskedasticity), would generate spurious results with these robust measures. This suggests that the observed correlation between the cross-sectional asymmetry and core inflation is indeed both positive and statistically significant, the Bryan and Cecchetti criticism notwithstanding.[3]

However, there is a sample characteristic (or problem) of a different nature that can readily generate this (or any) positive correlation, and which appears in this data: the cross-sectoral heteroskedasticity interacting with cross-sectoral covariance. Consider k random variables drawn i.i.d. from an $N(0, \Sigma)$ distribution. Given an appropriate choice of Σ, any correlation between the median and the asymmetry may be generated, even though both the unconditional median and unconditional third moment are zero. The intuition is provided by the following example. Imagine five random variables, labeled 1 to 5 and distributed normally with zero mean and with $\sigma_k^2 = 2\sigma_{k-1}^2$, $k = 5, \ldots, 2$. Select the off-diagonal elements of Σ so that the implied correlation matrix is given by

$$\Gamma := \begin{bmatrix} 1 & & & & \\ 0.95 & 1 & & & \\ 0.95 & 0.95 & 1 & & \\ -0.95 & -0.95 & -0.95 & 1 & \\ -0.95 & -0.95 & -0.95 & 0.95 & 1 \end{bmatrix}.$$

Most samples will be skewed one way or the other, and typically the median (frequently variable 3) will diverge from zero in the same direction, generating a high median-asymmetry correlation. Generalizations of this argument underscore the importance of computing the small sample distribution using Monte Carlo methods when analytical results are unavailable for the statistical measures of interest, even when the underlying multivariate distribution is free from kurtosis.[4] Furthermore, careful attention must be paid to selecting the appropriate model for this analysis.

Once the cross-sectoral covariance structure in the monthly CPI data is adequately approximated, the median-asymmetry correlation declines considerably in significance.[5] Further, no significant positive correlation is found between core inflation and cross-sectoral asymmetry at quarterly or annual frequencies.

Received for publication May 26, 1998. Revision accepted for publication December 15, 1998.

* Virginia Polytechnic Institute and State University.

Thanks are due to Mike Bryan and Steve Cecchetti for sharing their data, for useful e-mail interchanges, and for inspiring this paper; and to Rick Ashley, Clint Coakley, Russ Murphy, Stuart Rosenthal, two anonymous referees, and the editor. Any errors herein are entirely my own.

[1] The trimmed mean is computed by "trimming" the top and bottom fraction r of the observations prior to computing the mean. This measure has recently been advocated (see, e.g., Bryan et al., 1997) as a good high-frequency measure of core inflation. Here, $r = 0.08$.

[2] Virtually all industries feature significant asymmetry in their price change series (see Verbrugge, 1998).

[3] Note that it would be incorrect to conclude that this finding would support the sticky-price model of Ball and Mankiw (1995). In fact, their model predicts no positive relation between median inflation and cross-sectional asymmetry. (I am indebted to Steve Cecchetti for pointing this out to me.)

[4] Having said this, k-independent draws from a distribution featuring kurtosis is very similar to drawing a k-vector from a multivariate distribution featuring heteroskedasticity. (This insight is due to Steve Cecchetti.)

[5] Note that the covariance structure may well be of interest in its own light.

The Review of Economics and Statistics, May 1999, 81(2): 199–202

II. Methodology

The data used, and the definitions of inflation rates, are the same as those in Bryan and Cecchetti (1999); so, for example, quarterly inflation is defined as

$$\dot{p}_{it} := (1/k)\ \ln\,(p_{it}/p_{it-3}),$$

defined for $t = 4, 7, 10, \ldots$. Aside from the triples U-statistic (which has been infrequently used in the economics literature[6] despite its high power and prominent reputation amongst statisticians), the definitions of other statistics are standard. The intuitive basis of the triples statistic is the following: Take the sample of size N and examine all possible triples of members of the sample (i.e., examine $\binom{N}{3}$ combinations). If "most" of these triplets are right-skewed, infer that this is true of the underlying distribution. Formally, a triple of observations (X_i, X_j, X_k) is a *right triple* (is skewed to the right) if the middle observation is closer to the smaller observation than it is to the larger. An example of a right triple:

—XX———X—

Let

$$\begin{aligned} f^*(X_i, X_j, X_k) := \tfrac{1}{3}\,[&sign\,(X_i + X_j - 2X_k) \\ &+ sign\,(X_i + X_k - 2X_j) \\ &+ sign\,(X_j + X_K - 2X_i)]. \end{aligned} \qquad (2.1)$$

The range of this function is $\{\frac{1}{3}, 0, \frac{1}{3}\}$; a right triple is a triple which maps into 1/3, and a left triple is defined analogously. The triples U-statistic is given by

$$\hat{\eta} := \frac{1}{\binom{N}{3}} \sum_{i<j<k} f^*\,(X_i, X_j, X_k). \qquad (2.2)$$

One may generate an asymptotically normal test statistic by dividing $\hat{\eta}$ by various functions of the f^*. As the test statistic itself is not used in this paper, the formulae are omitted for brevity; the interested reader is referred to Randles et al. (1980). As there is no simple way to apply weights to the triples statistic, all statistics in this note are unweighted.

To test for significant correlation, I investigate two different measures of relation. First, I compute the simple correlation between the median and the triples U-statistics and determine whether it is significantly different from zero (computing p-values using Newey-West (1987) standard errors). Second, I conduct regression analysis (restricting attention to the median and the triples U-statistic); the usual practice is that a significant t-statistic is taken to be evidence of significant correlation. I regress the triples U-statistic on a constant, on the median, and on lags of the median and the triples U-statistic. In these regressions, the number of lags are chosen according to the Schwarz-Bayes information criterion. Finally, using five different Monte Carlo experiments, I examine the robustness of the correlation and the regression statistics. The first such experiment assumes that the component inflation series are all normally distributed, serially independent, and cross-sectionally independent; for each time period, each sector j receives a draw from a normal distribution $N(c_j, \sigma_j^2)$, where c_j and σ_j^2 are estimated from the CPI data. The second experiment alters this by allowing cross-sectoral correlation: each period, an N-vector of shocks is chosen from a multivariate normal distribution $N(c, \Sigma)$, where the vector of constants c and the variance-covariance matrix Σ are estimated from the data. The third uses a bootstrap methodology: an ARMA process is estimated for each sector (chosen according to the Schwarz-Bayes information criterion, subject to the constraint that the estimated process be stationary[7]), with draws taken from the actual innovation series (but independently across sectors). The fourth estimates an AR process for each sector (with number of lags chosen according to the Schwarz-Bayes information criterion, as above), allowing GARCH errors for all sectors that display conditional heteroskedasticity; all variance parameters are estimated from the data, with the number of variance terms chosen by a variant of step-down testing, subject to the constraint that estimated variance processes be sensible.[8,9] Finally, the fifth estimates AR processes for each sector, but in addition allows average inflation to enter with lags; the number of lags is chosen according to the Schwarz-Bayes information criterion as above. Each process is run long enough to shed the effects of initial conditions.

III. Empirical Results: Cross-Sectional Median Inflation and Inflation Asymmetry

Correlation results are reported in table 1. At the monthly frequency, the median-triples correlation is significant at conventional critical values; experiments 1 through 4 indicate that this correlation is significant[10]; only when cross-sectoral correlation is reasonably accounted for (in experiment 5) does the median-asymmetry correlation in the data decline in significance.[11] Using the triples U-statistic instead of skewness does not eliminate the small-sample problem

[6] However, see Verbrugge (1997, 1998).

[7] For each sectoral inflation series, the null of a unit root is rejected using standard unit root tests.

[8] Unrestricted regressions often estimate models in which the variance can become negative or is explosive.

[9] As almost no sectors display evidence of conditional heteroskedasticity at the annual frequency, GARCH models were estimated only for monthly and quarterly data.

[10] The experiments featuring cross-sectional independence invariably predict low median-triples correlations.

[11] However, I should point out that unreported analysis using a near-VAR (with thirteen own lags, and three lags of all other sectors) suggests that the analysis above might be generating a median-triples correlation that is too large. This discrepancy underscores the importance of carefully selecting one's Monte Carlo experiment(s).

TABLE 1.—MONTE CARLO RESULTS FOR CORRELATIONS CONSUMER PRICES, 36 COMPONENTS, 1967–1997

Experiment	Statistic	$k = 1$	$k = 3$	$k = 12$
Data	Correlation (median, triples U-statistic)	0.32	0.29	0.27
	Newey-West p-value	0.05	0.33	0.55
1. Serially independent, CS-independent shocks	Median correlation (median, triples-U)	0.05	0.04	−0.08
	Empirical p-value	0.00	0.01	0.07
2. Serially independent, CS-dependent shocks	Median correlation (median, triples-U)	0.17	0.23	0.59
	Empirical p-value	0.00	0.44	0.02
3. ARMA bootstrap	Median correlation (median, triples-U)	0.18	0.11	0.08
	Empirical p-value	0.01	0.04	0.27
4. Serially correlated, CSI-GARCH errors	Median correlation (median, triples-U)	0.10	0.08	—
	Empirical p-value	0.00	0.02	—
5. Lags of mean in AR, CS-dependent shocks	Median correlation (median, triples-U)	0.23	0.25	0.46
	Empirical p-value	0.17	0.62	0.15
Data	Correlation (mean, triples U-statistic)	0.67	0.59	0.49
	Newey-West p-value	0.00	0.01	0.26
1. Serially independent, CS-independent shocks	Median correlation (mean, triples-U)	0.72	0.71	0.60
	Empirical p-value	0.04	0.02	0.38
2. Serially independent, CS-dependent shocks	Median correlation (mean, triples-U)	0.59	0.57	0.77
	Empirical p-value	0.71	0.82	0.01
3. ARMA bootstrap	Median correlation (mean, triples-U)	0.78	0.76	0.76
	Empirical p-value	0.00	0.00	0.02
4. Serially correlated, CSI-GARCH errors	Median correlation (mean, triples-U)	0.74	0.69	—
	Empirical p-value	0.28	0.02	—
5. Lags of mean in AR, CS-dependent shocks	Median correlation (mean, triples-U)	0.67	0.61	0.70
	Empirical p-value	0.88	0.56	0.03
Data	Correlation (trimmed mean, triples-U)	0.48	0.42	0.37
	Newey-West p-value	0.00	0.09	0.42
1. Serially independent, CS-independent shocks	Median correlation (t-mean, triples-U)	0.45	0.42	0.29
	Empirical p-value	0.43	0.94	0.67
2. Serially independent, CS-dependent shocks	Median correlation (t-mean, triples-U)	0.37	0.37	0.68
	Empirical p-value	0.00	0.48	0.01
3. ARMA bootstrap	Median correlation (t-mean, triples-U)	0.59	0.52	0.48
	Empirical p-value	0.04	0.18	0.42
4. Serially correlated, CSI-GARCH errors	Median correlation (t-mean, triples-U)	0.51	0.47	—
	Empirical p-value	0.49	0.47	—
5. Lags of mean in AR, CS-dependent shocks	Median correlation (t-mean, triples-U)	0.46	0.41	0.58
	Empirical p-value	0.64	0.90	0.05

Notes: Reported are the time-series correlations of the (cross-sectional) sample median, mean, or trimmed mean (respectively) and the triples U-statistic. p-values for data are computed using the Newey-West procedure, with bandwidth chosen according to the Ljung-Box Q-statistic. "CS-dependent" refers to cross-sectionally dependent; "CSI" to cross-sectionally independent. Monte Carlo experiments differ according to the assumed data-generating process as described in the text; empirical p-values are two-sided p-values generated by 1,000 replications.

highlighted by Bryan and Cecchetti, nor does using the trimmed mean. At the quarterly frequency, no correlation is significant. Note that, as in Bryan and Cecchetti, correlations in the data are statistically significantly low at the annual frequency.

Regression results are reported in table 2, and empirical p-values for key regression statistics in table 3. Regression analysis suggests a significant positive relationship between the median and asymmetry (which is unaffected by inclusion of the standard deviation), as measured by low p-values and

TABLE 2.—ASYMMETRY, MEDIAN, AND STANDARD DEVIATION REGRESSIONS

	Dependent Variable: Triples U-statistic								
	$k = 1$			$k = 3$			$k = 12$		
	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
constant	0.01	−0.02	−0.03	0.01	−0.01	−0.01	0.02	−0.02	−0.02
	(0.02)	(0.00)	(0.00)	(0.00)	(0.14)	(0.16)	(0.01)	(0.32)	(0.15)
triples U_{t-1}	0.32	0.24	0.25	0.14	0.06	0.06	−0.23	−0.44	−0.44
	(0.00)	(0.00)	(0.00)	(0.21)	(0.58)	(0.57)	(0.23)	(0.03)	(0.03)
median$_t$		6.55	5.92		6.10	6.05		11.42	9.94
		(0.00)	(0.00)		(0.00)	(0.01)		(0.02)	(0.06)
std. deviation$_t$			1.52			0.09			3.02
			(0.04)			(0.97)			(0.51)
$\bar{R}^2$	0.098	0.157	0.172	0.011	0.073	0.065	0.019	0.191	0.174
Ljung-Box Q-statistic	35.54	30.16	25.07	30.18	23.33	23.51	5.54	7.32	6.67
	(0.54)	(0.74)	(0.91)	(0.41)	(0.76)	(0.75)	(0.59)	(0.39)	(0.46)

Notes: p-values are in parentheses (for regression coefficients, these are computed using the Newey-West (1987) procedure). In each case, the number of lags is chosen according to the Shwarz-Bayes information criterion.

TABLE 3.—MONTE CARLO RESULTS FOR REGRESSIONS

Experiment	Statistic	$k = 1$	$k = 3$	$k = 12$
2. Serially independent, CS-dependent shocks	Median coefficient on median	3.42	5.05	18.41
	p-value	0.00	0.56	0.16
	Median $\bar{R}^2$	0.03	0.04	0.32
	p-value	0.00	0.50	0.37
	Median $\Delta\bar{R}^2$	0.03	0.04	0.32
	p-value	0.14	0.64	0.32
3. ARMA bootstrap	Median coefficient on median	6.92	5.15	5.23
	p-value	0.87	0.88	0.62
	Median $\bar{R}^2$	0.04	0.01	−0.02
	p-value	0.03	0.08	0.07
	Median $\Delta\bar{R}^2$	0.03	0.00	−0.01
	p-value	0.18	0.07	0.04
5. Lags of mean in AR, CS-dependent shocks	Median coefficient on median	3.60	5.36	20.38
	p-value	0.02	0.73	0.15
	Median $\bar{R}^2$	0.10	0.06	0.25
	p-value	0.13	0.73	0.70
	Median $\Delta\bar{R}^2$	0.02	0.03	0.24
	p-value	0.02	0.39	0.61

Notes: Reported are results from regression (2) in table 3 on simulated data, with $\Delta\bar{R}^2$ referring to the increment in $\bar{R}^2$ moving from regression (1) to regression (2). Monte Carlo experiments differ according to the assumed data-generating process as described in the text. Empirical p-values are two-sided p-values generated by 1,000 Monte Carlo replications.

large contributions to $\bar{R}^2$; this holds up to Monte Carlo scrutiny at the monthly frequency, but not at the quarterly or annual frequencies.

I conclude that there is good evidence that core inflation is positively and significantly related to cross-sectional inflation asymmetry at the monthly frequency, and somewhat weaker evidence that core inflation is significantly negatively related to cross-sectional asymmetry at the annual frequency (once bias is accounted for). However, are these findings economically significant? Although it is perilous to try to answer such questions without the benefit of an appropriate model, I believe the null hypothesis should be that they are not. In absolute terms, particularly at the monthly frequency, the estimated coefficients are not very different from those of the Monte Carlo experiments. The gap in correlations is larger at the annual frequency, but this correlation is estimated very imprecisely in the data.

IV. Conclusion

Is higher inflation positively and significantly correlated with asymmetry in price changes across sectors? The evidence presented here suggests that this popular wisdom is correct, but only at the monthly frequency. At the annual frequency, the reverse appears to hold. However, it is unclear whether these findings have economic significance.

This analysis also highlights a sampling problem related to cross-sectoral heteroskedasticity and covariance that underscores the importance of careful Monte Carlo analysis when the exact small-sample distributions of the desired statistics are unknown.

REFERENCES

Ball, Laurence, and N. Gregory Mankiw, "Relative-Price Changes as Aggregate Supply Shocks," *Quarterly Journal of Economics* 110 (February 1995), 161–193.

Bryan, Michael F., and Stephen G. Cecchetti, "Inflation and the Distribution of Price Changes," *The Review of Economics and Statistics* 81 (May 1999), 188–196.

Bryan, Michael F., Stephen G. Cecchetti, and Rodney L. Wiggins, II, "Efficient Inflation Estimation," Federal Reserve Bank of Cleveland Working Paper No. 97.07 (1997).

Newey, Whitney K., and Kenneth D. West, "A Simple, Positive Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix," *Econometrica* 55 (May 1987), 703–708.

Randles, R., M. Flinger, G. Policello, and D. Wolfe, "An Asymptotically Distribution-Free Test for Symmetry versus Asymmetry," *Journal of the American Statistical Association* 75 (1980), 168–172.

Verbrugge, Randal, "Investigating Cyclical Asymmetries," *Studies in Nonlinear Dynamics and Econometrics* 2.1 (April 1997), 15–22.

——— "Longitudinal Inflation Asymmetries," Manuscript, VPI&SU (1998).