On the Granger causality between median inflation and price dispersion

Richard Ashley[a] and Haichun Ye[b,*]

[a] Department of Economics (0316), Virginia Tech, Blacksburg, Virginia 24061.
[b] School of Economics, Shanghai University of Finance and Economics, 777 Guoding Road, Shanghai 200433, China.

Running Head

Ashley and Ye: Median Inflation and Price Dispersion

Abstract

The Granger-causal relationship between the size and dispersion of fluctuations in sub-components of the U.S. Consumer Price Index (CPI) is examined using both in-sample and out-of-sample tests and data from January 1968 to December 2008. Strong in-sample evidence is found for feedback between median inflation and price dispersion; the evidence for Granger-causation from median inflation to price dispersion remains strong in out-of-sample testing, but is less strong for Granger-causation in the opposite direction. The implications of these results for the variety of price-level determination models in the literature are discussed.

Keywords: Granger causality, median inflation, price dispersion, out-of-sample testing

JEL Classification Code: C53, E31, E37

* Corresponding author. Email: haichunye@gmail.com

I. Introduction

The relationship between inflation and price dispersion has always been an important issue in macroeconomics. On the one hand, menu cost models, signal extraction models, and monetary search models have variously predicted a causal linkage from expected inflation, unexpected inflation, and inflation uncertainty to price dispersion. On the other hand, supply shock models argue that price dispersion can also have causal impacts on inflation. Surprisingly, while an overwhelming majority of the existing empirical work has examined the impacts of inflation on price dispersion, only limited research has been done on the potential reverse causality from price dispersion to inflation.

Our contribution in this paper is to provide a thorough investigation of Granger causality between inflation and price dispersion in both directions, employing both in-sample and out-of-sample Granger causality tests. Distinct from previous studies, which use the mean and standard deviation as measures of inflation and price dispersion, we measure inflation with the median change in the weighted log-price and price dispersion using the interquartile range of the weighted log-price; this choice is sensible because the cross-sectional distribution of prices is fat-tailed and skewed. Our in-sample Granger causality tests find strong evidence for feedback between median inflation and price dispersion. The in-sample evidence also suggests that the Granger causality from median inflation to price dispersion is mainly at high frequencies, corresponding to inflation fluctuations with periods of less than 3 to 6 months. Using a variety of out-of-sample Granger causality tests, we find strong evidence that median inflation has significant predictive content for price dispersion and very little evidence for Granger causality running from price dispersion to median inflation.

The rest of the paper is organized as follows. Section 2 provides a brief review of the literature on the relationship between inflation and price dispersion. In Sections 3 and 4, we describe the data used here and review the in-sample and out-of-sample Granger causality tests. Section 5 reports empirical results from our Granger causality tests; Section 6 offers concluding remarks.

II. Literature on Inflation and Price Dispersion

The theoretical literature on the impact of inflation on price dispersion is primarily built upon three types of models: menu cost models, signal extraction models and monetary search models.

In menu cost models (Sheshinski and Weiss, 1977, 1983; Rotemberg, 1983; Benabou, 1988; Diamond, 1992; Ball and Romer, 2003), firms are assumed to follow an (S,s) price adjustment rule. That is, a firm holds its nominal price constant as rising inflation erodes the real price, until – when the real price hits the lower bound ("s") – the firm adjusts its nominal price upward to restore the real price to the upper bound ("S"). Since firms are faced with different menu costs or firm-specific shocks, the (S,s) pricing policy can lead to staggered price changes in the economy. Under this staggered nonlinear price setting rule, an increase in inflation tends to raise relative price dispersion.[1]

In contrast to the menu cost models − which focus on expected inflation − the signal extraction models (Lucas, 1973; Barro, 1976; Hercowitz, 1981) emphasize the impacts of inflation uncertainty and unexpected inflation on price dispersion. According to the signal extraction models, higher inflation uncertainty and unexpected inflation make aggregate demand

[1] In fact, any nonlinear dynamic pricing rule can induce a feedback relationship between mean/median inflation and price dispersion. See Subramanian and Kawachi (2004, p. 79-80) for a similar argument in a different context.

shocks more unpredictable. Thus, firms respond less to all demand shocks (including idiosyncratic real demand shocks) with output adjustments, which, in turn, induce wider dispersion in relative prices.

In the monetary search models (Peterson and Shi, 2004; Head and Kumar, 2005), the effect of inflation on price dispersion works through the channel of consumers' search costs. On the one hand, higher expected inflation reduces the real value of fiat money and raises consumers' reservation price levels; this increases firms' market power and, consequently, price dispersion. On the other hand, widened price dispersion raises the gain from searching and thereby induces more consumer search, leading to a decrease in the dispersion of prices. In an environment of low inflation, a rise in inflation raises consumers' search intensity, and the second effect dominates, resulting in a fall in price dispersion. Where inflation rates are high, an increase in inflation dampens consumer's search intensity, and the first effect dominates, leading to a rise in price dispersion.

Compared to the abundance of theoretical work addressing the influence of inflation on the dispersion of prices, theories of the effect of price dispersion on inflation are more limited, coming mainly from the supply shock literature. Early studies – e.g., Tobin (1972) and Gordon (1975) – assume downward price rigidity and argue that, since nominal prices are rigid downward, an increase in price dispersion is inflationary. Ball and Mankiw (1995) later model firm's price-setting behavior under the New-Keynesian framework; they show that the skewness of price changes is positively associated with inflation and that larger price dispersion amplifies the effect of skewness on the inflation. Building upon the Ball-Mankiw model, Lourenco and Gruen (1995) further point out that the effect of price dispersion on inflation is contingent on the level of expected inflation: price dispersion is inflationary when expected inflation is higher, but not so when expected inflation is lower.

There are also a substantial number of empirical studies on the relationship between inflation and price dispersion. Notably, the majority of the existing empirical literature focuses almost entirely on the causal links from expected inflation, unexpected inflation and inflation uncertainty to price dispersion. Using different estimation methodologies, datasets and sample periods, these studies have variously found a positive effect, no effect, and even a negative effect of inflation on price dispersion.[2] In contrast, few studies have explored the potential causality from price dispersion to inflation.

Ashley (1981) tests for Granger causality between the CPI inflation and price dispersion in both directions by comparing the out-of-sample forecasting performances of univariate and bivariate time series models. He shows that inflation has predictive power for price dispersion but not vice-versa.[3] (In a related study, Fischer (1982) estimates vector autoregressive models for the United States and finds that relative price variability is an important determinant of inflation in the US.)

In this paper we provide a thorough investigation of the Granger causality between inflation and price dispersion (again in both directions), employing the latest kinds of both in-sample and out-of-sample tests and also taking advantage of the substantial amount of additional sample data now available.

[2] See, for example, Vining and Elwertowski (1976), Parks (1978), Reinsdorf (1994), Grier and Perry (1996), Parsley (1996), and Debelle and Lamont (1997).

[3] Ashley (1981) uses the mean and standard deviation as measures of inflation and price dispersion, respectively, and studies the Granger causality between these two variables over the period between January 1953 and June 1975 using a single out-of-sample test. In this study we use the sample median and interquartile range as measures of inflation and price dispersion, respectively, and look at a different sample period (January 1968 to December 2008) with both in-sample and a variety of out-of-sample tests. Despite the above differences, our findings with regard to the Granger causality between inflation and price dispersion are consistent with those in Ashley (1981), but richer due to the longer sample period and to the substantially more sophisticated testing techniques now available.

.

III. Measures of Inflation and Price Dispersion

While the sample mean of cross-sectional price changes is a conventional measure of inflation, Bryan and Cecchetti (1994) and Bryan, Cecchetti and Wiggins II (1997) argue that the weighted median is a better measure of inflation than the sample mean when the cross-sectional distribution of price changes is skewed and fat-tailed. Thus, we first examine the skewness and kurtosis of the cross-sectional distribution of price changes.

Using the 31 seasonally-unadjusted component price indices of the Consumer Price Index (CPI) that are available over the period from January 1968 to December 2008, we first compute the annualized monthly price growth rate as

$$\pi_{it} = 1200 * \ln(P_{it}/P_{it-1}) \qquad (1)$$

where P_{it} is the price index of component i at time t.[4] Next, we compute the skewness and kurtosis of the cross-sectional distribution of π_{it} as

$$S_t = \frac{\sum r_{it}(\pi_{it}-\pi_t)^3}{[\sum r_{it}(\pi_{it}-\pi_t)^2]^{\frac{3}{2}}} \text{ and } K_t = \frac{\sum r_{it}(\pi_{it}-\pi_t)^4}{[\sum r_{it}(\pi_{it}-\pi_t)^2]^2} \qquad (2)$$

where r_{it} is the relative importance of component i at time t and $\pi_t = \sum r_{it}\pi_{it}$.[5] The cross-sectional skewness (S_t) and kurtosis (K_t) are plotted against time in Figure 1. Several features are worth noting. First, the average skewness is close to zero (about 0.32) with a standard deviation

[4] These consumer price index components are obtained from the Bureau of Labor Statistics – see Appendix for details. The selection of the components used here is based on the availability of data over the entire sample period. Here we use seasonally unadjusted data simply because the usual seasonal adjustment procedures involve two-sided filtering which mixes together the past and future of a time series.
[5] Relative importance figures for the 31 components of the CPI index are obtained from the Bureau of Labor Statistics.

of 1.89. While there is thus little skewness in the distribution of π_t on average over the entire period, the sample distribution of π_{it} across the 31 individual components in any given month is generally skewed. Second, the average sample kurtosis of π_{it} (of 8.82) is large compared the value (of 3) expected for a Gaussian variate, implying that the cross-sectional distributions of monthly price growth rates generally have fat tails; in particular, the weighted kurtosis across the 31 components is in excess of 15 for about ten percent of the sample periods. Because the cross-sectional distributions of monthly CPI price growth rates across the components are generally skewed and fat-tailed, we follow Bryan and Cecchetti (1994) and Bryan, Cecchetti and Wiggins II (1997) and use the weighted sample median and interquartile range as our underlying measures of inflation and price dispersion, respectively, as these are the standard nonparametric measures of location and dispersion.[6] Because a histogram of the weighted sample interquartile range time series is highly skewed, with a shape resembling that of a chi-squared distribution, its logarithm is analyzed − and denoted "price dispersion" – below.

Figure 2 graphs the time paths of median inflation and price dispersion over the entire sample period (January 1968 to December 2008) using these measures. While there is a modest downward trend in median inflation, the price dispersion series generally fluctuates around a constant mean, with a couple of large spikes in the mid-1970s, the early 1980s and post-2005.[7]

---

[6] Given a series of N observations, v_1, v_2, …, v_N, ranked from the smallest to the largest, the weighted percentile for i-the observation is: $l_i = \frac{100}{CS_N}(CS_i - \frac{w_i}{2})$, where w_i is the weight of the i-th observation, $CS_N = \sum_{k=1}^{N} w_k$, and $CS_i = \sum_{k=1}^{i} w_k$. To find the value v corresponding to a given percentile l, we first find the observation number j where $l_j \leq l \leq l_{j+1}$ and then calculate the value as $v = v_j + \frac{l-l_j}{l_{j+1}-l_j}(v_{j+1} - v_j)$.

[7] The non-homgeneity in the variance of the median inflation variable evident in Figure 2a motivated some of the robustness checks described in Section 5.4 below but was, in the end, not problematic.

IV. Methodology

Sample Period Choices

Section IV describes the in-sample and out-of-sample Granger-causality tests which are used here to investigate the causal relationship between median inflation and price dispersion. With regard to sample period selection, the first twelve observations (comprising the sample data for 1968) are reserved for creating lagged variables; the 300 sample observations from January 1969 to December 1993 are used for model identification/estimation; and the remaining 180 observations, over the period from January 1994 to December 2008, are reserved for analyzing the out-of-sample forecasting performance of the models.[8]

Unrestricted and Restricted Models for the Two Time Series

Table 1 reports the results on unit root tests with regard to both median inflation and price dispersion over the sample period. All three unit root tests, including the Augmented Dickey-Fuller (ADF) test, Phillips-Perron test and Dickey-Fuller (DF) GLS test, suggest that the median inflation time series is trend stationary, whereas the price dispersion series is covariance stationary. For this reason a time trend is included in the models formulated below for the median inflation series.

To test for Granger causality from price dispersion to median inflation, we compare an unrestricted model of median inflation − which includes lags in price dispersion as explanatory variables − to a restricted model, in which lagged price dispersion variables are excluded. The change in the civilian unemployment rate is also included in both of these model specifications,

[8] This may seem like quite a large out-of-sample period (with 180 observations total), but this choice reflects the importance we attach to out-of-sample versus in-sample testing. Calculations in Ashley (2003) support the proposition that out-of-sample prediction period lengths in excess of 100 observations are worthwhile.

so as to control for potential Granger causality from changes in unemployment rate to both median inflation and price dispersion.[9]

The unrestricted model for median inflation is thus specified as follows:

$$y_t = \alpha_0 + \lambda t + \sum_{i=1}^{p} \alpha_i y_{t-i} + \sum_{j=1}^{q} \beta_j \Delta ur_{t-j} + \sum_{k=1}^{r} \gamma_k x_{t-k} + \varepsilon_{yt} \tag{1U}$$

where y_t and x_t are median inflation and price dispersion at time period t, respectively, and Δur_t is the change in unemployment rate. The restricted model for median inflation takes the form:

$$y_t = \alpha_0 + \lambda t + \sum_{i=1}^{p} \alpha_i y_{t-i} + \sum_{j=1}^{q} \beta_j \Delta ur_{t-j} + \nu_{yt} \tag{1R}$$

In a similar fashion we also estimate unrestricted and restricted models for price dispersion and test for Granger causality from median inflation to price dispersion:

$$x_t = \delta_0 + \sum_{i=1}^{h} \delta_i x_{t-i} + \sum_{j=1}^{m} \varphi_j \Delta ur_{t-j} + \sum_{k=1}^{s} \eta_k y_{t-k} + \varepsilon_{xt} \tag{2U}$$

$$x_t = \delta_0 + \sum_{i=1}^{h} \delta_i x_{t-i} + \sum_{j=1}^{m} \varphi_j \Delta ur_{t-j} + \nu_{xt} \tag{2R}$$

The lag lengths in these models are chosen so as to minimize the Bayesian information criterion (BIC) over the (in-sample) estimation period; the resulting estimated models are summarized in Table 2.

In-Sample Granger Causality Tests

While in-sample tests are clearly susceptible to pre-test distortion due to data mining, they are a useful first step in the Granger causality analysis, subject to confirmation (or dis-confirmation) via out-of-sample testing. Such out-of-sample testing − described in Section IV

[9] Seasonally un-adjusted monthly unemployment rate data are obtained from the Bureau of Labor Statistics.

below − both enhances the credibility of the in-sample results and also provides evidence that whatever effects founds are a "stable statistical regularity" over a substantially lengthier time period.

This in-sample testing step amounts to the usual F-test of the null hypothesis that a group of variables enters the unrestricted model with coefficients of zero; in-sample testing based on Sims (1972) is also reported below and also amounts to an F-test on a group of model coefficients. Testing for Granger causality dis-aggregated by frequency is described at the end of this section.

The in-sample test for Granger causality from price dispersion to median inflation is equivalent to testing the null hypothesis that the coefficients on the lagged values of the price dispersion variable entering the unrestricted model for the median inflation variable are all zero: a rejection of this null hypothesis indicates the existence of Granger causality running from price dispersion to median inflation.[10]

Similarly, the in-sample test for Granger causality from median inflation to price dispersion is equivalent to testing the null hypothesis that the coefficients on the lagged median inflation variables entering the unrestricted model for the price dispersion variable are all zero: a rejection of this null hypothesis indicates the existence of Granger causality running from median inflation to price dispersion.

[10] This test is, of course, only justified if the usual regression assumptions of homoskedastic and serially uncorrelated model errors are valid. Here sufficient lags are added to the model so that the correlogram of the fitting errors is consistent with serially uncorrelated model errors and the fitting errors are tested for heteroskedastictiy using both the Breusch-Pagan-Godfrey test and the White test. Because the homoskedasticity assumption is problematic, White-Eicker (robust) standard error estimates are used throughout.

The putatively 'causing' variables happen to enter the unrestricted models at only a single lag in the present instance. Consequently, the usual F statistic for the in-sample test has only one degree of freedom in the numerator and is just the square of the estimated t ratio with which the variable in question enters the unrestricted model.

In addition to the usual Granger causality test, we also implement the Sims (1972) causality test. Specifically, we regress median inflation on the lags and leads of price dispersion and then test the proposition that there is no Granger causality from median inflation to price dispersion with an F test of the null hypothesis that all of the coefficients on the leads of price dispersion are equal to zero. Rejection of this null hypothesis indicates there is a Granger causal link running from median inflation to price dispersion. A Sims test for Granger causality from price dispersion to median inflation is implemented here in a similar fashion.

We also assess the frequency dependence of the Granger causality between median inflation and price dispersion in the results reported below, by examining how the coefficient(s) on the putatively 'causing' variable(s) themselves vary across frequencies – i.e., across fluctuations in the 'causing' variable which are more (or less) persistent in nature. A number of methods for such assessment have been proposed in the literature – e.g., Breitung and Candelon (2006), Lemmens, et al. (2008), and Ashley and Verbrugge (2009) – but only the latter method remains valid in the presence of feedback between the time series. Because the in-sample results reported below in Section V indicate that feedback is likely present between median inflation and price dispersion, frequency dependence results are reported in Section V only for the Ashley and Verbrugge (2009) method.[11]

[11] "Feedback" between x_t and y_t is the case where both x_t fluctuations Granger cause future y_t fluctuations and y_t fluctuations Granger cause future x_t fluctuations. Because the Fourier transformation is a two-sided filter, it mixes

This method assesses the frequency dependence in a regression coefficient (on, for example, x_{t-1} in a model for y_t) by decomposing the data on x_{t-1} into *m* components, each associated with a sinusoidal frequency which corresponds to the reciprocal of the time period over which the fluctuations in this component tend to reverse themselves. Thus, the low-frequency (i.e., high period) components of x_{t-1} emphasize the highly persistent fluctuations in x_{t-1}, whereas the high frequency (i.e., low period) components of x_{t-1} emphasize the fluctuations in x_{t-1} which resemble random noise. (This is precisely the distinction between fluctuations in temporary versus permanent income in macroeconomic consumption theory.) The Ashley and Verbrugge (2009) method uses a backward-looking 36-month moving window to decompose the data on a monthly time series x_{t-1} into *m*=19 essentially uncorrelated components, corresponding to a zero-frequency (trend) component and 18 positive-frequency components, each of which corresponds to fluctuations with periods ranging from 2 to 36 months long. These 19 components precisely add up to x_{t-1}, so one need only replace x_{t-1} by these 19 components and estimate a coefficient (and standard error) for each.[12]

---

together both past and future values of a time series. Consequently, the application of frequency domain methods to data exhibiting feedback is fraught ─ see Ashley and Verbrugge (2007, Section 3.6) for a detailed exposition of this point.

[12] Windows software to easily perform this decomposition is available from the authors; see Ashley and Verbrugge (2007, 2009) for all of the analytical and calculational details. But it is worth mentioning here that the problematic effects of possible feedback are eliminated by only ever using the last filtered observation obtained from a 36-month window moving through the data set; thus, the decomposition is effectively a one-sided filter. There are only 18 possible non-zero frequencies possible with a 36-month long window (rather than 36 components) because the component related to the sine of each frequency can sensibly be aggregated with the corresponding cosine component. Thus, it is feasible to estimate all 19 possible component coefficients with a typical monthly data set. On other hand, this does 'use up' 18 additional 'degrees of freedom' in the regression and 36 sample observations are consumed by the initial window. Less importantly, fluctuations with periods in excess of 36 months cannot be distinguished from one another.

Out-of-Sample Tests for Improved Forecast Accuracy

As Ashley et al. (1980) points out, an out-of-sample comparison of forecasting performance is more in the spirit of the definition of Granger causality. The out-of-sample tests of Granger causality between price dispersion and median inflation used here are implemented in two steps. As a first step, we estimate both the restricted and unrestricted models for median inflation and for price dispersion. In the second step, we conduct formal statistical tests to examine whether the out-of-sample mean square forecast errors (MSFE) from the unrestricted models are smaller than those obtained using the restricted models.[13] If the unrestricted model for median inflation turns out to be superior over the restricted model in terms of forecast accuracy, price dispersion is then said to have predictive power for median inflation; this is considered to be evidence for Ganger causality running from price dispersion to median inflation. Granger causality from median inflation to price dispersion is tested similarly.

Five out-of-sample tests are used here: the Granger-Newbold (GN) test, the Diebold-Mariano (DM) test, Clark-West (CW) test, McCracken's (MSE-F) test, and also the Clark-McCracken (ENC-NEW) test. The first four of these are designed to test for equal mean squared forecast errors, while the last is a test for forecast encompassing. Each of these tests is briefly described below.

Granger and Newbold (1976) proposed a test based on the correlation between the sum of the restricted and unrestricted one-step-ahead forecast errors, $x_t = e_{r,t} + e_{u,t}$, and their difference, $z_t = e_{r,t} - e_{u,t}$, where $e_{r,t}$ and $e_{u,t}$ are the out-of-sample forecast errors from the restricted and unrestricted models, respectively. This test was first implemented in Ashley, et al. (1980) and

[13] As noted later in this section, we also test whether the out-of-sample forecasts obtained from the unrestricted model encompass those from the restricted one. Per comments in Rogoff and Stavrakeva (2008), however, we do not give results based on this test a causal interpretation.

Ashley (1981), and superseded by a direct bootstrap test described in Ashley (1998); these later versions relaxed the assumption that the errors were serially uncorrelated. In the direct bootstrap version of the test, a simple bivariate VAR (and the two corresponding univariate models) are estimated for $e_{r,t}$ and $e_{u,t}$, and the null hypothesis that E($e_{r,t}$)/E($e_{u,t}$) equals to one is directly tested by re-sampling from the fitting errors of these estimated models.[14]

Diebold and Mariano (1995) developed a test which also relaxes the assumption in the original version of the GN test that the forecast errors are serially uncorrelated. The Diebold-Mariano (DM) test statistic is defined as $DM = \bar{d}/\sqrt{LRVAR(\bar{d})}$, where $\bar{d} = \frac{1}{P}\sum[e_{r,t}^2 - e_{u,t}^2]$, P is the number of out-of-sample observations – here and below – and $LRVAR$(.) is the long-run variance function.

Recent work has shown, however, that these older forecast accuracy tests can suffer from serious size distortion problems when the models being compared are nested.[15] In particular, Clark and West (2006, 2007) point out that, under the null hypothesis of equal MSFE − i.e., where the restricted model is the actual data generating process − the unrestricted model is necessarily misspecified due to the inclusion of extraneous explanatory variables. While the population coefficients on these variables are zero, their sample estimates will be non-zero, leading to an upward bias in the sample MSFE for the unrestricted model. To correct for this upward bias, Clark and West (2006, 2007) proposed the modified Diebold-Mariano test statistic,

[14] This direct bootstrap version of this test explicitly allows for a substantial contemporaneous cross-correlation between the two forecast errors and also (through the VAR) for serial correlation, which might be present due to model misspecification. It does not allow for heteroskedasticity in the errors, however: nowadays a wild bootstrap would be used, as implemented below. It does, on the other hand, implement a double-bootstrap which roughly quantifies the uncertainty in the inference due to the bootstrap approximation itself, which is only justified for large samples. Thus, this direct bootstrap test is preferable for short out-of-sample periods, whereas the kind of bootstrapping implemented here − for the GN test and the four others − is preferable for longer out-of-sample periods. See also footnote #15 below.

[15] See West(1996), Clark and McCracken (2001, 2005), and West (2006), and Clark and West (2006).

$$CW = P^{1/2} \frac{\frac{1}{P}\sum[e_{r,t}^2 - e_{u,t}^2 + (\hat{f}_{r,t} - \hat{f}_{u,t})^2]}{LRVAR[e_{r,t}^2 - e_{u,t}^2 + (\hat{f}_{r,t} - \hat{f}_{u,t})^2]} \tag{3}$$

where $\hat{f}_{r,t}$ and $\hat{f}_{u,t}$ are the out-of-sample forecasts from the restricted and unrestricted models, respectively. They show that this test statistic is asymptotically a standard normal under the null hypothesis of equal MSFE for the two models, yielding a test with actual sizes close to but a little less than nominal size in finite samples.[16]

Another reason for size distortions in the GN and DM tests when the competing models are nested is that the asymptotic distributions of these forecast accuracy test statistics are significantly non-normal in that case. McCracken (2007)'s F-type test statistic,

$$MSE - F = P\sum(e_{r,t}^2 - e_{u,t}^2) \,/ \sum e_{u,t}^2. \tag{4}$$

is designed to correct for size distortions from this source. According to McCracken (2007), the asymptotic distribution of MSE-F is non-standard and depends on the forecasting scheme (fixed, rolling or recursive), the number of excess parameters in the nesting model, and also on the ratio of the number of out-of-sample observations to the number of in-sample observations. On the other hand, Clark and McCracken (2001) and McCracken (2007) have shown that this test is more powerful than the Diebold and Mariano test when the models are nested.

We also apply the Clark and McCracken (2001) test of forecast encompassing in the case of nested models:

$$ENC - NEW = P\sum(e_{u,t}^2 - e_{u,t}e_{b,t}) \,/ \sum e_{b,t}^2 \tag{5}$$

[16] This is for rolling forecasts; for forecasts calculated recursively, the limiting distribution of the CW statistic is a bit more complex; see Clark and West (2007) for details. Effectively, per Paye (2010), these tests are focusing on testing the underlying causal structure rather than simply testing whether the forecasts from one model are more accurate than those of a another, as in the direct bootstrap test of Ashley (1998) discussed in footnote #13 above.

where, as noted above, P is the number of out-of-sample observations. As pointed out by Rogoff and Stavrakeva (2008), forecast encompassing is arguably not quite the issue in testing for Granger-causality, because it focuses on testing whether the unrestricted model encompasses the restricted model rather than on whether the restricted model has smaller MSFE than the restricted one. Consequently, results with respect to this test are reported here, but are not emphasized in the discussion.

Bootstrap Implementation

Concerns regarding potential finite-sample size distortions in all of these tests based on the claimed asymptotic distributions of their test statistics prompt us to in each case use bootstrap replications to compute *p*-values for rejecting the null hypothesis of equal out-of-sample forecasting effectiveness for the restricted and unrestricted models. Simulated data for each of the three underlying time series (median inflation, price dispersion, and the change in unemployment rate) are generated by bootstrap re-sampling 3-vectors from the fitting errors of univariate autoregressive models for each of these variables.[17]

Because – as reported in Section 5.1 below − heteroskedasticity is an issue in this data set, the re-sampling was done using the 'wild' bootstrap proposed by Goncalves and Kilian (2004). Specifically, denoting the OLS fitting errors from the autoregressive models for median inflation, price dispersion, and the change in unemployment rate as τ_t, υ_t, and ω_t, respectively, we draw a sequence of *i.i.d.* innovations ε_t, t = 1, 2, … T, from the standard normal distribution and use $\varepsilon_t\tau_t$, $\varepsilon_t\upsilon_t$, and $\varepsilon_t\omega_t$ as the bootstrapped innovations to generate an artificial data set of 492

[17] The autoregression for the aggregate inflation equation includes a linear trend, its first, second, third and twelfth lags. The relative price dispersion equation is modeled as an AR(3) process and the seasonal difference of the change in the unemployment rate is modeled as an AR(4) process. These lag structures were chosen so as to minimize the BIC criterion.

observations.[18] The restricted and unrestricted models are then re-estimated and the six test statistics (F, GN, DM, CW, MSE-F and ENC-NEW) are calculated for the new data set. That completes one bootstrap replication. A total of 5,000 such replications are done, and the *p*-value reported in Table 5 for each of the tests is computed as the proportion of the generated test statistic values exceeding the test statistic value reported in Table 5 as having been observed using the actual sample data.

V. Empirical Results

In-Sample Model Estimation Results

We report the in-sample estimates of the restricted and unrestricted models for median inflation in Panel A of Table 2 and those for price dispersion in Panel B of Table 2. Lag lengths are chosen so as to minimize the Bayesian Information Criterion (BIC), leading to fairly simple models. Note that, because these are monthly time series which have not been subjected to seasonal adjustment, it is not odd to see terms included at lag twelve.

The Breusch-Pagan-Godfrey test and the White test both provide strong evidence that the errors in the model for median inflation are heteroskedastic, whereas the errors in the model for price dispersion appear to be homoskedastic. White-Eicker standard error estimates are therefore used throughout and the wild bootstrap is used in the re-sampling for the in-sample and out-of-sample forecasting test statistics, as described in Section IV.

In the unrestricted model for median inflation, the coefficient on the lagged price dispersion is highly significant, with a *t*-statistic of 2.9365, indicating strong in-sample predictive power of price dispersion for median inflation. In the unrestricted model for price dispersion, the

[18] For simplicity, we fix the values of initial observations at their actual sample values.

coefficient on the lagged median inflation is also found to be statistically significant at the 1% level, with a t-statistic of 3.7763, suggesting strong in-sample predictive power of median inflation for price dispersion.

With respect to the Sims test on causality from median inflation to price dispersion, we add leads of price dispersion to the unrestricted model for median inflation. The coefficients on the 12 leads of price dispersion are found to be jointly significant at the 0.01% level, with an F-statistic of 3.3147 and *p*-value of 0.0002. We consider this to be strong in-sample evidence for Granger causality running from median inflation to price dispersion. Similarly, we implement a Sims test with regard to Granger causality from price dispersion to median inflation by including 12 future values of median inflation in the unrestricted model for price dispersion and testing the null hypothesis that all twelve coefficients on the leads of median inflation are equal to zero; this null hypothesis is rejected at the 2% significance level. (The F-statistic = 2.1465 and the corresponding *p*-value is 0.0144.) Thus, the Sims causality tests confirm the result that there is feedback between median inflation and price dispersion.

Similarly, the Ashley and Verbrugge (2009) frequency dependent regression method finds that, while the relevant coefficient estimates in the two models − with the lagged putatively-causing variable replaced by its 19 frequency-specific components − are individually significant at some frequencies, the overall Granger causality evidence in these models is not as strong as in the models reported on above, which ignore possible frequency dependence. In particular, Table 3 shows that the null hypothesis that all of the coefficients on the various frequency components of lagged median inflation in the price dispersion equation are zero cannot be rejected at even the 10% level of significance; and a similar result is quoted for the null hypothesis that all of the coefficients on the various frequency components of lagged price dispersion in the equation for median inflation. This result is not unexpected, in view of the fact that it is not very easy to reject a null hypothesis embodying 19 parameter restrictions, even when a number of these restrictions are false.

On the other hand, restricting attention to just the high-frequency components of lagged median inflation, Table 3 shows that it is, in fact, possible to reject the null hypothesis that all of the coefficients on these high-frequency components of lagged median inflation are zero in the price dispersion equation – with a *p*-value of either 0.03 or 0.04, depending on whether "high-frequency" is defined as "fluctuations with period less than or equal to 6 months" or "fluctuations with period less than or equal to 3 months." This result is illustrated in Figure 3, which plots the 19 coefficient estimates (along with the ±1 standard deviation bands) against the period-length for the fluctuations in lagged median inflation included in the corresponding component. These frequency-dependence results are hardly compelling, but they do suggest that – in the sample period, at least – the observed causality from median inflation to price dispersion is primarily a high-frequency relationship. This is a useful insight if its credibility can be enhanced by finding unidirectional causality from median inflation to price dispersion in out-of-sample testing.

Forecasting Results

Using these model specifications, we then obtain recursive one-step-ahead out-of-sample forecast errors from the restricted and unrestricted models for median inflation and price dispersion, respectively. "Recursive," in this context, means that the model parameters are up-dated (re-estimated) using the additional data as the forecasting process moves through the out-of-sample period. The mean squared errors − both in-sample and out-of-sample – are tabulated and compared in Table 4 for the restricted and unrestricted models of median inflation (Panel A) and price dispersion (Panel B).

We find that, over the in-sample period, including price dispersion in the median inflation equation reduces the MSE by about 3% while including median inflation in the price dispersion equation reduces the MSE by over 4%. These results are, of course, consistent with the t-test results reported in Section V above.

The out-of-sample mean squared forecast errors (MSFE) from the unrestricted model for price dispersion are similarly around 3% smaller than those from the restricted model for price dispersion. In contrast, the unrestricted model for median inflation in fact provides *less* accurate forecasts than does the restricted model: its MSFE is actually 2% *larger*. Thus, the strength of the out-of-sample evidence for Granger causality from median inflation to price dispersion hinges on whether this 3% drop in MSFE is statistically significant. And the existence (much less the strength) of the out-of-sample evidence for Granger causality from price dispersion to median inflation rests on whether corrections for nesting are sufficient to credibly reverse the impact of this rise in the realized out-of-sample MSFE for the unrestricted model of median inflation.

Table 4 also reports estimates of the ratio, for each pair of models, of the MSE (or MSFE, using the out-of-sample period) to the sample variance of the time series (median inflation or price dispersion) being modeled. It is shown in Ashley (1983) that a model for a variable z_t that yields forecasts $\hat{z}_t$ which are so poor that MSE($\hat{z}_t$)/Var(z_t) exceeds one renders z_t useless as an input variable in models for forecasting other variables – regardless of the size and significance with which z_t enters those other models − if $\hat{z}_t$ will, in the end, be used to replace z_t. Reference to Table 4 shows that simple point estimates of these ratios are less than one except for the out-of-sample forecasts of aggregate inflation based on the unrestricted model, in which case the estimated ratio of the mean squared forecast errors to the variance of the median inflation variable itself is 1.0047. This result is another reflection of the dismal out-of-sample forecasting performance of the unrestricted model for median inflation.

In summary, there is strong in-sample evidence for Granger causality between these two time series in both directions. The out-of-sample evidence for Granger causality running from price dispersion to median inflation is very weak, to say the least. There is, however, out-of-sample evidence for Granger causality running from median inflation to price dispersion. The question is whether the out-of-sample forecasting improvement from including the median inflation time series in the model for price dispersion is statistically significant: that is addressed in the next section, using the tests described in Section IV.

Results from the Forecasting-Based Granger Causality Tests

Table 5 reports the results from both in-sample and out-of-sample tests of Granger causality between median inflation and price dispersion based on the relative forecasting effectiveness of the restricted versus unrestricted models for each time series. The reported p-values are for rejecting the null hypothesis of equal forecasting effectiveness, which corresponds to an absence of Granger causality running from the additional variable (or variables) included in the unrestricted model to the dependent variable in common to both models. In all cases these p-values are computed using the bootstrapped sampling distributions of the indicated test statistics, as described in Section IV.

In-sample F-test statistics and p-values are reported in the first row of Table 5. Since the p-values obtained are in both cases less than 0.005, both the null hypothesis of no Granger causality running from median inflation to price dispersion and the null hypothesis of no Granger causality running from price dispersion to median inflation can be rejected at the 0.5% significance level. If one is inclined to accept in-sample evidence as credible − which the authors are not − this would be strong evidence for bi-directional Granger causality (feedback) between these two time series.

The next five rows of Table 5 report null hypothesis rejection *p*-values and test statistic values for each of the out-of-sample tests described in Section IV. The left-most column is in each case reporting results for testing the null hypothesis that the out-of-sample forecast errors generated by the unrestricted model for the inflation series are no smaller than those generated by the restricted model, which excludes the price dispersion time series as an explanatory variable. Thus, a small *p*-value – allowing rejection of this null hypothesis − is evidence in favor of Granger causality running from price dispersion to median inflation.

Similarly, the right-most column is in each case reporting results for testing the null hypothesis that the out-of-sample forecast errors generated by the unrestricted model for the price dispersion series are no smaller than those generated by the restricted model, which excludes the median inflation time series as an explanatory variable. Thus, a small p-value – allowing rejection of this null hypothesis − is evidence in favor of Granger causality running from median inflation to price dispersion.

With regard to the null hypothesis ruling out Granger causality from price dispersion to median inflation, the results in the left-most column of Table 5 show that none of the out-of-sample test statistics is statistically significant at even the 10% level, except the Clark-McCracken (ENC-NEW) encompassing test statistic, which is significant at the 2% level. (This latter test is actually addressing forecast encompassing rather than forecast accuracy; consequently – per Rogoff and Stavrakeva (2008) − this result is of doubtful relevance to the Granger causality between these two time series.) In brief, there is no substantive out-of-sample evidence for Granger causality from price dispersion to median inflation.

The results are quite different with respect to Granger causality running from median inflation to price dispersion, however. For this set of out-of-sample tests, the results given in the right-most column of Table 5 indicate that the null hypothesis (of no causality) can be rejected at the 0.5% level using the Clark-West (CW) and McCracken (MSE-F) test statistics, at the 5% level using the Diebold-Mariano (DM) test statistic, and at the 10% level using the Granger-Newbold (GN) test statistic. We consider this to be strong out-of-sample evidence for Granger causality running from median inflation to price dispersion.

Robustness Checks

In this subsection, we evaluate the robustness of our Granger causality findings in several ways. First, additional control variables are allowed to enter both the restricted and unrestricted models, so as to diminish the chance that a fourth variable is driving the results. (A third variable, the change in the unemployment rate is already included in both models.) Second, so as to eliminate the possibility that our results might be driven by energy shocks impacting both time series, the analysis is repeated excluding three energy-related components from the construction of the median inflation and price dispersion series. Third, we allow for a possible structural break in each model, so as to rule out the possibility that our results are an artifact of this kind of model instability. Finally, we examine the robustness of our results over time in two ways: by partitioning the out-of-sample period into six 30-month sub-samples and repeating the out-of-sample testing for each, and by repeating the out-of-sample testing over a rolling forecast period. The results from these robustness checks are described below.

First, to ensure that our results of feedback between median inflation and price dispersion are not driven by an omitted variable, we allow for the introduction of three additional explanatory variables into both the restricted and unrestricted forecasting models for median inflation and price dispersion. Panel A of Table 6 reports the test results allowing for the inclusion of the growth rate in the broad money supply (M2); Panel B reports the analogous results, allowing for the inclusion of the growth rate in the Index of Industrial Production (IP); Panel C reports the results when the change in 3-month treasury bill rate (TB3) are included.[19] The inclusion of these additional control variables does not alter our main results. All in-sample test statistics are significant at the 0.3% level, indicating strong in-sample evidence for bi-directional Granger causality between median inflation and price dispersion. Regarding the out-of-sample test statistics associated with the null hypothesis of no Granger causality from price dispersion to median inflation, only the ENC-NEW test statistic is statistically significant when the M2 growth rate is included; when the growth rate in IP or the change in 3-month treasury bill rate is included, only the CW and ENC-NEW test statistics are significant. Therefore, there is still only very limited evidence for Granger causality from price dispersion to median inflation.

[19] The data on $M2_t$, on IP_t and on $TB3_t$ are both seasonally unadjusted and obtained from Board of Governors of the Federal Reserve System H.6, G.17 and H.15, respectively. The appropriate lag lengths with which these variables enter the models is determined by minimizing the BIC. The first lag of M2 growth is included in the forecasting models for median inflation while its 4th and 5th lags are included in the forecasting models for price dispersion. The 1st and 12th lags of IP growth are included in the forecasting models for median inflation, and its 6th lag is included in the forecasting models for price dispersion. The first lag of TB3 is included in the forecasting models for median inflation and its second lag is included in the forecasting models for price dispersion. These estimation results are available upon request.

Relative to the tests for Granger causality from median inflation to price dispersion, however, we note that all of the out-of-sample test statistics are still highly significant, except for the test based on the GN test statistic, which remains statistically significant at least at the 10% level. In short, the Granger causality results reported in Section V are quite robust to the inclusion of any one of these additional covariates.

Second, worried about the possibility that our Granger causality findings might be driven by energy shocks impinging on the two series at different relative lags, the analysis is repeated with the three (of thirty one, total) components which seem clearly energy-related omitted from the computation of the median and interquartile range of the of the monthly component-level growth rates.[20] The test results from this exercise are given in Panel A of Table 7. Again, there is still strong in-sample evidence for the feedback between the two series with the energy-related CPI components omitted. With respect to the out-of-sample Granger causality tests, while the evidence for Granger causality running from median inflation to price dispersion is somewhat diminished, it remains quite strong; and there is now also some distinct evidence for price dispersion Granger-causing median inflation, based on the Clark-West (CW) and McCracken (MSE-F) test statistics.

Third, we check to see if allowing for structural change in the models alters our results. Using Bai and Perron's (1998, 2003) and Andrews' (1993) procedures, we are able to identify a single structural break date in the coefficients of the unrestricted models for each variable.[21] Allowing for these structural shifts by including dummy variables on the coefficients yields the

[20] See, for example, Fischer (1981) and Taylor (1981), for discussions of the likely effect of energy price shocks on the relationship between inflation and price dispersion. The three CPI components excluded from the calculation of median inflation and price dispersion are "fuel oil & other fuels", "gas & electricity", and "motor fuel".

[21] The break dates identified by Andrews' supWald tests for the unrestricted aggregate inflation and relative price dispersion regressions are 1981M05 and 1973M01, respectively; more details on these results are available upon request.

test results reported in Panel B of Table 7, which are not appreciably different from the corresponding test results in Table 5.

Fourth, we examine the robustness of our out-of-sample test results to evaluating the forecasting effectiveness of the models over subsets of the full out-of-sample period. In particular, we divide the out-of-sample period into six 30-month sub-samples and conduct out-of-sample Granger causality tests for each. Panel A of Table 8 shows the testing results for Granger causality from price dispersion to median inflation. In general, we find very limited evidence for Granger causality from price dispersion to median inflation. In all six sub-samples, the GN and DM test statistics are not significant at the conventional levels at all; the CW and MSE-F statistics are significant at the 5% only in the first subsample period. Panel B reports the out-of-sample test results on Granger causality from median inflation to price dispersion. As compared to our previous results on Granger causality from median inflation to price dispersion, the results are now statistically weaker, but this is what one ought to expect using out-of-sample periods which are much shorter. While the DM and CW test statistics are not significant in any of the six subsamples, the GN test statistics are still significant at the 5% level in three subsamples, and the MSE-F test is still significant at the 1% level in two subsamples and at the 5% level in another two. Notably, however, none of the tests in Panel B rejects the null hypothesis (of no causality) in the final subsample, from July 2006 to December 2008.

Finally, we examined the manner in which the *p*-values for the GN, DM, CW, and MSE-F out-of-sample tests change as the starting month for the out-of-sample period is varied, beginning in January 1994. Figure 4 graphs the bootstrapped *p*-values for each of these tests as a function of this starting month; the horizontal line represents a p-value of 0.10.[22] These plots confirm the stability over time of our Granger causality results with respect to price dispersion (not) Granger causing median inflation. It also confirms the stability of our Granger causality results with respect to median inflation Granger causing price dispersion up until the middle of 2004, but indicates − consistent with the results in Panel B of Table 8 − that the relationship appears to diminish after that.

All in all, these checks indicate that our results are very robust in general, but that the Granger causation of median inflation on price dispersion probably dwindled after mid-2004.

VI. Conclusions

In this study we apply both in-sample and out-of-sample Granger causality tests to examine the causal relationship between median inflation and price dispersion. Given the fact that the cross-sectional distribution of the weighted CPI component price growth rates is skewed and fat-tailed, we use the median and the logarithm of the inter-quartile range of these component growth rates as measures of inflation and price dispersion, respectively.

Using a monthly dataset over the period January 1968 to December 2008, we find fairly strong in-sample evidence for bi-directional Granger causality – i.e., feedback – between median inflation and price dispersion. Absent confirmation of this feedback result via out-of-sample testing, however, we are somewhat skeptical of conclusions based solely on in-sample evidence, however, because it can easily be an artifact of the specification searches used in obtaining the

[22] The p-value plot for the ENC-NEW encompassing test is omitted because its interpretation in terms of Granger causality is murky. Qualitatively, for the test of price dispersion Granger-causing inflation, the ENC-NEW test p-value hovers in the range 0.05 to 0.10 until starting months in mid-2000 and then became even larger. For the ENC-NEW test of inflation Granger causing price dispersion, the p-value plot remains well below 0.05 until the middle of 2004 and then increases.

models. Results from a variety of out-of-sample Granger causality tests show that there is scant evidence for Granger causality from price dispersion to median inflation. In contrast, the out-of-sample tests strongly confirm the in-sample evidence for Granger causality running from median inflation to price dispersion. This evidence for uni-directional Granger causation from median inflation to price dispersion is very strong, at least until mid-2004. These out-of-sample Granger causality results are qualitatively stable across a varied set of robustness checks.

We also find modest in-sample evidence that the Granger causality running from median inflation to price dispersion is primarily a high-frequency phenomenon. In other words, it appears to be inflation fluctuations with periods of less than six months in duration which are driving the fluctuations in price dispersion.

Our results are thus not supportive of the theoretical models – Tobin (1972), Gordon (1975), Ball and Mankiw (1995), and Lourenco and Gruen (1995) – which predict Granger causation from price dispersion to inflation. In contrast, our results do support the array of menu cost, signal extraction, and monetary search models (surveyed in Section 2) which predict Granger causation from inflation to price dispersion, although further work is needed in order to distinguish between them.

Appendix

The 31 components of the consumer price index (CPI) include: cereal and bakery products, meats, poultry, fish and eggs, diary and related products, fruits and vegetables, other food at home, food away from home, alcoholic beverage, fuel oil and other fuels, gas and electricity, water, sewer and trash collection, household furnishings and operations, men's and boys' apparel, women's and girls' apparel, infants' and toddler' apparel, footwear, jewelry and watches, new vehicles, used cars and trucks, motor fuel, motor vehicle maintenance and repair, motor vehicle parts and equipment, motor vehicle insurance, public transportation, medical care commodities, medical care services, recreation, communication, tobacco and smoking products, personal care products, personal care services, education.

References

Ashley, Richard. 1981. “Inflation and the Distribution of Price Changes across Markets: A Causal Analysis.” *Economic Inquiry*, 19(4): 650-660.

Ashley, Richard. 1983. “On The Usefulness of Macroeconomic Forecasts as Inputs to Forecasting Models.” *Journal of Forecasting*, 2(3):211-223.

Ashley, Richard. 1998. “A New Technique for Postsample Model Selection and Validation.” *Journal of Economic Dynamics and Control*, 22(5): 647-665.

Ashley, Richard. 2003. “Statistically Significant Forecasting Improvements: How Much Out-of-Sample Data Is Likely Necessary?” *International Journal of Forecasting*, 19(2): 229-239.

Ashley, Richard, and Randall J. Verbrugge. 2007. “Mis-Specification in Phillips Curve Regressions: Quantifying Frequency Dependence in this Relationship While Allowing for Feedback” mimeo, http://ashleymac.econ.vt.edu/working_papers/ashley_verbrugge.pdf.

Ashley, Richard, and Randall J. Verbrugge. 2009 . “Frequency Dependence in Regression Model Coefficients: An Alternative Approach for Modeling Nonlinear Dynamic Relationships.” *Econometric Reviews*, 28: 4-20.

Ashley, Richard, Clive W.J. Granger, and Richard Schmalensee. 1980. “Advertising and Aggregate Consumption: An Analysis of Causality.” *Econometrica*, 48(5), 1149-1168.

Bai, Jushan, and Pierre Perron. 1998. “Estimating and Testing Linear Models with Multiple Structural Changes. *Econometrica*, 66 (1): 47-78.

Bai, Jushan, and Pierre Perron. 2003. “Computation and Analysis of Multiple Structural Change Models.” *Journal of Applied Econometrics*, 18(1): 1–22.

Ball, Laurence, and N. Gregory Mankiw. 1995. :Relative-Price Changes as Aggregate Supply Shocks.” *Quarterly Journal of Economics*, 110(1): 161-193.

Ball, Laurance, and David Romer. 2003. “Inflation and the Informativeness of Prices.” *Journal of Money, Credit and Banking*, 35(2): 177-196.

Barro, Robert. 1976. “Rational Expectations and the Role of Monetary Policy.” *Journal of Monetary Economics*, 2(1): 1-32.

Benabou, Roland. 1988. “Search, Price Setting and Inflation.” *Review of Economic Studies,* 55(3): 353-376.

Breitung, Jörg, and Bertrand Candelon. 2006. “Testing for short-and long-run causality: A frequency-domain approach.” *Journal of Econometrics*, 132: 363-378.

Bryan, Michael, and Stephen G. Cecchetti. 1994. "Measuring Core Inflation." In *Monetary Policy*, ed. Gregory Mankiw, 195-215. Chicago: University of Chicago Press.

Bryan, Michael, Stephen G. Cecchetti, and Rodney L. Wiggins II. 1997. "Efficient Inflation Estimators." National Bureau of Economic Research Working Paper No.6183.

Clark, Todd, and Michael McCracken. 2001. "Test of Equal Forecast Accuracy and Encompassing for Nested Models." *Journal of Econometrics,* 105(1): 85-110.

Clark, Todd, and Michael McCracken. 2005. "Evaluating Direct Multi-Step Forecasts." *Econometric Reviews,* 24(4): 369-404.

Clark, Todd, and Kenneth West. 2006. "Using Out-of-Sample Mean Squared Prediction Errors to Test the Martingale Difference Hypothesis." *Journal of Econometrics,* 135(1-2): 155-186.

Clark, Todd, and Kenneth West. 2007. "Approximately Normal Tests for Equal Predictive Accuracy in Nested Models." *Journal of Econometrics,* 138(1): 291-311.

De Abreu Lourenco, Richard, and David Gruen. 1995. "Price Stickiness and Inflation." Reserve Bank of Australia Research Discussion Paper No. 9502.

Debelle, Guy, and Owen Lamont. 1997. "Relative Price Variability and Inflation: Evidence from U.S. Cities." *Journal of Political Economy,* 105(1): 132-152.

Diamond, Peter. 1993. "Search, Sticky Prices and Inflation." *Review of Economic Studies,* 60(1): 53-69.

Fischer, Stanley. 1981. "Towards an Understanding of the Costs of Inflation: II." *Carnegie-Rochester Conference Series on Public Policy,* 15(1): 5-41.

Fischer, Stanley. 1982. "Relative Price Variability and Inflation in the United States and Germany." *European Economic Review, 18*(1): 171-196.

Goncalves, Silvia, and Lutz Kilian . 2004. "Bootstrapping Autoregressions with Conditional Heteroskedasticity of Unknown Form." *Journal of Econometrics,* 123(1): 89-120.

Gordon, Robert I.. 1977. "World Inflation and Monetary Accommodation in Eight Countries." *Brookings Papers on Economic Activity,* 8(2): 409-478.

Granger, Clive, and Paul Newbold. 1976. "Forecasting Transformed Series." *Journal of the Royal Statistical Society Series B,* 38(2): 189-203.

Grier, Kevin, and Mark Perry. 1996. "Inflation, Inflation Uncertainty, and Relative Price Dispersion: Evidence from Bivariate GARCH-M Models." *Journal of Monetary Economics,* 38(2): 391-405.

Head, Allen, and Alok Kumar. 2005. “Price Dispersion, Inflation, and Welfare.” *International Economic Review*, 46(2): 533-572.

Hercowitz, Zvi. 1981. “Money and Dispersion of Relative Prices.” *Journal of Political Economy,* 89(2): 328-356.

Lemmens, Aurélie, Christophe Croux and Marnik G. Dekimpe. 2008. “Measuring and testing Granger causality over the spectrum: An application to European production expectation surveys.” *International Journal of Forecasting*, 24: 414-431.

Lucas, Robert. 1973. “Some International Evidence on Output-Inflation Tradeoffs.” *American Economic Review,* 63(3): 326-334.

McCracken, Michael W. 2007. “Asymptotics for Out of Sample Tests of Granger Causality.” *Journal of Econometrics,* 140(2), 719-752.

Parks, Richard. 1978. “Inflation and Relative Price Variability.” *Journal of Political Economy,* 86(1): 79-95.

Parsley, David. 1996. “Inflation and Relative Price Variability in the Short and Long Run: New Evidence from the United States.” *Journal of Money, Credit and Banking,* 28(3), 323-341.

Paye, Bradley. 2010. “Do Macroeconomic Variables Forecast Aggregate Stock Market Volatility.” (Unpublished working paper.)

Peterson, Brian, and Shi, Shouyong. 2004. “Money, Price Dispersion and Welfare.” *Economic Theory,* 24(4): 907-932.

Reinsdorf, Marshall. 1994. “New Evidence on the Relation between Inflation and Price Dispersion.” *American Economic Review,* 84(3): 720-731.

Rogoff, Kenneth, and Vania Stavrakeva. 2008. “The Continuing Puzzle of Short Horizon Exchange Rate Forecasting.” National Bureau of Economic Research Working Paper No. 14071.

Rotemberg, Julio. 1983. “Aggregate Consequences of Fixed Costs of Price Adjustment.” *American Economic Review,* 73(3): 433-436.

Sheshinski, Eytan, and Yoram Weiss. 1977. “Inflation and Costs of Price Adjustment.” *Review of Economic Studies, 44*(2): 287-303.

Sheshinski, Eytan, and Yoram Weiss. 1983. “Optimum Pricing Policy under Stochastic Inflation.” *Review of Economic Studies,* 50(3): 513-529.

Sims, Christopher A. 1972. “Money, Income, and Causality.” *American Economic Review,* 62(4): 540-552.

Subramanian, S. V. and Ichiro Kawachi. 2004. “Income Inequality and Health: What Have We Learned So Far?” *Epidemiologic Reviews* 26: 78-91.

Taylor, John. 1981. “On the Relation between the Variability of Inflation and the Average Inflation Rate.” *Carnegie-Rochester Conference Series on Public Policy*, 15(1): 57-85.

Tobin, James. 1972. “Inflation and Unemployment.” *American Economic Review,* 62(1): 1-18.

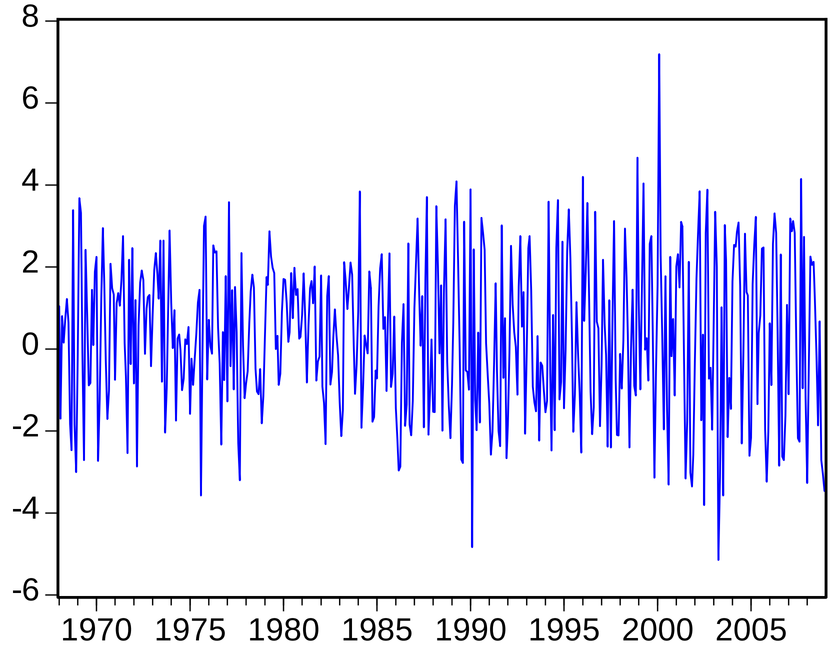
Vining, Daneil R. Jr, and Thomas C. Elwertowski. 1976. “The Relationship between Relative Prices and the General Price Level.” *American Economic Review,* 66(4): 699-708.

West, Kenneth. 1996. “Asymptotic Inference about Predictive Ability.” *Econometrica,* 64(5): 1067-1084.

West, Kenneth. 2006. “Forecast Evaluation.” In *Handbook of Economic Forecasting Vol.1*, ed. Graham Elliott, Clive William John Granger, and Allan Timmermann,100-134. Amsterdam: Elsevier.

Figure 1. Weighted Skewness and Kurtosis of Cross-Sectional Price-Change Distribution: January 1968 to December 2008

Panel A. Weighted Skewness



Panel B. Weighted Kurtosis

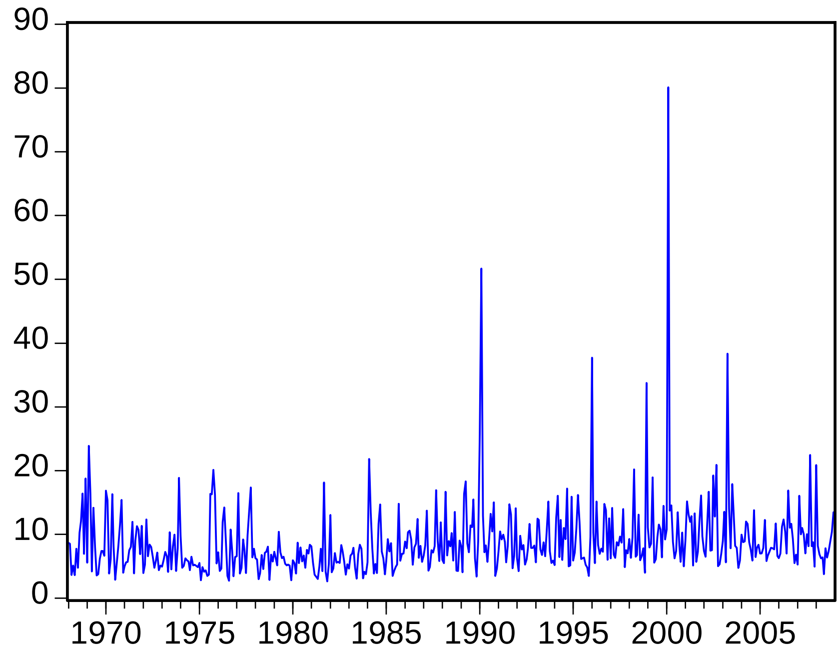
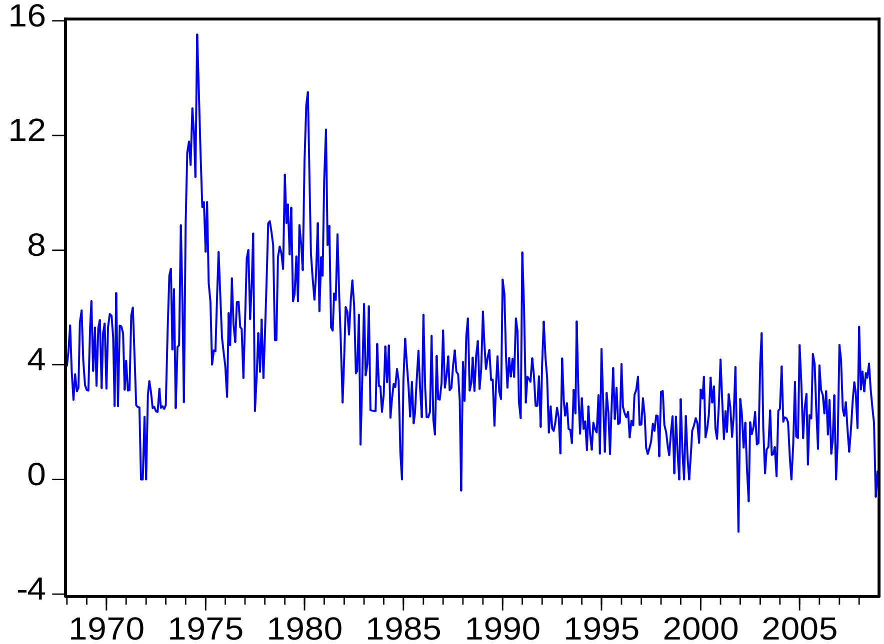


Figure 2. Monthly Median Inflation and Price Dispersion: January 1968 to December 2008

Panel A. Median Inflation



Panel B. Price Dispersion

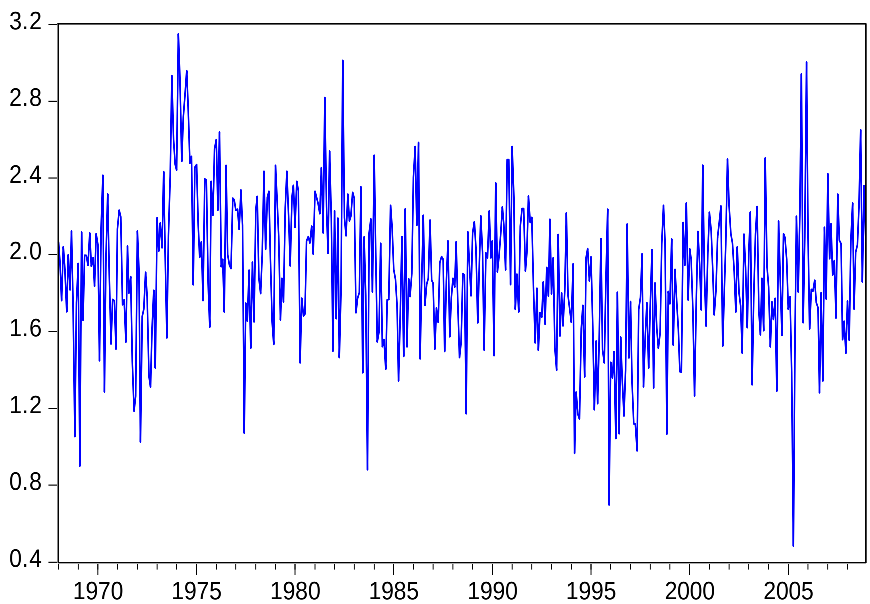
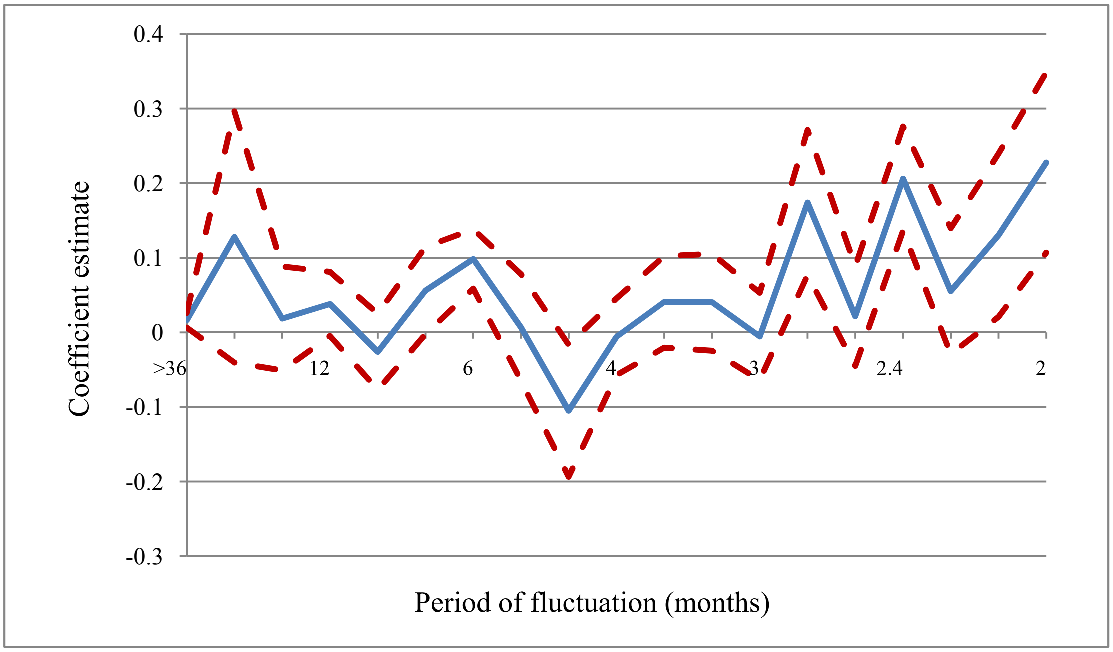


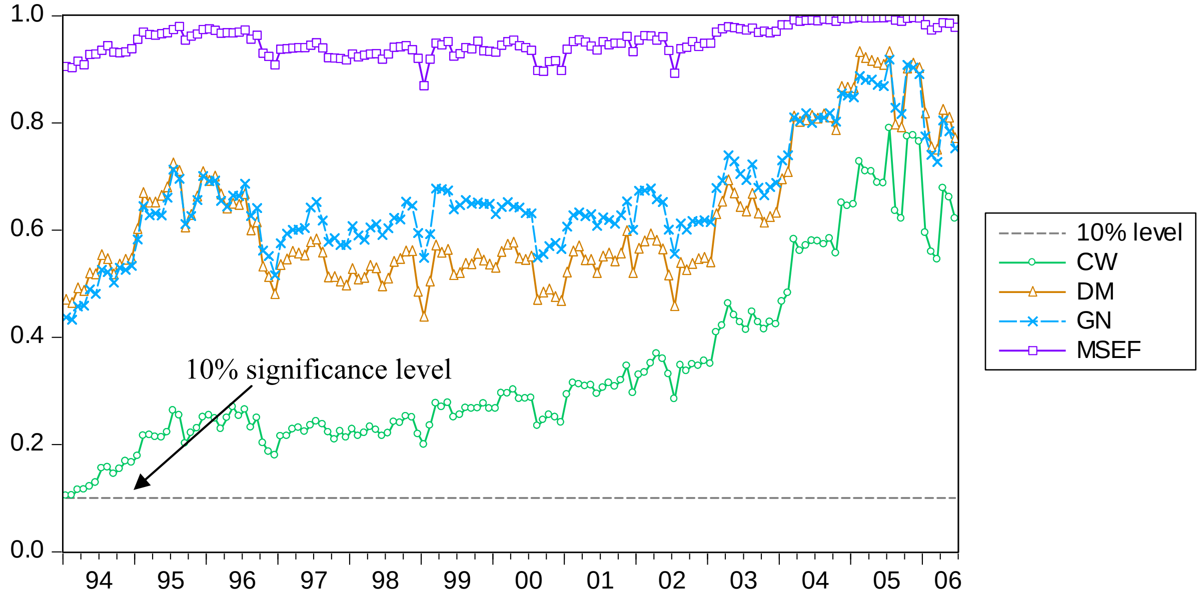
Figure 3. Parameter Estimates from the Frequency Dependence Regression of Price Dispersion



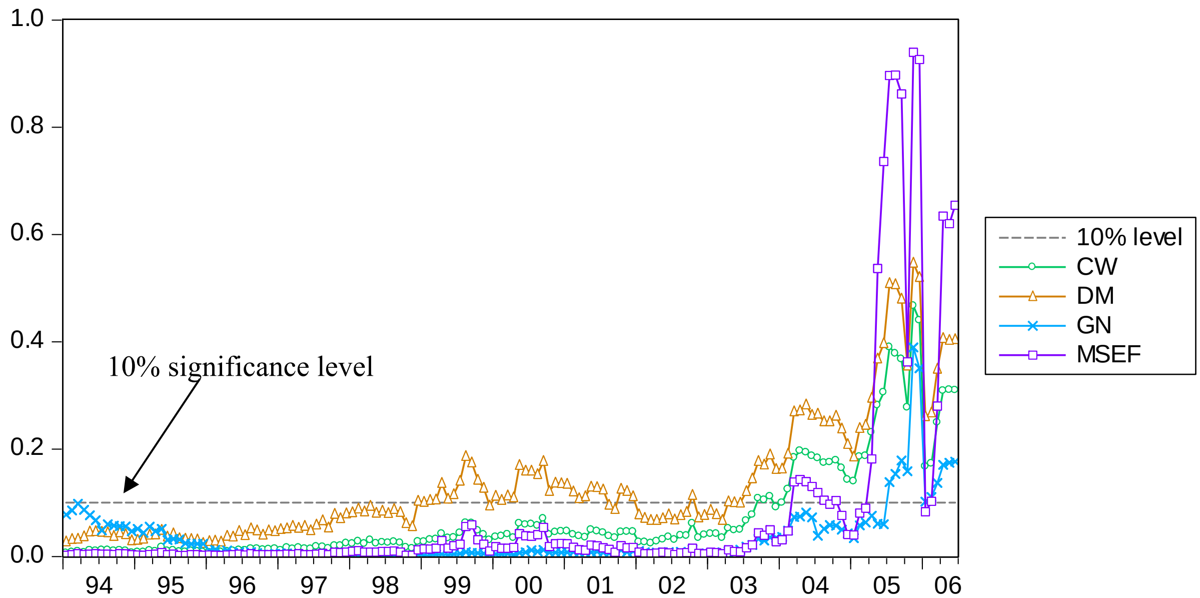
Notes: Solid line indicates the point estimates of parameters. Dashed lines indicate the ±1 standard deviation bands.

Figure 4. P-Values of Out-of-Sample Granger Causality Tests for Different Forecast Windows

Panel A. Testing for Granger Causality from Price Dispersion to Median Inflation



Panel B. Testing for Granger Causality from Median Inflation to Price Dispersion



Note: The green line with bubbles indicates the p-values of the CW statistics. The orange line with triangles indicates the p-values of the DM statistics. The blue line with cross indicates the p-values of the GN statistics. The purple line with squares indicates the p-values of the MSE-F statistics. All p-values are obtained using the bootstrap procedure described in Section 4.5.

Table 1. Testing for Unit Roots in Median Inflation and Price Dispersion

	Median Inflation	Price Dispersion
Augmented Dickey-Fuller (ADF) Test	-4.2245***	-6.0275***
Phillips-Perron Test	-6.9855***	--13.5826***
Dickey-Fuller (DF) GLS Test	-4.1967***	-4.5650***

Notes: Unit root tests are applied to the sample period 1969M1~1993M12. Both intercept and trend are included in the unit root test for median inflation. Only intercept is included in the unit root test for price dispersion. Bayesian information criterion (BIC) is used to select lag length in the ADF and DF GLS tests. 2 lags are used in the ADF and DF-GLS tests for both series, and 1 lag is used in the Phillips-Perron test for both series. The superscript *** indicates the significance level of 1%.

Table 2. In-Sample Model Estimates: January 1969 to December 1993

Panel A.	Dependent Variable: Median Inflation (y_t)		Panel B.	Dependent Variable: Price Dispersion (x_t)	
	Restricted	*Unrestricted*		*Restricted*	*Unrestricted*
constant	0.8527**	-0.4510	*constant*	0.8246***	0.9215***
	(0.3745)	(0.5497)		(0.1400)	(0.1422)
trend	-0.0017	-0.0021*	x_{t-1}	0.2635***	0.2272***
	(0.0011)	(0.0011)		(0.0541)	(0.0538)
y_{t-1}	0.5226***	0.4981***	x_{t-2}	0.1067**	0.0732
	(0.0648)	(0.0633)		(0.0520)	(0.0534)
y_{t-2}	0.0498	0.0389	x_{t-3}	0.2144***	0.1656***
	(0.0661)	(0.0640)		(0.0596)	(0.0601)
y_{t-3}	0.2034***	0.1748***	Δur_{t-12}	0.1254***	0.1041***
	(0.0694)	(0.0661)		(0.0404)	(0.0390)
y_{t-12}	0.1057**	0.0990**	y_{t-4}		0.0282***
	(0.0465)	(0.0464)			(0.0075)
Δur_{t-6}	-0.6961***	-0.6109***			
	(0.1821)	(0.1817)			
x_{t-3}		0.8611***			
		(0.2932)			
No. of Obs.	300	300	*No. of Obs.*	300	300
Adj. R^2	0.6273	0.6371	*Adj.* R^2	0.2120	0.2430
BIC	3.9180	3.9070	*BIC*	0.6651	0.6406
BPG(χ^2)	19.2830***	22.7909***	*BPG*(χ^2)	3.7227	3.5486
White (χ^2)	42.7917**	51.0438**	*White* (χ^2)	8.4276	10.5101
Durbin-Watson	1.9547	1.9592	*Durbin-Watson*	2.0866	2.0276

Notes: Δur_t denotes the change in the civilian unemployment rate from period *t*-1 to *t*. Robust standard errors are reported in parentheses. BPG denotes the Breusch-Pagan-Godfrey test statistic and White denotes the White test statistic. The superscripts ***, ** and * indicate significance at the 1%, 5%, and 10% levels, respectively.

Table 3. Results from Frequency Dependence Regression

Null Hypothesis	Dependent Variable: Price Dispersion (x_t)	Dependent Variable: Median Inflation (y_t)
H_0: Coefficients on the 19 frequency-specific components are all zero.	1.3445 (0.2539)	1.1493 (0.3030)
H_0: Coefficients on the fluctuations with periods greater than or equal to 12 months are all zero.	1.3450 (0.2539)	1.4671 (0.2128)
H_0: Coefficients on the fluctuations with periods less than or equal to 6 months are all zero.	1.9357** (0.0270)	0.8475 (0.6095)
H_0: Coefficients on the fluctuations with periods less than or equal to 3 months are all zero.	2.1462** (0.0397)	0.2529 (0.9709)

Notes: F-statistics are reported and their p-values are reported in parentheses. The superscripts ***, ** and * indicate that the null hypothesis can be rejected at the significance levels of 1%, 5% and 10%, respectively.

Table 4. In-Sample Fit and Out-of-Sample Forecasting Results

Panel A: Median Inflation (y)	Restricted Model (R)	Unrestricted Model (U)
In-Sample MSE	2.5781	2.5019
In-Sample Variance(y)	7.0833	7.0833
In-Sample MSE/Variance(y)	0.3640	0.3532
In-Sample MSE-U/MSE-R	0.9704	
Out-of-Sample MSFE	1.3361	1.3649
Out-of-Sample Variance(y)	1.3585	1.3585
Out-of-Sample MSFE/Variance(y)	0.9835	1.0047
Out-of-Sample MSFE-U/MSFE-R	1.0216	

Panel B: Price Dispersion (x)	Restricted Model (R)	Unrestricted Model (U)
In-Sample MSE	0.1035	0.0991
In-Sample Variance(x)	0.1336	0.1336
In-Sample MSE/Variance(x)	0.7747	0.7418
In-Sample MSE-U/MSE-R	0.9575	
Out-of-Sample MSFE	0.1319	0.1285
Out-of-Sample Variance(x)	0.1453	0.1453
Out-of-Sample MSFE/Variance(x)	0.9078	0.8844
Out-of-Sample MSFE-U/MSFE-R	0.9742	

Notes: The in-sample period is January 1968 to December 1993, which has a total of 300 observations; the out-of-sample period is January 1994 to December 2008, which has a total of 180 observations. MSE-R denotes the mean squared errors from the restricted model. MSE-U denotes the mean squared errors from the unrestricted model.

Table 5. Testing Granger Causality between Median Inflation and Price Dispersion

	H_0: No Granger causality from price dispersion to median inflation.	H_0: No Granger causality from median inflation to price dispersion.
In-Sample *F*-Test	8.6232*** (0.0040)	14.2606*** (0.0002)
Granger-Newbold (GN) Test	-0.6096 (0.4339)	0.5145* (0.0780)
Diebold-Mariano (DM) Test	-0.7105 (0.4631)	1.2864** (0.0260)
Clark-West (CW) Test	1.0748 (0.1098)	2.3591*** (0.0040)
McCracken (MSE-F) Test	-3.8032 (0.9010)	4.8704*** (0.0010)
Clark-McCracken (ENC-NEW) Test	3.1990** (0.0132)	5.2582*** (0.0004)

Notes: Sample test statistics are reported and their p-values (obtained from the wild bootstrap procedure described in Section 4.5) are reported in parentheses. The superscripts ***, ** and * indicate that the null hypothesis can be rejected at the significance levels of 1%, 5% and 10%, respectively.

Table 6. Including Additional Control Variables

Panel A. Adding the Growth Rate of Money Supply

	H_0: No Granger causality from price dispersion to median inflation.	H_0: No Granger causality from median inflation to price dispersion.
In-Sample *F*-Test	9.8832***	12.8911***
	(0.0026)	(0.0004)
Granger-Newbold (GN) Test	-0.7276	0.9035**
	(0.4939)	(0.0370)
Diebold-Mariano (DM) Test	-0.8683	1.6373***
	(0.5299)	(0.0086)
Clark-West (CW) Test	0.9660	2.5906***
	(0.1298)	(0.0036)
McCracken (MSE-F) Test	-4.8124	5.8670***
	(0.9382)	(0.0006)
Clark-McCracken (ENC-NEW) Test	2.9751**	5.5658***
	(0.0206)	(0.0010)

Panel B. Adding Industrial Production Growth

	H_0: No Granger causality from price dispersion to median inflation.	H_0: No Granger causality from median inflation to price dispersion.
In-Sample *F*-Test	12.0872***	8.6703***
	(0.0006)	(0.0032)
Granger-Newbold (GN) Test	-0.4951	0.8822**
	(0.3981)	(0.0360)
Diebold-Mariano (DM) Test	-0.5691	1.6923***
	(0.4177)	(0.0096)
Clark-West (CW) Test	1.5179**	2.5607***
	(0.0480)	(0.0038)
McCracken (MSE-F) Test	-3.6125	5.6536***
	(0.9024)	(0.0010)
Clark-McCracken (ENC-NEW) Test	5.2400***	4.9383***
	(0.0020)	(0.0008)

Panel C. Adding the Change in 3-month Treasury Bill Rate

	H_0: No Granger causality from price dispersion to median inflation.	H_0: No Granger causality from median inflation to price dispersion.
In-Sample *F*-Test	10.0639***	16.0604***
	(0.0018)	(0.0002)
Granger-Newbold (GN) Test	-0.6534	0.5144*
	(0.4665)	(0.0748)
Diebold-Mariano (DM) Test	-0.7474	1.2403**
	(0.4859)	(0.0292)
Clark-West (CW) Test	1.1224*	2.3468***
	(0.0994)	(0.0060)
McCracken (MSE-F) Test	-4.2920	4.7889***
	(0.9294)	(0.0028)
Clark-McCracken (ENC-NEW) Test	3.5223**	5.3411***
	(0.0146)	(0.0004)

Notes: Sample test statistics are reported and their p-values (obtained from the wild bootstrap procedure described in Section 4.5) are reported in parentheses. The superscripts ***, ** and * indicate that the null hypothesis can be rejected at the significance levels of 1%, 5% and 10%, respectively.

Table 7. Additional Robustness Checks

Panel A. Excluding Energy Prices

	H_0: No Granger causality from price dispersion to median inflation.	H_0: No Granger causality from median inflation to price dispersion.
In-Sample *F*-Test	6.1508***	14.0184***
	(0.0016)	(0.0070)
Granger-Newbold (GN) Test	0.2374	0.1539
	(0.1142)	(0.2649)
Diebold-Mariano (DM) Test	0.2623	2.3753**
	(0.1174)	(0.0340)
Clark-West (CW) Test	2.5577***	3.5810**
	(0.0066)	(0.0146)
McCracken (MSE-F) Test	2.1305**	11.0255***
	(0.0448)	(0.0062)
Clark-McCracken (ENC-NEW) Test	10.7474***	9.3605***
	(0.0002)	(0.0042)

Panel B. Allowing for Structural Breaks

	H_0: No Granger causality from price dispersion to median inflation.	H_0: No Granger causality from median inflation to price dispersion.
In-Sample *F*-Test	5.4364***	8.4283***
	(0.0048)	(0.0020)
Granger-Newbold (GN) Test	-0.1004	0.6590*
	(0.2196)	(0.0560)
Diebold-Mariano (DM) Test	-0.0709	1.7605***
	(0.2096)	(0.0062)
Clark-West (CW) Test	1.2733*	2.5754***
	(0.0728)	(0.0028)
McCracken (MSE-F) Test	-0.3427	5.7896***
	(0.2595)	(0.0012)
Clark-McCracken (ENC-NEW) Test	3.1778**	5.0384***
	(0.0204)	(0.0010)

Notes: Sample test statistics are reported and their p-values (obtained from the wild bootstrap procedure described in Section 4.5) are reported in parentheses. The superscripts ***, ** and * indicate that the null hypothesis can be rejected at the significance levels of 1%, 5% and 10%, respectively.

Table 8. Out-of-Sample Granger Causality Tests over Six 30-Month Sub-samples

Panel A. Testing for Granger causality from price dispersion to median inflation

Subperiod	Granger-Newbold (GN)	Diebold-Mariano (DM)	Clark-West (CW)	McCracken (MSE-F)	Clark-McCracken (ENC-NEW)
1994M01~1996M06	0.6307	0.7351	1.9634**	1.8690**	2.6257***
	(0.1160)	(0.1568)	(0.0316)	(0.0178)	(0.0032)
1996M07~1998M12	-2.0146	-1.4661	-0.7814	-5.8527	-1.2687
	(0.9436)	(0.8264)	(0.7127)	(0.9950)	(0.9760)
1999M01~2001M06	0.0470	0.4520	1.0982	1.0028*	1.2493**
	(0.3579)	(0.2484)	(0.1422)	(0.0592)	(0.0204)
2001M07~2003M12	0.4713	0.5213	1.0521	0.6631*	0.7550**
	(0.2302)	(0.2631)	(0.1600)	(0.0614)	(0.0184)
2004M01~2006M06	-0.6077	-0.4620	0.1808	-1.4994	0.3055
	(0.5799)	(0.5159)	(0.3817)	(0.9226)	(0.1978)
2006M07~2008M12	-1.2022	-1.1355	-0.6001	-1.8756	-0.5743
	(0.7630)	(0.7830)	(0.6399)	(0.9852)	(0.9628)

Panel B. Testing for Granger causality from median inflation to price dispersion

Subperiod	Granger-Newbold (GN)	Diebold-Mariano (DM)	Clark-West (CW)	McCracken (MSE-F)	Clark-McCracken (ENC-NEW)
1994M01~1996M06	-0.4917	0.6746	1.2556	0.6484*	0.5157*
	(0.5333)	(0.2012)	(0.1068)	(0.0696)	(0.0506)
1996M07~1998M12	0.3519	1.0264	1.0771	1.4934***	1.0148***
	(0.2178)	(0.1322)	(0.1344)	(0.0098)	(0.0078)
1999M01~2001M06	1.2974**	0.1722	0.8348	0.4441	1.1686**
	(0.0392)	(0.3303)	(0.1874)	(0.1256)	(0.0156)
2001M07~2003M12	1.9682**	0.4690	1.2550	1.0425**	1.6456***
	(0.0164)	(0.2464)	(0.1078)	(0.0426)	(0.0066)
2004M01~2006M06	0.7538*	0.7718	1.1130	1.1611**	0.9888**
	(0.0830)	(0.1822)	(0.1350)	(0.0222)	(0.0116)
2006M07~2008M12	0.4271	-0.1614	0.3019	-0.2681	0.2367
	(0.1656)	(0.4151)	(0.3217)	(0.6951)	(0.1566)

Notes: Sample test statistics are reported and their p-values (obtained from the wild bootstrap procedure described in Section 4.5) are reported in parentheses. The superscripts ***, ** and * indicate that the null hypothesis can be rejected at the significance levels of 1%, 5% and 10%, respectively.