N. Gregory Mankiw, Ricardo Reis, and Justin Wolfers

HARVARD UNIVERSITY AND NBER; HARVARD UNIVERSITY; AND STANFORD UNIVERSITY AND NBER

Disagreement About Inflation Expectations

1. Introduction

At least since Milton Friedman's renowned presidential address to the American Economic Association in 1968, expected inflation has played a central role in the analysis of monetary policy and the business cycle. How much expectations matter, whether they are adaptive or rational, how quickly they respond to changes in the policy regime, and many related issues have generated heated debate and numerous studies. Yet throughout this time, one obvious fact is routinely ignored: not everyone has the same expectations.

This oversight is probably explained by the fact that, in much standard theory, there is no room for disagreement. In many (though not all) textbook macroeconomic models, people share a common information set and form expectations conditional on that information. That is, we often assume that everyone has the same expectations because our models say that they should.

The data easily reject this assumption. Anyone who has looked at survey data on expectations, either those of the general public or those of professional forecasters, can attest to the fact that disagreement is substantial. For example, as of December 2002, the interquartile range of inflation expectations for 2003 among economists goes from 1½% to 2½%.

We would like to thank Richard Curtin and Guhan Venkatu for help with data sources, and Simon Gilchrist, Robert King, and John Williams for their comments. Doug Geyser and Cameron Shelton provided research assistance. Ricardo Reis is grateful to the Fundacao Ciencia e Tecnologia, Praxis XXI, for financial support.

Among the general public, the interquartile range of expected inflation goes from 0% to 5%.

This paper takes as its starting point the notion that this disagreement about expectations is itself an interesting variable for students of monetary policy and the business cycle. We document the extent of this disagreement and show that it varies over time. More important, disagreement about expected inflation moves together with the other aggregate variables that are more commonly of interest to economists. This fact raises the possibility that disagreement may be a key to macroeconomic dynamics.

A macroeconomic model that has disagreement at its heart is the sticky-information model proposed recently by Mankiw and Reis (2002). In this model, economic agents update their expectations only periodically because of the costs of collecting and processing information. We investigate whether this model is capable of predicting the extent of disagreement that we observe in the survey data, as well as its evolution over time.

The paper is organized as follows. Section 2 discusses the survey data on expected inflation that will form the heart of this paper. Section 3 offers a brief and selective summary of what is known from previous studies of survey measures of expected inflation, replicating the main findings. Section 4 presents an exploratory analysis of the data on disagreement, documenting its empirical relationship to other macroeconomic variables. Section 5 considers what economic theories of inflation and the business cycle might say about the extent of disagreement. It formally tests the predictions of one such theory—the sticky-information model of Mankiw and Reis (2002). Section 6 compares theory and evidence from the Volcker disinflation. Section 7 concludes.

2. *Inflation Expectations*

Most macroeconomic models argue that inflation expectations are a crucial factor in the inflation process. Yet the nature of these expectations—in the sense of precisely stating whose expectations, over which prices, and over what horizon—is not always discussed with precision. These are crucial issues for measurement.

The expectations of wage- and price-setters are probably the most relevant. Yet it is not clear just who these people are. As such, we analyze data from three sources. The Michigan Survey of Consumer Attitudes and Behavior surveys a cross section of the population about their expectations over the next year. The Livingston Survey and the Survey of Professional Forecasters (SPF) covers more sophisticated analysts—economists working

Table 1 SURVEYS OF INFLATION EXPECTATIONS

	Michigan survey	*Livingston survey*	*Survey of professional forecasters*
Survey population	Cross section of the general public	Academic, business, finance, market, and labor economists	Market economists
Survey organization	Survey Research Center, University of Michigan	Originally Joseph Livingston, an economic journalist; currently the Philadelphia Fed	Originally ASA/NBER; currently the Philadelphia Fed
Average number of respondents	Roughly 1000–3000 per quarter to 1977, then 500–700 per month to present	48 per survey (varies from 14–63)	34 per survey (varies from 9–83)
Starting date	Qualitative questions: 1946 Q1[1]; quantitative responses: January 1978	1946, first half (but the early data is unreliable)[1]	GDP deflator: 1968, Q4; CPI inflation: 1981, Q3
Periodicity	Most quarters from 1947 Q1 to 1977 Q4; every month from January 1978	Semi-annual	Quarterly
Inflation expectation	Expected change in prices over the next 12 months	Consumer Price Index (this quarter, in 2 quarters, in 4 quarters)	GDP deflator level, quarterly CPI level (6 quarters)

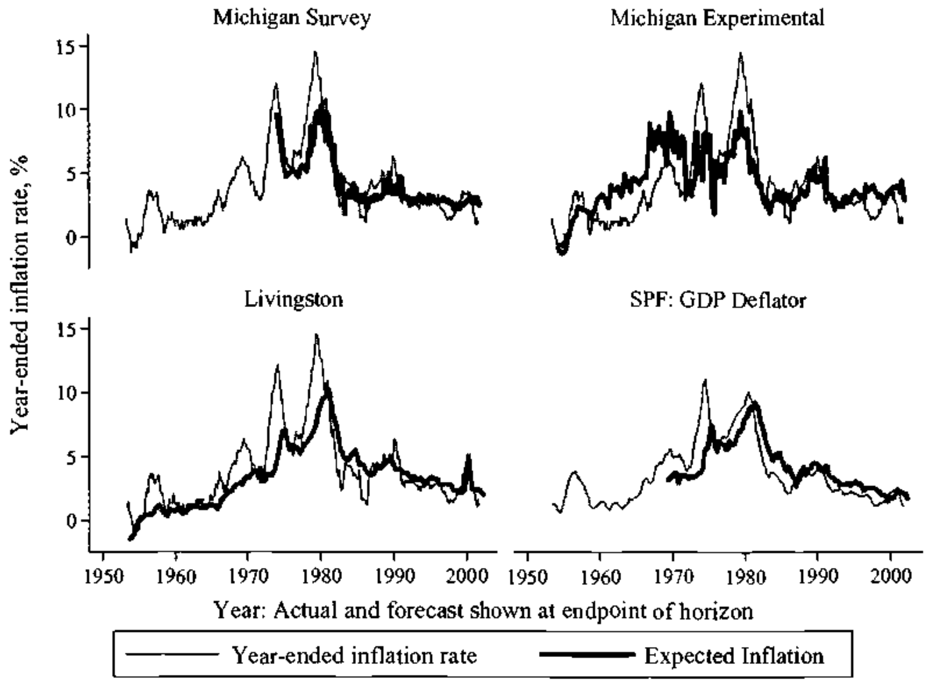
1. Our quantitative work focuses on the period from 1954 onward.

in industry and professional forecasters, respectively. Table 1 provides some basic details about the structure of these three surveys.[1]

Although we have three sources of inflation expectations data, throughout this paper we will focus on four, and occasionally five, series. Most papers analyzing the Michigan data cover only the period since 1978, during which these data have been collected monthly (on a relatively consistent basis), and respondents were asked to state their precise quantitative

1. For more details about the Michigan Survey, the Livingston Survey and the SPF, see Curtin (1996), Croushore (1997), and Croushore (1993), respectively.

Figure 1 MEDIAN INFLATION EXPECTATIONS AND ACTUAL INFLATION



inflation expectations. However, the Michigan Survey of Consumer Attitudes and Behaviors has been conducted quarterly since 1946, although for the first 20 years respondents were asked only whether they expected prices to rise, fall, or stay the same. We have put substantial effort into constructing a consistent quarterly time series for the central tendency and dispersion of inflation expectations through time since 1948. We construct these data by assuming that discrete responses to whether prices are expected to rise, remain the same, or fall over the next year reflect underlying continuous expectations drawn from a normal distribution, with a possibly time-varying mean and standard deviation.[2] We will refer to these constructed data as the Michigan experimental series.

Our analysis of the Survey of Professional Forecasters will occasionally switch between our preferred series, which is the longer time series of forecasts focusing on the gross domestic product (GDP) deflator (starting in 1968, Q4), and the shorter consumer price index (CPI) series (which begins in 1981, Q3).

Figure 1 graphs our inflation expectations data. The horizontal axis refers to expectations at the endpoint of the relevant forecast horizon

2. Construction of this experimental series is detailed in the appendix, and we have published these data online at www.stanford.edu/people/jwolfers (updated January 13, 2004).

rather than at the time the forecast was made. Two striking features emerge from these plots. First, each series yields relatively accurate inflation forecasts. And second, despite the different populations being surveyed, they all tell a somewhat similar story.

By simple measures of forecast accuracy, all three surveys appear to be quite useful. Table 2 shows two common measures of forecast accuracy: the square root of the average squared error (RMSE) and the mean absolute error (MAE). In each case we report the accuracy of the median expectation in each survey, both over their maximal samples and for a common sample (September 1982–March 2002).

Panel A of the table suggests that inflation expectations are relatively accurate. As the group making the forecast becomes increasingly sophisticated, forecast accuracy appears to improve. However, Panel B suggests that these differences across groups largely reflect the different periods over which each survey has been conducted. For the common sample that all five measures have been available, they are all approximately equally accurate.

Of course, these results reflect the fact that these surveys have a similar central tendency, and this fact reveals as much as it hides. Figure 2 presents simple histograms of expected inflation for the coming year as of December 2002.

Here, the differences among these populations become starker. The left panel pools responses from the two surveys of economists and shows some agreement on expectations, with most respondents expecting inflation in the 1½ to 3% range. The survey of consumers reveals substantially greater disagreement. The interquartile range of consumer expectations stretches from 0 to 5%, and this distribution shows quite long tails, with 5% of the population expecting deflation, while 10% expect inflation of at

Table 2 INFLATION FORECAST ERRORS

	Michigan	*Michigan experimental*	*Livingston*	*SPF–GDP deflator*	*SPF–CPI*
Panel A: maximal sample					
Sample	Nov. 1974–May 2002	1954, Q4–2002, Q1	1954, H1–2001, H2	1969, Q4–2002, Q1	1982, Q3–2002, Q1
RMSE	1.65%	2.32%	1.99%	1.62%	1.29%
MAE	1.17%	1.77%	1.38%	1.22%	0.97%
Panel B: common time period (September 1982–March 2002)					
RMSE	1.07%	1.24%	1.28%	1.10%	1.29%
MAE	0.85%	0.95%	0.97%	0.91%	0.97%

least 10%. These long tails are a feature throughout our sample and are not a particular reflection of present circumstances. Our judgment (following Curtin, 1996) is that these extreme observations are not particularly informative, and so we focus on the median and interquartile range as the relevant indicators of central tendency and disagreement, respectively.

The extent of disagreement within each of these surveys varies dramatically over time. Figure 3 shows the interquartile range over time for each of our inflation expectations series. A particularly interesting feature of these data is that disagreement among professional forecasters rises and falls with disagreement among economists and the general public. Table 3 confirms that all of our series show substantial co-movement. This table focuses on quarterly data—by averaging the monthly Michigan numbers and linearly interpolating the semiannual Livingston numbers. Panel A shows correlation coefficients among these quarterly estimates. Panel B shows correlation coefficients across a smoothed version of the data (a five-quarter centered moving average of the interquartile range). (The experimental Michigan data show a somewhat weaker correlation, particularly in the high-frequency data, probably reflecting measurement error caused by the fact that these estimates rely heavily on the proportion of the sample expecting price declines—a small and imprecisely estimated fraction of the population.)

Figure 2 DISTRIBUTION OF INFLATION EXPECTATIONS

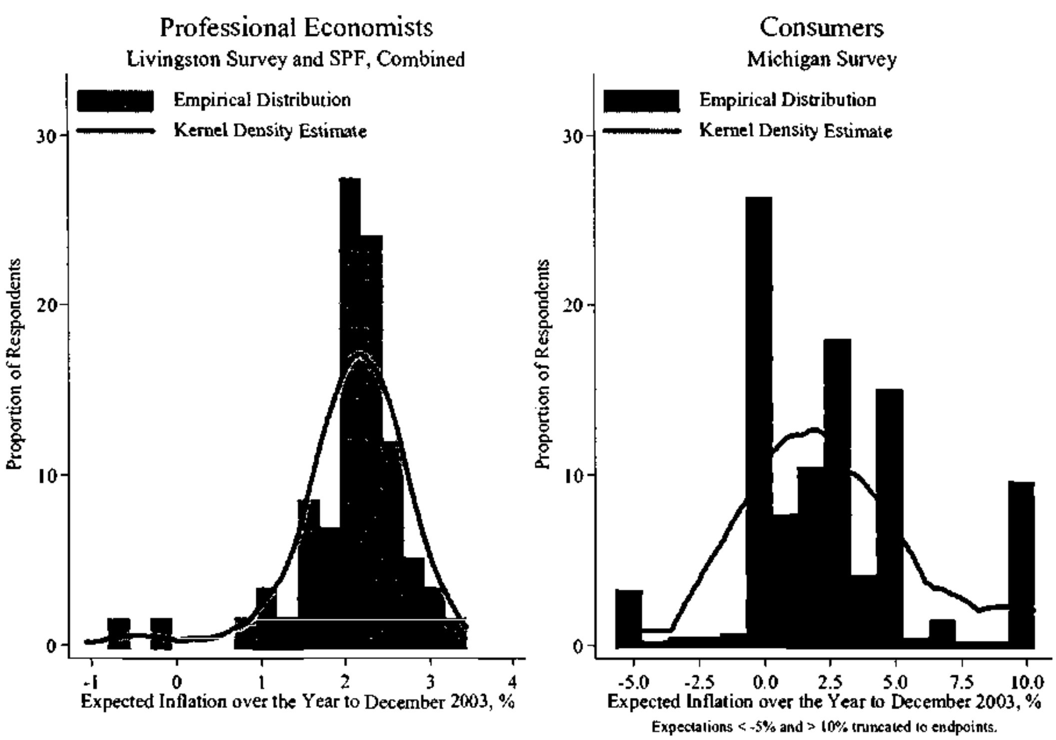
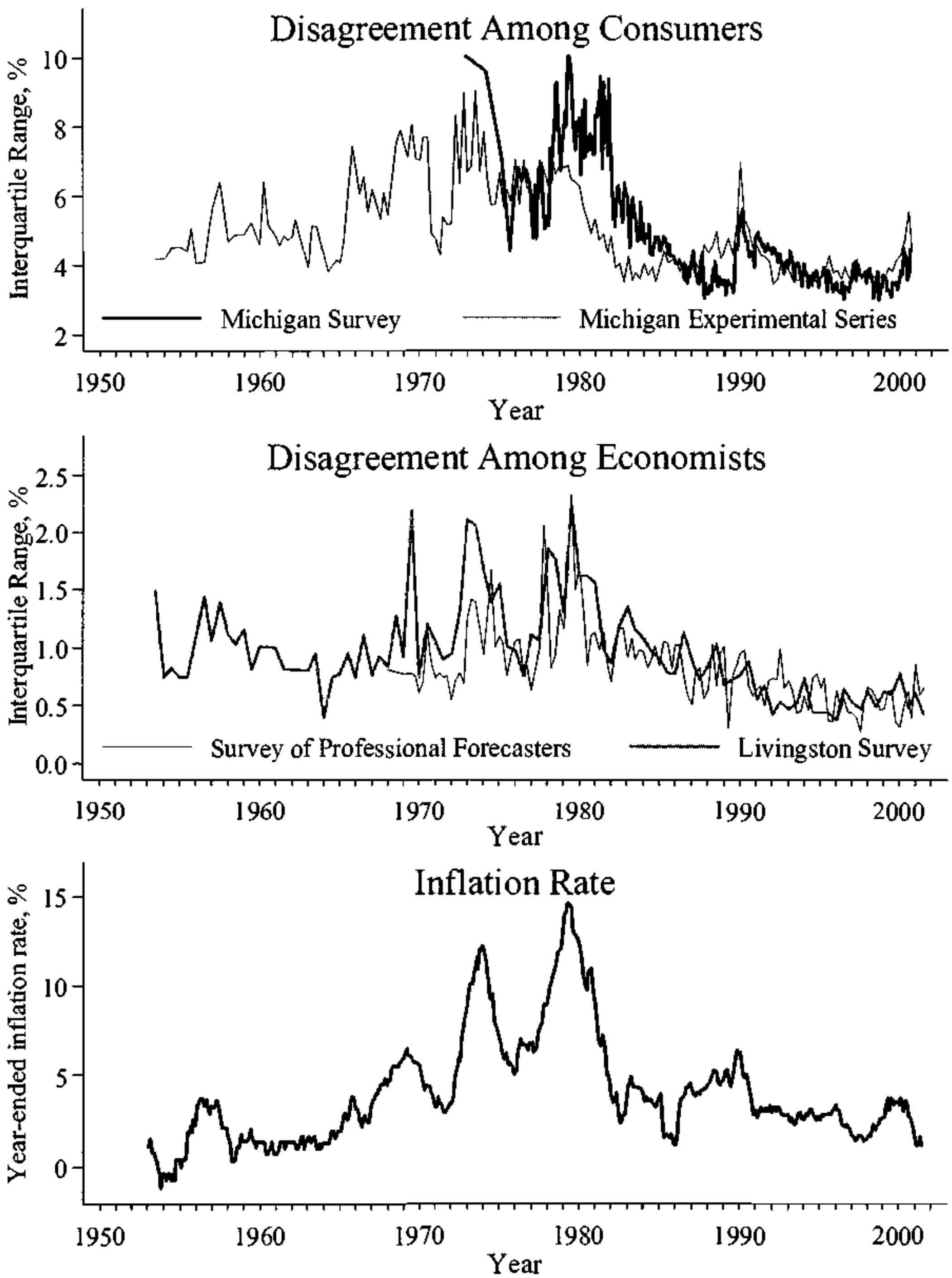


Figure 3 DISAGREEMENT OVER INFLATION EXPECTATIONS THROUGH TIME



Date reflects when the forecast is made.

Table 3 DISAGREEMENT THROUGH TIME: CORRELATION ACROSS SURVEYS[1]

	Michigan	*Michigan experimental*	*Livingston*	*SPF–GDP deflator*	*SPF–CPI*
Panel A: actual quarterly data					
Michigan	1.000				
Michigan experimental	0.682	1.000			
Livingston	0.809	0.391	1.000		
SPF–GDP deflator	0.700	0.502	0.712	1.000	
SPF–CPI	0.667	0.231	0.702	0.688	1.000
Panel B: 5 quarter centered moving averages					
Michigan	1.000				
Michigan experimental	0.729	1.000			
Livingston	0.869	0.813	1.000		
SPF–GDP deflator	0.850	0.690	0.889	1.000	
SPF–CPI	0.868	0.308	0.886	0.865	1.000

1. Underlying data are quarterly. They are created by taking averages of monthly Michigan data and by linearly interpolating half-yearly Livingston data.

A final source of data on disagreement comes from the range of forecasts within the Federal Open Market Committee (FOMC), as published biannually since 1979 in the Humphrey-Hawkins testimony.[3] Individual-level data are not released, so we simply look to describe the broad pattern of disagreement among these experts. Figure 4 shows a rough (and statistically significant) correspondence between disagreement among policymakers and disagreement among professional economists. The correlation of the *range* of FOMC forecasts with the interquartile range of the Livingston population is 0.34, 0.54 or 0.63, depending on which of the three available FOMC forecasts we use. While disagreement among Fed-watchers rose during the Volcker disinflation, the range of inflation forecasts within the Fed remained largely constant—the correlation between disagreement among FOMC members and disagreement among professional forecasters is substantially higher after 1982.

We believe that we have now established three important patterns in the data. First, there is substantial disagreement within both naïve and expert

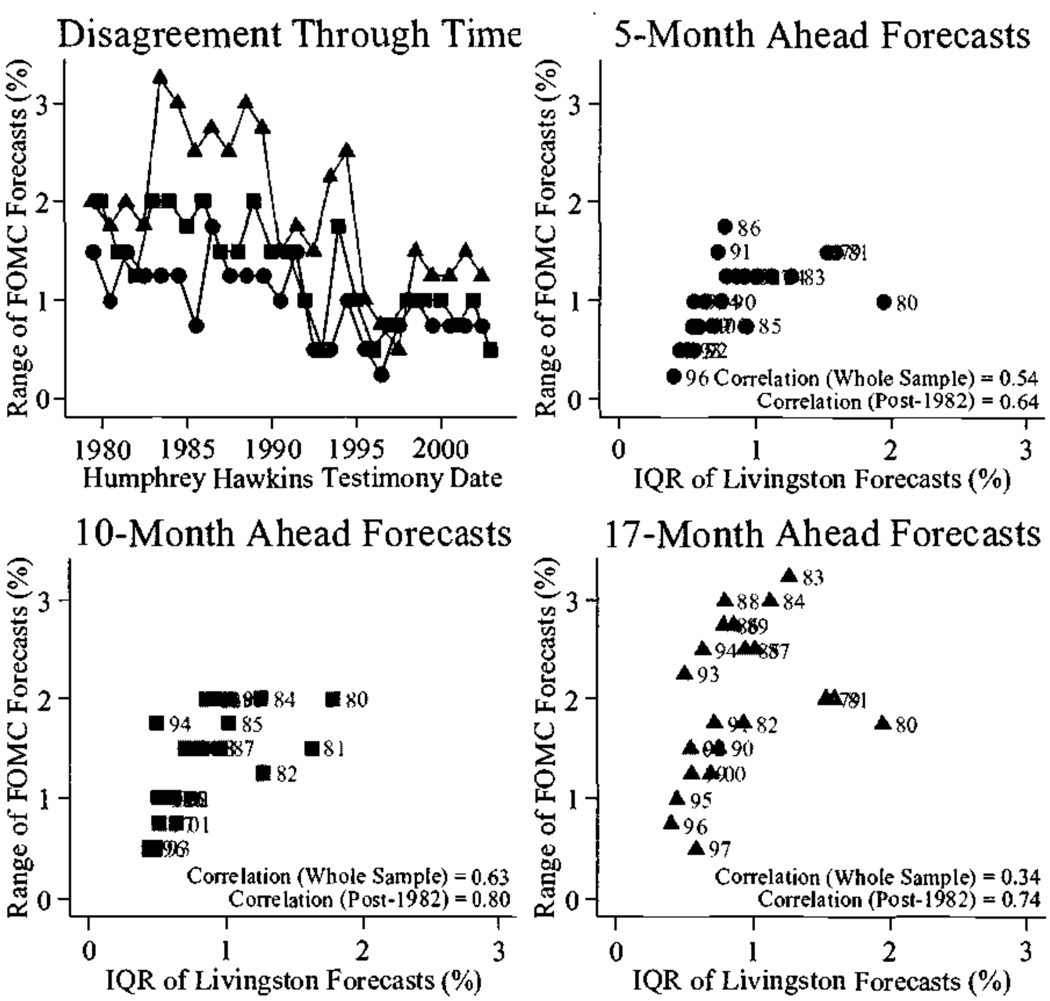
3. We are grateful to Simon Gilchrist for suggesting this analysis to us. Data were drawn from Gavin (2003) and updated using recent testimony published at http://www.federalreserve.gov/boarddocs/hh/ (accessed December 2003).

populations about the expected future path of inflation. Second, there are larger levels of disagreement among consumers than exists among experts. And third, even though professional forecasters, economists, and the general population show different degrees of disagreement, this disagreement tends to exhibit similar time-series patterns, albeit of a different amplitude. One would therefore expect to find that the underlying causes behind this disagreement are similar across all three datasets.

3. *The Central Tendency of Inflation Expectations*

Most studies analyzing inflation expectations data have explored whether empirical estimates are consistent with rational expectations. The rational expectations hypothesis has strong implications for the time series of

Figure 4 DISAGREEMENT AMONG THE FOMC



Humphrey-Hawkins testimony in February and July provides forecasts for inflation over the calendar year. Inflation concept varies.

expectations data, most of which can be stated in terms of forecast efficiency. More specifically, rational expectations imply (statistically) efficient forecasting, and efficient forecasts do not yield predictable errors. We now turn to reviewing the tests of rationality commonly found in the literature and to providing complementary evidence based on the estimates of median inflation expectations in our sample.[4]

The simplest test of efficiency is a test for bias: are inflation expectations centered on the right value? Panel A of Table 4 reports these results, regressing expectation errors on a constant. Median forecasts have tended to underpredict inflation in two of the four data series, and this divergence is statistically significant; that said, the magnitude of this bias is small.[5]

By regressing the forecast error on a constant and the median inflation expectation,[6] panel B of the table tests whether there is information in these inflation forecasts themselves that can be used to predict forecasting errors. Under the null of rationality, these regressions should have no predictive power. Both the Michigan and Livingston series can reject a rationality null on this score, while the other two series are consistent with this (rather modest) requirement of rationality.

Panel C exploits a time-series implication of rationality, asking whether today's errors can be forecasted based on yesterday's errors. In these tests, we regress this year's forecast error on the realized error over the previous year. Evidence of autocorrelation suggests that there is information in last year's forecast errors that is not being exploited in generating this year's forecast, violating the rationality null hypothesis. We find robust evidence of autocorrelated forecast errors in all surveys. When interpreting these coefficients, note that they reflect the extent to which errors made a year ago persist in today's forecast. We find that, on average, about half of the error remains in the median forecast. One might object that last year's forecast error may not yet be fully revealed by the time this year's forecast is made because inflation data are published with only one month lag. Experimenting with slightly longer lags does not change these results significantly.[7]

Finally, panel D asks whether inflation expectations take sufficient account of publicly available information. We regress forecast errors on recent macroeconomic data. Specifically, we analyze the inflation rate, the Treasury-bill rate, and the unemployment rate measured one month prior

4. Thomas (1999) provides a survey of this literature.

5. Note that the construction of the Michigan experimental data makes the finding of bias unlikely for that series.

6. Some readers may be more used to seeing regressions of the form $\pi = a + bE_{t-12}\pi_t$, where the test for rationality is a joint test of $a = 0$ and $b = 1$. To see that our tests are equivalent, simply rewrite $\pi_t - E_{t-12}\pi_t = a + (1 - b)E_{t-12}\pi_t$. A test of $a = 0$ and $b = 1$ translates into a test that the constant and slope coefficient in this equation are both zero.

7. Repeating this analysis with mean rather than median expectations yields weaker results.

Table 4 TESTS OF FORECAST RATIONALITY: MEDIAN INFLATION EXPECTATIONS[1]

	Michigan	*Michigan-experimental*	*Livingston*	*SPF (GDP deflator)*
Panel A: testing for bias: $\pi_t - E_{t-12}\,\pi_t = \alpha$				
α: mean error	0.42%	−0.09%	0.63%**	−0.02%
(Constant only)	(0.29)	(0.34)	(0.30)	(0.29)
Panel B: Is information in the forecast fully exploited? $\pi_t - E_{t-12}\,\pi_t = \alpha + \beta\, E_{t-12}\pi_t$				
β: $E_{t-12}\,[\pi_t]$	0.349**	−0.060	0.011	0.026
	(.161)	(.207)	(.142)	(.128)
α: constant	−1.016%*	−0.182%	0.595%	−0.132%
	(.534)	(.721)	(.371)	(.530)
Adj. R^2	0.197	−0.003	−0.011	−0.007
Reject eff.? $\alpha = \beta = 0$	Yes	No	Yes	No
(p-value)	(p = 0.088)	(p = 0.956)	(p = 0.028)	(p = 0.969)
Panel C: Are forecasting errors persistent? $\pi_t - E_{t-12}\,\pi_t = \alpha + \beta\,(\pi_{t-12} - E_{t-24}\pi_{t-12})$				
β: $\pi_{t-12} - E_{t-24}\,[\pi_{t-12}]$	0.371**	.580***	0.490***	0.640***
	(0.158)	(0.115)	(0.132)	(0.224)
α: constant	0.096%	0.005%	0.302%	−0.032%
	(0.183)	(0.239)	(0.210)	(0.223)
Adj. R^2	0.164	0.334	0.231	0.375
Panel D: Are macroeconomic data fully exploited? $\pi_t - E_{t-12}\pi_t = \alpha + \beta\, E_{t-12}\,[\pi_t] + \gamma\,\pi_{t-13} + \kappa\, i_{t-13} + \delta\, U_{t-13}$				
α: constant	−0.816%	0.242%	4.424%***	3.566%***
	(0.975)	(1.143)	(0.985)	(0.970)
β: $E_{t-12}\,[\pi_t]$	0.801***	−0.554***	0.295	0.287
	(0.257)	(0.165)	(0.283)	(0.308)
γ: inflation$_{t-13}$	−0.218*	0.610***	0.205	0.200
	(0.121)	(0.106)	(0.145)	(0.190)
κ: Treasury bill$_{t-13}$	−0.165**	−0.024	−0.319***	−0.321***
	(0.085)	(0.102)	(0.106)	(0.079)
δ: unemployment$_{t-13}$	0.017	−0.063	−0.675***	−0.593***
	(0.126)	(0.156)	(0.175)	(0.150)
Reject eff.? $\gamma = \kappa = \delta = 0$	Yes	Yes	Yes	Yes
(p-value)	(p = 0.049)	(p = 0.000)	(p = 0.000)	(p = 0.000)
Adjusted R^2	0.293	0.382	0.306	0.407
Sample	Nov. 1974– May 2002	1954, Q4– 2002, Q1	1954, H1– 2001, H2	1969, Q4– 2002, Q1
Periodicity	Monthly	Quarterly	Semiannual	Quarterly
N	290	169	96	125

1. ***, ** and * denote statistical significance at the 1%, 5%, and 10% levels, respectively (Newey-West standard errors in parentheses; correcting for autocorrelation up to one year).

to the forecast because these data are likely to be the most recent published data when forecasts were made. We also control for the forecast itself, thereby nesting the specification in panel B of Table 4. One might object that using real-time data would better reflect the information available when forecasts were made; we chose these three indicators precisely because they are subject to only minor revisions. Across the three different pieces of macroeconomic information and all four surveys, we often find statistical evidence that agents are not fully incorporating this information in their inflation expectations. Simple bivariate regressions (not shown) yield a qualitatively similar pattern of responses. The advantage of the multivariate regression is that we can perform an F-test of the joint significance of the lagged inflation, interest rates, and unemployment rates in predicting forecast errors. In each case the macroeconomic data are overwhelmingly jointly statistically significant, suggesting that median inflation expectations do not adequately account for recent available information. Note that these findings do not depend on whether we condition on the forecast of inflation.

Ball and Croushore (2003) interpret the estimated coefficients in a regression similar to that in panel D as capturing the extent to which agents under- or overreact to information. For instance, under the implicit assumption that, in the data, high inflation this period will tend to be followed by high inflation in the next period, the finding that the coefficient on inflation in panel D is positive implies that agents have underreacted to the recent inflation news. Our data support this conclusion in three of the four regressions (the Michigan series is the exception). Similarly, a high nominal interest rate today could signal lower inflation tomorrow because it indicates contractionary monetary policy by the Central Bank. We find that forecasts appear to underreact to short-term interest rates in all four regressions—high interest rates lead forecasters to make negative forecast errors or to predict future inflation that is too high. Finally, if in the economy a period of higher unemployment is usually followed by lower inflation (as found in estimates of the Phillips curve), then a negative coefficient on unemployment in panel D would indicate that agents are overestimating inflation following a rise in unemployment and thus are underreacting to the news in higher unemployment. We find that inflation expectations of economists are indeed too high during periods of high unemployment, again suggesting a pattern of underreaction; this is an error not shared by consumers. Our results are in line with Ball and Croushore's (2003) finding that agents seem to underreact to information when forming their expectations of inflation.

In sum, Table 4 suggests that each of these data series alternatively meets and fails some of the implications of rationality. Our sense is that these results probably capture the general flavor of the existing empirical literature, if not the somewhat stronger arguments made by individual authors. Bias exists but is typically small. Forecasts are typically inefficient, though not in all surveys: while the forecast errors of economists are not predictable based merely on their forecasts, those of consumers are. All four data series show substantial evidence that forecast errors made a year ago continue to repeat themselves, and that recent macroeconomic data is not adequately reflected in inflation expectations.

We now turn to analyzing whether the data are consistent with adaptive expectations, probably the most popular alternative to rational expectations in the literature. The simplest backward-looking rule invokes the prediction that expected inflation over the next year will be equal to inflation over the past year. Ball (2000) suggests a stronger version, whereby agents form statistically optimal univariate inflation forecasts. The test in Table 5 is a little less structured, simply regressing median inflation expectations against the last eight nonoverlapping, three-month-ended inflation observations. We add the unemployment rate and short-term interest rates to this regression, finding that these macroeconomic aggregates also help predict inflation expectations. In particular, it is clear that when the unemployment rate rises over the quarter, inflation expectations fall further than adaptive expectations might suggest. This suggests that consumers employ a more sophisticated model of the economy than assumed in the simple adaptive expectations model.

Consequently we are left with a somewhat negative result—observed inflation expectations are consistent with neither the sophistication of rational expectations nor the naïveté of adaptive expectations. This finding holds for our four datasets, and it offers a reasonable interpretation of the prior literature on inflation expectations. The common thread to these results is that inflation expectations reflect partial and incomplete updating in response to macroeconomic news. We shall argue in Section 5 that these results are consistent with models in which expectations are not updated at every instant, but rather in which updating occurs in a staggered fashion. A key implication is that disagreement will vary with macroeconomic conditions.

4. Dispersion in Survey Measures of Inflation Expectations

Few papers have explored the features of the cross-sectional variation in inflation expectations. Bryan and Venkatu (2001) examine a survey of inflation expectations in Ohio from 1998–2001, finding that women, singles, nonwhites, high school dropouts, and lower income groups tend to have higher inflation expectations than other demographic groups. They

Table 5 TESTS OF ADAPTIVE EXPECTATIONS: MEDIAN INFLATION EXPECTATIONS[1]

	Michigan	*Michigan-experimental*	*Livingston*	*SPF (GDP deflator)*
Adaptive expectations: $E_t\pi_{t+12} = \alpha + \beta(L)\,\pi_t + \gamma\,U_t + \kappa\,U_{t-3} + \delta\,i_t + \phi\,i_{t-3}$				
Inflation				
$\beta(1)$: sum of	0.706***	0.635***	0.530***	0.581***
8 coefficients	(0.037)	(0.085)	(0.048)	(0.054)
Unemployment				
γ: date of forecast	−0.633**	−1.237**	−0.755***	−0.405**
	(0.261)	(0.488)	(0.192)	(0.162)
κ: 3 months prior	0.585	0.555	1.055***	0.593***
	(0.231)	(0.467)	(0.185)	(0.171)
Treasury bill rate				
δ: date of forecast	0.035	−0.053	0.143**	0.069*
	(0.038)	(0.132)	(0.058)	(0.039)
ϕ: 3 months prior	−0.109**	−0.052	0.100**	0.144***
	(0.045)	(0.122)	(0.049)	(0.047)
Reject adaptive	$F_{4,277} = 9.94$***	$F_{4,156} = 6.67$***	$F_{4,83} = 24.5$***	$F_{4,112} = 13.4$***
expectations?	Yes	Yes	Yes	Yes
($\gamma = \kappa = \delta = \phi = 0$)				
Adjusted R^2	0.922	0.539	0.916	0.929
N	290	169	96	125
	(monthly)	(quarterly)	(semiannual)	(quarterly)

1. ***, ** and * denote statistical significance at the 1%, 5%, and 10% levels, respectively (Newey-West standard errors in parentheses; correcting for autocorrelation up to a year).

note that these differences are too large to be explained by differences in the consumption basket across groups but present suggestive evidence that differences in expected inflation reflect differences in the perceptions of current inflation rates. Vissing-Jorgenson (this volume) also explores differences in inflation expectations across age groups.

Souleles (2001) finds complementary evidence from the Michigan Survey that expectations vary by demographic group, a fact that he interprets as evidence of nonrational expectations. Divergent expectations across groups lead to different expectation errors, which he relates to differential changes in consumption across groups.

A somewhat greater share of the research literature has employed data on the dispersion in inflation expectations as a rough proxy for inflation uncertainty. These papers have suggested that highly dispersed inflation expectations are positively correlated with the inflation rate and, conditional on current inflation, are related positively to the recent variance of measured inflation (Cukierman and Wachtel, 1979), to weakness in the real economy (Mullineaux, 1980; Makin, 1982), and alternatively to lower interest rates (Levi and Makin, 1979; Bomberger and Frazer, 1981; and Makin, 1983), and to higher interest rates (Barnea, Dotan, and Lakonishok, 1979; Brenner and Landskroner, 1983). These relationships do not appear to be particularly robust, and in no case is more than one set of expectations data brought to bear on the question. Our approach is consistent with a more literal interpretation of the second moment of the expectations data: we interpret different inflation expectations as reflecting disagreement in the population; that is, different forecasts reflect different expectations.

Lambros and Zarnowitz (1987) argue that disagreement and uncertainty are conceptually distinct, and they make an attempt at unlocking the two empirically. Their data on uncertainty derives from the SPF, which asks respondents to supplement their point estimates with estimates of the probability that GDP and the implicit price deflator will fall into various ranges. These two authors find only weak evidence that uncertainty and disagreement share a common time-series pattern. Intrapersonal variation in expected inflation (uncertainty) is larger than interpersonal variation (disagreement), and while there are pronounced changes through time in disagreement, uncertainty varies little.

The most closely related approach to the macroeconomics of disagreement comes from Carroll (2003b), who analyzes the evolution of the standard deviation of inflation expectations in the Michigan Survey. Carroll provides an epidemiological model of inflation expectations in which expert opinion slowly spreads person to person, much as disease spreads through a population. His formal model yields something close to the Mankiw and Reis (2002) formulation of the sticky-information model. In an agent-based simulation, he proxies expert opinion by the average forecast in the Survey of Professional Forecasters and finds that his agent-based model tracks the time series of disagreement quite well, although it cannot match the level of disagreement in the population.

We now turn to analyzing the evolution of disagreement in greater detail. Figure 3 showed the inflation rate and our measures of disagreement. That figure suggested a relatively strong relationship between inflation and disagreement. A clearer sense of this relationship can be seen in Figure 5. Beyond this simple relationship in levels, an equally apparent fact from Figure 3 is that, when the inflation rate moves around a lot, dispersion appears to rise. This fact is illustrated in Figure 6.

In all four datasets, large changes in inflation (in either direction) are correlated with an increase in disagreement. This fanning out of inflation expectations following a change in inflation is consistent with a process of staggered adjustment of expectations. Of course, the change in inflation is (mechanically) related to its level, and we will provide a more careful attempt at sorting change and level effects below.

Figure 7 maps the evolution of disagreement and the real economy through time. The charts show our standard measures of disagreement,

Figure 5 INFLATION AND DISAGREEMENT

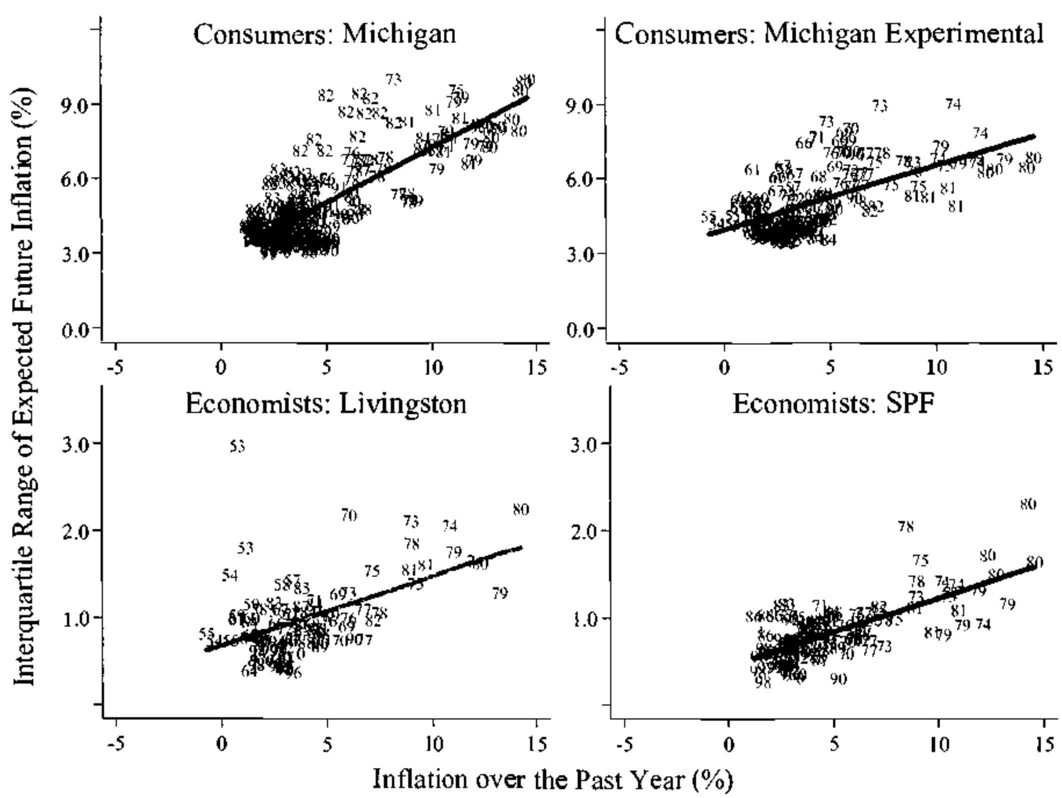
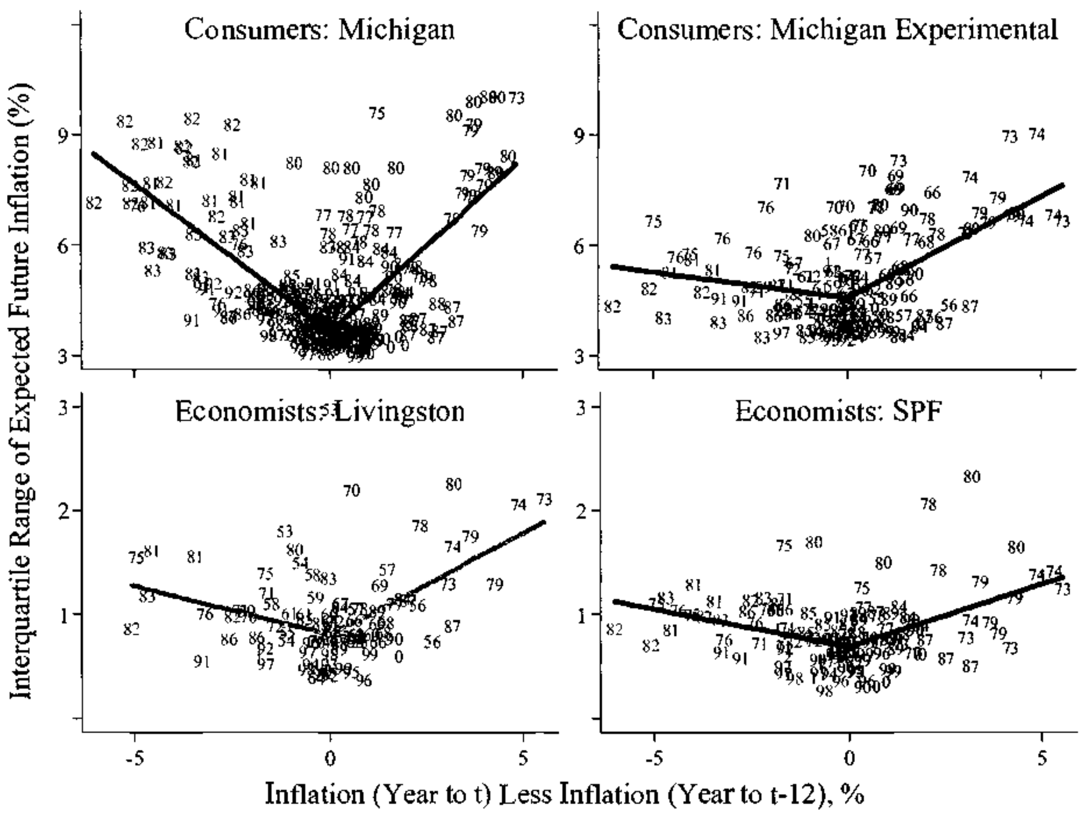


Figure 6 CHANGES IN INFLATION AND DISAGREEMENT



plus two measures of excess capacity: an output gap constructed as the difference between the natural logs of actual chain-weighted real output and trend output (constructed from a Hodrick-Prescott filter). The shaded regions represent periods of economic expansion and contraction as marked by the National Bureau of Economic Research (NBER) Business Cycle Dating Committee.[8]

The series on disagreement among consumers appears to rise during recessions, at least through the second half of the sample. A much weaker relationship is observed through the first half of the sample. Disagreement among economists shows a less obvious relationship with the state of the real economy.

The final set of data that we examine can be thought of as either a cause or consequence of disagreement in inflation expectations. We consider the dispersion in actual price changes across different CPI categories. That is, just as Bryan and Cecchetti (1994) produce a weighted median CPI by calculating rates of inflation across 36 commodity groups, we construct a weighted interquartile range of year-ended inflation rates across

8. We have also experimented using the unemployment rate as a measure of real activity and obtained similar results.

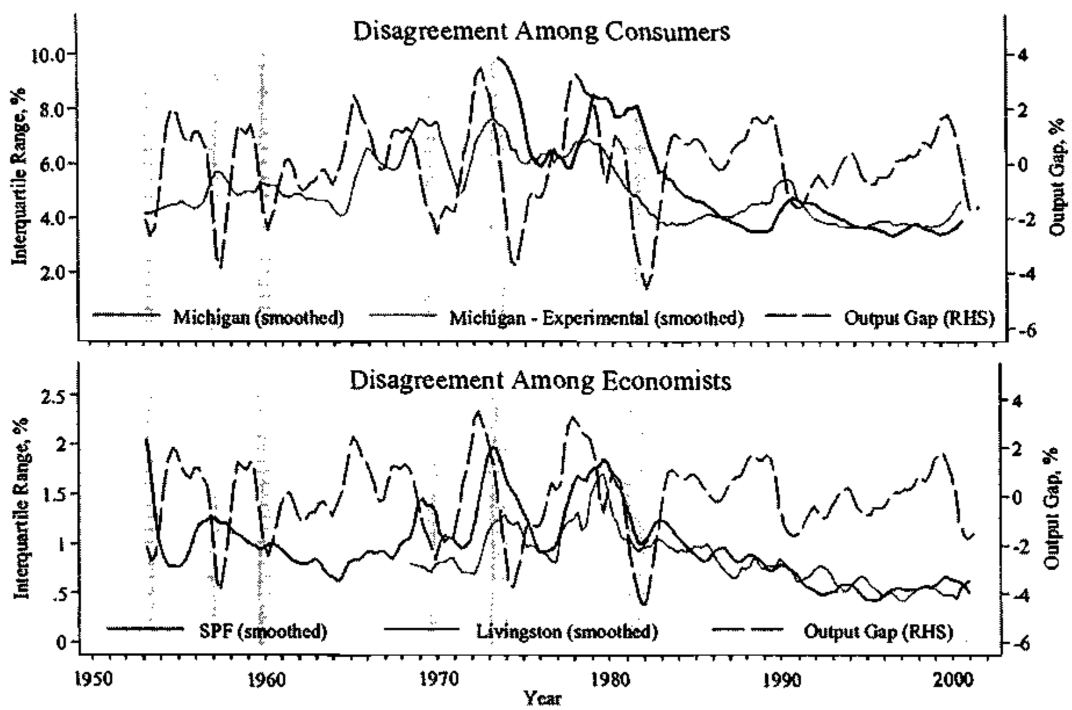
commodity groups. One could consider this a measure of the extent to which relative prices are changing. We analyze data for the period December 1967–December 1997 provided by the Cleveland Fed. Figure 8 shows the median inflation rate and the 25th and 75th percentiles of the distribution of nominal price changes.

Dispersion in commodity-level rates of inflation seems to rise during periods in which the dispersion in inflation expectations rises. In Figure 9, we confirm this, graphing this measure of dispersion in rates of price change against our measures of dispersion in expectations. The two look to be quite closely related.

Table 6 considers each of the factors discussed above simultaneously, reporting regressions of the level of disagreement against inflation, the squared change in inflation, the output gap, and the dispersion in different commodities' actual inflation rates. Across the four table columns, we tend to find larger coefficients in the regressions focusing on consumer expectations than in those of economists. This reflects the differences in the extent of disagreement, and how much it varies over the cycle, across these populations.

In both bivariate and multivariate regressions, we find the inflation rate to be an extremely robust predictor of disagreement. The squared change

Figure 7 DISAGREEMENT AND THE REAL ECONOMY



Shaded areas denoted NBER-dated recessions

Figure 8 DISTRIBUTION OF INFLATION RATES ACROSS CPI COMPONENTS

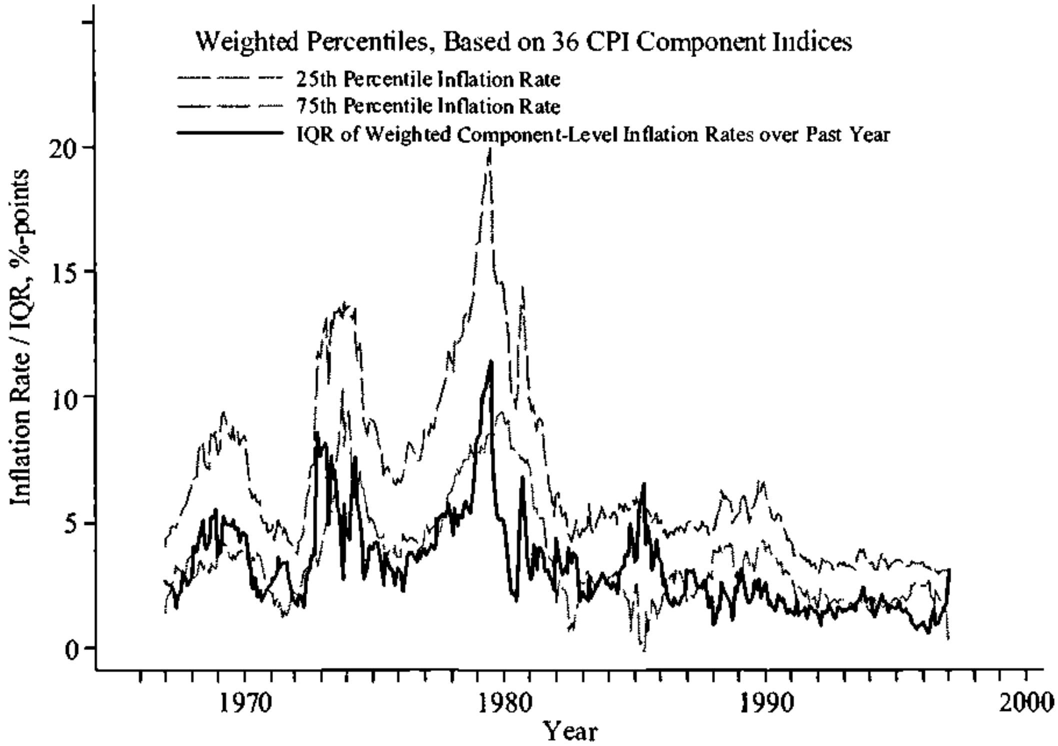


Figure 9 DISPERSION IN INFLATION EXPECTATIONS AND DISPERSION IN INFLATION RATES ACROSS DIFFERENT CPI COMPONENTS

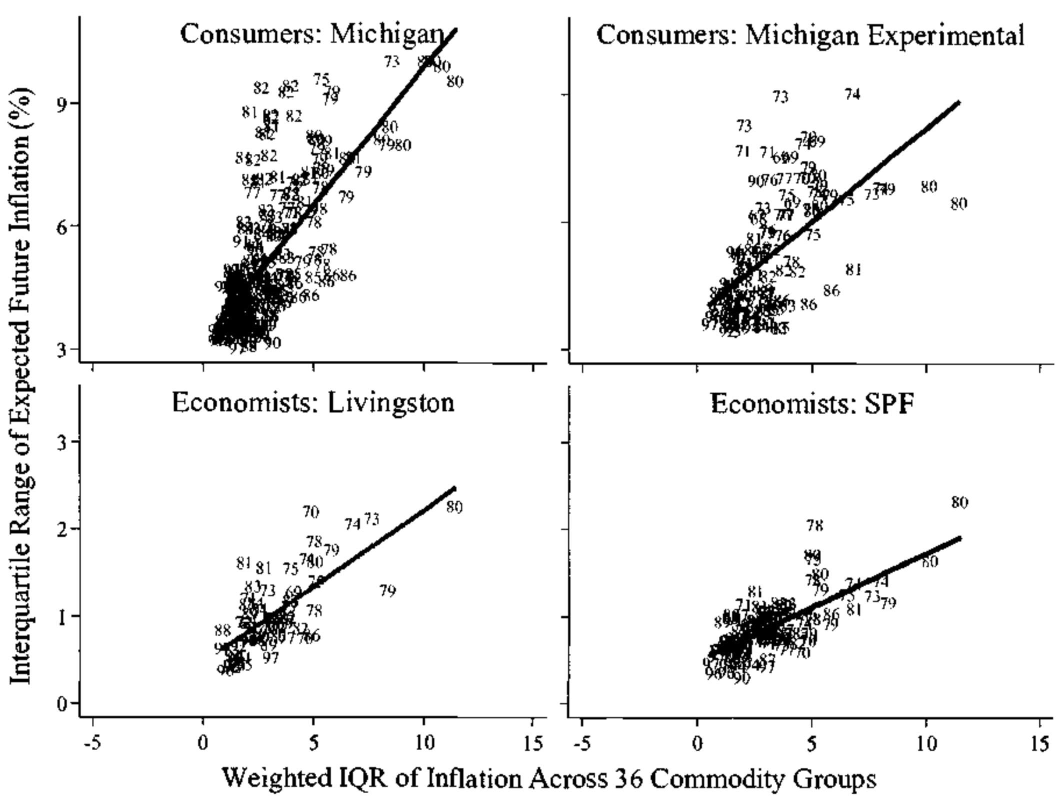


Table 6 DISAGREEMENT AND THE BUSINESS CYCLE: ESTABLISHING STYLIZED FACTS[1]

	Michigan	*Michigan-experimental*	*Livingston*	*SPF (GDP deflator)*
Panel A: bivariate regressions (each cell represents a separate regression)				
Inflation rate	0.441***	0.228***	0.083***	0.092***
	(0.028)	(0.036)	(0.016)	(0.013)
ΔInflation-squared	18.227***	1.259**	2.682***	2.292**
	(2.920)	(0.616)	(0.429)	(0.084)
Output gap	0.176	−0.047	0.070**	−0.001
	(0.237)	(0.092)	(0.035)	(0.029)
Relative price variability	0.665***	0.473***	0.117**	0.132
	(0.056)	(0.091)	(0.046)	(0.016)
Panel B: regressions controlling for the inflation rate (each cell represents a separate regression)				
ΔInflation-squared	10.401***	0.814	2.051***	−0.406
	(1.622)	(0.607)	(0.483)	(0.641)
Output gap	0.415***	0.026	−0.062**	−0.009
	(0.088)	(0.086)	(0.027)	(0.013)
Relative price variability	0.268***	0.210	0.085**	0.099***
	(0.092)	(0.135)	(0.042)	(0.020)
Panel C: multivariate regressions (full sample)				
Inflation rate	0.408***	0.217***	0.066***	0.095***
	(0.028)	(0.034)	(0.013)	(0.015)
ΔInflation-squared	7.062***	0.789	1.663**	−0.305
	(1.364)	(0.598)	(0.737)	(0.676)
Output gap	0.293***	0.017	0.020	−0.007
	(0.066)	(0.079)	(0.032)	(0.014)
Panel D: multivariate regressions (including inflation dispersion)				
Inflation rate	0.328***	0.204***	0.044**	0.037***
	(0.034)	(0.074)	(0.018)	(0.011)
ΔInflation-squared	5.558***	−0.320	1.398	−0.411
	(1.309)	(2.431)	(0.949)	(0.624)
Output gap	0.336***	−0.061	0.013	0.006
	(0.067)	(0.117)	(0.039)	(0.018)
Relative price variability	0.237***	0.210	0.062	0.100***
	(0.079)	(0.159)	(0.038)	(0.022)

1. *** and ** denote statistical significance at the 1% and 5% levels, respectively (Newey-West standard errors in parentheses; correcting for autocorrelation up to one year).

in inflation is highly correlated with disagreement in bivariate regressions, and controlling for the inflation rate and other macroeconomic variables only slightly weakens this effect. Adding the relative price variability term further weakens this effect. Relative price variability is a consistently strong predictor of disagreement across all specifications. These results are generally stronger for the actual Michigan data than for the experimental series, and they are generally stronger for the Livingston series than for the SPF. We suspect that both facts reflect the relative role of measurement error. Finally, while the output gap appears to be related to disagreement in certain series, this finding is not robust either across data series or to the inclusion of controls.

In sum, our analysis of the disagreement data has estimated that disagreement about the future path of inflation tends to:

- Rise with inflation.
- Rise when inflation changes sharply—in either direction.
- Rise in concert with dispersion in rates of inflation across commodity groups.
- Show no clear relationship with measures of real activity.

Finally, we end this section with a note of caution. None of these findings necessarily reflect causality and, in any case, we have deliberately been quite loose in even speaking about the direction of likely causation. However, we believe that these findings present a useful set of stylized facts that a theory of macroeconomic dynamics should aim to explain.

5. Theories of Disagreement

Most theories in macroeconomics have no disagreement among agents. It is assumed that everyone shares the same information and that all are endowed with the same information-processing technology. Consequently, everyone ends up with the same expectations.

A famous exception is the islands model of Robert Lucas (1973). Producers are assumed to live in separate islands and to specialize in producing a single good. The relative price for each good differs by island-specific shocks. At a given point in time, producers can observe the price only on their given islands and from it, they must infer how much of it is idiosyncratic to their product and how much reflects the general price level that is common to all islands. Because agents have different information, they have different forecasts of prices and hence inflation. Since all will inevitably make forecast errors, unanticipated monetary policy affects real output: following a change in the money supply, producers attribute some of the observed change in the price for their product to changes in relative rather than general prices and react by changing production.

This model relies on disagreement among agents and predicts dispersion in inflation expectations, as we observe in the data. Nonetheless, the extent of this disagreement is given exogenously by the parameters of the model. Although the Lucas model has heterogeneity in inflation expectations, the extent of disagreement is constant and unrelated to any macroeconomic variables. It cannot account for the systematic relationship between dispersion of expectations and macroeconomic conditions that we documented in Section 4.

The sticky-information model of Mankiw and Reis (2002) generates disagreement in expectations that is endogenous to the model and correlated with aggregate variables. In this model, the costs of acquiring and processing information and of reoptimizing lead agents to update their information sets and expectations sporadically. Each period, only a fraction of the population update themselves on the current state of the economy and determine their optimal actions, taking into account the likely delay until they revisit their plans. The rest of the population continues to act according to their pre-existing plans based on old information. This theory generates heterogeneity in expectations because different segments of the population will have updated their expectations at different points in time. The evolution of the state of the economy over time will endogenously determine the extent of this disagreement. This disagreement in turn affects agents' actions and the resulting equilibrium evolution of the economy.

We conducted the following experiment to assess whether the sticky-information model can capture the extent of disagreement in the survey data. To generate rational forecasts from the perspective of different points in time, we estimated a vector autoregression (VAR) on U.S. monthly data. The VAR included three variables: monthly inflation (measured by the CPI), the interest rate on three-month Treasury bills, and a measure of the output gap obtained by using the Hodrick-Prescott filter on interpolated quarterly real GDP.[9] The estimation period was from March 1947 to March 2002, and the regressions included 12 lags of each variable. We take this estimated VAR as an approximation to the model rational agents use to form their forecasts.

We follow Mankiw and Reis (2002) and assume that in each period, a fraction λ of the population obtains new information about the state of the economy and recomputes optimal expectations based on this new information. Each person has the same probability of updating their informa-

9. Using employment rather than detrended GDP as the measure of real activity leads to essentially the same results.

tion, regardless of how long it has been since the last update. The VAR is then used to produce estimates of future annual inflation in the United States given information at different points in the past. To each of these forecasts, we attribute a frequency as dictated by the process just described. This generates at each point in time a full cross-sectional distribution of annual inflation expectations. We use the predictions from 1954 onward, discarding the first few years in the sample when there are not enough past observations to produce nondegenerate distributions.

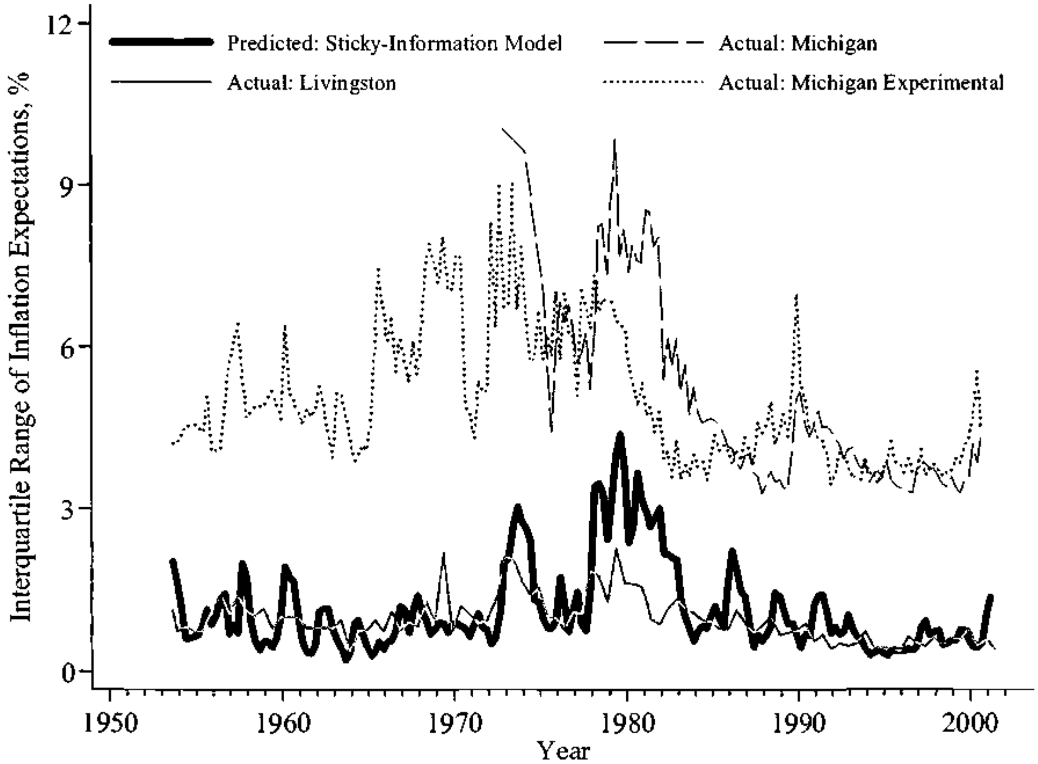
We compare the predicted distribution of inflation expectations by the sticky-information model to the distribution we observe in the survey data. To do so meaningfully, we need a relatively long sample period. This leads us to focus on the Livingston and the Michigan experimental series, which are available for the entire postwar period.

The parameter governing the rate of information updating in the economy, λ, is chosen to maximize the correlation between the interquartile range of inflation expectations in the survey data with that predicted by the model. For the Livingston Survey, the optimal λ is 0.10, implying that the professional economists surveyed are updating their expectations about every 10 months, on average. For the Michigan series, the value of λ that maximizes the correlation between predicted and actual dispersion is 0.08, implying that the general public updates their expectations on average every 12.5 months. These estimates are in line with those obtained by Mankiw and Reis (2003), Carroll (2003a), and Khan and Zhu (2002). These authors employ different identification schemes and estimate that agents update their information sets once a year, on average. Our estimates are also consistent with the reasonable expectation that people in the general public update their information less frequently than professional economists do. It is more surprising that the difference between the two is so small.

A first test of the model is to see to what extent it can predict the dispersion in expectations over time. Figure 10 plots the evolution of the interquartile range predicted by the sticky-information model, given the history of macroeconomic shocks and VAR-type updating, and setting $\lambda = 0.1$. The predicted interquartile range matches the key features of the Livingston data closely, and the two series appear to move closely together. The correlation between them is 0.66. The model is also successful at matching the absolute level of disagreement. While it overpredicts dispersion, it does so only by 0.18 percentage points on average.

The sticky-information model also predicts the time-series movement in disagreement among consumers. The correlation between the predicted and actual series is 0.80 for the actual Michigan data and 0.40 for the longer experimental series. As for the level of dispersion, it is 4 percentage points

Figure 10 ACTUAL AND PREDICTED DISPERSION OF INFLATION EXPECTATIONS



higher on average in the data than predicted by the model. This may be partially accounted for by some measurement error in the construction of the Michigan series. More likely, however, it reflects idiosyncratic heterogeneity in the population that is not captured by the model. Individuals in the public probably differ in their sources of information, in their sophistication in making forecasts, or even in their commitment to truthful reporting in a survey. None of these sources of individual-level variation are captured by the sticky-information model, but they might cause the high levels of disagreement observed in the data.[10]

Section 4 outlined several stylized facts regarding the dispersion of inflation expectations in the survey data. The interquartile range of expected inflation was found to rise with inflation and with the squared change in annual inflation over the last year. The output gap did not seem to affect significantly the dispersion of inflation expectations. We reestimate the regressions in panels A and C of Table 6, now using as the

10. An interesting illustration of this heterogeneity is provided by Bryan and Ventaku (2001), who find that men and women in the Michigan Survey have statistically significant different expectations of inflation. Needless to say, the sticky-information model does not incorporate gender heterogeneity.

Table 7 MODEL-GENERATED DISAGREEMENT AND MACROECONOMIC CONDITIONS[1]

	Multivariate regression	*Bivariate regressions*
Dependent Variable: Interquartile range of model-generated inflation expectations		
Constant	0.005***	
	(0.001)	
Inflation rate	0.127***	0.166***
	(0.028)	(0.027)
ΔInflation-squared	3.581***	6.702***
	(0.928)	(1.389)
Output gap	0.009	0.018
	(0.051)	(0.080)
Adjusted R^2	0.469	
N	579	579

1. *** denotes statistical significance at the 1% level (Newey-West standard errors in parentheses; correcting for autocorrelation up to one year).

dependent variable the dispersion in inflation expectations predicted by the sticky-information model with a λ of 0.1, the value we estimated using the Livingston series.[11] Table 7 presents the results. Comparing Table 7 with Table 6, we see that the dispersion of inflation expectations predicted by the sticky-information model has essentially the same properties as the actual dispersion of expectations we find in the survey data. As is true in survey data, the dispersion in sticky-information expectations is also higher when inflation is high, and it is higher when prices have changed sharply. As with the survey data, the output gap does not have a statistically significant effect on the model-generated dispersion of inflation expectations.[12]

We can also see whether the model is successful at predicting the central tendency of expectations, not just dispersion. Figure 11 plots the median expected inflation, both in the Livingston and Michigan surveys and as predicted by the sticky-information model with $\lambda = 0.1$. The Livingston and predicted series move closely with each other: the correlation is 0.87. The model slightly overpredicts the data between 1955 and

11. Using instead the value of λ that gave the best fit with the Michigan series (0.08) gives similar results.
12. The sticky-information model can also replicate the stylized fact from Section 5 that more disagreement comes with larger relative price dispersion. Indeed, in the sticky-information model, different price-setters choose different prices only insofar as they disagree on their expectations. This is transparent in Ball, Mankiw, and Reis (2003), where it is shown that relative price variability in the sticky-information model is a weighted sum of the squared deviations of the price level from the levels expected at all past dates, with earlier expectations receiving smaller weights. In the context of the experiment in this section, including relative price dispersion as an explanatory variable for the disagreement of inflation expectations would risk confounding consequences of disagreement with its driving forces.

1965, and it underpredicts median expected inflation between 1975 and 1980. On average these two effects cancel out, so that over the whole sample, the model approximately matches the level of expected inflation (it overpredicts it by 0.3%). The correlation coefficient between the predicted and the Michigan experimental series is 0.49, and on average the model matches the level of median inflation expectations, underpredicting it by only 0.5%.

In Section 3, we studied the properties of the median inflation expectations across the different surveys, finding that these data were consistent with weaker but not stronger tests of rationality. Table 8 is the counterpart to Table 4, using as the dependent variable the median expected inflation series generated by the sticky-information model. Again, these results match the data closely. We cannot reject the hypothesis that expectations are unbiased and efficient in the weak sense of panels A and B. Recall that, in the data, we found mixed evidence regarding these tests. Panels C and D suggest that forecasting errors in the sticky-information expectations are persistent and do not fully incorporate macroeconomic data, just as we found to be consistently true in the survey data.

Table 9 offers the counterpart to Table 5, testing whether expectations can be described as purely adaptive. This hypothesis is strongly rejected—sticky-information expectations are much more rational than

Figure 11 ACTUAL AND PREDICTED MEDIAN INFLATION EXPECTATIONS

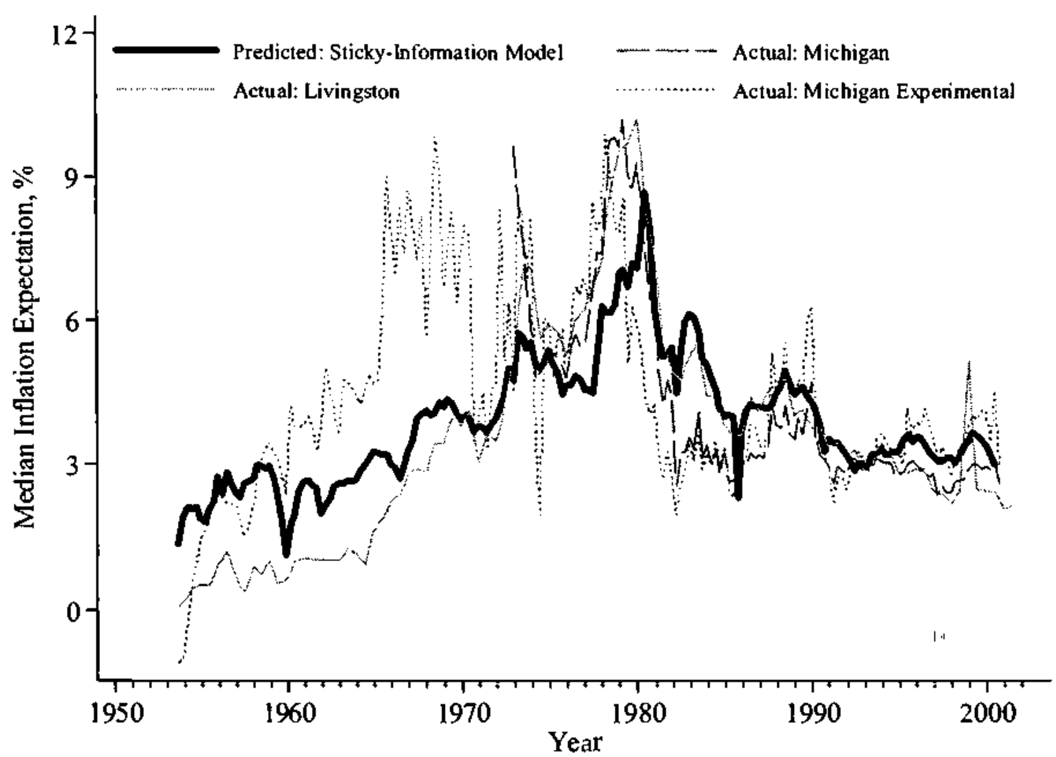


Table 8 TESTS OF FORECAST RATIONALITY: MEDIAN INFLATION EXPECTATIONS PREDICTED BY THE STICKY-INFORMATION MODEL[1]

Panel A: Testing for bias: $\pi_t - E_{t-12}\,\pi_t = \alpha$	
Mean error	0.262%
(Constant only)	(0.310)
Panel B: Is information in the forecast fully exploited? $\pi_t - E_{t-12}\,\pi_t = \alpha + \beta\, E_{t-12}\,\pi_t$	
β: $E_{t-12}[\pi_t]$	0.436*
	(0.261)
α: constant	−1.416%*
	(0.822)
Adj. R^2	0.088
Reject efficiency?	No
$\alpha = \beta = 0$	$p = 0.227$
Panel C: Are forecasting errors persistent? $\pi_t - E_{t-12}\,\pi_t = \alpha + \beta\,(\pi_{t-12} - E_{t-24}\,\pi_{t-12})$	
$\pi_{t-12} - E_{t-24}[\pi_{t-12}]$	0.604***
	(0.124)
Constant	0.107%
	(0.211)
Adj. R^2	0.361
Panel D: Are macroeconomic data fully exploited? $\pi_t - E_{t-12}\,\pi_t = \alpha + \beta\, E_{t-12}[\pi_t] + \gamma\,\pi_{t-13} + \kappa\, i_{t-13} + \delta\, U_{t-13}$	
α: constant	1.567%*
	(0.824)
β: $E_{t-12}[\pi_t]$	0.398
	(0.329)
γ: inflation$_{t-13}$	0.506***
	(0.117)
κ: Treasury bill$_{t-13}$	−0.413**
	(0.139)
δ: unemployment$_{t-13}$	−0.450***
	(0.135)
Reject efficiency?	Yes
$\gamma = \kappa = \delta = 0$	$p = 0.000$
Adjusted R^2	0.369

1. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively (Newey-West standard errors in parentheses; correcting for autocorrelation up to one year).

simple, backward-looking adaptive expectations. Again, this finding matches what we observed in the survey data.

Given how closely the predicted and actual dispersion of expectations and median expected inflation co-move, it is not surprising to find that the results in Tables 4, 5, and 6 are closely matched by the model-generated time series for disagreement in Tables 7, 8, and 9. A stronger test in the tradition of moment-matching is to see whether the sticky-information model can robustly generate the stylized facts we observe in the data. We verify this by implementing the following exercise. Using the residuals from our estimated VAR as an empirical distribution, we randomly draw 720 residual vectors and, using the VAR parameter estimates, use these draws to build hypothetical series for inflation, the output gap, and the Treasury-bill rate. We then employ the sticky-information model to generate a predicted distribution of inflation expectations at each date, using the procedure outlined earlier. To eliminate the influence of initial conditions, we discard the first 10 years of the simulated series so that we are left with 50 years of simulated data. We repeat this procedure 500 times, thereby generating 500 alternative 50-year histories for inflation, the output gap, the Treasury-bill rate, the median expected inflation, and the interquartile range of inflation expectations predicted by the sticky-information model with $\lambda = 0.1$. The regressions in Tables 4, 5, and 6, describing the relationship of disagreement and forecast errors

Table 9 TESTS OF ADAPTIVE EXPECTATIONS: MEDIAN INFLATION EXPECTATIONS PREDICTED BY THE STICKY-INFORMATION MODEL[1]

Adaptive expectations: $E_t\pi_{t+12} = \alpha + \beta(L)\,\pi_t + \gamma\, U_t + \kappa\, U_{t-3} + \delta\, i_t + \phi\, i_{t-3}$	
Inflation	1.182***
$\beta(1)$: sum of 8 coefficients	(0.100)
Unemployment	
γ : date of forecast	−0.561***
	(0.087)
κ : 3 months prior	0.594***
	(0.078)
Treasury bill rate	
δ : date of forecast	0.117***
	(0.026)
ϕ : 3 months prior	0.160***
	(0.027)
Reject adaptive expectations?	Yes
($\gamma = \kappa = \delta = \phi = 0$)	p = 0.000
Adjusted R^2	0.954
N	579

1. *** denotes statistical significance at the 1% level. (Newey-West standard errors in parentheses; correcting for autocorrelation up to a year).

with macroeconomic conditions, are then reestimated on each of these 500 possible histories, generating 500 possible estimates for each parameter.

Table 10 reports the mean parameter estimates from each of these 500 histories. Also shown (in parentheses) are the estimates at the 5th and 95th percentile of this distribution of coefficient estimates. We interpret this range as analogous to a bootstrapped 95% confidence interval (under the null hypothesis that the sticky-information model accurately describes expectations). These results suggest that the sticky-information model robustly generates a positive relationship between the dispersion of inflation expectations and changes in inflation, as we observe in the data. Also, as in the data, the level of the output gap appears to be related only weakly to the dispersion of expectations.

At odds with the facts, the model does not suggest a robust relationship between the level of inflation and the extent of disagreement. To be sure, the relationship suggested in Table 6 does occur in some of these alternative histories, but only in a few. In the sticky-information model, agents disagree in their forecasts of future inflation only to the extent that they have updated their information sets at different points in the past. Given our VAR model of inflation, only changes over time in macroeconomic conditions can generate different inflation expectations by different people. The sticky-information model gives no reason to find a systematic

Table 10 MODEL-GENERATED DISAGREEMENT AND MACROECONOMIC CONDITIONS[1]

	Multivariate regression	Bivariate regressions
(Dependent Variable: Interquartile range of model-generated inflation expectations)		
Constant	1.027***	
	(0.612; 1.508)	
Inflation rate	−0.009	−0.010
	(−0.078; 0.061)	(−0.089; 0.071)
ΔInflation-squared	0.029***	0.030***
	(0.004; 0.058)	(0.005; 0.059)
Output gap	−0.019	−0.023
	(−0.137; 0.108)	(−0.163; 0.116)
Joint test on macro data	Reject at 5% level in 98.2% of histories	
Adjusted R^2	0.162	
N	588	588

1. *** denotes statistical significance at the 1% level. (The 5^{th} and 95^{th} percentile coefficient estimates across 500 alternative histories are shown in parentheses.) Adjusted R^2 refers to the average adjusted R^2 obtained in the 500 different regressions.

relationship between the level of inflation and the extent of disagreement. This does not imply, however, that for a given history of the world such an association could not exist, and for the constellation of shocks actually observed over the past 50 years, this was the case, as can be seen in Table 7. Whether the level of inflation will continue to be related with disagreement is an open question.

Table 11 compares the median of the model-generated inflation expectations series with the artificial series for inflation and the output gap. The results with this simulated data are remarkably similar to those obtained earlier. Panel A shows that expectations are unbiased, although there are many possible histories in which biases (in either direction) of up to one-quarter of a percentage point occur. Panel B shows that sticky-information expectations are typically inefficient, while panel C demonstrates that they induce persistent forecast errors. Panel D shows that sticky-information expectations also fail to exploit available macroeconomic information fully, precisely as we found to be true in the survey data on inflation expectations. The precise relationship between different pieces of macroeconomic data and expectation errors varies significantly across histories, but in nearly all of them there is a strong relationship. Therefore, while the coefficients in Table 11 are not individually significant across histories, within each history a Wald test finds that macroeconomic data are not being fully exploited 78.6% of the time. That is, the set of macro data that sticky-information agents are found to underutilize depends on the particular set of shocks in that history.

Table 12 tests whether sticky-information expectations can be confused for adaptive expectations in the data. The results strongly reject this possibility. Sticky-information expectations are significantly influenced by macroeconomic variables (in this case, the output gap and the Treasury-bill rate), even after controlling for information contained in past rates of inflation.

The sticky-information model does a fairly good job at accounting for the dynamics of inflation expectations that we find in survey data. There is room, however, for improvement. Extensions of the model allowing for more flexible distributions of information arrival hold the promise of an even better fit. An explicit microeconomic foundation for decisionmaking with information-processing costs would likely generate additional sharp predictions to be tested with these data.

6. *A Case Study: The Volcker Disinflation*

In August 1979, Paul Volcker was appointed chairman of the Board of Governors of the Federal Reserve Board, in the midst of an annual inflation

Table 11 TESTS OF FORECAST RATIONALITY: MEDIAN INFLATION EXPECTATIONS PREDICTED BY THE STICKY-INFORMATION MODEL OVER SIMULATED HISTORIES[1]

Panel A: Testing for bias: $\pi_t - E_{t-12}\,\pi_t = \alpha$	
Mean error	0.057%
(Constant only)	(–0.264; 0.369)
Panel B: Is information in the forecast fully exploited? $\pi_t - E_{t-12}\,\pi_t = \alpha + \beta\, E_{t-12}\,\pi_t$	
$\beta : E_{t-12}\,[\pi_t]$	0.308**
	(0.002; 0.6971)
α : constant	–1.018%
	(–2.879; 0.253)
Adjusted R^2	
Reject efficiency? $\alpha = \beta = 0$	Reject at 5% level in 95.4% of histories
Panel C: Are forecasting errors persistent? $\pi_t - E_{t-12}\,\pi_t = \alpha + \beta\,(\pi_{t-12} - E_{t-24}\,\pi_{t-12})$	
$\beta : \pi_{t-12} - E_{t-24}\,[\pi_{t-12}]$	0.260***
	(0.094; 0.396)
α : constant	0.039%
	(–0.237; 0.279)
Adjusted R^2	0.072
Panel D: Are macroeconomic data fully exploited? $\pi_t - E_{t-12}\,\pi_t = \alpha + \beta\, E_{t-12}\,[\pi_t] + \gamma\,\pi_{t-13} + \kappa\, i_{t-13} + \delta\, U_{t-13}$	
α : constant	–0.617%
	(–3.090; 1.085)
$\beta : E_{t-12}\,[\pi_t]$	0.032
	(–0.884; 0.811)
γ : inflation_{t-13}	0.064
	(–0.178; 0.372)
κ : $\text{Treasury bill}_{t-13}$	0.068
	(–0.185; 0.385)
δ : output gap_{t-13}	0.170
	(–0.105; 0.504)
Joint test on macro data ($\gamma = \kappa = \delta = 0$)	Reject at 5% level in 78.6% of histories
Adjusted R^2	0.070
N	569

1. *** and ** denote statistical significance at the 1% and 5% levels, respectively. (The 5th and 95th percentile coefficient estimates across 500 alternative histories are shown in parentheses.) Adjusted R^2 refers to the average adjusted R^2 obtained in the 500 different regressions.

Table 12 TESTS OF ADAPTIVE EXPECTATIONS: MEDIAN INFLATION EXPECTATIONS PREDICTED BY THE STICKY-INFORMATION MODEL OVER SIMULATED HISTORIES[1]

Adaptive expectations: $E_{t-12}\,\pi_t = \alpha + \beta\,(L)\,\pi_t + \gamma\,U_t + \kappa\,U_{t-3} + \delta\,i_t + \phi i_{t-3}$	
Inflation	1.100**
$\beta(1)$: sum of 8 coefficients	(0.177; 2.082)
Output gap	
γ : Date of forecast	0.380**
	(0.064; 0.744)
κ : 3 months prior	−0.300
Treasury bill rate	(−0.775; 0.190)
δ : Date of forecast	0.063
	(−0.042; 0.165)
ϕ : 3 months prior	0.149
	(−0.111; 0.371)
Reject adaptive expectations?	Reject at 5% level
($\gamma = \kappa = \delta = \phi = 0$)	in 100% of histories
Adjusted R^2	0.896
N	569

1. ** denotes statistical significance at the 5% level. (The 5th and 95th percentile coefficient estimates across 500 alternative histories are shown in parentheses.) Adjusted R^2 refers to the average adjusted R^2 obtained in the 500 different regressions.

rate of 11%, one of the highest in the postwar United States. Over the next three years, using contractionary monetary policy, he sharply reduced the inflation rate to 4%. This sudden change in policy and the resulting shock to inflation provides an interesting natural experiment for the study of inflation expectations. The evolution of the distribution of inflation expectations between 1979 and 1982 in the Michigan Survey is plotted in Figure 12.[13] For each quarter there were on average 2,350 observations in the Michigan Survey, and the frequency distributions are estimated nonparametrically using a normal kernel-smoothing function.

Three features of the evolution of the distribution of inflation expectations stand out from Figure 12. First, expectations adjusted slowly to this change in regime. The distribution of expectations shifts leftward only gradually over time in the data. Second, in the process, dispersion increases and the distribution flattens. Third, during the transition, the distribution became approximately bimodal.

We now turn to asking whether the sticky-information model can account for the evolution of the full distribution of expectations observed in the survey data during this period. Figure 13 plots the distribution of

13. The Livingston and SPF surveys have too few observations at any given point in time to generate meaningful frequency distributions.

Figure 12 THE VOLCKER DISINFLATION: THE EVOLUTION OF INFLATION EXPECTATIONS IN THE MICHIGAN SURVEY

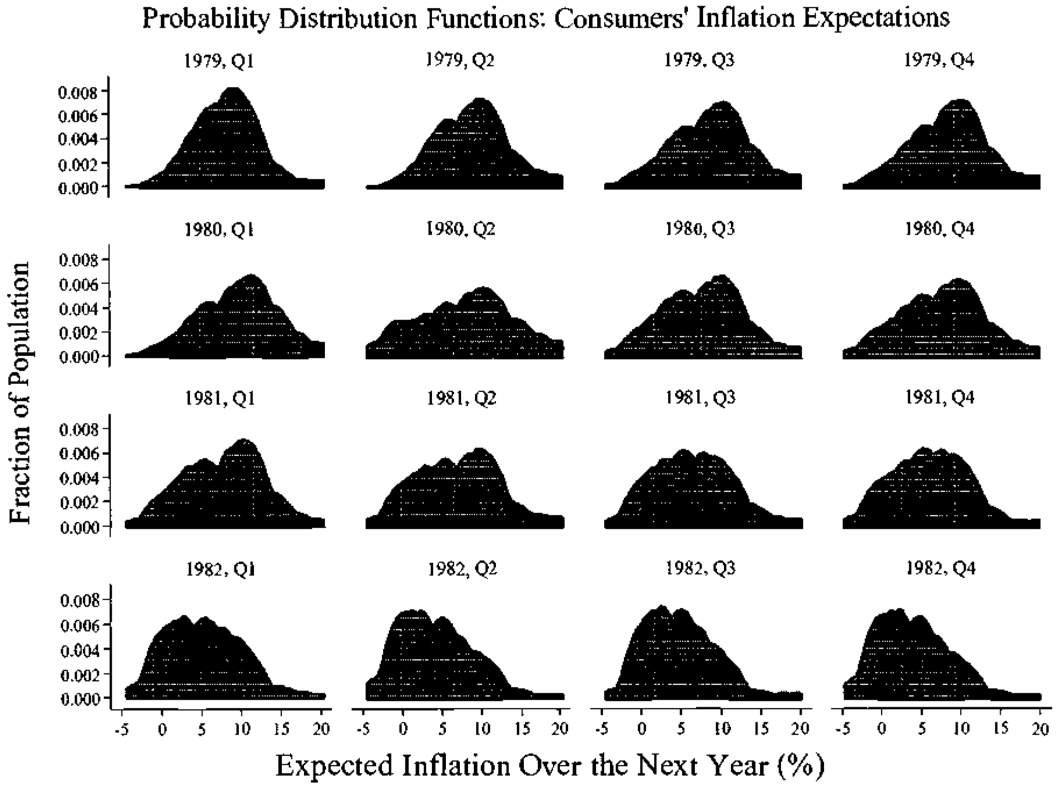
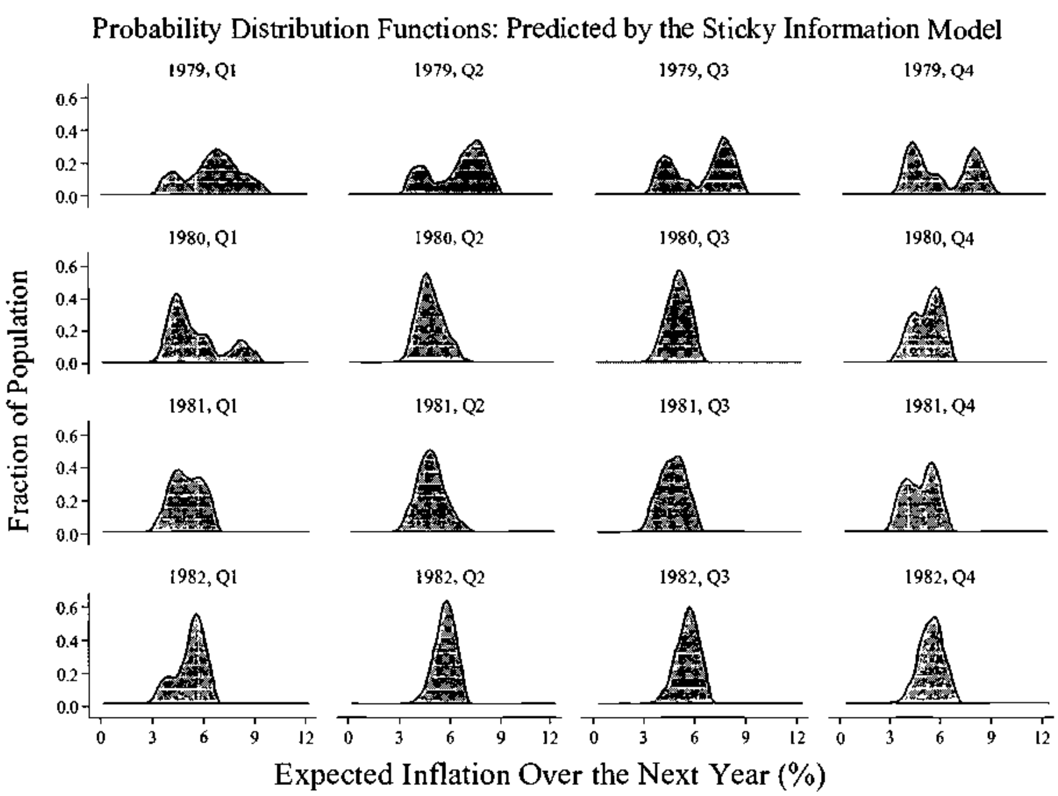


Figure 13 THE VOLCKER DISINFLATION: THE EVOLUTION OF INFLATION EXPECTATIONS PREDICTED BY THE STICKY-INFORMATION MODEL



inflation expectations predicted by the VAR application of the sticky-information model described in Section 5.

In the sticky-information model, information disseminates slowly throughout the economy. As the disinflation begins, a subset of agents who have updated their information sets recently lower their expectation of inflation. As they do so, a mass of the cross-sectional distribution of inflation expectations shifts leftward. As the disinflation proceeds, a larger fraction of the population revises its expectation of the level of inflation downward, and thus a larger mass of the distribution shifts to the left. The distribution therefore flattens and dispersion increases, as we observed in the actual data.

The sudden change in inflation isolates two separate groups in the population. In one group are those who have recently updated their information sets and are now expecting much lower inflation rates. In the other are those holding to pre-Volcker expectations, giving rise to a bimodal distribution of inflation expectations. As more agents become informed, a larger mass of this distribution shifts from around the right peak to around the left peak. Ultimately, the distribution resumes its normal single peaked shape, now concentrated at the low observed inflation rate.

Clearly the sticky-information model generates predictions that are too sharp. Even so, it successfully accounts for the broad features of the evolution of the distribution of inflation expectations during the Volcker disinflation.

7. Conclusion

Regular attendees of the NBER Macroeconomics Annual conference are well aware of one fact: people often disagree with one another. Indeed, disagreement about the state of the field and the most promising avenues for research may be the conference's most reliable feature. Despite the prevalence of disagreement among conference participants, however, disagreement is conspicuously absent in the theories being discussed. In most standard macroeconomic models, people share a common information set and form expectations rationally. There is typically little room for people to disagree.

Our goal in this paper is to suggest that disagreement may be a key to macroeconomic dynamics. We believe we have established three facts about inflation expectations. First, not everyone has the same expectations. The amount of disagreement is substantial. Second, the amount of disagreement varies over time together with other economic aggregates. Third, the sticky-information model, according to which some people form expectations based on outdated information, seems capable of explaining many features of the observed evolution of both the central tendency and the dispersion of inflation expectations over the past 50 years.

We do not mean to suggest that the sticky-information model explored here is the last word in inflation expectations. The model offers a good starting point. It is surely better at explaining the survey data than are the traditional alternatives of adaptive or rational expectations, which give no room for people to disagree. Nonetheless, the model cannot explain all features of the data, such as the positive association between the level of inflation and the extent of disagreement. The broad lesson from this analysis is clear: if we are to understand fully the dynamics of inflation expectations, we need to develop better models of information acquisition and processing. About this, we should all be able to agree.

8. Appendix: An Experimental Series for the Mean and Standard Deviation of Inflation Expectations in the Michigan Survey from 1946 to 2001

The Michigan Survey of Consumer Expectations and Behavior has been run most quarters since 1946, Q1, and monthly since 1978. The current survey questions have been asked continuously since January 1978 (see Curtin, 1996, for details):

Qualitative: *"During the next 12 months, do you think that prices in general will go up, or go down, or stay where they are now?"*
Quantitative: *"By about what percent do you expect prices to go (up/down) on the average, during the next 12 months?"*

For most of the quarterly surveys from June 1966–December 1976, a closed-ended version of the quantitative question was instead asked as:

Closed: *"How large a price increase do you expect? Of course nobody can know for sure, but would you say that a year from now prices will be about 1 or 2% higher, or 5%, or closer to 10% higher than now, or what?"*

Prior to 1966, the survey did not probe quantitative expectations at all, asking only the qualitative question.

Thus, for the full sample period, we have a continuous series of only qualitative expectations. Even the exact coding of this question has varied through time (Juster and Comment, 1978):

- 1948 (Q1)–1952 (Q1): "What do you think will happen to the prices of the things you buy?"
- 1951 (Q4), 1952 (Q2)–1961 (Q1): "What do you expect prices of household items and clothing will do during the next year or so—stay where they are, go up or go down?"

- 1961 (Q2)–1977 (Q2): "Speaking of prices in general, I mean the prices of the things you buy—do you think they will go up in the next year or go down?"
- 1977 (Q3)–present: "During the next 12 months, do you think that prices in general will go up, or go down, or stay where they are now?"

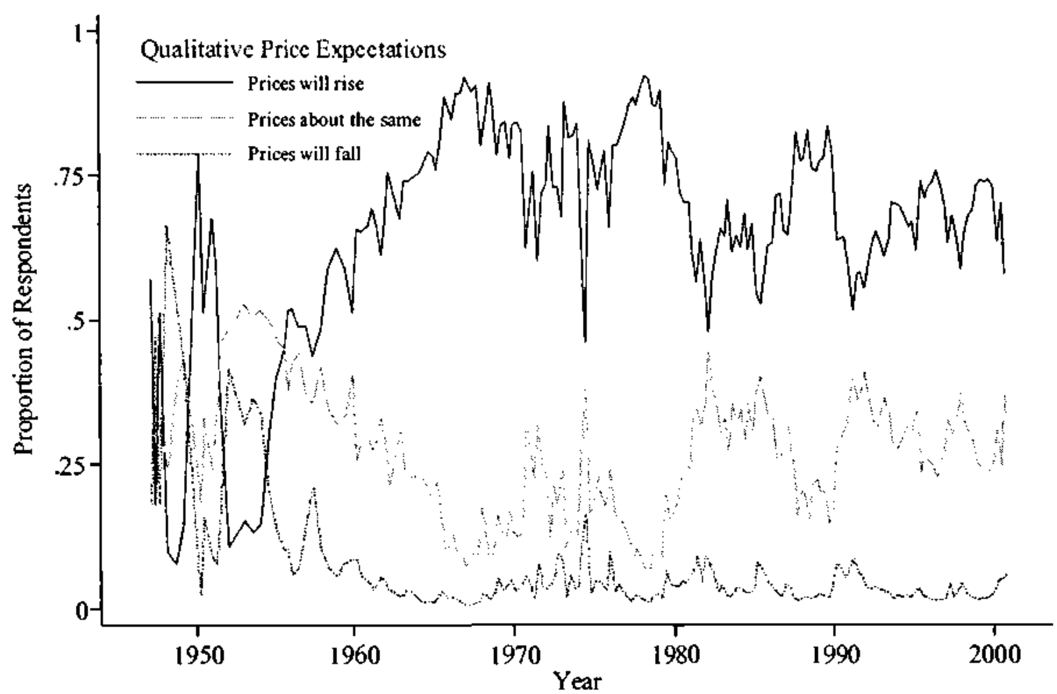
Lacking a better alternative, we proceed by simply assuming that these different question wordings did not affect survey respondents.

We compile raw data for our experimental series from many different sources:

- 1948 (Q1)–1966 (Q1): unpublished tabulations put together by Juster and Comment (1978, Table 1).
- 1966 (Q2)–1977 (Q2): tabulations from Table 2 of Juster and Comment (1978).
- 1967 (Q2), 1977 (Q3)–1977 (Q4): data were extracted from Inter-university Consortium for Political and Social Research (ICPSR) studies #3619, #8726, and #8727, respectively.
- January 1978–August 2001: a large cumulative file containing microdata on all monthly surveys. These data were put together for us by the Survey Research Center at the University of Michigan, although most of these data are also accessible through the ICPSR.

These raw data are shown in Figure 14.

Figure 14 QUALITATIVE RESPONSES TO THE MICHIGAN SURVEY—LONG HISTORY



To build a quantitative experimental series from these qualitative data, we make two assumptions. First, note that a relatively large number of respondents expect no change in prices. We should probably not interpret this literally but rather as revealing that they expect price changes to be small. We assume that when respondents answer that they expect no change in prices, they are stating that they expect price changes to be less than some number, c%. Second, we assume that an individual i's expectation of inflation at time t, π_{it}, is normally distributed with mean μ_t and standard deviation σ_t. Note especially that the mean and standard deviation of inflation expectations are allowed to shift through time, but that the width of the band around zero for which inflation expectations are described as unchanged shows no intertemporal variation (that is, there is no time subscript on c).

Consequently, we can express the observed proportions in each category as a function of the cumulative distribution of the standard normal distribution F_N; the parameter c; and the mean and standard deviation of that month's inflation expectations, μ_t, and σ_t,

$$\%Down_t = F_N\left(\frac{-c-\mu_t}{\sigma_t}\right)$$

$$\%Up_t = 1 - F_N\left(\frac{c-\mu_t}{\sigma_t}\right)$$

Thus, we have two independent data points for each month (*%Same* is perfectly collinear with *%Up+%Down*), and we would like to recover two time-varying parameters. The above two expressions can be solved simultaneously to yield:

$$\mu_t = c\left[\frac{F_N^{-1}(\%Down_t) + F_N^{-1}(1-\%Up_t)}{F_N^{-1}(\%Down_t) - F_N^{-1}(1-\%Up_t)}\right]$$

$$\sigma_t = c\left[\frac{2}{F_N^{-1}(1-\%Up_t) - F_N^{-1}(\%Down_t)}\right]$$

Not surprisingly, we can recover the time series of the mean and standard deviation of inflation expectations up to a multiplicative parameter, c; that is, we can describe the time series of the mean and dispersion of inflation expectations, but the scale is not directly interpretable. To recover a more interpretable scaling, we can either make an ad hoc assumption about the width of the zone from which same responses are drawn, or fit some other feature of the data. We follow the second approach and equate the sample mean of the experimental series and the corresponding quantitative estimates of median inflation expectations from the same survey over the shorter 1978–2001 period when both quantitative and qualitative data are available. (We denote the median inflation expectation by $\tilde{\pi}$.)[14] formally, this can be stated:

$$\sum_{t}^{1978-2001} \mu_t = \sum_{t}^{1978-2001} \tilde{\pi} \quad \text{which solves to yield:}$$

$$c = \frac{\sum_{t}^{1978-2001} \tilde{\pi}}{\sum_{t}^{1978-2001} \frac{F_N^{-1}(\%Down_t) + F_N^{-1}(\%1 - Up_t)}{F_N^{-1}(\%Down_t) - F_N^{-1}(\%1 - Up_t)}}$$

This assumption yields an estimate of $c = 1.7\%$. That is, the specific scaling adopted yields the intuitively plausible estimate that those expecting inflation between −1.7% and +1.7% respond that prices will stay where they are now. More to the point, this specific scaling assumption is not crucial to any of our regression estimates. It affects the interpretation of the magnitude of coefficients but not the statistical significance.

Thus, for our sample of T periods, with $2T + 1$ parameters and $2T + 1$ unknowns, we can estimate the time series of the mean and standard deviation of inflation expectations. As a final step, we rely on the assumption of normality to convert our estimate of the sample standard deviation into an estimate of the interquartile range.

Figures 1 and 3 show that the median and interquartile range of the constructed series move quite closely with the quantitative estimates over the period from 1978. Table 2 reports on the correlation of this series with other estimates.

REFERENCES

Ball, Laurence. (2000). Near-rationality and inflation in two monetary regimes. Cambridge, MA: National Bureau of Economic Research. NBER Working Paper 7988.

Ball, Laurence, and Dean Croushore. (2003). Expectations and the effects of monetary policy. *Journal of Money, Credit and Banking* 35(4): 473–484.

Ball, Laurence, N. Gregory Mankiw, and Ricardo Reis. (2003). Monetary policy for inattentive economies, *Journal of Monetary Economics*, forthcoming.

Barnea Amir, Amihud Dotan, and Josef Lakonishok. (1979). The effect of price level uncertainty on the determination of nominal interest rates: Some empirical evidence. *Southern Economic Journal* 46(2): 609–614.

Bomberger, William, and William Frazer. (1981). Interest rates, uncertainty and the Livingston data. *Journal of Finance* 36(3): 661–675.

14. It is just as valid to refer to the mean of this experimental series as the median expectation, given the assumption of normality.

Brenner, Menachem, and Yoram Landskroner. (1983). Inflation uncertainties and returns on bonds. *Economica* 50(200):463–468.

Bryan, Michael, and Stephen Cecchetti. (1994). Measuring core inflation. In *Monetary Policy*, N. Gregory Mankiw (ed.). Chicago: University of Chicago Press for NBER.

Bryan, Michael, and Guhan Venkatu. (2001). The curiously different inflation perspectives of men and women. Federal Reserve Bank of Cleveland Economic Commentary Series. Available at http://www.clev.frb.org/Research/Com2001/1015.pdf.

Carroll, Christopher. (2003a). Macroeconomic expectations of households and professional forecasters. *Quarterly Journal of Economics* 118(1):269–298.

Carroll, Christopher. (2003b). The epidemiology of macroeconomic expectations. In *The Economy as an Evolving Complex System, III*, Larry Blume and Steven Durlauf (eds.). Oxford, England: Oxford University Press.

Croushore, Dean. (1993). Introducing: The Survey of Professional Forecasters. Federal Reserve Bank of Philadelphia. *Business Review* November/December: 3–15.

Croushore, Dean. (1997). The Livingston Survey: Still useful after all these years. Federal Reserve Bank of Philadelphia. *Business Review* March/April: 15–27.

Cukierman, Alex, and Paul Wachtel. (1979). Differential inflationary expectations and the variability of the rate of inflation: Theory and evidence. *American Economic Review* 69(4):595–609.

Curtin, Richard. (1996). Procedure to estimate price expectations. University of Michigan Survey Research Center. Mimeo.

Friedman, Milton. (1968). The role of monetary policy. *American Economic Review* 58(1):1–17.

Gavin, William T. (2003). FOMC forecasts: Is all the information in the central tendency? Federal Reserve Bank of St Louis. Working Paper 2003-002A.

Juster, F. Thomas, and Robert Comment. (1978). A note on the measurement of price expectations. Institute for Social Research, University of Michigan. Unpublished manuscript.

Khan, Hashmat, and Zhenhua Zhu. (2002). Estimates of the sticky-information Phillips curve for the United States, Canada, and the United Kingdom. Bank of Canada Working Paper 2002–19.

Levi, Maurice, and John Makin. (1979). Fisher, Phillips, Friedman and the measured impact of inflation on interest. *Journal of Finance* 34(1):35–52.

Lambros, Louis, and Victor Zarnowitz. (1987). Consensus and uncertainty in economic prediction. *Journal of Political Economy* 95(3):591–621.

Lucas, Robert E., Jr. (1973). Some international evidence on inflation-output tradeoffs. *American Economic Review* 63:326–334.

Makin, John. (1982). Anticipated money, inflation uncertainty and real economic activity. *Review of Economics and Statistics* 64(1):126–134.

Makin, John. (1983). Real interest, money surprises, anticipated inflation and fiscal deficits. *Review of Economics and Statistics* 65(3):374–384.

Mankiw, N. Gregory, and Ricardo Reis. (2002). Sticky information versus sticky prices: A proposal to replace the new Keynesian Phillips curve. *Quarterly Journal of Economics* 117(4):1295–1328.

Mankiw, N. Gregory, and Ricardo Reis. (2003). Sticky information: A model of monetary non-neutrality and structural slumps. In *Knowledge, Information and Expectations in Modern Macroeconomics: In Honor of Edmund S. Phelps*, P. Aghion, R. Frydman, J. Stiglitz, and M. Woodford (eds.). Princeton, NJ: Princeton University Press.

Mullineaux, Donald. (1980). Inflation expectations and money growth in the United States. *American Economic Review* 70(1):149–161.
Souleles, Nicholas. (2001). Consumer sentiment: Its rationality and usefulness in forecasting expenditure—Evidence from the Michigan micro data. *Journal of Money, Credit and Banking* forthcoming.
Thomas, Lloyd, Jr. (1999). Survey measures of expected U.S. inflation. *Journal of Economic Perspectives* 13(4):125–144.
Vissing-Jorgenson, Annette. (2003). Perspectives on behavioral finance: Does "irrationality" disappear with wealth? Evidence from expectations and actions. In *NBER Macroeconomics Annual 2003*, Mark Gertler and Kenneth Rogoff (eds.). Cambridge, MA: MIT Press.

Comment

ROBERT G. KING*
Boston University, NBER, and Federal Reserve Bank of Richmond

1. *Expectations and Macroeconomics*

Disagreement about inflation expectations, particularly controversies over the importance of these expectations for the relationship between real and nominal variables, has been a central topic in macroeconomics during the last three decades. Most analyses have taken inflation expectations—and other expectations about macroeconomic variables—as identical across agents, or they at least have taken the view that cross-sectional differences in beliefs are second-order for macroeconomic phenomena.

The view that average expectations are sufficient for most macroeconomic purposes is present in many diverse lines of research. In the early studies of Gordon (1970) and Solow (1969), inflation expectations were viewed as adaptive, but differences across agents in the speed of expectation adjustment were not stressed. Instead, this viewpoint was made operational by using simple distributed lag specifications as proxies for expectations, making beliefs about inflation depend only on a subset of available data despite the fact that it was generally more complicated in macroeconomic models. Famously criticized by Lucas (1972) and Sargent (1971), who employed rational expectations models with homogenous beliefs in their arguments, the adaptive expectations viewpoint has largely been replaced by rational expectations modeling. Following Lucas and Sargent, the specific form of rational expectations employed most frequently is that all information is common to agents.

* The views expressed in this comment are not necessarily these of The Federal Reserve Bank of Richmond or of The Federal Reserve System.

While the homogeneous expectations model has been dominant, it is important to remember that an earlier line of flexible price macroeconomic research during this period sought to use limited information constructs—the imperfect information models developed by Lucas (1973), Barro (1976), and others during the late 1970s and early 1980s—to rationalize monetary nonneutrality. In contrast to macroeconomic models incorporating rational expectations with common information, these setups featured incomplete adjustment of average beliefs precisely because individuals had limited and disparate information sets. The profession ultimately turned away from these models, however, for two reasons. First, their implications were fragile with respect to the specification of the nature and evolution of information sets. Second, it was difficult to believe that they explained the apparent nonneutrality of money stock measures in an economy like that of the United States, with readily available monetary statistics.

1.1 (MY) EXPECTATIONS

Recently, Mankiw and Reis (2002) have resurrected the idea that limited information—particularly infrequent adjustment of expectations—is important for the interplay of real and nominal variables. They do so within models in which firms are price-setters rather than participants in competitive markets like those envisioned by Lucas and Barro.

Because I cut my research teeth on the earlier generation of imperfect information models, I was delighted when Mark Gertler asked me to discuss a prospective paper in which these authors would explore the evidence for sticky expectations. I thought that it would be an excellent opportunity to think further about an important topic: what macroeconomic and microeconomic implications most sharply distinguish the sticky-expectations model of Mankiw and Reis (MR) from the popular sticky-price model that has been much employed in recent macroeconomic research. So I was excited to have the opportunity.

1.2 EXPECTATIONS ARE NOT ALWAYS FULFILLED

Conference organizers suggest topics to authors and sometimes get papers that are very different from those expected. Karl Brunner once asked Robert Lucas to write a survey of empirical evidence on the Phillips curve and got "Econometric Policy Evaluation: A Critique." My expectations were not fulfilled with this paper, but I am not disappointed. The Mankiw-Reis-Wolfers (MRW) paper is a fascinating description of various measures of survey inflation expectations. It documents how these measures vary through time; how they are related to the level of inflation; how they move over the course of business cycles; and how they evolved during an important episode, the Volcker deflation. It is sure to stimulate much interesting future research.

1.3 LINK TO THE STICKY-EXPECTATIONS MODEL

The sticky-expectations model of MR implies that macroeconomic shocks—particularly monetary policy shocks—have real effects because some agents adjust expectations and others don't. It also has the effect that monetary shocks cause dispersion in inflation expectations because some agents adjust their forecasts immediately when a shock occurs and others do so only gradually. This implication motivates the current paper: MRW want to find out whether there are important changes over time in the cross-sectional variability of expectations.

Now, there are other empirical implications of the sticky-expectations model that one might want to explore, both at the micro and macro levels. In the MR model, when a firm gets an opportunity to update its information, it chooses an entire *path* for future nominal prices that it will charge until its next information update.[1] To me, this *micro implication* flies in the face of one of the central facts that new Keynesian macroeconomics has long stressed, which is the tendency for many nominal prices to stay constant for substantial periods of time.[2]

And there are also *macro implications* of this adjustment pattern. Ball, Mankiw, and Romer (1988) use data on a large number of different countries to argue for sticky-price models rather than the alternative information-confusion model of Lucas. They argue that sticky-price models imply that the output-inflation trade-off should depend negatively on the average rate of inflation because high rates of inflation would induce firms to undertake more frequent adjustments. They argue that cross-country evidence strongly supports this implication, rather than Lucas's implication that the slope should depend on variability of inflation. Now, because a Mankiw-Reis firm sets a path of prices, it can neutralize the effects of the average inflation rate on its real revenues. Its incentives for frequency of information adjustment would therefore be unaffected by average inflation, just like Lucas's flexible price firm, because the MR firm's price is flexible with respect to forecasted inflation.

1. If a firm chose a nominal price that would be held fixed until the next receipt of information, then the MR model has the attractive characteristic that it simply collapses to the well-known Calvo (1983) model. Thus, the essential feature of the model is that the firm chooses a price plan rather than a price.
2. A notable and important exception is the practices of selective discounts, such as sales.

Each of these two implications of the MR model seems inconsistent with key facts long stressed by Keynesian macroeconomics, old and new. So if the MR model is the principal motivation for the MRW investigation, then the link is less than fully satisfactory.

1.4 INTELLECTUAL CURRENTS

There are other reasons for studying disagreement about inflation expectations. Stepping back, the MRW paper is part of recent work that is sometimes called behavioral macroeconomics and at other times called macroeconomics and individual decisionmaking. This work aims at (1) taking a careful look at how individuals actually make decisions at the individual level, and (2) developing hypotheses about behavior that are well-specified alternatives to those explored in earlier neoclassical studies of micro data. Work on behavioral macroeconomics in general and the MRW paper in particular is thus a timely and welcome contribution. But one must also bear in mind that survey reports of expectations are not quite the sort of behavior about which most economists—neoclassical or not—are prone to theorize about. They are not market actions, just statements.

1.5 REMEMBERING MUTH

Even if we take these measures as accurate indicators of individual expectations, it is important to remember Muth's (1961) original description of rational expectations. He noted that:

- "The hypothesis (is) . . . that expectations of firms . . . tend to be distributed, for the same information set, about the prediction of the theory . . . ."
- "The hypothesis asserts three things: (i) that information is scarce and the economic system generally does not waste it; (ii) the way that expectations are formed depends specifically on the structure . . . of the economy; (and) (iii) a 'public prediction' will have no effect . . . (unless it is based on inside information)."
- "It does *not assert* that the scratch work of entrepreneurs resembles the system of equations in any way; nor does it state that the predictions of entrepreneurs are perfect or that their expectations are all the same."

So Muth was comfortable with deviations of individual expectations from average, potentially of a systematic, type. He nevertheless chose to construct an economic model in which only average expectations mattered and to explore the implications of this model for the dynamics of agricultural prices.

1.6 CRITICAL QUESTION

As macroeconomists, we know that there are lots of types of heterogeneity in the world. We abstract from many in building our models, as Muth did. The key to successful macro model building is to put in heterogeneity that is important for the issue at hand and to leave out the rest. For example, in studying capital formation, one might think that it is important to take careful account of the age distribution of capital stocks because this distribution could aid in predicting the timing of firms' upgrades and replacements. One would like this age distribution of capital stocks to reflect underlying costs, presumably of a fixed sort, that keep firms from rapidly adjusting their capital stocks.

More specifically, one might think—as I did—that lumpy investment at the micro level would produce an important set of distributed lag effects on aggregate investment not present in standard neoclassical models. But the general equilibrium analysis of Thomas (2002) shows that this need not be the case: one carefully constructed model with a rich age distribution of capital stocks does not produce very different investment behavior than the simplest neoclassical model. So this form of heterogeneity did not turn out to be important in one particular setting.

By contrast, modern sticky-price models take heterogeneity in nominal prices, a comparable age distribution of prices, as a first-order phenomenon. Many such models produce *very* different real responses from those in flexible price models without a distribution of prices. While most of these studies impose a time-dependent pattern of price adjustment, the nonneutrality results of some sticky-price models survive the introduction of state dependent pricing.

So the critical question becomes, Is heterogeneity in beliefs important for macroeconomic models of the Phillips curve?

2. *Expectations, Credibility, and Disinflation*

To make the question asked above concrete within a particular model, I now consider a stylized model of the Volcker deflation, stimulated by the Mankiw-Reis-Wolfers discussion of this topic in Section 6 of their paper.

2.1 A STICKY-PRICE MODEL

For this purpose, I use a simple macroeconomic model consisting of an inflation equation for the private sector and a monetary policy rule that involves a specification of an inflation path. The examples are simplifications of the analyses of Ball (1994, 1995).

2.1.1 *Private Behavior* The private sector inflation equation is:

$$\pi_t = E_t \pi_{t+1} + \phi y_t$$

where π_t is the inflation rate, $E_t\pi_{t+1}$ is the expectation of future inflation, y_t is a measure of the output gap, and ϕ is a slope coefficient that captures the structural effect that the output gap has on the inflation rate at a given expected future inflation rate.[3] As is well known, this specification—sometimes called the new Keynesian Phillips curve—can be derived from underlying microeconomic foundations with a stochastic price adjustment mechanism of the Calvo (1983) form.

2.1.2 *A Policy of Gradual Disinflation* I also assume that the monetary authority takes whatever monetary actions are necessary to produce a gradual disinflation path:

$$\bar{\pi}_t = \begin{Bmatrix} \pi^h - gt \text{ for } t = 1,2,\ldots T \\ \pi^l \text{ for } t > T \end{Bmatrix}$$

with $g = (\pi^h - \pi^l)/T > 0$. This rule specifies that inflation gradually declines from the high level π^h to the low level π^l over the course of T periods, with an identical change in the inflation rate taking place in each period.

2.1.3 *Imperfect Credibility and Expected Inflation* I finally specify a sense in which the representative agent sees the policy as imperfectly credible. In each period, there is a probability α_t that the disinflation will be continued next period. If the disinflation is terminated, then inflation will return to the high level π^h and will stay there in all future periods. With this specification, expected future inflation takes the following form:

$$\begin{aligned} E_t \pi_{t+1} &= \alpha_t \bar{\pi}_{t+1} + (1 - \alpha_t)\pi^h \\ &= \alpha_t(\bar{\pi}_t - g) + (1 - \alpha_t)\pi^h \end{aligned}$$

Within a successful deflation, the path of the output gap is therefore:

$$\begin{aligned} y_t &= \frac{1}{\phi}[\pi_t - E_t \pi_{t+1}] \\ &= \frac{1}{\phi}[\alpha_t g + (1 - \alpha_t)(\bar{\pi}_t - \pi^h)] \end{aligned}$$

3. A coefficient β is sometimes inserted before expected future inflation. Because it is a quarterly discount factor, its value is just below 1 and it is omitted for simplicity in the discussion below.

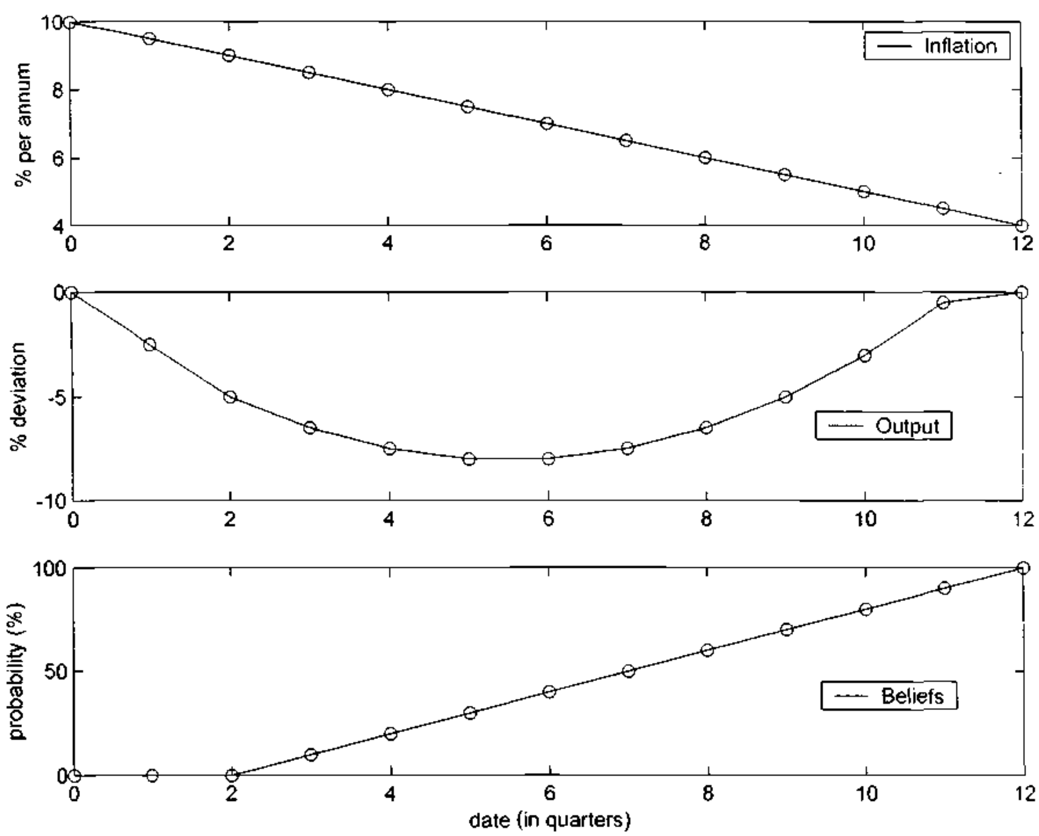
This expression captures two effects familiar from Ball's work on sticky prices and deflation. First, a perfectly credible ($\alpha_t = 1$) deflation produces a boom in output. Second, if inflation is reduced in an imperfectly credible manner, then a recession will occur. For example, if the disinflation is completely incredible ($\alpha_t = 0$), then output declines in lock-step with inflation. It also highlights that the behavior of output depends importantly on the dynamics of beliefs that are here taken to be common across all agents, i.e., on the behavior of the α_t over time.

2.2 DYNAMICS OF A SUCCESSFUL DEFLATION

I now use this simple model to analyze an example of dynamics within a successful disinflation, which is assumed to take three years to complete and to reduce the inflation rate from 10% per year to 4% per year. In particular, I suppose that the beliefs about the disinflation are initially stubborn, with $\alpha_t = 0$ for three quarters, and then gradually rise until the disinflation is fully credible at its endpoint.

The particular assumptions and their implications for output are displayed in Figure 1. Inflation is assumed to decline gradually, as displayed in the top panel. The credibility of the disinflation rises through time.

Figure 1 DYNAMICS OF AN IMPERFECTLY CREDIBLE BUT ULTIMATELY SUCCESSFUL DEFLATION



Output is initially not much affected but then declines because the disinflation is incredible.[4] As its credibility rises, the disinflation's real consequences evaporate.

2.3 REINTERPRETING THE MODEL

To this point, we have assumed that there is no disagreement about inflation expectations in the sense of Mankiw, Reis, and Wolfers. Suppose, however, that we now assume that there is such disagreement. A fraction α_t of the population is optimistic, believing that the disinflation will continue, while the remaining population members are pessimistic. Under this alternative interpretation, the dynamics of inflation and output are unaffected, but there would be variability in measures of disagreement similar to those considered in this paper: disagreement would be small at the beginning and end of the disinflation, while it would be higher in the middle. So, in this model, disagreement about inflation expectations can occur and evolve over time. But modeling these disagreements does not seem essential to understanding the episode.

2.4 CONNECTING WITH THE ACTUAL DISINFLATION EXPERIENCE

I think that the actual disinflation experience involved the following four features:

First, it was *widely discussed*: it is hard to imagine that agents didn't know that something was up.[5]

Second, it was *widely debated* on two dimensions. People disagreed about whether it would work and whether it was a good idea. The former suggests disagreement about expectations.

Third, there was some *uncertainty* about what was going on: people were not sure what the Federal Reserve was up to in terms of its long-range objectives for inflation.

Fourth, it was *imperfectly credible*. As Shapiro (1994) notes, the Volcker deflation is very different from the prior disinflation attempts by the

4. As the reader will note, the scale of the output effect depends entirely on the choice of the parameter ϕ. In drawing the graph, I chose a $\phi = .2$, which meant that a maximum output decline of about 2.5% occurred, although no vertical scale is included in the diagram. A choice of $\phi = .05$ would have alternatively brought about a maximum 10% decline in output. In models that derive ϕ from underlying micro structure, it is related to two deeper parameters: the effect of real marginal cost on inflation (which Gali and Gertler [1999] estimate to be about .05) and the elasticity of real marginal cost to the output gap. Dotsey and King (2001) discuss how some structural features of the underlying economy affect this latter feature. Values of this elasticity much less than 1 arise from models with elastic labor supply, variable capacity utilization, and intermediate inputs. Hence, small values of ϕ and large output effects are not hard to generate from modern sticky-price models.

5. This restates a common criticism of the Barro-Lucas–type incomplete information models, which my then-colleague Stan Engerman once summarized as "Don't the people in your economies have telephones?" and Ed Prescott later put as "People read newspapers."

Federal Reserve. Within a few years after each of the four prior episodes, inflation was reduced only temporarily and then returned to an even higher level within a few years.

During the Volcker deflation, long-term interest rates stayed high for a long time, much longer than any modern pricing model—including that of Mankiw and Reis—would predict, if there was not imperfect credibility about long-term inflation. Unraveling the nature of this episode is an important topic for research in monetary economics, but I am not convinced that understanding the dynamics of measures of disagreement about expectations is important for understanding the episode.

2.5 IMPERFECT CREDIBILITY VERSUS STICKY EXPECTATIONS

In terms of practical macroeconomics, one might ask whether I have drawn a distinction without a difference in my discussion. I don't think so. Sticky expectations are a structural feature of price dynamics for Mankiw and Reis and describe both normal situations and unusual events. Imperfect credibility is a feature of the macroeconomy and monetary policy, and is likely to be more important in some situations than others. So, the inflation-output trade-off during the Volcker deflation might give a poor guide to the nature of that trade-off in a current monetary policy context.

REFERENCES

Ball, Laurence. (1994). Credible disinflation with staggered price-setting. *American Economic Review* 84(1): 282–289.

Ball, Laurence. (1995). Disinflation with imperfect credibility. *Journal of Monetary Economics* 35(1): 5–23.

Ball, Laurence, N. Gregory Mankiw, and David Romer. (1988). The new Keynesian economics and the output-inflation trade-off. *Brookings Papers on Economic Activity* 1988(1): 1–82.

Barro, Robert J. (1976). Rational expectations and the role of monetary policy. *Journal of Monetary Economics* 2(1): 1–32.

Calvo, Guillermo. (1983). Staggered prices in a utility-maximizing framework. *Journal of Monetary Economics* 12(3): 383–398.

Dotsey, Michael, and Robert G. King. (2001). Production, pricing and persistence. Cambridge, MA: National Bureau of Economic Research. NBER Working Paper No. 8407.

Gali, Jordi, and Mark Gertler. (1999). Inflation dynamics: A structural econometric analysis. *Journal of Monetary Economics* 44(2): 195–222.

Gordon, Robert J. (1970). The recent acceleration of inflation and its lessons for the future. *Brookings Papers on Macroeconomics* 1:8–41.

Lucas, Robert E., Jr. (1972). Econometric testing of the natural rate hypothesis. In *The Econometrics of Price Determination*, Otto Eckstein (ed.). Washington, DC: Board of Governors of the Federal Reserve System.

Lucas, Robert E., Jr. (1973). Some international evidence on output inflation trade-offs. *The American Economic Review* 63(3): 326–334.

Mankiw, N. Gregory, and Ricardo Reis. (2002). Sticky information versus sticky prices: A proposal to replace the new Keynesian Phillips curve. *Quarterly Journal of Economics* 117(4): 1295–1328.

Muth, John F. (1961). Rational expectations and the theory of price movements. *Econometrica* 29(3):315–335.

Sargent, Thomas J. (1971). A note on the accelerationist controversy. *Journal of Money, Credit and Banking* 3(August):50–60.

Shapiro, Matthew D. (1994). Federal Reserve policy: Cause and effect. In *Monetary Policy*, N. G. Mankiw (ed.). Cambridge, MA: MIT Press.

Solow, Robert M. (1969). *Price Expectations and the Behavior of the Price Level.* Mancheser, UK: Manchester University Press.

Thomas, Julia K. (2002). Is lumpy investment relevant for the business cycle? *Journal of Political Economy* 110:508–534.

Comment

JOHN C. WILLIAMS
Federal Reserve Bank of San Francisco

This is an excellent paper that uncovers several novel and fascinating "stylized facts" about cross-sectional dispersion of inflation expectations based on surveys of households and economists in the United States. Of particular interest is the finding that the degree of dispersion is positively correlated with both the inflation rate and dispersion in relative prices. In addition, the authors propose a theory that can account for many of the empirical time-series regularities related to both the median and dispersion in surveys of inflation expectations. Given the central role of expectations in modern macroeconomic theory, this paper is certain to stimulate a wide range of theoretical and empirical research aimed at understanding heterogeneous expectations and their implications for the behavior of the economy. A question of immediate interest is whether the stylized facts regarding inflation expectations in the United States also describe expectations of other key macroeconomic variables, such as gross domestic product (GDP) and interest rates, and expectations data in other countries.

Throughout my discussion I will follow the authors' lead and treat surveys as representing reasonably accurate measures of agents' expectations. But it is worth keeping in mind that expectations derived from financial market data can differ in important ways from expectations taken from surveys. For example, as noted in the paper, Ball and Croushore (2001) document the insensitivity of the median value from surveys of inflation expectations to economic news. In contrast,

I would like to thank Kirk Moore for excellent research assistance

Gürkaynak et al. (2003) find that forward nominal interest rates are highly sensitive to economic news, and they provide evidence that this sensitivity is primarily related to the inflation component of interest rates. Because financial market participants "put their money where there mouths are," one is tempted to put greater faith in estimates taken from financial market data, even while recognizing the difficult measurement problems associated with extracting expectations from these data. Still, additional study and comparison of both sources of expectations data is needed to form a more complete picture of the properties of expectations.

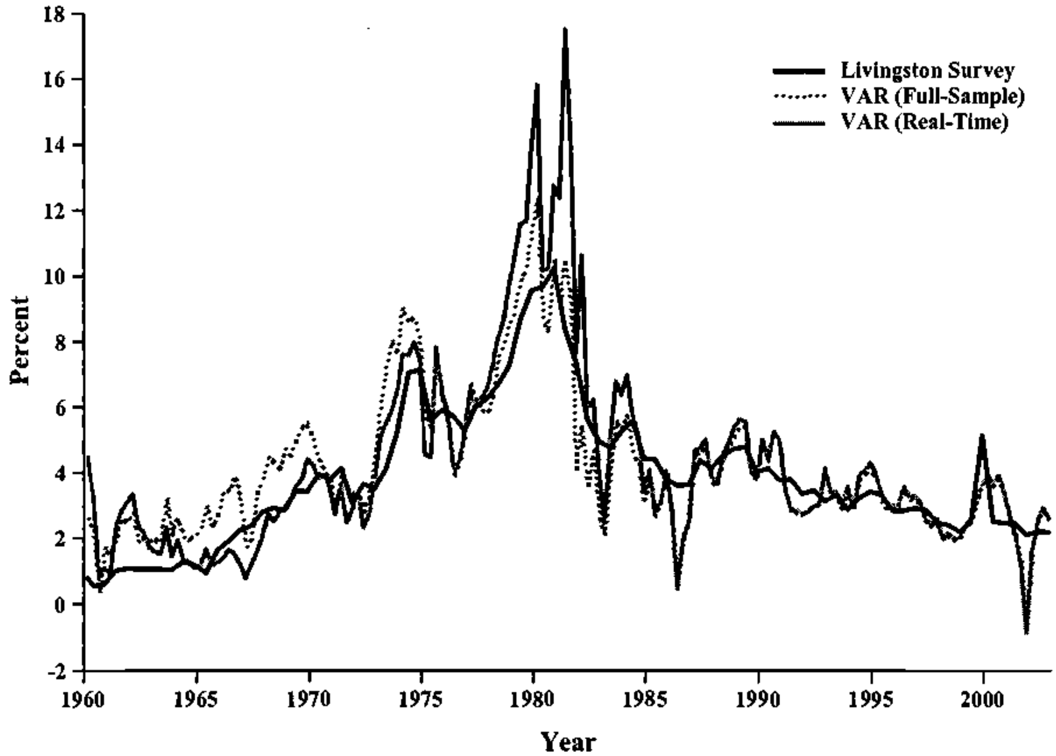
The remainder of my discussion will focus on two topics: learning and model uncertainty. The first relates to is the real-time information that forecasters are assumed to possess. The second provides an alternative explanation of the evidence on dispersion in forecasters' inflation expectations based on the notion that there exists a range of competing forecast models. I find that model uncertainty provides an intuitively more appealing description of the form of disagreement among economists than that proposed in the paper.

The authors argue that the Mankiw and Reis (2002) sticky-information model can explain many of the properties of the median values and dispersion of surveys of inflation expectations. In this model, agents use a three-variable, 12-lag monthly vector auto regression (VAR), that includes inflation, the output gap, and the three-month T-bill rate, estimated over the entire postwar sample to generate forecasts.[1] Individual agents, however, update their expectations only at random intervals, with a 10% probability of an update in each month. Given this structure, the resulting median sticky-information forecast is closely related to the median of a geometrically weighted average of past inflation forecasts. The cross-sectional dispersion in forecasts reflects the dispersion of forecasts across vintages. In fact, there is absolutely no disagreement about forecasting methods; all the differences arise from the differences across vintages of forecasts that people are assumed to use.

I find the assumption that households and economists had access to the full-sample VAR estimates to be unrealistic: people in 1960 simply did not possess the knowledge of forecast model specification and

1. The use of the output gap in the forecasting VAR is problematic because of well-documented problems with real-time estimates of GDP and potential (or the natural rate of) output, issues emphasized by Orphanides (2001), Orphanides and van Norden (2002), and others. A preferable approach would be to use the unemployment rate or the real-time estimates of capacity utilization, which is the approach I follow in the model-based exercises reported in this discussion. As noted by the authors, their results are not sensitive to the use of the output gap in this application, so this criticism is intended more as a general warning.

Figure 1 IN SAMPLE VERSUS OUT-OF-SAMPLE FORECASTS



parameter estimates that we have accumulated over the subsequent 40 or more years. Instead, they needed to estimate and test models based on the limited data that they had on hand and that ran only back to the late 1940s.[2] By constructing the sticky-information model forecasts from a VAR estimated over the full sample, the authors are giving agents far more information, especially during the earlier part of the sample, than they had. A more palatable procedure would be to assume that agents estimate their models in real time, using the vintage of data available to them, and form forecasts accordingly.

VAR forecasts using the data available at the time track the Livingston Survey pretty closely through the 1960s and 1970s and do a much better job at this over that period than forecasts based on a VAR estimated over the full sample. The shaded gray line in Figure 1 shows the forecasts of consumer price index (CPI) inflation over the next four quarters from a real-time, three-variable VAR estimated over data from 1950 through

2. Data from the period of World War II was arguably of limited value for forecasting inflation because of the stringent price controls in place during the war.

the current quarter.[3] The dashed black line reports the corresponding forecasts from a VAR estimated over the full sample. For comparison, the solid black line shows the Livingston Survey of expected price increases over the next 12 months.

The forecasts of the real-time and full-sample VARs are nearly identical from the mid-1980s on, but in the earlier period they display sizable differences. In the 1960s and 1970s, the real-time VAR tracks the Livingston Survey closely, while the full-sample VAR significantly overpredicts inflation expectations nearly throughout the period. In contrast, in the early 1980s, during the Volcker disinflation, the real-time VAR severely overpredicts inflation expectations. This discrepancy may reflect judgmental modifications to forecasts that incorporate extra-model knowledge of the Fed's goals and actions at the time (as well as other influences on inflation) not captured by the simple VAR.

The issue of real-time forecasts versus forecasts after the fact also has implications for the interpretation of forecast rationality tests reported in the paper. In describing the properties of median forecasts, the authors apply standard tests of forecast rationality to the four survey series that they study. Such tests boil down to looking for correlations between forecast errors and observable variables, the existence of which implies that forecast errors are predictable and therefore not rational. They consider four such tests. The simplest test is that for forecast bias, i.e., nonzero mean in forecast errors. A second test is for serial correlation in forecast errors in nonoverlapping periods. A third test, which I call the forecast information test in the following, is a test of correlation between forecast errors and a constant and the forecast itself. The final test, which I call the all information test, is a joint test of the correlation between forecast errors and a set of variables taken from the VAR described above, assumed to be in forecasters' information set.

They find mixed results on bias and forecast information tests, but rationality of the median value of surveys is rejected based on the serial correlation and all information tests. They then show that the median forecast predicted by the sticky-information model yields similar results—with forecast errors exhibiting positive serial correlation and a high rate of rejection of the forecast information and all information tests—providing support for that model.

An alternative interpretation of these results is that forecasters have been learning about quantitative macroeconomic relationships over time. The

3. I chose the CPI for this analysis because it sidesteps the issue of differences between real-time and final revised data in national income account price indexes, such as the GDP deflator. The CPI is not revised except for seasonal factors, and because I am focusing on four-quarter changes in prices, seasonal factors should be of little importance to the analysis.

tests are based on the correlations in the full sample and ignore the fact that forecasters, even those with the correct VAR model, had inaccurate estimates of these relationships at the time of their forecasts. Because of sampling errors in the forecaster's model, these tests are biased toward rejecting the null of rationality.[4]

Indeed, a wide variety of reasonable forecasting models, including a quarterly version of the VAR used in the paper (with the unemployment rate substituting for the output gap), yield a pattern of rejections similar to those seen in the survey data when one assumes that the forecasts were constructed in real-time using knowledge of the data correlations available at the time. Table 1 reports the results from rationality tests from several simple forecasting models. The first line of the table reports the results from the three-variable VAR with four quarterly lags, where the VAR is reestimated each period to incorporate the latest observed data point. This real-time VAR exhibits no bias over the past 40 years, but forecast rationality is rejected based on positive forecast error serial correlation and the two information tests. (For this test, I include the most recent observed value of the inflation rate, the unemployment rate, and the 3-month T-bill rate.) The results for other forecast models (the details of which I describe below) are also consistent with the evidence from the surveys. And as indicated in the bottom line of the table, the median forecast among these 10 forecasting models also exhibits the pattern of rejections seen in the survey data. This evidence suggests that forecasters use models in which the parameters change over time.

I now turn to the second half of the paper. The authors show that the sticky-information model provides a parsimonious theory of disagreement that is in accord, at least qualitatively, with the time-series pattern of disagreement seen in the survey data. In addition, the model can generate a positive correlation between disagreement and the inflation rate, also a prominent feature of the survey data.

The evidence for the sticky-information model from inflation expectations disagreement from household surveys, however, is not clear cut. The model cannot come close to matching the magnitude of the dispersion in household inflation expectations, for which the interquartile range (IQR)—that already excludes one half of the sample as outliers—can reach 10 percentage points! In Figure 10 of the paper, the difference between the measure of disagreement in the data and that predicted by

4. In addition to uncertainty about model parameters, the specification of forecasting models changes over time in response to incoming data, driving another wedge between the information set available to real-time forecasters and after-the-fact calculations of what forecasters "should have known." In models with time-varying latent variables such as the natural rates of unemployment and interest, this problem also extends to the real-time specification and estimation of the latent variable data-generating processes, as discussed in Orphanides and Williams (2002).

Table 1 REAL-TIME MODEL-BASED FORECASTS[1]

	Forecast model				Tests of forecast rationality		
Description	*RHS variables*	*Discounting of past data*	*RMSE*	*Bias*	*Error serial correlation*	*Forecast info.*	*All info.*
VAR	π, u, r	None	2.28	−0.11	0.42 ***	***	***
		0.015	2.06	−0.16	0.11	***	***
		0.030	2.09	−0.11	−0.01	***	***
Phillips curve	π, u	None	1.96	0.20	0.50 ***		***
		0.015	1.69	0.02	0.19	***	***
		0.030	1.72	−0.01	0.03	***	***
Autoregressive	π	None	1.92	0.46 *	0.46 ***		***
		0.015	1.73	0.09	0.16	**	***
		0.030	1.73	0.05	0.13	***	***
Random walk	π	—	1.79	0.01	0.17	***	***
Memo							
Median forecast			1.67	0.02	0.22 *	**	***

1. Forecast variable is the four-quarter percentage change in the CPI. Sample: 1960–2002; quarterly data. RMSE denotes root mean squared forecast error. Bias is the mean forecast error. One asterisk indicates an asymptotic p-value between 5 and 10 percent; two asterisks, between 1 and 5 percent; and three asterisks, below 1 percent. Error serial correlation reports the estimated coefficient on lagged forecast error in regression of forecast error on a constant and the lagged forecast error. Test I is a regression of forecast errors on a constant and the forecasted value. Test II is a regression of the forecast error on a constant, the forecasted value, the inflation rate, the Treasury-bill rate, and the unemployment rate. P-values based on Newey-West HAC standard errors.

the model is not constant over time and appears to be highly persistent. There's clearly something else going on here, with the sticky-information model capturing only part of the process of households' expectations formation.

The sticky-information model is closely linked to Chris Carroll's (2003) model, whereby households randomly come into contact with professional forecasts. I find his model to be a highly plausible description of expectations formation by households, who are unlikely to keep in constant touch with the latest macroeconomic data. But as a macroeconomic forecaster myself, I find it entirely implausible as a description of the behavior of business economists and professional forecasters surveyed in the Livingston Survey and the Survey of Professional Forecasters (SPF), respectively. Professional forecasters update their forecasts regularly and update their forecasting models at frequent intervals. The primary reason economists' forecasts disagree is not due to lags in formulating new forecasts, but instead is because economists themselves disagree about how to model the economy best![5] This is an aspect of dispersions in expectations entirely absent from the model in the paper.

In fact, there already exist theories of rational heterogeneity of beliefs that naturally yield expectations disagreement (see, for example, Brock and Hommes [1997], Branch [2003], and Branch and Evans [2003]). These theories assume that agents have at their disposal a range of forecasting models but are uncertain about which model or models to use. They update their model choice or priors over the various models based on forecasting performance. Idiosyncratic differences in agents' characteristics, say, different initial conditions in model priors and the costs for learning new models, implies that a range of models will be in use at any point in time.

This description matches closely the real-world practice of economic forecasting that recognizes the high degree of model uncertainty in forecasting. There exists many competing inflation forecast models, including time-series models, Bayesian VARs, reduced-form Phillips curves, and large-scale macroeconometric models in use at the same time. And for each model, several variants differ with respect to model specification and estimation, including the lag length of explanatory variables; treatment of latent variables; sample size; and the inclusion or exclusion of additional explanatory variables such as energy prices, import prices, price control dummies, sample, wages, productivity, etc. (compare Brayton et al. [1999] and Stock and Watson [1999]). These models have similarly good track

5. The disagreement seen in published forecasts is an imperfect measure of true disagreement. There are incentives both not stray too far from the consensus (Scharfstein and Stein, 1990; Lamont, 2002) as well as to stand out from the crowd (Laster et al., 1999).

records in terms of forecasting performance, but at times they can yield strikingly different forecasts. In practice, forecasters combine the forecasts from a subset of these models, along with extra-model information, to arrive at a point estimate forecast.

But why don't all forecasters arrive at the same forecast? For the same reason as in the theory sketched above: idiosyncratic differences between economists imply persistent deviations in modeling choices. One source of such differences might originate during graduate school training. For example, economists trained in the 1960s likely rely more heavily on structural macroeconometric models for forecasting, while those trained during the 1990s probably place more weight on Bayesian VARs. Because these models are about equally good in terms of forecasting accuracy, the pace of convergence to a single best mix of models is likely to be slow.

To illustrate how model uncertainty can lead to forecast disagreement conforming to the evidence presented in the paper, I construct an artificial universe of forecasters, each of whom is assumed to use one of the 10 forecasting models listed in Table 1. As seen in the table, these models have roughly similar out-of-sample forecast accuracy. In each case, the model is reestimated each period. The parameters are unrestricted. For each of the three main types of models, I consider three variants: one is estimated by standard ordinary least squares (OLS) and the second and third are estimated by weighted least squares (WLS), with the weights declining geometrically using the values 0.985 or 0.970, as indicated in the table. WLS estimation is designed as protection against structural change by downweighting old data (Evans and Honkapohja, 2001). The first main model is the VAR described above. The second model is a Phillips curve model that includes a constant, four lags of inflation, and two lags of the unemployment rate. The third model is a fourth-order autoregression. The set of models is completed with a simple random walk model where the forecast inflation over the next four quarters equals the inflation rate over the past four quarters.

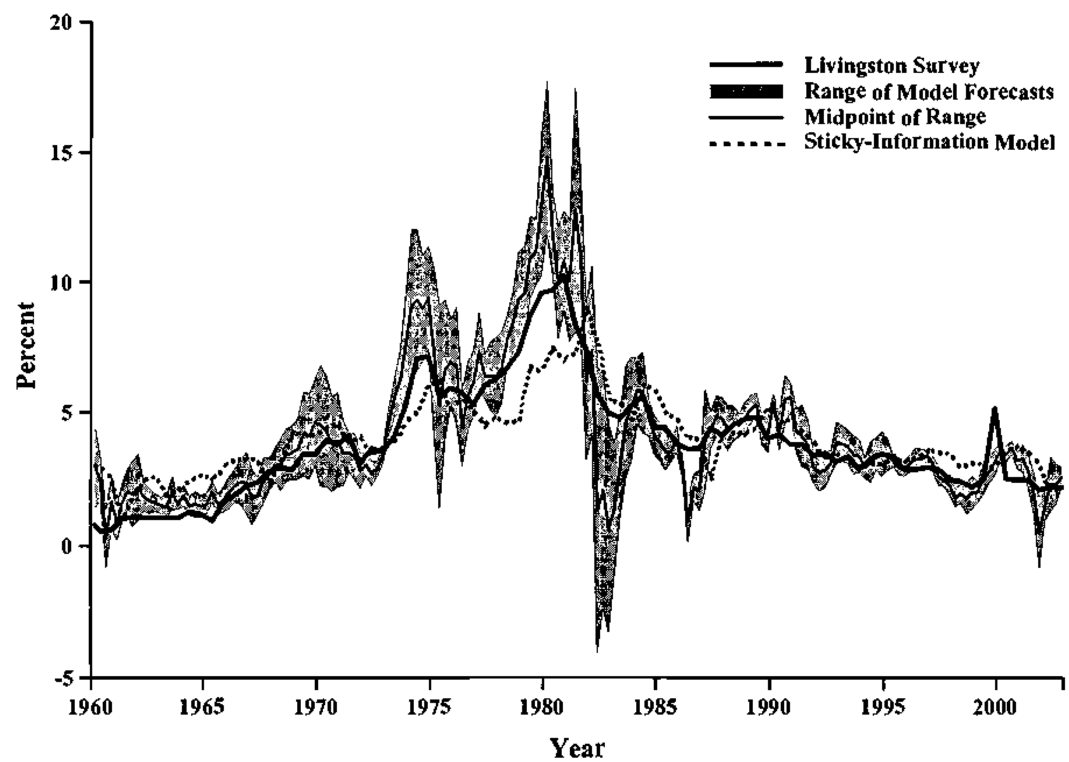
The out-of-sample forecasting performance (measured by the root mean squared forecast error) of these various models is quite similar. The random walk model beats the other models without discounting, consistent with the findings of Atkeson and Ohanion (2001). The VAR model is the worst performer of the group, supporting Fair's (1979) finding that unrestricted VAR models tend to perform poorly out of sample. The Phillips curve model, with discounting, is the best performer of the group, and in all three cases the model with discounting outperforms the OLS version. Evidently, structural breaks are of great enough importance in this sample that protecting against such breaks outweighs the efficiency loss associated with discarding data (see Orphanides and Williams

[2004]). Finally, the median forecast from this set of 10 models performs better than any individual forecast, in line with the literature that averaging across forecast models improves performance (see Clement [1989] and Granger [1989] for surveys of this topic).

At times, these forecasting models yield different forecasts of inflation, as shown by the shaded region in Figure 2. The degree of disagreement across models widens appreciably around 1970, again in the mid-1970s, and most strikingly during the period of the disinflation commencing at the end of the 1970s and continuing into the early 1980s. The magnitude of disagreement across models is much smaller during the 1960s and from the mid-1980s through the end of the sample.

The time-series pattern seen in the interquartile range from these forecast models is similar to that seen in the Livingston Survey IQR, as depicted in Figure 3. During periods of stable and low inflation, forecast dispersion among these models is modest, but it spikes when inflation is high or changing rapidly, exhibiting the same pattern as in the IQR from the survey data. Of course, the purpose of this exercise is only to illustrate how model uncertainty can give rise to forecast disagreement similar to that seen in the survey data. As noted above, the extent of model disagreement extends beyond the set of models I have considered here,

Figure 2 MODEL UNCERTAINTY AND FORECAST DISAGREEMENT

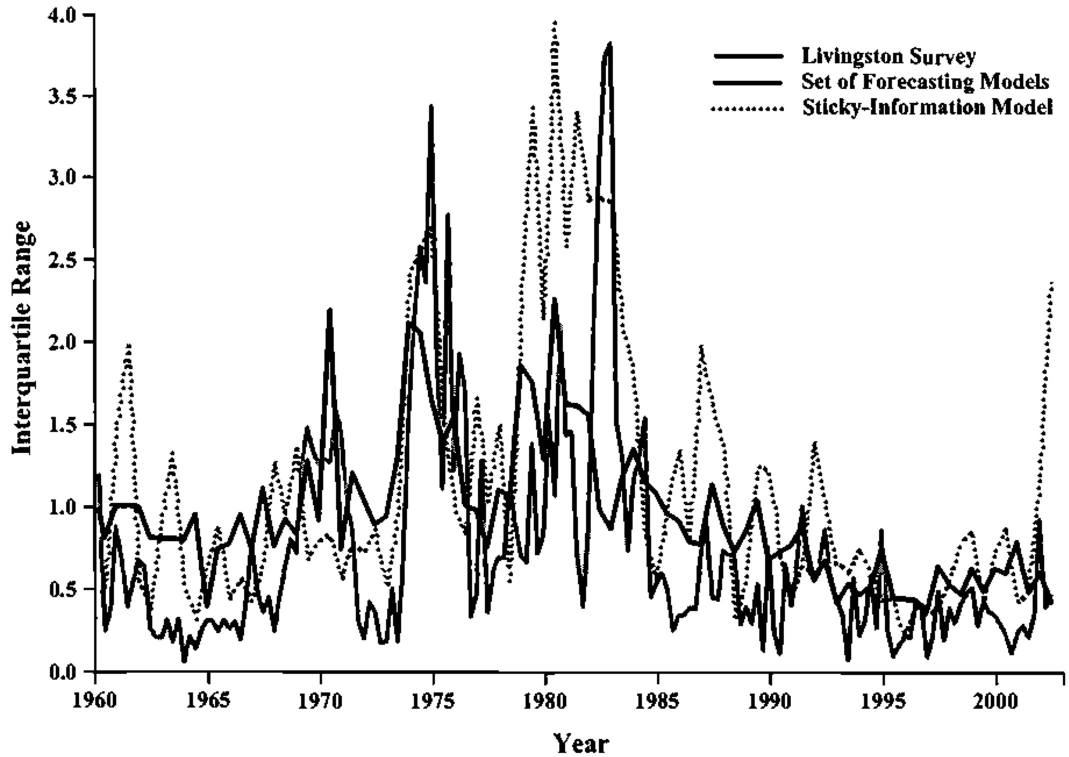


suggesting a wider range of disagreement than reported here. On the other hand, good forecasters will average various models, reducing the range of disagreement implied by individual models. Careful quantitative analysis of model-based forecast disagreement is left to future work.

Of course, there are many reasonable alternative explanations, in addition to the sticky-information model, for forecast disagreement among economists that conform to the main properties of inflation expectations highlighted in this paper. My goal was to provide an illustrative example of one such case. (Robert King's discussion of public disagreement regarding the Federal Reserve's ultimate inflation objective in this volume provides another.) To test these alternative theories against each other, it will be useful to examine evidence from expectations of a wider set of variables, including long-run inflation expectations, and from surveys in other countries. In addition, this paper has focused on a few moments in the survey data. The Livingston Survey and SPF are panel datasets that track the responses through time of individual forecasters. The panel aspect of this data is an untapped well of information that may help discern different theories of disagreement.

Finally, although expectations disagreement may be useful in discerning alternative models of expectations formation, it is unclear how quantitatively important disagreement on its own, even time-varying

Figure 3 FORECAST DISAGREEMENT



disagreement, is for the evolution of the aggregate economy, given a path for the mean expectation. The macroeconomic implications of time-varying disagreement provide another avenue of future research that is sure to be stimulated by this fine paper.

REFERENCES

Atkeson, Andrew, and Lee E. Ohanian. (2001). Are Phillips curves useful for forecasting inflation? *Federal Reserve Bank of Minneapolis Quarterly Review* 25(1): 2–11.

Ball, Laurence, and Dean Croushore (2001). Expectations and the effects of monetary policy. Federal Reserve Bank of Philadelphia. Research Working Paper: 01/12.

Branch, William A. (2003). The theory of rationally heterogeneous expectations: Evidence from survey data on inflation expectations. Department of Economics, College of William and Mary. Manuscript.

Branch, William A., and George W. Evans. (2003). Intrinsic heterogeneity in expectation formation. Department of Economics, College of William and Mary. Manuscript.

Brayton, Flint, John M. Roberts, and John C. Williams. (1999). What's happened to the Phillips curve? Board of Governors of the Federal Reserve System. Finance and Economics Discussion Paper Series: 99/49.

Brock, William A., and Cars H. Hommes. (1997). A rational route to randomness. *Econometrica* 65(5);1059–1160.

Carroll, Christopher D. (2003). Macroeconomic expectations of households and professional forecasters. *Quarterly Journal of Economics* 118(1):269–298.

Clement, Robert T. (1989). Combining forecasts: A review and annotated bibliography. *International Journal of Forecasting* 5(4):559–583.

Evans, George W., and Seppo Honkapohja. (2001). *Learning and Expectations in Macroeconomics*. Princeton, NJ: Princeton University Press.

Fair, Ray C. (1979). An analysis of the accuracy of four macroeconometric models. *Journal of Political Economy* 87(4):701–718.

Granger, Clive. (1989) Combining forecasts—Twenty years later. *Journal of Forecasting* 8:167–174.

Gürkaynak, Refet, Brian Sack, and Eric Swanson. (2003). The excess sensitivity of long-term interest rates: Evidence and implications for macroeconomic models. Board of Governors of the Federal Reserve System. Manuscript.

Lamont, Owen A. (2002). Macroeconomic forecasts and microeconomic forecasters. *Journal of Economic Behavior and Organization* 48(3):265–280.

Laster, David, Paul Bennett, and In Sun Geoum. (1999). Rational bias in macroeconomic forecasts. *Quarterly Journal of Economics* 114(1):293–318.

Mankiw, N. Gregory, and Ricardo Reis. (2002). Sticky information versus sticky prices: A proposal to replace the new Keynesian Phillips curve. *Quarterly Journal of Economics* 117(4):1295–1328.

Orphanides, Athanasios. (2001). Monetary policy rules based on real-time data. *American Economic Review* 91(4):964–985.

Orphanides, Athanasios, and Simon van Norden. (2002). The unreliability of output-gap estimates in real time. *Review of Economics and Statistics* 84(4):569–583.

Orphanides, Athanasios, and John C. Williams. (2002). Robust monetary policy rules with unknown natural rates. *Brookings Papers on Economic Activity* 2:63–145.

Orphanides, Athanasios, and John C. Williams. (2004). Imperfect knowledge, inflation expectations, and monetary policy. In *Inflation Targeting*, Michael Woodford, (ed.). Chicago, IL: University of Chicago Press.

Scharfstein, David S., and Jeremy C. Stein. (1990). Herd behavior and investment. *American Economic Review* 80(3):465–479.

Stock, James H., and Mark W. Watson. (1999). Forecasting inflation. *Journal of Monetary Economics* 44(2):293–335.

Discussion

A number of participants suggested alternative explanations for the disagreement across agents in inflation expectations documented by the authors, and they recommended that the authors test their theory against plausible alternatives. Olivier Blanchard suggested the possibility that people form inflation expectations based on small samples of goods. Inflation variability across goods could then drive dispersion in expectations. He noted that there is indeed a lot of inflation variability across goods. He pointed out that this story is consistent with the fact that there is more dispersion in inflation expectations when inflation is changing. He also speculated that women might buy different baskets of goods from men, hence explaining the difference in expected inflation between women and men in the Michigan Survey. Justin Wolfers explained that, in favor of Blanchard's story, people's assessment of past inflation is a good predictor of their inflation expectation. The caveat is that the extent of variation in the inflation that they think they have experienced is not rationalizable. Ken Rogoff pointed out that there is evidence of substantial inflation dispersion across cities within the United States. He suggested that this might favor the hypotheses of Blanchard and Williams: that individual forecasts might be made on the basis of different baskets.

Another explanation was suggested by Mike Woodford. He remarked that Chris Sims's theory of limited information-processing capacity could endogenously generate the co-movement of disagreement and other macro variables. Greg Mankiw responded that the sticky-information model is a close relation of Chris Sims's work, but that it is analytically more tractable and generates testable predictions more easily than the work of Sims.

Ken Rogoff asked whether a story about policy credibility could explain the fact that there is more disagreement when inflation is changing. He hypothesized that fat tails in the distribution could be due to expectations of more extreme histories than the United States has experienced. He noted that, in view of the historically higher inflation experienced by both developing countries and other OECD countries, such expectations were not necessarily unreasonable. He also remarked that few plays of the game have been experienced so far. Greg Mankiw objected that a credibility story would not be able to explain positive autocorrelated errors in forecasts. Robert Shimer responded that Markov switching in inflation regime and a small sample is sufficient to explain such autocorrelation. Ricardo Reis agreed with Rogoff that the inflation history of the United States was uneventful from the point of view of identification.

Athanasios Orphanides was concerned that the authors' sticky-information explanation for disagreement cannot reasonably explain disagreement among professional forecasters who have strong incentives to update their information regularly. He suggested that a more reasonable explanation would be disagreement about the model used to forecast inflation or about estimates of variables such as the natural rate of unemployment. As evidence that such disagreement is widespread, he cited the response of many forecasters in the SPF that they did not use an estimate of the NAIRU to forecast inflation. Robert Shimer agreed with Orphanides that professional forecasters were unlikely to have sticky information. Justin Wolfers responded to Orphanides that disagreement about models or the NAIRU is merely a richer version of the sticky-information explanation of why people stick to bad forecasts for long periods of time.

Mark Gertler said that examining the source of cyclical dispersion in expectations is important. He hypothesized that it may be due to something beyond sticky information or imperfect monetary policy credibility. He noted that, although the Fed is unlikely to suffer from sticky information or imperfect credibility, its Green Book forecasts in the late 1970s missed both the upsurge in inflation and the disinflation. He remarked that this forecast error is correlated with dispersion in inflation expectations and suggested that this avenue would be interesting to explore. Greg Mankiw asked Gertler whether the Green Book forecast errors were autocorrelated over the period covered by the paper. Gertler responded that, because inflation was relatively flat over the period, it is hard to distinguish autocorrelation.

The reliability of the data used by the authors concerned some participants. Rick Mishkin contended that household surveys should be regarded with skepticism. He noted that, in contrast to forecasters, who make their living from their expectations, households have no incentive to think hard about their survey responses. He suspected that respondents claiming to expect 10% inflation were unlikely to be behaving in a way consistent with this expectation, and that the level of dispersion in the survey responses was exaggerated. Mishkin also noted that there is a literature that documented incentives for forecasters to make extreme predictions to attract attention. Greg Mankiw responded that the fact that the cyclical response of expected inflation is as predicted by the model, and that disagreement co-moves strongly across different surveys suggests that they are picking up more than just noise. Mankiw agreed that the fact that private-sector forecasters are selling a product does affect the incentives they face compared with, for example, the Fed's Green Book forecast. Ken Rogoff noted that forecasters may try to avoid changing their predictions to maintain credibility.

Ricardo Reis stressed that the main point of the paper was to demonstrate the extent of disagreement across agents in inflation expectations. He remarked that nothing in macro theory so far can explain why there should be so much disagreement, and the paper explored a first possible explanation. Reis said that there is a large middle ground between adaptive and hyperrational expectations, and that the paper is an attempt to explore one possible alternative. Greg Mankiw concluded that the bottom line of the paper is that disagreement is widespread, and it varies over time with variables that interest macroeconomists. Hence, macro models should be consistent with disagreement of this type.