TESTING THE RATIONALITY OF PANEL DATA FORECASTS USING A GMM ESTIMATOR

Richard H. Cohen, Economics Department, University of Hawaii

I. Introduction

The search for rational (i.e., unbiased and conditionally efficient) forecasts of macroeconomic variables is a challenging and significant endeavor for planners in the private and public sectors: major investment decisions may hinge on an assessment of the macroeconomic environment. In contrast, the conventional wisdom among academics testing models of optimizing agents treats the existence of rational forecasts as a maintained assumption, rather than as a hypothesis to be tested. However, should the model fail, it could be instructive to be able to attribute the failure to a misspecification of the behavioral model or a misspecification of the expectation assumption. To separate these two potential causes, it would be necessary to undertake an independent test of the expectation assumption.

Using recently developed panel data methods, we analyze forecast errors and conduct rationality tests on real GNP growth predictions of professional forecasters who participated in the American Statistical Association-National Bureau of Economic Research quarterly survey from its inception in 1968 to its termination in 1989. The conventional approach to dealing with pooled cross-section and time series data assumes that 1) there is an equal number of observations for each decision unit and 2) there are no missing observations. The specification of the behavior of the disturbances frequently assumes that they are cross-sectionally heteroscedastic and time-wise autoregressive. In contrast the approach we use in this paper allows for 1) a different number of observations for each decision unit and 2) a random pattern of missing observations. In addition, following Keane and Runkle (KR, 1990), our test statistics take into account not only an individual forecaster's prediction errors due to lagged data availability, but also contemporaneous and lagged cross-correlations among forecasters. Our programs highlight the power and flexibility of SAS/IML. Our main findings are that, while forecasts are unbiased and efficient across the entire sample period, there are significant overforecasts of real GNP growth for all recessionary quarters through the fifth. Further, there is some evidence of violations of conditional efficiency during recessions, suggesting that the overestimates are potentially avoidable.

Section II gives a general overview of rationality tests. Section III describes our adaptation of a panel data methodology developed by KR to testing the rationality of real income growth forecasts during recessions. Section IV describes each test and reports our findings. Section V briefly concludes.

II. General Overview of Rationality Tests

According to Brown and Matal (1979, pp.493), "full rationality implies that all available information has been used in an optimal manner." Assuming a symmetric quadratic loss function, it is straightforward to show that the optimal (in the sense of minimum MSE) use of available information is achieved when the agent's anticipation of real income growth is set equal to the conditional mathematical expectation of the realized real income growth process.

In general, then, full rationality implies that the conditional forecast errors have zero means, i.e., they are orthogonal to each of the variables in the information set. Defining $y_{t+1} = \chi_t$ as the optimal forecast of real income growth, conditional on all relevant information available at $t$. Obviously, it is virtually impossible to demonstrate full rationality, inasmuch as the latter requires that the orthogonality condition holds for every subset $\mathcal{X}_t$ of $\chi_t$. If the mean conditional forecast error equals zero for a given subset of information set variables, the forecast is said to satisfy the condition for partial rationality, i.e., with respect to that subset. It then follows that partial rationality is a necessary condition for full rationality. For example, if we assume that the expectation of real income growth is a linear function of information set variables, then the null hypothesis is that, in the regression

$$y_{t+1} = \beta Z_t + \epsilon_{t+1}$$

$\beta = 0$ (where $\beta$ is a vector of coefficients and $Z_t$ is a vector of information set variables). In this sense rationality tests "are more suited to give evidence against the rational expectations assumption rather than in its favor" (Visco, 1984, p. 140).

We will test for bias as well as conditional efficiency of real income growth forecasts for the entire sample period as well as during recessions. In the tests that follow, we will be careful to distinguish between systematic unconditional forecast errors at a given stage of the business cycle, which have no necessary implication for violation of the rationality assumption, and systematic forecast errors that are conditional upon information available to the forecaster at the time the forecast is made.

III. Implementing the Keane-Runkle Methodology

We base our real income rationality tests on the same data base and methodology that KR employ to test the rationality of survey price level forecasts. According to these authors, their approach remedies at least four flaws common in the literature on testing rationality. First, their subjects (polled in the ASA-NBER survey) are professional forecasters, who, unlike typical households, possess both the knowledge base and economic incentive to produce accurate forecasts. Second, KR test forecast rationality by comparing their respondents' predictions to the initial observations available in real time, rather than the final revised estimates. This is intended to avoid biasing the tests in favor of rejection due to unforeseeable revisions. Third, KR test the rationality of individual forecasts using panel data methods. This avoids the aggregation bias inherent in the use of sample means of forecasts at each point in time. Fourth, in order to adjust for cross-forecaster correlation of prediction errors, which had confounded previous attempts to employ panel data techniques, KR develop a consistent estimator for the covariance matrix. This estimator takes into account not only an individual forecaster's MA(1) errors due to lagged data availability, but also contemporaneous and lagged cross-correlations among forecasters, as well. With these modifications KR are unable to reject unbiasedness or informational efficiency of price level forecasts.

In what follows we describe in some detail how the above innovations are incorporated in our rationality tests.

A. The Prediction Series

For each quarter from 1968:IV to 1989:IV Victor Zarnowitz coordinated the ASA-NBER Survey of Forecasts by Economic Statisticians. Professional economic forecasters were polled for their predictions of approximately twenty macroeconomic variables. The forecasters were asked for their predictions for the current quarter (since the relevant data is unavailable by the end of the second month of the quarter, when questionnaires must be returned) and for the next four quarters.

We construct our prediction series by taking the difference between the logarithms of the one-step-ahead real GNP forecasts and the current real GNP forecasts. We choose to express the forecast objective in growth rate terms to lessen the potential for statistical problems due to 1) nonstationarity of the forecast objective; and 2) heteroscedasticity resulting from redefinitions of the base year figure used to standardize the level of real GNP. Because these base year changes are also associ-
ated with unforecastable definitional changes, we delete from our analysis the three quarters during which major definitional benchmark changes occurred: 1975:IV, 1980:IV, and 1985:IV. Like KR, we include all forecasters who responded to the survey for at least twenty quarters, not necessarily consecutive. This produces a sample size (N) equal to 78 forecasters.

As in the case with KR's level series of price forecasts, there is still a two-quarter lag between the one-step-ahead prediction of real GDP growth and the first revised data. The latter is available 45 days after the end of the forecast quarter—approximately the time when the next quarter's questionnaires must be returned. Thus, we assume that the prediction error process follows (at least) an MA(1) process. This implies, of course, that rational forecasts cannot be expected to produce independent, identically distributed errors: there will generally be a nonzero covariance between the forecast error for t+1 and information pertaining to time t but not publicly available until t+1.

B. The Forecast Objective

As KR point out (1990, p. 724), "whether revised data should be used [as a forecast objective] depends on two issues: First do forecasters try to predict the initial or the revised data? Second, are there significant and predictable data revisions?" To the extent that the preliminary data are rational forecasts of the final revised data, forecast revisions will be unpredictable. This means that the preliminary release itself does not contribute any additional information relevant to forecasting the final revised data, beyond that already contained in the information set used to construct the preliminary release. In this case the choice of dependent variable should have little impact on tests of forecast rationality.

KR cite a study by Mankiw and Shapiro (1985) which shows that, for the period 1976:1 to 1982:IV, early estimates of real GDP growth were rational forecasts of subsequently revised estimates. The intuition underlying Mankiw and Shapiro's test is as follows: Revisions from early to later releases of macroeconomic data arise from either of two sources—measurement error or forecast error. Measurement error occurs if the earlier data reflects sampling errors. (One obvious source of measurement error in the preliminary estimate is that it is based on incomplete data). In this case the revision should be uncorrelated with the revised value but correlated with the early estimate. For example, if the early estimate turned out to be unusually high, the later estimate would probably be revised downward. Thus, in general we would expect the correlation between early estimates and their revisions to be negative. On the other hand, if the early data is relatively free from measurement error, we would expect the earlier estimates to represent the best available forecasts of the later estimates. The earlier estimates then contain only forecast errors, which are assumed to be uncorrelated with the earlier estimate. Instead, revisions are correlated with the later estimates only.

Another implication of the measurement error/forecast error dichotomy is based on a comparison of the variances of the respective releases. "Efficient forecasts are necessarily smoother than the object being forecast" (Mankiw and Shapiro [1986], p. 22). An efficient forecast, therefore, has a lower variance than the forecast objective. Conversely, forecasts characterized by measurement error have higher variances than the forecast objective.

In panel 9(a) of Table 2 we examine these two implications with respect to the 45-day and prebenchmark revised data. For the entire sample period the data strongly support Mankiw and Shapiro's conclusion that the earlier release is at least an unbiased forecast of the later release. (We will see, below, that this conclusion does not hold for the subset of observations corresponding to recessions). First, the standard deviation of the prebenchmark revised data exceeds the standard deviation of the 45-day release. Second, the correlation of the revision with the prebenchmark revised figure is substantial (0.347) and statistically significant (p = 0.0018), while the correlation of the revision with the 45-day release is much lower (0.092) and insignificant (p = 0.423). Thus, to the extent that we believe that the true forecast objective is the final prebenchmark revised value, we can use these data as our dependent variable without being concerned that our tests will be biased toward rejection of the null hypothesis of rationality, as would have been the case if the 45-day data contained information which helps to predict the prebenchmark revised value. For this reason we adopt the final prebenchmark revised data (as defined in Zarnowitz [1985]) as our forecast objective.

C. The Use of Panel Data versus Survey Means

KR identify two problems with using consensus forecasts in tests of rationality. First, if some forecasters are irrepressibly optimistic about the state of the economy, while others are just as inherently pessimistic, then averaging their expectations will tend to cancel their individual biases. This tends to produce false acceptance of the null hypothesis of unbiasedness. Put differently, we are not testing the rationality of individuals' forecasts, but rather demonstrating the likely superiority of a combined forecast.

A second problem is that use of the survey means as an independent variable creates a specification error that leads to upward bias of the slope coefficient. To our knowledge KR are the first to discover this aggregation bias, which increases with the dispersion of the individual forecasts. While KR do not provide an explanation for this result, one possible explanation appears to be that, even if the covariance of the survey mean forecast with the forecast objective equals the average covariance of the individual forecasts with the forecast objective, the variance of the survey mean forecast is necessarily less than the variance of the individual forecasts. Thus, least squares estimation produces an inflated value for the coefficient of the forecast. (KR cite another study in which the bias is on the order of 40 percent).

Both these problems are manifestations of the ambiguity involved in identifying the specific information set which conditions a consensus forecast. Rational forecasts differ only if respondents' predictions are based, in part, on private information. However, one respondent's forecast error may be orthogonal to both public and private components of his own information set, but not to the private component of another respondent's information set. In general the correlation of regressor and error will produce biased and inconsistent parameter estimates, even if the regressor is publicly available. This specification bias is avoided if tests of bias are undertaken on individual forecasts or pooled cross-section-time-series forecasts.

Because pooling increases the degrees of freedom compared to running separate regressions for each forecaster, tests of hypotheses using panel data will be substantially more powerful than tests using individual data. This consideration is especially significant for the current project, because most of our tests of conditional rationality use a recession dummy variable for which there are relatively few observations in the 20 years of data. (Only 20 out of 82 quarters are classified as recessions).

D. Incorporating Cross-Covariances of Forecast Errors

However, the existence of aggregate shocks creates a distinct specification problem when panel data are used in OLS regressions. (See Zarnowitz [1985]). KR agree with Zarnowitz that cross-correlated shocks render the covariance matrix of the parameter estimates inconsistent. To remedy this problem, they construct a GMM covariance estimator, based on Hansen and Hodrick (1980), that takes into account cross-section correlation, both contemporaneous and at a one-quarter lag. (An MA(1) error structure is also incorporated into each forecaster's prediction errors). We follow KR's specification of the GMM covariance matrix by assuming that
The error structure is homoscedastic with a one-period memory, necessitated by the publication lag described in the previous section. Thus, every forecaster has the same individual T x T covariance matrix $Q_i$. With elements $q_i(m,n) = \delta_{mn}$ and $q_i(m,n-1) = \delta_{mn}^1 = 0$ otherwise.

This similarity holds for any pair of forecasters $i$ and $j$:

$$E[e_i \epsilon_j] = \delta_{ij}, \quad E[e_i \epsilon_j^*] = \delta_{ij}^1, \quad E[e_i^* \epsilon_j^*] = 0 \quad \text{for all } i, j, k \quad \text{such that } k > 1.$$

Thus, every pair of forecasters has the same $T \times T$ cross-covariance matrix.

Similarly, for each pair of forecasters $i$ and $j$:

$$E[q_i(m,n)q_j(n,m)] = \delta_{mn}, \quad E[q_i(m,n)q_j(n,m-1)] = \delta_{mn}^1, \quad E[q_i(m,n)q_j(n,m+1)] = 0 \quad \text{otherwise.}$$

Thus, $E[q_i(m,n)] = \delta_{mn}$ for all $i, j$.

The cross-covariance matrix has dimension $NT \times NT$. (For the subset of recessionary observations, the mean forecast error of -3.1% is composed of a mean realization of -2.2% and a mean forecast of 0.9%). Statistically significant mean forecast errors occurred in each recessionary quarter except for the second and sixth. Interestingly, there appears to be no strong systematic relationship between length of recession and magnitude of forecast error. (Using information set variables, we test this formally in Table 2, panel (5).) On average, consistent declines in the magnitude of forecast errors are not evident until the fifth quarter of a recession. (The longest recession is only six quarters). Because a relatively large component of the total decline in real GNP growth during recessions is unanticipated, the remaining task is to determine to what extent the recessionary forecast errors were unavoidable, given publicly available information.

### Table 1

<table>
<thead>
<tr>
<th>Test</th>
<th>Hypothesis</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$H_0: \beta_0 = 0, \beta_1 = 1$ in $Y_{t+1} = \beta_0 + \beta_1 Y_t + \varepsilon_{t+1}$</td>
<td>$p = 0.860$, indicating that the null hypothesis is not rejected.</td>
</tr>
<tr>
<td>2</td>
<td>$H_0: \beta_0 = 0, \beta_1 = 0$ in $Y_{t+1} = \beta_0 + \beta_1 Y_t + \varepsilon_{t+1}$</td>
<td>$p = 0.391$, indicating that the null hypothesis is rejected.</td>
</tr>
<tr>
<td>3</td>
<td>$H_0: \beta_0 = 0, \beta_1 = 0$ in $Y_{t+1} = \beta_0 X_t + \varepsilon_{t+1}$</td>
<td>$p = 0.036$, indicating that the null hypothesis is rejected.</td>
</tr>
</tbody>
</table>

The tests for bias in panel (1) of Table 2 shows that, despite the significant prediction errors during recessions, the null hypothesis of no bias cannot be rejected for the entire sample. Therefore, in the following tests for conditional efficiency, we impose the unbiasedness restriction and make our dependent variable the forecast error.

Next, we need to determine the lag at which information about the onset of a recession becomes part of the forecasters' information sets. If Ferson and Merrick's assumption is true, then rational forecasters should not be able to use information about the current stage of the business cycle to improve their forecasts of real GNP growth. Thus, our first test of orthogonality of the forecast error is with respect to Ferson and Merrick's contemporaneous recession dummy variable:

$\text{Test (3): } H_0: \beta_0 = 0, \beta_1 = 0 \text{ in } Y_{t+1} = \beta_0 X_t + \varepsilon_{t+1}$

where $D_{t+1}$ is a panel data context.

The p-value for this test (in panel (3)) is 0.860, indicating that this
variable cannot be used to improve forecasts.

As a further check that no past information about the existence of the current recession can be used to improve forecasts, we test for the collective significance of lagged recession dummies:

Test (4): \( H_0: \beta_0 = \beta_1 = \ldots = \beta_p = 0 \) in
\[ Y_{t+1} - \gamma = \beta_0 + \beta_1 D_{t-1} + \epsilon_{t+1} \]
where \( D_{t-1} = 1 \) if the economy was in recession \( p \) quarters prior to the one in which the forecast is made \((i = 1, \ldots, 5)\); 0 otherwise.

The results in panel (4) show that forecasters could not use information that the economy was in recession \( p \) quarters ago to improve their forecasts.

One particular way of combining information about the stage of the business cycle at recent lags is to conjecture that the behavior of real GNP growth is related to the length of the current recession. Thus, in the next test we examine whether this function of lagged dummy variables helps to improve forecasts.

Test (5): \( H_0: \beta_0 = \beta_1 = 0 \) in
\[ Y_{t+1} - \gamma = \beta_0 + \beta_1 D_{t-1} + \epsilon_{t+1} \]
where \( D_{t-1} = 1 \) if the economy was in the \( x \)th quarter of a recession at \( t-1 \), 0 otherwise.

In this case nonrejection occurs with a p-value of 0.313. From the survey forecasts, Zamowitz and Moore (1982) have developed a series of "sequential signals of recession and recovery." These signals the current stage of the business cycle. For the tests that follow, we adopt what we believe is a more reasonable assumption and include recession dummies with a lag of one period or longer.

C. Testing Conditional Efficiency With Respect to the Zamowitz-Moore Filter for Predicting Recessions

Instead of relying on the "naive" or "no-change" forecast based on the stage of the business cycle in the quarter preceding the forecast, we decided to investigate whether a particular rule for forecasting stages of the business cycle contains information which could be used to explain survey forecast errors. Zamowitz and Moore (1982) have developed a series of "sequential signals of recession and recovery." These signals are based on smoothed rates of change in the composite indexes of leading and coincident indicators. For example, in their band approach (1982, pp.76-77),

\[ \text{[the expected sequence of signals at business cycle peaks...is when each of the following signals is first observed:} \]

First signal (P1): The leading index rate falls below 2.3%; the coincident index rate will usually be higher than 2.3%, but we require only that it be nonnegative \((L < 2.3; C \geq 0)\).

Second signal (P2): The leading index rate falls below -1.0% and the coincident rate falls below 2.3% \((L < -1.0; C < 2.3)\).

Third signal (P3): The coincident index rate falls below -1.0%, while the leading index rate is still negative \((L < 0; C < -1.0)\).

...the expected sequence of signals at business cycle troughs is given by the first occurrence of the following:

First signal (T1): The leading index rate rises above 1.0%, while the coincident index rate is less than 1.0% \((L > 1.0; C < 1.0)\). A T1 must follow a P1.

Second signal (T2): The leading index rate rises above 4.3% and the coincident index rate rises above 1.0% \((L > 4.3; C > 1.0)\).

Third signal (T3): Both the leading index rate and the coincident index rate rise above 4.3% \((L > 4.3; C > 4.3)\).

We want to choose the filters that provided the most reliable and timely forecasts of business cycle changes. The three signals of recession and the three signals of recovery correctly called all postwar peaks and troughs. However, P1 provided a false signal on three occasions, and P2 provided a false signal on one occasion. However, P3 lagged business cycle peaks by an average of three months, while P2 led peaks by an average of three months. We decided to use P2 as our criterion for predicting peaks. Because T1 provided no false signals and had the longest lead time of the three signals of recovery (i.e., minus one month, which is the shortest lag time), we chose T1 as our criterion for predicting troughs. Thus, the recession dummy equals one when P2 is satisfied, changes to zero when T1 is satisfied, and so on.

Thus, in Test (6) we examine whether forecasters made efficient use of the Zamowitz-Moore filter:

Test (6): \( H_0: \beta_0 = \beta_1 = 0 \) in
\[ Y_{t+1} - \gamma = \beta_0 + \beta_1 D_{t-1} + \epsilon_{t+1} \]
where \( D_{t-1} = 1 \) if the economy is forecast at time \( t \) to be in recession, based on the application of the Zamowitz-Moore (1982) band filter.

The forecasters evidently made good use of the information in these band filters; the p-value for this test of conditional efficiency is 0.770.

D. Testing Conditional Efficiency with Respect to Forecasts Generated from a Univariate Time Series Model

As a basis for comparison with survey forecast accuracy as well as rationality, we will test the efficiency of the survey forecasts with respect to a univariate time series model for real income growth. We will estimate this model via recursive least squares, i.e., using data prior to the relevant forecast. A stationary univariate relation can serve as an appropriate benchmark for other forecasts, whether generated from survey data or from behavioral models. This is because, as Nerlove (1983) has pointed out, in the absence of structural change, univariate models satisfy at least two conditions necessary for partial rationality: unbiasedness of forecasts and (if the forecast interval coincides with the measurement interval) lack of serially correlated errors. (Of course, these models may omit relevant variables from the true behavioral model, and thus fail to be fully efficient).

Thus, we investigate whether the panel's forecasts could have benefited from (two-step-ahead) recursive predictions generated by a univariate time series model. We purposely construct the data set so that, if the test is biased, it is biased in favor of rejecting the null hypothesis of conditional efficiency. (For details on the construction of the data set, see Cohen (1991)). For data prior to the initial forecast of 1968:IV, examination of autocorrelation, partial autocorrelation and inverse autocorrelation functions suggest that the real GNP growth series is best modeled as a stationary AR(1) process. We also investigate whether the stationarity property also carries over to the end of the forecast sample period. Using a Dickey-Fuller test (not reported) we are able to reject the null hypothesis of a unit root for both the data set ending in 1968:III and the one ending in 1989:II. Also, based on a comparison of the Akaike Information Criterion, we confirm that, relative to its main competitor, an MA(1) model, the AR(1) specification proves to be the more appropriate structure for each of the sequentially updated data sets from 1968:III to 1989:II.

Thus, consider the following formulation:

Test (7): \( H_0: \beta_0 = \beta_1 = 0 \) in
\[ Y_{t+1} - \gamma = \beta_0 + \beta_1 Z_t + \epsilon_{t+1} \]
where \( Z_t \) is the 2-step-ahead recursive prediction of real GNP growth generated from a univariate model.

Despite "rigging" the test so as to favor rejection of conditional efficiency, the p-value of 0.544 (shown in panel (7)) provides no evidence of inefficiency. Further, the forecasting accuracy of the univariate model is consistently inferior to that of the professional forecasters. (The relevant table is in Cohen (1991)). This inferiority is especially evident during the first quarter of recessions, when the mean forecast error for the univariate model is -9.9%, versus -5.9% for the forecasters. Failure to predict turning points is a well-known shortcoming of AR models.
E. Testing Conditional Efficiency with Respect to the Consumption-Based CAPM's Lagged Dependent Variables

Since real income growth, real consumption growth and the real rate of return are jointly determined in the CCAFPM, and since the model's restrictions are decisively rejected during recessions, it may be of interest to investigate whether forecasts of real GNP growth take into account the recessionary behavior of the other two endogenous variables. Our information set variables are the cross-products of the first and second lags of real consumption growth and the real rate of return with the respective lags of the recession dummy, constructed so as to replicate the most recent data available to the forecasters in real time.

Test (8): $H_0: \alpha = \beta = 0$ in

$$y_{t+1} = z_0 + \beta_1 D_{1,t} + \beta_2 C_{t-1} + \beta_3 D_{1,t}C_{t-1} + \beta_4 D_{2,t} + \beta_5 C_{t-1}$$

$D_{1,t}$ is as defined as in test (4), $C_{t-1}$ is the annualized growth of real consumption, a weighted average of nondurables and services, based on the 45-day data available for the previous quarter. $C_{t-1}$ is the same variable, based on the 75-day data available for two quarters previous. $r_{t+1}$ is the real 90-day Treasury bill rate, using CRSP Government Bond File data for the previous quarter, deflated using a weighted average of the 45-day implicit price deflators for the previous quarter, $r_{t+1}$ is the same variable, using CRSP data from two quarters previous, deflated using a weighted average of the 75-day implicit price deflators available for two quarters previous.

As can be seen from panel (8), this test just barely fails to reject conditional efficiency. (The p-value is 0.052).

F. A Digression: Accuracy and Conditional Efficiency of the Department of Commerce's 45-Day Release of Real GNP Growth

In Section III(B) above we concluded that, over the entire sample period, revisions behave more like forecast errors than measurement errors, supporting the hypothesis that the 45-day data are efficient forecasts of the final revised prebenchmark figures. (See Table 2, panel (9a)). Examining these same diagnostics during recessions gives some insight as to the types of errors Department of Commerce statisticians made in attempting to forecast the current (i.e., zero-period-ahead) value for real GNP growth. The results are compatible with the hypothesis that the revisions contain both measurement and forecast error. The standard deviation of the 45-day release exceeds that of the final prebenchmark revised data, and both correlations are significant at less than the 0.02 level, even though there are only 13 recessionary observations. Consistent with a mean revision during recessions of -0.64%, the correlation between the 45-day figure and the revision is negative (-0.64%) and highly significant (p<0.004), implying an adjustment in the opposite direction. (See panel (9b)).

Just as with the private forecasters, we want to investigate to what extent this finding of systematic overforecasting of real GNP growth during recessions represents unavoidable forecast errors (i.e., from efficient forecasts) as opposed to informational inefficiency. Therefore, in panel (9b) we compute the same test statistics, this time conditional upon the economy being in recession the previous quarter, since that information seems to have been in the private forecasters' information sets. The mean forecast revision during recessions drops (in absolute value) to 0.69%; this indicates an increase in forecast accuracy. The standard deviation of the later release slightly exceeds that of the earlier release; this indicates an increase in forecast efficiency. Finally, both correlations fall in absolute value, and neither correlation is statistically significant. Most dramatic is the drop in magnitude and statistical significance of the correlation between the revision and the provisional estimate: lagging the recessionary observations one quarter reduces this correlation from -0.643 to -0.005 and increases the p-value from 0.004 to 0.989. Our inference is that the existence of one additional quarter of information about the stage of the business cycle reduces much of the measurement error and a sizable amount of the forecast error, as well. In other words, assuming that the current stage of the business cycle is not in the information set of government statisticians markedly improves our evaluation of the efficiency of their provisional estimates of real GNP growth. Importantly, the findings of this section are consistent with the results of the rationality tests on private (one-step-ahead) forecasts conducted in Section B above.

V. Conclusion

While informational efficiency can only be rejected in certain cases, never established once and for all, we conclude that the panel's forecasts are unbiased and generally efficient. The only near exception to efficiency that we find is with respect to previous periods' real consumption growth and the real return during recessions.

<table>
<thead>
<tr>
<th>Qtr of Recession</th>
<th>Variable</th>
<th>Sample Mean</th>
<th>T-Stat</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>First</td>
<td>$X_{1,t}$</td>
<td>-0.000</td>
<td>-25.72</td>
<td>128</td>
</tr>
<tr>
<td></td>
<td>$Y_{t}$</td>
<td>-0.001</td>
<td>-0.41</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$X_{1,t}Y_{t}$</td>
<td>-0.059</td>
<td>-18.47</td>
<td></td>
</tr>
<tr>
<td>Second</td>
<td>$X_{1,t}$</td>
<td>-0.012</td>
<td>-6.48</td>
<td>143</td>
</tr>
<tr>
<td></td>
<td>$Y_{t}$</td>
<td>-0.008</td>
<td>-2.74</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$X_{1,t}Y_{t}$</td>
<td>0.004</td>
<td>1.07</td>
<td></td>
</tr>
<tr>
<td>Third</td>
<td>$X_{1,t}$</td>
<td>0.000</td>
<td>0.14</td>
<td>97</td>
</tr>
<tr>
<td></td>
<td>$Y_{t}$</td>
<td>0.020</td>
<td>8.22</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$X_{1,t}Y_{t}$</td>
<td>-0.020</td>
<td>-6.12</td>
<td></td>
</tr>
<tr>
<td>Fourth</td>
<td>$X_{1,t}$</td>
<td>-0.008</td>
<td>-0.91</td>
<td>117</td>
</tr>
<tr>
<td></td>
<td>$Y_{t}$</td>
<td>0.020</td>
<td>4.32</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$X_{1,t}Y_{t}$</td>
<td>-0.028</td>
<td>-4.49</td>
<td></td>
</tr>
<tr>
<td>Sixth</td>
<td>$X_{1,t}$</td>
<td>0.025</td>
<td>28.38</td>
<td>70</td>
</tr>
<tr>
<td></td>
<td>$Y_{t}$</td>
<td>0.007</td>
<td>1.06</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$X_{1,t}Y_{t}$</td>
<td>0.018</td>
<td>3.18</td>
<td></td>
</tr>
<tr>
<td>Reccession (All qtrs)</td>
<td>$X_{1,t}$</td>
<td>-0.022</td>
<td>-11.00</td>
<td>671</td>
</tr>
<tr>
<td></td>
<td>$Y_{t}$</td>
<td>0.009</td>
<td>6.07</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$X_{1,t}Y_{t}$</td>
<td>-0.031</td>
<td>-14.82</td>
<td></td>
</tr>
<tr>
<td>Nonrecession</td>
<td>$X_{1,t}$</td>
<td>0.043</td>
<td>57.22</td>
<td>1830</td>
</tr>
<tr>
<td></td>
<td>$Y_{t}$</td>
<td>0.036</td>
<td>51.48</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$X_{1,t}Y_{t}$</td>
<td>0.007</td>
<td>8.59</td>
<td></td>
</tr>
</tbody>
</table>

Table 2

<table>
<thead>
<tr>
<th>Panel</th>
<th>Test</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1)</td>
<td>$H_0: \beta = \gamma = 0$</td>
<td>0.1143</td>
</tr>
<tr>
<td></td>
<td>$x^2$</td>
<td>0.565</td>
</tr>
<tr>
<td></td>
<td>(0.017)</td>
<td>(0.414)</td>
</tr>
<tr>
<td>(2)</td>
<td>$H_0: \beta = \gamma = 0$</td>
<td>0.036</td>
</tr>
<tr>
<td></td>
<td>$x^2$</td>
<td>0.007</td>
</tr>
<tr>
<td></td>
<td>-0.033</td>
<td>0.632</td>
</tr>
<tr>
<td></td>
<td>(0.008)</td>
<td>(0.014)</td>
</tr>
<tr>
<td>(3)</td>
<td>$H_0: \beta = \gamma = 0$</td>
<td>0.003</td>
</tr>
<tr>
<td></td>
<td>$x^2$</td>
<td>0.007</td>
</tr>
<tr>
<td></td>
<td>51.48</td>
<td>0.003</td>
</tr>
</tbody>
</table>

435
### Notes for Table 2

\( y_{t+1} \) is the annualized growth rate of real GNP from quarter \( t \) to quarter \( t+1 \). Unless indicated otherwise, \( y_{t+1} \) is the most revised vintage available before the benchmark revisions of 1975:IV, 1980:IV, and 1985:IV. The sample period is from 1968:IV to 1989:4, except for the exclusion of the three quarters in which benchmark revisions occurred. \( \gamma^* \) is the annualized growth rate of the real GNP forecast made at time \( t \), corresponding to \( y_{t+1} \). Unless indicated otherwise, coefficients are estimated via OLS. Standard errors calculated from GMM covariance matrices are given in parentheses below the \( \beta \)'s. Similarly, the \( p \)-values are calculated using test statistics incorporating GMM covariance matrices. The \( p \)-value represents the probability of obtaining a \( \chi^2 \) value at least as high as the reported value, if the null hypothesis is true.

I thank Shirley Yee of the University of Hawaii Computing Center for expert programming assistance.

### References


