DATA REVISIONS ARE NOT WELL-BEHAVED

S. Borağan Aruoba*

University of Maryland

Department of Economics

May 19, 2005

Abstract

We document the empirical properties of revisions to major macroeconomic variables in the United States. Our findings suggest that they do not satisfy simple desirable statistical properties. In particular, we find that these revisions do not have a zero mean, which indicates that the initial announcements by statistical agencies are biased. We also find that the revisions are quite large compared to the original variables and they are predictable using the information set at the time of the initial announcement, which means that the initial announcements of statistical agencies are not rational forecasts. We also provide evidence that professional forecasters ignore this predictability.

Key Words : Forecasting, news and noise, real-time data, NIPA variables

JEL Codes : C22, C53, C82

*Department of Economics, University of Maryland, College Park, MD 20742. Email: aruoba@econ.umd.edu

1 Introduction

Most macroeconomic variables are substantively revised by statistical agencies in the subsequent months after their initial announcements. These revisions generally reflect the arrival of new information which wasn't available at the time of the initial announcement. At any given time the revisions obviously refer to past information, i.e. to information based on which the agent made decisions in the past. One might think that since the agent can't change the decision he took in the past, these revisions may seem irrelevant for the agent's current decision problem. This is true if revisions are well-behaved.

To facilitate the discussion, we will use the following notation. Let y_t^{t+1} denote a statistical agency's initial announcement of a variable that was realized at time t and y_t^f denote the *final* or *true* value of the same variable. The two objects will be related by the following identity

$$y_t^f = y_t^{t+1} + r_t^f$$

where r_t^f is the final revision which can potentially be never observed.

From a statistical point of view, we would expect the final revision to satisfy three properties in order to consider them well-behaved. First, we expect its mean to be zero. This would imply that the initial announcement of the statistical agency is an unbiased estimate of the final value. Second, we expect the variance of the final revision to be small, compared to the variance of the final value. Finally, we expect the final revision to be unpredictable given the information set at the time of the initial announcement. When the final revision is predictable, the initial announcement of the statistical agency is not an optimal forecast of the final value. We summarize these three properties as follows:

$$
\begin{aligned}
(P1): &\quad E\left(r_t^f\right) = 0 \\
(P2): &\quad var\left(r_t^f\right) \text{ is small} \\
(P3): &\quad E\left(r_t^f | I_{t+1}\right) = 0
\end{aligned}
$$

where I_{t+1} is the information set at the time of the initial announcement. Our goal in this paper is to investigate the validity of these properties for revisions to macroeconomic variables in the United States.

We are certainly not the first to analyze the statistical properties of data revisions. Indeed, revisions to macroeconomic data are well understood by economists and have been studied for decades. An important part of the literature on data revisions considers the question we devote most of this paper to, the predictability of data revisions. Mankiw et al. [1984] assess whether the preliminary announcements of money stock are rational forecasts of the final announcements (news hypothesis) or are observations of the revised series, measured with error (noise hypothesis). A similar analysis was applied to GNP data by Mankiw and Shapiro [1986] (henceforth MS). The conclusion from these two studies is that while the revisions to GNP can be considered to be news, those of money stock data are better characterized as noise. In other words, they found evidence of predictability for the revisions to the money stock data while revisions to GNP data seems to be unpredictable. Mork [1987] and Mork [1990] consider the same question and find predictability in both GNP and money stock revisions using a slightly different methodology.

In a recent paper Faust et al. [forthcoming] look at the revisions to the GDP growth rates for the G-7 countries and find that while for the United States, revisions are very slightly predictable, for Italy, Japan and United Kingdom, about half the variability of subsequent revisions can be accounted for by information available at the time of the preliminary announcement by using methods similar to Mankiw et al. [1984] and MS.

Using the revisions to a variety of important macroeconomic variables, we find strong evidence against the three properties outlined above. In particular, we find that the unconditional mean of revisions are positive for all variables – significantly so for a majority of them. Moreover, we find that variance of the revisions are quite large compared to the variance of the original data series. We also show that the zero forecast implied by $(P3)$ can be improved significantly in an ex-post forecasting exercise and we demonstrate that the predictability found in the ex-post exercise can also be exploited for a number of variables in real time. We find that these results are robust in subsamples, if not stronger in the more recent periods. That is, we find a bigger bias, a larger variability in revisions and a larger degree of predictability in periods which coincide with the decline in volatility that is well-documented for the U.S. economy.

In an effort to investigate how much of our findings are understood by the private sector, we analyze the responses in the Survey of Professional Forecasters. We find that these responses are consistent with revisions having a zero mean and being unpredictable, contrary to our findings.

The rest of this paper is organized as follows. In Section 2 we describe the data used in the paper. In Section 3 we report the unconditional properties of revisions, investigating the validity of (P1) and (P2). In Section 4 we turn to predictability of revisions and consider the validity of (P3). In Section 5 we summarize our results from different subsamples. We also provide a simple example showing the importance of data revisions. In Section 6 we discuss our findings from the Survey of Professional Forecasters regarding the forecasters' perceptions of data revisions. We conclude in Section 7. A technical appendix provides details of the analysis and some additional results.

2 Data

2.1 Data Sources

In order to conduct our empirical analysis, we need to have observations of data revisions. This would require observing both the current or revised version of the data as well as the data as it was initially announced by statistical agencies. The latter was not readily available to researchers until recently since the statistical agencies typically report only the most recent revised version. Researchers could only go to a library and collect the historical data from publications.

In 1999, the Federal Reserve Bank of Philadelphia released and made publicly available the "Real-Time Data Set" (RTDS)[1] which records the information set that would be available to an agent on the 15^{th} day of the middle month of a quarter starting from the last quarter of 1965. The RTDS includes quarterly data for major National Income and Product Account

(NIPA) variables such as real and nominal output, consumption, investment and their subcategories, monetary measures, banking system data, price level and unemployment rate. It also includes monthly data on capacity utilization, industrial production and employment.

The majority of the analysis in this paper is based on data from the RTDS. Our primary analysis will focus on two original NIPA variables (nominal and real output)[2], six variables derived from them (growth of real output, nominal output and inflation based on output deflator, annual and quarterly), unemployment rate and levels and growth rates of employment, capacity utilization and industrial production. In Section 5.2 we also report our results about the revisions to the components of real output (both in levels and growth rates) in order to understand which components are responsible for the results we report in the paper regarding revisions to real output.

We also put together a small-scale real-time data set for this paper using labor productivity as announced by the Bureau of Labor Statistics (BLS) in the Monthly Labor Review (MLR) covering 1971-2004. This data set, similar to the RTDS, consists of observations of the quarterly and annual growth rates of labor productivity as they are observed in the middle month of each quarter.[3]

There are two dimensions that we can use to categorize our data. First, we group the variables according to their sampling frequency and as explained above, we have a mixture of monthly and quarterly variables. We also group them according to their units of measurement. Variables like nominal output or employment grow over time and so do the revisions to these variables. Therefore, when we analyze the revisions to these variables we use the revisions relative to the initial announcements and refer to such variables *relative* variables in what follows. On the other hand, variables such as growth of output or capacity utilization are expressed in percentage units and will not generally have a nonstationarity problem. Accordingly, we use the levels of the revisions to these variables and refer to them as *level* variables.

2.2 Defining the Revisions

The raw data that we obtained from the RTDS or from our labor productivity data set is a two-dimensional object for each variable where for variable y at time t, we have observations starting in $t+1$ and continuing until the end of the sample, T. We use the notation y_t^s to represent the value of the variable y of time t, as announced at time s, where $s \geq t+1$ since by convention $t+1$ is the first date the time t variables are announced. Using this notation, y_t^{t+1} denotes the initial announcement of the statistical agency.

We can define four different revision concepts that will be relevant for our analysis:

$$
\begin{array}{ll}
\text{Cumulative Revision after } h \text{ periods :} & r_t^{(1)h} = y_t^{(t+1)+h} - y_t^{t+1} \\
\text{Relative Cumulative Revision after } h \text{ periods :} & r_t^{(2)h} = \dfrac{r_t^{(1)h}}{y_t^{t+1}}
\end{array}
$$

The first object, $r_t^{(1)h}$, is the revision to the time t variable made at $t+1+h$ relative to its initial announcement, y_t^{t+1}. It measures the cumulative revision up to time $t+1+h$, which is h periods after the initial announcement. $r_t^{(2)h}$ is the cumulative revision at horizon h relative to the initial announcement. We focus on $r_t^{(1)h}$ for *level* variables and on $r_t^{(2)h}$ for *relative* variables so that the revisions we consider will be stationary. We use r_t^h to refer to $r_t^{(1)h}$ for revisions to *level* variables and $r_t^{(2)h}$ for revisions to *relative* variables to save notation.

2.3 "Uninformative" Revisions

We can broadly divide revisions into two categories: *informative* and *uninformative* revisions. The former carry "informational content", i.e. they reflect the incorporation of the new information received by the statistical agency which were not available previously. The latter change the definition of the variable or make statistical changes such as the change of base year or seasonal weights. We will consider these revisions to be *uninformative* for the users of the data since they would not be able to extract information from this revision that they can compare with their old information set.[4] For our analysis we identify the uninformative revisions and remove them from the sample. The details are provided in the technical appendix.

2.4 Defining the Final Revision

Some previous papers define the final revision as the difference between the latest available observation for the variable and its initial announcement. But using this method might be wrong for *relative* variables because of the benchmark revisions that might have occurred since the first announcement of the variable. Also, for *level* variables, the analysis may also be distorted since the growth rates that go in to the calculation of the revision would come from data with different underlying statistical definitions. This would suggest, therefore, instead of using the latest available revision as the final revision, we should use the last *informative* revision. In other words, we want to include as many revisions as possible in our final revision in order to include all relevant informative revisions, but we want to avoid including any uninformative revisions.

In order to define the final revision we determine the numbers of periods after which there are no more informative revisions for each variable. For some variables such as the NIPA variables, the statistical agencies follow a very specific schedule for revisions which makes it very easy for us to define the final revisions. For other variables, we look at the incremental revisions at different horizons and find a pattern in revisions. Essentially, for each variable we find a finite number K, and define the K^{th} revision of the variable as the final revision. The details are provided in the technical appendix.

3 Unconditional Properties of Final Revisions

In this section we first consider whether the macroeconomic data revisions in the United States satisfy the first two of the three properties we listed in Section 1 by computing the relevant statistics and running some statistical tests.

In the first panel of Table I we report results for the *relative* variables, and in the second panel we report results for the *level* variables. The first column of Table I reports the number of observations for each variable. For quarterly variables we have about 35 years of data while for the monthly variables we have between 16 and 37 years of data. The next column reports the mean of the final revision for each variable. We use Newey-West Heteroskedasticity and Autocorrelation Consistent standard errors (Newey and West [1987]) in computing the test of significance for these means due to the apparent autocorrelated structure of revisions. The results indicate that the mean of final revisions for all variables are positive and except for four variables (quarterly growth of labor productivity, unemployment rate and two different measures of capacity utilization) they are statistically different from zero.[5] The interpretation of this result is that the initial announcements of the statistical agencies are biased estimates of the final values. In addition to being statistically significant, the means of final revisions are quite large. Level of nominal GDP is revised upwards by 0.60% of its initial announcement, which is an average revision of about $67 billion in terms of 2004 prices. While the mean relative revision of real output and employment are slightly lower, it is close to 1% for industrial production index. As for *level* variables, the numbers range from virtually zero to 1.2%. It is worth noting that the average revision for real output growth is between 20-25 basis points, considering that average growth rate of real output in this period is about 2%. We can conclude that there is strong evidence against $(P1)$, i.e. the revisions do not have a zero mean.

The next two columns report the minimum and maximum final revision for each variable. We see that the range of final revisions, both for *relative* and *level* variables are quite large. For example, the final revision of annual real output growth fluctuates between -1.6% and 2% while the final revision of annual labor productivity growth fluctuates between -3.1% and 3.3%. The only possible exception is the final revision to unemployment rate which only fluctuates between -0.2% and 0.2%, which is consistent with the observation that the revisions to the unemployment rate are small and confined to changes in seasonal factors.

Next, we report the standard deviation of final revisions. Since the standard deviation of final revisions by itself may not be very informative of the size of final revisions, we also report the noise-to-signal ratio for final revisions for *level* variables, which is defined as the standard deviation of final revisions divided by the standard deviation of the final value of the variable.[6] This statistic, along with the minimum and maximum final revisions, will give us an idea about the size of final revisions relative to the size of the original variables. The numbers we find range from 0.13 to 0.94, excluding the unemployment rate. There are two extreme cases: the growth of labor productivity (0.82 and 0.94) and unemployment rate (0.05). As explained above, the latter is expected given the nature of the revisions to unemployment rate while the former is due in part to the strong negative correlation of the initial announcement and the final revision for the growth of labor productivity. Even after removing these three extreme variables, the average noise-to-signal ratio is 0.33. Such large numbers suggest that the final revisions are sizable compared to the original variables, and we conclude that $(P2)$ is not supported by the data. It is interesting to note that the signal-to-noise ratios for annual growth variables are about half of their counterparts for monthly or quarterly growth variables.

The next column reports the simple correlation of the final revision with the initial announcement for *level* variables. We compute asymptotic standard errors for these correlations (not reported) and test for their significance. It is not possible to talk about a general pattern in terms of sign of the correlations but most of the significant correlations are negative. They are as large as -0.5 and the average absolute correlation is 0.19. This is our first evidence that $(P3)$ may not be consistent with the data since the final revisions are correlated with the initial announcements. The next set of statistics will shed some further light on the issue and we take it up more rigorously in the next section.

The last two columns report the first order autocorrelation coefficients for final revisions and the p-values from a Q-statistic at 20 lags.[7] We also test the significance of the former by the appropriate Q-statistic (not reported). The final revisions to all relative variables and all annual growth variables show strong signs of persistence, with positive autocorrelation coefficients between 0.55 and 0.95, with an average of 0.77. Similar results hold for both definitions of capacity utilization. On the other hand, the persistence of the revisions to the quarterly and monthly growth variables is quite weak. One explanation of the persistence in revisions is the particular schedule that revisions follow. We often see revisions effecting a number of consecutive periods announced on the same date, therefore using the same information set. If a common information shock causes the revisions to the variable in these periods, the final revisions will appear correlated. While the apparent persistence in final revisions suggests the possibility of their predictability, this cannot be used as direct evidence to that effect. The autocorrelated structure documented here cannot be exploited to provide a forecast of r_t^f, since r_{t-1}^f is not realized until $t+K$ and thus is not in the information set of $t+1$.

To summarize our results from Table I, we find that the mean final revision is positive and statistically significant for most of the variables and final revisions are large relative to the original variables. We also have some evidence that suggests predictability of revisions.

4 Forecastability of Final Revisions

Having analyzed the unconditional properties of data revisions in the previous section, we now turn to investigating the validity of (P3), which states that the revisions must be unpredictable given the information set at the time of the initial announcement. We start our analysis by revisiting an old methodology which attempts to label data revisions as “news” or “noise”. Next we conduct two forecasting exercises, an ex-post exercise which looks at the predictability of final revisions using the full sample and a real-time exercise which attempts to mimic the forecasting problem of a user of statistical data who is trying to forecast final revisions in real time.

4.1 News vs. Noise Revisited

Two of the most important papers in the literature that analyze the nature of the revisions to macroeconomic variables is MS and Mankiw et al. [1984] where the authors analyze whether the preliminary announcements of GNP and money stock are rational forecasts of the true, or “final” announcements. In this section we replicate some of their analysis for our new (and longer) data set in order to provide a comparison between results from our new data set and the old and well-known results.

In the framework of the aforementioned papers, final revisions can classified into two categories:

- **Noise**: The initial announcement is an observation of the final series, measured with error. This means that the revision is uncorrelated with the final value but correlated with the data available when the estimate is made (e.g. preliminary announcement).

- **News**: The initial announcement is an efficient forecast that reflects all available information and subsequent estimates reduce the forecast error, incorporating new information. The revision is correlated with the final value but uncorrelated with the data available when the estimate is made, i.e. unpredictable with using the information set at the time of the initial announcement.

To classify revisions as noise or news, they consider the regressions

$$y_t^{t+1} = \alpha_1 + \beta_1 y_t^f + \nu_t^1 \tag{1}$$

$$y_t^f = \alpha_2 + \beta_2 y_t^{t+1} + \nu_t^2 \tag{2}$$

where the joint hypothesis $\alpha_1 = 0$, $\beta_1 = 1$ would test the noise hypothesis, and the joint hypothesis $\alpha_2 = 0$, $\beta_2 = 1$ would test the news hypothesis. As can be easily shown, these hypotheses are mutually exclusive but, they are not collectively exhaustive, that is, we can reject both hypotheses, especially when the unconditional mean of revisions is not equal to zero.[8] In this case, we can reject both hypotheses and there is no guidance in the original MS methodology when this happens. Using this framework, they conclude that the revisions to GNP (both as level in constant dollars and growth in current dollars) can be considered to be news and those of money stock data are better characterized as noise, since they reject one and fail to reject the other hypothesis in each case.

Using a particular subsample of our data set, we are able to replicate the results of MS, that is we reject the noise hypothesis and fail to reject the news hypothesis for real output growth. However, this conclusion is not robust, even within the same subsample. If news hypothesis was true, that is if revisions were errors from a rational forecast, then any other explanatory variable that was observed at the time of the initial announcement included in (2) should have a coefficient of zero. When we add r^3_{t-3}, which is announced at time $t+1$, to the right hand side of (2) we find that the estimated coefficient of r^3_{t-3} is statistically significant and, more importantly, the F-test with null hypothesis setting all coefficients to zero is now rejected. Therefore, this small change, which simply follows from the statement of the news hypothesis, leads to the rejection of the news hypothesis as well.

Next, we repeat the same analysis for all level variables using the longest available sample for each variable. We find that for all but three variables we reject the noise hypothesis and for all but one variable we reject the news hypothesis. Therefore using only the original MS methodology we are able to reject both hypotheses and thus are unable to classify revisions as neither optimal forecast errors or measurement errors for all but three variables. On the other hand, we reach an equally ambiguous conclusion for unemployment rate where we fail to reject both hypothesis. When we look at the source of the rejection of both hypotheses, we see that in almost all regressions the constant is statistically significant, and mostly positive. This is of course related to the observation from Table I that all revisions have positive means.

To sum up, we find that the original MS results are very special since introducing a small variation in the methodology or looking at a longer sample[9] reverses the results.

4.2 An Ex-Post Forecastability Exercise

In this section we turn to testing if $(P3)$ is supported by the data, that is, if the conditional mean of final revisions with respect to the information set at the time of the initial announcement is zero. This will be identical to testing the news hypothesis since both are a restatement of efficiency (rationality) of preliminary announcements since rational forecast errors must be orthogonal to the information set at the time of the forecast.

To that end, we estimate the following equations. For *relative* variables, we estimate

$$r_t^f = \alpha + \sum_{i=1}^{s} \beta_i r_{t-i}^{(2)i} + I_Q \sum_{i=1}^{4} \lambda_i Q_t^i + \delta t + \varepsilon_t \tag{3}$$

where the dependent variable, r_t^f is the final revision relative to the initial announcement and the explanatory variables are a constant, relative revisions to past months or quarters announced at time $t+1$, quarterly dummy variables, Q_t^i and a linear trend. The indicator I_Q is one for quarterly variables and zero for monthly variables as we consider seasonality for only quarterly variables in an effort to limit the number of coefficients estimated. For *level* variables we estimate

$$r_t^f = \alpha + \gamma y_t^{t+1} + \sum_{i=1}^{s} \beta_i r_{t-i}^{(1)i} + I_Q \sum_{i=1}^{4} \lambda_i Q_t^i + \delta t + \varepsilon_t \tag{4}$$

where the only differences are that we use the level of revisions, $r_t^{(1)h}$ and we add the initial announcement as an explanatory variable. For NIPA variables we choose $s = 10$, for all monthly variables we use $s = 14$ and for labor productivity we use $s = 6$.

Except for the presence of past revisions as explanatory variables, these equations are very similar to forecasting equations considered in similar studies that analyze the predictability of revisions. We include these revisions to analyze the predictive power of past revisions in explaining future revisions. We also include seasonal dummies in our estimations since there might be some seasonality in the final revisions due to the specific revision schedules of statistical agencies, even though the original series might be deseasonalized.

It is important to note that all explanatory variables, including past revisions, are chosen such that they are all known at time $t+1$. However, since r_t^f is not observed until $t+K+1$, this exercise cannot be implemented in real time. In other words, we would not be able to exploit the forecastability we might find in this section in real time. Nevertheless this is still a valid forecasting exercise as all explanatory variables are measurable at time $t+1$.

We want to stress that by estimating this equation we are not trying to find the best model for revisions. If that were the case, one would imagine many other variables potentially being relevant, or a multivariate analysis would be warranted.[10] Our aim by estimating these equations is to show that we can find *a* forecasting model that can perform better than the model implied by (P3), one that has a zero conditional mean.

We conduct the exercise using the following algorithm. For each variable, we estimate the relevant equation by considering all possible combinations of explanatory variables.[11]

Using both Akaike Information Criterion (AIC) and Schwartz Information Criterion (SIC) as a guide we choose the best model for each variable and label this model as Model 1. Using the parameter estimates of this model we get the fitted value of r_t^f which we denote as $\hat{r}_t^1$.

To understand the marginal contribution of the past revisions and the initial announcement to forecasting the final revision, we eliminate these terms from the model and re-estimate the simple linear regression with only seasonal dummy variables. We label this model Model 2 and denote the fitted value as $\hat{r}_t^2$. Finally, we consider the forecast of r_t^f based on (P3) and define this case as Model 3 with the forecast given by $\hat{r}_t^3 = 0$ for all t.

Given the forecasts from the three models, we conduct two tests. First, along the lines of our test of rational data revisions, we test for the joint significance of all coefficients in (3) and (4). This test will essentially have Model 3 or (P3) as its null hypothesis. We also compare the predictive powers of $\hat{r}_t^1$ and $\hat{r}_t^2$ versus $\hat{r}_t^3$. In order to do so, we compute the root mean squared errors (RMSE) of forecasts from Model 1 and 2, relative to the RMSE of the forecast from Model 3. To the extent that the relative RMSE is less than one, Model 1 or Model 2 provides a better forecast.[12]

The results from this exercise are summarized in Table II. In the forth column, we list the explanatory variables Model 1 as picked by SIC and AIC. Almost all models picked by the two criteria include at least one past revision which demonstrates the importance of including these variables in predictive regressions. Interestingly the linear trend is important for 11 of the 22 variables we consider, even though the dependent variables are stationary. This suggests a potentially time-varying pattern in revisions and we take up this issue in

Section 5.1.

The fifth and sixth columns report the R^2 and the adjusted R^2 $\left(\bar{R}^2\right)$ for each regression. The R^2's range from zero (none of the explanatory variables except the constant are relevant) to 0.37 for industrial production index for manufacturing. The average R^2 for all variables is about 0.14 and the average $\bar{R}^2$ is about 0.12. For important variables such as real output growth, inflation and labor productivity growth, the R^2's are 0.16, 0.07 and 0.23, respectively. These numbers may not seem too large in other regression contexts, but considering the widespread belief that revisions are unpredictable, they are extremely large.

Next column reports the p-value of the Wald statistic testing the significance of all coefficients in the regressions. All p-values are zero, indicating that we can reject this null hypothesis at any conventional significance level. In the terminology of the previous section, this means a rejection of the news hypothesis for all of the variables we consider.[13]

The last two columns report the RMSE of Model 1 and Model 2, relative to Model 3. All relative RMSE's are less than unity indicating that our forecasting models perform better than a zero forecast for all variables. The average relative RMSE is 0.88 and 0.87 for Model 1 and Model 2, respectively, which mean that on average our forecasting models provide a $12 - 13\%$ improvement over the zero forecast.

In Figure I we plot the final revision, r_t^K, and the fitted values from the ex-post forecasting exercise in this section and the forecasted values from the real-time exercise from the next section for Nominal Output and Quarterly Real Output Growth. In each panel we also show the zero line (solid) and the unconditional mean of the final revisions (dotted). The fitted values from the ex-post exercise are given by the thick dotted lines, and the forecasts from the real-time exercise are shown by the dotted lines. We see that the fitted values from this exercise is quite a good forecast of revisions as they track them closely.

To sum up our findings from this ex-post forecastability exercise, we find that using a very limited information set that is known at time $t+1$, we are able to predict the final revision that will be realized at $t+K+1$. Using three different statistics, goodness-of-fit, a Wald test and relative RMSE, we find that the forecasting model we estimate performs significantly better than a zero forecast (P3) would imply. We conclude that (P3) is not supported by the data and that the initial announcements of statistical agencies are not rational forecasts of the true value of variables.

4.3 A Real-Time Forecastability Exercise

In this section, we conduct a very simple real-time forecasting exercise. Our goal here is not to find the best model to forecast revisions in real time but rather to demonstrate that even a simple scheme can produce a better forecast than a zero forecast in real time.[14]

In this exercise, at every period we simply look at the time series of final revisions for each variable that has been realized, and we compute their mean. We use this mean as our forecast of this period's revision. In other words, our forecast is

$$\hat{r}_t^{RT} = \mu_{t-K}$$

where μ_s is the sample mean for $\left\{r_1^K, ..., r_s^K\right\}$. To assure parameter stability, we start forecasting in the second halves of the relevant samples for each variable.

The results from this exercise are reported in Table III. We find that for 13 variables out of the 22 variables, the RMSE of the real-time forecast is lower than the RMSE of the zero forecast as evidenced by a relative RMSE of less than unity. Since we are considering out-of-sample forecasts, we also compute the Diebold Mariano [1995] (henceforth DM) statistic which has the null hypothesis of equal forecast accuracy between two forecasts. We define

$$d_t = \left(r_t - \hat{r}_t^{RT}\right)^2 - \left(r_t - 0\right)^2$$

as the difference between the loss functions of the two forecasts where we choose the squared error as the loss function. The DM test amounts to testing the significance of the mean of d_t.[15] The DM test results show that out of the 12 variables where the real-time forecast was better than the zero forecast, for two variables (annual and monthly growth of industrial production in manufacturing) the difference between forecast accuracies is statistically significant. It is important to note that the zero forecast is not statistically better than this simple real-time forecast for any of the variables.

In Figure I, we plot the forecast from this exercise for revisions to nominal output and quarterly output growth. We see that the forecast is very close to the unconditional mean of revisions and given that the latter is significantly different from zero, it is not surprising that the former is better than a zero forecast.

Due to the extremely simple nature of the forecast, it is not possible to capture all the dynamics in final revisions in this exercise. However, we interpret these results as suggestive of potential gains of forecasting revisions, even in real time.

5 Sensitivity Analysis and Further Results

In this section, we summarize our results from the sensitivity analysis where we repeat most of the analysis carried out in the previous sections in subsamples. We also analyze the revisions to the components of output. Finally, we provide an example that show the relevance of data revisions for policy.

5.1 Subsamples

For each variable, we divide the sample in to two and repeat the analysis in the subsamples. It may be the case that the three desirable properties of data revisions, (P1), (P2) and (P3) may not hold in the full sample due to some problems early on in our full sample and if we consider only the second half of our sample, we may find that they hold. This may be the case, for example, as a result of improvements in data collection due to technological progress. However, another equally plausible argument is that technological progress makes data collection harder due to increased variety of goods. This would suggest that as the statistical agencies are struggling to make the necessary corrections, they might create revisions which do not satisfy these three properties. The detailed results are provided in the technical appendix.

We find that the means of final revisions for all variables are positive in both halves of their samples, while the mean revision for the first half of the sample is bigger than that of the second half for 14 out of 22 variables. On the other hand, the noise-to-signal ratio for all variables increase in the second half of their samples, indicating that the statistical agencies make larger revisions. It is interesting to note that the second halves of samples roughly correspond to the period where real economic activity in the U.S. is much less volatile. (See, for example, Stock and Watson [2003]) Analyzing the link between these two observations, if any, would be an interesting exercise. Finally, for all but five variables, final revisions are more persistent in the second half of the sample.

In the ex-post forecasting exercise, we find that while all qualitative results are valid for both subsamples, we find evidence of increased predictability for revisions in the second half of the sample.

We conclude that the failure of the three properties (P1), (P2) and (P3) we documented in the full sample is not necessarily due to a certain part of the sample. However, we find increased evidence against these properties in the second half of the sample which lends support for the second view about the effect of technological progress on the quality of data described above.

5.2 Components of Real Output

We repeat our analysis for components of real output in order to identify the source of the results we find for revisions to real output. We find that consumption and exports are the two components of output that has a significantly positive mean for revisions. When we break up consumption further, durables consumption seems to be the component that is responsible for a significant part of this finding. Analyzing the revisions to the growth rates of major components of real output, we get similar results. We also find that all components of output have a larger noise-to-signal ratio than output itself. This shows that the large revisions to real output are not caused by one or two components. Similarly, revisions to all components have either no correlation or a significant negative correlation with the initial announcements and these correlation coefficients are generally larger than that of output. Finally, revisions to all components of output show significant positive autocorrelation, except for revisions to quarterly growth variables which show no persistence.

In the ex-post forecasting exercise, we find that, with just a few exceptions, revisions to all components of output are predictable at least as much as output itself and most components are more predictable than output, as measured by the magnitudes of R^2's and relative RMSEs.

Overall, our results from this section indicate that the failure of (P1), (P2) and (P3) for revisions to real output is not entirely due to one or a few of its components but rather a general phenomenon which is valid for almost all of its components. Consumption, in particular durables consumption seems to be the component that contributes most to these results. This result is quite significant given the debate concerning measurement of consumer electronics and similar goods whose quality changes quite remarkably in short amounts of time.[16] Our results are at least suggestive that the revisions to components of output which are arguably harder to measure contribute to the results we find in this paper regarding revisions to output.

5.3 Business Cycles and Final Revisions

In Figure II, we plot the final revisions of annual real output growth and annual inflation, with shaded areas to reflect the recessions in the United States as determined by the National Bureau of Economic Research (NBER). It is interesting to note that the final revision to real output growth are negative (positive) before (after) recessions for all five recessions in our sample.[17] This means that the initial announcement is overstated before recessions and understated after recessions. In other words, looking at the initial numbers, it is harder to realize the decline in real output in real time at the beginning of the recessions (and accordingly label the period as a recession) and similarly to realize the increase in real output at the end of recessions.[18] This fact may have big implications for policy. If we look at a simple Taylor [1993] rule for monetary policy, keeping the level of inflation constant it prescribes a cut in interest rates when real output is below potential output and an increase in the interest rates in the opposite case. If the central bank has growth data which are overstated before recessions and understated after recessions in its information set, then it will be less aggressive in cutting interest rates before recessions and it will be late in increasing interest rates to reduce inflationary pressures after the recessions.

The plot of final revisions for inflation tells a similarly interesting story. Before all five recessions, final revision to inflation is positive which indicates that initial numbers are understated. Moreover, there are periods of extended positive final revisions during expansions, especially between the 1974 and 1980 recessions and the second half of the period between the 1983 and 1991 recessions. If the central bank is worried about inflation during expansions since inflation is understated during the periods listed above, it will be slow to increase interest rates.

Understanding where the economy stands is undoubtedly of great interest for policy makers. These two examples show the importance of data revisions for policy.

6 Results from the Survey of Professional Forecasters

So far we showed that revisions to some of the major macroeconomic variables in the United States don't have a zero mean, they are quite large and they are predictable. Next, we want to understand how much of these findings are understood by the private sector. To that end, we turn to the Survey of Professional Forecasters (SPF).

The SPF is a quarterly survey conducted since 1968, first by the American Statistical Association and NBER and after 1990 by the Federal Reserve Bank of Philadelphia.[19] In the survey carried out at time $t+1$, among other questions regarding the future, the forecasters are given the initial announcements of the variables for time t and asked to change the numbers if they so desire. They are, in essence, asked to forecast the final revision at the time of the initial announcement, i.e. to compute $E\left(r_t^f | I_{t+1}\right)$.

We use three *relative* variables that we used in our empirical analysis (real output, nominal output and real consumption expenditures) in the previous sections along with the level of the Price Index based on GDP deflator. The choice of these variables is due to data restrictions since the questions in the relevant part of the survey were in terms of levels and not growth rates. However, note that if the forecasters don't believe that there is a revision in the *level* of the variable, then their forecast of the growth of the variable will be identical to that of the BEA, implying no revision in the *growth* of the variable.

The first row of Table IV reports the fraction of periods where the median forecast reported in the SPF is equal to the initial announcement of the BEA, which are given to the forecasters. This is the case for more than 85% of periods for all 4 variables, with even stronger results in the post-1990 era. Next, we consider the fraction of forecasters who report a forecast that is within one point of the initial announcement.[20] In the second row of Table IV, we report the average of this fraction. In the pre-1990 era, about 70% of forecasters report a very small deviation from the initial announcement. In the post-1990 era, this average fraction is about 95%. We also report the fraction of forecasters that report a revision with the correct sign, that is the same sign as the final revision as we defined in the previous sections. The average fraction is about 30% in the pre-1990 era and slightly higher in the post-1990 era. Finally, we compute the mean revision reported by the forecasters, averaged over time and over forecasters, as percentage of the initial announcement. Note that this corresponds to the unconditional mean of final revisions we report in Table I. The forth row of Table IV reports our results from SPF and we report the results from the RTDS for the relevant time period for comparison on the next row. The numbers we get from the SPF are very small and for most of the variables in the post-1990 sample they are negligible. These numbers are clearly at odds with the results we compute from the data.

Overall, the results from SPF suggest that the forecasters' responses are consistent with $(P1)$ and $(P3)$, that is the revisions have a zero mean and they are not forecastable.

7 Conclusion

Macroeconomic data revisions are innocuous if they are well-behaved. In this paper, we postulate three properties that we expect these revisions to satisfy and we find that none of them is satisfied. In particular, we find that the means of final revisions are not zero, indicating that the initial announcements of statistical agencies are biased. We also find that the magnitudes of revisions are quite large compared to the original variables. We further show that the forecast from a forecasting equation is significantly better than a naive zero-forecast, which would be optimal if initial announcements of statistical agencies are optimal forecasts of the final values.

We repeat our analysis for two subsamples and find that while all the findings go through in both samples, the evidence against the three properties seems to be stronger in the second half of the sample. This finding is consistent with the view that technological progress makes collecting data harder due to the difficulty in adjusting the quality of goods in the economy. Another piece of evidence that supports this view is that revisions to durables consumption seem to be the source of the problem for the results we get regarding the revisions to real output.

Finally, we turn to the Survey of Professional Forecasters to analyze the perception of the private sector regarding the three desirable properties of data revisions. We find significant evidence that suggests that the respondents of the survey believe that these properties are indeed satisfied by the data since their responses imply a zero unconditional and conditional mean for revisions.

We do not wish to interpret the findings in this paper as failures of the statistical agencies. We believe that these institutions have certain loss functions and use their resources for producing the best possible data and they may be avoiding some other problems at the expense of the problems we outline in this paper.[21] In fact, a recent paper by economists at the BEA (Fixler and Grimm [2002]) analyzes the reliability of NIPA data for the period 1983-2000 and they report mean revisions that are close to those we find in this paper and yet they conclude, without any statistical justification, that these are not statistically significant. As for forecastability, they only consider forecasting a vintage of the data using an earlier vintage, and they conclude that revisions are not predictable. Therefore we have some evidence that the BEA does not think that the fact that these three properties are not satisfied by the data is an important problem.

In our analysis, we explicitly ignored the information set and decisions of the policy makers. Although we showed that private agents seem to be unaware of the bias and predictability of revision, it will be hard to reach the same conclusion for major economic policy makers such as the Federal Reserve. It is a well known fact (e.g. Romer and Romer [2000]) that the forecasts of the Federal Reserve are consistently better than those of private forecasters. While this may be due to informational superiority (having access to a larger set of data), explicit consideration of revisions may very well be responsible. A rigorous analysis in this direction will be the subject of future research.

Another interesting topic of future research is extending the forecastability analysis to a multi-variate framework. There are some interesting and unexpected cross-correlations between revisions to unrelated variables and it would be interesting to explore whether these correlations can be exploited to add to the predictability and forecastability results we obtain in this paper. Moreover, there might be some more expected links between revisions to related variables such as monthly industrial production and quarterly GDP.

Finally, one of the interesting observations from our analysis is the apparent concurrent reduction in the variance of major macroeconomic variables and the increase in the noise-to-signal ratios. The former is an important observation that has big implications for policy and economic research and any potential links between these two observations will be of interest to many.

References

[1] Conrad, W., and C. Corrado, "Application of the Kalman Filter to Revisions in Monthly Retail Sales Estimates," *Journal of Economic Dynamics and Control,* 1 (1979), 177-98.

[2] Croushore, D. and T. Stark, "A Real-Time Data Set for Macroeconomists: Does the Data Vintage Matter?", Federal Reserve Bank of Philadelphia Working Paper #99-21 (1999).

[3] Croushore, D. and T. Stark, "A Real-Time Data Set for Macroeconomists", *Journal of Econometrics*, 105 (2001), 111-130.

[4] Diebold, Francis X and Roberto Mariano, "Comparing Predictive Accuracy," *Journal of Business and Economic Statistics*, 13 (1995), 253-265

[5] Faust, J., J.H. Rogers and J. Wright, "News and Noise in G-7 GDP Announcements", *Journal of Money, Credit and Banking*, (forthcoming).

[6] Fixler, D.J and B.T. Grimm, "Reliability of GDP and Related NIPA Estimates", *Survey of Current Business,* 82 (2002), 9-27.

[7] Howrey, E. Phillip, "The Use of Preliminary Data in Economic Forecasting", Review of Economics and Statistics, 60 (1978), 193-200.

[8] Ljung, G. and G. Box, "On a Measure of Lack of Fit in Time Series Models," *Biometrika*, 66 (1979), 265-270.

[9] Landefeld, J.S. and B.T. Grimm, "A Note on the Impact of Hedonics and Computers on Real GDP," *Survey of Current Business*, (December 2000), 17-22.

[10] Mankiw, N.G., D.E. Runkle and M.D. Shapiro, "Are Preliminary Announcements of the Money Stock Rational Forecasts?", *Journal of Monetary Economics*, 14 (1984), 15-27.

[11] Mankiw, N.G. and M.D. Shapiro, "News or Noise: An Analysis of GNP Revisions", *Survey of Current Business*, 66 (1986), 20-25.

[12] Mork, K.A., "Ain't Behavin': Forecast Errors and Measurement Errors in Early GNP Estimates", *Journal of Business and Economic Statistics*, 5 (1987), 165-175.

[13] Mork, K.A., "Forecastable Money-Growth Revisions: A Closer Look at the Data," *Canadian Journal of Economics*, 23 (1990), 593-616.

[14] Newey, W. and K. West, "A Simple Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix," *Econometrica*, 55 (1987), 703–708.

[15] Romer, C.D. and D.H. Romer, "Federal Reserve Information and the Behavior of Interest Rates," *American Economic Review*, 90 (2000), 429-457.

[16] Stock, J.H. and M.W. Watson, "Has the Business Cycle Changed? Evidence and Explanations", Working Paper, (2003).

[17] Tanizaki, H. and R.S. Mariano, "Prediction, Filtering and Smoothing in Non-Linear and Non-Normal Cases Using Monte Carlo Integration," *Journal of Applied Econometrics*, 9 (1995), 163-179.

[18] Taylor, J.B., "Discretion versus Policy Rules in Practice," *Carnegie-Rochester Conference Series on Public Policy*, 39 (1993), 195–214.

Notes

[1] See Croushere and Stark [2001] for the details of the data set. An earlier version, Croushere and Stark [1999], provide some examples of empirical applications using this data set. The data set is publicly available on the internet at http://www.phil.frb.org/econ/forecast/reaindex.html. A bibliography of relevant papers as well as detailed documentation about the data set is also available from the same internet page.

[2] The RTDS uses GNP before 1992 and GDP afterwards, following the "headline variable" announced by the Bureau of Economic Analysis. As such, we will use the term "output" instead of GNP or GDP.

[3] Unlike the RTDS, which has access to "deep" data sources, each vintage in our data set has information about only approximately ten quarters preceding the date of the vintage.

[4] For example if we are following a variable which was 100 at its initial announcement, was revised to 102 the next quarter, and was revised to 150 following a base year change, we wouldn't be able to compare 150 with 102 since they are in two different scales.

[5]All statistical tests in this paper uses 10% significance. In some tables we also report the p-values for reference and where relevant mark the coefficients with p-values less than 10% with boldface.

[6]Note that this number is bounded below by zero but unbounded from above due to the possible correlation between r_t^f and y_t^{t+1}.

[7]Q-statistic (Ljung and Box [1979]) tests the null hypothesis that all autocorrelation coefficients up to a certain lag is equal to zero. The test statistic is a chi-square random variable with degrees of freedom equal to the number of lags.

[8]Note that all these statements are made in the population. Due to sampling errors, we can reject or fail to reject both hypotheses in small samples.

[9]One factor behind our results may be our increased power in the tests arising from our sample size. Our regressions have at least 120 observations whereas the regressions in MS have 32 quarterly observations (8 years).

[10]For example, one can imagine using information from monthly industrial production to forecast revisions to quarterly output.

[11]For example, for a level NIPA variable we have 17 explanatory variables with $2^{17} = 131,072$ possible combinations.

[12]The relative RMSEs $RMSE_1/RMSE_3$ and $RMSE_2/RMSE_3$ are in fact identical to Theil's U-statistic since the forecast associated with the third model is zero at all periods.

[13]Remember that we already established this result by using only the initial announcement as the explanatory variable for level variables in Section 4.1. Here we show that this result also extends to relative variables.

[14]A much more sophisticated forecast can be obtained by using state space methods. This general idea has been previously pursued in the literature. Howrey [1978], is one of the first papers to show how one can use the preliminary announcements to get an optimal prediction of the true variable. Conrad and Corrado [1978] apply the Kalman filter for getting better estimates for the monthly retail sales. Finally, Tanizaki and Mariano [1994] derive a non-linear and non-gaussian filter using importance sampling and Monte Carlo integration methods with Kalman Filter and apply this filter to the per capita consumption of the US.

[15]This will be a simple t-test except for the possible auto-correlation in d_t. We use Newey-West standard errors with appropriate lags for this test.

[16]According to Landefeld and Grimm [2000], about 18% of GDP is deflated by hedonic techniques, which means that quality-adjusted prices are used. Similarly, BLS uses hedonic prices for a significant number of goods to construct CPI.

[17]We count the "double-dip recession" of the 1980s as one recession.

[18]This is also the reason for the delay between the actual turning points and NBER's announcement, which is occasionally more than a year.

[19]The survey data and supporting documentation is publicly available on the internet at
http://www.phil.frb.org/econ/spf/index.html.

[20]Note that the magnitudes of the objects that are forecast are in the order of thousands. Therefore a

one-point change would correspond to a revision of less than 0.1% of the initial announcement.

[21] As Mork [1987] notes, the downward bias in initial announcements (or a positive mean of revisions) can be interpreted as a conscious conservativisim on the part of the statistical agency that avoids issuing over-optimistic figures. The fact that this motive is not symmetric can be explained by the similar asymmetry in the performance of the economy, which has a positive average growth rate.

Table I - Summary Statistics of Final Revisions - Full Sample

	N	Mean	Minimum	Maximum	Std. Dev.	Noise / Signal	Corr. with Initial	A/C (1)	Q-stat (20)
Revisions as Percentage of Initial Announcements									
Nominal Output	141	**0.60%**	-1.87%	3.12%	0.75%	-	-	**0.68**	**0.00**
Real Output	141	**0.43%**	-1.74%	2.82%	0.79%	-	-	**0.72**	**0.00**
Non-Farm Payroll Employment	422	**0.26%**	-1.08%	2.22%	0.70%	-	-	**0.95**	**0.00**
Industrial Production Index (Total Industry)	445	**0.94%**	-3.52%	7.17%	1.86%	-	-	**0.94**	**0.00**
Industrial Production Index (Manufacturing)	247	**0.74%**	-4.13%	2.33%	2.33%	-	-	**0.95**	**0.00**
Level of Revisions									
Annual Growth of Real Output	137	**0.18%**	-1.62%	2.00%	0.61%	0.24	-0.14	**0.55**	**0.00**
Annual Growth of Nominal Output	137	**0.29%**	-1.74%	2.57%	0.67%	0.25	-0.01	**0.60**	**0.00**
Annual Inflation (Output Deflator)	137	**0.10%**	-0.65%	1.12%	0.33%	0.13	**0.32**	**0.67**	**0.00**
Annual Growth of Labor Productivity	123	**0.35%**	-3.12%	3.34%	1.34%	0.82	**-0.51**	**0.67**	**0.00**
Annual Growth of Non-Farm Payroll Employment	414	**0.15%**	-0.83%	1.22%	0.39%	0.14	**0.23**	**0.92**	**0.00**
Annual Growth of Industrial Production (Total Industry)	433	**0.48%**	-2.66%	5.40%	0.80%	0.20	0.06	**0.82**	**0.00**
Annual Growth of Industrial Production (Manufacturing)	235	**0.56%**	-2.48%	2.93%	1.11%	0.23	-0.10	**0.83**	**0.00**
Quarterly Growth of Real Output	137	**0.25%**	-2.85%	5.12%	1.51%	0.41	-0.02	**-0.15**	0.15
Quarterly Growth of Nominal Output	137	**0.43%**	-3.60%	6.66%	1.59%	0.44	-0.07	-0.06	0.19
Quarterly Inflation (Output Deflator)	137	**0.17%**	-1.90%	3.15%	0.73%	0.27	0.13	**0.15**	0.84
Quarterly Growth of Labor Productivity	123	0.30%	-8.94%	7.02%	2.97%	0.94	**-0.42**	**-0.17**	0.50
Monthly Growth of Non-Farm Payroll Employment	414	**0.37%**	-4.85%	5.19%	1.42%	0.64	0.00	**0.10**	**0.00**
Monthly Growth of Industrial Production (Total Industry)	433	**1.11%**	-20.28%	24.12%	4.70%	0.48	**-0.08**	**0.11**	**0.00**
Monthly Growth of Industrial Production (Manufacturing)	235	**1.21%**	-12.81%	14.61%	4.58%	0.52	**-0.24**	0.01	0.41
Civilian Unemployment Rate	131	0.01%	-0.20%	0.17%	0.08%	0.05	0.04	-0.06	**0.00**
Capacity Utilization (Total Industry)	202	0.14%	-1.50%	2.30%	0.84%	0.46	**-0.38**	**0.85**	**0.00**
Capacity Utilization (Manufacturing)	249	0.14%	-2.10%	2.40%	0.94%	0.28	**-0.41**	**0.87**	**0.00**

Notes: All monthly and quarterly growth variables are annualized. Boldface denote significance at 10% level. A/C(1) column reports the first order autocorrelation coefficient. Q-stat(20) reports the p-value associated with the Q-statistic at 20 lags.

TABLE II - Results of the Ex-Post Forecasting Exercise - Full Sample

Dependent Variable (Final Revision of)	Criterion	N	Explanatory Variables	R^2	$\overline{R}^2$	Wald Test p-value	RMSE1 / RMSE3	RMSE2 / RMSE3
Revisions as Percentage of Initial Announcements								
Nominal Output	SIC	125	Cons, R5, R10, Trend	0.14	0.12	0.00	0.75	0.78
	AIC	125	Cons, R7, R10, Q1, Trend	0.16	0.13	0.00	0.74	
Real Output	SIC AIC	125	Cons, R10, Trend	0.09	0.07	0.00	0.86	0.88
Non-Farm Payroll Employment	SIC	412	Cons, Trend	0.03	0.03	0.00	0.93	0.82
	AIC	412	Cons, R6, Trend	0.03	0.03	0.00	0.93	
Industrial Production Index (Total Industry)	SIC	427	Cons, R14	0.05	0.04	0.00	0.87	0.47
	AIC	427	Cons, R14, Trend	0.06	0.05	0.00	0.86	
Industrial Production Index (Manufacturing)	SIC	230	Cons, R12, Trend	0.36	0.35	0.00	0.76	0.47
	AIC	230	Cons, R4, R12, R14, Trend	0.37	0.36	0.00	0.76	
Level of Revisions								
Annual Growth of Real Output	SIC	121	Cons, R10	0.00	-0.01	0.00	0.95	0.95
	AIC	121	Cons, Init, R4, R10	0.04	0.01	0.00	0.93	
Annual Growth of Nominal Output	SIC	121	Cons, R10, Trend	0.05	0.03	0.00	0.90	0.90
	AIC	121	Cons, R2, R4, R10, Trend	0.08	0.04	0.00	0.88	
Annual Inflation (Output Deflator)	SIC	125	Init, R5	0.10	0.09	0.00	0.91	0.93
	AIC	122	Init, R5, R9, Trend	0.13	0.10	0.00	0.90	
Annual Growth of Labor Productivity	SIC	110	Cons, Init, R6	0.28	0.27	0.00	0.83	0.97
	AIC	110	Cons, Init, R3, R6	0.30	0.28	0.00	0.82	
Annual Growth of Non-Farm Payroll Employment	SIC	400	Init	0.08	0.08	0.00	0.90	0.78
	AIC	400	Init, R12, R14	0.09	0.08	0.00	0.90	
Annual Growth of Industrial Production (Total Industry)	SIC	415	Trend	0.02	0.02	0.00	0.89	0.70
	AIC	415	Cons, Init, R14, Trend	0.04	0.03	0.00	0.88	
Annual Growth of Industrial Production (Manufacturing)	SIC	218	Cons, R8, R10, R14, Trend	0.40	0.38	0.00	0.69	0.73
	AIC	218	Cons, Init, R8, R10, R12, R14, Trend	0.41	0.40	0.00	0.68	
Quarterly Growth of Real Output	SIC	122	Cons, R1, R9	0.09	0.07	0.00	0.93	0.96
	AIC	121	Cons, R1, R3, R9, R10, Q3	0.15	0.12	0.00	0.90	
Quarterly Growth of Nominal Output	SIC	125	Cons, R2, Q3	0.06	0.04	0.00	0.93	0.93
	AIC	122	Cons, R1, R2, R9, Q3, Trend	0.11	0.07	0.00	0.91	
Quarterly Inflation (Output Deflator)	SIC	125	Init, R1	0.05	0.04	0.00	0.95	0.96
	AIC	122	Cons, R1, R9, Trend	0.07	0.05	0.00	0.94	
Quarterly Growth of Labor Productivity	SIC	114	Init, R1, Q2	0.21	0.19	0.00	0.89	0.96
	AIC	114	Cons, Init, R1, Q2	0.23	0.21	0.00	0.88	
Monthly Growth of Non-Farm Payroll Employment	SIC	400	Cons, Init, R6, R12	0.15	0.14	0.00	0.90	0.80
	AIC	400	Cons, Init, R5, R6, R12	0.15	0.14	0.00	0.89	
Monthly Growth of Industrial Production (Total Industry)	SIC	415	Cons	0.00	0.00	0.00	0.97	0.92
	AIC	415	Cons, R1, R2, R4, R7, R9, R12	0.05	0.03	0.00	0.95	
Monthly Growth of Industrial Production (Manufacturing)	SIC	218	Init, R4, Trend	0.08	0.08	0.00	0.93	0.97
	AIC	218	Cons, Init, R1, R2, R3, R4, R7, R14	0.14	0.11	0.00	0.90	
Civilian Unemployment Rate	SIC	121	R6, R7, Trend	0.17	0.16	0.00	0.90	0.97
	AIC	121	Cons, R4, R5, R8, R9, Trend	0.23	0.19	0.00	0.87	
Capacity Utilization (Total Industry)	SIC AIC	188	Cons, Init, R7, Trend	0.29	0.28	0.00	0.82	0.96
Capacity Utilization (Manufacturing)	SIC AIC	235	Cons, Init, R7, R14, Trend	0.28	0.27	0.00	0.84	0.94

Notes: RMSE1: Root mean squared error of forecasting model. RMSE2: Root mean squared error of the model with only trend and seasonal variables. RMSE3: Root mean squared of zero forecast. Explanation of variables: "Cons": Constant, Rs: sth revision to the variable at pe
Linear trend, Qs: Dummy for sth quarter, "Init": Initial announcement.

Table III - Results of the Real-Time Forecasting Exercis

	RMSE1	RMSE2	RMSE1/RMSE2	DM
Revisions as Percentage of Initial Announcements				
Nominal Output	0.79	0.83	**0.96**	-0.06
Real Output	0.86	0.90	**0.96**	-0.06
Non-Farm Payroll Employment	0.64	0.55	1.16	0.10
Industrial Production Index (Total Industry)	2.15	2.18	**0.98**	-0.15
Industrial Production Index (Manufacturing)	3.39	3.37	1.01	0.20
Level of Revisions				
Annual Growth of Real Output	0.65	0.67	**0.96**	-0.03
Annual Growth of Nominal Output	0.61	0.64	**0.95**	-0.04
Annual Inflation (Output Deflator)	0.34	0.29	1.20	0.04
Annual Growth of Non-Farm Payroll Employment	0.41	0.40	1.01	0.00
Annual Growth of Industrial Production (Total Industry)	1.00	1.10	**0.91**	-0.22
Annual Growth of Industrial Production (Manufacturing)	1.48	1.62	**0.91**	**-0.44**
Quarterly Growth of Real Output	1.26	1.28	**0.98**	-0.06
Quarterly Growth of Nominal Output	1.18	1.21	**0.97**	-0.09
Quarterly Inflation (Output Deflator)	0.59	0.55	1.06	0.04
Monthly Growth of Non-Farm Payroll Employment	1.16	1.16	**1.00**	0.00
Monthly Growth of Industrial Production (Total Industry)	3.92	4.00	**0.98**	-0.60
Monthly Growth of Industrial Production (Manufacturing)	4.16	4.37	**0.95**	**-1.81**
Capacity Utilization (Total Industry)	0.98	0.95	1.03	0.06
Capacity Utilization (Manufacturing)	0.99	0.96	1.03	0.05

Notes: RMSE1: Root mean squared error from the real-time forecast. RMSE2: Root mean squared error from the zero forecast. In the RMSE1/RMSE2 column, entries less than unity are denoted by boldface. DM column reports the Diebold-Mariano statistic and the statistics with *p*- values less than 0.10 are denoted by boldface.

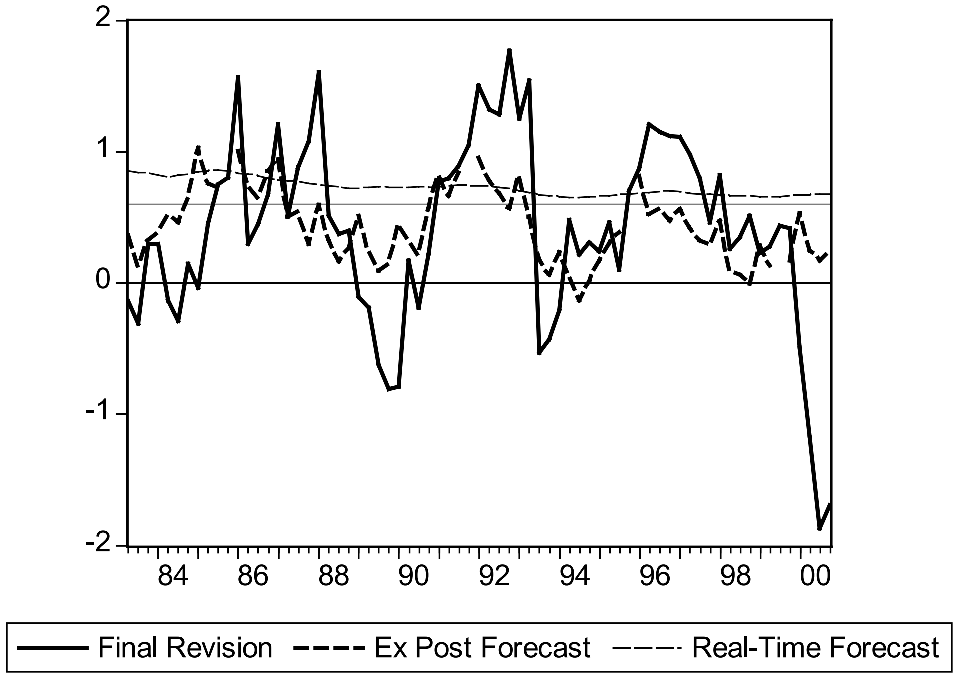
TABLE IV - Results from the Survey of Professional Forecasters

	Nominal Output	Real Output	Real Consumption	Price Index
Pre-1990				
Median Equals Initial	88.5%	86.5%	92.6%	92.0%
Within One Point	77.0%	67.9%	76.2%	97.4%
Correct Sign	20.7%	23.7%	22.1%	38.9%
Mean Revision (SPF)	0.03%	0.00%	0.08%	0.07%
Mean Revision (RTDS)	0.60%	0.46%	0.47%	0.14%
Post-1990				
Median Equals Initial	97.8%	95.4%	95.5%	91.1%
Within One Point	92.0%	94.0%	94.4%	100.0%
Correct Sign	26.1%	35.2%	34.8%	43.3%
Mean Revision (SPF)	0.00%	0.00%	0.00%	0.04%
Mean Revision (RTDS)	0.60%	0.51%	0.40%	0.08%

Note: See the text for the definitions of each row.

Figure I – Final Revisions and Various Forecasts

Nominal Output



Quarterly Output

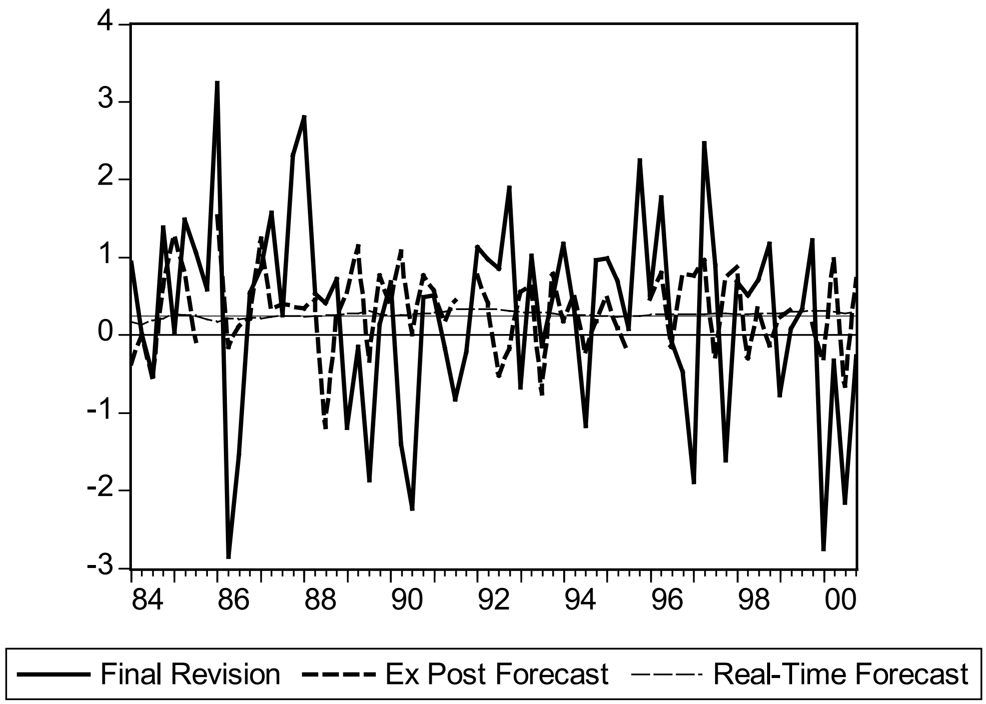
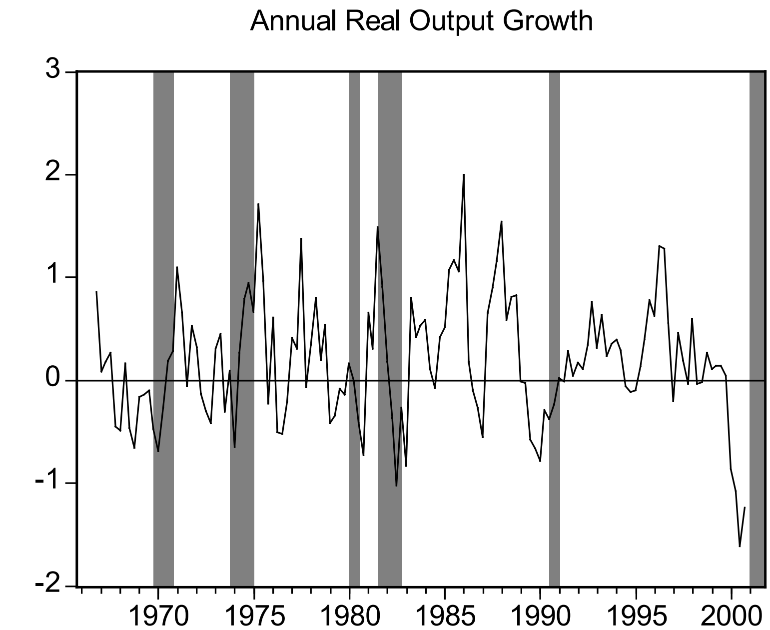
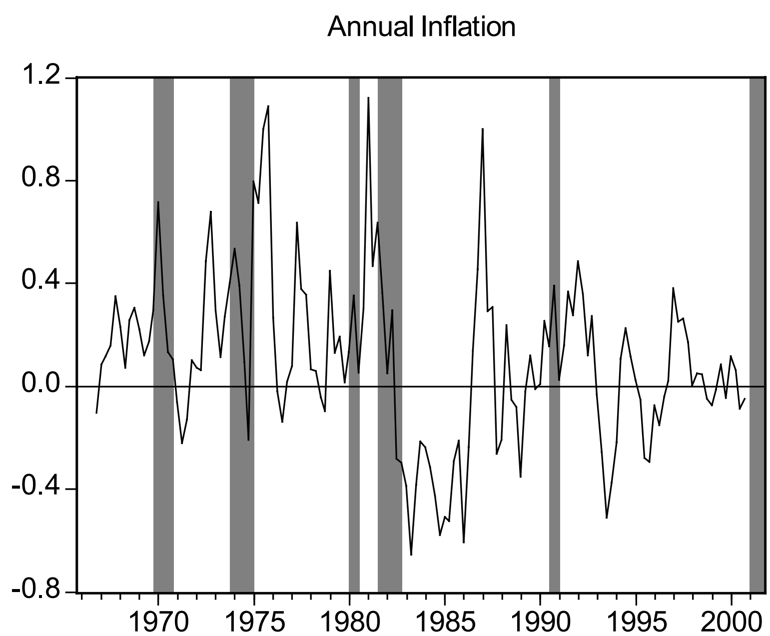


Figure II – Final Revisions for Real Output Growth and Inflation





Note: The shaded areas are the business cycle recessions as determined by the National Bureau of Economic Research.