



ELSEVIER

Journal of Econometrics 105 (2001) 111–130

**JOURNAL OF
Econometrics**

www.elsevier.com/locate/econbase

A real-time data set for macroeconomists

Dean Croushore*, Tom Stark

Research Department, Federal Reserve Bank of Philadelphia, Philadelphia, PA 19106, USA

Abstract

This paper describes a real-time data set for macroeconomists that can be used for a variety of purposes, including forecast evaluation. The data set consists of quarterly vintages, or snapshots, of the major macroeconomic data available at quarterly intervals in real time. The paper explains the construction of the data set, examines the properties of several of the variables in the data set across vintages, and provides an example showing how data revisions can affect forecasts. © 2001 Elsevier Science S.A. All rights reserved.

JEL classification: C82; E30

Keywords: Real-time data; Forecasting; Data revisions

1. Introduction

In creating models to use for forecasting, economists use the most recent vintage of historical data available to them to develop and test alternative models. They often compare the forecasts from a new model to forecasts from alternative models or to forecasts that were made by others in real time. However, since the analysis of the new forecasts is often based on the final, revised data, rather than the data that were available to economic agents who were making forecasts in real time, the results of such exercises may be misleading.

To avoid such problems in creating forecasting models, we have developed a data set that gives a modeler a snapshot of the macroeconomic data available at any given date in the past. We call the information set available at a

* Corresponding author. Tel.: +1-215-574-3809; fax: +1-215-574-4364.

E-mail address: dean.croushore@phil.frb.org (D. Croushore).

particular date a “vintage”, and we call the collection of such vintages a “real-time data set”.

This paper explains the reasons for the construction of this data set, describes the data set, and provides an empirical example demonstrating the extent to which the data vintage matters for evaluating forecasts.¹ In creating our real-time data set, our goal is to provide a basic foundation for research on issues related to data revision by allowing researchers to use a standard data set, rather than collecting real-time data themselves for every different study. We begin with a brief discussion of related research in Section 2. We provide details about the data set in Section 3. In Section 4, we examine several different variables, showing the degree to which the data are affected by revisions. Section 5 provides an empirical forecasting example demonstrating how forecasts may be sensitive to the choice of data vintage. We draw conclusions from these results in Section 6.

2. Related research

Research that deals with the fact that data are revised has been ongoing since the seminal studies of Zellner (1958), Morgenstern (1963), and Cole (1969). Diebold and Rudebusch (1991) provided a dramatic example of the importance of data revisions by showing that the index of leading indicators does a much worse job of predicting future movements of industrial production in real time than it does after the data are revised.²

The importance of being careful about the preliminary nature of some data has been emphasized by numerous researchers, beginning with Stekler (1967), who found value in early data releases even though they contained errors. Howrey (1978) showed how to adjust for the fact that data within a particular vintage have been revised to differing degrees, while Harvey et al. (1981) showed how to deal with revisions that occur at irregular intervals.

A substantial body of research has been devoted to examining the efficiency of data production or investigating the nature and statistical properties of data revisions. Conrad and Corrado (1979) showed how a data user could improve on published data on retail sales. Mankiw et al. (1984) explored the efficiency of early releases of the money supply. Mankiw and Shapiro (1986) did the same for real output data, and Mork (1987) found an improved empirical technique for performing such analysis. Pierce (1981) and Sargent (1989)

¹ In a companion paper, we analyze the degree to which data vintage matters for the robustness of empirical studies in macroeconomics. See Croushore and Stark (1999b).

² In support of this result, Robertson and Tallman (1998a) show that a VAR that uses real-time data from the index of leading indicators produces no better forecasts for industrial production than an AR model using just lagged data on industrial production. However, they also show that the leading indicators may be useful in forecasting real output (GNP/GDP) in real time.

described some theories on how to think about the data revision process. Patterson and Heravi (1991), Swanson et al. (1999), and Swanson (1996) have examined different aspects of data revisions, finding, among other results, that data revisions are somewhat forecastable.

Data revisions are important because they may affect policy decisions or the manner in which such decisions depend on the most recent data, as Maravall and Pierce (1986) investigated some years ago. Revisions may also affect people's expectations, as Boschen and Grossman (1982) found. Recently, a number of authors have investigated the role of data revisions in affecting monetary policy rules (Orphanides, 1998, 2001; Evans, 1998; Ghysels et al., 1999), how measures of monetary policy shocks are affected by data revisions (Rudebusch, 1998; Croushore and Evans, 1999), the empirical relationship between money and output (Amato and Swanson, 2001), the impact on policy research (Runkle, 1998), and how monetary policy responds to uncertainty (Rudebusch, 2001).

Forecasting models may be particularly sensitive to data revisions, as Fair and Shiller (1990) pointed out. Denton and Kuiper (1965) found that the use of preliminary (rather than final) data led to large forecast errors, but Trivellato and Rettore (1986) found effects that were much more modest. Howrey (1996) showed that forecasts of the level of GNP are much more sensitive to data revisions than forecasts of growth rates. Swanson and White (1997) used real-time data to investigate optimal model selection for a number of variables. Robertson and Tallman (1998b) used a real-time data set to evaluate alternative VAR model specifications for forecasting unemployment, inflation, and output growth. Koenig and Dolmas (1997) developed a method for forecasting real output growth using monthly data based on real-time analysis. A further development of that idea in a paper by Koenig et al. (2000) examined the question of what sets of real-time data forecasters should use: the fully revised series available at each date, or some unrevised series of initial releases, again in the context of forecasting real output growth.

Our real-time data set promises to facilitate all the types of empirical research discussed above, by making available a single source that is well documented and thoroughly investigated, containing real-time data for a large number of macroeconomic variables.

3. The data set

In concept, developing a real-time data set is simple—just enter old data into spreadsheets. But in reality, producing our real-time data set required a substantial amount of effort, including digging through old source data and figuring out what data were available at what time, a procedure that was not trivial, considering the lack of documentation for much of the data. As a

result, the data-collection phase of this project has been going on for the past eight years.

Our real-time data set now includes data as they existed in the middle of each quarter (on the 15th day of the month, to be precise), from November 1965 to the present.³ For each vintage date, the observations are identical to those one would have seen in published sources at that time. For example, if you want to know what the data looked like on August 15, 1968, just pull down the data set from our web site and look at the vintage for August 1968, and you will find the relevant data—a time series for each variable from the first quarter of 1947 to the second quarter of 1968.⁴ See Table 1 for a list of the variables included in the data set and the Internet address where the data can be found.

A few notes about some of the variables are worth mentioning here, though more complete details can be found in the documentation files on our web page.⁵ First, even though the interest-rate variables are never revised, they are included for completeness. The other variables are revised to some degree over time, though some, like the CPI, are revised only through changes in seasonal adjustment factors or changes in the base year. Note that the data set includes the chain-weighted GDP price index in vintages beginning with February 1996 but does not include a price-level measure before that.⁶ However, for vintages prior to February 1996, a price level can be constructed by taking the ratio of nominal GNP/GDP to real GNP/GDP. The data set is mostly complete, but some data are missing for the money stock, monetary base, and reserves variables; we are in the process of adding some of the missing data from obscure sources.

Though the project of collecting these data seems simple, it turned out that finding old data was not easy. Further, since the critical element for economic research is the timing of the data (was it released during the second week of February or the third?), we were very careful to include in each vintage only the observations we knew were available at the time. In many cases data

³ Why the middle of each quarter? Because one of the original motivations for this project came from research on the forecast efficiency of the Survey of Professional Forecasters (see Keane and Runkle, 1990; Croushore, 1993), whose forecasters make their forecasts in the middle of each quarter.

⁴ More precisely, there are two data sets at each date, one containing quarterly variables, such as real GDP, and another containing monthly variables, such as the unemployment rate.

⁵ From links on the web page, you can download the data, read documentation about the data, and find out when new data will become available. We plan to add new vintages shortly after the 15th day of the middle month of each quarter.

⁶ The Bureau of Economic Analysis switched from a fixed-weight methodology to a chain-weight methodology in early 1996. That change was significant in a number of ways because it included substantial revisions in past data and because chain weighting is very different conceptually from fixed-weighting. Under chain weighting, the components of real GDP do not sum up to real GDP.

Table 1

Basic information about the real-time data set

 Web site: <http://www.phil.frb.org/econ/forecast/reaindex.html>

Variables included in the data set

Quarterly observations:

Nominal GNP (vintages before 1992) or GDP (vintages in 1992 and after)

Real GNP (vintages before 1992) or GDP (vintages in 1992 and after) and components:

Consumption and its components:

Durables

Nondurables

Services

Components of Investment:

Business Fixed Investment

Residential Investment

Change in Business Inventories (Change in Private Inventories after
vintage August 1999)Government Purchases (Government Consumption and Gross Investment,
vintages in 1996 and after)

Exports

Imports

Chain-Weighted GDP Price Index (vintages in 1996 and after)

Corporate Profits

Import Price Index

Monthly observations:

(Quarterly averages of these variables are also available in the quarterly data sets)

Money supply measures:

M1

M2

Reserve measures (data from Board of Governors):

Total reserves

Nonborrowed reserves

Nonborrowed reserves plus extended credit

Monetary base

Civilian Unemployment Rate

Consumer Price Index (seasonally adjusted)

3-month T-bill Rate

10-year T-bond Rate

were revised, but the publications that detailed the revisions did not always say when the revisions were made available. So it took a substantial amount of effort to figure out exactly what observations should have been, or should not have been, included in each vintage. A comprehensive set of notes about the data set is available on our web site to help researchers understand our conventions for including or excluding particular observations. Also, some of the data have been collected in real time since this project began in 1991, though the scope of the project has expanded since that time.

After entering all the data into a set of database worksheets, we ran a number of editing checks to try to ensure the quality of the data.⁷ In some cases this was easy. For example, we tested a large sample of vintages to make sure that (prior to the switch to chain weighting) the sum of the components of real output added up to total real output. In other cases, where there was no adding-up constraint, we plotted growth rates of the variables to ensure that they looked sensible. This helped tremendously in finding typographical errors in the data set.

4. Data revisions

We know that the economic data are revised, but are such revisions large enough to worry about? For example, when one compares forecasts from a new model estimated on today's data with forecasts from a model estimated on old data, how different are the data sets and the resulting forecasts? To investigate this question, we will look at a few selected variables: real output, business fixed investment, and consumption spending on durable goods.⁸

First, to see how much the data vintage matters for fairly long horizons, we will examine revisions to five-year average growth rates.⁹ We look at the annual average growth rate over five-year periods in Table 2 for data from vintages (hereafter called "benchmark vintages") dated November 1975, November 1980, November 1985, November 1991, November 1995, and August 1999. These vintages were chosen because they were the last vintages prior to a comprehensive revision of the national income and product accounts.

When the government made comprehensive revisions to the national income and product account (NIPA) data following our benchmark vintage dates, they often made significant changes, including modifying the definitions of variables and incorporating new source data.¹⁰ The base year was changed for real variables in January 1976 (from 1958 to 1972), in December 1985

⁷ Most of the data entry was done over the last six years by a small army of undergraduate students working as interns. From Princeton University: Michael Hodge, Ron Patrick, Adam Stark, Jason Harvey, Jake Erhard, Keith Wilbur, and Andrew Stern. From the University of Pennsylvania: Peter High, Lisa Forman, and Bill Wong. However, from 1997 to 1998 the lion's share of the work was done by Bill Wong, who hammered the data set into shape under the supervision of one of the authors, Tom Stark. Our thanks to all these wonderful students who produced a high-quality product! Research assistants at the Philadelphia Fed continue to collect new data.

⁸ See Croushore and Stark (1999b) for a similar analysis of the revisions to nominal output, real consumption spending, and the price level.

⁹ Patterson and Heravi (1991) examine cointegration between successive benchmark revisions in data from the UK, finding that cointegration sometimes exists.

¹⁰ The construction of the NIPA data are explained in Ritter (2000).

Table 2
Average growth rates over five years for benchmark vintages (annualized percentage points)

Period	Vintage year					
	1975	1980	1985	1991	1995	1999
<i>Real output</i>						
49Q4 to 54Q4	5.2	5.1	5.1	5.5	5.5	5.3
54Q4 to 59Q4	2.9	3.0	3.0	2.7	2.7	3.2
59Q4 to 64Q4	4.1	4.0	4.0	3.9	4.0	4.2
64Q4 to 69Q4	4.3	4.0	4.1	4.0	4.0	4.4
69Q4 to 74Q4	2.1	2.2	2.5	2.1	2.3	2.6
74Q4 to 79Q4	NA	3.7	3.9	3.5	3.4	3.9
79Q4 to 84Q4	NA	NA	2.2	2.0	1.9	2.2
84Q4 to 89Q4	NA	NA	NA	3.2	3.0	3.2
89Q4 to 94Q4	NA	NA	NA	NA	2.3	1.9
<i>Real business fixed investment</i>						
49Q4 to 54Q4	3.9	5.0	5.0	4.9	4.8	4.7
54Q4 to 59Q4	2.5	3.1	3.2	2.7	3.0	3.0
59Q4 to 64Q4	5.9	5.6	5.7	5.3	5.7	6.2
64Q4 to 69Q4	6.3	6.2	6.4	6.0	6.0	7.1
69Q4 to 74Q4	2.0	1.7	1.9	2.5	2.9	3.6
74Q4 to 79Q4	NA	3.9	5.8	5.2	5.7	6.5
79Q4 to 84Q4	NA	NA	4.6	2.4	2.4	3.5
84Q4 to 89Q4	NA	NA	NA	2.7	1.0	1.1
89Q4 to 94Q4	NA	NA	NA	NA	5.7	2.7
<i>Real consumption spending on durables</i>						
49Q4 to 54Q4	4.1	3.8	3.6	3.9	4.3	3.8
54Q4 to 59Q4	3.0	2.0	2.0	2.0	1.7	2.4
59Q4 to 64Q4	6.3	5.2	5.4	4.7	4.5	6.0
64Q4 to 69Q4	7.8	7.0	7.2	6.7	6.4	7.3
69Q4 to 74Q4	1.7	2.6	2.7	2.6	2.0	2.8
74Q4 to 79Q4	NA	7.1	7.0	6.8	6.3	6.6
79Q4 to 84Q4	NA	NA	4.6	4.8	4.1	4.9
84Q4 to 89Q4	NA	NA	NA	4.9	4.7	5.0
89Q4 to 94Q4	NA	NA	NA	NA	4.9	3.2

Note: Each table entry shows the average annual growth rate of the variable over the five-year period shown in the first column, as recorded in the benchmark data vintage at the top of each column. Reading across columns within a given row illustrates the degree to which five-year annual average growth rates have been revised over time.

(from 1972 to 1982), in late November 1991 (from 1982 to 1987), and in January 1996 (from 1987 to 1992). As a result, some of the differences across the benchmark vintages we look at (1975–1980, 1985–1991, 1991–1995, and 1995–1999) incorporate base-year changes, which affect real variables. Most importantly, since the base-year changes in 1976, 1985, and 1991 used the

old fixed-weighted index methodology, the change of base year alters the timing of substitution bias; this bias is large for dates further away from the base year.

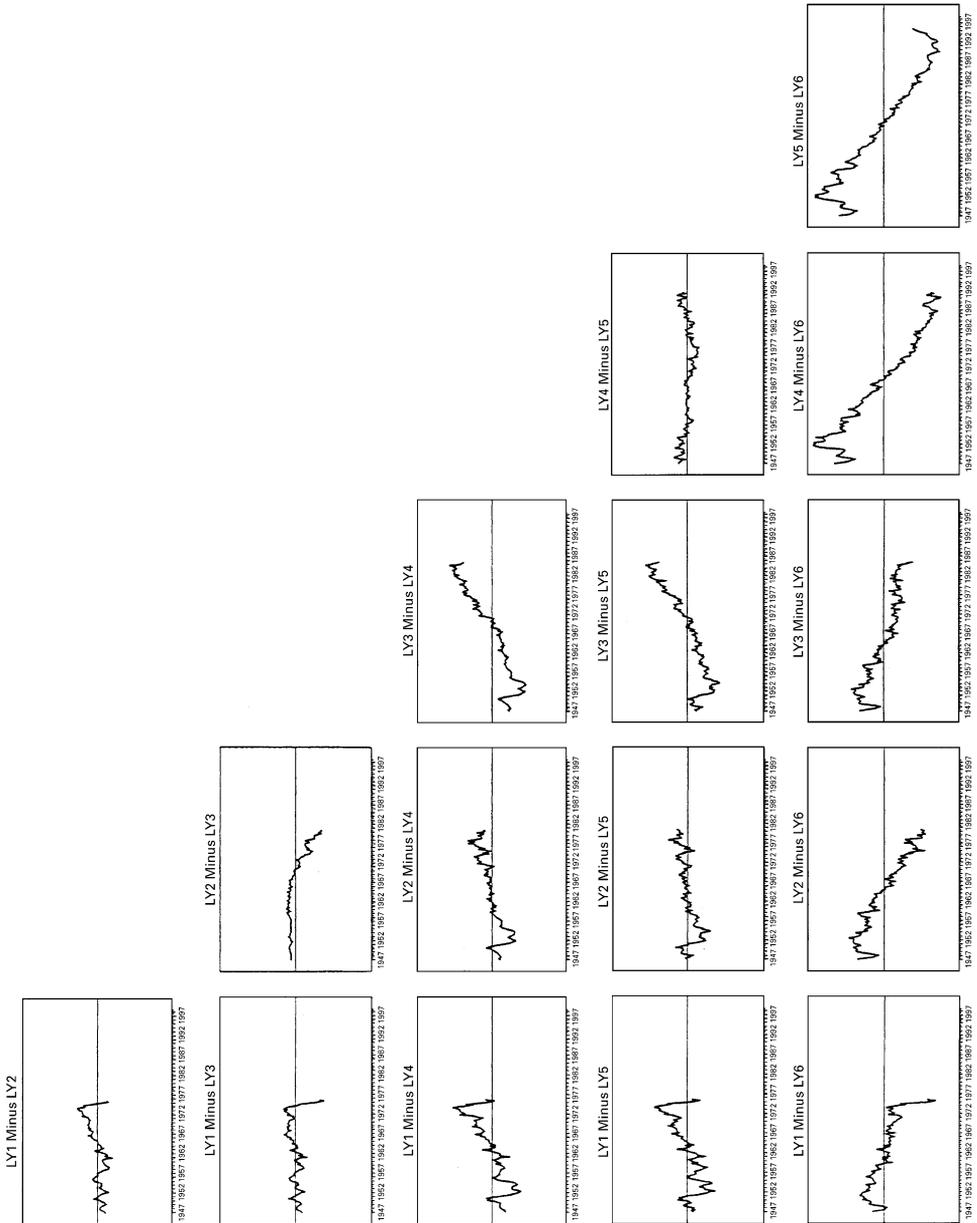
There is a potentially significant change in one of our variables across the benchmark vintages. The real output variable is GNP in vintages before February 1992 but GDP in vintages from February 1992 on. Our data set is consistent with the “headline” variable (that is, the variable that is listed in Tables 1.1 and 1.2 in the *Survey of Current Business* and is the focus of the discussion about aggregate economic activity), but users need to be aware of this change, since the differences between GNP and GDP are not random; they are persistent in sign. So some of the differences across vintages in real output arise because of this definitional change.¹¹

A major change in the methodology of the national income accounts arose in 1996, when the government switched from fixed-weighted indexes to chain-weighted indexes, to eliminate substitution bias. Under the fixed-weight methodology, such a change in the base year led to significant changes in the growth rates of real variables, often with large changes for years in the distant past. Under chain weighting, however, a change of base year has no impact on the growth rates of real variables.

As we look across the columns of Table 2, we can see how the five-year annual average growth rate has changed across benchmark vintages. For real output, the vintage makes a difference, especially when the base year is changed. Especially large changes show up in moving from the 1985 to the 1991 benchmark vintage (reflecting the base-year shift of December 1985) and moving from the 1995 to the 1999 benchmark vintage (reflecting the

Fig. 1. Real output. Vertical scale: -0.08 to 0.08 . Note: Plots show differences between log levels of variables with the mean difference subtracted, that is, $\log[X(t,a)] - \log[X(t,b)] - m$, where m is the mean of $\log[X(\tau,a)/X(\tau,b)]$ over the largest sample of τ contained in both vintages, and where b is a later vintage than a . The notation on each plot follows the convention $Lz\#$, where L means the logarithm of the variable, z represents the variable ($z=Y$ for real output, $z=IB$ for business fixed investment, and $z=CD$ for consumption of durables), and where $\#$ represents the benchmark vintage, with $\#=1$ for the November 1975 vintage, $\#=2$ for November 1980, $\#=3$ for November 1985, $\#=4$ for November 1991, $\#=5$ for November 1995, and $\#=6$ for August 1999. Each plot shows dates along the horizontal axis from 1947Q1 to 1998Q3. The last data point plotted is 1975Q3 in column 1, 1980Q3 in column 2, 1985Q3 in column 3, 1991Q3 in column 4, and 1995Q3 in column 5. The vertical axis in each plot is listed in the legend for each figure; these are demeaned log differences. Plots show trends (an upward trend means later data points were revised upward relative to earlier data, reflecting faster growth in the revised data), spikes (which arise when data for a particular data are revised sharply up or down), and deviations from trend (which reflect low frequency differences between vintages).

¹¹ We could create a data set with all GNP data, but GNP data are no longer released at the same time as the headline number (GDP); so the timing in all the data sets would change.



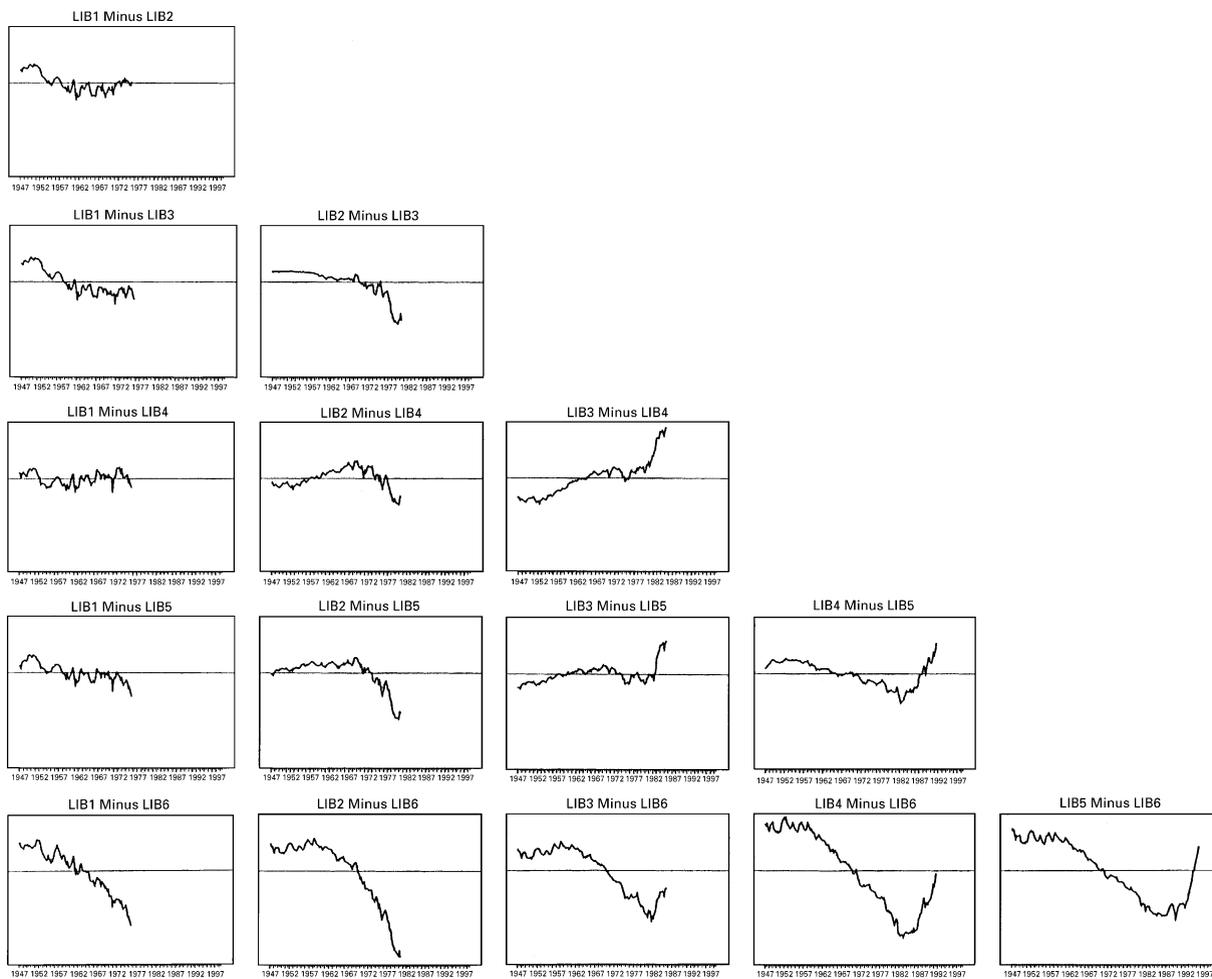


Fig. 2. Real business fixed investment. Vertical scale: -0.25 to 0.15 . *Note:* See notes for Fig. 1.

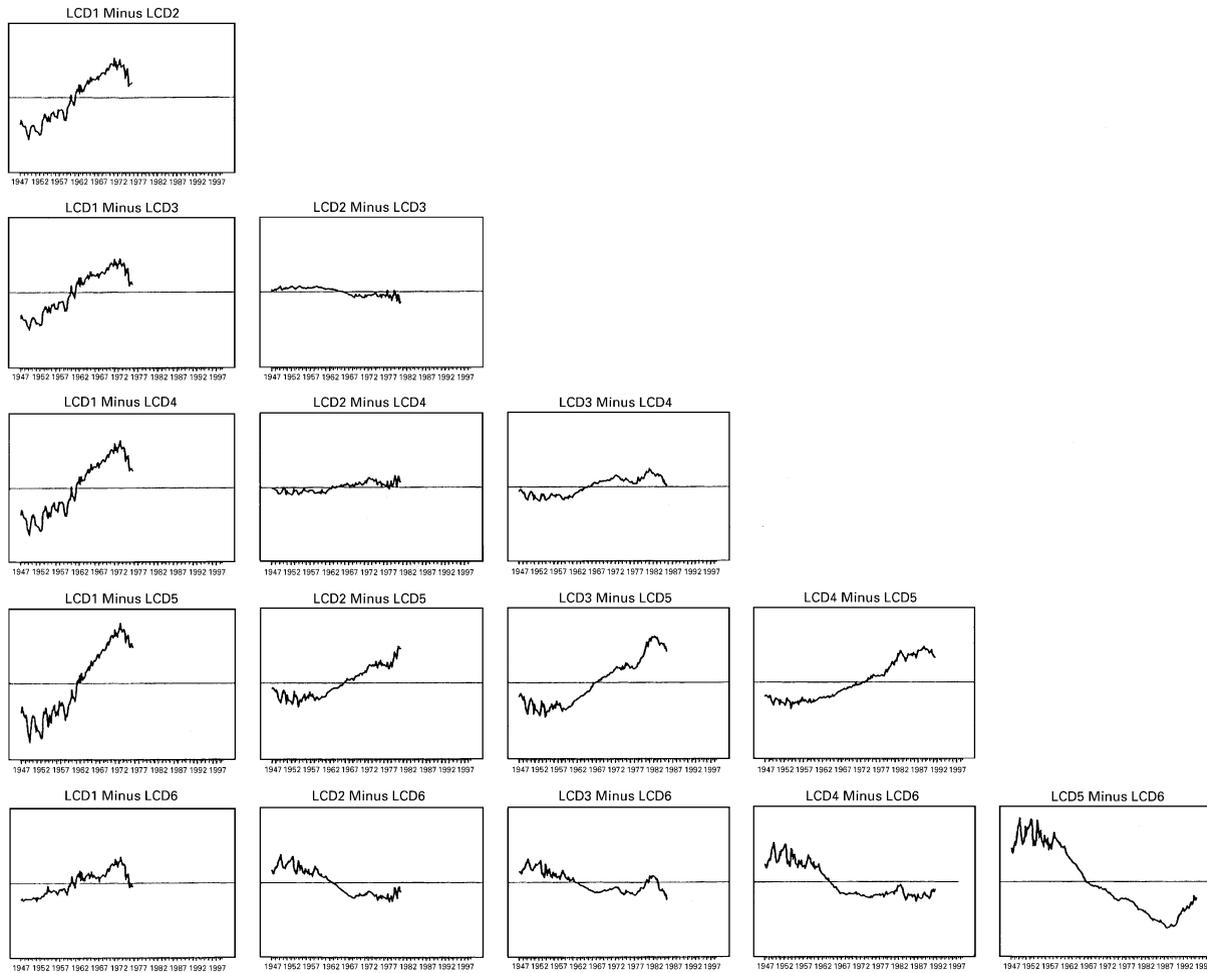


Fig. 3. Real consumption spending on durable goods. Vertical scale: -0.20 to 0.20 . Note: See notes for Fig. 1.

move to chain weighting in 1996). But those differences in real output growth are tiny compared with what we see for various components of output. The growth rates for business fixed investment have changed dramatically across vintages. For example, business investment grew 3.9 percent per year in the first half of the 1950s, according to the 1975 benchmark vintage, but grew 5.0 percent per year according to the 1980 vintage. Even more dramatic was the slowdown in the measured growth rate in the first half of the 1990s, from 5.7 percent in the 1995 vintage to just 2.7 percent in the 1999 vintage (owing, in large part, to the switch to chain weighting and the impact that had on investment in producers' durable equipment, especially for computers). Consumer spending on durables is also quite volatile as measured across the benchmark vintages, especially in the second half of the 1950s and again in the second half of the 1960s.

One way to examine how revisions affect the data is to plot differences in the data across vintages for the same date. Figs. 1–3 show plots of the differences between the log levels of the variables, with the mean difference (over the common time period) subtracted, because it reflects mainly base-year changes. Let $X(t, s)$ represent the level of a variable for time t in vintage s . We plot, for each date t that is common to vintages a and b , the value of $Z_t \equiv \log[X(t, a)/X(t, b)] - m \equiv \log[X(t, a)] - \log[X(t, b)] - m$, where m is the mean of $\log[X(\tau, a)/X(\tau, b)]$ over the largest sample of τ contained in both vintages, and where b is a later vintage than a .¹²

In the figures, the first column of plots compares the first benchmark vintage ($a = 1975$) to each later benchmark vintage. So, the upper left plot is the second benchmark vintage ($b = 1980$) compared to the first; the plot below compares the third benchmark vintage ($b = 1985$) to the first; and so on. The second column does the same for the second benchmark vintage ($a = 1980$), and so on, and the final column, which has just one entry, compares the fifth benchmark vintage ($a = 1995$) to the sixth ($b = 1999$). The notation on each plot follows the convention $Lz\#$, where L means the logarithm of the variable, z represents the variable ($z = Y$ for real output, $z = IB$ for business fixed investment, and $z = CD$ for consumption of durables), and where $\#$ represents the benchmark vintage, with $\# = 1$ for the November 1975 vintage, $\# = 2$ for November 1980, $\# = 3$ for November 1985, $\# = 4$ for November 1991, $\# = 5$ for November 1995, and $\# = 6$ for August 1999.

If you look at plots on the main diagonal of the figures, you are comparing adjacent benchmark vintages. The plots below the main diagonal show comparisons across two or more benchmark vintages. Each plot shows dates along the horizontal axis from 1947Q1 to 1998Q3. The last data point plotted is 1975Q3 in column 1, 1980Q3 in column 2, 1985Q3 in column 3, 1991Q3

¹² Since we have removed the mean, we will not capture any mean shifts in variables, but those are illustrated in Table 2.

in column 4, and 1995Q3 in column 5. The vertical axis in each plot is listed at the bottom of each figure; these are demeaned log differences.

Three major features of the plots are apparent: (1) trends; (2) spikes; and (3) other deviations from a linear trend. First, the dominant feature of the plots is the presence of trends. A downward tilt means later data points were revised upward relative to earlier data, reflecting faster trend growth; similarly, an upward tilt means that later data points were revised downward relative to earlier data. Second, a spike in a plot means that data for a particular date or series of dates were revised significantly in one direction relative to other dates in the sample. The third source of difference in the plots is the presence of long-lived deviations from a linear trend (or, when no trend is evident, from zero), suggesting that there are low frequency differences between vintages. Taken together, the plots point to cross-vintage differences at many frequencies.¹³

In Fig. 1, the effects of substitution bias on real output growth rates are apparent by looking at the tilt in the plots. For example, in moving from benchmark vintage 3 to benchmark vintage 4, the log ratio series tilts upward, because the fixed-weight method with a change in base year from 1972 to 1982 greatly changes the relative pricing relationships between energy and other goods. Thus, the plot is tilted, as even data from long before were affected significantly. But moving from vintage 5 to vintage 6 reverses that effect, due to chain weighting. Notice also that the movement from GNP to GDP (from vintage 4 to vintage 5) did not cause much effect.

In Fig. 2, showing business fixed investment, the most striking result is the steepness of the plots in the bottom row. This represents a change in methodology when chain weighting was introduced in 1996. As a result, changes to investment spending estimates were particularly pronounced, because of large changes in the price indexes for investment in technology (especially computers), and hence in the real value of investment.¹⁴ In addition, the changes in the measurement of investment spending when benchmark revisions occur (columns 1, 3, 4, and 5) are remarkable, especially because they are nonlinear. They come from a variety of sources, including new data from censuses, changes in estimated prices, and changes in procedures for calculating values. This suggests that, in analyzing forecasts, one should be very careful about what vintage of the data one uses as “actual”, since redefinitions, changes in methodology, and changes in relative prices seem to have dramatic effects on both the levels and the growth rates of business fixed investment.

Fig. 3 shows that real consumption on durables is strongly affected by some revisions, but less so by others. As is the case with some of the other plots, chain weighting leads to the opposite tilt direction of fixed weighting,

¹³ See Croushore and Stark (1999b) for some spectral analysis of the revisions.

¹⁴ For more on these issues, see Landefeld and Parker (1995, 1997).

because of distortions of relative prices under fixed weighting. In the first column, the decline in durables from the benchmark revision came about from the reclassification of some items from consumption to residential investment, plus reclassification of a portion of autos from consumption to business investment (depending on personal versus business ownership). In the fourth column, the declines in the growth rates of durables arise because of changed depreciation assumptions, new source data, and quality adjustments. And in the last column, it is mostly the change to chain weighting that affects the pattern of revisions. Chain weighting reverses some of the earlier effects of fixed weighting, so the lower left-hand plot is basically flat, though other plots have a substantial tilt to them.

Having documented that data revisions are potentially large for a variety of variables, we now pose the question: do such revisions matter for forecasting?

5. An empirical forecasting example

To illustrate how the data vintage matters in forecasting, we run some simple empirical exercises. We estimate and forecast real output growth with an ARIMA model and compare the forecasts generated from models estimated on latest-available data (from our August 1999 vintage) to those generated from models estimated on our real-time data.¹⁵ We proceed in the following manner: (1) estimate a model for real output growth using data for the second quarter of 1948 through the third quarter of 1965 that was known in November 1965; (2) forecast quarter-over-quarter real output growth for the fourth quarter of 1965 and the following three quarters to the third quarter of 1966, then form a four-quarter average growth-rate forecast over that time span; (3) repeat parts (1) and (2) in a rolling procedure, going forward one quarter each step, adding one more observation to the sample used for estimation; and (4) calculate the forecast errors for the four-quarter-average forecasts.¹⁶ We follow this procedure once using the real-time data set (for which data revisions are possible as we roll forward each quarter using the next vintage associated with that quarter), and a second time using only the latest vintage available when this analysis was undertaken (vintage August 1999, which contains no data revisions as we roll forward each quarter).

¹⁵ See Croushore and Stark (1999a) for similar results with additional forecasting models, including a univariate Bayesian model, and with the multivariate quarterly Bayesian vector error-correction (QBVEC) model of Stark (1998). Similar results to those discussed in this paper obtain, though Bayesian methods seem to reduce the impact of data revision on forecasts.

¹⁶ This procedure could also be carried out to examine forecasts one quarter ahead, two quarters ahead, three quarters ahead, or four quarters ahead, rather than the average of the four quarterly forecasts, with similar results. Note that we use the latest vintage (August 1999) to represent the actuals in this exercise; later, we discuss an alternative choice for actuals.

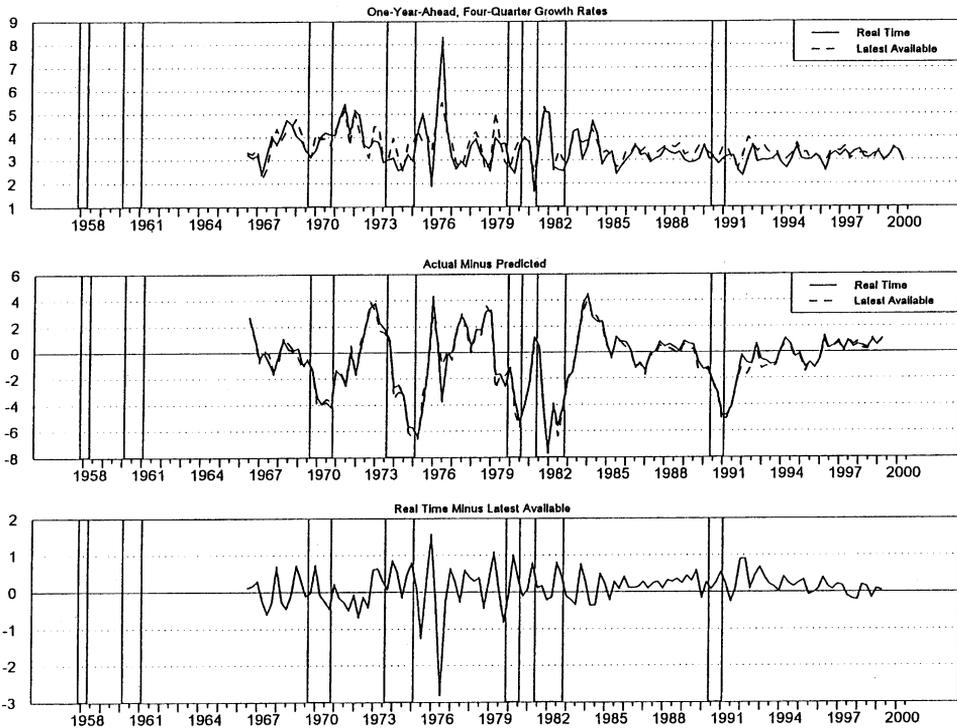


Fig. 4. (a) A comparison of two real GDP forecasts from a rolling AR(4) model. (b) A comparison of two forecast errors from a rolling AR(4) model. (c) Difference between errors. *Note:* These three plots compare forecasts made using an AR(4) model in a rolling fashion, adding an additional quarterly data point each period to generate a forecast for the average annual growth rate of real output over the next year. The top plot shows the forecasted growth rates, one from generating forecasts using the real-time data set and the other from generating forecasts using latest available data (vintage August 1999). The middle plot compares the forecast errors from the two different forecasts. The bottom plot shows the difference between the forecast errors, illustrating how much the use of real-time data matters. The latest available vintage (August 1999) is used as “actual” in calculating forecast errors.

Using an AR(4) model on the first-difference of log real output, we find that the two forecasts look somewhat different over time but not dramatically so (top panel of Fig. 4). Forecast errors appear similar (middle panel) but can be quite different at certain times (bottom panel), such as in 1976. A scatter plot shows a positive relationship between the two sets of forecasts, but there are differences between the forecasts (Fig. 5). Evidently, the vintage of the data matters even for such simple forecasts as these AR(4) forecasts. However, taking the August 1999 vintage as representing the actual value for the data, the root-mean-squared-forecast error (RMSE) is not very different when forecasts are based on real-time data (2.38) as opposed to final revised

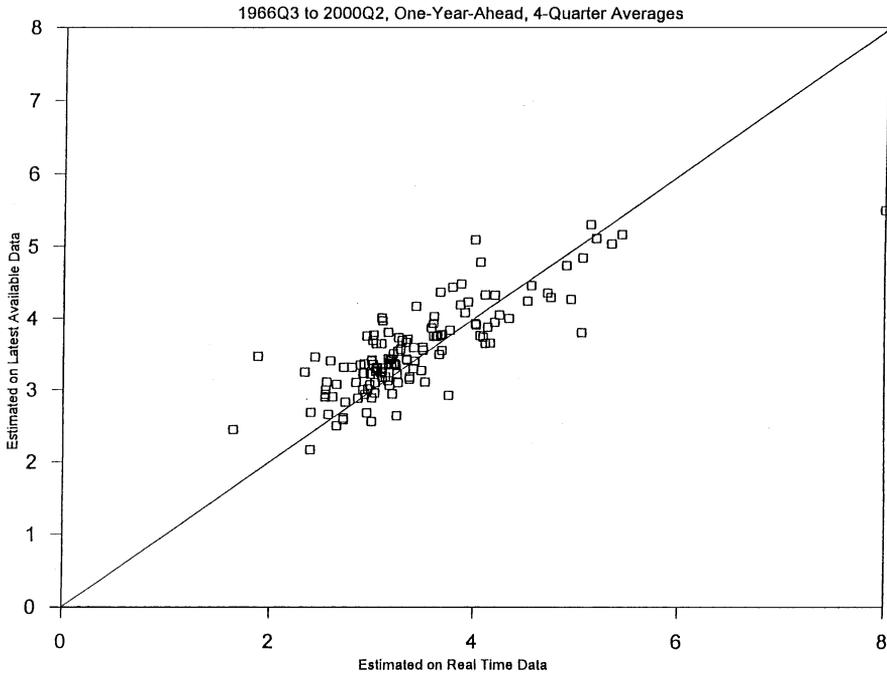


Fig. 5. Two real GDP growth forecasts from a rolling AR(4) specification. *Note:* This scatter plot shows the forecasts from Fig. 4 along a 45-degree line so that the differences in forecasts based on the alternative data sets are readily apparent.

data from the August 1999 vintage (2.40). That is actually quite surprising because it says that having today's vintage gives no better forecast performance than having available just real-time data, when the goal is to forecast the data as they appear today; or it may simply mean that forecasting a variable such as real output growth using time-series methods is not a very productive enterprise in the sense that the forecast errors associated with the AR(4) model are large relative to the revisions of the data.

An additional benefit of the real-time data set is that it can be used by a researcher to choose a different set of values to use as "actuals" in calculating forecast errors. In the results we have discussed so far, the latest vintage (August 1999) was used to represent the actuals, but alternative choices are sensible as well. One could argue that a forecaster should not need to forecast changes in NIPA methodology, and thus that data from not long after the forecast was made should be used as actuals. Alternatively, one could argue that forecasters should be expected to forecast variables reasonably well until a comprehensive NIPA revision is made. We can easily use the real-time data set to calculate the forecast errors defined on these alternative actuals.

If we use as actuals the real-time values of real output growth one year after the end of the forecast period, we get an RMSE of 2.67 forecasting with real-time data as opposed to 2.70 forecasting with final revised data. If we use as actuals the data from the benchmark vintage just before the next comprehensive revision, we find an RMSE of 2.80 forecasting with real-time data as opposed to 2.85 forecasting with final revised data.¹⁷ Somewhat larger differences (around 0.2) between the forecasts using the alternative data sets occur if we compute the RMSEs over nonoverlapping five-year intervals, rather than over the entire period.

This exercise illustrates the idea that forecasts for selected sample periods may be substantially affected by data vintage, as Figs. 4 and 5 show. We have done other experiments (not reported here) that show similar results with other forecasting models (see Croushore and Stark, 1999a).

6. Conclusions

This paper describes our real-time data set for macroeconomists, explains how the data were assembled, and shows the extent to which some data revisions are potentially large enough to matter for forecasting. Forecasts based on real-time data are certainly correlated positively with forecasts based on final data, but data revisions to real output may cause forecasts based on latest-available data to be considerably different from forecasts based on real-time data over selected sample periods. To be fair, the results of our empirical exercise suggest that when evaluated over long periods, forecast-error statistics are not sensitive to the distinction between real-time and latest-available data, even though forecasts for isolated periods can diverge. We suspect, however, that this result may not generalize to other classes of models, and in particular, depends crucially on the persistence of the process of revisions, the persistence of the process describing the variable being modeled, and on the type of model.¹⁸ At a minimum, we believe further research is warranted and that our empirical findings sound a cautionary note for studies claiming that some new, improved forecasting method is superior to other methods, if the study presents only evidence based on latest-available data rather than real-time data.

Our hope is that the real-time data set presented in this paper and available on our web site will serve as a standard for forecasters and others engaged in research that may be affected by data revisions.

¹⁷ None of these differences in root mean square errors are large enough to be statistically significant, based on tests for comparing forecast accuracy, like those of Diebold and Mariano (1995).

¹⁸ See, for example, the results of Cole (1969), who found dramatic differences in forecasts because of data revisions.

Acknowledgements

We thank Jim Sherma, Bill Wong, and Bill Dalasio for their fine research assistance. We thank Athanasios Orphanides, Ellis Tallman, participants in seminars at the Federal Reserve Bank of Philadelphia, the University of Pennsylvania, the International Finance Division of the Federal Reserve Board, and George Washington University, as well as those at the Midwest Macroeconomics meetings, the Pennsylvania Economics Association, the Federal Reserve System Committee on Macroeconomics, and the National Bureau of Economics Summer Institute, for their comments. Thanks also to three referees and the journal editors (Frank Diebold and Ken West) for their comments. The views expressed in this paper are those of the authors and do not necessarily reflect those of the Federal Reserve Bank of Philadelphia or of the Federal Reserve System.

References

- Amato, J.D., Swanson, N.R., 2001. The real time predictive content of money for output. *Journal of Monetary Economics*, forthcoming.
- Boschen, J.F., Grossman, H.I., 1982. Tests of equilibrium macroeconomics using contemporaneous monetary data. *Journal of Monetary Economics* 10, 309–333.
- Cole, R., 1969. Data errors and forecasting accuracy. In: Jacob Mincer (Ed.), *Economic Forecasts and Expectations: Analyses of Forecasting Behavior and Performance*. National Bureau of Economic Research, New York, pp. 47–82.
- Conrad, W., Corrado, C., 1979. Application of the Kalman filter to revisions in monthly retail sales estimates. *Journal of Economic Dynamics and Control* 1, 177–198.
- Croushore, D., 1993. Introducing: the survey of professional forecasters. *Federal Reserve Bank of Philadelphia Business Review* Nov./Dec. 1993, 3–13.
- Croushore, D., Evans, C.L., 1999. Data revisions and the identification of monetary policy shocks. Manuscript, Federal Reserve Bank of Philadelphia, December.
- Croushore, D., Stark, T., 1999a. Does data vintage matter for forecasting? Federal Reserve Bank of Philadelphia Working Paper 99-15, October.
- Croushore, D., Stark, T., 1999b. A real-time data set for macroeconomists: does the data vintage matter? Federal Reserve Bank of Philadelphia Working Paper 99-21, December.
- Denton, F.T., Kuiper, J., 1965. The effect of measurement errors on parameter estimates and forecasts: a case study based on the Canadian preliminary national accounts. *Review of Economics and Statistics* 47, 198–206.
- Diebold, F.X., Mariano, R.S., 1995. Comparing predictive accuracy. *Journal of Business and Economic Statistics* 13, 253–263.
- Diebold, F.X., Rudebusch, G.D., 1991. Forecasting output with the composite leading index: a real-time analysis. *Journal of the American Statistical Association* 86, 603–610.
- Evans, C.L., 1998. Real-time Taylor rules and the federal funds futures market. *Federal Reserve Bank of Chicago Economic Perspectives*, Third Quarter 1998, 44–55.
- Fair, R.C., Shiller, R.J., 1990. Comparing information in forecasts from econometric models. *American Economic Review* 80, 375–389.
- Ghysels, E., Swanson, N.R., Callan, M., 1999. Monetary policy rules and data uncertainty. manuscript, Texas A&M University.

- Harvey, A.C., McKenzie, C.R., Blake, D.P.C., Desai, M.J., 1981. Irregular data revisions. In: Arnold Zellner (Ed.), *Applied Time Series Analysis of Economic Data*. U.S. Department of Commerce, Washington, DC, Economic Research Report ER-5, pp. 329–339.
- Howrey, E.P., 1978. The use of preliminary data in econometric forecasting. *Review of Economics and Statistics* 60, 193–200.
- Howrey, E.P., 1996. Forecasting GNP with noisy data: a case study. *Journal of Economic and Social Measurement* 22, 181–200.
- Keane, M.P., Runkle, D.E., 1990. Testing the rationality of price forecasts: new evidence from panel data. *American Economic Review* 80, 714–735.
- Koenig, E.F., Dolmas, S., 1997. Real-time GDP growth forecasts. Federal Reserve Bank of Dallas Working Paper 97-10, December.
- Koenig, E.F., Dolmas, S., Piger, J., 2000. The use and abuse of ‘real-time’ data in economic forecasting. Federal Reserve Bank of Dallas Working Paper 00-04, September.
- Landefeld, J.S., Parker, R.P., 1995. Preview of the comprehensive revision of the national income and product accounts: BEA’s new featured measures of output and prices. *Survey of Current Business*, July 1995, 31–38.
- Landefeld, J.S., Parker, R.P., 1997. BEA’s chain indexes, time series, and measures of long-term economic growth. *Survey of Current Business*, May 1997, 58–68.
- Mankiw, N.G., Runkle, D.E., Shapiro, M.D., 1984. Are preliminary announcements of the money stock rational forecasts? *Journal of Monetary Economics* 14, 15–27.
- Mankiw, N.G., Shapiro, M.D., 1986. News or noise: an analysis of GNP revisions. *Survey of Current Business*, May 1986, 20–25.
- Maravall, A., Pierce, D.A., 1986. The transmission of data noise into policy noise in U.S. monetary control. *Econometrica* 54, 961–979.
- Morgenstern, O., 1963. *On the Accuracy of Economic Observations*. Princeton University Press, Princeton.
- Mork, K.A., 1987. Ain’t Behavin’: forecast errors and measurement errors in early GNP estimates. *Journal of Business and Economic Statistics* 5, 165–175.
- Orphanides, A., 2001. Monetary rules based on real-time data. *American Economic Review*, forthcoming.
- Orphanides, A., 1998. Monetary policy evaluation with noisy information. Manuscript, Federal Reserve Board of Governors, October.
- Patterson, K.D., Heravi, S.M., 1991. Data revisions and the expenditure components of GDP. *Economic Journal* 101, 887–901.
- Pierce, D.A., 1981. Sources of error in economic time series. *Journal of Econometrics* 17, 305–321.
- Ritter, J.A., 2000. Feeding the national accounts. *Federal Reserve Bank of St. Louis Review* 82, March/April 2000, 11–20.
- Robertson, J.C., Tallman, E.W., 1998a. Data vintages and measuring forecast model performance. *Federal Reserve Bank of Atlanta Economic Review*, Fourth Quarter, 1998, 4–20.
- Robertson, J.C., Tallman, E.W., 1998b. Real-time forecasting with a VAR model. Manuscript, Federal Reserve Bank of Atlanta, November.
- Rudebusch, G.D., 1998. Do measures of monetary policy in a VAR make sense? *International Economic Review* 39, 907–931.
- Rudebusch, G.D., 2001. Is the fed too timid? Monetary policy in an uncertain world. *Review of Economics and Statistics* 83, 203–217.
- Runkle, D.E., 1998. Revisionist history: how data revisions distort economic policy research. *Federal Reserve Bank of Minneapolis Quarterly Review*, Fall 1998, 3–12.
- Sargent, T., 1989. Two models of measurements and the investment accelerator. *Journal of Political Economy* 97, 251–287.

- Stark, T., 1998. A Bayesian vector error corrections model of the U.S. economy. Federal Reserve Bank of Philadelphia Working Paper 98-12, June.
- Stekler, H.O., 1967. Data revisions and economic forecasting. *Journal of the American Statistical Association* 62, 470–483.
- Swanson, N., 1996. Forecasting using first-available versus fully revised economic time-series data. *Studies in Nonlinear Dynamics and Econometrics* 1, 47–64.
- Swanson, N.R., Ghysels, E., Callan, M., 1999. A multivariate time series analysis of the data revision process for industrial production and the composite leading indicator. In: Robert F. Engle, Halbert White (Eds.), *Cointegration, Causality, and Forecasting*. Oxford University Press, Oxford.
- Swanson, N.R., White, H., 1997. A model selection approach to real-time macroeconomic forecasting using linear models and artificial neural networks. *Review of Economics and Statistics* 79, 540–550.
- Trivellato, U., Rettore, E., 1986. Preliminary data errors and their impact on the forecast error of simultaneous-equations models. *Journal of Business and Economic Statistics* 4, 445–453.
- Zellner, A., 1958. A statistical analysis of provisional estimates of gross national product and its components, of selected national income components, and of personal saving. *Journal of the American Statistical Association* 53, 54–65.