

Revisionist History: How Data Revisions Distort Economic Policy Research

David E. Runkle
Research Officer
Research Department
Federal Reserve Bank of Minneapolis

Abstract

This article describes how and why official U.S. estimates of the growth in real economic output and inflation are revised over time, demonstrates how big those revisions tend to be, and evaluates whether the revisions matter for researchers trying to understand the economy's performance and the contemporaneous reactions of policymakers. The conclusion may seem obvious, but it is a point ignored by most researchers: To have a good chance of understanding how policymakers make their decisions, researchers must use not the final data available, but the data available initially, when the policy decisions are actually made.

The views expressed herein are those of the author and not necessarily those of the Federal Reserve Bank of Minneapolis or the Federal Reserve System.

During 1974 and 1975, the U.S. economy reeled from the effects of huge oil price increases. At that time, data suggested that the economy was in a recession, a recession by far more severe than any since shortly after the end of World War II. Initial estimates suggested that real output (adjusted for inflation) fell 5.8 percent between the second quarter of 1974 and the second quarter of 1975. But those estimates have been greatly revised during the past 20 years. Today the best estimate is that real output declined only 2.0 percent between mid-1974 and mid-1975.

The fact of data revision is generally well known, but few academic studies have considered the effects that it can have on the conduct and understanding of economic policy.¹ This is an unfortunate omission. Revisions in estimates of real output growth and inflation have historically been large, and these revisions can cause at least two types of significant distortions. Most obviously, the data initially available provide neither an accurate nor an unbiased prediction of the final, revised data. That is, at the time the data are first released, they do not provide an accurate picture of how the economy is actually performing. Most important here, data revisions can significantly distort economic research. Using final data rather than initial data will mislead anyone trying to understand the historical relationship between the economy's performance and contemporaneous economic policy decisions.

Measuring Economic Conditions

Before demonstrating how data revisions can distort policy research, let's examine how and why U.S. economic data are revised and evaluate how big those revisions tend to be.

Data Series

Every quarter, the Bureau of Economic Analysis (BEA) of the U.S. Department of Commerce releases new estimates of the growth in real economic output and inflation in the United States. These data are crucial, both for economic policymakers and for economists who want to analyze the relationship between economic conditions and policymakers' decisions.

Over the last 30 years, the primary measures of U.S. real growth and inflation have changed. Until late in 1991, the main series used to measure real growth was growth in the inflation-adjusted level of the gross national product (GNP). In the fourth quarter of 1991, attention shifted to the gross domestic product (GDP). (See Parker 1991a, b.) A similar shift occurred for inflation. Until late 1991, the BEA's primary inflation measure was growth in the implicit price deflator for GNP. (The *implicit price deflator* is the measure that translates nominal values into values adjusted for inflation, or *real* values.) In the fourth quarter of 1991, the primary inflation measure became growth in the implicit GDP price deflator.²

Late in 1995, the measures of real growth and inflation changed again. The BEA changed an assumption behind its computation of the implicit price deflator. Until the fourth quarter of 1995, the way the BEA computed the deflator implicitly assumed that the market basket of goods and services remained unchanged across time. In that quarter, the BEA attempted to rectify that distortion by shifting to what is called a *chain-weighted implicit price deflator* (discussed in Landefeld and Parker 1995).

In this study, I use the measures of real output growth and inflation that were emphasized by the BEA at each particular time.³ The data sample contains 144 observations on each series from the first quarter of 1961 through the fourth quarter of 1996. All national income and product account (NIPA) data for this study are available starting in 1960. Since I look at growth in output and the deflator over as many as four quarters, the data sample starts in 1961. For every quarter, I compute both the *initial* and the *most recent* estimates of growth in real output and inflation over the past one, two, and four quarters. I call the most recent estimate the *final* estimate even though that estimate may eventually be revised again.

For many years, growth and inflation numbers have been used as benchmarks both by policymakers and by economists who analyze policy.⁴ The initial estimates of real output growth and inflation used here are the first data for each quarter published in the BEA's monthly publication, the *Survey of Current Business*. These first data are typically released in the first month after a calendar quarter. For example, the estimates of real growth for the one, two, and four quarters ending in the second quarter of 1975 were released near the end of July 1975. (All the data used to compute the initial estimates of real growth and inflation have been collected from past issues of the *Survey of Current Business*.⁵ Final estimates have been electronically retrieved from the BEA.)

Regular Revisions

Between the time that the BEA makes its initial and final estimates of real growth and inflation, the data have been revised many times.

The first set of revisions over the months after the initial estimates take into account data not available when the initial estimates were made. (For a description of this revision process, see Young 1987, 1993.) In addition to those early revisions, the BEA also revises all data three times, in July of the three years after the initial estimate. These July revisions use additional information, such as tax data, that is only available annually, and these estimates are more reliable than the initial estimates.

Data are not revised simply because more information becomes available, however. Revisions also occur because of changed definitions and classifications and regular comprehensive renovations in the NIPA. Comprehensive revisions have historically been completed every five or ten years, and before 1995, they involved completely rewriting the history of real growth and inflation.

Comprehensive revisions involve changing the base year for determining real output. For instance, the most recent comprehensive revision changed the reference year for GDP dollars from 1987 to 1992. When comprehensive revisions are made, the BEA also examines whether definitions and classifications used in the NIPA are still the best available. Often, changes in definitions and classifications can have significant effects on the measurement of real growth.⁶ These revisions provide a different understanding of real growth and inflation than that provided by the initial estimates.⁷

Almost all empirical economic research involving data on real growth and inflation uses final data. Note, however, that those final estimates are not available to policymakers earlier, when they have to make decisions. Do the revisions really matter for this research or for policymaking?

So What?

Yes, data revisions matter. There are at least three ways to demonstrate this. A simple graphical display of the differences between the initial and final estimates can show that the revisions have been large. Statistical tests can show, contrary to what many economists presume, that initial estimates of real growth and inflation are not unbiased forecasts of the final estimates and that the two estimates can differ substantially. And a comparison of revision uncertainty and forecast uncertainty can show that revision uncertainty is a significant fraction of the quite large uncertainty about forecasts of real growth and inflation.

□ Size

Charts 1 and 2 show the difference between the BEA's initial and final estimates of growth in real output and the deflator for each quarter, expressed as annual rates. The charts show that the quarterly revisions can be quite large. Between 1961 and 1996, quarterly real growth estimates were revised upward by as much as 7.5 percentage points and downward by as much as 6.2 percentage points. This can make the difference between a recession and a simple slowdown. For example, real GNP was initially thought to have dropped between the third and fourth quarters of 1974 at an annual rate of 9.1 percent, as severe a fall in output as occurred during the Great Depression. But the final estimate indicates that between those quarters, real GNP dropped at an annual rate of only 1.6 percent. This is an upward revision of 7.5 percentage points.

Similarly, between 1961 and 1996, inflation estimates for individual quarters were revised upward by as much as 5.5 percentage points and downward by as much as 4.0 percentage points. The absolute value of the difference between the initial estimate and the final estimate, on an annual basis, was more than 1.6 percentage points for real growth and 1.0 percentage point for inflation.

Charts 3 and 4 show that revisions can be very large even for growth over a four-quarter horizon. These charts show the difference between the final estimate of how much real output and the deflator grew over any particular four-quarter period and the initial estimates of those growth rates that were made just after a quarter ended. Clearly, the revisions of growth over four quarters can be substantial.

□ Bias

A second way to evaluate the importance of data revisions is to test statistically whether the initial estimates of real growth and inflation are *unbiased* predictors of the final estimates in any particular period. This means to ask, Are the initial estimates neither too high nor too low on average?⁸ A systematic error in the revision process would greatly compromise the ability to judge historic policy decisions based on today's estimates of what happened in the past. Such a bias would challenge most of the academic evaluations of economic policy: researchers generally use final data, under the assumption that the decision to use initial or final data doesn't matter much.

Here I test whether the initially published one-, two-, and four-quarter growth rates for real output and the deflator are on average neither too high nor too low compared to the final data. I run the following regression:

$$(1) \quad F_t^i = \alpha_0 + \alpha_1 I_t^i + \varepsilon_t$$

where F_t^i is the final estimate of growth in the initial quarters ending at t , I_t^i is the initial estimate of that growth, ε_t is the error in the regression, and the α 's are coefficients to be estimated. In equation (1), if $\alpha_0 = 0$ and $\alpha_1 = 1$, then the initial estimate can be considered an *unbiased predictor* of the final estimate of real growth and inflation in the past.

The hypothesis that the initial estimates are unbiased is my null hypothesis. I test that hypothesis by comparing the estimated regression coefficients to their hypothesized values. If that difference is large, relative to the amount of uncertainty about the coefficients, then I must reject the null hypothesis. A chi-square test lets me make exactly this comparison. (For details on chi-square tests, see, for example, Theil 1971.) If the chi-square test statistic is too large, then the hypothesis of unbiasedness can be rejected.

The accompanying table shows the results of running this regression, with the coefficients estimated by the generalized method-of-moments procedure.⁹ Clearly, this null hypothesis can be rejected at the 1 percent level. The initial estimates of real growth and inflation are *biased*, not unbiased, forecasts of the final estimates of real growth and inflation.¹⁰

Charts 5 and 6 confirm what the table shows: Initial and final estimates of these variables can be quite different. Plotted against each other, the initial and final estimates are scattered all over the charts, not nicely positioned along the 45-degree line, as they would be if the initial and final estimates were the same.¹¹ Thus, the researcher's choice of which estimates to use does matter.

□ Uncertainty

A third way to evaluate the importance of data revisions is to show that revision error is significant compared to forecast error for the key data series. That is, data revisions are important when the uncertainty about where the economy has been is a significant fraction of the uncertainty about where the economy is going.

One simple way to predict either real growth or inflation is to use only the past values of that particular variable. A model that does that is called a *univariate autoregression*. An example of this model is to predict quarterly real growth using only the past two values of quarterly real growth:

$$(2) \quad X_t = \alpha_0 + \alpha_1 X_{t-1} + \alpha_2 X_{t-2} + \varepsilon_t$$

where X_t is the growth rate of real output in quarter t , ε_t is the forecast error in the model, and again the α 's are coefficients to be estimated.

The uncertainty about economic forecasts of real output growth can be measured by the estimated standard deviation of ε_t . In general, this forecast uncertainty is known to be fairly large. Based on an estimation of equation (2) for quarterly real output with annualized data on one-quarter real growth from 1961 through 1996, the estimated standard deviation of the forecast error is 3.11 percent.

Uncertainty about data revisions is nearly that large. According to equation (2), the standard deviation of the revision error in annualized quarterly real growth (that is, the standard deviation of the final quarterly data minus the ini-

tial quarterly data) is 2.21 percent, or 71 percent of the standard deviation of the one-quarter-ahead forecast error in quarterly real growth. Thus, uncertainty about the size of the revision in quarterly growth between the time of the initial estimate and the final estimate is a fairly large fraction of the uncertainty about one-quarter-ahead real growth.¹²

The standard deviations of the forecast error and the revision error for inflation are both smaller than those for real output growth. However, for inflation, the revision error is still a significant fraction of the one-quarter-ahead forecast error. For inflation, the standard deviation of the forecast error from the univariate autoregression model is 1.59 percent, while that of the revision error is 1.22 percent. In terms of standard deviations, therefore, the revision error is 77 percent of the one-quarter-ahead forecast error.¹³

Analyzing Economic Policies

So, data revisions can have a large effect on the perceived history of real growth and inflation in the United States. Data revisions can be large, and initial estimates of real growth and inflation are not rational forecasts of final estimates. Now let's examine how these misperceptions can distort views about how economic policy is made.

One active area of research in macroeconomics has been attempts to determine how economic conditions have affected economic policymaking. For example, researchers have analyzed how changes in real growth and inflation appear to have affected the federal funds rate targeted by the monetary policymaking arm of the Federal Reserve System, the Federal Open Market Committee (FOMC) (Taylor 1993; Cochrane 1994; Christiano and Eichenbaum 1995; and Leeper, Sims, and Zha 1996). But evaluating how the FOMC has actually determined economic policy requires examining how the FOMC members have responded to the data they had when they made their decisions, not to the final data they did not have. As obvious as that may seem, few, if any, researchers take it into account.

Let's see what difference the initial and final data can make in policy analysis. Probably the best-known example of how researchers have used economic data to study FOMC policy is Taylor's (1993) model of economic policy. This model is an attempt to describe the rule that the FOMC has historically used when it has decided where to aim the federal funds rate. The model is an equation that has become known as the *Taylor rule*:

$$(3) \quad r = p + 0.5y + 0.5(p - 2) + 2$$

where

r = the federal funds rate

p = the rate of inflation over the preceding four quarters

y = the percentage deviation of real output from a target.

That is,

$$(4) \quad y = 100(Y - Y^*)/Y^*$$

where

Y = real output

Y^* = trend real output.

Taylor's study shows that this rule does an extraordinarily good job of describing FOMC policy from 1987 through 1992 when the rule is given final estimates available in 1993.¹⁴ But giving the Taylor rule initial estimates yields very different results.

Chart 7 shows how different these results are over the entire sample period (1961–96).¹⁵ Note that several times during this period, the difference between the federal funds rate predictions based on initial and final data exceeds two percentage points, and once it exceeds four percentage points. Since FOMC members often debate whether to aim to raise or lower the federal funds rate just one-quarter of a percentage point, these differences are substantial.

Using initial data to construct the Taylor rule appears to be a more accurate technique than using final data. The mean error between the Taylor rule's prediction and the actual federal funds rate is 83 percent larger when final data are used than when initial data are. The variance of that error is 48 percent larger using final data than initial data.¹⁶

What does this difference mean for an analysis of the relationship between economic conditions and policies? Let's look at a particular period. Recall the period of the oil price shocks, in the mid-1970s. We have seen that the U.S. economy's downturn during that time was initially thought to be much more severe than it is now believed to have been. This difference between the initial and final data has a big effect on the Taylor rule's estimate of how the FOMC would have responded to economic conditions from the third quarter of 1974 to the third quarter of 1975. Applied to that period and given the initial data, the Taylor rule predicts that the federal funds rate would have been reduced 5.9 percentage points. However, given the final data, the Taylor rule predicts a reduction of only 1.9 percentage points.

The FOMC actually did act to reduce the federal funds rate 5.8 percentage points between the third quarters of 1974 and 1975. This should not be surprising because, again, in that period, much of the data the FOMC had suggested that the economy was in a severe recession.¹⁷ However, anyone using the Taylor rule and the final data to understand how the FOMC generally responds to economic conditions would be seriously misled.

Concluding Remarks

Initial views of economic activity at any particular time can differ substantially from what will become the historical views of that period. Consequently, anyone trying to understand recent economic history and the reaction of policymakers must be careful about which data they use. To have a good chance of understanding, they must use not the final data, but the data available initially, when the policy decisions were actually made.

Note: I do not mean to be critical of the data collection and processing efforts of the BEA or of the policymaking efforts of the FOMC. Both institutions do the best they can, given the available information. Naturally, more information about the economy becomes available over time. The members of the FOMC cannot know exactly how data will later be revised, so they must form policy based on

the best information available when decisions must be made. My message here is primarily to research economists: Don't assume that policymakers' foresight about data revisions is 20/20.

*Also, Adjunct Professor of Finance, University of Minnesota.

¹Mankiw and Shapiro (1986) and Young (1987, 1993) examine the statistical properties of data revisions over much shorter periods than I discuss. None of that work examines the effect that data revisions can have on the understanding of economic conditions and policymaking.

²The GDP series measures output by domestic workers in the United States. The GNP series measures the output of domestically owned factors of production, including production abroad.

³For data through the third quarter of 1991, I use GNP; for data starting in the fourth quarter of 1991, I use GDP; and starting with the fourth quarter 1995 data, I switch to real GDP measured with a chain-weighted deflator.

⁴For example, both real growth and nominal growth in GNP (which takes into account both real growth and inflation) are predicted by the Federal Reserve Board in the chair's twice-annual testimony to Congress, mandated by the Humphrey-Hawkins Act (the Full Employment and Balanced Growth Act of 1978).

⁵The implicit price deflator did not consistently appear in the *Survey of Current Business* in the earliest years of my data sample. Therefore, to obtain the maximum sample for this study, I collected data on the initial estimates of the past five levels of nominal GNP (GDP) and real GNP (GDP) and used them to compute the deflator.

⁶For example, in the last 15 years, the BEA has changed the way it measures government investment, residential investment, and retirement funds for the military.

⁷Revisions are not necessarily improvements in the data. Conceptual problems in measuring real output, for example, can cause potentially serious measurement problems for data in the remote past, when a then-recent base year was used for prices.

⁸This issue was originally discussed by Mankiw, Runkle, and Shapiro (1984) and Mankiw and Shapiro (1986). The latter study, using data from 1975 to 1982, examines whether initial estimates of GNP growth were unbiased forecasts of revised estimates and finds that they were. Since Mankiw and Shapiro's study was published, many economists have assumed that its results justify using final data for policy analysis. The difference between Mankiw and Shapiro's results and mine are likely due to the fact that my sample is more than four times longer than theirs. In addition, the data in my study are subject to many more comprehensive (benchmark) revisions and changes in definitions and classifications than are those in the Mankiw and Shapiro study. There is no reason to assume that the initial estimates actually are unbiased forecasts of the final revised data. But even if they are not, using final data to understand how policymakers perceived economic conditions at the time of initial data releases can be misleading.

⁹For a discussion of generalized method-of-moments estimation, see Hansen 1982. Note that the number of moving-average (MA) terms differs depending on the number of quarters over which real growth and inflation are calculated. If, for each one-quarter period, the initial growth or inflation estimate was a rational forecast of the final estimate and all revisions were independent by quarter, then the error term in the regression would be $MA(I - 1)$, where I is the number of quarters over which growth or inflation is computed (Hansen and Hodrick 1980).

¹⁰Of course, if these biases in the initial data could have been predicted when the initial data were released, then policymakers could have taken the biases into account when making their decisions. However, it is highly unlikely that policymakers can accurately estimate the changes that will be made in comprehensive revisions released many years later.

¹¹The correlation between the one-quarter initial and final growth rates is 0.79 for real output and 0.90 for the deflator. Data from the first quarter of 1961 to the fourth quarter of 1996 were used in these computations and in the charts.

¹²The standard deviation of the one-quarter-ahead forecast error for real growth reported here comes from using two lags of growth in the forecasting regression. The results show little sensitivity to the lag length of the regression.

There are two problems with calling the standard deviation of the residuals from this model a true measure of the uncertainty in predicting real growth. One problem is that the data used for this model are all final revised growth rates—data not available to forecasters when they are trying to make predictions. The other problem is that the residuals from this model are all in-sample forecasts. However, this is how time series analysis of forecast uncertainty is actually done.

¹³As with the data for real growth, these results show little sensitivity to the lag length of the regression.

¹⁴Neither Taylor nor I assume that FOMC members actually use the Taylor rule to determine policy. FOMC members use a wide range of information on many more economic variables than those included in Taylor's simple equation.

¹⁵Here the final and initial data were estimated to have different trend rates of real output growth.

¹⁶Researchers who believe that reaction functions, such as the Taylor rule, describe FOMC policy often estimate those functions. To correctly estimate a reaction function with initial data, researchers need to reestimate it for every period, using only the data that were available in that period. I don't do that here because I am merely illustrating the differences between using initial and final data in the reaction function that Taylor estimated.

¹⁷Even though FOMC members do not have final data when they make their decisions, they clearly understand that initial data are noisy. Therefore, they use data from many sources to estimate the ways in which initial data will be revised in the months to come. However, FOMC members cannot know how comprehensive revisions will eventually affect initial data. Economists who don't acknowledge that will surely misunderstand how the FOMC made its decisions based on available data.

References

- Christiano, Lawrence J., and Eichenbaum, Martin. 1995. Liquidity effects, monetary policy, and the business cycle. *Journal of Money, Credit, and Banking* 27 (November, Part 1): 1113–36.
- Cochrane, John H. 1994. Shocks. *Carnegie-Rochester Conference Series on Public Policy* 41 (December): 295–364.
- Hansen, Lars Peter. 1982. Large sample properties of generalized method of moments estimators. *Econometrica* 50 (July): 1029–54.
- Hansen, Lars Peter, and Hodrick, Robert J. 1980. Forward exchange rates as optimal predictors of future spot rates: An econometric analysis. *Journal of Political Economy* 88 (October): 829–53.
- Landefeld, J. Steven, and Parker, Robert 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* 75 (July): 31–38. Bureau of Economic Analysis, U.S. Department of Commerce.
- Leeper, Eric M.; Sims, Christopher A.; and Zha, Tao. 1996. What does monetary policy do? *Brookings Papers on Economic Activity* 2: 1–63.
- Mankiw, N. Gregory; Runkle, David E.; and Shapiro, Matthew D. 1984. Are preliminary announcements of the money stock rational forecasts? *Journal of Monetary Economics* 14 (July): 15–27.
- Mankiw, N. Gregory, and Shapiro, Matthew D. 1986. News or noise: An analysis of GNP revisions. *Survey of Current Business* 66 (May): 20–25. Bureau of Economic Analysis, U.S. Department of Commerce.
- Parker, Robert P. 1991a. A preview of the comprehensive revision of the national income and product accounts: Definitional and classificational changes. *Survey of Current Business* 71 (September): 23–31. Bureau of Economic Analysis, U.S. Department of Commerce.
- _____. 1991b. A preview of the comprehensive revision of the national income and product accounts: New and redesigned tables. *Survey of Current Business* 71 (October): 20–28. Bureau of Economic Analysis, U.S. Department of Commerce.
- Taylor, John B. 1993. Discretion versus policy rules in practice. *Carnegie-Rochester Conference Series on Public Policy* 39 (December): 195–214.
- Theil, Henri. 1971. *Principles of econometrics*. New York: Wiley.
- Young, Allan H. 1987. Evaluation of the GNP estimates. *Survey of Current Business* 67 (August): 18–42. Bureau of Economic Analysis, U.S. Department of Commerce.
- _____. 1993. Reliability and accuracy of the quarterly estimates of GDP. *Survey of Current Business* 73 (October): 29–43. Bureau of Economic Analysis, U.S. Department of Commerce.

Testing Initial Data Estimates for Bias

Based on Annualized Growth Rates
Over One, Two, and Four Quarters, 1961–96*

Variable	Equation	Horizon (No. of Quarters)	Coefficient Value (and Standard Error)		No. of Moving Average Terms	χ^2 for Null†
			α_0	α_1		
Real Growth	(1)	1	.0092 (.0026)	.7358 (.0665)	0	16.29**
	(2)	2	.0076 (.0022)	.7685 (.0563)	1	16.95**
	(3)	4	.0066 (.0018)	.7947 (.0473)	3	19.07**
Deflator Growth	(4)	1	.0083 (.0020)	.9242 (.0505)	0	38.47**
	(5)	2	.0071 (.0018)	.9382 (.0420)	1	33.16**
	(6)	4	.0058 (.0021)	.9588 (.0502)	3	20.48**

*The number of observations in the data sample is 144 for both variables.

†Null hypothesis = Initial estimates are unbiased predictors of final estimates.

**Chi-square test statistic is significant at the 1 percent level.

Source of basic data: U.S. Department of Commerce,
Bureau of Economic Analysis

Charts 1–4

How Much U.S. Data Estimates Change

Differences Between Initial and Final Estimates of Annualized Growth in Real Output and the Price Deflator, 1961–96*

Charts 1–2 Revisions in One-Quarter Growth Rates

Chart 1 Real Output

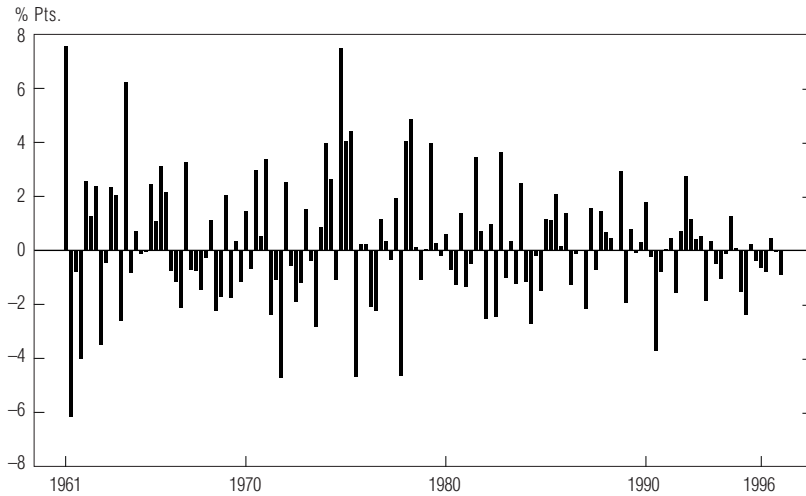
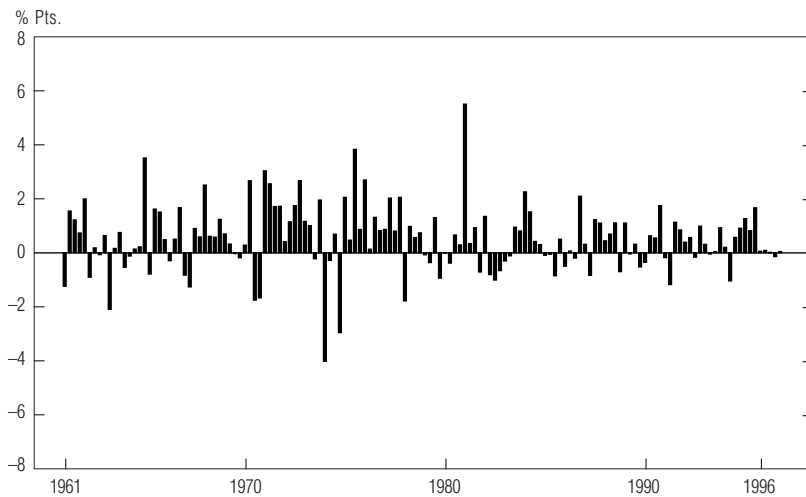


Chart 2 Price Deflator



*Output= Gross national product from 1st quarter 1961 through 3rd quarter 1991;
gross domestic product from 4th quarter 1991 through 4th quarter 1996.

Deflator= The measure appropriate for the current output measure.
Source: U.S. Department of Commerce, Bureau of Economic Analysis

Charts 3-4 Revisions in Four-Quarter Growth Rates

Chart 3 Real Output

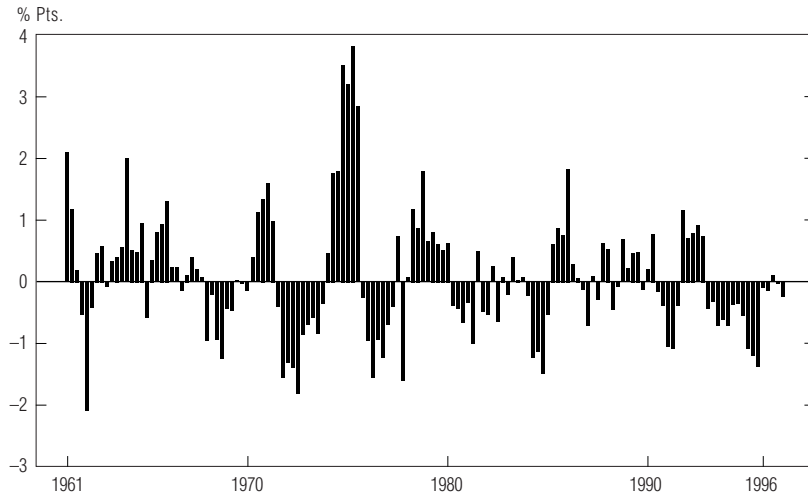
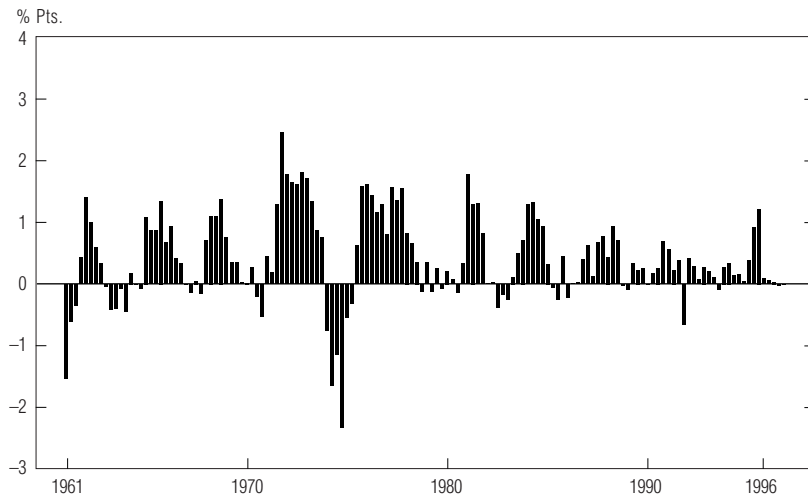


Chart 4 Price Deflator



Charts 5-6

Initial vs. Final Estimates of Quarterly Growth

Annualized Growth Rates, 1961-96

Chart 5 Real Output

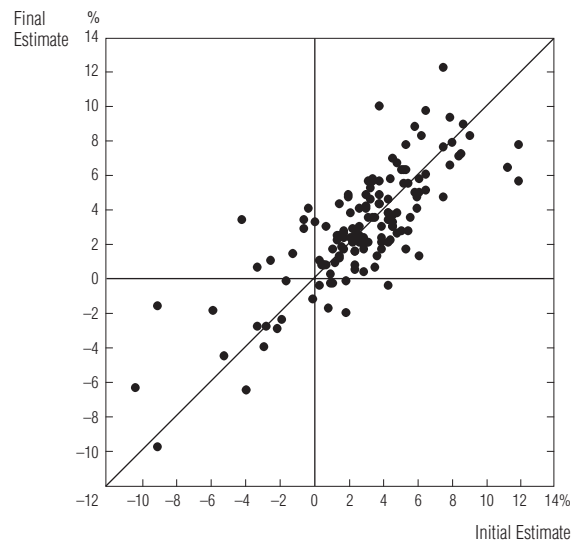
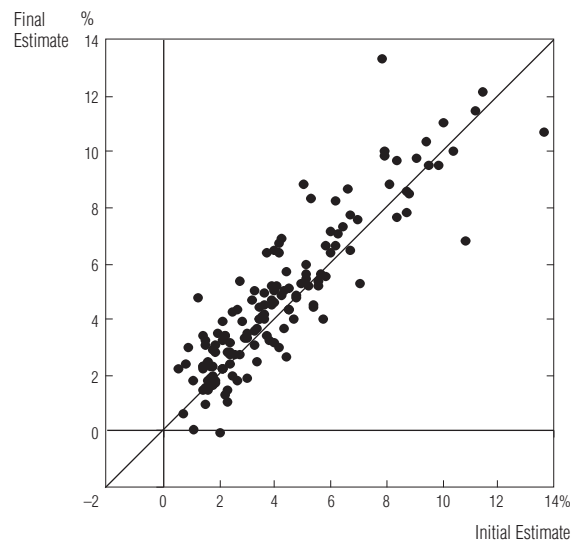


Chart 6 Price Deflator

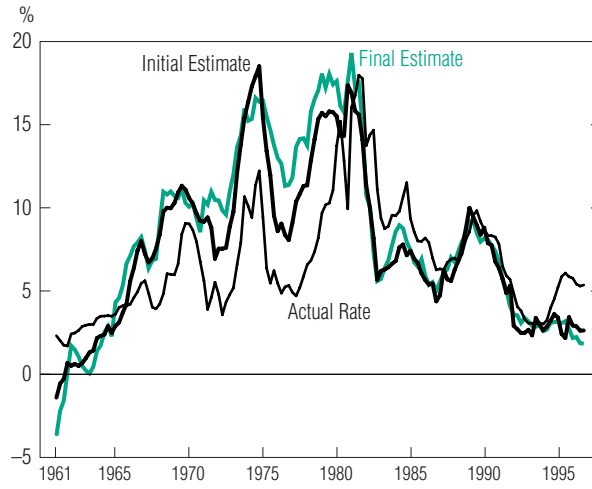


Source: U.S. Department of Commerce, Bureau of Economic Analysis

Chart 7

A Demonstration That Data Revisions Matter

Federal Funds Rate, Actual and Predicted by Taylor Rule
Using Initial and Final Data Estimates, Quarterly 1961–96



Sources: U.S. Department of Commerce, Bureau of Economic Analysis;
Board of Governors of the Federal Reserve System
