FEDERAL RESERVE BANK OF PHILADELPHIA

Ten Independence Mall
Philadelphia, Pennsylvania 19106-1574
(215) 574-6428, www.phil.frb.org

Working Papers

Research Department

WORKING PAPER NO. 99-4

A REAL-TIME DATA SET FOR MACROECONOMISTS

Dean Croushore and Tom Stark
Research Department
Federal Reserve Bank of Philadelphia

June 1999

WORKING PAPER NO. 99-4

A REAL-TIME DATA SET FOR MACROECONOMISTS

Dean Croushore and Tom Stark
Research Department
Federal Reserve Bank of Philadelphia

June 1999

We thank Jim Sherma, Bill Wong, and Bill Dalasio for their fine research assistance. We thank participants in seminars at the Federal Reserve Bank of Philadelphia and the University of Pennsylvania, as well as those at the Midwest Macroeconomics meetings, 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.

A REAL-TIME DATA SET FOR MACROECONOMISTS

Abstract

This paper presents the concept and uses of a real-time data set that can be used by economists for testing the robustness of published econometric results, for analyzing policy, and for forecasting. The data set consists of vintages, or snapshots, of the major macroeconomic data available at quarterly intervals in real time. The paper illustrates why such data may matter, explains the construction of the data set, examines the properties of several of the variables in the data set across vintages, examines key empirical papers in macroeconomics and investigates their robustness to different vintages, looks at how policy analysis may be affected by data revisions, and shows how forecasts can be affected by data revisions.

A REAL-TIME DATA SET FOR MACROECONOMISTS

I. INTRODUCTION

Macroeconomists use historical data for a variety of purposes: to test models, to analyze economic events, and to forecast. In many cases, however, the data that should be used in these studies are not the (final, revised) data available from government statistical agencies today, but rather the original, unrevised data available to economic agents who were around at the time. In other cases, the ability to verify published findings and to check the robustness of those findings to different data sets is an important test of the validity of the results.

These reasons motivated us to create a data set that gives a researcher a snapshot of the macroeconomic data available to an academic researcher, policymaker, or forecaster at any given date in the past. We refer to each data set corresponding to the information set at a particular date as a "vintage" and to the collection of such vintages as a "real-time data set."

This paper explains the reasons for the construction of this data set, describes the data set, and provides some empirical demonstrations of cases when the vintage matters and when it doesn't matter.

The type of analysis we perform in this paper is related to earlier literature. The most well-known study that compared results based on real-time data with later data was Diebold and Rudebusch (1991), who showed that the index of leading indicators does a much worse job of predicting future movements of output in real time than it does after the data are revised. More recently, Runkle (1998) has developed some ideas similar to those in this paper, using a real-time data set on real output to show how much vintage matters. There have been a number of attempts to examine how preliminary and incomplete data affect monetary policy, beginning with the seminal study of Maravall and Pierce (1986), who showed that even though the revisions to measures of the money supply are large, monetary policy wouldn't have been much different if more accurate data had been known. Recently, a number of studies have analyzed similar issues in the context of Taylor's rule for setting monetary policy. These studies include: (1) Orphanides (1997), who showed that Taylor's rule doesn't fit nearly as well in real time as it does using revised data; (2) Ghysels, Swanson, and Callan (1998), who showed that, contrary to the results of Maravall and Pierce, if the Fed used a Taylor-type rule and based policy decisions on changes in the index of industrial production, policy would have improved significantly if policymakers waited for data to be revised, rather than reacting to newly released data; (3) Evans (1998), who found that the federal funds futures market does a better job of forecasting the federal funds rate than does a Taylor-type rule, using real-time data; and (4) Rudebusch (1998b), who showed that although some research (which assumed that data don't get revised) suggests that the optimal coefficients in a Taylor-type rule are much bigger than Taylor originally suggested, data uncertainty potentially plays an important role in reducing the coefficients in the rule. Other approaches to using real-time data in the analysis of monetary policy include Amato (1998), who looks at the predictive power of M2 using real-time data, and Orphanides (1998), who uses a variety of interest-rate-based policy rules to examine the impact of data uncertainty on the optimal rule. In addition, Rudebusch (1998a) questions the value of VAR estimates of monetary-policy shocks because they aren't based on real-time data. There has been very little work done that uses real-time data for forecasting; two papers by Robertson and Tallman (1998a and 1998b) and a paper by Koenig and Dolmas (1997) are the only exceptions known to us.[1]

---

[1] In some papers, as in Koenig and Dolmas (1997), the term real time has a somewhat different meaning. In this paper, real time means that the forecasts made at any date are based on the currently available data set for all past dates. Koenig and Dolmas use real time to mean a data set consisting

Our goal is to provide a basic foundation for these types of studies and to provide benefits by allowing researchers to use a standard data set, rather than being forced to collect all the real-time data themselves for every different study. Section II of this paper provides details about the data set, including how it was constructed, which variables are available for the full time period and which are incomplete, and how the data set was checked for quality. In section III, we look at the properties of selected variables across vintages, to illustrate how much vintage matters for raw data. In section IV, we look more closely at real consumption data, examining the structure of the revisions. In section V, we look at some key empirical papers in macroeconomics and explore the degree to which vintage matters for their results. In section VI, we look at how policy analysis could be shaped by data vintage. Section VII examines how forecasts of different types may be sensitive to the choice of data vintage. We draw conclusions from these results in section VIII.

II. THE DATA SET

To help macroeconomic researchers avoid the types of problems discussed above, we have put together a sequence of real-time data sets. Conceptually, it's simple. We've gone back through original source documents of statistics and compiled a listing of the data as it existed in the middle of each quarter (on the 15th day of the month, to be precise), from November 1965 to the present.[2] So, for example, if you want to know what the data looked like in August 1968,

---

of preliminary data from each data set in the past--these data have been revised, but Koenig and Dolmas assume the forecaster ignores the revised data and uses just the preliminary data.

[2] 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. The

you'd simply need to pull down our data set for August 1968, and you would find data that looked exactly like you would see in published sources at that time: There is a time series for each variable from the first quarter of 1947 to the second quarter of 1968.[3]

The variables included in the data set are nominal and real GNP (GDP after 1991); the components of real GNP/GDP, including total personal consumption expenditures, broken down into durables, nondurables, and services; business fixed investment; residential investment; the change in business inventories; government purchases (government consumption and government investment since 1996); exports and imports; the chain-weighted GDP price index (since 1996); the M1 and M2 measures of the money supply; total reserves at banks (adjusted for changes in reserve requirements); nonborrowed reserves; nonborrowed reserves plus extended credit; the adjusted monetary base (the reserves measures and monetary base measures are from the Federal Reserve Board, not the versions from the St. Louis Fed); the civilian unemployment rate; the consumer price index (CPI-U); the three-month T-bill interest rate; and the 10-year Treasury bond interest rate. The interest rates are included for completeness, even though they are never revised. 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. The data sets are mostly complete; there are some missing data for the money stock variables and a lot of missing data for the monetary base and reserves variables.[4]

---

forecasters in that survey make their forecasts in the middle of each quarter.

[3] Actually, 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.

[4] The consumer price index is available on a seasonally adjusted basis only in the more recent data sets. However, since the seasonally unadjusted CPI series is not revised, it can be used without

Though the project of collecting these data seems simple, but arduous, it turned out that finding old data is 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 tried very carefully to include in the data sets only the data we knew were available at the time. In many cases, data were revised, but the publications that detailed the revisions did not always say when the data were made available. So it took a substantial amount of time to figure out exactly what data should have been, or should not have been, included in each data set. A comprehensive set of notes about the data sets is available to help researchers understand our conventions on including or excluding particular data. Also, some of the data have been collected in real time since this project began in 1991.

Who did all this work? Some of it was done over the last five 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!

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. In some cases, this was easy. For example, we made sure that the sum of the components of real GNP added up to total GNP at all dates, in a

---

concern about revisions. For complete notes on all the variables and any missing data, see the documentation files on our web page.

random sample of data sets. 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 typos in the data set.

Where can you find these data? The data set is easily accessible on the Philadelphia Fed's web site at http://www.phil.frb.org/econ/. From links on that 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 data sets shortly after the 15th day of the middle month of each quarter.

III. DATA REVISIONS

How big are the revisions to the data? We don't have space here to describe the revisions to all the data, so we'll look at certain key variables, including nominal output, real output, real consumption spending, and the price level.

First, let's see how much vintage matters for the medium run, that is, five-year average growth rates. Table III.1 shows the annual average growth rate over five-year periods from 1949Q4 to 1994Q4 for data from vintages dated November 1975, 1980, 1985, 1991, 1995, and 1998. The first five of these vintages were chosen because they were the last vintages prior to a comprehensive revision of the national income and product accounts; the last vintage, November 1998, is the latest available data. For ease of exposition, we'll call these benchmark vintages. Each of the comprehensive revisions that were made after our benchmark vintage dates incorporated major changes to the data, including new source data and definitional changes. In addition, the base year was changed for real variables in January 1976 (from 1958 to 1972), in December 1985 (from 1972 to 1982), in late November 1991 (from 1982 to 1987), and in

January 1996 (from 1987 to 1992), so some of the differences across the benchmark vintages we look at (1980, 1991, 1995, and 1998) incorporate base-year changes, which affect real variables. In particular, 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 are two other changes of note regarding the comprehensive revisions. First, the output variable (both real and nominal) is GNP before 1992, but GDP during and after 1992. Our data set is consistent with the "headline" variable, 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 nominal and real output arise because of this definitional change.[5] In the current exercise, keep in mind that differences in benchmark vintages before and after 1992 reflect this change.

The second major change in methodology comes from the switch to chain weighting in vintages during or after 1996. This represented a significant change in how real variables were constructed, one that greatly reduces the substitution bias. In particular, the switch to chain weighting means that a change of base year (which is arbitrary under chain weighting) will have no effect on the growth rates of variables, whereas the growth rates changed significantly under the old fixed-weighting method.

Reading across the columns of Table III.1 shows how the five-year annual average growth rate has changed across benchmark vintages. Nominal output from the 1950s and 1960s

[5] 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.

wasn't revised too much, but the data from the 1970s and early 1980s show changes of as much as 0.5 percentage point across vintages. Real output is strongly affected by changes in benchmark vintage, especially when the base year is changed. The differences are, on average, much larger than they are for nominal output. Especially large changes show up in the November 1991 benchmark vintage (reflecting the base-year shift of December 1985) and the November 1998 benchmark vintage (reflecting the move to chain weighting). Shifts similar to those of real output, but in the opposite direction, show up in the data on the price level.[6] Finally, changes across benchmark vintages in growth rates for real consumption are usually in the same direction as changes in real output growth rates but of smaller magnitude.

To investigate these issues further, we examine plots (Figures III.1 to III.4) of the same data, where we show differences between the log levels of the variables, with the mean difference subtracted (since it reflects mainly base-year changes). Define the variable X(t,s) as the level of the data for time t in vintage s. The plots show, for each date t, the log [X(t,a)/X(t,b)] - m, where m is the mean of log[X(τ,a)/X(τ,b)] over the largest sample of τ contained in both vintages, and where b is a later vintage than a.[7]

In the figures, each column of plots represents a particular benchmark vintage, where in each row the data from that vintage are subtracted from all subsequent benchmark vintages. The labels on each plot follow the structure Lz#, where L means the logarithm of the variable, z represents the variable (z=N for nominal output, z=Y for real output, z=P for the price level, z=C

[6] Note that the price level in the November 1998 vintage is the chain-weighted price index; in earlier vintages, it's the deflator. But the differences between the two are trivial.

[7] Since we've removed the mean, we won't capture any mean shifts in variables, but those are illustrated in Table III.1.

for real consumption), and where # represents the benchmark vintage, with #=1 for the November 1975 vintage, #=2 for 1980, #=3 for 1985, #=4 for 1991, #=5 for 1995, and #=6 for 1998. Reading along the main diagonal of the plots reflects a comparison of subsequent benchmark vintages; the plots below the main diagonal reflect comparisons across more than one benchmark vintage. 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, 1995Q3 in column 5, and 1998Q3 in column 6. The vertical axis in each plot runs from -0.08 to +0.08; these are demeaned log differences.

There are three major features to note about the plots: (1) trends; (2) spikes; and (3) persistent deviations from a linear trend. First, the dominant feature of the plots is the presence of trends. A downward tilt means that 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. Unit root tests find some of the plots exhibiting stationarity, while others do not. Taken together, the plots point to cross-vintage differences at many frequencies, an observation we plan to explore in the frequency domain in future work.

In Figure III.1, the most striking result is the downward spike in all the plots in the first column. This arose because the original estimate of nominal GNP in late 1974 through the third quarter of 1975 was too low. Data used in the comprehensive revision of January 1976 raised nominal GNP substantially, especially in 1975Q3. But real GNP for that date was also increased substantially in the annual revisions that occurred in July 1976 and July 1977. So the spike is attributable to a series of new source data over time that made a substantial difference in the level of nominal GNP over the course of several quarters.

In Figure III.2, the effects of substitution bias are apparent. You'll note that the real output series, especially moving from vintage 3 to vintage 4, is tilted upward. This arises because the fixed-weighted method using the 1982 base year greatly changes the relative pricing relationships between energy and other goods. Thus, even data from long before were affected in a strong way, leading to a tilt in the plot. But note that when we move from vintage 5 to vintage 6, chain weighting reverses that effect. Notice also that the movement from GNP to GDP (from vintage 4 to vintage 5) didn't cause much effect in real output.

Figure III.3 shows that the price level is affected quite a bit by vintage changes. As with real output, note the substantial tilt between vintages 3 and 4. The downward tilt shows the large change in relative prices over time reflected in the price index. This tilt was reversed when we moved to chain-weighting, as the lower right-hand plot between vintages 5 and 6 shows. Note that the net effect on long-ago data, shown in the lower left-hand plot between vintages 1 and 6, is relatively small.

Figure III.4 shows that real consumption doesn't mirror real output terribly closely, so it has its own unique differences across vintages. There are substantial tilts in the plots, but many of them reverse direction, which means there's more going on than substitution bias, as was the case for real output. The move to chain weighting, shown in the lower right-hand plot, shows up as a reversal of the tilt in earlier plots. In the plot showing the differences between vintages 2 and 3, we see that there was very little difference at all across the vintages for data between 1947 and 1968.

All these differences across vintages point to the fact that the data are revised substantially. If we look at quarterly log differences from one quarter to the next in the variables (Table III.2), we find that while most of the correlations across these vintages are above 0.9, the correlations aren't as high as one might expect, given that these are different measurements of data over the same period. Thus, growth rates from quarter to quarter can change substantially; they may even be large from one year to the next. To sharpen our focus on these issues, we now take a particular variable, real consumption, and run some additional tests to illustrate how much vintage matters for growth rates.

IV. PROPERTIES OF REAL CONSUMPTION DATA ACROSS VINTAGES

The examples given in the introduction were illustrative of the types of issues for which having a real-time data set may be important. But how much does it really matter? Are the differences between the real-time data and the final revised data trivial? Or do they matter economically?

To further investigate the degree to which having a real-time data set matters, we begin by looking at real consumption spending from the national income accounts.[8] We select three data sets, dated February 1986, November 1993, and February 1998, and plot the data on real consumption growth (quarterly, at annual rates) from 1947Q2 to 1985Q4 (Figure IV.1). There are substantial differences between the growth rates, especially in the 1950s. One important difference between the vintages is that the 1986 and 1993 vintage data sets use a fixed base year to calculate real consumption spending, whereas the 1998 vintage data set uses chain weighting. To demonstrate this more clearly, we plot the differences between the growth rates across each pair of vintages of the data (Figure IV.2). You can see that in some quarters the growth rates of consumption change nearly 5 percentage points, and differences of more than 2 percentage points are not uncommon. Moreover, in many instances, significant differences of the same sign persist for more than a quarter, and the variance of the differences in the growth rates appears to change over time. Notice, though, that there's not as much difference between the February 1986 and November 1993 vintages as there is between either of those and the February 1998 chain-weighted vintage.

---

[8] We don't examine real output, as Runkle (1998) did, because the switch from GNP to GDP in 1992 led to systematic differences, which may affect some of the tests we perform later.

A slightly different way of looking at the data is to compare how the data change from when they are first released to later versions. Because we collect the data in mid-quarter, the first time an additional observation appears in the data set for a particular quarter, it is the version of the data known as the “advance” release. We can track the value of the observation from its advance value to its latest (most recent) value. One reason for doing this is to see the extent to which the revisions are characterized as containing news or reducing noise, as suggested by Mankiw and Shapiro (1986). The idea is that if the revisions are characterized as containing news, subsequent releases of the data for that date contain new information that was not available in the earlier releases. As a result, the advance release is an efficient estimate of later data. This implies that the revision to the data is correlated with the revised data but not the earlier data. It also implies that the variance of the data should increase as we look at later and later vintages, since an optimal forecast is smoother than the data. On the other hand, if data are characterized as reducing noise, then subsequent releases of the data just eliminate noise in the earlier release, so the earlier release is the true value plus measurement error that gets reduced over time. In this case, the revision should be uncorrelated with the revised data, but correlated with the advance data. In addition, the variance of the data should decline as it is further revised. In running tests for news and noise, Mankiw and Shapiro found that the revisions to real GNP data from 1976 to 1982 were best characterized as containing news, not reducing noise.

To formalize this, we use the following notation. Let X(t, s) represent the data for date t as of vintage s. Then a revision of the data from vintage i to vintage j (where j > i) is $e(t, i, j) \equiv X(t, j) - X(t, i)$. For example, e(93Q4, Feb. ‘94, Feb. ‘95) ≡ X(93Q4, Feb. ‘95) - X(93Q4, Feb. ‘94). To say that a revision is characterized as containing news means that the revision is uncorrelated (orthogonal) to earlier vintage data, so that $e(t, i, j) \perp X(t, i)$. To say that a revision is characterized as reducing noise means that the revision is uncorrelated with later vintage data, so that $e(t, i, j) \perp X(t, j)$.

We begin by looking at figures that show four different data sets, each consisting of the four-quarter moving-average growth rate of real consumption (had we looked at quarter-to-quarter growth rates, graphical analysis would have been impossible because there's too much quarterly variation in the data). One data set (labeled initial) consists of the growth rate each quarter as shown in the advance release made available one month after the end of a quarter, which is X(t, t+1), where t+1 refers to the vintage 1 quarter after date t. The second (labeled 1-year-later estimate) consists of the growth rate for a quarter based on a data set with a vintage one year after the initial vintage or five quarters after date t, X(t, t+5); the third (3-year-later estimate) is based on a vintage three years after the initial vintage or 13 quarters after date t, X(t, t+13). The fourth data set (latest) consists of the November 1998 vintage of data, X(t, Nov. 1998). A time-series plot of the real consumption growth rates from these four different data sets shows that although the qualitative movements of the different series are similar, growth rates across the series can vary by significant amounts--as much as two percentage points (Figure IV.3). That's a big difference, because this is not just quarter-to-quarter variation, but a difference in the four-quarter average growth rate.

It's also instructive to examine the revisions to the data from the initial release to 1 year later, from 1 year to 3 years later, and from 3 years later to the latest data (Figure IV.4). Revisions to the four-quarter growth rates are often quite large from one of our data sets to the next, with many revisions exceeding 1 percentage point. The standard deviation of all the revisions is in the neighborhood of one-half of a percentage point. In going from the initial release to the final data, the revisions to the annual growth rates are even larger, with a standard deviation of 0.8 percentage point (Figure IV.5).

Are the revisions to real consumption data best characterized as containing news or reducing noise? To find out, we use a quarterly version (not the four-quarter moving average, but just quarter-to-quarter growth rates) of our four data sets to run tests like those of Mankiw and Shapiro. First, we examine the standard deviation of the real consumption growth rates from the four different data sets in Table IV.1. If the revisions contain news, these should increase from initial, to 1-year, to 3-year, to latest data sets; if the revisions reduce noise, the standard deviations should decline as we move down the rows from initial to latest. As the table shows, the standard deviation rises from initial to 1 year, then falls in each successive series. So, the initial to 1-year revision contains news, while the 1-year to 3-year and 3-year to latest revisions reduce noise.

Next, we examine the correlation between the revisions and the growth rates (Table IV.2). Consistent with the earlier result, only the initial to 1-year revision can be characterized as containing news because it is correlated with later data and uncorrelated with earlier data. The other five revisions can be characterized as reducing noise because they are correlated with some earlier data and uncorrelated with later data. Overall, one could argue that revisions to the initial consumption data contain news and that subsequent revisions simply reduce noise.

The results shown in Figures IV.4 and IV.5 and Tables IV.1 and IV.2 suggest that revisions to the data can be substantial, so they could potentially influence the outcomes of research studies. The extent to which they do so is our next subject.

V. DOES VINTAGE MATTER FOR KEY MACROECONOMIC RESULTS?

It's clear that the vintage of the data makes a difference for quarterly and annual growth rates, but does it matter for empirical work? After all, the data shown in Figure IV.2 suggest that the differences in growth rates between the vintages may well be totally random. Thus, the choice of vintage may have little significant effect on tests of economic hypotheses. To see whether that's true, we now take a number of empirical exercises from the economic literature, rerun them with differing vintages of data, and see how much the vintage matters. We examine empirical work by Kydland and Prescott (1990), Hall (1978), Beveridge and Nelson (1981), and Blanchard and Quah (1989).

Kydland and Prescott (1990)

Kydland and Prescott examine the correlation of real GNP with lags and leads of itself and other variables. They filter the data with an HP filter, then calculate the cross correlations. They use data from a 1990 vintage; we compare our results for data vintages from February 1990, February 1994, and February 1998 to their results (Table V.1). As the table shows, although there are some quantitative differences, the qualitative pattern is quite similar across all the vintages. A plot of the HP-filtered cyclical data from the three vintages shows little difference across vintages (Figure V.1). The biggest differences across vintages are on the order of one percentage point and occur only in the 1950s (Figure V.2). Trend real output growth also behaves similarly across vintages, though the four-quarter average of trend output growth can differ as much as 0.5 percentage point at times (Figure V.3). Part of the differences across vintages for real output could be attributable to the switch between GNP and GDP that occurred between the 1990 and 1994 vintages. So it's useful to also examine other variables, for which the revision pattern may be different. Figure V.4 shows results for real consumption, showing much smaller revisions between the 1990 and 1994 vintages. Altogether, however, since the purpose of Kydland and Prescott's research was to establish general business-cycle facts, it's hard to conclude that the data vintage matters.

Hall (1978)

Hall found evidence supporting the life-cycle/permanent-income hypothesis using data on U.S. consumption spending. Although Hall's results have been challenged and modified in a variety of ways, in such papers as those by Flavin (1981) and Deaton (1987), an even more fundamental question is: are Hall's empirical results robust to different data sets? That is, would we get significantly different outcomes depending on what vintage of data we used?

Hall's original data set included observations on consumption from 1948Q1 to 1977Q1; so we assume that he had data of vintage 1977Q2. Hall begins by testing to see if consumption can be predicted from its own past values. Under the pure life-cycle/permanent-income hypothesis, only the first lagged value of consumption should help predict current consumption. Hall regresses consumption on four lags of consumption, testing to see if the last three lags are jointly zero.[9] His original result is shown in the first line of Table V.2. In the table, the coefficient estimates are given, with standard errors in parentheses. The column labeled s shows the standard error of estimate, DW is the Durbin-Watson statistic, and F is the value of the F-statistic testing the hypothesis that the coefficients on the second, third, and fourth lags of consumption are jointly zero, with the p-value for the test shown in parentheses. The F test shows that you can't reject the hypothesis at the 5 percent level.

[9] The variable used is real consumption of nondurables and services divided by the population.

Using our real-time data set with consumption data from the vintage of the second quarter of 1977, we are able to replicate Hall's results fairly closely, as the second line of the table shows. The only major difference is the sign of the constant term. Our replication confirms Hall's finding that the coefficients on the second, third, and fourth lagged terms are jointly zero.

However, when we rerun the test on the same sample period (1948Q1 to 1977Q1) using vintage data from the first quarter of 1998, the coefficients change dramatically, and the F-test now rejects the hypothesis that the second-through-fourth lagged consumption terms are jointly zero. The p-value for the test is only .02, so we reject the hypothesis at the 5 percent level.

Further, when we update the sample to include data through 1997, we reject the hypothesis even more convincingly. Again, the coefficient estimates change dramatically, and the F-statistic rises to 8.1, with a p-value of less than 0.005.

These results mean that Hall's original hypothesis, that only the first lag of consumption matters in determining contemporaneous consumption, is not well supported by the data. Hall's test was legitimate, but his empirical result does not stand the test of time, either in terms of revisions to the data, or in terms of additional data.

Beveridge and Nelson (1981)

In their classic 1981 paper, Beveridge and Nelson introduced a procedure for decomposing a time series into permanent and transitory components, in which both components were stochastic. The methodology depends only on past data, but revisions to the data could well make the vintage of the data matter. The question we pose is: does a change in the vintage of the data set make a significant difference to how a time series is decomposed?

We apply the Beveridge-Nelson procedure to data on real output and compare the results across vintages. We begin by assuming that their data, which included GNP data through 1977Q1, were the data available in May 1977. We run their procedure first on the May 1977 data set, again on the data set of May 1987, and again on the data set from August 1997, to see how vintages of data a decade apart, but covering the same period (1947Q2 to 1977Q1), are decomposed. The original Beveridge-Nelson paper includes a decomposition of real GNP but doesn't indicate the time-series process used. Based on our implementation of Box-Jenkins methods, and comparing our results to those of Beveridge and Nelson, we think they used an ARIMA(1,1,2) process for real GNP, so we use that as well.[10]

The results show that the transitory components (Figure V.5) are not affected very much by the vintage of the data set. As the figure shows, the lower frequency movements of the transitory components are similar in all three vintages of the data. There are a few periods in which the transitory component differs in magnitude, such as in 1950, 1957, and 1968. But, overall, the vintage of the data set doesn't matter very much, at least at lower frequencies.

Repeating this exercise for other variables, such as real consumption shown in Figure V.6, shows similar patterns to that of real output, with the main differences across vintages coming when the data spike up or down.

Blanchard and Quah (1989)

---

[10] Similarly, an ARIMA(1,1,2) process is used by Blanchard and Fischer (1989), page 16, in their general characterization of the business-cycle facts.

Blanchard and Quah use a structural VAR in output and unemployment to define supply disturbances as shocks that have a permanent effect on output and demand disturbances as shocks that have a temporary effect on output. They examine U.S. data from 1950 to 1987, calculating impulse responses and variance decompositions based on a VAR model in output and unemployment. We examine how changes in the vintage of the data affect the decomposition of shocks into supply disturbances versus demand disturbances, how the impulse responses change across data vintages, and how the cumulative effects of demand and supply shocks vary with the data vintage.

We compare Blanchard and Quah's results to ours using the February 1988 version of our data set, then comparing those results in turn to our November 1993 data set and our February 1998 data set. First, using our February 1988 data set, we are able to replicate the results of Blanchard and Quah fairly precisely. When we plot the impulse responses to supply and demand shocks (Figures V.7 and V.8), the plots are quite similar to Figures 1 and 2 in Blanchard and Quah, both qualitatively and quantitatively.[11]

When we look at the decomposition of shocks into demand and supply shocks for the three different vintages of the data (Figure V.9), we notice there are substantial differences across data vintages. The differences are particularly noticeable for demand shocks, as many of the local peaks and troughs are largest in magnitude using the '88 vintage data and smallest in magnitude when using the '98 vintage data. However, demand shocks are temporary, so these

---

[11]To measure the unemployment rate, Blanchard and Quah use the seasonally adjusted rate for males, age 20 and over. Because this rate does not appear in our data set, we substitute the total civilian rate of unemployment for the Blanchard/Quah measure. On the basis of our replication using the February '88 vintage, this substitution has little effect on the results.

differences in magnitude don't seem to matter as much when we look at the cumulative effect of the shocks (Figure V.10). As this figure shows, even the fairly small differences across vintages in the measured supply shocks make a large difference on the cumulative effect of supply shocks on output and unemployment. The impact of demand shocks on output isn't strongly affected by the vintage of the data. But the data vintage matters a lot for the impact of demand shocks on unemployment, especially at turning points.

The other way in which the method of Blanchard and Quah is often used is to establish stylized facts about how economic variables respond to shocks. These are generally shown in figures that illustrate the impulse responses to a shock. Using the Blanchard and Quah method, and the same three vintages of data used above, we calculate the impulse responses for demand and supply shocks (Figure V.11). Note that the impulse responses are very sensitive to vintage, especially for demand shocks. The response of output or unemployment to a demand shock is sometimes as much as five times as large, using 1998 data, than when using 1988 data. So the vintage of the data set seems to matter quite significantly for impulse responses. Why this is so is difficult to determine, but the estimated variance-covariance matrix shows a much different variance of the structural shocks, along with a substantially different parameter estimate of the coefficient on output in the unemployment equation. This occurs despite the fact that differences in the data don't seem large (Figures V.12 and V.13). This suggests that there may be something about the procedure for estimating a structural VAR that makes it very sensitive to small changes in the data.[12]

[12] In preliminary work on Cochrane's (1994) output-consumption structural VEC, we find similar large differences in the cross-vintage impulse responses.

Can we be more precise? As noted above, in examining the estimated coefficients of the structural VAR representation, we notice particularly large differences in the estimated coefficient on contemporaneous output growth in the structural unemployment equation as we move from vintages February '88 and November '93 to February '98. The coefficient estimate is 4.62 in the February '88 data, 2.45 in the November '93 data, and 0.63 in the February '98 data, with output growth measured in log first differences and the unemployment rate expressed as a percent, rather than in percentage points.

In a recent paper, Sarte (1997) shows that standard structural VAR instrumental variables (IV) techniques—which use structural shock estimates as instruments—can fail over certain ranges of the parameter space, which depends, in part, on the assumed identification scheme. The key condition for such a failure is a low pairwise correlation between the instrument/structural shock and the variable instrumented. In estimating the model, we employ the standard IV approach and use the estimated structural shock attached to the output equation as an instrument for contemporaneous output growth in the unemployment equation. We then checked Sarte's key condition for IV failure by computing for each vintage the correlation coefficient between the output-equation structural shock and output growth. For vintages February '88 and November '93, those correlations border on zero: 0.04 and 0.08, respectively. Such low correlations call into question the usefulness of structural shocks as instruments and, by implication, the just-identified structural VAR methodology. Indeed, a reasonable conclusion is that the SVAR is unidentified empirically in the first two vintages. In contrast, the pairwise correlation in the February '98 data rises significantly, to 0.23, suggesting a higher possibility that the model is identified empirically.

We view these results as an extension of Sarte's. Sarte showed that alternative identification schemes, holding constant the data vintage, may fail empirically. Our results indicate that a given identification scheme may fail empirically in some vintages but not in others. On the basis of these results, we recommend that structural VAR users check their results for robustness along the lines suggested by Sarte and across different vintages of data.

VI. HOW VINTAGE MATTERS FOR POLICY ANALYSIS

The vintage of the data being used can be quite important for analyzing the response of policymakers to economic events. On many occasions, the preliminary data turn out to be misleading indicators.

As an example, consider the monetary-policy decision made in early October 1992. The transcripts of the Federal Open Market Committee meeting show that policymakers were very concerned about economic weakness.[13] But someone looking at the statistics on GDP or industrial production today might think that it was obvious that the economy was recovering nicely. Real GDP growth in the first three quarters of 1992 was 4.7 percent, 2.5 percent, and 3.0 percent (at annual rates), while industrial production rose 1.5 percent, 6.2 percent, and 1.8 percent (at annual rates). So why were policymakers concerned? At the time of the FOMC meeting, the statistics available showed GDP growth of 2.9 percent in the first quarter (versus 4.7 percent in today's statistics), and 1.5 percent in the second quarter (vs. 2.5 percent); the third quarter number hadn't been released yet. The statistics available at the time showed that

[13] The committee voted not to ease policy at that meeting, but there were a large number (four) of dissents. Many members of the committee felt that easing would be necessary soon.

industrial production had declined 3.3 percent in the first quarter (versus an increase of 1.5 percent in today's statistics), risen 5.3 percent in the second quarter (versus 6.2 percent), and had declined at a 1.1 percent annual rate from May to August (versus 0.0 percent growth in today's statistics). Although GDP in today's statistics rose at a 3.0 percent annual rate in the third quarter of 1992, the statistics available in early October suggested that GDP would grow only about 1.6 percent (Federal Reserve Board staff forecast), and many monthly indicators--the leading and coincident indexes, industrial production, capacity utilization, employment and hours worked, retail sales, auto sales, consumer credit, consumer confidence, residential construction, new home sales, building permits, durable goods orders, factory orders, the purchasing managers' index, and nonresidential construction--showed declines. Most of these indicators were later revised up substantially. So an economist looking at today's data would be surprised that the FOMC was worried at all; but at the time, fears that the economy was slowing down or even slipping into another recession were clearly justified.

To illustrate how the vintage of data can be useful in analyzing monetary policy, we're going to examine the now-popular rule for setting monetary policy that was suggested by John Taylor (1993). The Taylor rule suggests that monetary policy should be set according to the equation:

$$i_t = r_{t-1}^* + \pi_{t-1} + \tfrac{1}{2}\, ygap_{t-1} + \tfrac{1}{2}\, \pi gap_{t-1}, \qquad (1)$$

where i_t is the nominal federal funds rate set at date t, r_{t-1}^* is the equilibrium real federal funds rate estimated with data through time t-1, π_{t-1} is the average inflation rate (using the GDP deflator) over the past year to time t-1, $ygap_{t-1}$ is the percentage output gap, which is the level of real output (y) minus the level of potential output (y^*) divided by the level of potential output at time t-1, and πgap_{t-1} is the amount by which inflation (π_{t-1}) over the past year differs from its target (π_{t-1}*).

Recently, Orphanides (1997) illustrated how important the data vintage is for analyzing Taylor's rule, showing that the rule implies a quite different time path for the federal funds rate when based on real-time data as opposed to final revised data. However, Orphanides proceeds by assuming that the Fed's inflation target is set at 2 percent throughout the sample period. This isn't a bad hypothesis for the period that he analyzed, 1987 to 1992. But if we want to think about Taylor's rule over longer periods, it is useful to allow for different inflation targets over time. For example, in the late 1970s, the Fed's inflation target must certainly have been significantly higher than 2 percent. If we assumed a 2 percent inflation target and applied Taylor's rule to previous periods, we'd find that it fit badly.

To get a perspective on how Taylor's rule would behave in different periods, and to see if using real-time data would yield different results than using today's data in analyzing Taylor's rule, we perform the following experiment. For the period over which we have vintages of real-time data (since the third quarter of 1965), we run the following experiment. First, we operationalize Taylor's rule by making some assumption about how people might have estimated the real (equilibrium) federal funds rate and the output gap. Second, we ask the question: suppose the Fed had been following Taylor's rule since 1965; what was the implicit inflation target it must have had at the time? That is, we take equation (1) and solve it for the inflation target, taking as given the federal funds rate over time, our estimates of the output gap, and the inflation rate. Third, we do this experiment first using current (vintage November 1998) data, and using real-time data--the data the Fed would have had at the time. If the final data are not significantly different from the real-time data, the implicit inflation targets will be fairly close. But if the real-time data give a significantly different picture of the state of the economy, the implicit inflation targets will be quite far apart.

Inverting Taylor's rule from equation (1) yields an expression for the implicit inflation target:

$$\pi_{t-1}^* = 2[r_{t-1}^* - (i_t - \pi_{t-1})] + \pi_{t-1} + ygap_{t-1}. \tag{2}$$

As Taylor did in his original article describing the rule, we use the GNP/GDP deflator as our measure of inflation. To calculate the output gap, we face the difficulty that, prior to the mid-1970s, an economist was likely to have used a trend line for estimating potential GDP growth, but after the oil-price shocks of the 1970s, the economist would likely have taken into account the significant differences in productivity growth and output growth that became apparent in the data. To allow for this without taking a strong stand about how economists changed their views about potential GDP, we use an HP filter to determine potential GDP.[14] Finally, to estimate the equilibrium real federal funds rate, we proceed as Taylor suggested--we simply take the average real federal funds rate since data on the federal funds rate began to be collected in 1954.

We begin by looking at the estimates of the inflation target using the November 1998 vintage data (Figure VI.1), then run the exercise again using real-time data (Figure VI.2). The results are a bit surprising: if the Fed were following the Taylor rule for the last 33 years, it would have frequently been trying to drive inflation below zero. Another way of saying this is: if Taylor's rule really represented a reaction function for the Fed in response to economic events,

---

[14] Though the use of the HP filter may be questioned, other methods are likely to provide similar results. The main issue is how different views of the data might have led to different policy actions, not how potential GDP is determined.

the Fed acted in the 1980s and part of the 1990s as if it were trying to drive the economy into deflation. This may simply mean that the Taylor rule is a poor description of monetary policy or that our measure of the inflation target is missing something important.

It is interesting to note, however, that the target for the inflation rate is above actual inflation in periods in which economists generally think Federal Reserve policy was easy (middle 1970s, early 1990s), and the target for the inflation rate is below actual inflation in periods in which Federal Reserve policy was tight (from 1979 through the early 1980s, the mid-1990s). Some periods don't match up as well, however (mid-1980s, late 1960s). To investigate this in more detail, we compare the difference between inflation and its target to the index of monetary policy created by Boschen and Mills (1995) (Figure VI.3). We see that periods when the inflation rate is above the target ($\pi - \pi^*$ is positive) generally correspond to periods when the Boschen-Mills index is negative, indicating tight policy. The inflation rate is usually below the target ($\pi - \pi^*$ is negative) when the Boschen-Mills index is positive, indicating loose policy. The two indicators have a clear negative correlation; the correlation between them is -0.46.

The key question we're interested in, however, is the difference between the time path of the inflation target that comes from real-time data compared to the time path of the inflation target that comes from today's data--that difference is illustrated in Figure VI.4. Though the overall movements in the two target inflation rates are roughly the same, they differ substantially from time to time. In the first quarter of 1972, for example, the real-time target is 5.5 percent while the target based on today's data is 9.3 percent. As recently as the first quarter of 1994, the difference is significant: the target based on real-time data is 5.1 percent, while the target based on today's data is 6.4 percent. For the entire sample period, the real-time target averages 1.7 percent, while the target based on today's data is 2.6 percent.

Many factors contribute to the differences between the inflation target series. The main difference is attributable to revisions to the price deflator, although revisions to real output data also contribute small variations in the calculated output gaps. Changes in the price deflator are key for two reasons: (1) they affect the inflation term in the rule; and (2) they affect the calculation of the average real federal funds rate. The first effect is usually fairly small, though the difference averaged 0.2 percentage point over the sample period. But the second effect, the impact on the calculation of the average real federal funds rate, makes a large difference in some periods (Figure VI.5). For example, in December 1985, the BEA rebenchmarked the national income accounts, changing the base year from 1972 to 1982. Because this was in the days before chain-weighting, rebenchmarking changed the entire past history of inflation. The average inflation rate from 1972 to 1984 rose from 6.9 percent to 7.3 percent, purely as a result of the change in the base year and the reweighting of the components. The effects were even more dramatic for the period from 1972 to 1977, in which the average inflation rate rose from 7.0 percent to 7.7 percent. The results are apparent in the sudden drop in the estimate of the real federal funds rate in the first quarter of 1986.

Finally, we consider one last conceptual experiment to illustrate the differences between the real-time data set and today's data. Suppose the Fed had, in fact, been using Taylor's rule from 1965 to 1998 and that its inflation target over time was exactly the target shown in Figure VI.2. Now suppose an econometrician in November 1998 were trying to replicate Taylor's rule, using current vintage data instead of real-time data, the same technique the Fed used, and the

Fed's announced inflation targets. What nominal federal funds interest rate series would the 1998 econometrician calculate as what the Fed should have done? The results are shown in Figure VI.6. The federal funds rate path calculated by the econometrician in 1998 is considerably different from what the Fed actually did, with the differences (Figure VI.7) entirely attributable to changes in the data.

In sum, the vintage of data being used makes a tremendous difference in analyzing policy. Economists analyzing policy need to realize that policymakers had different data sets that gave a different picture of the state of the economy. Basing estimates of policy reaction functions on the latest version of the data is likely to be misleading.[15]

VII. HOW VINTAGE MATTERS FOR FORECASTING

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, with a univariate Bayesian model and with a multivariate quarterly Bayesian vector error-correction (QBVEC) model, and compare the results based on using real-time data to those based on current-vintage data. We have complete data on all variables used in the three models in real-time data sets with vintages beginning in February 1975 and data in each vintage data set going back to 1959, the

---

[15] McNees (1992) develops a Fed reaction function based on real-time data. For an interesting analysis of how policymakers should respond to current data that they know are going to be revised, see Ghysels et al. (1998). For an exercise in measuring shocks to monetary policy when data are revised, see Croushore and Evans (1999).

limiting variable being M2 in the QBVEC model. We proceed in the following manner: (1) estimate a model for real output growth using data from the first quarter of 1959 through the fourth quarter of 1974 that was known in February 1975; (2) forecast quarter-over-quarter real output growth over the next five quarters (from the first quarter of 1975 to the first quarter of 1976), 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; and (4) calculate the forecast errors based on the four-quarter-average forecasts. We follow this procedure once using the real-time data set (for which data revisions are possible as we roll forward each quarter), and a second time using today's data (vintage November 1998, which contains no data revisions as we roll forward each quarter).

When we run this exercise first with an AR(4) model on real output, we find that the two forecasts look somewhat different over time, but not dramatically so (Figure VII.1). There's certainly a lot more variation in the actual data than there is across forecasts, as can be seen in Figure VII.2. A scatter plot of the two forecast series shows a positive relationship between the two sets of forecasts, but there are systematic differences between the forecasts (Figure VII.3). Evidently, the vintage of the data matters even for such simple forecasts as these AR(4) forecasts. Taking the November 1998 vintage data set as representing the actual value for the data, we show in the first two rows of Table VII.1 that the root-mean-square-forecast error is not very different when forecasts are based on real-time data as opposed to final revised data. That's actually quite surprising because it says that having today's data set gives no better forecast performance than having available just real-time data; or it may simply mean that forecasting a variable such as real output growth using time-series methods isn't a very productive enterprise; all forecasts are pretty much returns to trend.

An alternative forecasting method, which helps to filter shocks (or revisions) to the data, is to use Bayesian methods for estimation. Following Litterman (1986), we forecast output growth with a Bayesian AR(4) process, with standard Minnesota priors.[16] This treatment means our prior is that real output is represented by a random walk with drift or that output growth is a constant plus white noise, but the priors are not tight. With very loose priors, we'd have the AR(4) model discussed above. If we made the priors very tight, we'd impose the random walk exactly. Instead, we choose a relatively loose value of the tightness parameter of 0.2, which is the standard deviation around our prior that the coefficient on the first lag of output growth is zero. Standard deviations on our priors for the remaining coefficients fall with the lag.

Following the same procedure as in the non-Bayesian AR(4) exercise above, we obtain somewhat different results. Figure VII.4 shows some substantial differences between the forecasts made with real-time data and those made with today's data. Those differences in forecasts relate more to the estimate of the average growth rate of real output (in levels of real output, the drift), as Figure VII.5 shows. The scatterplot in Figure VII.6 shows that the forecasts are more tightly clustered than they were in the non-Bayesian AR(4) case, but the majority are above the 45-degree line, showing persistent differences in the forecasts. Despite that, the real-time forecasts again have about the same root-mean-squared error as those made with today's

[16] Unlike Litterman, we estimate the model in first differences, thereby imposing with certainty a unit root on the process.

data, even though today's data are used as actuals in calculating the forecast error, as can be seen in Table VII.1.

Our third forecasting exercise uses the quarterly Bayesian vector error corrections model of Stark (1998). This model was designed for forecasting and for evaluating monetary policy issues at the Philadelphia Fed. It applies Litterman's techniques to a VAR composed of real output, the GDP price index, the federal funds rate, real import prices, the unemployment rate, real M2, and the interest rate on 10-year Treasury bonds. The multivariate aspect of this model, compared to the earlier models, leads to substantially reduced forecast errors. The model differs from Litterman's approach in two major ways: (1) it imposes unit roots in the model, which others have found improves the forecasting ability of this class of model; and (2) it adds an error-correction term consisting of the spread between the federal funds rate and the long-bond rate, with a diffuse prior. Thus, any long-run forecast obeys a cointegrating relationship between short-term and long-term interest rates. In this model, we can examine forecasts of variables other than real output.

It's instructive to first examine a couple of episodes in which the real-time data differ substantially from today's data. Figure VII.7 shows history and forecasts from data vintage August 1975. Solid lines to the left of the break in each line represent history as seen in real time, while the dashed lines to the left of the break are what is seen in today's data. To the right of the break in each line are the forecasts. Note that real output growth looks substantially different today than it did in real time (the numbers on the vertical axis are log changes in the level of real output from one quarter to the next, so take the differences between the two lines and multiply by 400 to get the difference in growth rates). As a result, the forecast for real output growth in the future is substantially higher in real time, since the model predicts a greater rebound from a deeper recession. Note also the substantial differences in the forecast for inflation. Similarly, in the third quarter of 1992 (Figure VII.8), output appeared to be growing much more slowly than it does in today's data, resulting in substantial differences in the forecasts.

Looking at figures like those we used for the AR(4) model and BAR(4) model, we see that the forecasts for real output aren't too different (Figure VII.9) except in certain periods. The forecasts are substantially more variable (Figure VII.10) than was the case with the other forecast models. The real-time forecast lines up more closely (Figure VII.11) with the forecast using today's data than was the case for the AR(4), but the differences are larger than was the case with the BAR(4). We can also see that although the differences between real-time forecasting and forecasting with today's data aren't very large for real output growth, they are substantially larger for other variables, such as inflation, as can be seen in Figures VII.12 to VII.14. As with the other variables, there isn't much difference in the root-mean-square errors between using real-time or latest data. But note that the root-mean-square errors are lower for the QBVEC compared to univariate methods (Table VII.1).

VIII. CONCLUSIONS

This paper reports on how differing vintages of data can lead to somewhat different results for major studies in macroeconomics. It is somewhat reassuring that for many of the studies we examine, the results are generally robust, at least qualitatively, for different vintages of the data. But in some cases, the empirical results are quite sensitive to the exact vintage of the data.

What can we conclude from these results? In practice, economists run thousands of empirical exercises each day, some of which get reported in academic journals and influence economists' thoughts about the structure of the economy. Our exercise is really one in the spirit of checking such results for robustness and can thus be used to confirm some results in the literature, such as those of Kydland and Prescott. But when empirical results are sensitive to the vintage of the data, economists should be more cautious about accepting those particular results or perhaps about accepting the empirical methods that led to those results. If an empirical method is robust to data vintage, as in the case of Kydland and Prescott, an empirical researcher can have more confidence that the method itself is sound and not overly sensitive to minor variations in the data. But if the empirical method is one that leads to very different results for minor variations in the data, a researcher should be skeptical. Or, certainly, further research is needed to establish the validity of the research method.

In addition, we found that the data vintage matters significantly for both policy analysis and forecasting. Using a recent vintage of data to analyze what policymakers should have done in the past is problematic because it suggests that policymakers should have foreseen how data would be revised in the future. In examining monetary policy rules, for example, researchers need to base their analysis on the real-time data that were available to policy-makers, if they want to understand their actions. The differences in policy implications between using real-time data and current-vintage data can be substantial.

The same is true in forecasting. Forecasts based on real-time data are certainly correlated with forecasts based on final data, but data revisions to real output are so large that they may cause forecasts based on current-vintage data to be considerably different from forecasts based on real-time data. This sounds a cautionary note for studies claiming that some new, improved forecasting method beats other methods, if the study presents only evidence based on current-vintage 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 macroeconomic researchers, economists engaged in policy analysis, and forecasters.

Table III.1
Average Growth Rates Over Five Years
For Benchmark Vintages
Annualized percentage points

Vintage Year:	'75	'80	'85	'91	'95	'98
Period						
			Real Output			
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
			Real Consumption			
49Q4 to 54Q4	3.6	3.3	3.3	3.7	3.9	3.8
54Q4 to 59Q4	3.4	3.3	3.3	3.3	3.4	3.5
59Q4 to 64Q4	4.1	3.8	3.8	3.7	3.8	4.0
64Q4 to 69Q4	4.5	4.3	4.4	4.4	4.5	4.8
69Q4 to 74Q4	2.3	2.6	2.6	2.5	2.6	2.8
74Q4 to 79Q4	NA	4.4	4.4	3.9	3.9	4.1
79Q4 to 84Q4	NA	NA	2.8	2.5	2.5	2.6
84Q4 to 89Q4	NA	NA	NA	3.2	3.1	3.4
89Q4 to 94Q4	NA	NA	NA	NA	2.3	2.1
			Prices			
49Q4 to 54Q4	2.6	2.7	2.7	2.5	2.4	2.6
54Q4 to 59Q4	2.6	2.6	2.6	2.9	2.9	2.4
59Q4 to 64Q4	1.4	1.5	1.5	1.6	1.6	1.3
64Q4 to 69Q4	3.6	3.9	3.9	4.1	4.1	3.7
69Q4 to 74Q4	6.3	6.5	6.2	6.8	6.5	6.3
74Q4 to 79Q4	NA	7.1	7.0	7.5	7.7	7.2
79Q4 to 84Q4	NA	NA	6.1	6.1	6.4	6.2
84Q4 to 89Q4	NA	NA	NA	3.3	3.6	3.4
89Q4 to 94Q4	NA	NA	NA	NA	2.9	3.1

			Nominal Output			
49Q4 to 54Q4	7.9	7.9	7.9	8.1	8.0	8.0
54Q4 to 59Q4	5.6	5.6	5.7	5.7	5.7	5.7
59Q4 to 64Q4	5.6	5.5	5.6	5.6	5.7	5.6
64Q4 to 69Q4	8.0	8.1	8.2	8.3	8.2	8.2
69Q4 to 74Q4	8.6	8.8	8.9	9.1	9.0	9.1
74Q4 to 79Q4	NA	11.1	11.2	11.3	11.4	11.4
79Q4 to 84Q4	NA	NA	8.5	8.2	8.5	8.6
84Q4 to 89Q4	NA	NA	NA	6.5	6.7	6.7
89Q4 to 94Q4	NA	NA	NA	NA	5.2	5.1

Table III.2
Contemporaneous Correlations Across Benchmark Vintages
Quarterly changes in logarithms of variables

Vintage Year:	‘75	‘80	‘85	‘91	‘95	‘98
			Real Output			
‘75	1.00					
‘80	0.97	1.00				
‘85	0.97	0.99	1.00			
‘91	0.92	0.95	0.96	1.00		
‘95	0.91	0.94	0.95	0.97	1.00	
‘98	0.92	0.92	0.94	0.94	0.95	1.00
			Real Consumption			
‘75	1.00					
‘80	0.96	1.00				
‘85	0.96	0.99	1.00			
‘91	0.94	0.97	0.98	1.00		
‘95	0.94	0.96	0.97	0.98	1.00	
‘98	0.96	0.96	0.97	0.97	0.97	1.00
			Prices			
‘75	1.00					
‘80	0.94	1.00				
‘85	0.94	0.99	1.00			
‘91	0.87	0.96	0.96	1.00		
‘95	0.85	0.94	0.95	0.95	1.00	
‘98	0.94	0.93	0.93	0.90	0.90	1.00
			Nominal Output			
‘75	1.00					
‘80	0.98	1.00				
‘85	0.98	0.99	1.00			
‘91	0.96	0.98	0.98	1.00		
‘95	0.96	0.97	0.98	0.99	1.00	
‘98	0.95	0.97	0.98	0.99	0.99	1.00

Table IV.1
Standard Deviations of Real Consumption Growth Rates
1965Q3 to 1995Q3
In percentage points

Data Set	Standard Deviation
Initial Release	3.40
1-Year Later	3.57
3-Years Later	3.17
Latest	3.10

Table IV.2
Correlations of Revisions with Growth Rates
Consumption Data
1965Q3 to 1995Q3

Revisions/Data Set	Initial	1-Year	3-Year	Final
Initial to 1-Year	-0.02	0.30*	0.27*	0.20*
	(0.18)	(3.45)	(3.02)	(2.25)
1-Year to 3-Year	-0.44†	-0.48†	-0.15	-0.14
	(5.37)	(5.98)	(1.69)	(1.56)
3-Year to Final	-0.16	-0.22†	-0.23†	0.11
	(1.79)	(2.45)	(2.58)	(1.18)
Initial to 3-Year	-0.37†	-0.17	0.08	0.04
	(4.39)	(1.84)	(0.82)	(0.40)
1-Year to Final	-0.43†	-0.50†	-0.26?	-0.04
	(5.21)	(6.25)	(2.95)	(0.40)
Initial to Final	-0.42†	-0.28?	-0.08	0.10
	(5.08)	(3.16)	(0.83)	(1.06)

Absolute values of t-statistics are in parentheses below each correlation coefficient.
An asterisk (*) means there's a significant (at the 5% level) correlation between the revision and the later data, implying "news."
A dagger (†) means there's a significant (at the 5% level) correlation between the revision and the earlier data, implying "noise."
A question mark (?) means there's a significant correlation that doesn't fit easily into the news/noise dichotomy.

Table V.1
Kydland-Prescott Cross-Correlations

		Cross Correlation of Real GNP/GDP With										
Vintage	Variable x	x(t-5)	x(t-4)	x(t-3)	x(t-2)	x(t-1)	x(t)	x(t+1)	x(t+2)	x(t+3)	x(t+4)	x(t+5)
KP 1990	Real GNP/GDP	-0.03	0.15	0.38	0.63	0.85						
Feb. 1990		-0.03	0.15	0.37	0.62	0.85						
Feb. 1994		-0.02	0.15	0.36	0.61	0.84						
Feb. 1998		-0.09	0.11	0.34	0.60	0.84						
KP 1990	GNP/GDP deflator	-0.50	-0.61	-0.68	-0.69	-0.64	-0.55	-0.43	-0.31	-0.17	-0.04	0.09
Feb. 1990		-0.49	-0.60	-0.67	-0.69	-0.64	-0.56	-0.43	-0.31	-0.18	-0.05	0.08
Feb. 1994		-0.51	-0.60	-0.66	-0.66	-0.59	-0.48	-0.36	-0.26	-0.15	-0.05	0.07
Feb. 1998		-0.35	-0.49	-0.60	-0.68	-0.70	-0.66	-0.55	-0.40	-0.22	-0.04	0.12
KP 1990	Real Consumption	0.25	0.41	0.56	0.71	0.81	0.82	0.66	0.45	0.21	-0.02	-0.21
Feb. 1990		0.24	0.40	0.55	0.70	0.80	0.82	0.65	0.44	0.21	-0.02	-0.21
Feb. 1994		0.18	0.35	0.53	0.71	0.84	0.87	0.70	0.48	0.25	0.02	-0.17
Feb. 1998		0.12	0.31	0.50	0.69	0.84	0.88	0.71	0.48	0.23	-0.02	-0.20
KP 1990	M2	0.48	0.60	0.67	0.68	0.61	0.46	0.26	0.05	-0.15	-0.33	-0.46
Feb. 1990		0.46	0.57	0.64	0.66	0.60	0.46	0.25	0.05	-0.14	-0.31	-0.42
Feb. 1994		0.44	0.57	0.65	0.69	0.64	0.50	0.29	0.08	-0.10	-0.28	-0.41
Feb. 1998		0.44	0.58	0.66	0.69	0.63	0.48	0.27	0.07	-0.12	-0.28	-0.40

Table V.2
Hall's Tests on Consumption

Regression 1: $c_t = \hat{a}_0 + \hat{a}_1 c_{t-1} + \hat{a}_2 c_{t-2} + \hat{a}_3 c_{t-3} + \hat{a}_4 c_{t-4} + e_t$

	$\hat{a}_0$	$\hat{a}_1$	$\hat{a}_2$	$\hat{a}_3$	$\hat{a}_4$	R^2	s	DW	F
Sample 1948Q1 to 1977Q1									
Hall's results	8.2 (8.3)	1.130 (0.092)	-0.040 (0.142)	0.030 (0.142)	-0.113 (0.093)	.9988	14.5	1.96	1.7 (0.17)
Replication vintage 1977Q2	-8.122 (8.489)	1.130 (0.092)	-0.024 (0.142)	-0.004 (0.143)	-0.095 (0.094)	.9988	14.7	1.97	1.7 (0.17)
Replication vintage 1998Q1	-9.859 (27.498)	1.102 (0.093)	0.166 (0.138)	-0.256 (0.137)	-0.007 (0.094)	.9988	57.5	2.00	3.5 (0.02)
Sample 1948Q1 to 1997Q4									
Vintage 1998Q1	15.589 (14.296)	1.153 (0.070)	0.163 (0.108)	-0.011 (0.108)	-0.157 (0.070)	.9997	57.0	1.97	8.1 (0.00)

Table VII.1
Forecast Errors from Rolling Regressions
Forecast Horizons 1976Q1 to 1998Q3
Four-quarter average forecasts of real output

Forecast	Data Set	Mean Error	Mean Absolute Error	Root Mean Square Error
AR(4)	Real Time	-0.19	1.74	2.49
	November 1998	-0.48	1.70	2.40
BAR(4)	Real Time	-0.24	1.67	2.35
	November 1998	-0.54	1.67	2.34
QBVEC	Real Time	-0.70	1.41	1.90
	November 1998	-0.79	1.46	1.88

REFERENCES

Amato, Jeffery D. "The Real-Time (In)Significance of M2," Manuscript, Federal Reserve Bank of Kansas City, September 1998.

Beveridge, Stephen, and Charles R. Nelson. "A New Approach to Decomposition of Economic Time Series into Permanent and Transitory Components with Particular Attention to Measurement of the Business Cycle," *Journal of Monetary Economics* 7 (March 1981), pp. 151-74.

Blanchard, Olivier Jean, and Stanley Fischer. *Lectures on Macroeconomics*. Cambridge, Mass.: MIT Press, 1989.

Blanchard, Olivier Jean, and Danny Quah. "The Dynamic Effects of Aggregate Demand and Supply Disturbances," *American Economic Review* 79 (September 1989), pp. 655-73.

Boschen, John F., and Leonard O. Mills. "The Relation Between Narrative and Money Market Indicators of Monetary Policy," *Economic Inquiry* 33 (January 1995), pp. 24-44.

Campbell, John Y., and N. Gregory Mankiw. "Are Output Fluctuations Transitory?" *Quarterly Journal of Economics* 102 (November 1987), pp. 857-80.

Cochrane, John H. "Permanent and Transitory Components of GNP and Stock Prices," *Quarterly Journal of Economics* (February 1994), pp. 241-65.

Croushore, Dean, and Charles L. Evans. "Data Revisions and the Identification of Monetary Policy Shocks," Manuscript, Federal Reserve Bank of Chicago, February 1999.

Deaton, Angus S. "Life-Cycle Models of Consumption: Is the Evidence Consistent with the Theory?" in T.F. Bewley, ed., *Advances in Econometrics: Fifth World Congress, Vol. II* (Cambridge: Cambridge University Press, 1987), pp. 121-48.

Diebold, Francis X., and Glenn D. Rudebusch. "Forecasting Output With the Composite Leading Index: A Real-Time Analysis," *Journal of the American Statistical Association* 86 (September 1991), pp. 603-10.

Evans, Charles L. "Real-Time Taylor Rules and the Federal Funds Futures Market," Federal Reserve Bank of Chicago *Economic Perspectives* (Third Quarter 1998), pp. 44-55.

Flavin, Marjorie A. "The Adjustment of Consumption to Changing Expectations about Future Income," *Journal of Political Economy* 89 (1981), pp. 974-1009.

Ghysels, Eric, Norman R. Swanson, and Myles Callan. "Monetary Policy Rules and Data Uncertainty," Manuscript, Penn State University, September 1998.

Hall, Robert E. "Stochastic Implications of the Life Cycle-Permanent Income Hypothesis: Theory and Evidence," *Journal of Political Economy* 86 (December 1978), pp. 971-87.

Koenig, Evan F., and Sheila Dolmas. "Real-Time GDP Growth Forecasts," Federal Reserve Bank of Dallas Working Paper 97-10, December 1997.

Kydland, Finn E., and Edward C. Prescott. "Business Cycles: Real Facts and a Monetary Myth," Federal Reserve Bank of Minneapolis *Quarterly Review* (Spring 1990), pp. 3-18.

Litterman, Robert B. "Forecasting with Bayesian Vector Autoregressions--Five Years of Experience," *Journal of Business and Economic Statistics* 4 (January 1986), pp. 25-38.

Mankiw, N. Gregory, and Matthew D. Shapiro. "News or Noise: An Analysis of GNP Revisions," *Survey of Current Business* (May 1986), pp. 20-5.

Maravall, Augustin, and David A. Pierce. "The Transmission of Data Noise into Policy Noise in U.S. Monetary Control," *Econometrica* 54 (July 1986), pp. 961-79.

McNees, Stephen K. “A Forward-Looking Monetary Policy Reaction Function: Continuity and Change,” Federal Reserve Bank of Boston *New England Economic Review* (Nov./Dec. 1992), pp. 3-13.

Orphanides, Athanasios. “Monetary Rules Based on Real-Time Data,” Federal Reserve Board of Governors, Finance and Economics Discussion Series 1998-03, December 1997.

Orphanides, Athanasios. “Monetary Policy Evaluation With Noisy Information,” Manuscript, Federal Reserve Board of Governors, October 1998.

Robertson, John C., and Ellis W. Tallman. “Real-Time Forecasting with a VAR Model,” Manuscript, Federal Reserve Bank of Atlanta, November 1998a.

Robertson, John C., and Ellis W. Tallman. “Data Vintages and Measuring Forecast Model Performance,” Federal Reserve Bank of Atlanta *Economic Review* (Fourth Quarter 1998b), pp. 4-20.

Rudebusch, Glenn D. “Do Measures of Monetary Policy in a VAR Make Sense?” *International Economic Review* 39 (1998a), pp. 907-31.

Rudebusch, Glenn D. “Is the Fed Too Timid? Monetary Policy in an Uncertain World,” Manuscript, Federal Reserve Bank of San Francisco, October 1998b.

Runkle, David E. “Revisionist History: How Data Revisions Distort Economic Policy Research,” Federal Reserve Bank of Minneapolis *Quarterly Review* (Fall 1998), pp. 3-12.

Sarte, Pierre-Daniel G. “On the Identification of Structural Vector Autoregressions,” Federal Reserve Bank of Richmond *Economic Quarterly* (Summer 1997), pp. 45-67.

Stark, Tom. “A Bayesian Vector Error Corrections Model of the U.S. Economy,” Federal Reserve Bank of Philadelphia Working Paper 98-12, June 1998.

Taylor, John B. "Discretion versus Policy Rules in Practice," *Carnegie-Rochester Conference Series on Public Policy* 39 (December 1993), pp. 195-214.

Figure III.1
Log Nominal Output Ratios

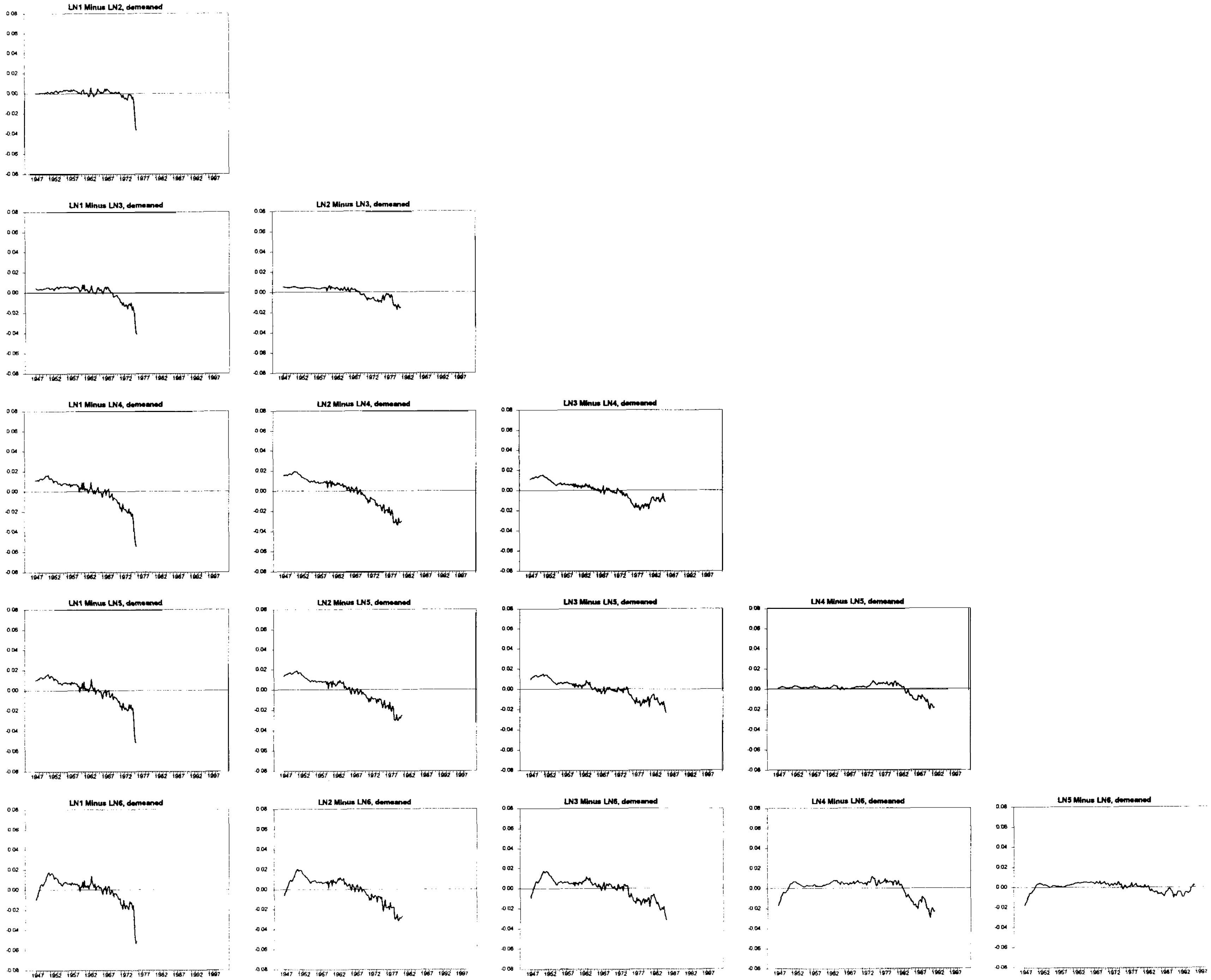


Figure III.2
Log Real Output Ratios

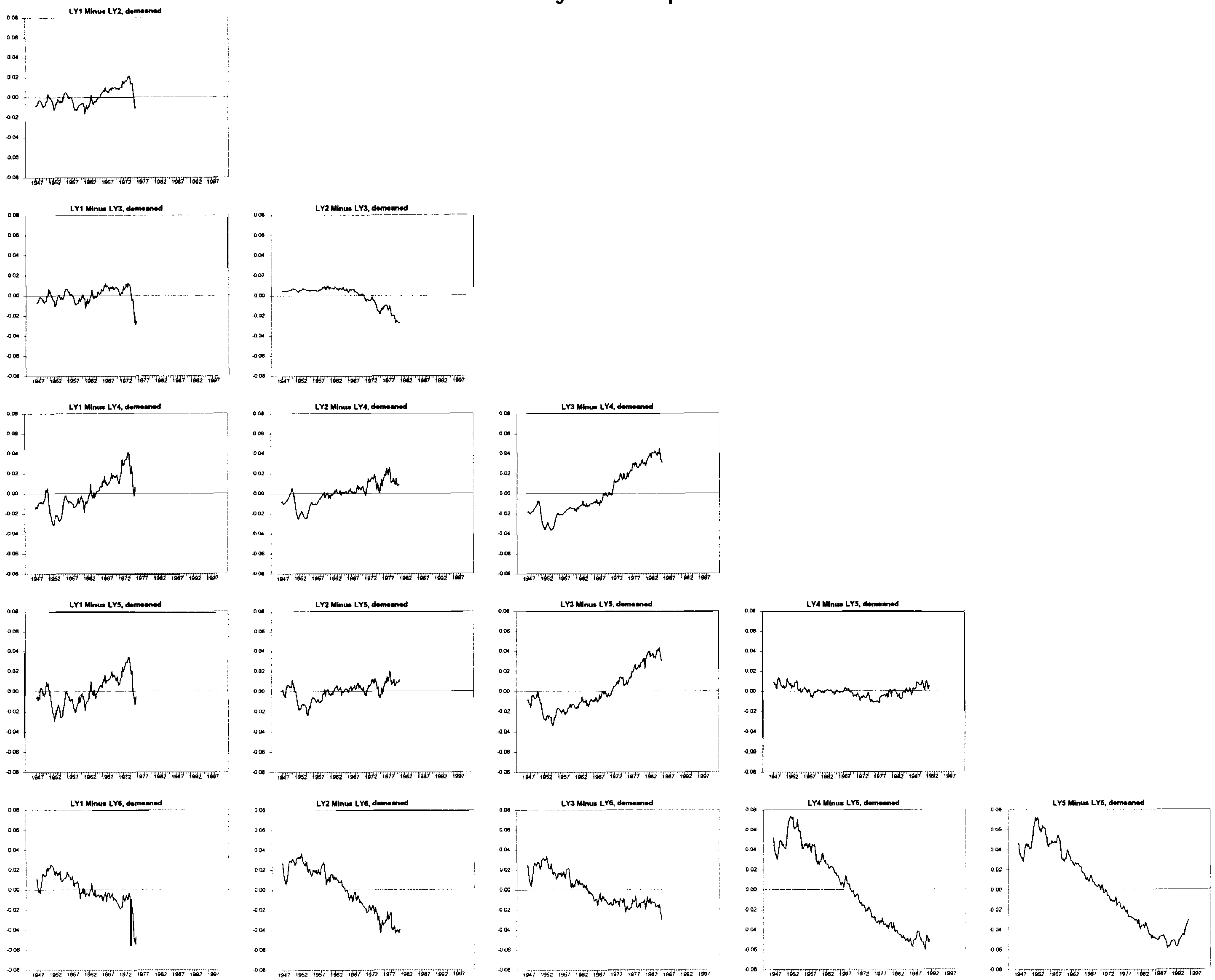


Figure III.3
Log Price Level Ratios

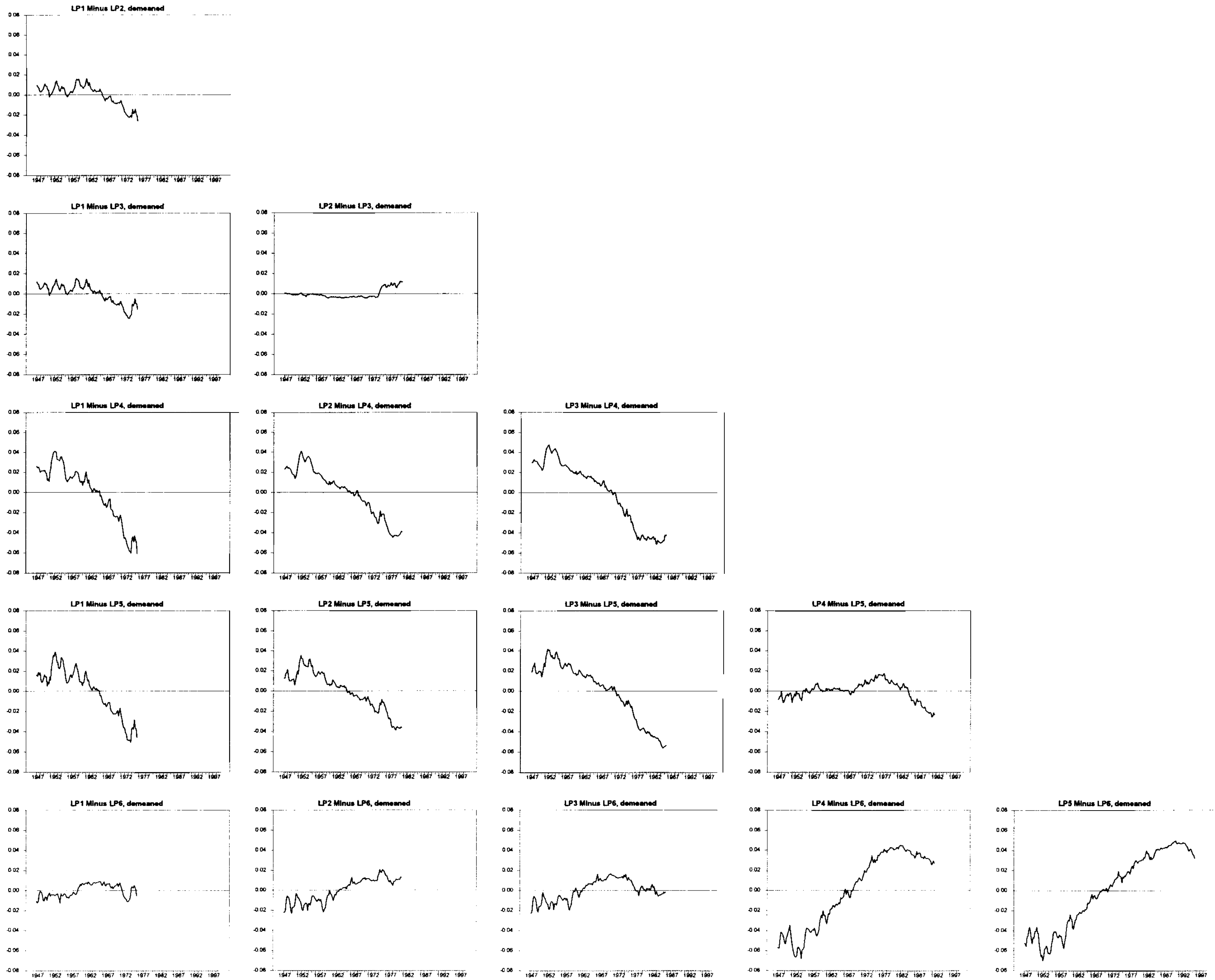


Figure III.4
Log Real PCE Ratios

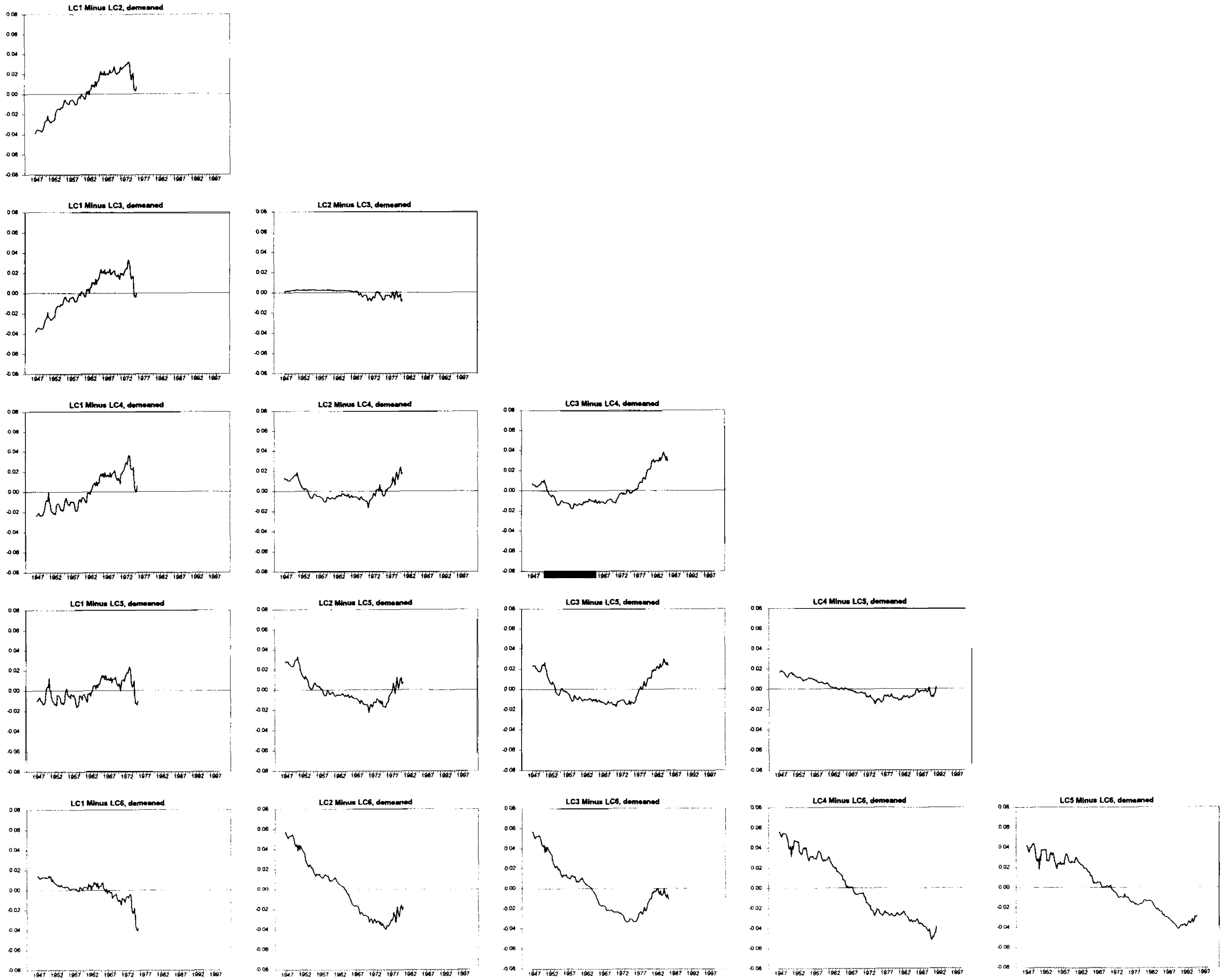


Figure IV.1

Annualized Quarter-Over-Quarter Real PCE Growth Rates

1947Q2 to 1985Q4

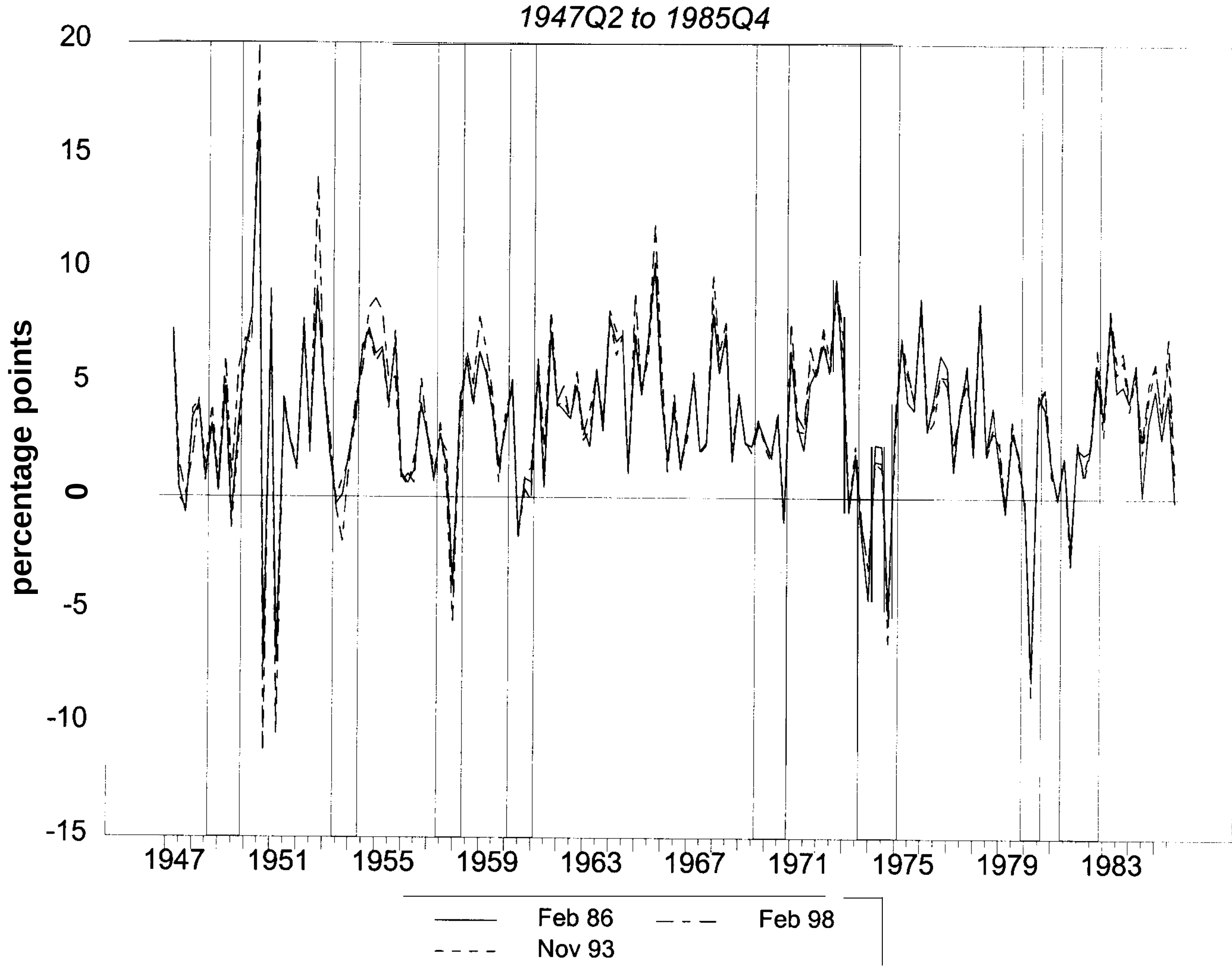
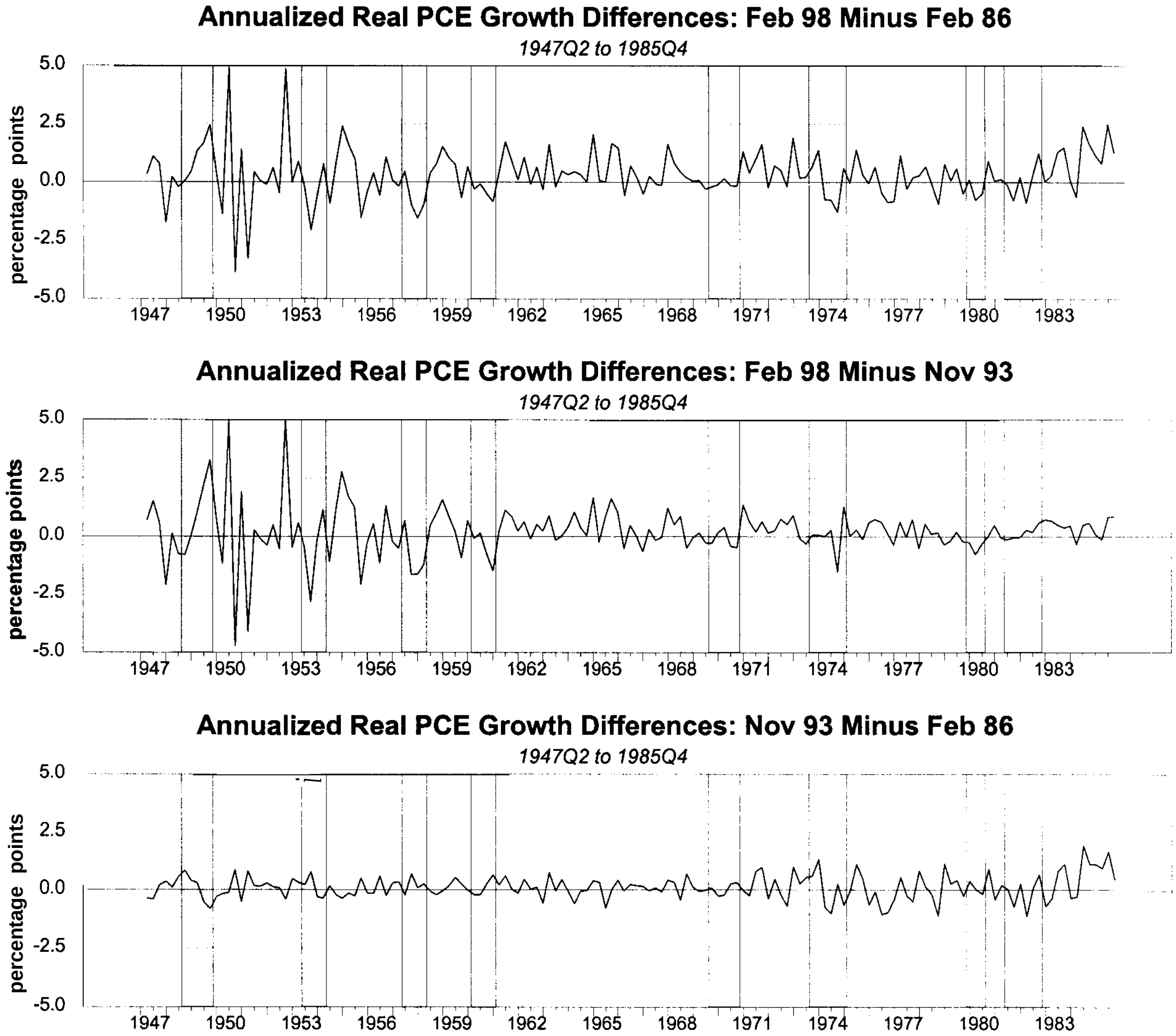


Figure IV.2



Annualized Real PCE Growth Differences: Feb 98 Minus Feb 86
1947Q2 to 1985Q4
percentage points
5.0
2.5
0.0
-2.5
-5.0
1947
1950
1953
1956
1959
1962
1965
1968
1971
1974
1977
1980
1983
Annualized Real PCE Growth Differences: Feb 98 Minus Nov 93
1947Q2 to 1985Q4
percentage points
5.0
2.5
0.0
-2.5
-5.0
1947
1950
1953
1956
1959
1962
1965
1968
1971
1974
1977
1980
1983
Annualized Real PCE Growth Differences: Nov 93 Minus Feb 86
1947Q2 to 1985Q4
percentage points
5.0
2.5
0.0
-2.5
-5.0
1947
1950
1953
1956
1959
1962
1965
1968
1971
1974
1977
1980
1983

Figure IV.3

Real Consumption Growth Rates:

Initial Estimate

1 -Year Later Estimate

3-Years Later Estimate

Latest Estimate

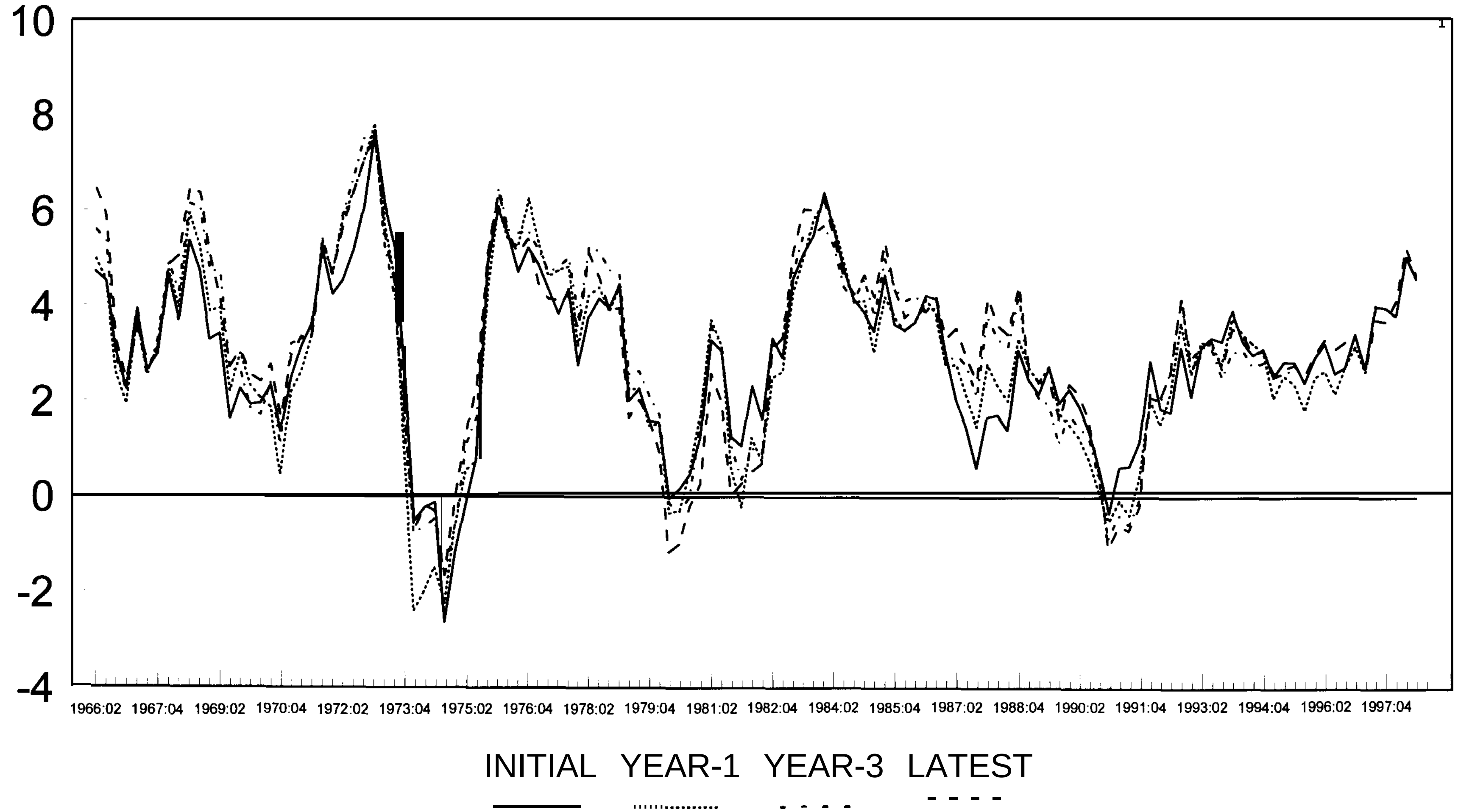


Figure IV.4

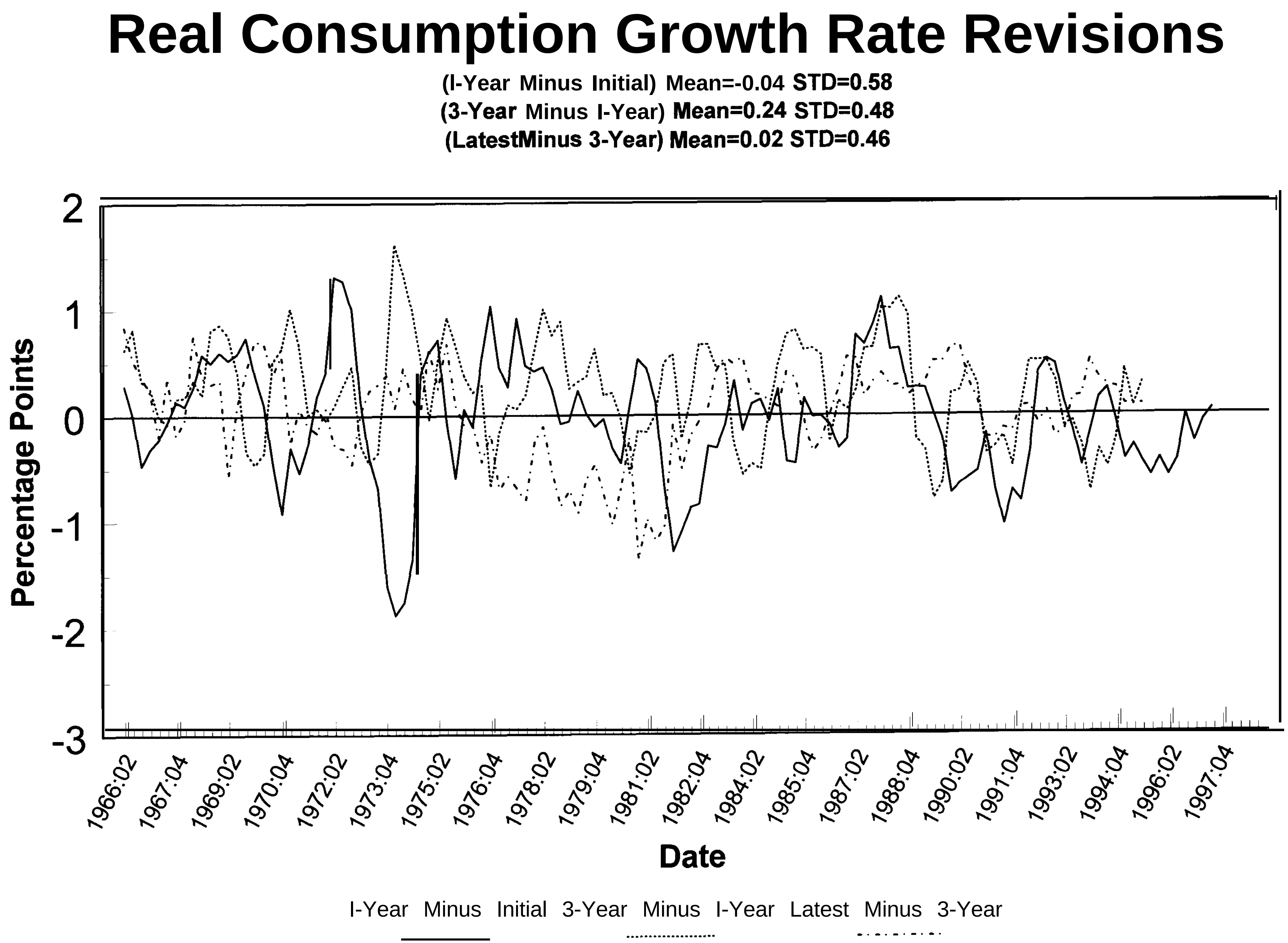

Real Consumption Growth Rate Revisions
(I-Year Minus Initial) Mean=-0.04 STD=0.58
(3-Year Minus I-Year) Mean=0.24 STD=0.48
(LatestMinus 3-Year) Mean=0.02 STD=0.46
2
1
0
-1
-2
-3
Percentage Points
1966:02
1967:04
1969:02
1970:04
1972:02
1973:04
1975:02
1976:04
1978:02
1979:04
1981:02
1982:04
1984:02
1985:04
1987:02
1988:04
1990:02
1991:04
1993:02
1994:04
1996:02
1997:04
Date
I-Year Minus Initial 3-Year Minus I-Year Latest Minus 3-Year

Figure IV.5

Real Consumption Growth Rate Revisions

(FinalMinus Initial Revision) Mean=0.21 STD=0.78

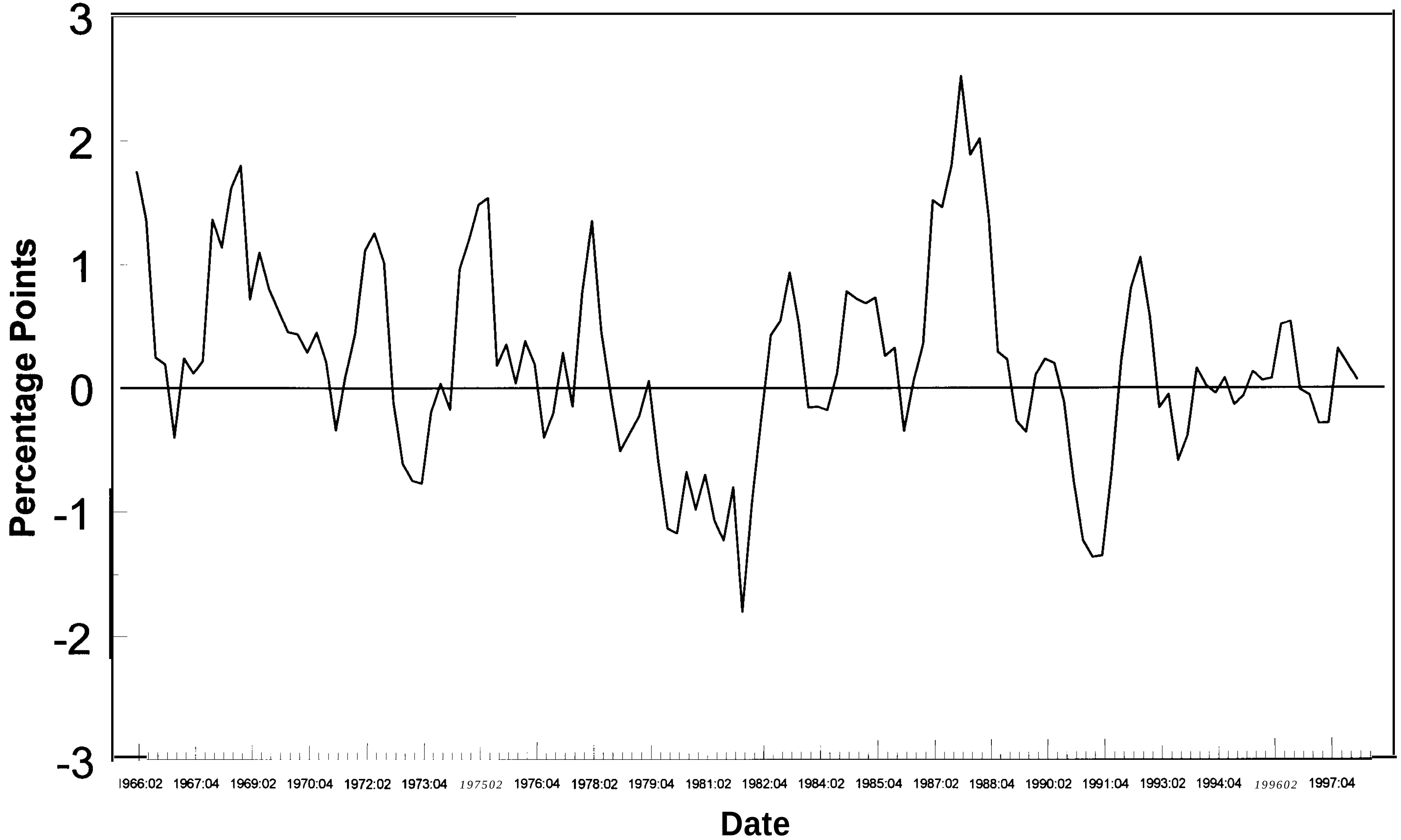
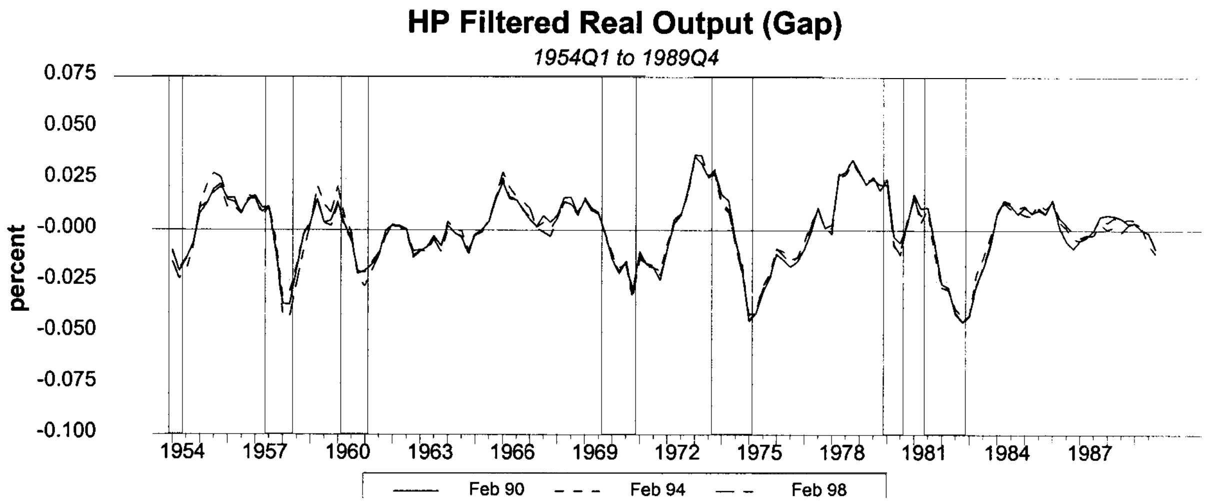


Figure V.1


HP Filtered Real Output (Gap)
1954Q1 to 1989Q4
percent
0.075
0.050
0.025
-0.000
-0.025
-0.050
-0.075
-0.100
1954
1957
1960
1963
1966
1969
1972
1975
1978
1981
1984
1987
Feb 90
Feb 94
Feb 98

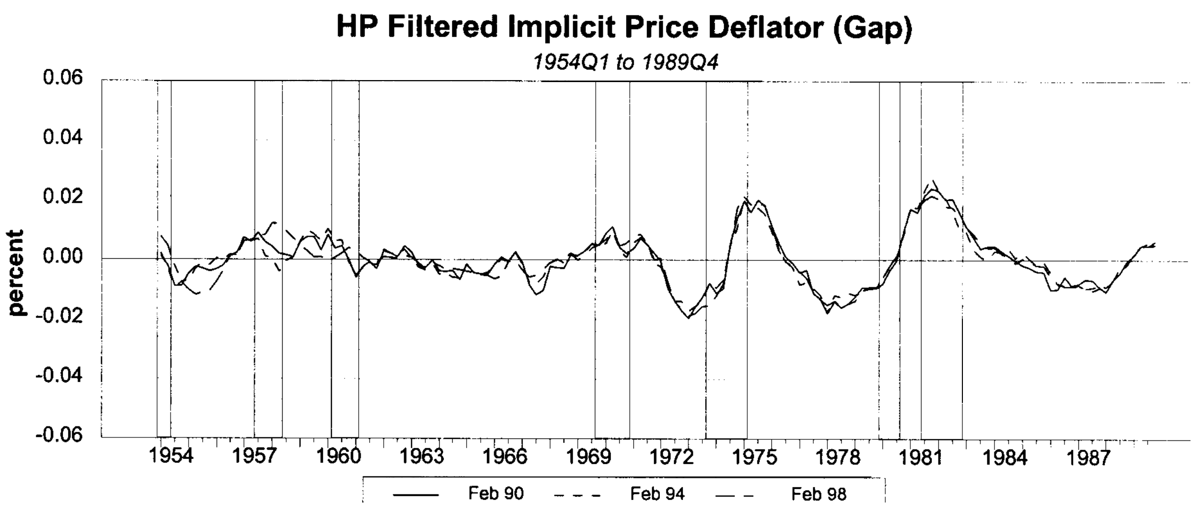

HP Filtered Implicit Price Deflator (Gap)
1954Q1 to 1989Q4
percent
0.06
0.04
0.02
0.00
-0.02
-0.04
-0.06
1954
1957
1960
1963
1966
1969
1972
1975
1978
1981
1984
1987
Feb 90
Feb 94
Feb 98

Figure V.2

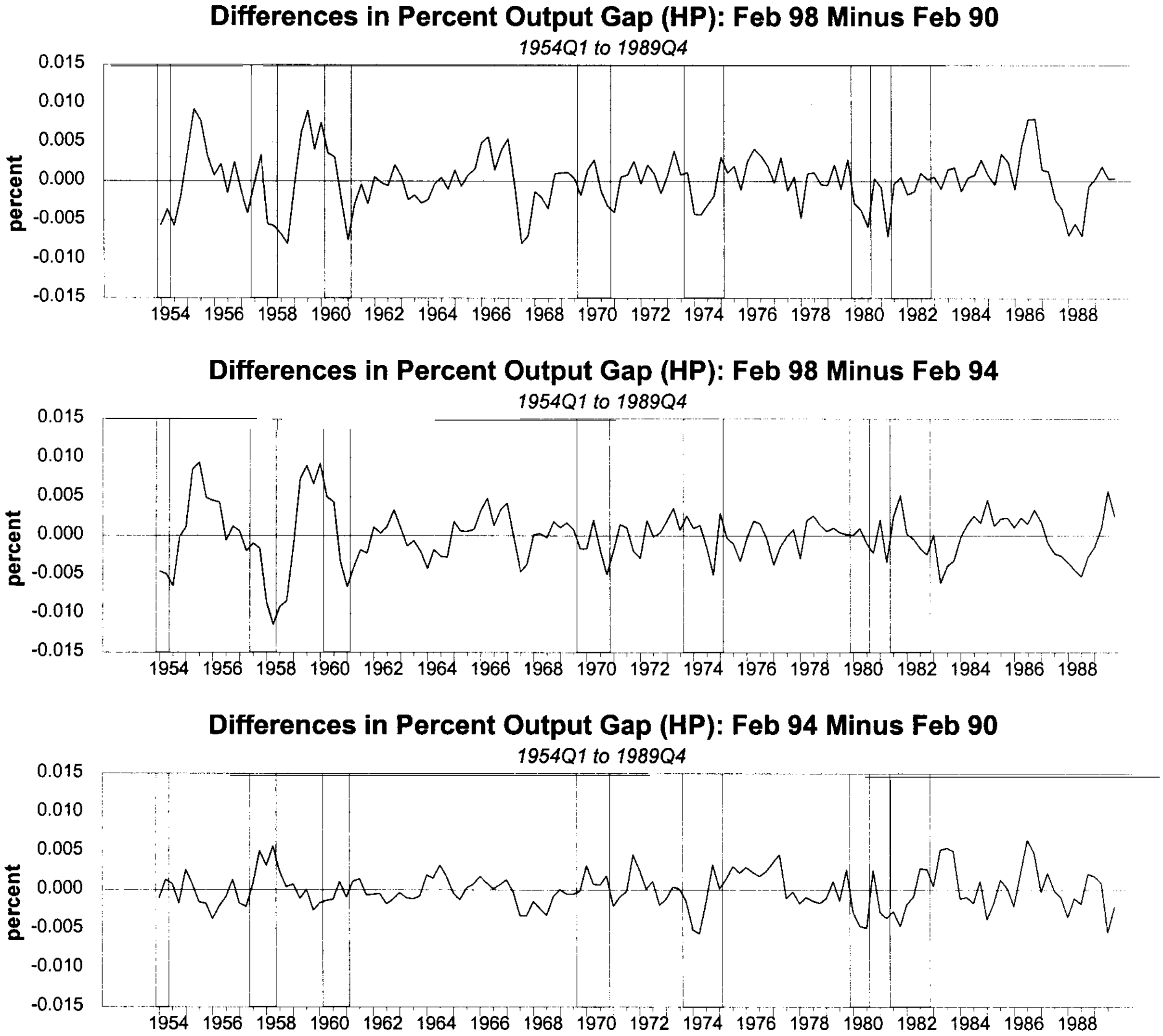

Differences in Percent Output Gap (HP): Feb 98 Minus Feb 90
1954Q1 to 1989Q4
percent
0.015
0.010
0.005
0.000
-0.005
-0.010
-0.015
1954 1956 1958 1960 1962 1964 1966 1968 1970 1972 1974 1976 1978 1980 1982 1984 1986 1988
Differences in Percent Output Gap (HP): Feb 98 Minus Feb 94
1954Q1 to 1989Q4
percent
0.015
0.010
0.005
0.000
-0.005
-0.010
-0.015
1954 1956 1958 1960 1962 1964 1966 1968 1970 1972 1974 1976 1978 1980 1982 1984 1986 1988
Differences in Percent Output Gap (HP): Feb 94 Minus Feb 90
1954Q1 to 1989Q4
percent
0.015
0.010
0.005
0.000
-0.005
-0.010
-0.015
1954 1956 1958 1960 1962 1964 1966 1968 1970 1972 1974 1976 1978 1980 1982 1984 1986 1988

Figure V.3

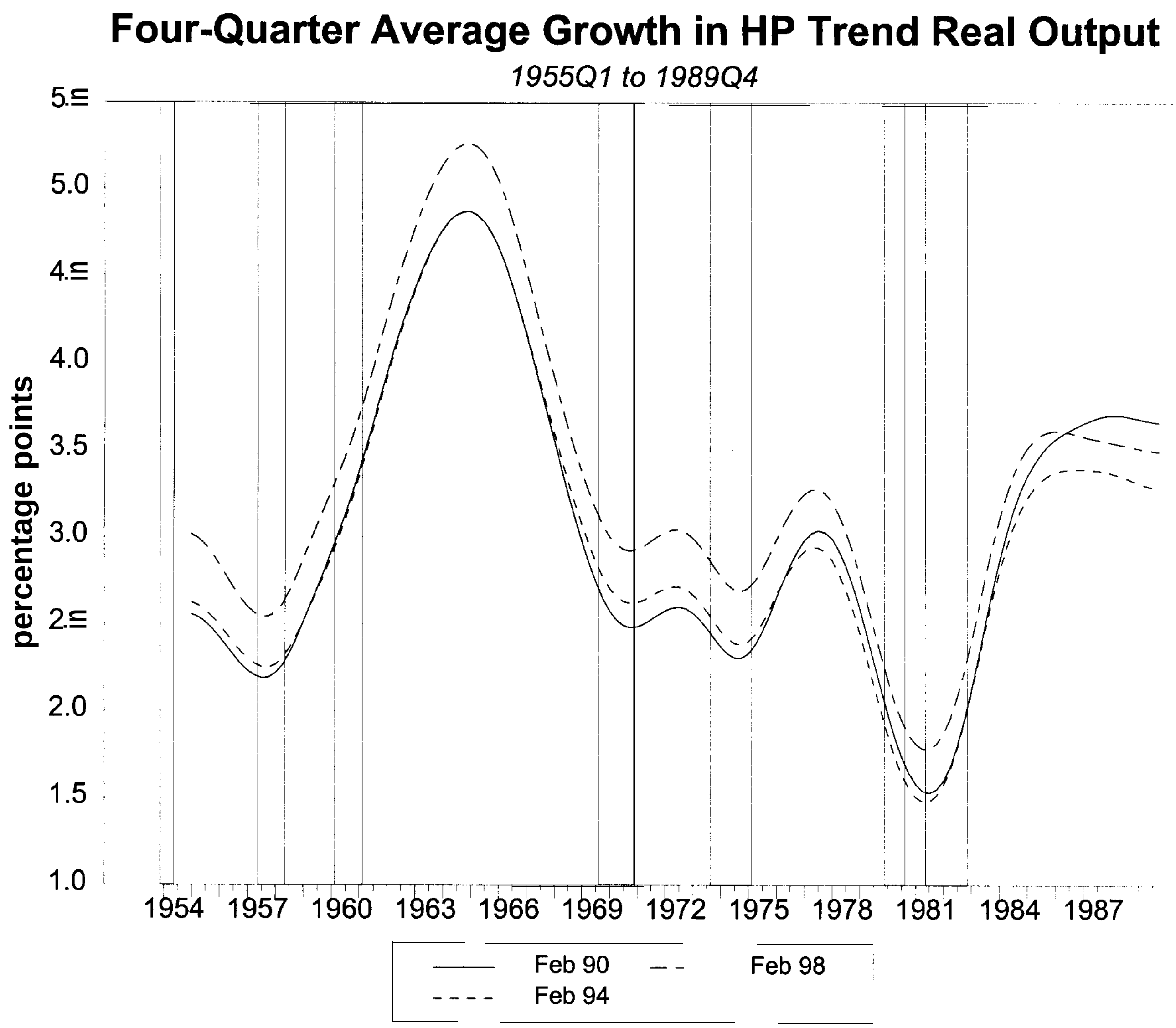

Four-Quarter Average Growth in HP Trend Real Output
1955Q1 to 1989Q4
percentage points
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
1954
1957
1960
1963
1966
1969
1972
1975
1978
1981
1984
1987
Feb 90
Feb 98
Feb 94

Figure V.4

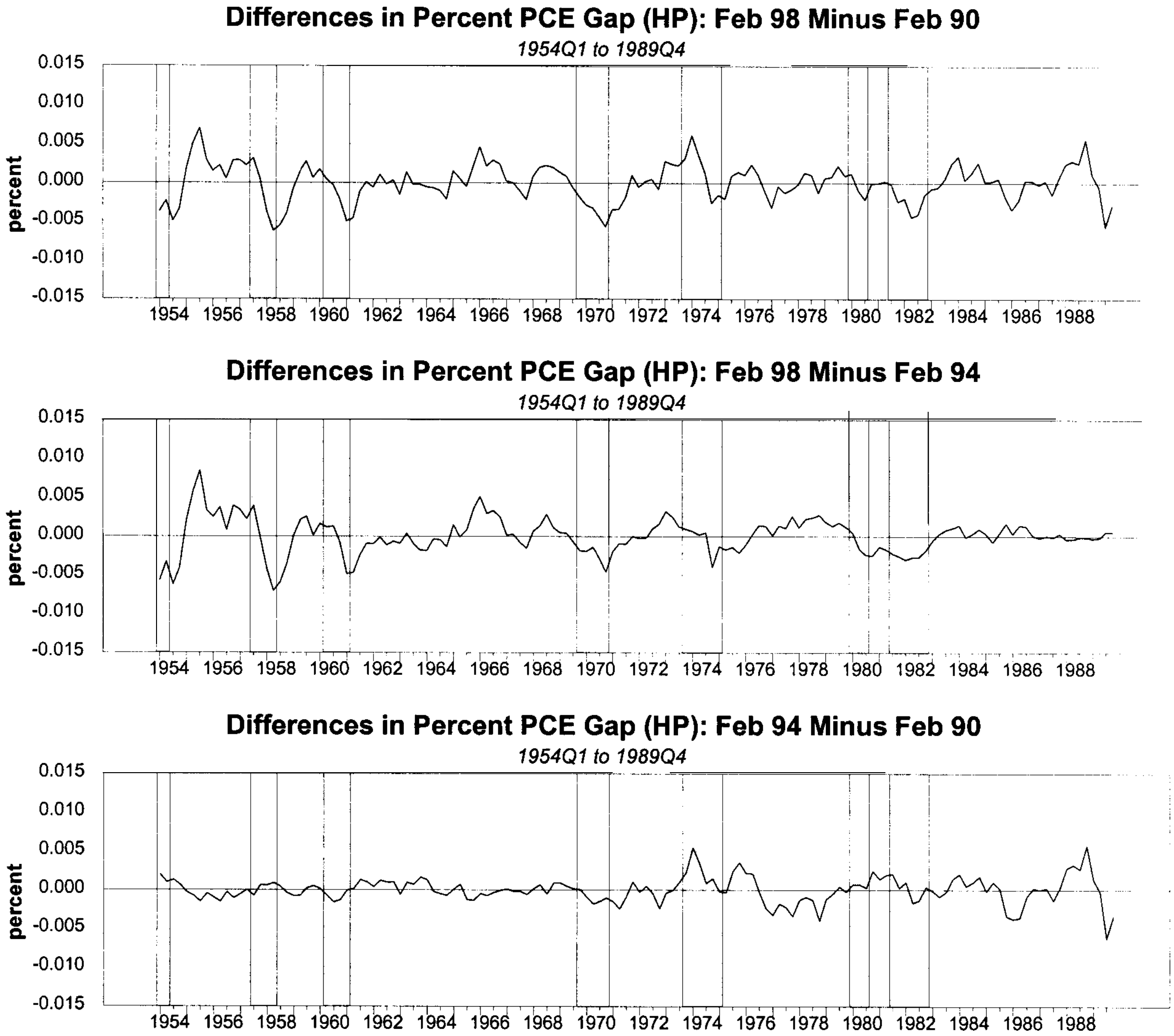

Differences in Percent PCE Gap (HP): Feb 98 Minus Feb 90
1954Q1 to 1989Q4
percent
0.015
0.010
0.005
0.000
-0.005
-0.010
-0.015
1954 1956 1958 1960 1962 1964 1966 1968 1970 1972 1974 1976 1978 1980 1982 1984 1986 1988
Differences in Percent PCE Gap (HP): Feb 98 Minus Feb 94
1954Q1 to 1989Q4
percent
0.015
0.010
0.005
0.000
-0.005
-0.010
-0.015
1954 1956 1958 1960 1962 1964 1966 1968 1970 1972 1974 1976 1978 1980 1982 1984 1986 1988
Differences in Percent PCE Gap (HP): Feb 94 Minus Feb 90
1954Q1 to 1989Q4
percent
0.015
0.010
0.005
0.000
-0.005
-0.010
-0.015
1954 1956 1958 1960 1962 1964 1966 1968 1970 1972 1974 1976 1978 1980 1982 1984 1986 1988

Figure V.5

A Comparison of Beveridge/Nelson Transitory Real Output

Specification: ARIMA(1,1,2), 1947Q3 to 1977Q1

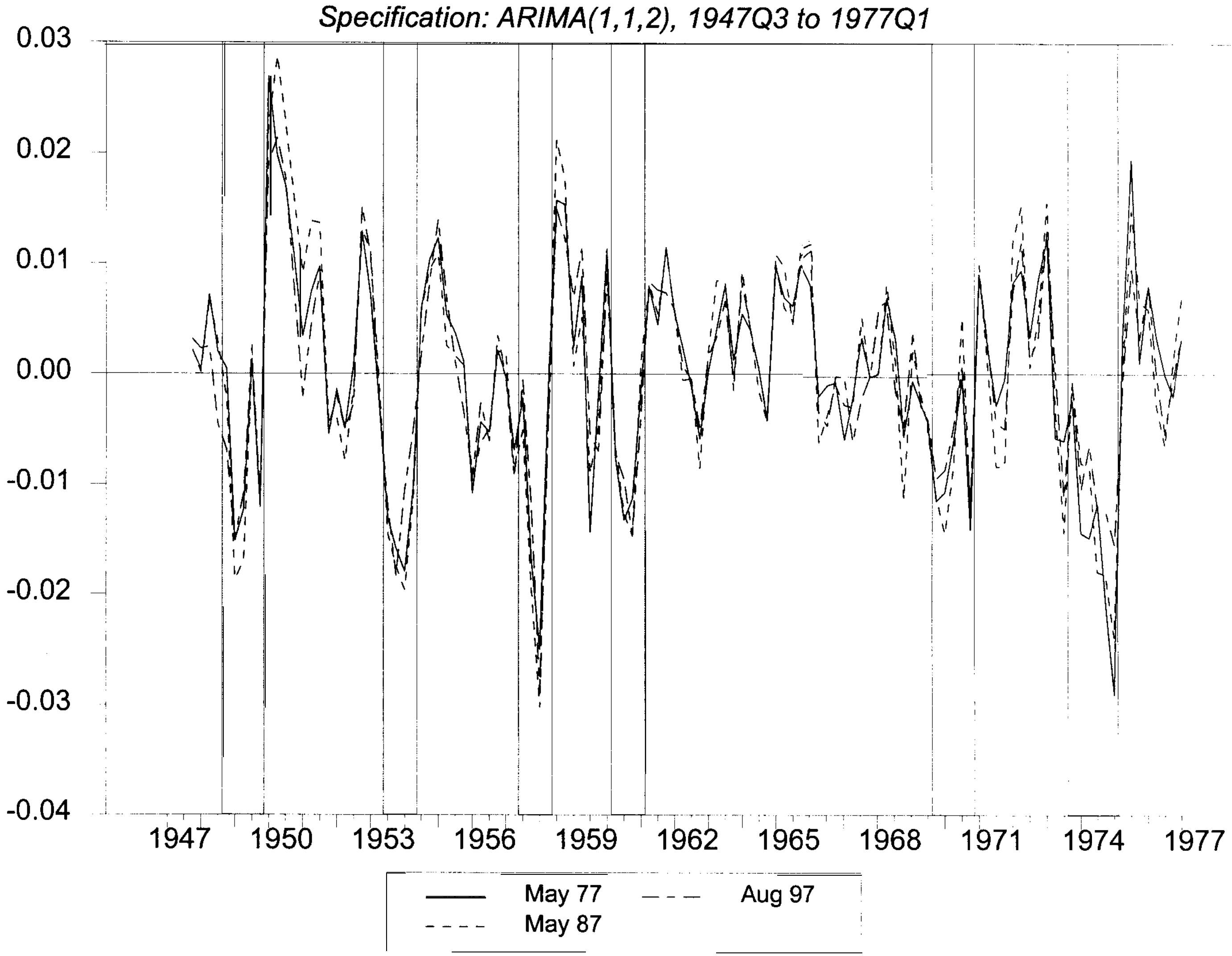


Figure V.6

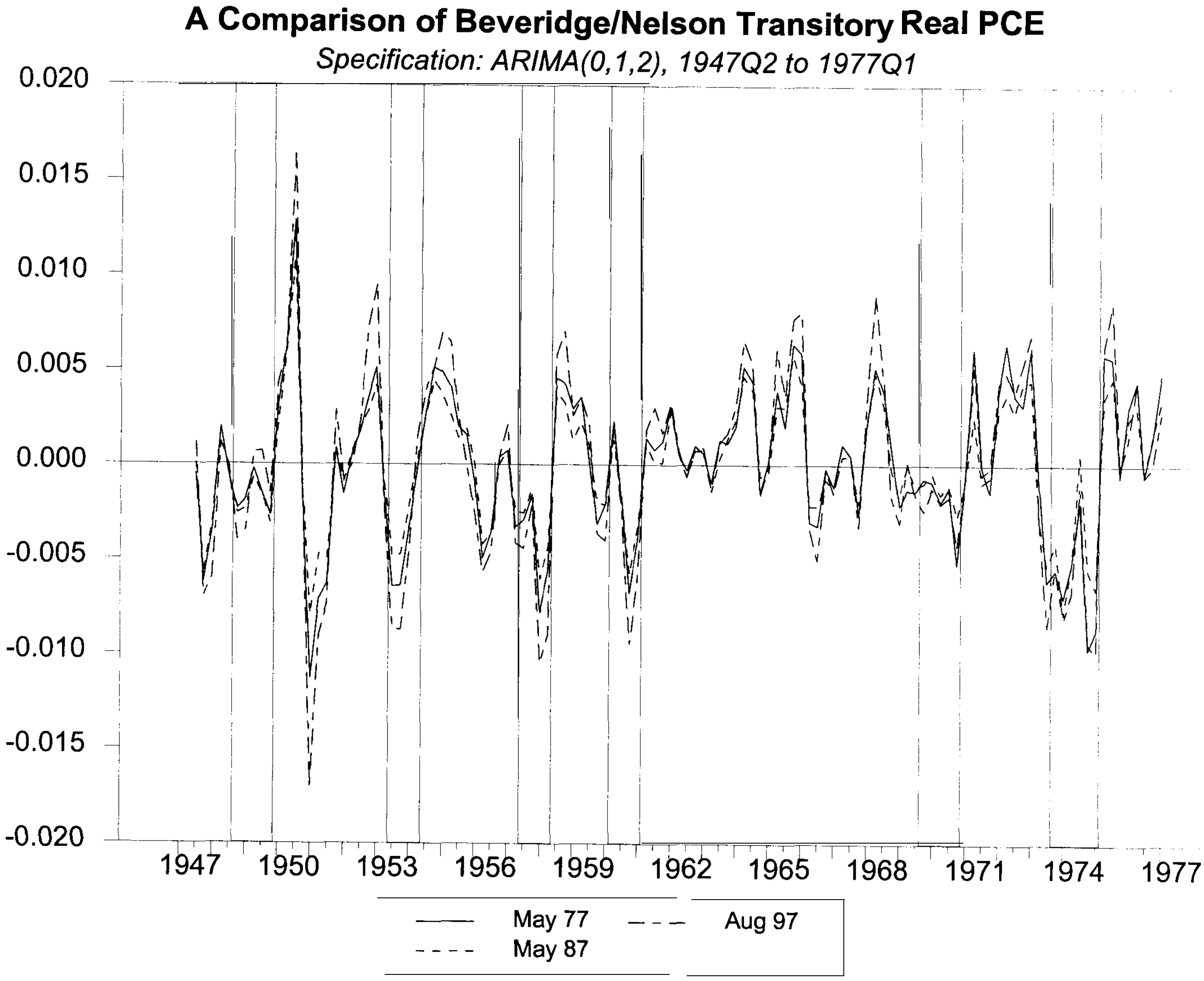
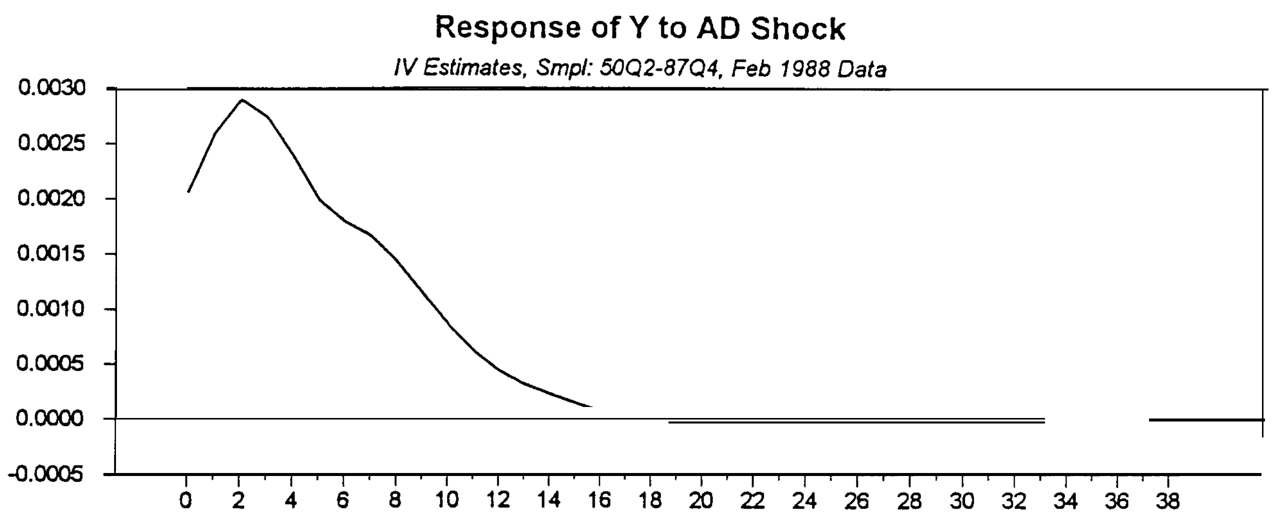
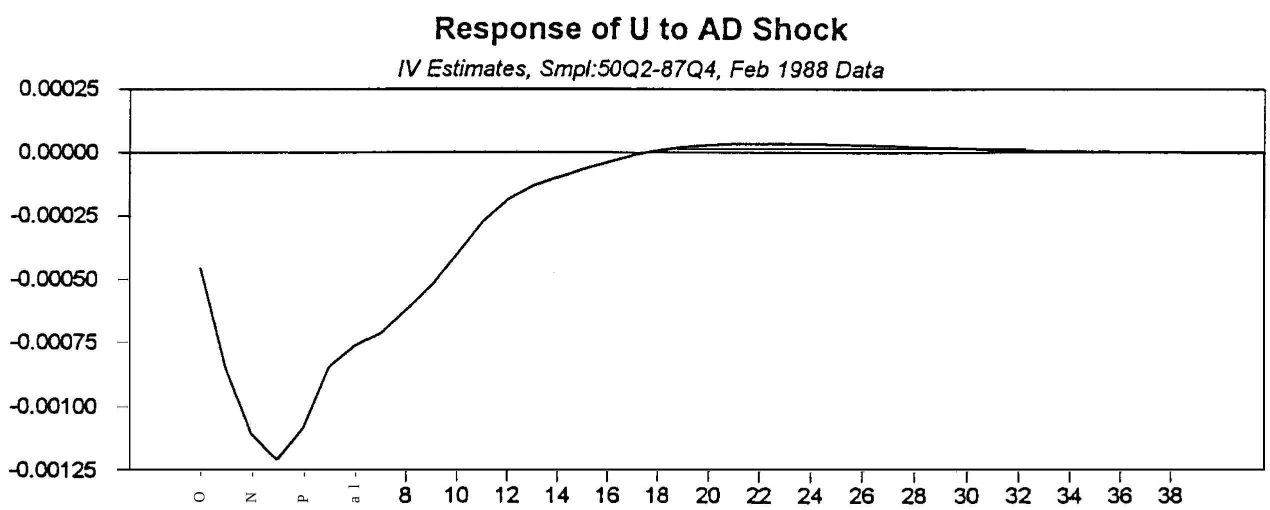

A Comparison of Beveridge/Nelson Transitory Real PCE
Specification: ARIMA(0,1,2), 1947Q2 to 1977Q1
0.020
0.015
0.010
0.005
0.000
-0.005
-0.010
-0.015
-0.020
1947
1950
1953
1956
1959
1962
1965
1968
1971
1974
1977
May 77
Aug 97
May 87

Figure V.7

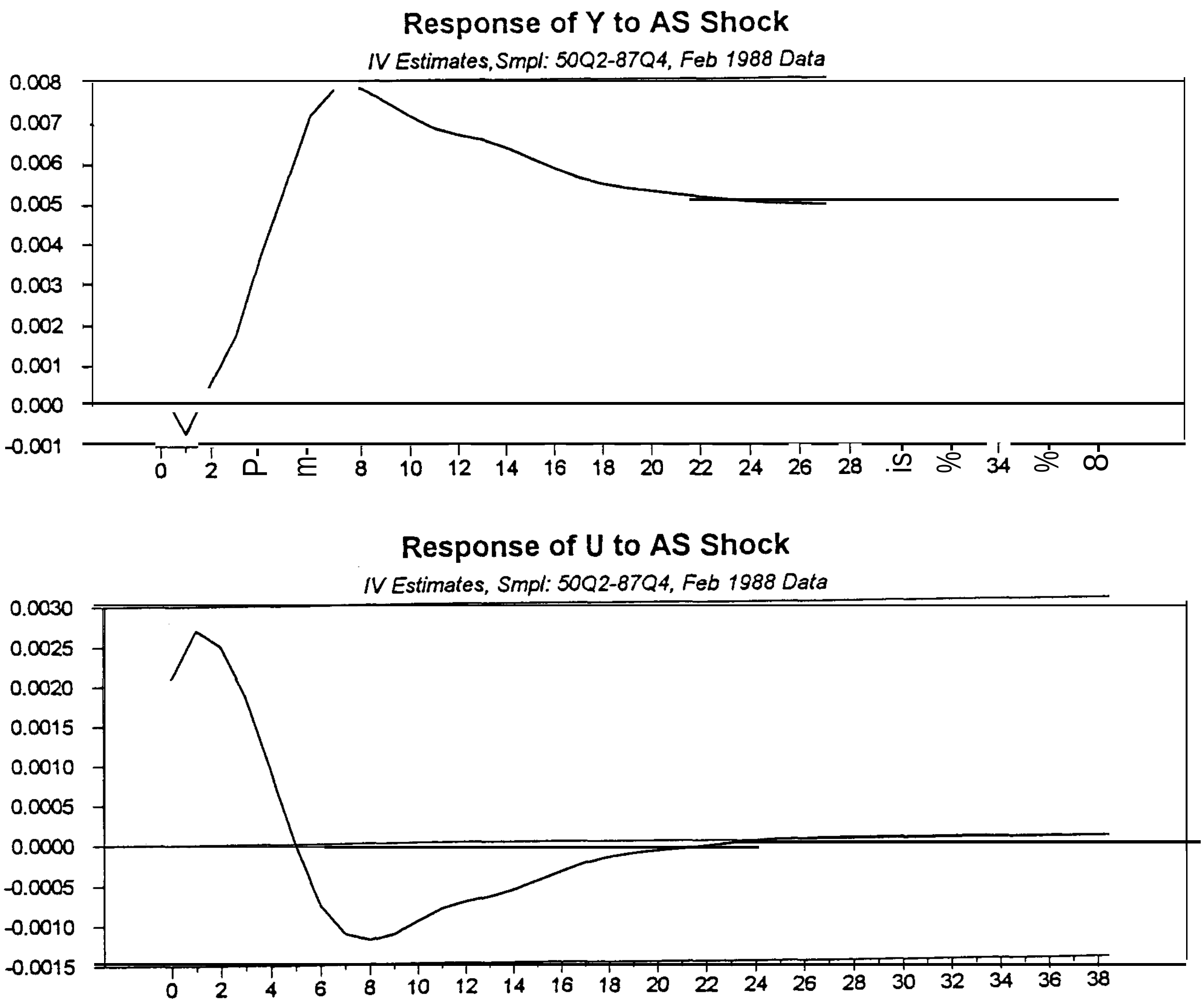


Response of Y to AD Shock
IV Estimates, Smpl: 50Q2-87Q4, Feb 1988 Data
0.0030
0.0025
0.0020
0.0015
0.0010
0.0005
0.0000
-0.0005
0 2 4 6 8 10 12 14 16 18 20 22 24 26 28 30 32 34 36 38



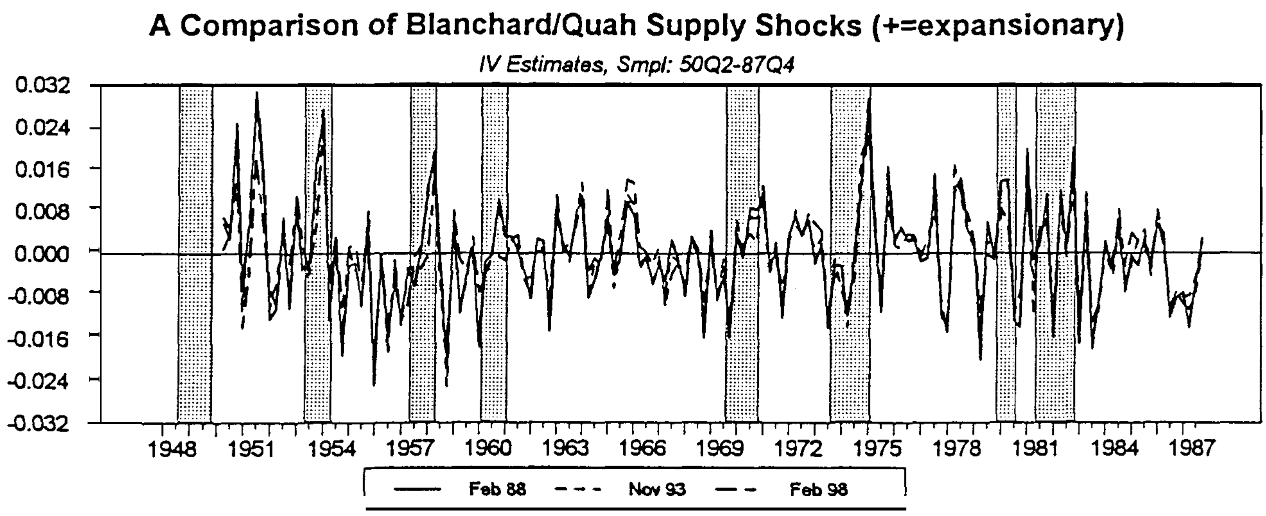
Response of U to AD Shock
IV Estimates, Smpl:50Q2-87Q4, Feb 1988 Data
0.00025
0.00000
-0.00025
-0.00050
-0.00075
-0.00100
-0.00125
8 10 12 14 16 18 20 22 24 26 28 30 32 34 36 38

Figure V.8



Response of Y to AS Shock
IV Estimates,Smpl: 50Q2-87Q4, Feb 1988 Data
0.008
0.007
0.006
0.005
0.004
0.003
0.002
0.001
0.000
-0.001
0
2
8
10
12
14
16
18
20
22
24
26
28
34
Response of U to AS Shock
IV Estimates, Smpl: 50Q2-87Q4, Feb 1988 Data
0.0030
0.0025
0.0020
0.0015
0.0010
0.0005
0.0000
-0.0005
-0.0010
-0.0015
0
2
4
6
8
10
12
14
16
18
20
22
24
26
28
30
32
34
36
38

Figure V.9


A Comparison of Blanchard/Quah Supply Shocks (+=expansionary)
IV Estimates, Smpl: 50Q2-87Q4
0.032
0.024
0.016
0.008
0.000
-0.008
-0.016
-0.024
-0.032
1948
1951
1954
1957
1960
1963
1966
1969
1972
1975
1978
1981
1984
1987
Feb 88
Nov 93
Feb 98

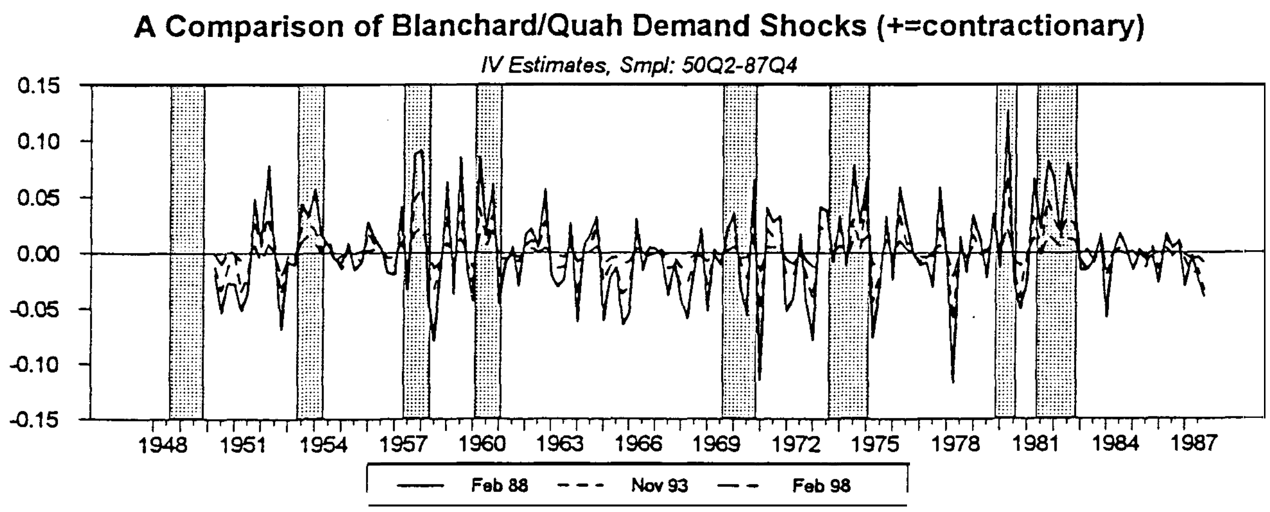

A Comparison of Blanchard/Quah Demand Shocks (+=contractionary)
IV Estimates, Smpl: 50Q2-87Q4
0.15
0.10
0.05
0.00
-0.05
-0.10
-0.15
1948
1951
1954
1957
1960
1963
1966
1969
1972
1975
1978
1981
1984
1987
Feb 88
Nov 93
Feb 98

Figure V.10

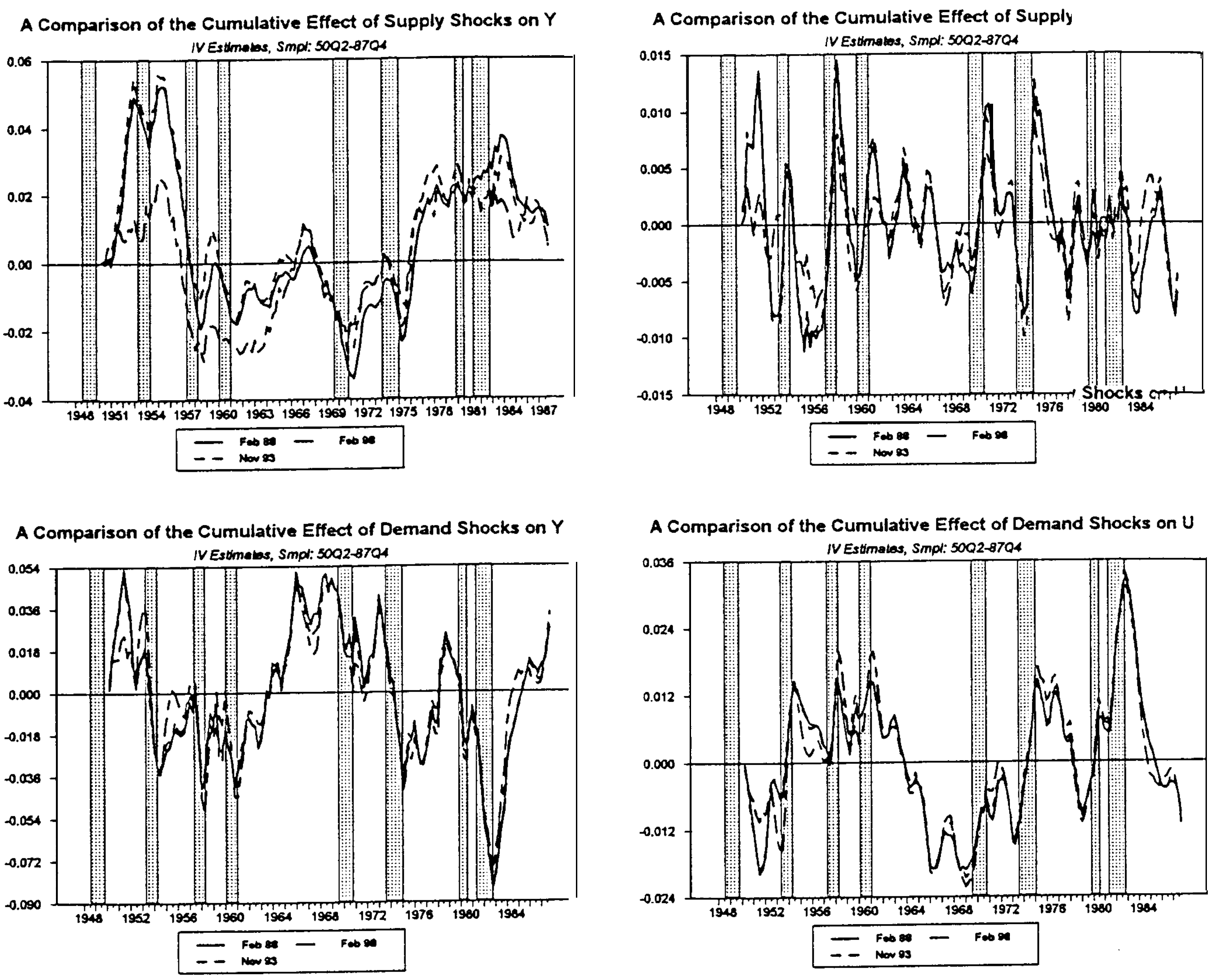
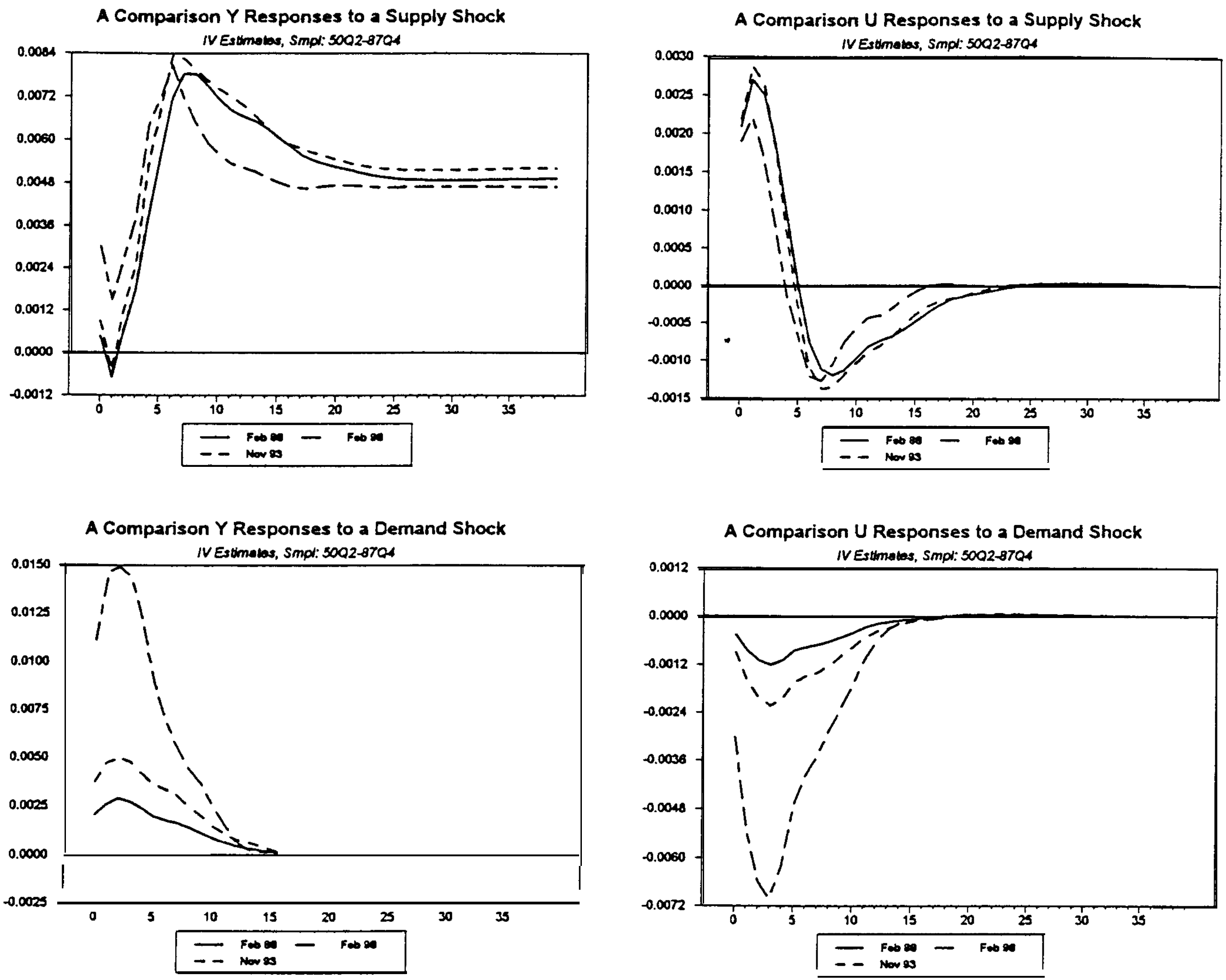

A Comparison of the Cumulative Effect of Supply Shocks on Y
IV Estimates, Smpl: 50Q2-87Q4
Feb 88
Feb 98
Nov 93
A Comparison of the Cumulative Effect of Supply
IV Estimates, Smpl: 50Q2-87Q4
Shocks
Feb 88
Feb 98
Nov 93
A Comparison of the Cumulative Effect of Demand Shocks on Y
IV Estimates, Smpl: 50Q2-87Q4
Feb 88
Feb 98
Nov 93
A Comparison of the Cumulative Effect of Demand Shocks on U
IV Estimates, Smpl: 50Q2-87Q4
Feb 88
Feb 98
Nov 93

Figure V.11

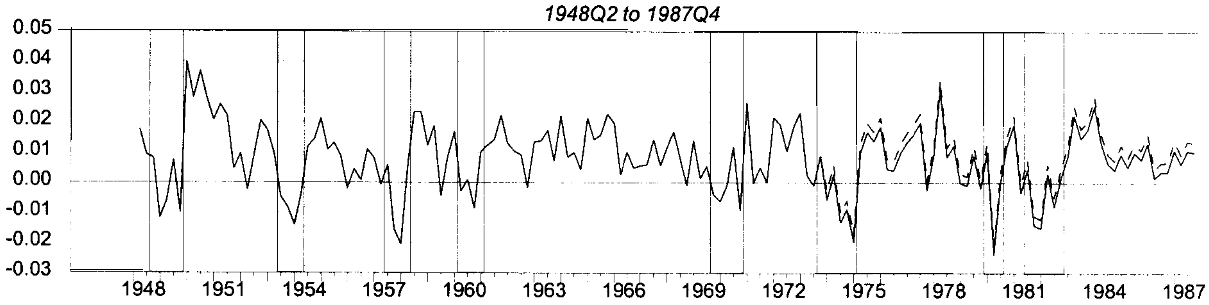


A Comparison Y Responses to a Supply Shock
IV Estimates, Smpl: 50Q2-87Q4
0.0084
0.0072
0.0060
0.0048
0.0036
0.0024
0.0012
0.0000
-0.0012
0 5 10 15 20 25 30 35
Nov 93
A Comparison U Responses to a Supply Shock
IV Estimates, Smpl: 50Q2-87Q4
0.0030
0.0025
0.0020
0.0015
0.0010
0.0005
0.0000
-0.0005
-0.0010
-0.0015
0 5 10 15 20 25 30 35
Nov 93
A Comparison Y Responses to a Demand Shock
IV Estimates, Smpl: 50Q2-87Q4
0.0150
0.0125
0.0100
0.0075
0.0050
0.0025
0.0000
-0.0025
0 5 10 15 20 25 30 35
Nov 93
A Comparison U Responses to a Demand Shock
IV Estimates, Smpl: 50Q2-87Q4
0.0012
0.0000
-0.0012
-0.0024
-0.0036
-0.0048
-0.0060
-0.0072
0 5 10 15 20 25 30 35
Nov 93

Figure V.12

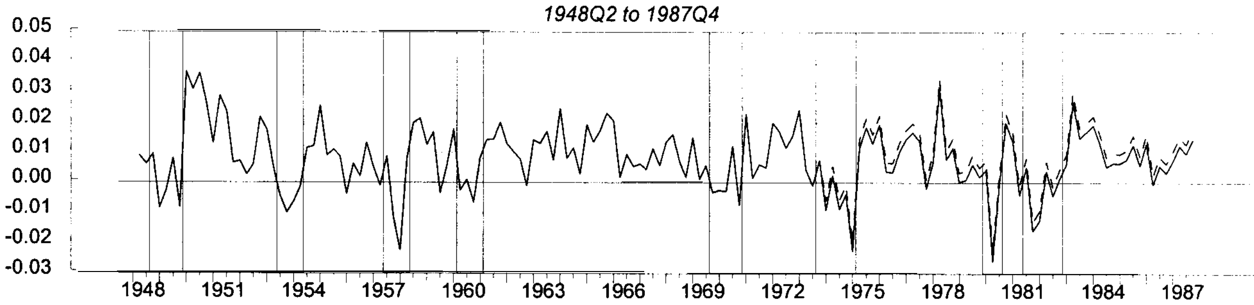
Actual & Trend Break Adjusted Real Output Growth: Feb 1988 Data

1948Q2 to 1987Q4



Actual & Trend Break Adjusted Real Output Growth: Nov 1993 Data

1948Q2 to 1987Q4



Actual & Trend Break Adjusted Real Output Growth: Feb 1998 Data

1948Q2 to 1987Q4

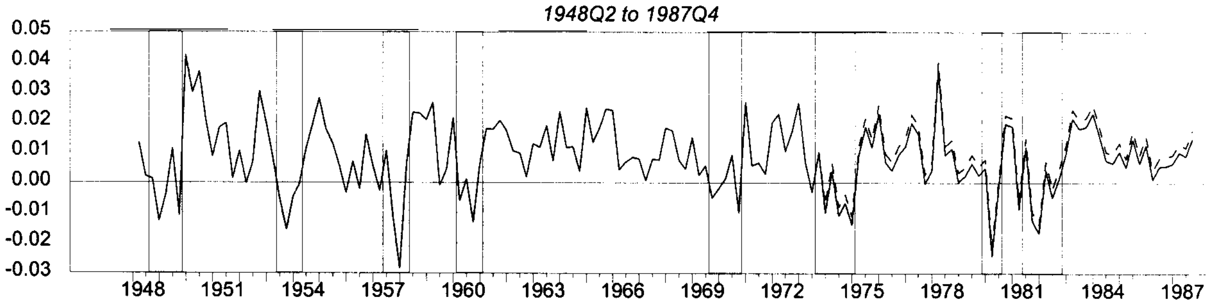
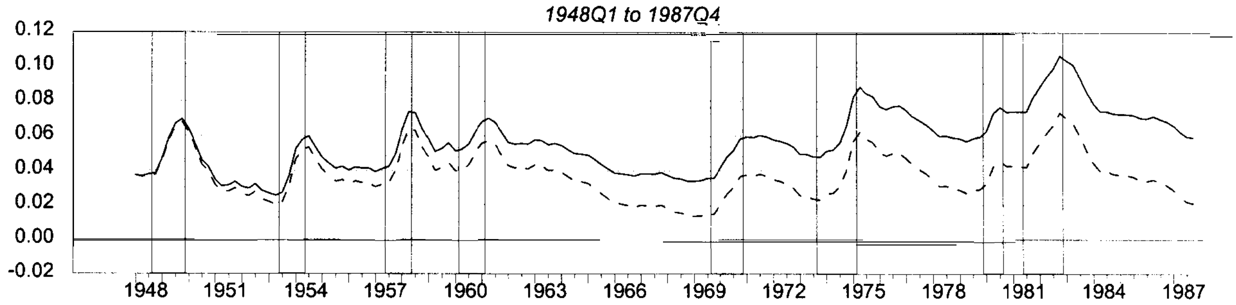


Figure V.13

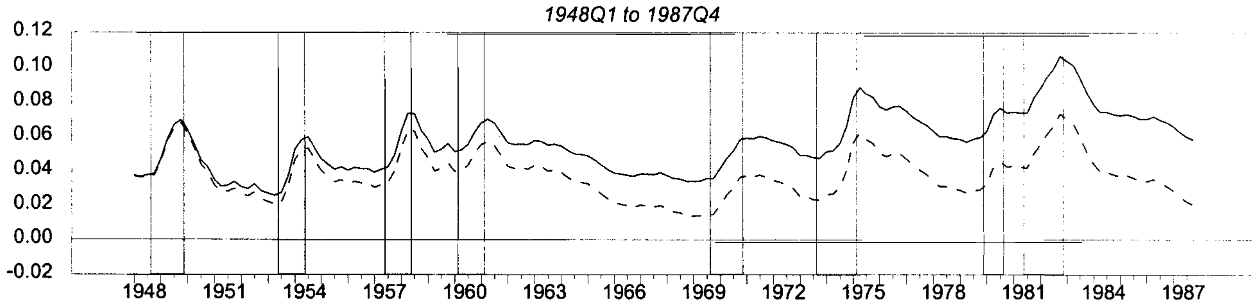
Actual & Detrended Unemployment Rate: Feb 1988 Data

1948Q1 to 1987Q4



Actual & Detrended Unemployment Rate: Nov 1993 Data

1948Q1 to 1987Q4



Actual & Detrended Unemployment Rate: Feb 1998 Data

1948Q1 to 1987Q4

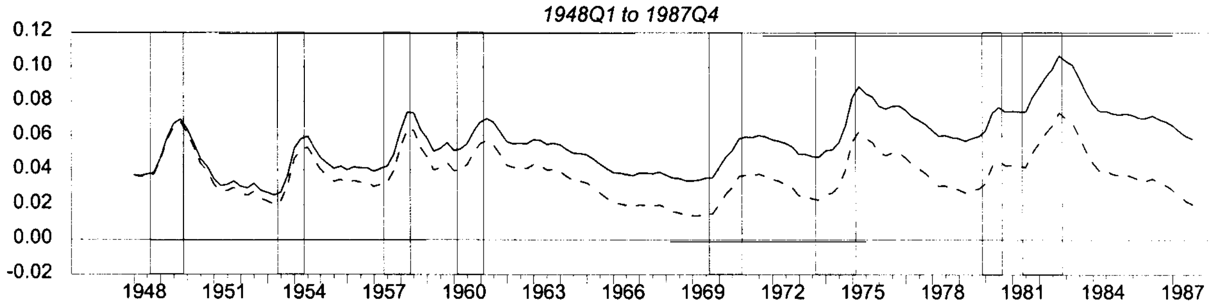


Figure VI.1

Actual Inflation and Inflation Target
Based On November 1998 Vintage Data

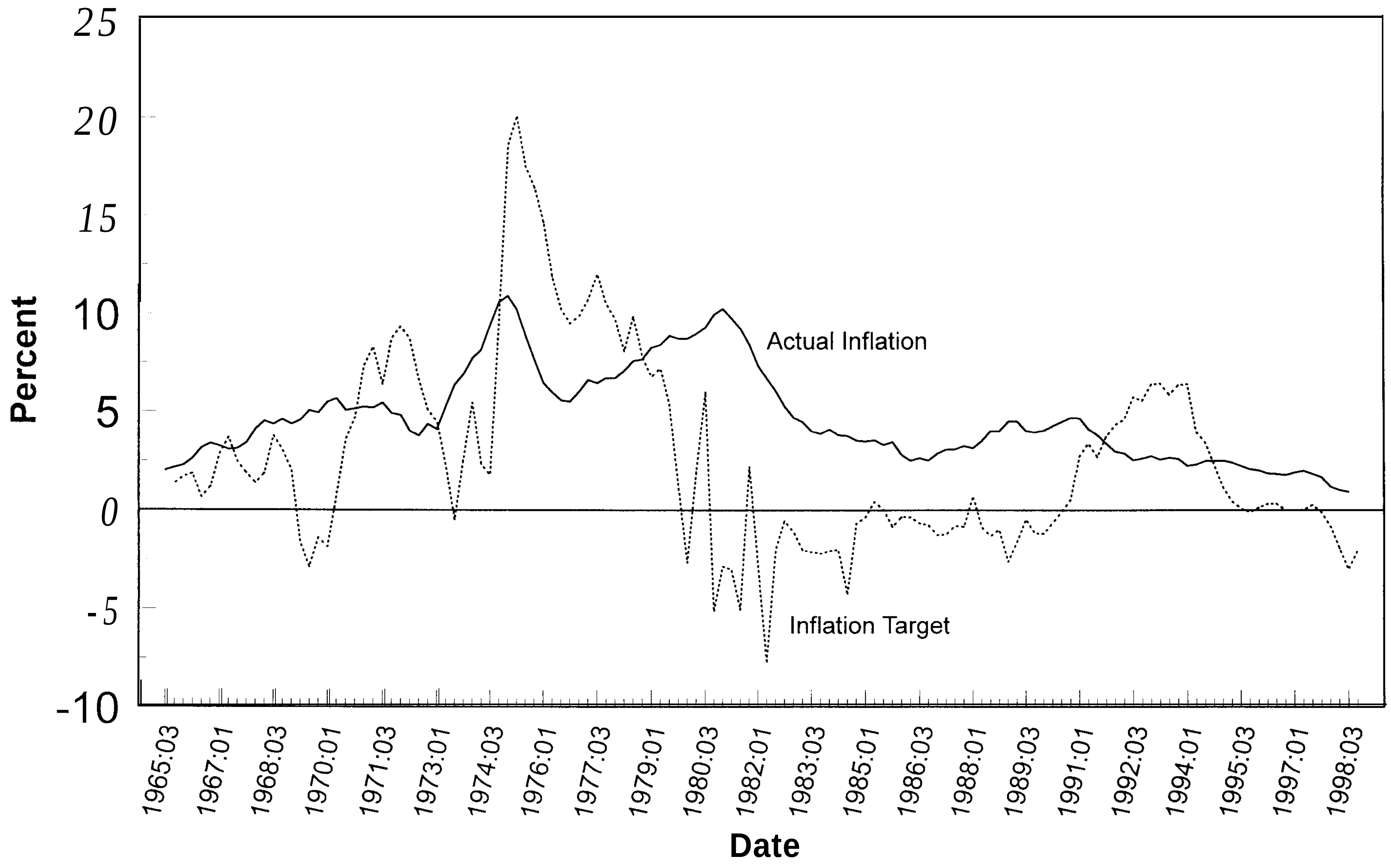


Figure VI.2

Actual Inflation and Inflation Target Based On Real-Time Vintage Data

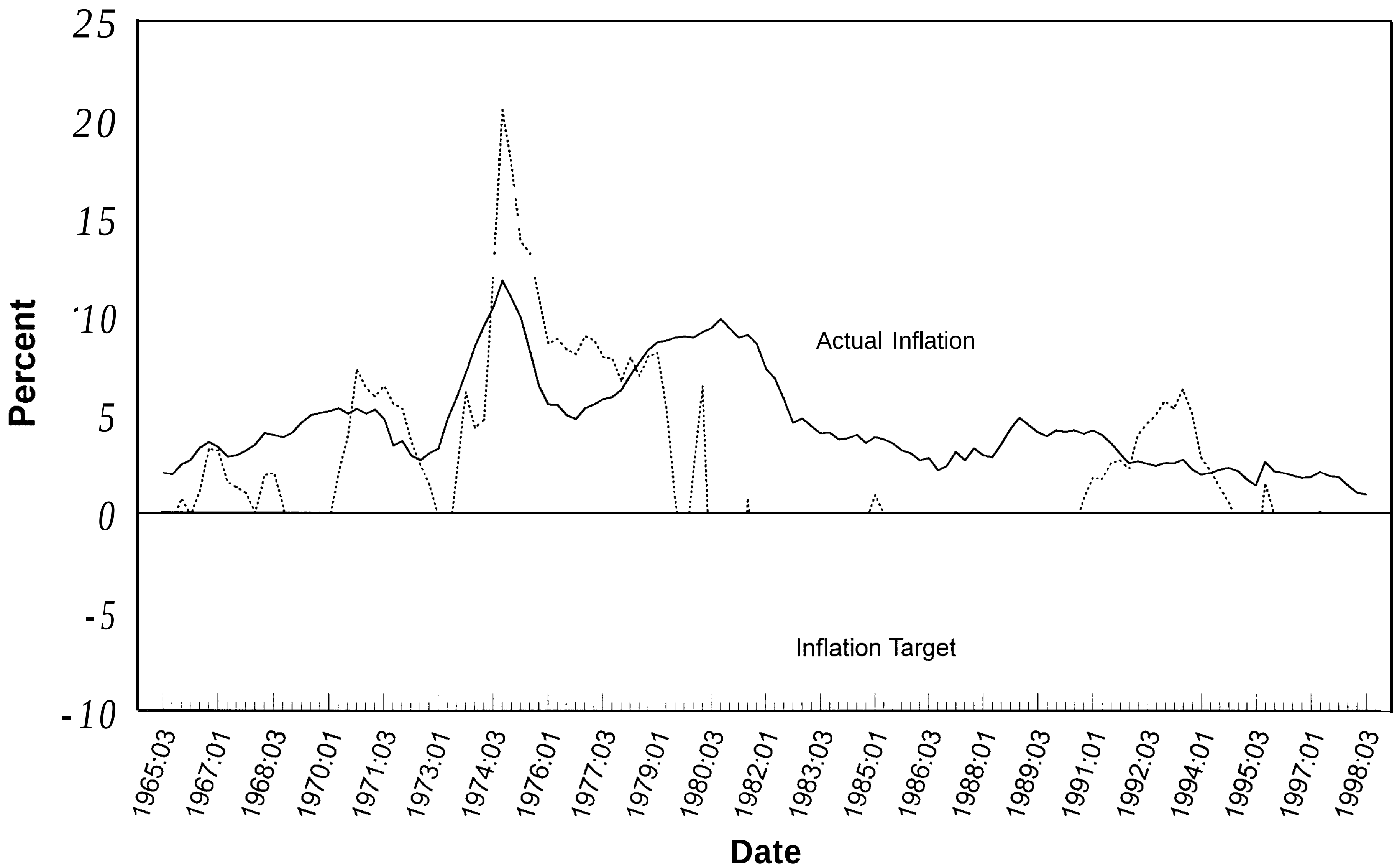


Figure VI.3

Inflation Minus Target Versus Boschen-Mills Index

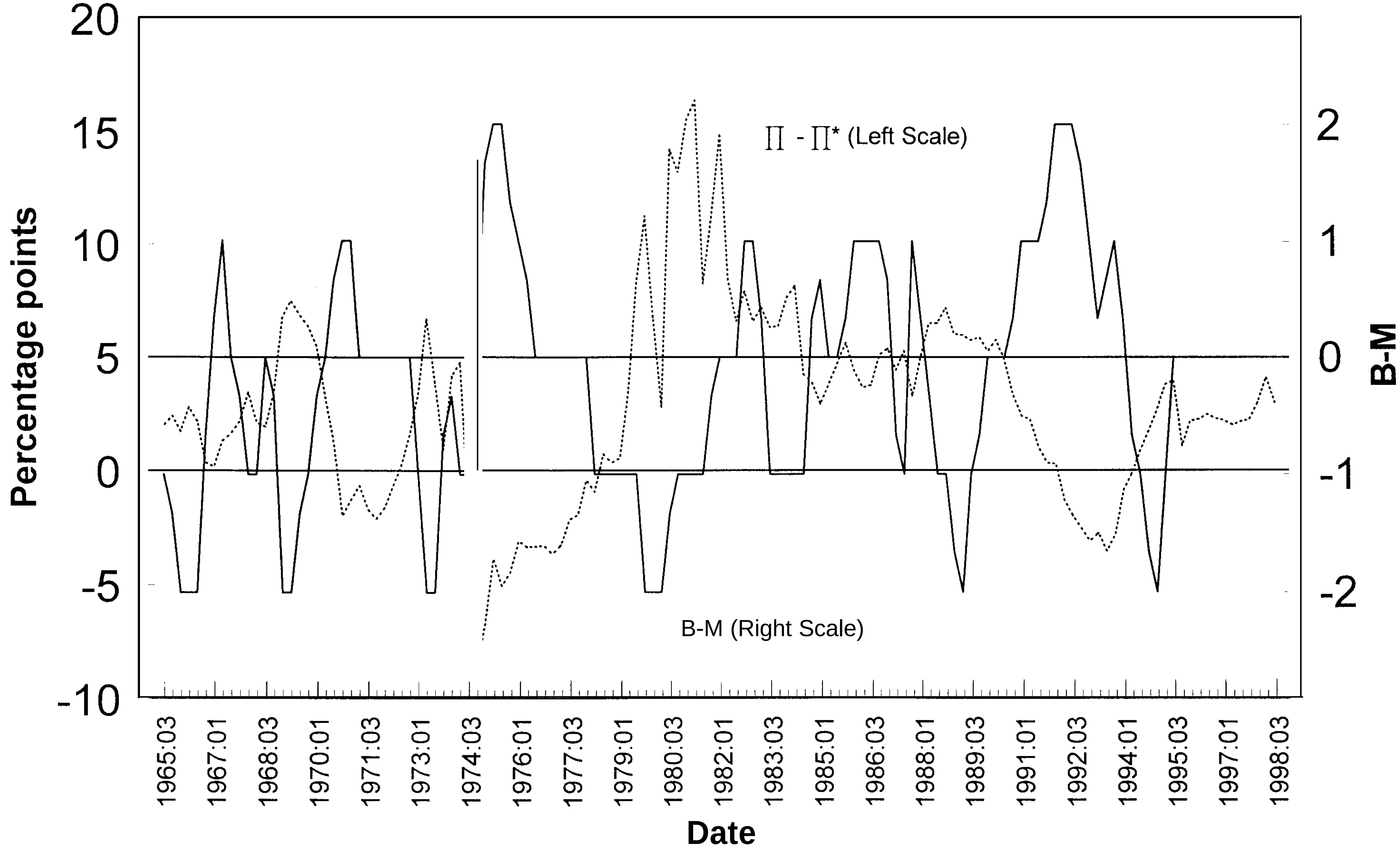


Figure VI.4

Inflation Target
Real-Time Versus November 1998

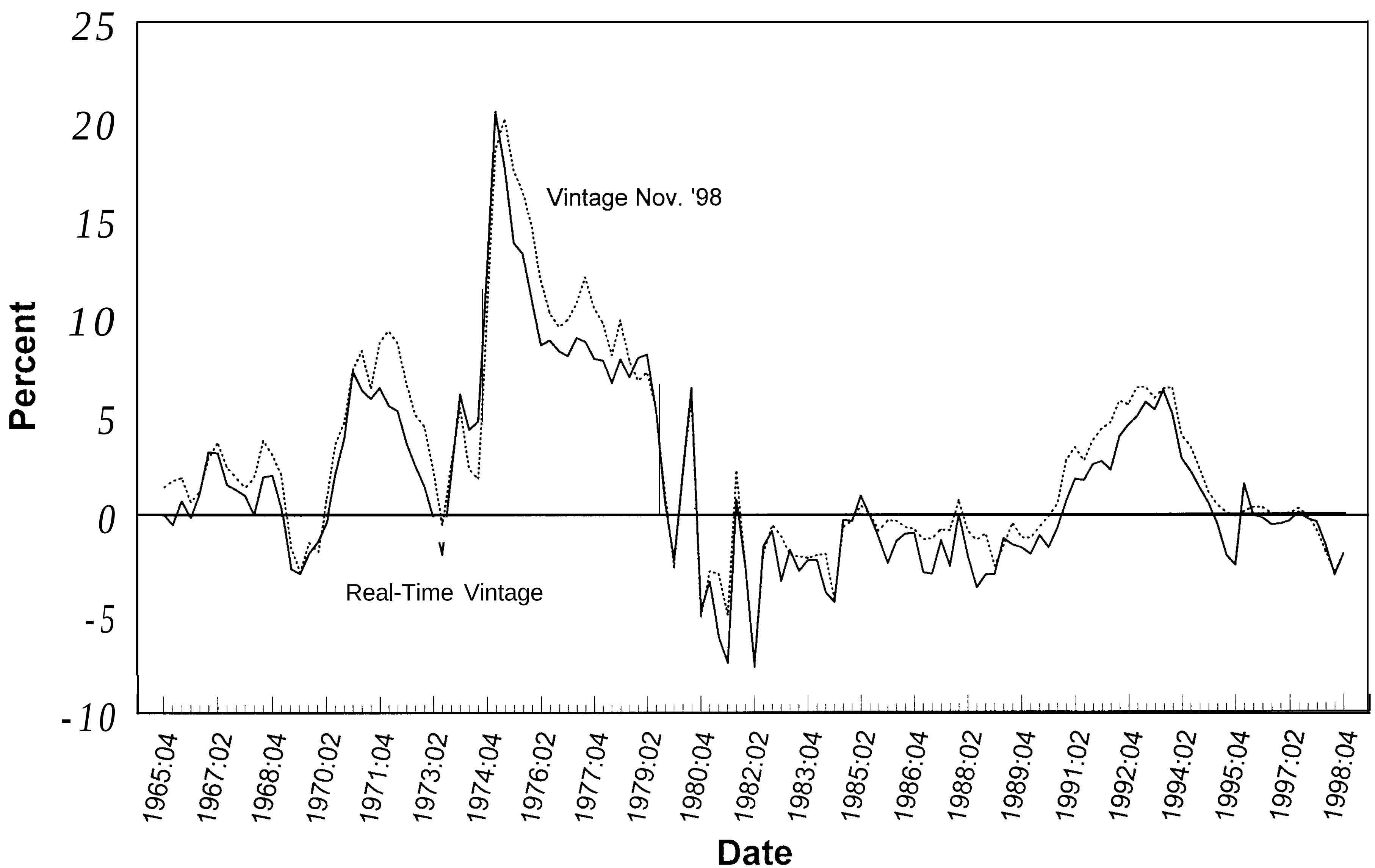


Figure VI.5

Estimated Equilibrium Real Federal Funds Rate

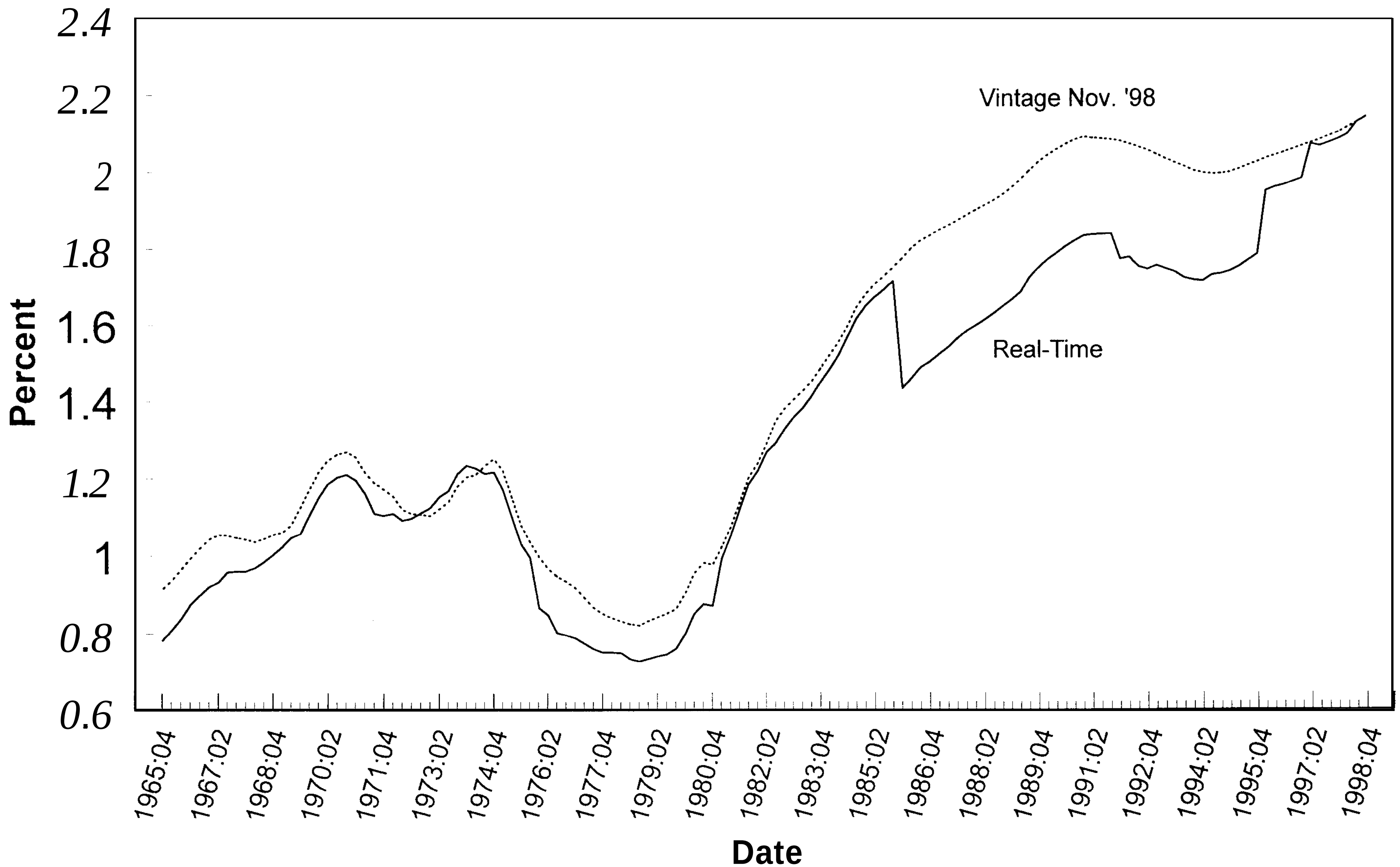


Figure VI.6

Actual Fed Funds Rate Compared to Fed Funds Target Based On November 1998 Data

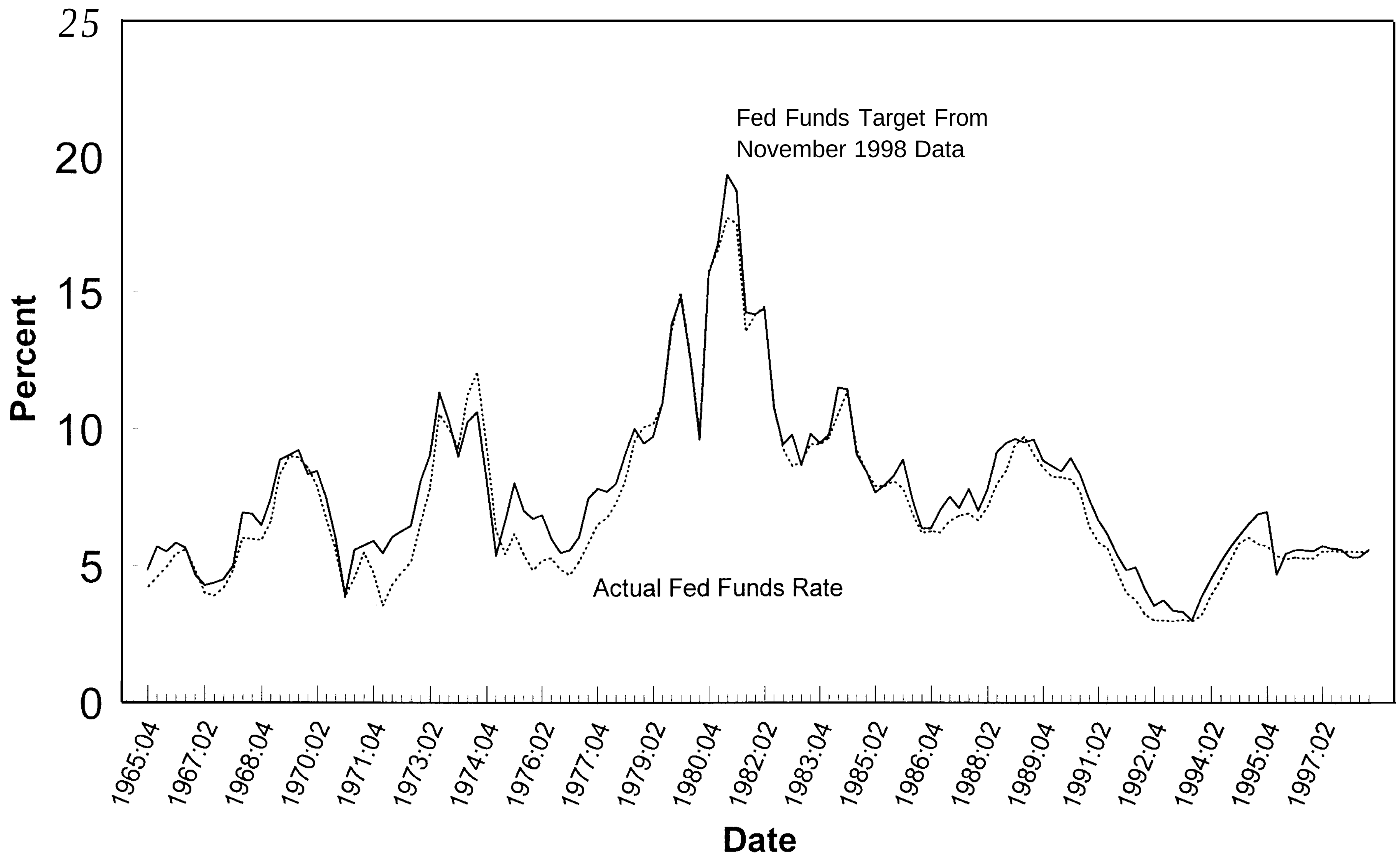


Figure VI.7

Fed Funds Target Based On November 1998 Data Minus Actual Fed Funds Rate

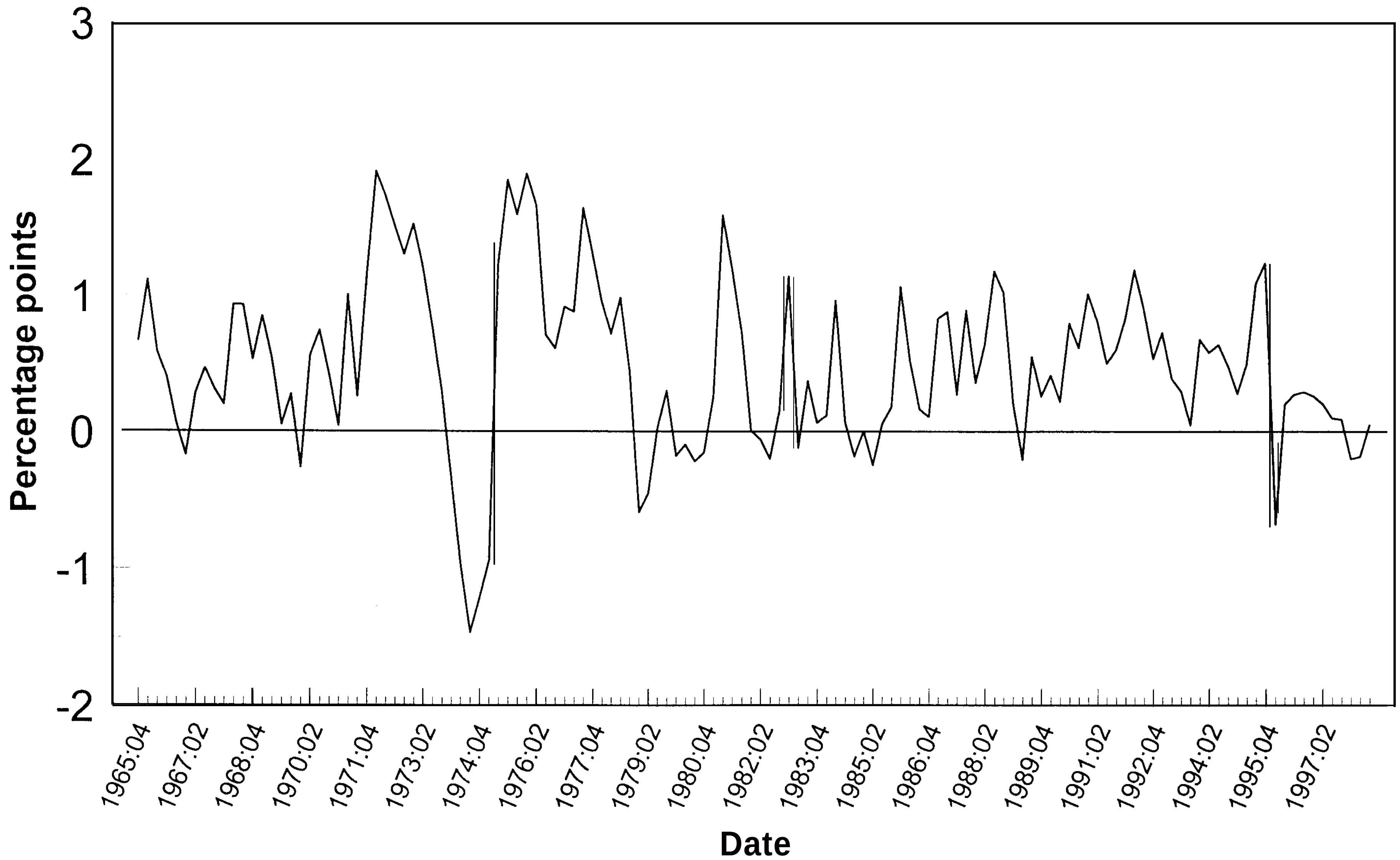


Figure VII.1

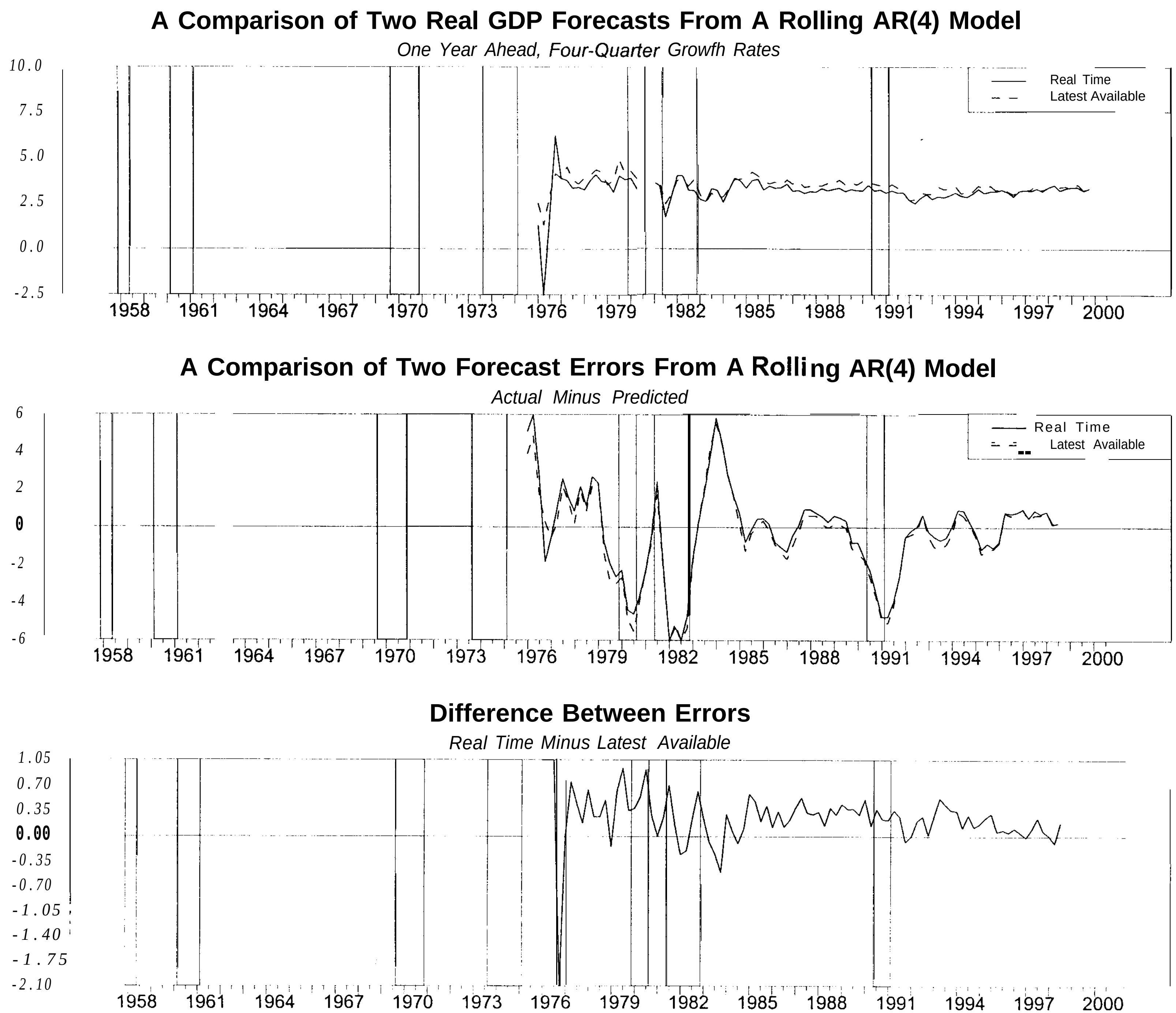
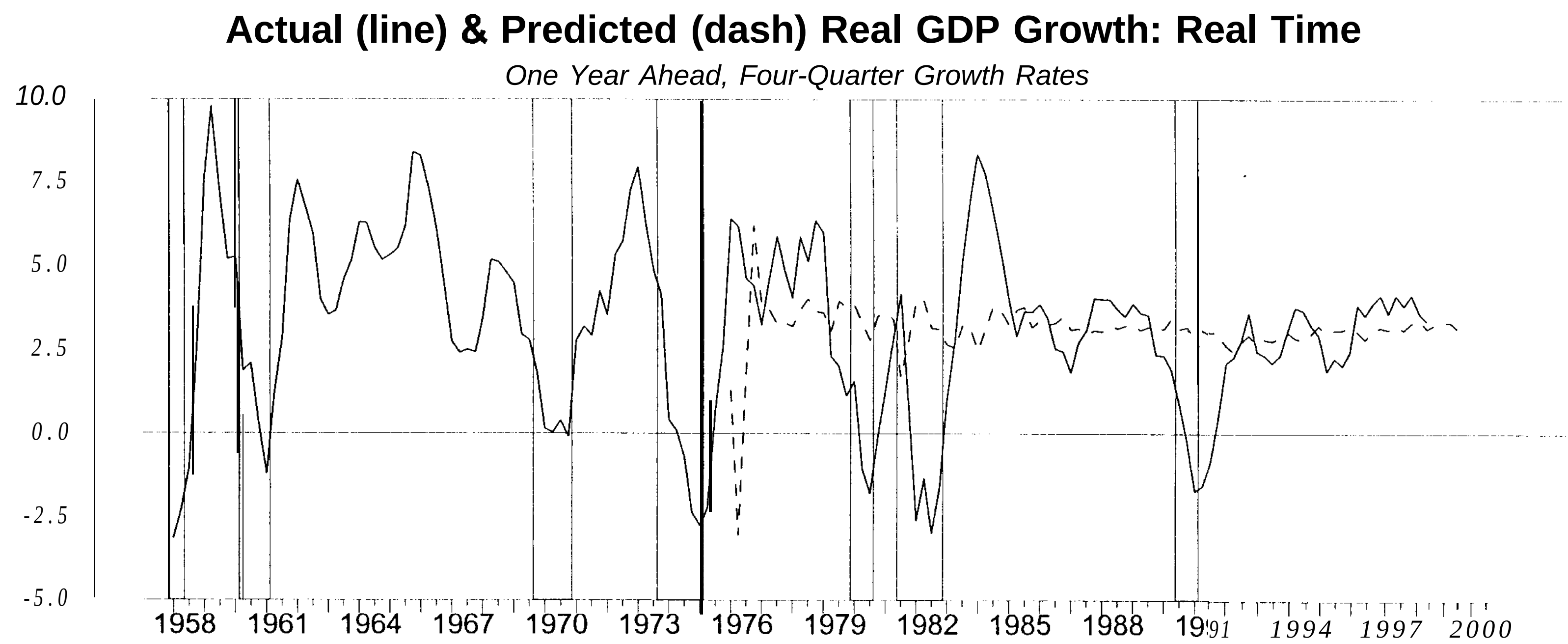

A Comparison of Two Real GDP Forecasts From A Rolling AR(4) Model
One Year Ahead, Four-Quarter Growfh Rates
Real Time
Latest Available
10.0
7.5
5.0
2.5
0.0
-2.5
1958
1961
1964
1967
1970
1973
1976
1979
1982
1985
1988
1991
1994
1997
2000
A Comparison of Two Forecast Errors From A Rolling AR(4) Model
Actual Minus Predicted
Real Time
Latest Available
6
4
2
0
-2
-4
-6
1958
1961
1964
1967
1970
1973
1976
1979
1982
1985
1988
1991
1994
1997
2000
Difference Between Errors
Real Time Minus Latest Available
1.05
0.70
0.35
0.00
-0.35
-0.70
-1.05
-1.40
-1.75
-2.10
1958
1961
1964
1967
1970
1973
1976
1979
1982
1985
1988
1991
1994
1997
2000

Figure VII.2


Actual (line) & Predicted (dash) Real GDP Growth: Real Time
One Year Ahead, Four-Quarter Growth Rates
10.0
7.5
5.0
2.5
0.0
-2.5
-5.0
1958
1961
1964
1967
1970
1973
1976
1979
1982
1985
1988
1991
1994
1997
2000

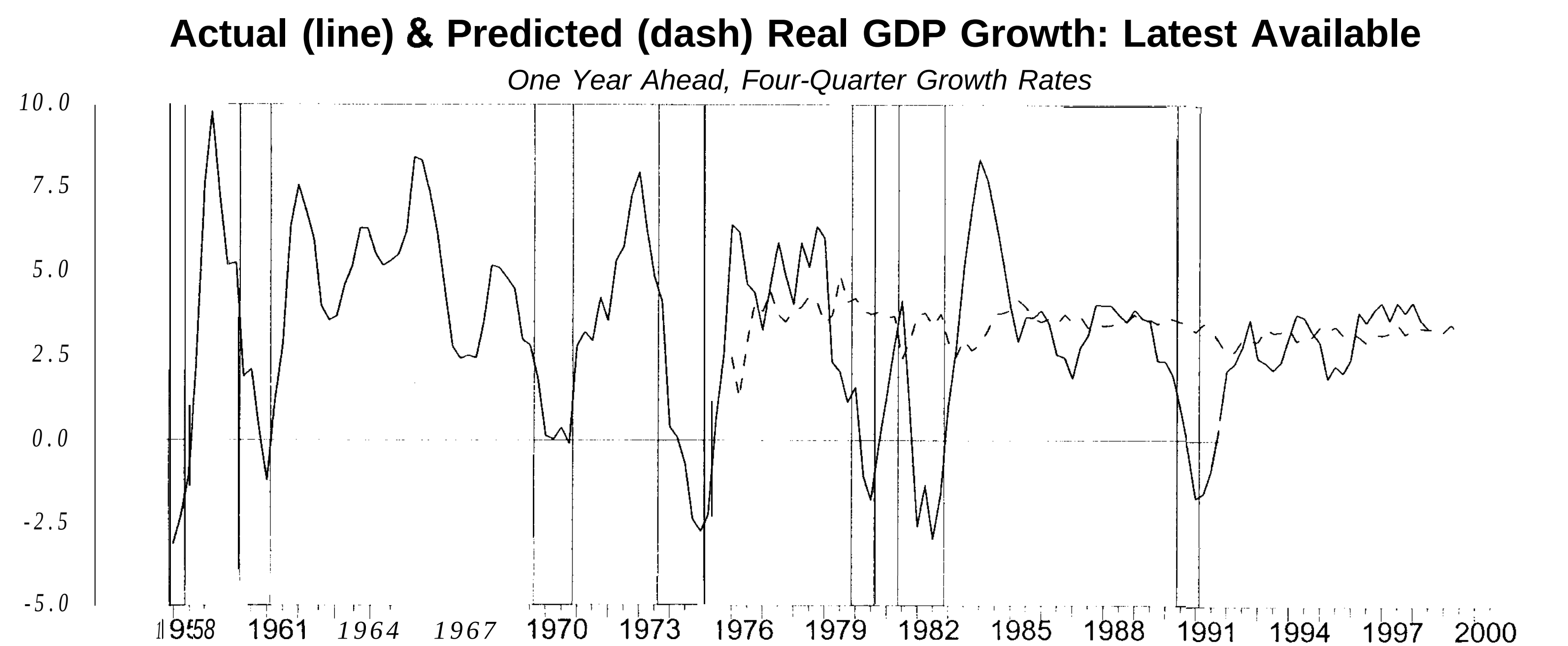

Actual (line) & Predicted (dash) Real GDP Growth: Latest Available
One Year Ahead, Four-Quarter Growth Rates
10.0
7.5
5.0
2.5
0.0
-2.5
-5.0
1958
1961
1964
1967
1970
1973
1976
1979
1982
1985
1988
1991
1994
1997
2000

Figure VII.3

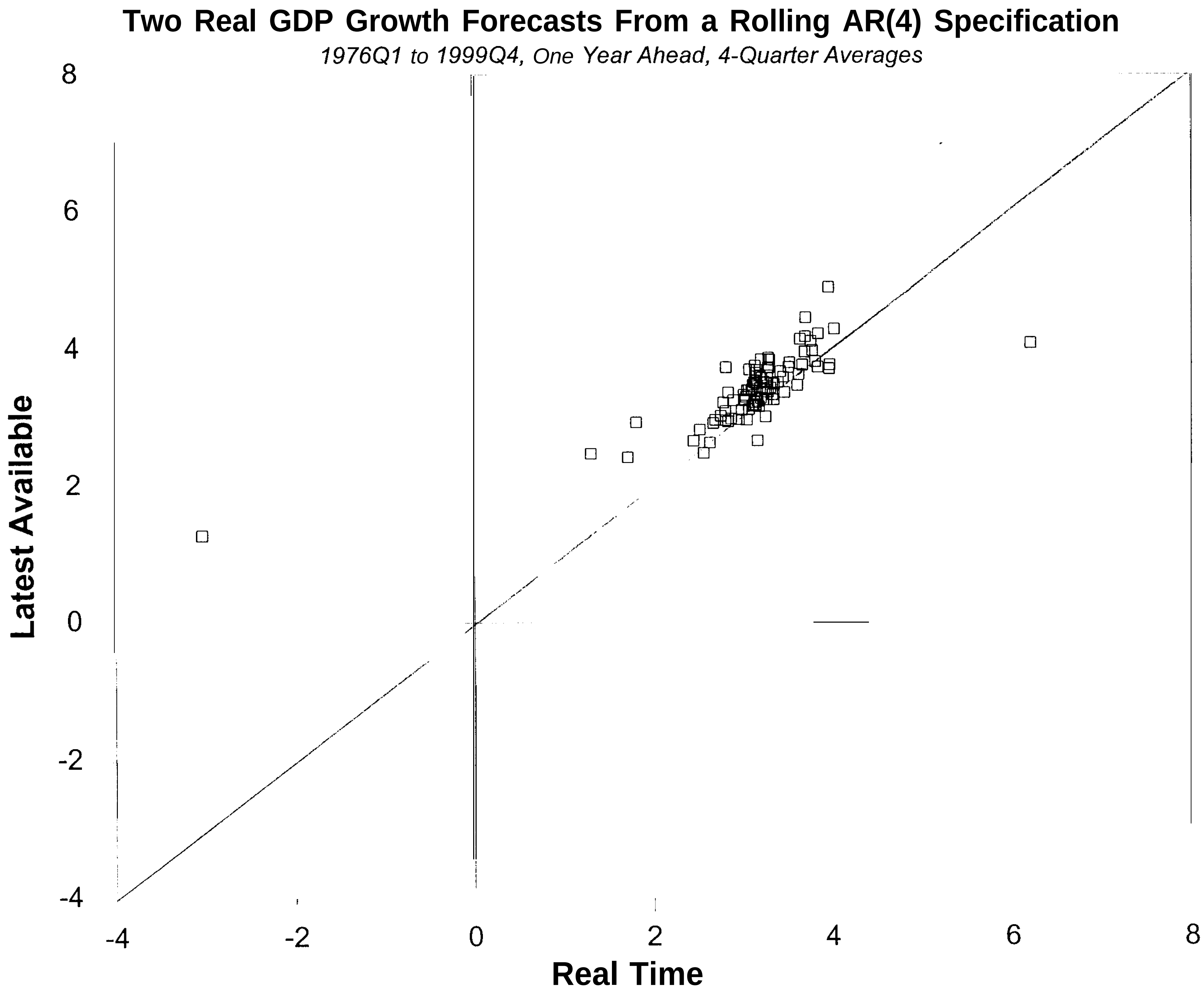

Two Real GDP Growth Forecasts From a Rolling AR(4) Specification
1976Q1 to 1999Q4, One Year Ahead, 4-Quarter Averages
Latest Available
Real Time
8
6
4
2
0
-2
-4
-4
-2
0
2
4
6
8

Figure VII.4

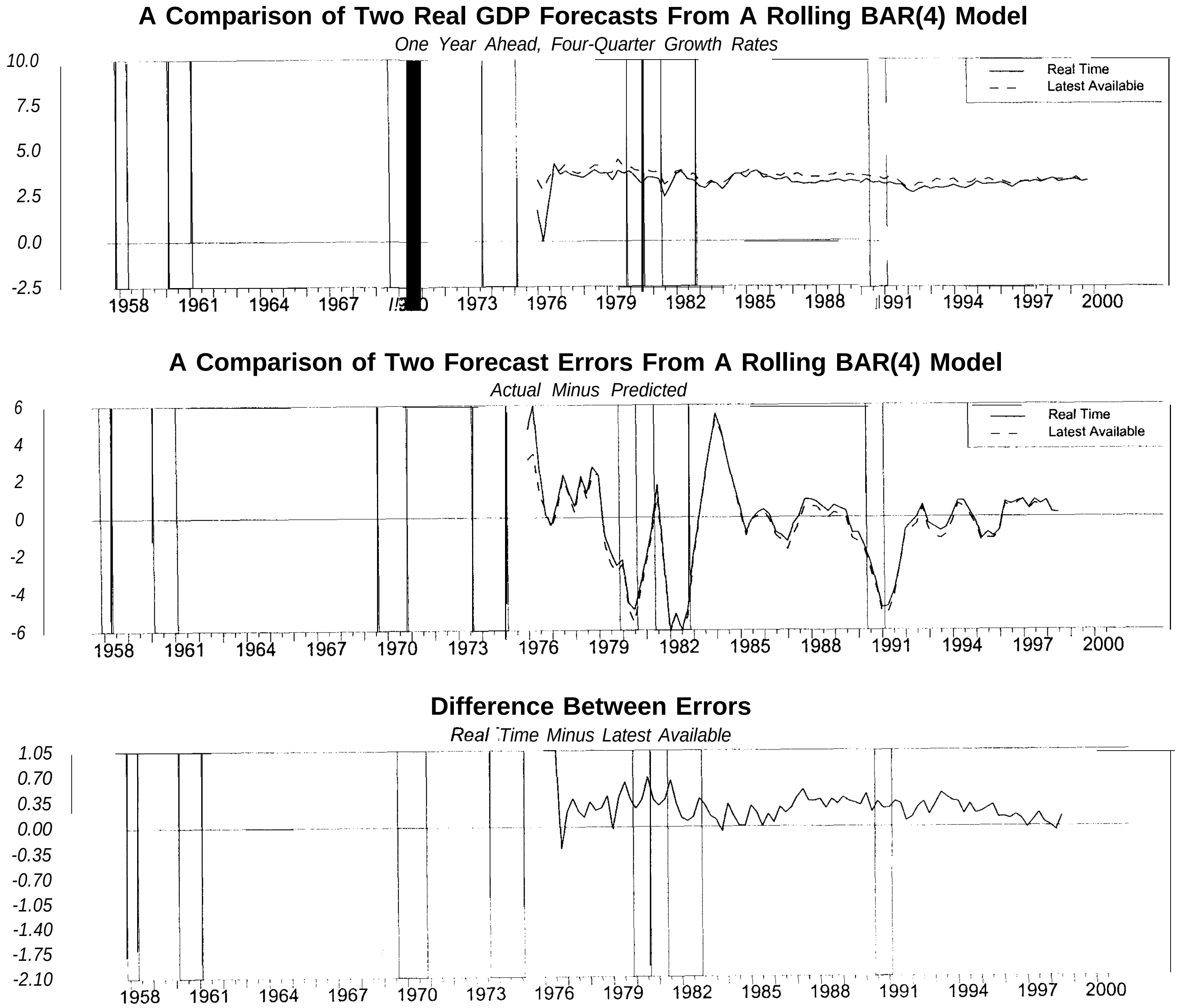
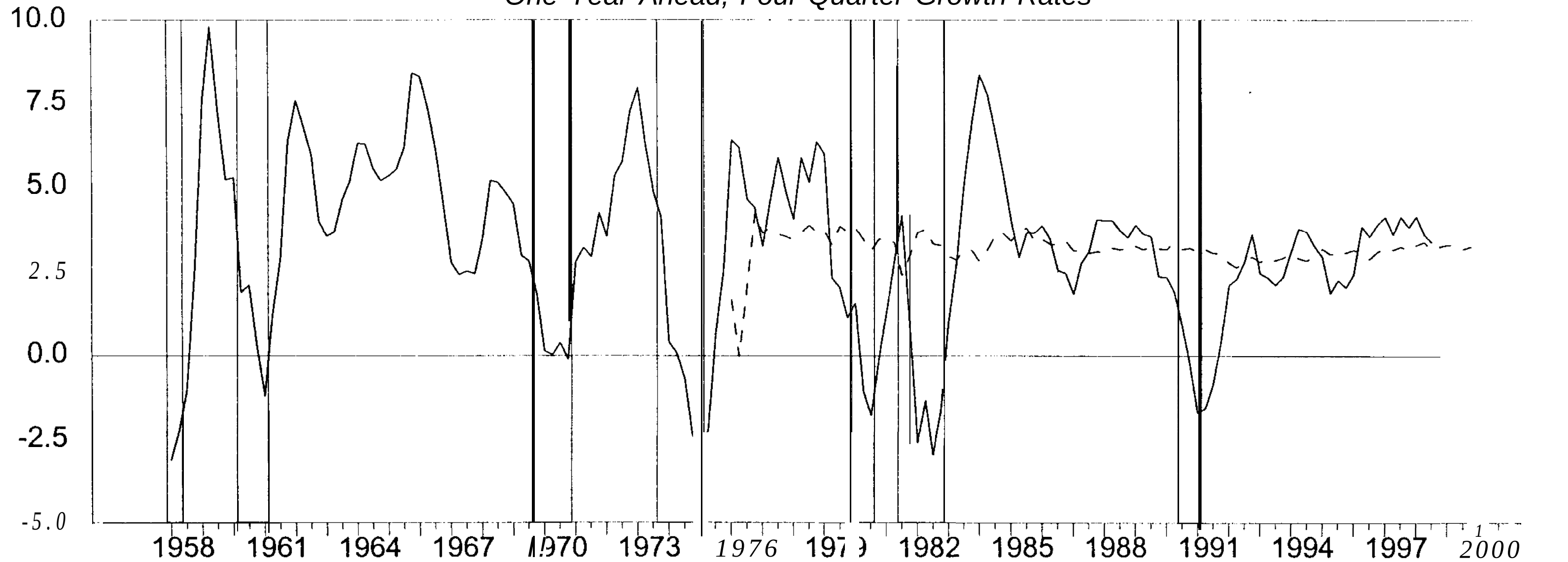

A Comparison of Two Real GDP Forecasts From A Rolling BAR(4) Model
One Year Ahead, Four-Quarter Growth Rates
10.0
7.5
5.0
2.5
0.0
-2.5
1958 1961 1964 1967 1973 1976 1979 1982 1985 1988 1991 1994 1997 2000
Real Time
Latest Available
A Comparison of Two Forecast Errors From A Rolling BAR(4) Model
Actual Minus Predicted
6
4
2
0
-2
-4
-6
1958 1961 1964 1967 1970 1973 1976 1979 1982 1985 1988 1991 1994 1997 2000
Real Time
Latest Available
Difference Between Errors
Real Time Minus Latest Available
1.05
0.70
0.35
0.00
-0.35
-0.70
-1.05
-1.40
-1.75
-2.10
1958 1961 1964 1967 1970 1973 1976 1979 1982 1985 1988 1991 1994 1997 2000

Figure VII.5

Actual (line) & Predicted (dash) Real GDP Growth: Real Time

One Year Ahead, Four-Quarter Growth Rates



Actual (line) & Predicted (dash) Real GDP Growth: Latest Available

One Year Ahead, Four-Quarter Growth Rates

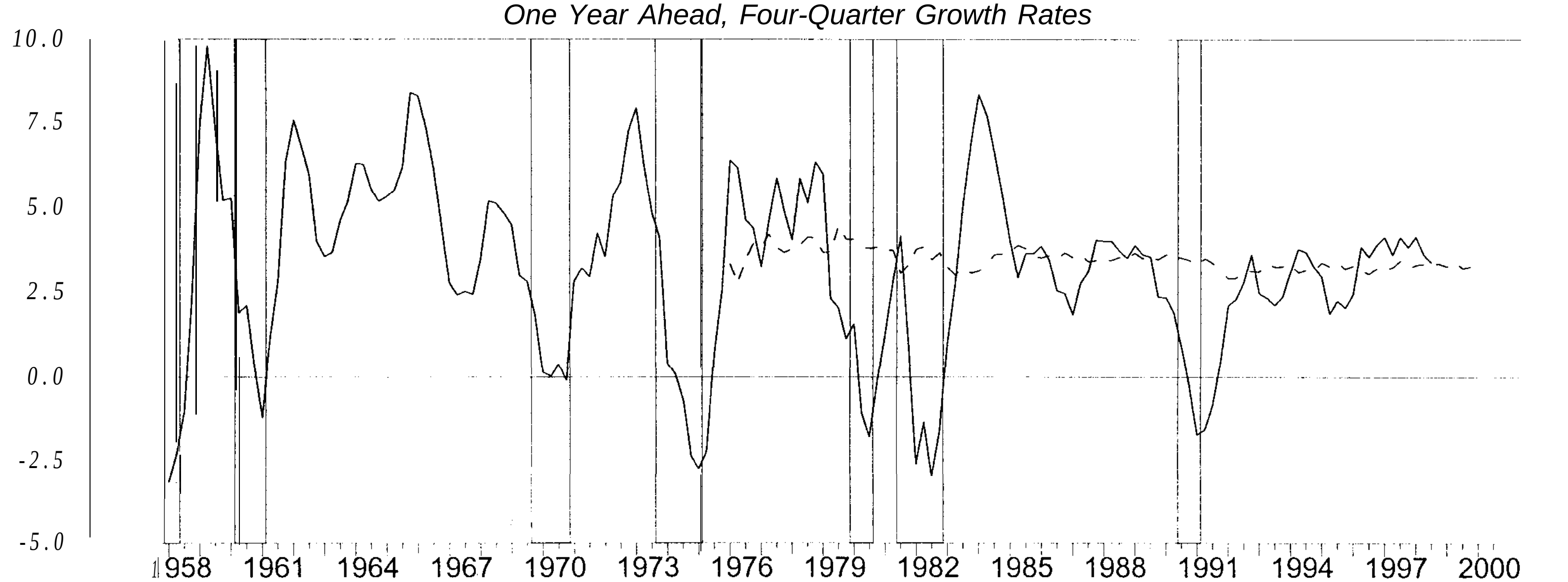


Figure VII.6

Two Real GDP Growth Forecasts From a Rolling BAR(4) Specification

197601 to 199904, One Year Ahead, 4-Quarter Averages

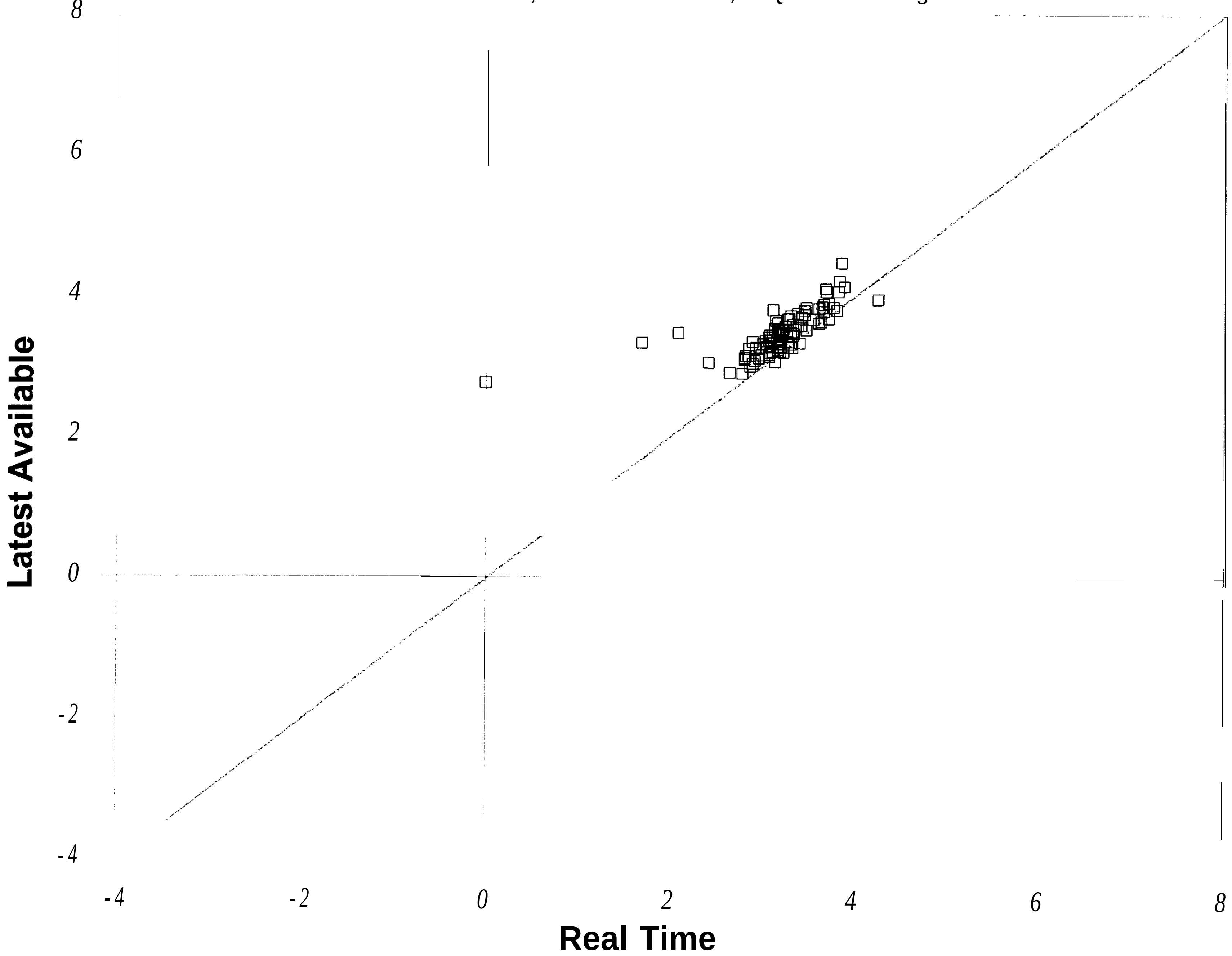


Figure VII.7

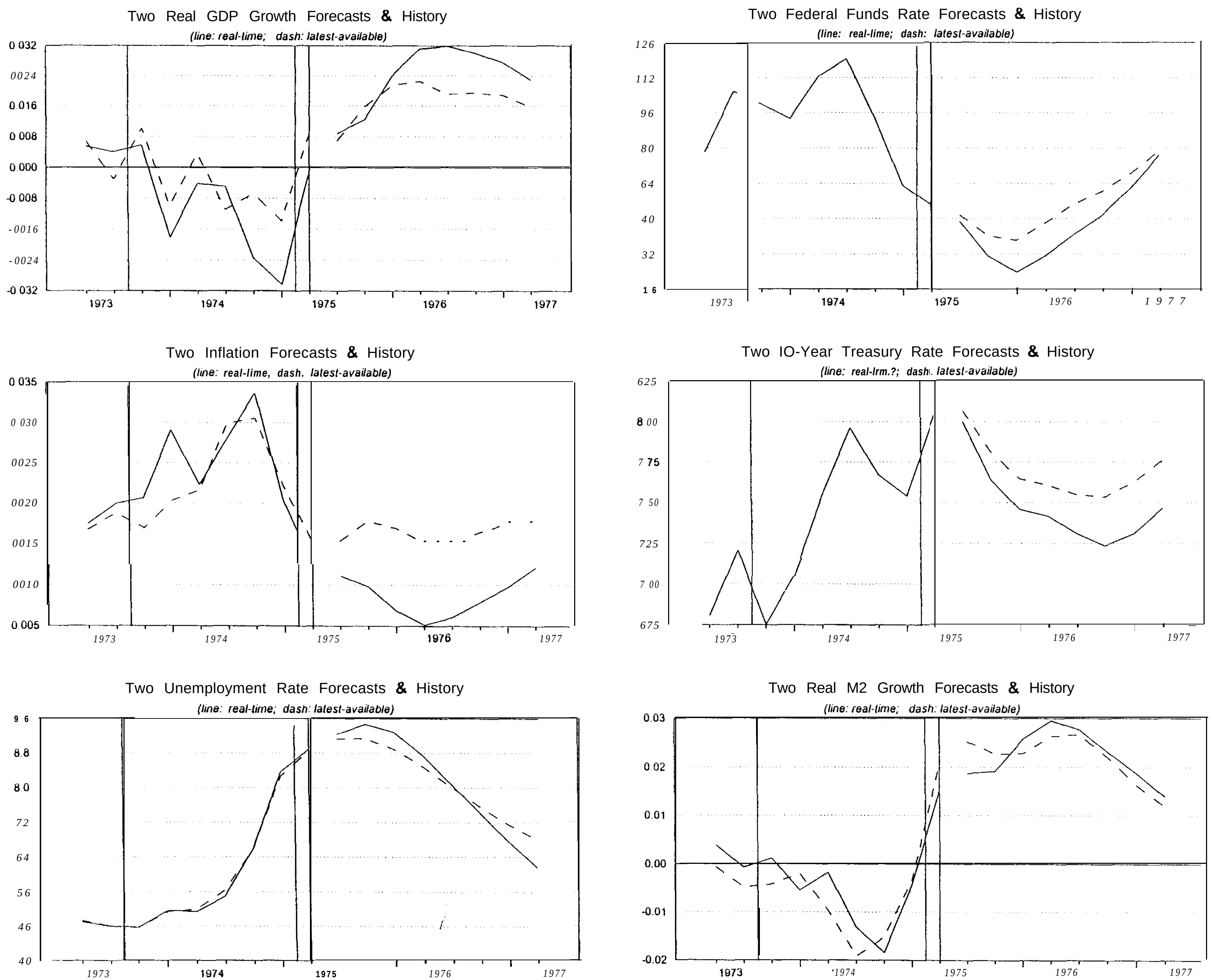

Two Real GDP Growth Forecasts & History
(line: real-time; dash: latest-available)
0.032
0.024
0.016
0.008
0.000
-0.008
-0.016
-0.024
-0.032
1973
1974
1975
1976
1977
Two Federal Funds Rate Forecasts & History
(line: real-time; dash: latest-available)
126
112
96
80
64
40
32
16
Two Inflation Forecasts & History
(line: real-time, dash: latest-available)
0.035
0.030
0.025
0.020
0.015
0.010
0.005
Two IO-Year Treasury Rate Forecasts & History
(line: real-time; dash: latest-available)
625
8.00
7.75
7.50
725
7.00
675
Two Unemployment Rate Forecasts & History
(line: real-time; dash: latest-available)
96
8.8
8.0
72
64
56
46
40
Two Real M2 Growth Forecasts & History
(line: real-time; dash: latest-available)
0.03
0.02
0.01
0.00
-0.01
-0.02

Figure VII.8

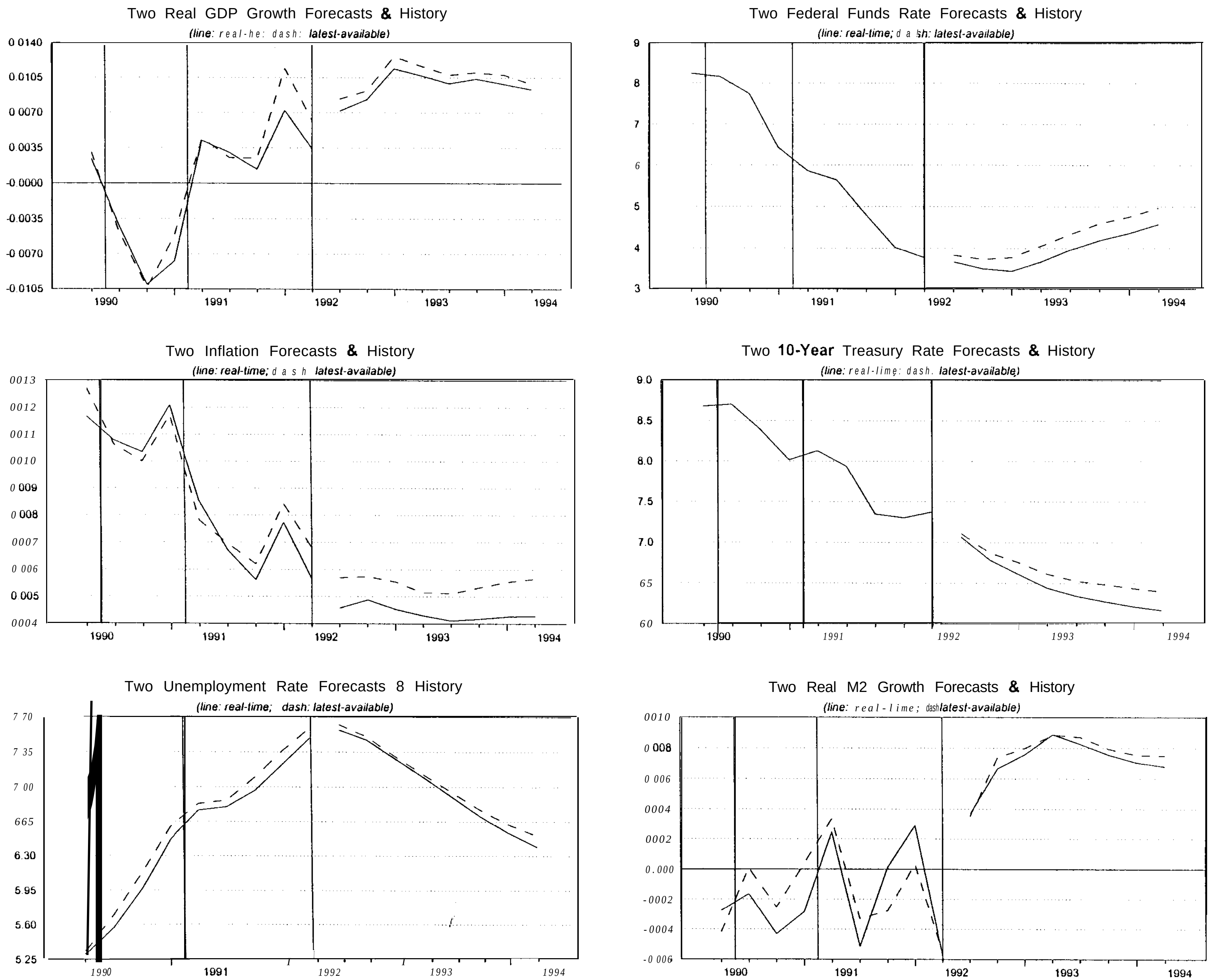
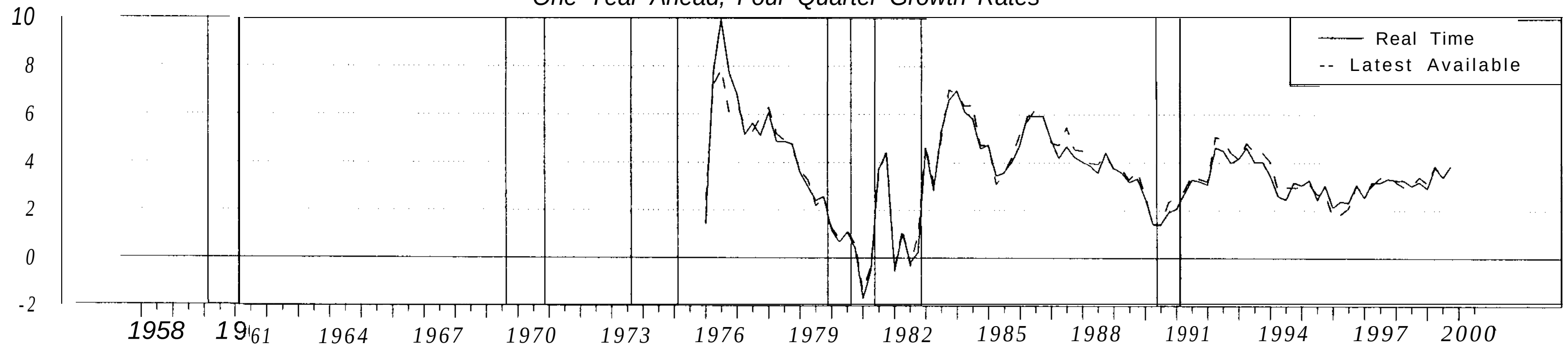

Two Real GDP Growth Forecasts & History
(line: real-time; dash: latest-available)
Two Federal Funds Rate Forecasts & History
(line: real-time; dash: latest-available)
Two Inflation Forecasts & History
(line: real-time; dash: latest-available)
Two 10-Year Treasury Rate Forecasts & History
(line: real-time; dash: latest-available)
Two Unemployment Rate Forecasts 8 History
(line: real-time; dash: latest-available)
Two Real M2 Growth Forecasts & History
(line: real-time; dash: latest-available)
1990
1991
1992
1993
1994

Figure VII.9

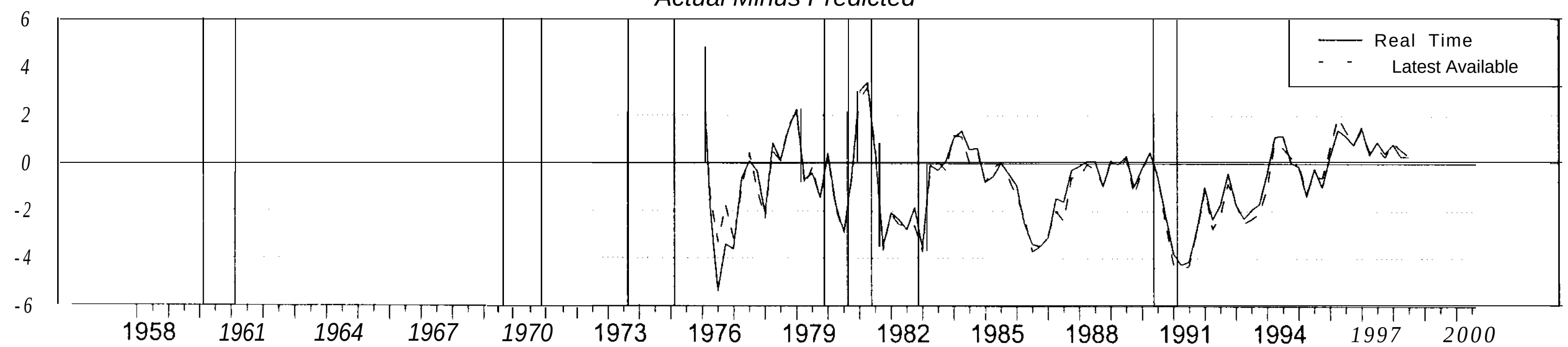
A Comparison of Two Real GDP Forecasts From a Rolling QBVEC(5) Model

One Year Ahead, Four Quarter Growth Rates



A Comparison of Two Forecast Errors From a Rolling QBVEC(5) Model

Actual Minus Predicted



Difference Between Errors

Real Time Minus Latest Available

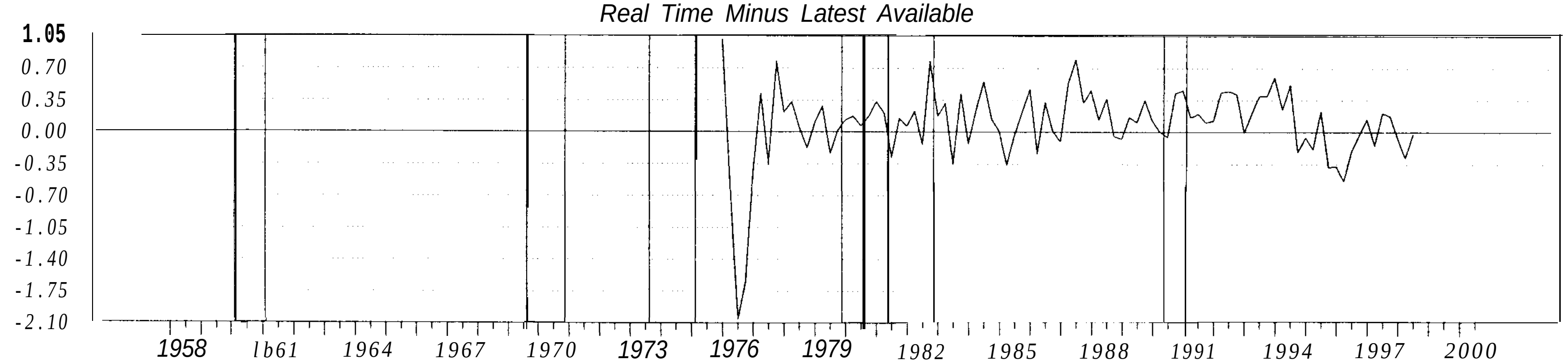
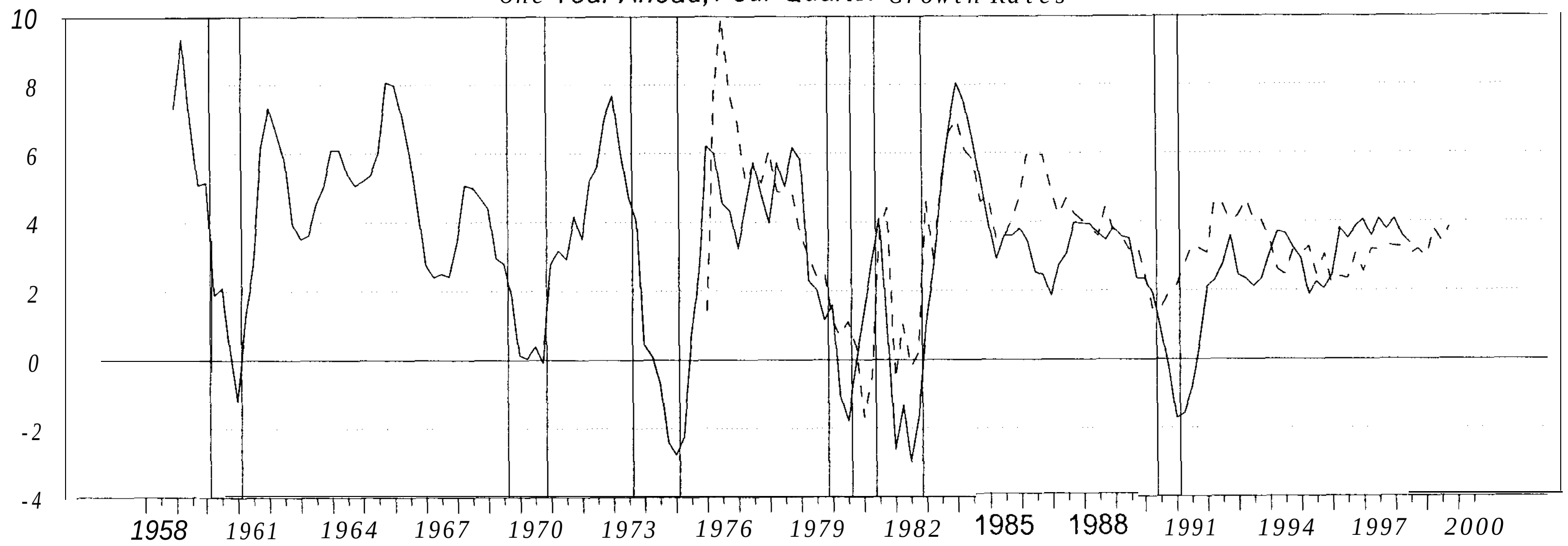


Figure VII.10

Actual (line) & Predicted (dash) Real GDP: Real Time

One Year Ahead, Four Quarter Growth Rates



Actual (line) & Predicted (dash) Real GDP: Latest Available

One Year Ahead, Four Quarter Growth Rates

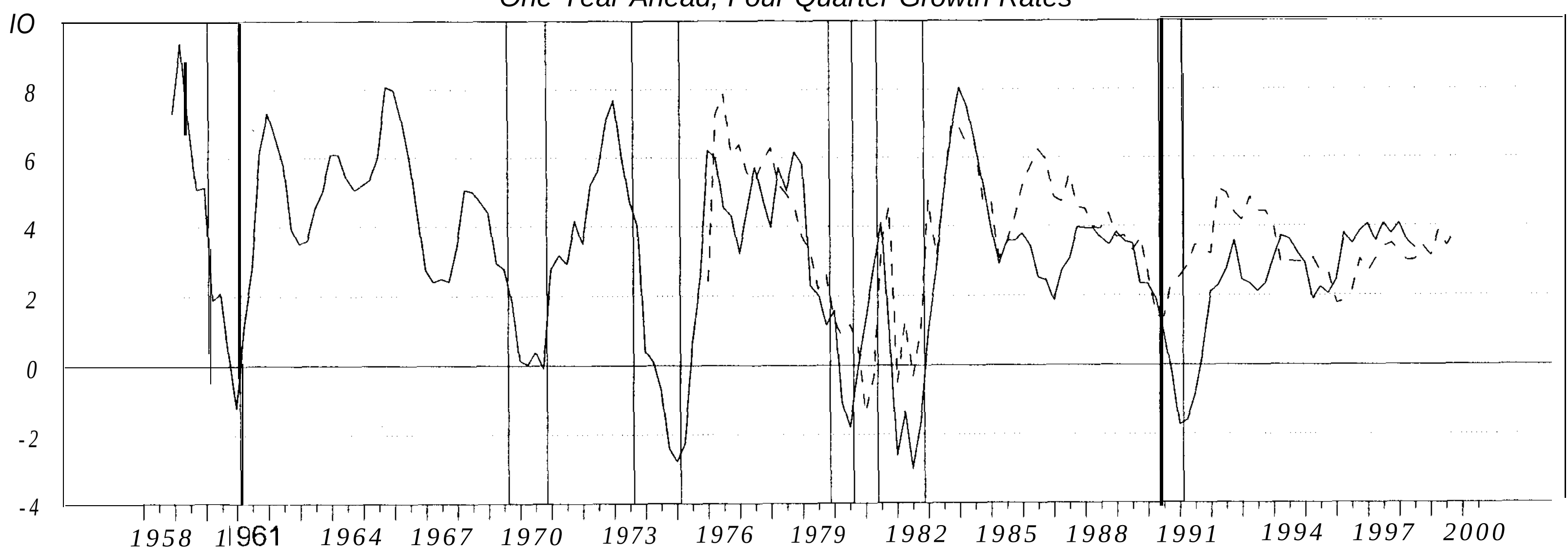


Figure VII.11

Two Real GDP Growth Forecasts From a Rolling QBVEC(5)

197601 to 199904, One Year Ahead, 4-Quarter Averages

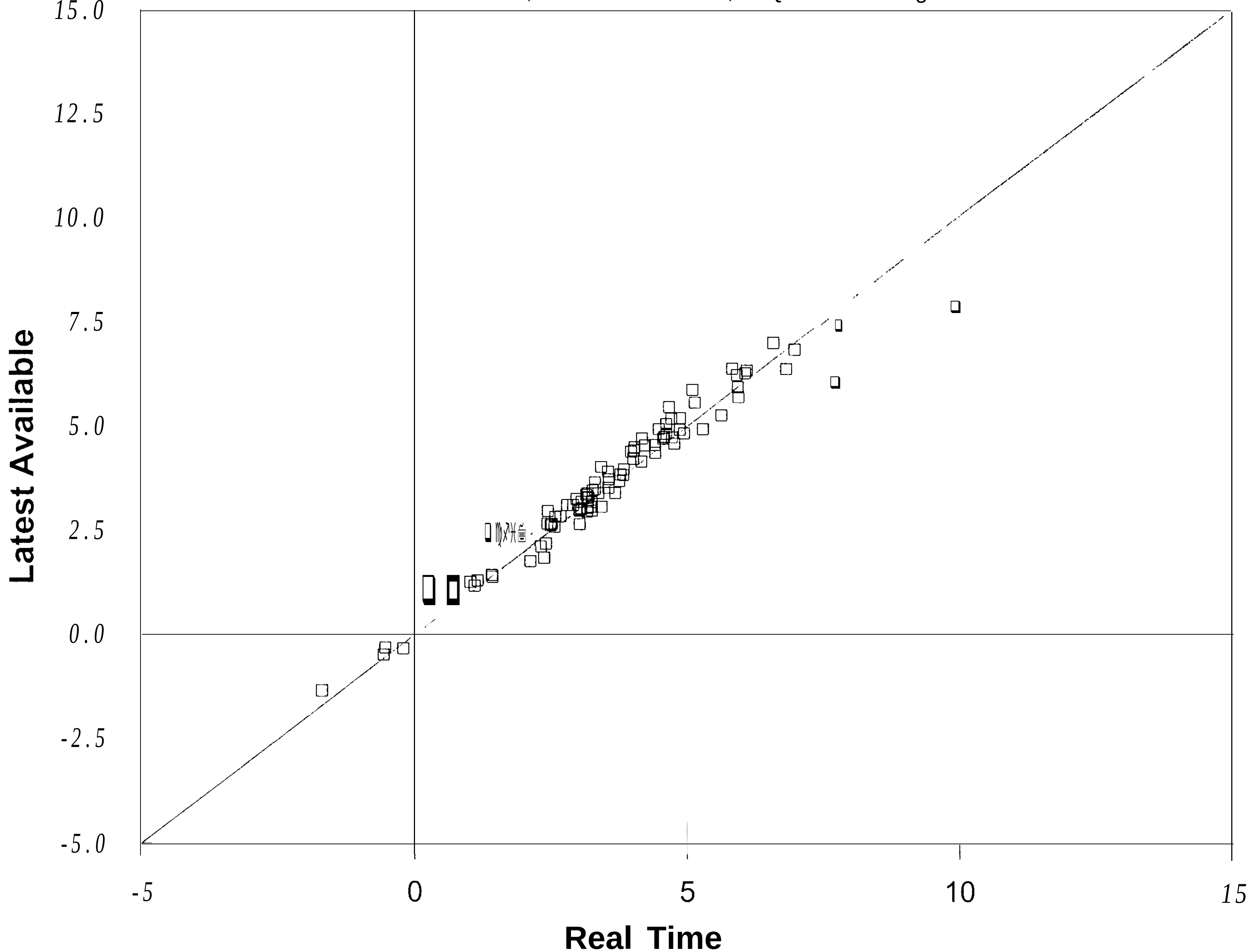


Figure VII.12

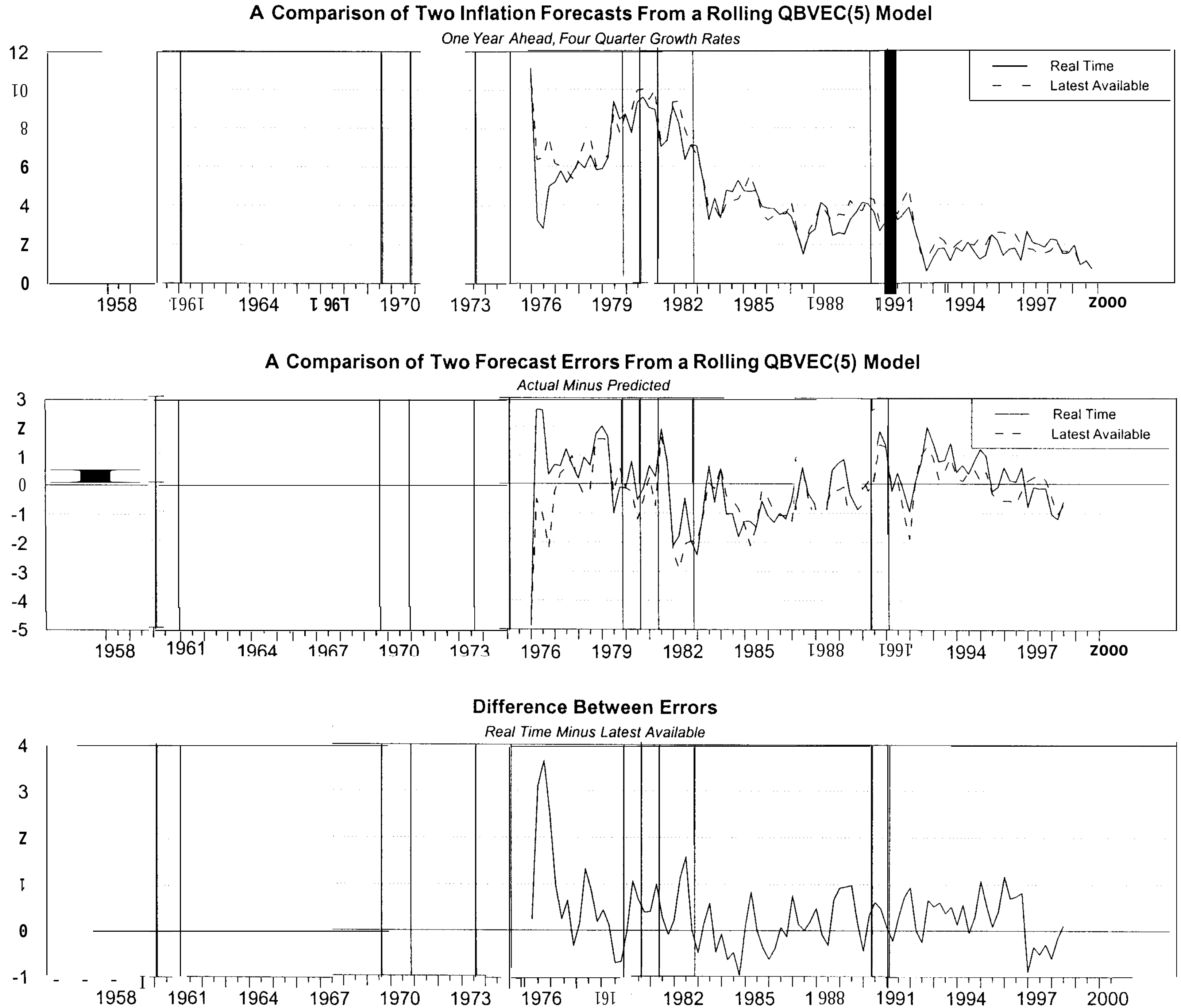
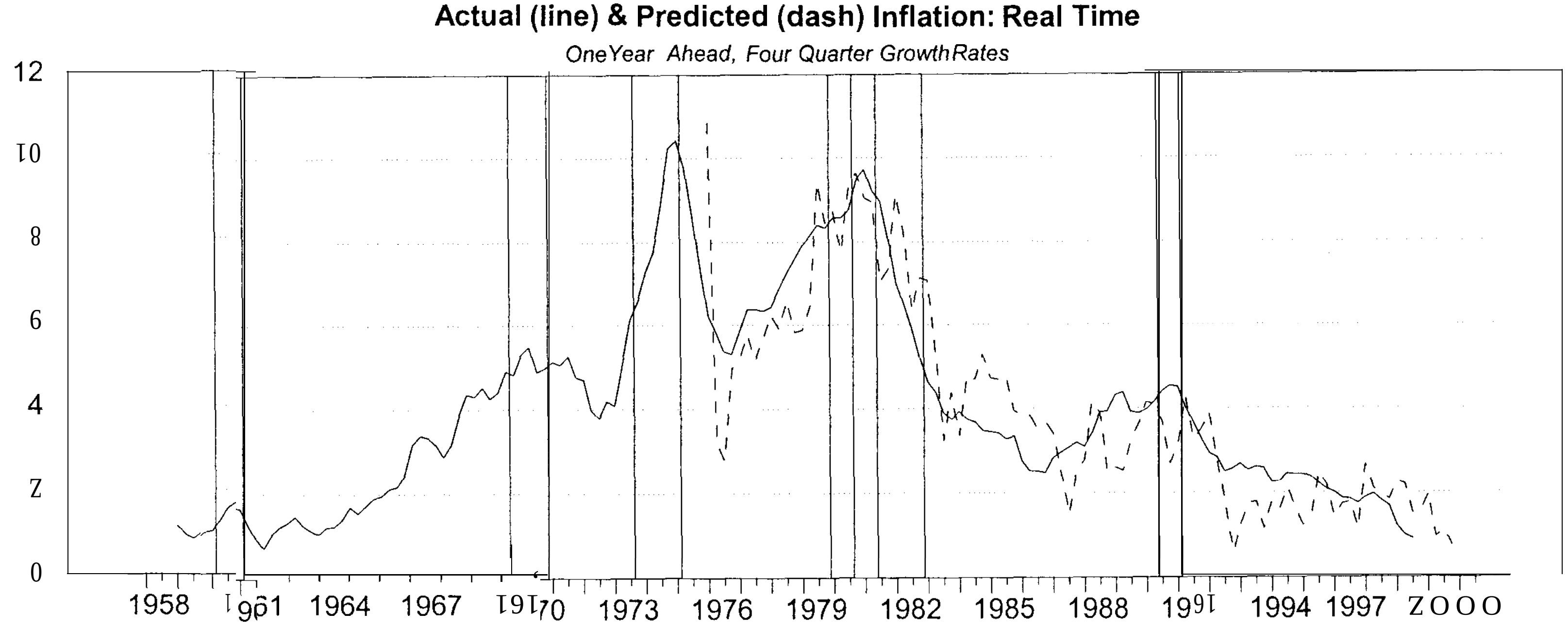


Figure VII.13



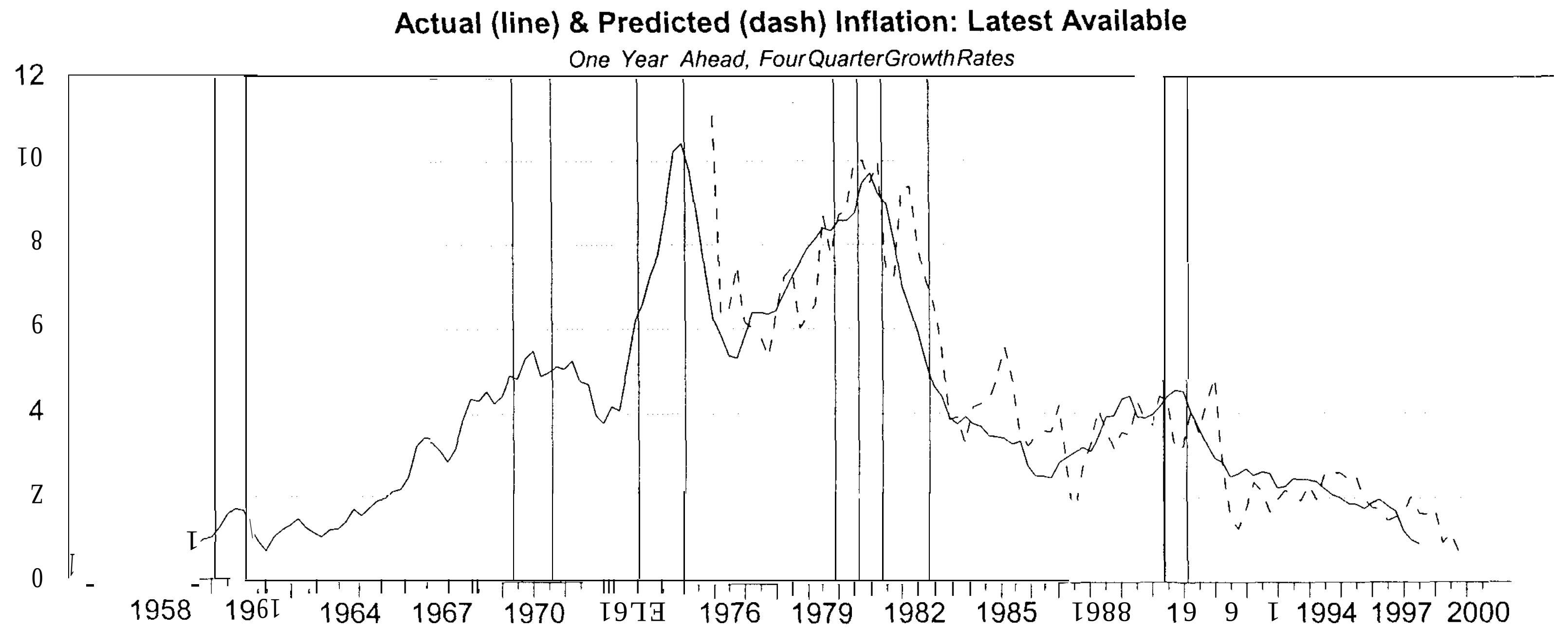


Figure VII.14