Real-Time GDP Growth Forecasts

Evan F. Koenig

And

Shelia Dolomas

December 1997

Research Department

Working Paper

97-10

Federal Reserve Bank of Dallas

REAL-TIME GDP GROWTH FORECASTS

by

Evan F. Koenig
and
Sheila Dolmas

Research Department
Federal Reserve Bank of Dallas
2200 N. Pearl St.
Dallas, TX 55201

November 1997

This paper benefited greatly from comments and suggestions offered by Preston Miller and, especially, Ken Emery. The views expressed are those of the authors, and should not be taken to represent the official views of the Federal Reserve Bank of Dallas or the Federal Reserve System.

Abstract

We forecast current-quarter real GDP growth using monthly data that would have been available to an analyst in real time. We demonstrate that using real-time data is of major importance both when estimating GDP forecasting models and when evaluating their performance. Moreover, we show that the out-of-sample forecasting performance of our model is comparable or superior to that of the Blue-Chip consensus forecast provided that more than one month of current-quarter data are available.

Introduction

Both when making business plans and when formulating monetary policy, it is essential to have as clear a picture as possible of current economic conditions. In this regard, an important summary statistic is the growth rate of real gross domestic product (GDP). Economists devote substantial time and effort to constructing early estimates of current-quarter GDP growth, and their prognostications receive much press attention. Despite this effort and scrutiny, GDP forecasts are not very accurate. For example, since 1990, the root-mean-square error of the highly respected Blue Chip consensus forecast of current-quarter GDP has been 1.6 percentage points based on forecasts published in the second month of the quarter, and 1.2 percentage points based on forecasts published in the first month *after* the quarter. A 95% confidence interval for an early estimate of real GDP growth is fully 6.2 percentage points wide, while a 95% confidence interval for an end-of-quarter estimate of real GDP growth is 4.8 percentage points wide.[1]

This paper reports on an effort to use monthly, coincident indicators of real economic activity to forecast current-quarter GDP growth. In large part, the motivation for this effort is a desire to obtain more accurate and more timely forecasts than those currently available from private forecasting firms. Recent research suggests that the predictions of individual private analysts may have an irrational element (Lamont 1995, Ehrbeck and Waldmann 1996) or be rationally inaccurate (Laster, Bennett, and Geoum 1997). Consensus forecasts have a better record than most individual analysts (Graham 1996, McNees 1987), but often do not reflect all the information that one might wish. For example, the Blue Chip newsletter that a subscriber receives during the second week of a given month contains forecasts based on information that was available within the first week of that month. As a result, the forecasts contained in the July newsletter do not reflect industrial production and retail sales data for June, and may or may not reflect the June employment numbers, even though all these data are released by mid July.

---

[1] These root-mean-square errors and confidence bounds assume that one is trying to forecast the Commerce Department's "final" GDP growth estimate, which becomes available with a three-month lag. If, instead, one is trying to predict the Commerce Department's "advance" estimate (available with a one-month lag), the early-quarter and end-of-quarter errors are 1.3 percentage points and 0.9 percentage points, respectively. The corresponding 95% confidence bounds are 5.1 and 3.6 percentage points wide.

Finally, it is not clear whether private analysts are trying to forecast the variables that are of greatest concern to policy makers. Within the current quarter, private analysts appear to focus on predicting the Commerce Department's first, or "advance" estimate of GDP growth (Trehan 1989). However, this initial government estimate is based on incomplete data. Arguably, policy makers are more interested in the third, or "final" estimate of GDP growth, which more accurately measures the actual behavior of the economy.[2] The model that does the best job of forecasting the advance GDP estimate is not necessarily the model that does the best job of forecasting the final estimate.

Our study is unique in its extensive use of real-time data. For each variable in our model at each month in our sample, we have a 12-month history of the data that were available at the time. This data set allows us to obtain an accurate assessment of how well our model is likely to perform in actual use. Moreover, it allows us to achieve a level of forecasting performance markedly superior to that which would have been possible had we estimated the model conventionally, using today's data. To minimize the dangers of over fitting, we rely heavily on rolling, out-of-sample forecast exercises when evaluating the performance of our model and when comparing its forecasts to the forecasts of others.

When we say we use real-time data, we mean that at every point in the sample, the data used in the estimation is always the data that would have been available to a private forecaster at the time. For example, when the left-hand-side variable is 1985:Q1 GDP growth, all right-hand-side variables are measured as they appeared in the first quarter of 1985. Thus, we use data of as many vintages as there are data points in the sample.

Closely related work includes Braun (1990), Trehan (1989, 1992), Fitzgerald and Miller

---

[2] Of course, even the "final" estimate is not really final: it is followed by benchmark revisions and rebasings. However, both types of revision become available only so far after the fact as to be largely irrelevant to policy makers. Moreover, it is not at all clear that rebased statistics give a more accurate picture of GDP movements than do earlier releases. For example, 1985 real GDP growth is probably better measured in 1982 dollars than in 1992 dollars. Of course, rebasings are much less of an issue for chain-weight measures of real GDP than for fixed-weight measures.

(1989), and Miller and Chin (1996).[3] Braun predicts current-quarter output growth using monthly labor-market data. His procedure has two steps: estimating a relationship between output growth and the quarterly average of either aggregate hours or the unemployment rate, and forecasting the quarterly average of the relevant labor-market variable from available monthly observations of that variable. Although Braun is careful to use real-time hours and unemployment data in his estimations, the output-growth data are not real time. Moreover, only in-sample forecasting results are reported.[4]

Trehan, like Braun, uses a two-step approach to forecasting current-quarter aggregate output. Three monthly indicator variables are included in the model: non-farm employment, industrial production, and real retail sales. When complete data for a given quarter are unavailable, a Bayesian vector autoregression (BVAR) is used to fill in the missing information. Unfortunately, the model is estimated using current data rather than data that would have been available to an analyst forecasting in real time. Moreover, Trehan takes only a cursory look at the real-time performance of his model in comparison to the forecasting performance of private analysts.[5]

Miller and Chin take Trehan's approach one step farther, combining the GDP forecasts generated by a model that uses monthly data with those generated by a more conventional quarterly model. Like Trehan, Miller and Chin do most of their analysis using currently-available data and take only a brief look at the real-time performance of their model.

Unlike Trehan and Miller-Chin, Fitzgerald and Miller use only real-time data. However, the Fitzgerald-Miller definition of real-time data differs from the definition used here. Thus,

---

[3] Zadrozny (1990) and Rathjens and Robins (1993) are somewhat less closely related to the current paper, as they use monthly data to improve forecasts of *next* quarter's output growth. Moreover, neither paper uses real-time data.

[4] The distinction between in-sample and out-of-sample results is potentially quite important in Braun's framework, because his output-growth forecasts are contingent on estimates of trend productivity growth (in the hours model) or the NAIRU and potential output (in the unemployment model). All three of these estimates are notoriously subject to revision.

[5] Table 3 in Trehan (1992) reports the real-time mean errors, mean absolute errors, and root mean square errors generated over a four-year period by the Trehan model and the Blue Chip consensus forecast.

Fitzgerald and Miller use data that is of a single vintage in each of their estimations: each right-hand-side variable is measured as it would have been at the end-date of the sample period. In contrast, we have as many vintages as data points: at each date within our sample, we use only data that would have been available at the time. Moreover, Fitzgerald and Miller limit themselves to predicting the advance estimate of output growth using monthly aggregate hours data. Forecasts from the Fitzgerald-Miller model are compared with those from the Minneapolis Fed's quarterly model, but not with the monthly forecasts of private analysts.

We have blended aspects of the Braun, Trehan, and Fitzgerald-Miller approaches to forecasting aggregate output. Like Trehan, we look to monthly employment, industrial production, and real retail sales for information on current-quarter real GDP. As in Braun, our right-hand-side variables are all measured as they would have been in real time. However, as in Fitzgerald and Miller, our aggregate output data is also real time, being either real GDP growth as initially reported or as reported in the Commerce Department's final (third) release. Real-time data sets are tedious to assemble. To keep the data requirements of the current exercise manageable, we do not follow Braun, Trehan, and Miller-Chin in estimating a separate model for forecasting missing monthly data.[6] Instead, we regress GDP growth directly on monthly employment, production, and sales data, and on lagged quarterly GDP growth rates.

Our principal findings are as follows. First, provided that we have two or three months of current-quarter data, the Blue Chip forecast contains no information beyond that already contained in the forecasts of our model, and our root-mean-square errors are substantially lower than those reported by Miller and Chin and Fitzgerald and Miller. On the other hand, our model does rather poorly when only one month of current quarter data are available. This comparative weakness probably reflects the fact that our model contains only coincident--not *leading*--indicators of real economic activity. Second, our out-of-sample predictions of the advance GDP estimate are somewhat more accurate than our predictions of the final GDP estimate. Both sets

[6] One might be tempted to include a short-term interest rate or a long-short interest-rate spread in the forecasting model, on the grounds that such variables are not subject to ex post revisions and tend to move in advance of real activity. However, any such forecasting relationship would likely be sensitive to the monetary authority's policy rule and, hence, unreliable.

of forecasts pass simple efficiency and stability tests, provided that two or three months of current-quarter data are available. Finally, we demonstrate how important it is that the estimation and evaluation of GDP forecasting models be conducted using data that would have been available to an analyst in real time. Out-of-sample forecasting exercises that use currently-available data rather than real-time data can give a very misleading impression of how well a forecasting model will do in real time. For the particular forecasting model developed in this paper, taking the conventional approach markedly *under*states real-time performance.[7]

The following section describes our model in detail. Next, the real-time data set is discussed and empirical results are presented. Concluding remarks complete the paper.

The Model

We actually estimate three completely separate models: one using a single month of current-quarter data, a second using two months of current-quarter data, and a third using a full three months of current-quarter data. In principle, there are restrictions that one could impose across the models to improve the efficiency of the estimation. We chose, instead, to focus our efforts on collecting an unusually complete set of real-time data (described below) and conducting a thorough set of out-of-sample real-time forecasting experiments.

Following Trehan (1992), our initial set of monthly indicator variables included non-farm employment, real retail sales (nominal sales deflated by the consumer price index), and industrial production. These variables are all important and closely-watched direct measures of current real economic activity. Non-farm employment and industrial production are among only four variables included in the Conference Board's composite coincident index, and real retail sales serve as a timely proxy for a third component of that index (real manufacturing and trade sales).[8,9]

---

[7] Braun (1990) finds that exactly the opposite is true for his models.

[8] The fourth component of the coincident index--real personal income--is released substantially later than the employment, retail sales, CPI, and industrial production reports.

[9] Based on findings reported in Koenig (1996) and Fitzgerald and Miller (1989), we tried including manufacturing capacity utilization, the aggregate hours of workers in the service-

To obtain our forecasting models, we regressed the annualized quarter-to-quarter percentage change in real GDP on a constant, four lagged percentage changes in real GDP, and five annualized month-to-month percentage changes in each of our three coincident indicators. To be precise, we estimated equations the form:

$$\Delta y_t = \alpha_{0,s} + \Sigma_{i=1}^{4} \alpha_{i,s} \Delta y_{t-i} + \Sigma_{j=1}^{5} \beta_{j,s} \delta em_{t,s-j} + \Sigma_{j=1}^{5} \gamma_{j,s} \delta ip_{t,s-j} + \Sigma_{j=1}^{5} \epsilon_{j,s} \delta rs_{t,s-j},$$

where Δy_t denotes the annualized quarterly percentage change in real GDP in quarter t , and where $\delta em_{t,s}$, $\delta ip_{t,s}$, and $\delta rs_{t,s}$ are the annualized monthly percentage changes in non-farm employment, industrial production, and real retail sales, respectively, in month s of quarter t.[10] When s = 1, all right-hand-side variables are as they would have appeared to an analyst immediately after the release of the industrial production, retail sales, and CPI reports for the first month of quarter t. Similarly, when s = 2, all right-hand-side variables are as they would have appeared to an analyst after the release of the industrial production, retail sales, and CPI reports for the second month of quarter t. Finally, when s = 3, all right-hand-side variables are as they would have appeared to an analyst after the release of the industrial production, retail sales, and reports for the third month of quarter t. As alternative left-hand-side variables we used real GDP growth as estimated in the Commerce Department's "advance" report (generally released during the first month after the end of the quarter) and real GDP growth as estimated in the Commerce Department's "final" report (released during the third month after the end of the quarter).[11]

---

producing sector, and the ratio of goods-producing to service-producing hours as additional right-hand-side variables. However, none of these variables was statistically significant, and we dropped them from our analysis. Below, we compare the out-of-sample forecasting performance of our model to that of the Fitzgerald-Miller model.

[10] If $x_{t,s}$ is a monthly variable, we define $x_{t,0} \equiv x_{t-1,3}$, $x_{t,-1} \equiv x_{t-1,2}$, $x_{t,-2} \equiv x_{t-1,1}$, etc.

[11] During the three-year period from 1984 through 1986, the Commerce Department released a "flash" current-quarter GNP estimate in the third month of each quarter. Our analysis ignores this estimate.

Equation 1 can be rationalized as follows. Let y_t denote the logarithm of quarterly aggregate output and suppose that there is a monthly measure of current real economic activity, $z_{t,s}$, such that $y_t = (z_{t,3} + z_{t,2} + z_{t,1})/3$ for all t. Then

$$y_t - y_{t-1} = [(z_{t,3} - z_{t,2}) + 2(z_{t,2} - z_{t,1}) + 3(z_{t,1} - z_{t-1,3}) + 2(z_{t-1,3} - z_{t-1,2}) + (z_{t-1,2} - z_{t-1,1})]/3. \quad (2)$$

Thus, the quarter-to-quarter percentage change in real GDP is a weighted average of five month-to-month percentage changes in the coincident indicator. In practice, one or more of the monthly percentage changes on the right-hand side of equation 2 will be either a preliminary estimate or entirely unavailable. If a preliminary estimate, then a regression will de-emphasize that percentage change in favor of others, measured more accurately. If entirely unavailable, then other lagged monthly changes in the coincident indicator may capture some of the missing information. For these reasons, when estimating equation 1 we do not restrict the coefficient weights attached to monthly percentage changes in employment, industrial production, and retail sales. Moreover, when s = 1 or s = 2, we extend the distributed lags in the coincident indicators back in time to include monthly changes from two quarters prior to t.

Data and Estimation Methodology

General Discussion. Table 1 illustrates how we went about estimating our models, using the 1997:Q1 GDP growth forecast as an example. As shown in the top third of the table, all data used in the 1-month model were available by February 19, when the last of the monthly data for January (the CPI) were released. In addition to January data, our forecast is based on lagged monthly growth rates of employment, sales, production and prices extending from September through December of 1996--all measured as of February, 1997--and on GDP growth rates over the period from 1996:Q1-1996:Q4. These lagged GDP growth rates are measured as of January 31, 1997, when the earliest estimate of 1996:Q4 GDP growth was released. Two different versions of the model are estimated. In one version, the left-hand-side variable is the advance estimate of 1997:Q1 GDP growth. In the other version, the left-hand-side variable is the final estimate of 1997:Q1 GDP growth.

As we move to the 2-month and 3-month GDP growth models, notice three things. First, the left-hand-side variables do not change. Second, all three forecasting equations have the same lags of GDP growth on their right-hand sides (1996:Q1 through 1996:Q4). However, the GDP data undergo revisions as we move from the 1-month model to the 2-month model to the 3-month model. Third, the time period covered by the monthly variables on the right-hand sides of the forecasting equations changes as we move from one model to the next. In particular, the range of months over which growth in employment, retail sales, industrial production, and the CPI are measured shifts forward by one month, and all these data go through an additional month of revisions.

Chain-Weight GDP. In constructing the data sets used to forecast chain-weight GDP growth, we treated the switch to chain-weight numbers just like any other GDP data revision or rebasing. In particular, the data sets begin with fixed-weight GDP numbers, and then change over to chain-weight numbers as they become available. We constructed two different data sets for each of the models. The first--used in forecasting "final" chain-weight GDP--switches to the chain-weight numbers when the "final" chain-weight numbers were first released, in the first quarter of 1993. The second, used in forecasting "advance" chain-weight GDP, switches to chain-weight numbers when the "advance" chain-weight numbers were first released. The Commerce Department did not begin publishing its "advance" estimates of chain-weight GDP growth until October of 1994, for the third quarter of 1994. In other words, prior to October of 1994, the chain-weight GDP numbers were released two and three months after each quarter, with no one-month estimate.

The Results

Forecasting Fixed-Weight GDP. We estimated our fixed-weight GDP forecasting equations using data from 1980:Q1 through 1989:Q4 and again using data from 1980:Q1 through 1996:Q4. As noted above, separate models were estimated for predicting the advance estimate of real GDP and predicting the final estimate of real GDP. Moreover, separate models were estimated for the cases in which the analyst would have had one-month of current-quarter data available, two months of current-quarter data available, and three-months of current-quarter data available. All data were real time--exactly the data that would have been available to a private forecaster over this period. For example, when predicting 1985:Q1 GDP growth, we measure all of our right-hand-side variables as they were measured in 1985:Q1.

Tables 2A-C present summary statistics for the in-sample regressions, including the joint statistical significance of the lags of each of the right-hand-side variables, the sum of the coefficients attached to the lags of each of the right-hand-side variables, and the statistical significance of the sum of the coefficients attached to each of the right-hand-side variables. Collectively, the monthly percentage changes in employment, industrial production, and retail sales are always highly statistically significant. (See the F-test results toward the bottom of the tables.) However, due to colinearity between the three indicators, the monthly percentage changes in any particular indicator are sometimes not significant. Advance GDP is consistently easier to predict than final GDP. Possible explanations of this result are discussed below. In predicting final GDP, the overall weight placed on monthly employment data noticeably increases as one goes from forecasts based on one month of current-quarter data to forecasts based on two months of current-quarter data to forecasts based on three months of current-quarter data. Serial correlation is a significant problem only in the model that predicts final real GDP using two months of current-quarter data.

Our out-of-sample forecasting exercises were conducted using rolling samples. Thus, coefficient estimates obtained using data through 1989:Q4 were used to forecast real GDP growth in 1990:Q1. The sample period was then extended by one quarter, the models re-estimated, and the new coefficient estimates were used to forecast 1990:Q2 GDP growth. In this way, we obtained forecasts running from 1990:Q1 through 1995:Q3. The ending date was chosen to preserve comparability with the Miller-Chin and Blue-Chip consensus forecasts. (Miller-Chin's real-time results are confined to 1990:Q1-1995:Q3, and Blue-Chip participants abandoned fixed-weight GDP forecasting in favor of chain-weight GDP forecasting beginning in 1996:Q1.) As always, at each date we used only data that would actually have been available to a private forecaster. Summary statistics from these rolling, out-of-sample forecasting exercises are displayed in Table 3A, in the rows labeled "KD." Plots of actual and forecasted GDP growth are displayed in Figure 1A (the advance GDP estimate) and Figure 1B (the final GDP estimate).

In two important respects, our results are similar to those reported by other analysts. First, we find that it is easier to predict the advance estimate of GDP growth than it is to predict the final estimate. For example, with three months of current-quarter data, the root-mean-square error of our forecast of advance GDP is 0.82 percentage points--1/3 smaller than the 1.23-percentage-point-root-mean-square error that we obtain when forecasting final GDP. Second, we find that the improvement in forecasting performance that is achieved by going from one month of current-quarter data to two months of current quarter data is much larger than that achieved by going from two months of current-quarter data to three months of current-quarter data. Thus, the root-mean-square error of our forecasts of advance GDP drop from 1.63 to 0.93 to 0.82 as we move from 1 month to 2 months to 3 months of current-quarter data. In predicting the final estimate of GDP growth, the root-mean-square error is cut by over 1/3 as a result of adding a second month of data, and not at all as a result of adding a third month of data.[12]

Our first thought was that the relative ease with which we are able to predict the advance GDP estimate might reflect our use of real-time data, rather than revised data, for our right-hand-side variables. For example, our 3-month forecasts are based on data of the same vintage as that available to the Commerce Department when it was preparing the advance estimate of GDP. The data used by the Commerce Department to construct the final GDP estimate, in contrast, is at least two months older than ours. In an effort to test the importance of this "vintage effect" we estimated a version of our 3-month model of final GDP in which the right-hand-side variables

---

[12] We experimented with a model intermediate between the 1-month and 2-month Koenig-Dolmas models described above. It used two months of current-quarter employment data, but only one month of current-quarter sales and production data. (The rationale is that sales and production data are not released until about two weeks after the employment data become available.) As might be expected, out-of-sample performance was intermediate between that of our 1-month and 2-month models. However, performance was not as good as the Miller-Chin, Fitzgerald-Miller, and Blue Chip consensus forecasts that would have been available at about the same time (the first week of the third month of the quarter). A model intermediate between our 2-month and 3-month models performed no better than our 2-month model. Given that our 2-month and 3-month models perform about equally well, this result is also not particularly surprising.

were measured three months after the close of the quarter (matching the vintage of the final GDP estimate). Surprisingly, the in-sample and out-of-sample forecasting performance of the model *deteriorated* slightly rather than improved. Apparently, revisions to our monthly indicators are not highly correlated with revisions to the Commerce Department's GDP estimates. The monthly data that probably *are* correlated with GDP revisions are data for variables like inventory investment and net export growth, that are not included in our set of indicators.

As for the result that the third month of current-quarter data has a smaller impact on forecast performance than does the second month, a large part of the explanation is apparent in equation 2: in calculating the quarter-to-quarter change in real activity, the third month of current-quarter data receives only ½ as much weight as the second month of current-quarter data.

Our out-of-sample forecast period includes one outright recession and several quarters in which estimated GDP growth dropped below 1%, but remained positive. For a policy maker, the distinction between outright recessions and growth recessions is important, and it is essential that a forecasting model not confuse the two. In this regard, Figure 1 suggests that our 1-month model is much less satisfactory than our 2-month and 3-month models. The 1-month model often recognizes recessions and slowdowns after the fact, and tends to convert quarters of weak but positive growth into quarters of GDP decline.

How do our forecasts stack up against the real-time forecasts of others? In addition to summary performance measures for the Koenig-Dolmas model, Table 3A gives comparable measures of the performance for the Miller-Chin and Fitzgerald-Miller models and the Blue Chip consensus forecast. The table lists the various forecasts in the order in which they become available. For example, the first forecast listed is the Blue Chip consensus forecast published in the second week of the first month of the quarter, before any current-quarter data are available. The second and third forecasts listed are those obtained from the Miller-Chin and Fitzgerald-Miller models in the first week of the second month of the quarter, just after the release of the employment report for the first month of the quarter. The final forecast is the Commerce Department's own "advance" GDP estimate, released toward the end of the first month of the following quarter. (In the table, the first month of the following quarter is labeled "month four" of the current quarter.)

Forecasting performance ought to improve as more current-quarter data become available. A general tendency in this direction is apparent in the root-mean-square errors reported in Table 3A, but there are notable exceptions. First, the root-mean-square error of each Miller-Chin and each Fitzgerald-Miller forecast is never lower than the root-mean-square error of the Blue Chip forecast released the previous month. Second, the Koenig-Dolmas and Fitzgerald-Miller forecasts that become available during the second month of the quarter yield root-mean-square errors that are strikingly higher than those of the Blue Chip and Miller-Chin forecasts. This poor performance probably reflects the fact that the Koenig-Dolmas and Fitzgerald-Miller forecasts are based solely on coincident indicators of economic activity. In contrast, the Miller-Chin and Blue Chip forecasts incorporate information on variables that tend to lead the business cycle.

One would expect the importance of leading indicators to diminish as more and more current-quarter data become available. Consistent with this expectation, the performance of the Koenig-Dolmas models improves relative to the performance of the Blue Chip forecasts as we move from 1-month results to 2-month and 3-month results. Indeed, our 2-month and 3-month models nearly always yield root-mean-square errors that are lower than those obtained from the Blue Chip newsletter released the same month. The forecasting performance of our 2-month model is nearly as good as that of the Blue Chip newsletter released the *following* month.

In predicting final GDP, the Commerce Department's own advance GDP estimate clearly dominates all challengers.

How is it that our 2-month and 3-month models perform so well, despite their limited information sets and relatively unsophisticated econometrics? We think that the key is our real-time data set. Evidence consistent with this hypothesis is contained in Table 3B. The "KD (rev.)" results in this table are from models estimated using today's data (specifically, data as they appeared in March, 1997), but used to forecast in real time (that is, real-time data are plugged into the estimated equations to generate forecasts). The effect of using today's data in the estimation of the 2-month and 3-month models is to increase their root-mean-square errors by about 50% when predicting advance GDP and by between 22% and 35% when predicting final

GDP.[13] Clearly, the real-time forecasting performance of these models is quite sensitive to how they are estimated: for optimal performance it is important that at each date within the sample period, the data contained in the sample be exactly the data that would have been available to an analyst at the time.

Suppose that we not only estimate our models' equations using today's data, but also plug today's data into the estimated equations to generate forecasts of GDP growth. Moreover, suppose that we compare our forecasts with GDP growth as it is currently reported. In other words, suppose that we do what analysts *usually* do when estimating and evaluating their models and reporting their results. In Table 3B, this exercise is labeled "Naive KD."[14] For the 2-month and 3-month models, root-mean-square errors are 40% to 50% higher than those recorded for the same models estimated and evaluated using real-time data. Root-mean-square errors are between 10% and 15% higher than those reported in the lines labeled "KD (rev.)," where the models are estimated using today's data but evaluated using real-time data. The lesson is that one must use real-time data in both estimation and evaluation if one is to get an accurate sense of how well a given forecasting model is capable of performing in actual use. For our models, the usual approach--which only makes use of today's data--markedly understates actual performance.

Tables 4A-C present results from efficiency tests and tests of marginal predictive power. First, we regressed Commerce Department GDP estimates on a constant and each of several out-of-sample forecasts, including forecasts generated by our own real-time models. A forecast is called efficient if the constant term in this regression is not significantly different from 0 and the coefficient attached to the forecast is not significantly different from 1. The only forecasts that are consistently inefficient are those that our model generates when it is estimated using today's data. [See the results labeled "KD(rev.)."] In addition, our 3-month model estimated with real-

---

[13] We conducted a similar exercise in which our models were estimated using data as they appeared at the start of the out-of-sample forecast period, in 1989:Q4. Errors were even larger than those generated by the models estimated with 1997 data.

[14] We report these results in columns headed "Predicting Final Fixed-Weight GDP" even though in this particular case we are comparing the forecasts generated by our model to *today's* GDP growth data rather than real-time Commerce Department "final" estimates.

time data appears to be inefficient when used to predict the final GDP release.[15]

Next, we regressed Commerce Department GDP estimates on a constant, the forecasts of one of our real-time models, and an alternative forecast, such as the Blue Chip consensus. If the alternative forecast has predictive power beyond that of the forecast generated by our real-time model, then the alternative forecast will enter this regression with a statistically significant coefficient. According to Table 4A, not only do the Blue Chip forecasts have predictive power beyond those of our 1-month model, they totally dominate our 1-month forecasts. (Our model fails to have any marginal predictive power beyond the Blue Chip forecasts.) Finally, the entries in the bottom row of Table 4A indicate that there is no advantage to using real-time data when estimating our 1-month model.

Results for our 2-month and 3-month models are considerably more encouraging. According to Tables 4B and 4C, these models, estimated using real-time data, dominate the Blue Chip forecasts released the same month. (In row 5 of the tables, our models have marginal predictive power and the Blue Chip forecasts do not.) Indeed, in predicting the Commerce Department's advance GDP estimate, the performance of our 2-month model compares favorably with that of the Blue Chip newsletter released the *following* month. (In row 6 of Table 4B, our 2-month forecast and the Blue Chip forecast each receive about 50% weight. Multicolinearity prevents either coefficient from achieving statistical significance.) The importance of using real-time data when estimating GDP forecasting equations is illustrated by the results reported in the very last rows of Tables 4B and 4C, which show that our 2-month and 3-month real-time models dominate the same models estimated using today's data.

Row 6 of Table 4C pits our 3-month model of final GDP against the Commerce Department's advance GDP release. One cannot reject the hypothesis that our forecast contains no information beyond that included in the official advance estimate. In contrast, Trehan (1989)

---

[15] For this model, the joint hypothesis that the constant term in the efficiency regression is 0 and the slope coefficient is 1 has marginal probability .033.

presents evidence that the advance GDP estimate was inefficient during the 1980s.[16] In an effort to shed further light on the efficiency of the official advance estimate, we again compared the Commerce Department's advance estimates with the forecasts of our 3-month model, this time using in-sample forecasts extending back all the way to 1980:Q1. Results are reported in Table 5, row 1. In results similar to those reported by Trehan and strikingly different from the results reported in Table 4C, our model's forecasts receive 50% weight over the extended sample period, and are highly statistically significant.

A clue to what is happening is displayed in Figure 2, which shows the number of days delay with which the advance GDP estimate was released, beginning in 1980:Q1 and running through 1996:Q4. Over the early part of the sample (through 1987:Q3) the advance GDP estimate was released with an average lag of about 20 days. Beginning in 1987:Q4 the release date was shifted back by a week. A second, smaller shift appears to have occurred in 1996, so that the average lag is now in excess of 30 days.[17] These shifts suggest that since 1988 the Commerce Department has been exercising more care in the preparation of its advance GDP estimates, and that the advance estimates of the late 1980s and early 1990s incorporate more complete information than do the advance estimates of the early-and-mid 1980s. Rows 2, 3, and 4 of Table 5 present evidence consistent with this conjecture. These rows show what happens when the sample period for the efficiency-test regression is split in two, with 1987:Q4 as the dividing point. Quite clearly, the weight attached to the Commerce Department's estimate rises relative to that attached to our model's forecasts as the sample period is extended. Our model's forecast is statistically significant in the late sample period, but its coefficient is cut nearly in half.

In summary, the information content of the Commerce Department's advance GDP estimate has increased, over the years, relative to that of our model's forecasts. However, this increase in relative information content has come at a price. During most of the 1980s, the

---

[16] Trehan pits the advance GNP estimate against what appear to be in-sample predictions from his forecasting model. The GNP-GDP distinction is inconsequential for his results.

[17] The 54-day delay in the release of the initial estimate of 95:Q4 GDP was due to the January, 1996 government funding crisis.

Commerce Department's advance estimate was released at about the same time that our 3-month forecast would have been available. Now, the advance estimate is typically not available until fully two weeks after our forecast.

Forecasting Chain-Weight GDP. Real-time Commerce Department chain-weight national income accounts data are available for only a few years, complicating the estimation and evaluation of forecasting models for chain-weight GDP. We experimented with several approaches to estimating such forecasting models. Ultimately, we decided to handle the switch from fixed-weight to chain-weight GDP exactly as if it were a change in the base year of the fixed-weight GDP statistics. Thus, when estimating a model designed to predict the Commerce Department's advance chain-weight GDP release, each of our samples contains nothing but fixed weight data until 1994:Q3 (when advance chain-weight estimates first become available) and uses chain-weight data thereafter. When estimating a model designed to predict the Commerce Department's final chain-weight GDP release, each of our samples contains fixed-weight data through 1992, and chain-weight data from 1993:Q1 onward.

Figures 3A and 3B are the chain-weight counterparts of Figures 1A and 1B. They show actual GDP growth estimates along with forecasts generated by our 1-month, 2-month, and 3-month models of chain-weight GDP. Similarly, Table 6 is the chain-weight counterpart of Table 3. It gives the mean errors, mean absolute errors, and root-mean-square errors generated by our forecasting models. As before, our models are estimated using only real-time data and forecasts are obtained by substituting real time data into the right-hand sides of the estimated equations. Then the sample period is extended by one quarter and the process is repeated.

Both qualitatively and quantitatively, results are little changed by the move from fixed-weight to chain-weight GDP. Comparing Tables 3 and 6, mean absolute errors and root-mean-square errors are quite similar. Moreover, Table 6, like Table 3, suggests that it is generally easier to predict the advance GDP release than to predict the final GDP release; Table 6, like Table 3, suggests that obtaining a second month of current-quarter data has a much larger impact on forecast accuracy than does obtaining a third month of current-quarter data; and Table 6, like Table 3, shows that in predicting the final GDP release, even our 3-month model is no match for the Commerce Department's advance estimate.

Table 7 presents tests of the efficiency with which our models predict chain-weight GDP. Here, as in Table 4, we regress actual GDP growth on a constant and our forecast of GDP growth. Forecasts are efficient if the estimated constant is not significantly different from 0 and the estimated slope coefficient is not significantly different from 1. For the 2-month and 3-month models, efficiency cannot be rejected. However, the constant term in the 1-month regressions is too large to be consistent with efficiency.

Unfortunately, few analysts bothered to forecast chain-weight GDP until 1996, leaving us with too short a track record to meaningfully compare our models' predictions to the real-time predictions of others.

Stability of the Forecasting Models. In an effort to test the stability of our forecasting models, we estimated a series of regressions in which we included one or more dummy variables on the right-hand side of our forecasting equations. Specifically, for each model we estimated one regression in which we included a separate dummy variable for each quarter of our out-of-sample forecast period, and another regression in which we included a single dummy variable defined to equal 1 over the *entire* out-of-sample forecast period. The joint significance of the quarterly dummies in the first regression is a test of whether or not the model's out-of-sample forecasting performance is significantly poorer than its in-sample performance (Dufour 1980). The t statistic of the single dummy in the second regression provides a test for systematic bias in the out-of-sample forecasts of the model.

In Table 8, the probability values for F tests of the joint significance of the separate quarterly dummies are reported in the rows labeled "Dufour," while the P values for the t test of the single dummies are reported in the rows labeled "Single." None of the P values falls below the 0.05 cutoff for statistical significance. The only test statistic that comes close to statistical significance is that for the single dummy in the 3-month model of final, fixed-weight GDP. The suggestion is that the out-of-sample forecasts of this model may exhibit systematic bias.

To provide the reader with an alternative, informal sense of how stable our models are during the 1990s, Table 8 also reports two root-mean-square error statistics for each model.

Specifically, we compare the root-mean-square error that each of our models would have generated had we held its coefficients fixed over the out-of-sample forecast period to the root-mean-square errors that the same model generates when we allow quarter-by-quarter re-estimation of the forecast equations. The first of these root-mean-square errors is labeled "RMS." The second is labeled "Rolling RMSE." For a given model, when these two numbers are close, re-estimation of the model's coefficients is not important to its out-of-sample forecast performance. Without exception, the two root-mean-square errors are within 10% of one another.

Concluding Remarks

The results of this paper are generally encouraging. They suggest that a simple forecasting model is capable of matching the near-term GDP forecasting performance of private analysts (as captured in the Blue Chip consensus forecast). The key to successful forecasting is that the forecasting equations be estimated with real-time data. By this we mean that at each date within each sample period, the model-builder must not use any data that would have been unavailable to an analyst at the time. In our estimations, for example, whenever we are predicting 1985:Q1 GDP growth, it is always using only employment, sales, and industrial production data that were released within the first quarter of 1985 (or, in the case of our 3-month model, released within a few weeks of the end of the first quarter). Most forecasting models are not estimated in this way. Instead, analysts estimate and re-estimate their models using the most up-to-date data.

Clearly, there is room for improvement in our model and our estimation procedures. We have not made more than a cursory effort to search over alternative coincident indicators of real activity. We have made no effort at all to include leading indicators in our analysis--an omission that especially limits the performance of our 1-month model. Finally, we have not imposed any of the cross-equation restrictions that might be expected to improve the efficiency of our estimations.

Of necessity, our forecast comparisons are limited to fixed-weight measures of GDP. With the passage of time, it should be possible to extend these comparisons to the new chain-weight measures.

References

Blue Chip Economic Indicators (various issues), Capitol Publications, Alexandria, VA.

Braun, Steven N. (1990) "Estimation of Current-Quarter Gross National Product by Pooling Preliminary Labor-Market Data," Journal of Business and Economic Statistics 8, pp. 293-304.

Dufour, Jean-Marie (1980) "Dummy Variables and Predictive Tests for Structural Change," *Economics Letters* 41, pp. 241-7.

Ehrbeck, Tilman and Robert Waldmann (1996) "Why Are Professional Forecasters Biased? Agency versus Behavioral Explanations," *Quarterly Journal of Economics* 111, pp. 21-40.

Fitzgerald, Terry J. and Preston J. Miller (1989) "A Simple Way to Estimate Current-Quarter GNP," Federal Reserve Bank of Minneapolis *Quarterly Review*, Fall, pp. 27-31.

Graham, John R. (1996) "Is a Group of Economists Better than One? Than None?" *Journal of Business* 69, pp. 193-232.

Koenig, Evan F. (1996) "Capacity Utilization as a Real-Time Predictor of Manufacturing Output," Federal Reserve Bank of Dallas *Economic Review*, Third Quarter, pp. 16-23.

Lamont, Owen (1995) "Macroeconomic Forecasts and Microeconomic Forecasters," NBER Working Paper #5284.

Laster, David, Paul Bennett, and In Sun Geoum (1997) "Rational Bias in Economic Forecasts," Federal Reserve Bank of New York *Staff Reports*, no. 21.

McNees, Stephen K. (1987) "Consensus Forecasts: Tyranny of the Majority," *New England Economic Review*, Nov./Dec., pp. 15-21.

Miller, Preston J. and Daniel M. Chin (1996) "Using Monthly Data to Improve Quarterly Model Forecasts," Federal Reserve Bank of Minneapolis *Quarterly Review*, Spring, pp. 16-33.

Rathjens, Peter and Russell P. Robins (1993) "Forecasting Quarterly Data Using Monthly Information," *Journal of Forecasting* 12, pp. 321-330.

Trehan, Bharat (1989) "Forecasting Growth in Current Quarter Real GNP," Federal Reserve Bank of San Francisco *Economic Review*, Winter, pp. 39-52.

___ (1992) "Predicting Contemporaneous Output," Federal Reserve Bank of San Francisco *Economic Review*, No. 2, pp. 3-11.

Zadrozny, Peter A. (1990) "Forecasting U.S. GNP at Monthly Intervals with an Estimated Bivariate Time Series Model," Federal Reserve Bank of Atlanta *Economic Review*, Nov./Dec., pp. 2-15.

TABLE 1. Data Used for Predicting 1997;Q1 Real GDP Growth

	L-H-S Variable	Right-Hand-Side Variables				
	ΔGDP	ΔGDP	δEmployment	δRetail Sales	δIndustrial Prod.	δCPI
1-Month Model						
Data included:	97:Q1	96:Q1-96:Q4	96:09-96:12, 97:01	96:09-96:12, 97:01	96:09-96:12, 97:01	96:09-96:12, 97:01
Release Date:	4-28-97 Adv. 6-30-97 Final	1-31-97	2-7-97	2-15-97	2-15-97	2-19-97
2-Month Model						
Data included:	97:Q1	96:Q1-96:Q4	96:10-96:12, 97:01-97:02	96:10-96:12, 97:01-97:02	96:10-96:12, 97:01-97:02	96:10-96:12, 97:01-97:02
Release Date:	4-28-97 Adv. 6-30-97 Final	2-28-97	3-7-97	3-13-97	3-13-97	3-19-97
3-Month Model						
Data included:	97:Q1	96:Q1-96:Q4	96:11-96:12, 97:01-97:03	96:11-96:12, 97:01-97:03	96:11-96:12, 97:01-97:03	96:11-96:12, 97:01-97:03
Release Date:	4-28-97 Adv. 6-30-97 Final	3-28-97	4-4-97	4-11-97	4-16-97	4-16-97

TABLE 2A. Summary of Estimation Results--1 Month of Current-Quarter Data

	Predicting Advance Fixed-Wt. GDP		Predicting Final Fixed-Wt. GDP	
	1980:Q1-1989:Q4	1980:Q1-1996:Q4	1980:Q1-1989:Q4	1980:Q1-1996:Q4
Employment				
Joint Signif.	0.061	0.125	0.161	0.181
Sum of Coeff.	0.040	0.208	-0.201	0.055
Signif. of Sum	0.906	0.432	0.598	0.850
Industrial Prod.				
Joint Signif.	0.231	0.257	0.219	0.357
Sum of Coeff.	0.155	0.140	0.156	0.120
Signif. of Sum	0.153	0.125	0.198	0.232
Real Retail Sales				
Joint Signif.	0.019	0.000	0.022	0.001
Sum of Coeff.	0.170	0.184	0.122	0.157
Signif. of Sum	0.042	0.004	0.178	0.026
Overall				
Adjusted R^2	0.820	0.725	0.768	0.672
Std. Error of Est.	1.484	1.537	1.662	1.698
Significance of F	0.000	0.000	0.000	0.000
Significance of Q	0.855	0.722	0.870	0.342

TABLE 2B. Summary of Estimation Results--2 Months of Current-Quarter Data

	Predicting Advance Fixed-Wt. GDP		Predicting Final Fixed-Wt. GDP	
	1980:Q1-1989:Q4	1980:Q1-1996:Q4	1980:Q1-1989:Q4	1980:Q1-1996:Q4
Employment				
Joint Signif.	0.345	0.033	0.533	0.057
Sum of Coeff.	0.359	0.291	0.301	0.289
Signif. of Sum	0.290	0.110	0.464	0.188
Industrial Prod.				
Joint Signif.	0.111	0.000	0.034	0.000
Sum of Coeff.	0.271	0.260	0.280	0.263
Signif. of Sum	0.002	0.000	0.008	0.000
Real Retail Sales				
Joint Signif.	0.013	0.000	0.065	0.000
Sum of Coeff.	0.159	0.145	0.101	0.120
Signif. of Sum	0.017	0.000	0.188	0.014
Overall				
Adjusted R^2	0.832	0.847	0.742	0.780
Std. Error of Est.	1.433	1.146	1.751	1.390
Significance of F	0.000	0.000	0.000	0.000
Significance of Q	0.669	0.270	0.022	0.044

TABLE 2C. Summary of Estimation Results--3 Months of Current-Quarter Data

	Predicting Advance Fixed-Wt. GDP		Predicting Final Fixed-Wt. GDP	
	1980:Q1-1989:Q4	1980:Q1-1996:Q4	1980:Q1-1989:Q4	1980:Q1-1996:Q4
Employment				
Joint Signif.	0.365	0.083	0.201	0.033
Sum of Coeff.	0.421	0.307	0.711	0.478
Signif. of Sum	0.291	0.135	0.087	0.043
Industrial Prod.				
Joint Signif.	0.141	0.000	0.283	0.011
Sum of Coeff.	0.229	0.249	0.156	0.192
Signif. of Sum	0.024	0.000	0.117	0.002
Real Retail Sales				
Joint Signif.	0.074	0.000	0.142	0.001
Sum of Coeff.	0.138	0.130	0.102	0.118
Signif. of Sum	0.011	0.000	0.057	0.002
Overall				
Adjusted R^2	0.801	0.842	0.789	0.800
Std. Error of Est.	1.558	1.163	1.585	1.326
Significance of F	0.000	0.000	0.000	0.000
Significance of Q	0.104	0.132	0.916	0.448

TABLE 3A. Summary Statistics for Out-of-Sample Forecasting Exercise, 90:Q1 - 95:Q3

Forecast	Release Date: Month, Week	Predicting Advance Fixed-Weight GDP			Predicting Final Fixed-Weight GDP		
		Mean Error	Mean Ab. Er.	RMSE	Mean Error	Mean Ab. Er.	RMSE
BC	M1, W2	-0.02	1.06	1.32	0.05	1.32	1.58
MC	M2, W1	-0.43	1.04	1.36	---	---	---
FM	M2, W1	-0.01	1.34	1.58	0.19	1.51	1.93
BC	M2, W2	0.12	1.05	1.28	0.19	1.29	1.56
1-Month KD	**M2, W3**	**0.07**	**1.28**	**1.63**	**0.16**	**1.60**	**1.94**
MC	M3, W1	0.10	0.99	1.34	---	---	---
FM	M3, W1	0.40	1.22	1.53	0.67	1.40	1.81
BC	M3, W2	0.23	0.97	1.14	0.30	1.16	1.44
2-Month KD	**M3, W3**	**0.07**	**0.79**	**0.93**	**0.31**	**1.10**	**1.24**
MC	M4, W1	0.10	0.92	1.15	---	---	---
FM	M4, W1	0.24	0.96	1.27	0.49	1.20	1.57
BC	M4, W2	0.29	0.76	0.89	0.36	0.94	1.20
3-Month KD	**M4, W3**	**0.26**	**0.63**	**0.82**	**0.56**	**0.98**	**1.23**
Advance	M4, W4,5	---	---	---	0.07	0.56	0.64

Notes:

"BC" is the Blue Chip consensus forecast published during the second week of each month.

"MC" is the Miller-Chin real-time forecast of current-quarter advance GDP, available in the first week of each month from the second month of the quarter through the first month of the following quarter.

"FM" is the Fitzgerald-Miller real-time forecast of current-quarter GDP, available in the first week of each month from the second month of the quarter through the first month of the following quarter.

"KD" is the Koenig-Dolmas real-time forecast of current-quarter GDP, available in the third week of each month from the second month of the quarter through the first month of the following quarter.

"Advance" is the Commerce Department's advance (first) estimate of real GDP growth, available one full month after the close of the quarter.

TABLE 3B. Summary Statistics for Out-of-Sample Forecasting Exercise, 90:Q1 - 95:Q3

Forecast	Predicting Advance Fixed-Weight GDP			Predicting Final Fixed-Weight GDP		
	Mean Error	Mean Ab. Er.	RMSE	Mean Error	Mean Ab. Er.	RMSE
1-Month						
KD	0.07	1.28	1.63	0.16	1.60	1.94
KD (rev.)	0.65	1.31	1.58	0.73	1.58	1.91
Naive KD	---	---	---	-0.39	1.58	1.97
2-Month						
KD	0.07	0.79	0.93	0.31	1.10	1.24
KD (rev.)	0.40	1.14	1.37	0.47	1.42	1.67
Naive KD	---	---	---	-0.47	1.48	1.83
3-Month						
KD	0.26	0.63	0.82	0.56	0.98	1.23
KD (rev.)	0.47	1.00	1.28	0.54	1.26	1.50
Naive KD	---	---	---	-0.34	1.39	1.75

Notes:

"KD" is the Koenig-Dolmas real-time forecast of current-quarter GDP.

"KD (rev.)" is the Koenig-Dolmas model estimated with today's data, and used to forecast GDP in real time.

"Naive KD" is the Koenig-Dolmas model estimated with today's data, and used to forecast GDP growth as currently estimated.

TABLE 4A. Tests of Efficiency and Marginal Predictive Power
1 Month of Current-Quarter Data, 90:Q1 - 95:Q3

Predicting Advance Fixed-Weight GDP					Predicting Final Fixed-Weight GDP				
Constant	KD	Blue Chip	BC + 1	KD(rev.)	Constant	KD	Blue Chip	BC + 1	KD(rev.)
0.758 (0.533)	0.666+ (0.202)	---	---	---	0.779 (0.654)	0.695* (0.250)	---	---	---
-0.185 (0.528)	---	1.152+ (0.225)	---	---	-0.421 (0.628)	---	1.306+ (0.268)	---	---
-0.018 (0.414)	---	---	1.130+ (0.178)	---	-0.198 (0.511)	---	---	1.264+ (0.219)	---
1.111+ (0.351)	---	---	---	0.689+ (0.141)	1.151* (0.452)	---	---	---	0.711+ (0.182)
-0.271 (0.532)	0.230 (0.210)	0.961+ (0.284)	---	---	-0.605 (0.642)	0.269 (0.230)	1.124+ (0.308)	---	---
-0.083 (0.434)	0.117 (0.193)	---	1.038+ (0.236)	---	-0.355 (0.544)	0.190 (0.216)	---	1.142+ (0.260)	---
1.033* (0.470)	0.070 (0.272)	---	---	0.644+ (0.226)	0.962 (0.598)	0.158 (0.321)	---	---	0.619* (0.263)

Notes:

* Significant at the 5% level.

+ Significant at the 1% level.

"KD" is the Koenig-Dolmas real-time forecast of current-quarter GDP.

"Blue Chip" is the Blue Chip consensus forecast published in the month during which the Koenig-Dolmas forecast becomes available.

"BC + 1" is the Blue Chip consensus forecast published in the month *following* the availability of the Koenig-Dolmas forecast.

"KD(rev.)" is the Koenig-Dolmas model estimated with today's data, but used to forecast GDP in real time.

TABLE 4B. Tests of Efficiency and Marginal Predictive Power
2 Months of Current-Quarter Data, 90:Q1 - 95:Q3

Predicting Advance Fixed-Weight GDP					**Predicting Final Fixed-Weight GDP**				
Constant	KD	Blue Chip	BC + 1	KD(rev.)	Constant	KD	Blue Chip	BC + 1	KD(rev.)
0.496 (0.245)	0.791+ (0.082)	---	---	---	0.467 (0.357)	0.917+ (0.130)	---	---	---
-0.018 (0.414)	---	1.130+ (0.178)	---	---	-0.198 (0.511)	---	1.264+ (0.219)	---	---
0.158 (0.281)	---	---	1.070+ (0.116)	---	0.003 (0.371)	---	---	1.194+ (0.153)	---
0.977+ (0.281)	---	---	---	0.666+ (0.096)	0.995* (0.388)	---	---	---	0.698+ (0.132)
0.380 (0.324)	0.708+ (0.169)	0.151 (0.269)	---	---	-0.010 (0.428)	0.637+ (0.197)	0.530 (0.291)	---	---
0.290 (0.271)	0.456 (0.229)	---	0.487 (0.312)	---	0.048 (0.352)	0.410 (0.222)	---	0.748* (0.281)	---
0.527 (0.254)	0.690+ (0.188)	---	---	0.102 (0.171)	0.483 (0.356)	0.723+ (0.219)	---	---	0.204 (0.186)

Notes:

* Significant at the 5% level.

+ Significant at the 1% level.

"KD" is the Koenig-Dolmas real-time forecast of current-quarter GDP.

"Blue Chip" is the Blue Chip consensus forecast published in the month during which the Koenig-Dolmas forecast becomes available.

"BC + 1" is the Blue Chip consensus forecast published in the month *following* the availability of the Koenig-Dolmas forecast.

"KD(rev.)" is the Koenig-Dolmas model estimated with today's data, but used to forecast GDP in real time.

TABLE 4C. Tests of Efficiency and Marginal Predictive Power
3 Months of Current-Quarter Data, 90:Q1 - 95:Q3

Predicting Advance Fixed-Weight GDP					Predicting Final Fixed-Weight GDP				
Constant	KD	Blue Chip	Advance	KD(rev.)	Constant	KD	Blue Chip	Advance	KD(rev.)
0.451 (0.228)	0.898+ (0.085)	---	---	---	0.213 (0.329)	1.210+ (0.144)	---	---	---
0.158 (0.281)	---	1.070+ (0.116)	---	---	0.003 (0.371)	---	1.194+ (0.153)	---	---
---	---	---	---	---	-0.170 (0.198)	---	---	1.114+ (0.070)	---
0.923+ (0.282)	---	---	---	0.727+ (0.103)	0.877* (0.364)	---	---	---	0.798+ (0.133)
0.312 (0.257)	0.633* (0.246)	0.344 (0.300)	---	---	0.045 (0.344)	0.755* (0.357)	0.497 (0.359)	---	---
---	---	---	---	---	-0.198 (0.189)	0.273 (0.154)	---	0.917+ (0.130)	---
0.403 (0.242)	1.057+ (0.248)	---	---	-0.149 (0.219)	0.216 (0.351)	1.201+ (0.346)	---	---	0.007 (0.252)

Notes:

* Significant at the 5% level.

+ Significant at the 1% level.

"KD" is the Koenig-Dolmas real-time forecast of current-quarter GDP.

"Blue Chip" is the Blue Chip consensus forecast published in the month during which the Koenig-Dolmas forecast becomes available.

"Advance" is the Commerce Department's advance (first) estimate of real GDP growth.

"KD(rev.)" is the Koenig-Dolmas model estimated with today's data, but used to forecast GDP in real time.

TABLE 5. Does Our 3-Month Model Contain Information Beyond that in the Commerce Department's Advance GDP Estimate? Additional Tests

Final GDP Growth Regressed on a Constant and Alternative Forecasts

Sample Period	Constant	KD (3-month)	Advance GDP
80:Q1-96:Q4	-0.033 (0.156)	0.511+ (0.106)	0.501+ (0.099)
87:Q4-96:Q4	-0.176 (0.272)	0.386* (0.172)	0.739+ (0.151)
80:Q1-87:Q3	-0.211 (0.204)	0.693+ (0.133)	0.298* (0.126)
Coeff. Change (row2 - row3)	0.035 (0.340)	-0.307 (0.218)	0.440* (0.196)

* Significant at the 5% level. + Significant at the 1% level.

TABLE 6. Summary Statistics for Out-of-Sample Forecasting Exercise--Chain-Weight GDP

	Predicting Advance Chain-Weight GDP			Predicting Final Chain-Weight GDP		
1-Month:	Mean Error	Mean Abs. Er.	RMSE	Mean Error	Mean Abs. Er.	RMSE
KD 94:3 - 96:4	0.78	1.41	1.79	1.25	1.39	1.86
KD 93:1 - 96:4	---	---	---	0.82	1.41	1.82
2-Month:						
KD 94:3 - 96:4	-0.24	0.88	1.16	0.10	0.90	1.01
KD 93:1 - 96:4	---	---	---	0.14	1.08	1.28
3-Month:						
KD 94:3 - 96:4	0.04	0.75	0.86	0.27	0.80	1.12
Advance 94:3 - 96:4	---	---	---	-0.03	0.53	0.58
KD 93:1 - 96:4	---	---	---	0.23	0.93	1.19

TABLE 7. Tests of Predictive Efficiency--Chain-Weight GDP

	Predicting Advance GDP, 94:Q3-96:Q4		Predicting Final GDP, 93:Q1 - 96:Q4	
	Constant	KD	Constant	KD
1-Month	2.10* (0.78)	0.38 (0.43)	1.87* (0.73)	0.47 (0.31)
2-Month	0.74 (0.80)	0.66* (0.25)	0.78 (0.73)	0.76+ (0.25)
3-Month	-0.15 (0.83)	1.07+ (0.29)	-0.00 (0.86)	1.09+ (0.31)

* Significant at 5% level. + Significant at 1% level.

TABLE 8. **Testing the Stability of the Forecasting Model**

1-Month:	Advance Fixed-Weight GDP, 90:Q1-96:Q4 Mean	P-Value	RMS	Rolling RMSE	Final Fixed-Weight GDP, 90:Q1-96:Q4 Mean	P-Value	RMS	Rolling RMSE
Dufour	0.36	0.401	1.86	1.82	0.22	0.442	2.15	2.03
Single	0.41	0.394	---	---	0.55	0.301	---	---
2-Month:								
Dufour	0.15	0.990	1.02	1.01	0.37	0.993	1.23	1.22
Single	0.18	0.606	---	---	0.47	0.267	---	---
3-Month:								
Dufour	0.34	1.000	0.85	0.87	0.67	0.961	1.34	1.24
Single	0.36	0.299	---	---	0.69	0.079	---	---

1-Month:	Advance Chain-Weight GDP, 94:Q3-96:Q4 Mean	P-Value	RMS	Rolling RMSE	Final Chain-Weight GDP, 93:Q1-96:Q4 Mean	P-Value	RMS	Rolling RMSE
Dufour	0.75	0.284	1.76	1.79	0.87	0.537	1.92	1.82
Single	0.77	0.184	---	---	0.75	0.169	---	---
2-Month:								
Dufour	-0.34	0.622	1.16	1.16	0.03	0.849	1.27	1.32
Single	-0.38	0.422	---	---	0.13	0.779	---	---
3-Month:								
Dufour	-0.01	0.934	0.86	0.86	0.19	0.823	1.17	1.19
Single	-0.09	0.844	---	---	0.27	0.522	---	---

Figure 1A. Advance Fixed-Weight GDP

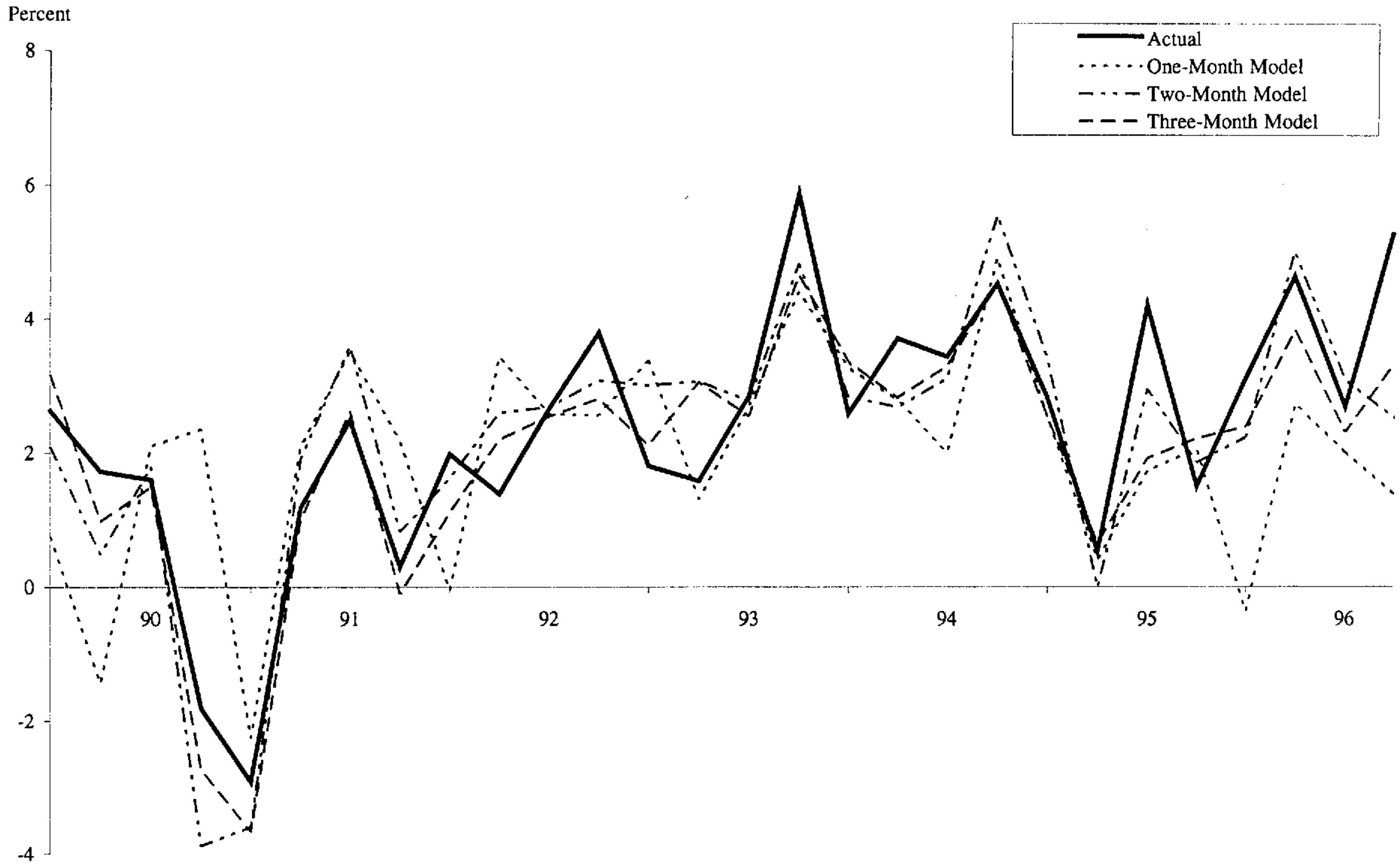


Figure 1B. Final Fixed-Weight GDP

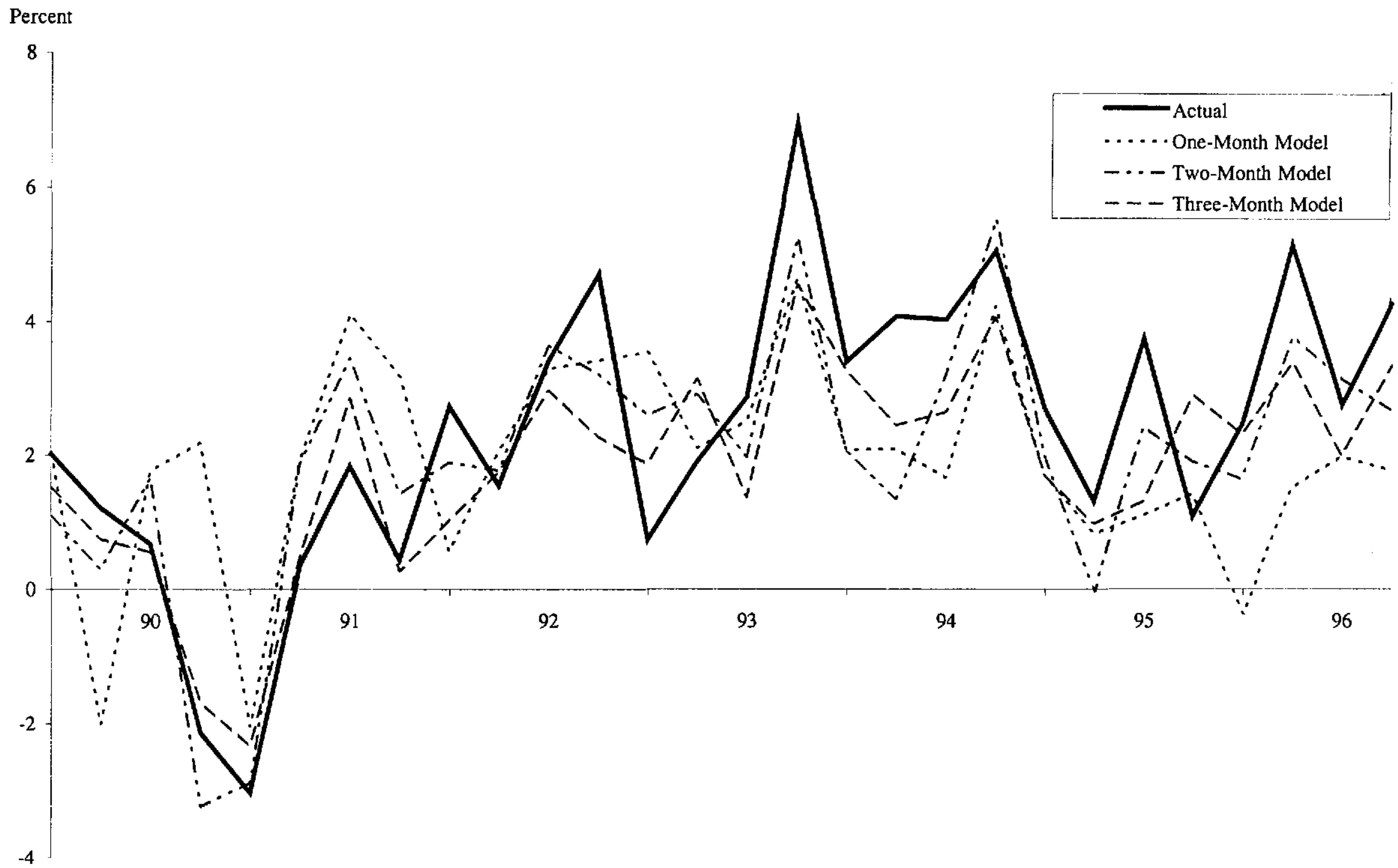


Figure 2. Timing of Advance GDP Release
Number of Days after Quarter Ends

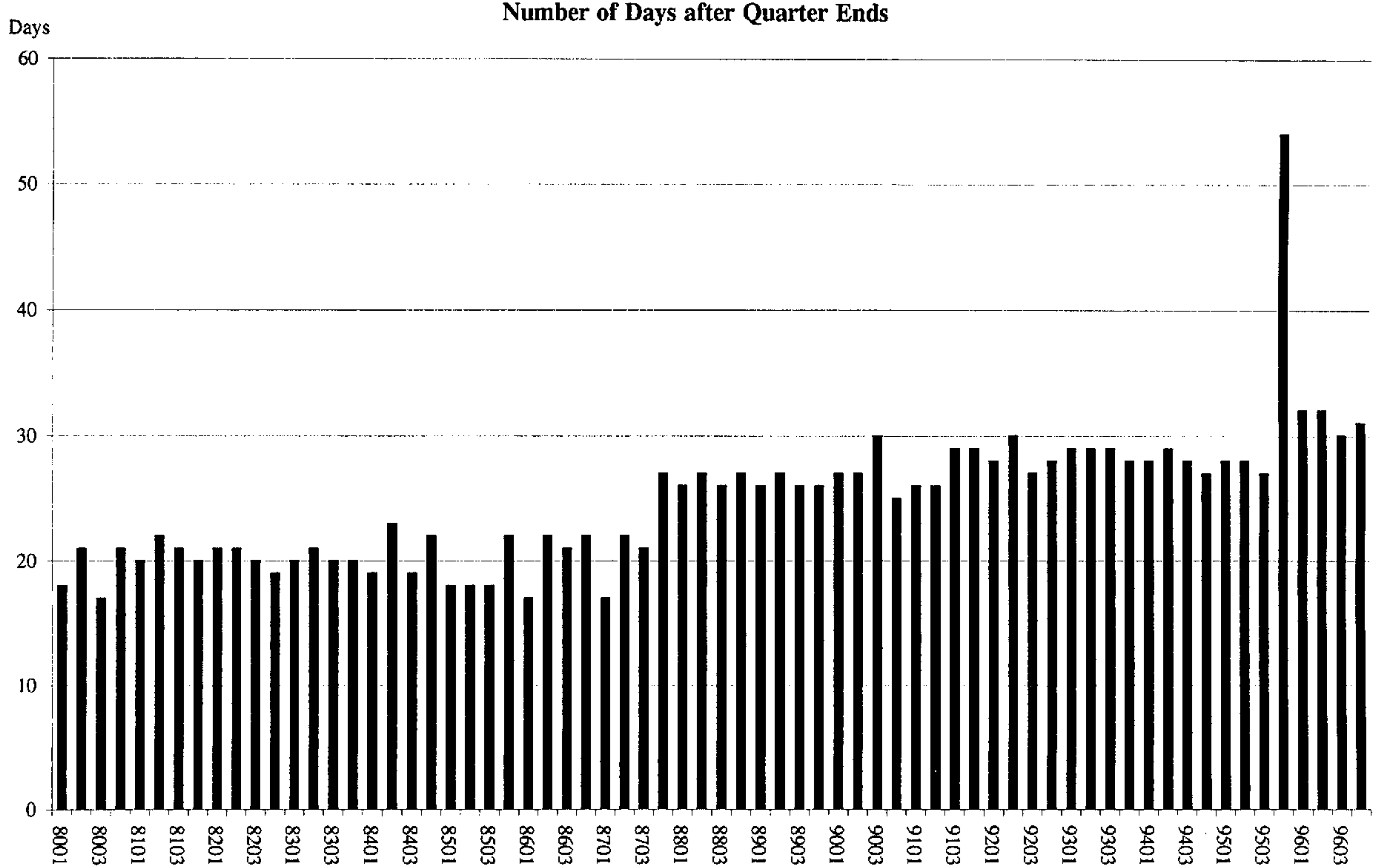
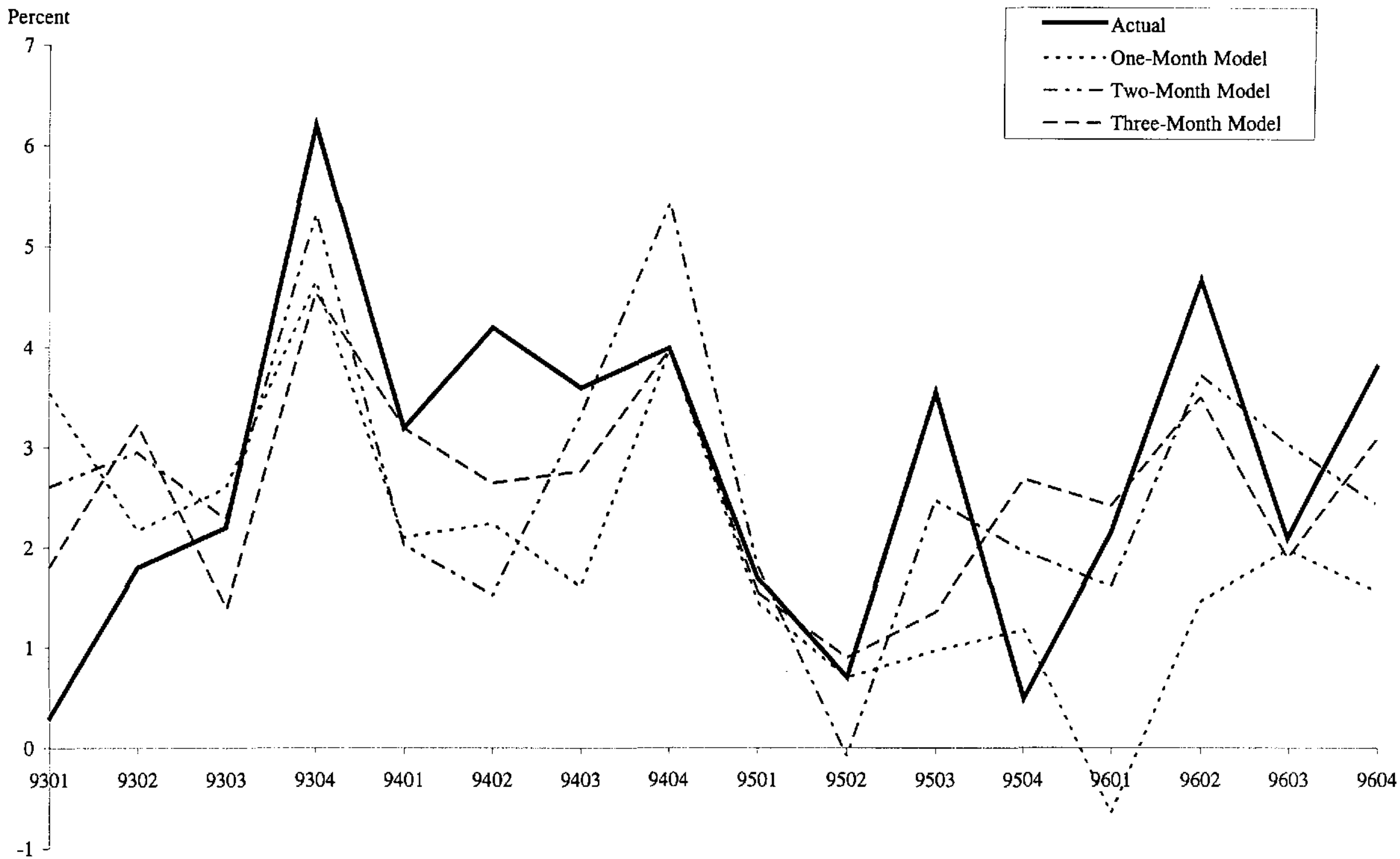


Figure 3B. Final Chain-Weight GDP



Research Papers Presented at the
1994 Texas Conference on Monetary Economics
April 23-24, 1994
held at the Federal Reserve Bank of Dallas, Dallas, Texas

Available, at no charge, from the Research Department
Federal Reserve Bank of Dallas, P. O. Box 655906
Dallas, Texas 75265-5906

Please check the titles of the Research Papers you would like to receive:

____ 1 A Sticky-Price Manifesto (Laurence Ball and N. Gregory Mankiw)

____ 2 Sequential Markets and the Suboptimality of the Friedman Rule (Stephen D. Williamson)

____ 3 Sources of Real Exchange Rate Fluctuations: How Important Are Nominal Shocks? (Richard Clarida and Jordi Gali)

____ 4 On Leading Indicators: Getting It Straight (Mark A. Thoma and Jo Anna Gray)

____ 5 The Effects of Monetary Policy Shocks: Evidence From the Flow of Funds (Lawrence J. Christiano, Martin Eichenbaum and Charles Evans)

Name:	Organization:
Address:	City, State and Zip Code:
Please add me to your mailing list to receive future Research Papers:	Yes ____ No ____

RESEARCH PAPERS OF THE RESEARCH DEPARTMENT
FEDERAL RESERVE BANK OF DALLAS

Available, at no charge, from the Research Department
Federal Reserve Bank of Dallas, P. O. Box 655906
Dallas, Texas 75265-5906

Please check the titles of the Research Papers you would like to receive:

___ 9201 Are Deep Recessions Followed by Strong Recoveries? (Mark A. Wynne and Nathan S. Balke)
___ 9202 The Case of the "Missing M2" (John V. Duca)
___ 9203 Immigrant Links to the Home Country: Implications for Trade, Welfare and Factor Rewards (David M. Gould)
___ 9204 Does Aggregate Output Have a Unit Root? (Mark A. Wynne)
___ 9205 Inflation and Its Variability: A Note (Kenneth M. Emery)
___ 9206 Budget Constrained Frontier Measures of Fiscal Equality and Efficiency in Schooling (Shawna Grosskopf, Kathy Hayes, Lori L. Taylor, William Weber)
___ 9207 The Effects of Credit Availability, Nonbank Competition, and Tax Reform on Bank Consumer Lending (John V. Duca and Bonnie Garrett)
___ 9208 On the Future Erosion of the North American Free Trade Agreement (William C. Gruben)
___ 9209 Threshold Cointegration (Nathan S. Balke and Thomas B. Fomby)
___ 9210 Cointegration and Tests of a Classical Model of Inflation in Argentina, Bolivia, Brazil, Mexico, and Peru (Raul Anibal Feliz and John H. Welch)
___ 9211 Nominal Feedback Rules for Monetary Policy: Some Comments (Evan F. Koenig)
___ 9212 The Analysis of Fiscal Policy in Neoclassical Models[1] (Mark Wynne)
___ 9213 Measuring the Value of School Quality (Lori Taylor)
___ 9214 Forecasting Turning Points: Is a Two-State Characterization of the Business Cycle Appropriate? (Kenneth M. Emery & Evan F. Koenig)
___ 9215 Energy Security: A Comparison of Protectionist Policies (Mine K. Yücel and Carol Dahl)
___ 9216 An Analysis of the Impact of Two Fiscal Policies on the Behavior of a Dynamic Asset Market (Gregory W. Huffman)
___ 9301 Human Capital Externalities, Trade, and Economic Growth (David Gould and Roy J. Ruffin)
___ 9302 The New Face of Latin America: Financial Flows, Markets, and Institutions in the 1990s (John Welch)
___ 9303 A General Two Sector Model of Endogenous Growth with Human and Physical Capital (Eric Bond, Ping Wang, and Chong K. Yip)
___ 9304 The Political Economy of School Reform (S. Grosskopf, K. Hayes, L. Taylor, and W. Weber)
___ 9305 Money, Output, and Income Velocity (Theodore Palivos and Ping Wang)
___ 9306 Constructing an Alternative Measure of Changes in Reserve Requirement Ratios (Joseph H. Haslag and Scott E. Hein)
___ 9307 Money Demand and Relative Prices During Episodes of Hyperinflation (Ellis W. Tallman and Ping Wang)
___ 9308 On Quantity Theory Restrictions and the Signalling Value of the Money Multiplier (Joseph Haslag)
___ 9309 The Algebra of Price Stability (Nathan S. Balke and Kenneth M. Emery)
___ 9310 Does It Matter How Monetary Policy is Implemented? (Joseph H. Haslag and Scott Hein)
___ 9311 Real Effects of Money and Welfare Costs of Inflation in an Endogenously Growing Economy with Transactions Costs (Ping Wang and Chong K. Yip)
___ 9312 Borrowing Constraints, Household Debt, and Racial Discrimination in Loan Markets (John V. Duca and Stuart Rosenthal)
___ 9313 Default Risk, Dollarization, and Currency Substitution in Mexico (William Gruben and John Welch)
___ 9314 Technological Unemployment (W. Michael Cox)
___ 9315 Output, Inflation, and Stabilization in a Small Open Economy: Evidence from Mexico (John H. Rogers and Ping Wang)
___ 9316 Price Stabilization, Output Stabilization and Coordinated Monetary Policy Actions (Joseph H. Haslag)
___ 9317 An Alternative Neo-Classical Growth Model with Closed-Form Decision Rules (Gregory W. Huffman)
___ 9318 Why the Composite Index of Leading Indicators Doesn't Lead (Evan F. Koenig and Kenneth M. Emery)
___ 9319 Allocative Inefficiency and Local Government: Evidence Rejecting the Tiebout Hypothesis (Lori L. Taylor)
___ 9320 The Output Effects of Government Consumption: A Note (Mark A. Wynne)
___ 9321 Should Bond Funds be Included in M2? (John V. Duca)

___ 9322 Recessions and Recoveries in Real Business Cycle Models: Do Real Business Cycle Models Generate Cyclical Behavior? (Mark A. Wynne)
___ 9323* Retaliation, Liberalization, and Trade Wars: The Political Economy of Nonstrategic Trade Policy (David M. Gould and Graeme L. Woodbridge)
___ 9324 A General Two-Sector Model of Endogenous Growth with Human and Physical Capital: Balanced Growth and Transitional Dynamics (Eric W. Bond, Ping Wang,and Chong K. Yip)
___ 9325 Growth and Equity with Endogenous Human Capital: Taiwan's Economic Miracle Revisited (Maw-Lin Lee, Ben-Chieh Liu, and Ping Wang)
___ 9326 Clearinghouse Banks and Banknote Over-issue (Scott Freeman)
___ 9327 Coal, Natural Gas and Oil Markets after World War II: What's Old, What's New? (Mine K. Yücel and Shengyi Guo)
___ 9328 On the Optimality of Interest-Bearing Reserves in Economies of Overlapping Generations (Scott Freeman and Joseph Haslag)
___ 9329* Retaliation, Liberalization, and Trade Wars: The Political Economy of Nonstrategic Trade Policy (David M. Gould and Graeme L. Woodbridge) (Reprint of 9323 in error)
___ 9330 On the Existence of Nonoptimal Equilibria in Dynamic Stochastic Economies (Jeremy Greenwood and Gregory W. Huffman)
___ 9331 The Credibility and Performance of Unilateral Target Zones: A Comparison of the Mexican and Chilean Cases (Raul A. Feliz and John H. Welch)
___ 9332 Endogenous Growth and International Trade (Roy J. Ruffin)
___ 9333 Wealth Effects, Heterogeneity and Dynamic Fiscal Policy (Zsolt Becsi)
___ 9334 The Inefficiency of Seigniorage from Required Reserves (Scott Freeman)
___ 9335 Problems of Testing Fiscal Solvency in High Inflation Economies: Evidence from Argentina, Brazil, and Mexico (John H. Welch)
___ 9336 Income Taxes as Reciprocal Tariffs (W. Michael Cox, David M. Gould, and Roy J. Ruffin)
___ 9337 Assessing the Economic Cost of Unilateral Oil Conservation (Stephen P.A. Brown and Hillard G. Huntington)
___ 9338 Exchange Rate Uncertainty and Economic Growth in Latin America (Darryl McLeod and John H. Welch)
___ 9339 Searching for a Stable M2-Demand Equation (Evan F. Koenig)
___ 9340 A Survey of Measurement Biases in Price Indexes (Mark A. Wynne and Fiona Sigalla)
___ 9341 Are Net Discount Rates Stationary?: Some Further Evidence (Joseph H. Haslag, Michael Nieswiadomy, and D. J. Slottje)
___ 9342 On the Fluctuations Induced by Majority Voting (Gregory W. Huffman)
___ 9401 Adding Bond Funds to M2 in the P-Star Model of Inflation (Zsolt Becsi and John Duca)
___ 9402 Capacity Utilization and the Evolution of Manufacturing Output: A Closer Look at the "Bounce-Back Effect" (Evan F. Koenig)
___ 9403 The Disappearing January Blip and Other State Employment Mysteries (Frank Berger and Keith R. Phillips)
___ 9404 Energy Policy: Does it Achieve its Intended Goals? (Mine Yücel and Shengyi Guo)
___ 9405 Protecting Social Interest in Free Invention (Stephen P.A. Brown and William C. Gruben)
___ 9406 The Dynamics of Recoveries (Nathan S. Balke and Mark A. Wynne)
___ 9407 Fiscal Policy in More General Equilibriium (Jim Dolman and Mark Wynne)
___ 9408 On the Political Economy of School Deregulation (Shawna Grosskopf, Kathy Hayes, Lori Taylor, and William Weber)
___ 9409 The Role of Intellectual Property Rights in Economic Growth (David M. Gould and William C. Gruben)
___ 9410 U.S. Banks, Competition, and the Mexican Banking System: How Much Will NAFTA Matter? (William C. Gruben, John H. Welch and Jeffery W. Gunther)
___ 9411 Monetary Base Rules: The Currency Caveat (R. W. Hafer, Joseph H. Haslag, andScott E. Hein)
___ 9412 The Information Content of the Paper-Bill Spread (Kenneth M. Emery)
___ 9413 The Role of Tax Policy in the Boom/Bust Cycle of the Texas Construction Sector (D'Ann Petersen, Keith Phillips and Mine Yücel)
___ 9414 The P* Model of Inflation, Revisited (Evan F. Koenig)
___ 9415 The Effects of Monetary Policy in a Model with Reserve Requirements (Joseph H. Haslag)
___ 9501 An Equilibrium Analysis of Central Bank Independence and Inflation (Gregory W. Huffman)
___ 9502 Inflation and Intermediation in a Model with Endogenous Growth (Joseph H. Haslag)
___ 9503 Country-Bashing Tariffs: Do Bilateral Trade Deficits Matter? (W. Michael Cox and Roy J. Ruffin)
___ 9504 Building a Regional Forecasting Model Utilizing Long-Term Relationships and Short-Term Indicators (Keith R. Phillips and Chih-Ping Chang)

___ 9505 Building Trade Barriers and Knocking Them Down: The Political Economy of Unilateral Trade Liberalizations (David M. Gould and Graeme L. Woodbridge)
___ 9506 On Competition and School Efficiency (Shawna Grosskopf, Kathy Hayes, Lori L. Taylor and William L. Weber)
___ 9507 Alternative Methods of Corporate Control in Commercial Banks (Stephen Prowse)
___ 9508 The Role of Intratemporal Adjustment Costs in a Multi-Sector Economy (Gregory W. Huffman and Mark A. Wynne)
___ 9509 Are Deep Recessions Followed By Strong Recoveries? Results for the G-7 Countries (Nathan S. Balke and Mark A. Wynne)
___ 9510 Oil Prices and Inflation (Stephen P.A. Brown, David B. Oppedahl and Mine K. Yücel)
___ 9511 A Comparison of Alternative Monetary Environments (Joseph H. Haslag))
___ 9512 Regulatory Changes and Housing Coefficients (John V. Duca)
___ 9513 The Interest Sensitivity of GDP and Accurate Reg Q Measures (John V. Duca)
___ 9514 Credit Availability, Bank Consumer Lending, and Consumer Durables (John V. Duca and Bonnie Garrett)
___ 9515 Monetary Policy, Banking, and Growth (Joseph H. Haslag)
___ 9516 The Stock Market and Monetary Policy: The Role of Macroeconomic States (Chih-Ping Chang and Huan Zhang)
___ 9517 Hyperinflations and Moral Hazard in the Appropriation of Seigniorage: An Empirical Implementation With A Calibration Approach (Carlos E. Zarazaga)
___ 9518 Targeting Nominal Income: A Closer Look (Evan F. Koenig)
___ 9519 Credit and Economic Activity: Shocks or Propagation Mechanism? (Nathan S. Balke and Chih-Ping Chang)
___ 9601 The Monetary Policy Effects on Seignorage Revenue in a Simple Growth Model (Joseph H. Haslag)
___ 9602 Regional Productivity and Efficiency in the U.S.: Effects of Business Cycles and Public Capital (Dale Boisso, Shawna Grosskopf and Kathy Hayes)
___ 9603 Inflation, Unemployment, and Duration (John V. Duca)
___ 9604 The Response of Local Governments to Reagan-Bush Fiscal Federalism (D. Boisso, Shawna Grosskopf and Kathy Hayes)
___ 9605 Endogenous Tax Determination and the Distribution of Wealth (Gregory W. Huffman)
___ 9606 An Exploration into the Effects of Dynamic Economic Stabilization (Jim Dolmas and Gregory W. Huffman)
___ 9607 Is Airline Price Dispersion the Result of Careful Planning or Competitive Forces? (Kathy J. Hayes and Leola B. Ross)
___ 9608 Some Implications of Increased Cooperation in World Oil Conservation (Stephen P.A. Brown and Hillard G. Huntington)
___ 9609 An Equilibrium Analysis of Relative Price Changes and Aggregate Inflation (Nathan S. Balke and Mark A. Wynne)
___ 9610 What's Good for GM...? Using Auto Industry Stock Returns to Forecast Business Cycles and Test the Q-Theory of Investment (Gregory R. Duffee and Stephen Prowse)
___ 9611 Does the Choice of Nominal Anchor Matter? (David M. Gould)
___ 9612 The Policy Sensitivity of Industries and Regions (Lori L. Taylor and Mine K. Yücel)
___ 9613 Oil Prices and Aggregate Economic Activity: A Study of Eight OECD Countries (Stephen P.A. Brown, David B. Oppedahl and Mine K. Yücel)
___ 9614 The Effect of the Minimum Wage on Hours of Work (Madeline Zavodny)
___ 9615 Aggregate Price Adjustment: The Fischerian Alternative (Evan F. Koenig)
___ 9701 Nonlinear Dynamics and Covered Interest Rate Parity (Nathan S. Balke and Mark E. Wohar)
___ 9702 More on Optimal Denominations for Coins and Currency (Mark A. Wynne)
___ 9703 Specialization and the Effects of Transactions Costs on Equilibrium Exchange (James Dolmas and Joseph H. Haslag)
___ 9704 The Political Economy of Endogenous Taxation and Redistribution (Jim Dolmas and Gregory W. Huffman)
___ 9705 Inequality, Inflation, and Central Bank Independence (Jim Dolmas, Gregory W. Huffman, and Mark A. Wynne)
___ 9706 On The Political Economy of Immigration (Jim Dolmas and Gregory W. Huffman)
___ 9707 Business Cycles Under Monetary Union: EU and US Business Cycles Compared (Mark A. Wynne and Jahyeong Koo)
___ 9708 Allocative Inefficiency and School Competition (Shawna Grosskopf, Kathy Hayes, Lori L. Taylor and William L. Weber)

___ 9709 Goods-Market Competition and Profit Sharing: A Multisector Macro Approach (John V. Duca and David D. VanHoose)

___ 9710 Real-Time GDP Growth Forecasts (Evan F. Koenig and Sheila Dolmas)

Name:	Organization:
Address:	City, State and Zip Code:
Please add me to your mailing list to receive future Research Papers: ____ Yes ____ No	