THE USE AND ABUSE OF "REAL-TIME" DATA IN ECONOMIC FORECASTING

Evan F. Koenig
Sheila Dolmas
Jeremy Piger

Research Department
Working Paper 0004

FEDERAL RESERVE BANK OF DALLAS

THE USE AND ABUSE OF "REAL-TIME" DATA IN ECONOMIC FORECASTING

by

Evan F. Koenig,*
Sheila Dolmas,**
and Jeremy Piger†

Research Department
Federal Reserve Bank of Dallas
2200 N. Pearl St.
Dallas, TX 75201
e-mail: Evan.F.Koenig@dal.frb.org

August 2001

This paper had its origins in a GDP forecasting project undertaken jointly with Ken Emery. Helpful comments and suggestions were offered by Jim Stock, an anonymous referee, Nathan Balke, Todd Clark, Dean Croushore, Jon Faust, Preston Miller, John Robertson, Ellis Tallman, Jonathan Wright and attendees of the November 1999 meeting of the Federal Reserve System Committee on Macroeconomics. Dean Croushore and Ellis Tallman were kind enough to share their data and Dong Fu provided research assistance. Early drafts of the paper were completed when all three authors were affiliated with the Federal Reserve Bank of Dallas. The opinions expressed do not necessarily represent the views of the Federal Reserve Bank of Dallas, the Federal Reserve Bank of St. Louis, or the Federal Reserve System.

* Federal Reserve Bank of Dallas
** Unaffiliated
† Federal Reserve Bank of St. Louis

Weapaper2.091/8-17-01

Abstract

We distinguish between three different ways of using real-time data to estimate forecasting equations and argue that the most popular approach should generally be avoided. The point is illustrated with a model that uses monthly industrial production, employment, and retail sales data to predict real GDP growth. When the model is estimated using our preferred method, its out-of-sample forecasting performance is superior to that obtained using conventional estimation and compares favorably with that of the Blue-Chip consensus.

JEL: C13, C82

I. Introduction

Many economic time series are subject to revision. Revisions to measures of real economic activity–such as employment, sales, and production–are sometimes large, and may occur years after official figures are first released. Nevertheless, analysts typically use only data of the most recent vintage when estimating and evaluating their economic forecasting models. Current-vintage data are commonly used even in *ex-post* forecasting exercises that are labeled "out-of-sample." For example, 1999:Q1-vintage data (the most recent available at the time this paper was written) might be used to estimate recursive regressions and forecasts of real GDP growth running from 1990:Q1 through 1997:Q4. In general, the use of current-vintage data can lead an analyst to include variables on the right-hand-side of his forecasting equation that, in real time, have little marginal predictive power (Diebold and Rudebusch 1991, Swanson 1996). It can lead to an exaggerated assessment of the forecasting performance of a model relative to alternative models and relative to predictions that were actually available at the time (Fair and Shiller 1990, Orphanides 1999).

In those studies where the potential pitfalls of relying on current-vintage data have been taken seriously, a common response has been to use *end-of-sample*-vintage data for estimation and evaluation purposes instead of current-vintage data.[1] As the sample period over which the forecasting equation is estimated is extended, the vintage of the data used to estimate the equation is updated. Thus, rather than forecast 1990:Q1-1997:Q4 GDP growth using recursive regressions all estimated with 1999:Q1-vintage data, one uses the prediction of an equation estimated with 1990:Q1-vintage data to forecast GDP growth in 1990:Q1, the prediction of an equation estimated with 1990:Q2-vintage data to forecast GDP growth in 1990:Q2, and so forth. This procedure mimics the actual practice of many professional forecasters and provides a level playing field for comparing their performance with that of the model. However, there is reason to suspect that this "conventional approach" to real-time estimation and forecasting is often suboptimal. *A fortiori*, the predictions made by professional forecasters may also often be suboptimal.[2]

Rather than use end-of-sample-vintage data to estimate their forecasting equations, we argue that analysts should generally use data of as many different vintages as there are dates in their samples. More specifically, at every date within a sample, right-hand-side variables ought to be measured as they would have been at that time. We call these "real-time-vintage data." For example, when the left-hand-side variable is 1990:Q1 GDP growth, all right-hand-side variables should be measured as they appeared in 1990:Q1. Only 1990:Q1-vintage data should be used in forecasting 1990:Q1 GDP growth *regardless of whether or not the sample period extends beyond 1990:Q1*. Thus, we argue that when a data point is added to the end of the sample period, the data that appear on the right-hand-side of the equation earlier on in the sample ought *not* to be updated. We further argue that first-available official estimates should be used for the left-hand-side variable when estimating the forecasting equation. First-available estimates ought to be used on the equation's left-hand side for estimation purposes even if one is ultimately interested in predicting final-revised data.

The intuition underlying our arguments is simple. The empirical relationship between GDP growth and early estimates of (say) employment growth will typically be different from that between GDP growth and estimates of employment growth available several years after the fact. It is the former relationship that is of interest to the economist trying to forecast growth in aggregate output. The trouble with the conventional approach to real-time estimation is that the data on the right-hand-side of the forecasting equation range from extensively revised (early in the sample) to nearly unrevised (at the end of sample). What we call real-time-vintage estimation avoids this problem by including at each point in the sample only right-hand-side data that would have been available to a forecaster at that date.

What of our argument that initial-release data, rather than end-of-sample data, should be used for the left-hand-side variable (real GDP growth) when estimating the forecasting equation? The argument is based on the empirically plausible assumption that the government's initial data release is an efficient estimate of subsequent releases, meaning that revisions to the initial release are completely unpredictable using data available at the time of the initial release. If this assumption holds, unbiased estimates of the parameters linking true GDP growth to the real-time-vintage-right-hand-side variables will be obtained regardless of the vintage of the data on the left-hand side of the estimated equation. Less obviously, the parameter estimates obtained using initial-release data as the dependent variable will be more accurate, in finite samples, than those obtained using revised data. Any revised GDP release will incorporate information that cannot, under our efficiency assumption, be predicted using data available at the time of the initial release. In particular, any revised GDP release will incorporate information that is uncorrelated with the real-time-vintage right-hand-side variables. To the econometrician estimating the forecasting equation, this additional information is extraneous noise. Even if the government's initial estimates are not fully efficient, they may be good enough that our proposed approach will perform well in practice.

It might appear that collecting the data required to implement our version of real-time forecasting would be prohibitively difficult. However, for most variables of interest it is easier to obtain short data series of many vintages than it is to reconstruct long series of a few vintages (such as are required for real-time estimation using end-of-sample-vintage data).

The specific application we consider is forecasting same-quarter real GDP growth using monthly data on employment, industrial production, and retail sales.[3] Economists devote substantial time and effort to constructing early estimates of 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 Blue Chip consensus forecast of GDP growth has been 1.5 percentage points based on forecasts published in the first month after the quarter (about three weeks prior to the release of the government's official "advance" GDP estimate). The corresponding 95% confidence interval is 5.8 percentage points wide. Consensus forecasts are known to be more accurate than those of most individuals (Graham 1996, McNees 1987).

Despite the limited set of monthly indicators included in our model and our relatively short sample period, we are able to achieve an out-of-sample forecasting performance that is as good as, or better than, that of the Blue Chip consensus. An important contributor to our model's strong performance is the fact that we estimate it with real-time rather than end-of-sample-vintage data. When our model is estimated conventionally, its forecasting performance suffers.

II. The Forecasting Problem

Consider the problem of forecasting a single variable, y, using time-series observations on a $1 \times k$ vector, $\boldsymbol{x}$, of other variables (which might possibly include lagged values of y). Official estimates of both $\boldsymbol{x}$ and y are subject to revision. The initial official estimate of y is based on at least as much information as is available to the econometrician. We adopt the following notation:

$y(t) \equiv$ the "true" period-t value of y, $0 \leq t$

$\boldsymbol{x}(t) \equiv$ the "true" period-t value of $\boldsymbol{x}$, $0 \leq t$

$y(t)_s \equiv$ the official estimate of $y(t)$ released at time s, where $s \geq t$.

$\boldsymbol{x}(t)_s \equiv$ the official estimate of $\boldsymbol{x}(t)$ available at time $s \geq t$.

For concreteness, the reader may want to think of $y(t)_s$ as an official estimate of GDP growth. The econometrcian is trying to forecast GDP growth in quarter $T > 0$ ($y(T)$). No official estimate of quarter-T GDP growth has yet been released ($y(T)_T$ is, as yet, unavailable). However, an initial estimate of, (say) quarter-T employment growth *is* available and is included in the vector of variables observed by the econometrician ($\boldsymbol{x}(T)_T$).[4] Also available and included in $\boldsymbol{x}(T)_T$ might be partially revised official estimates of lagged GDP and employment growth.

Since the goal is to forecast $y(T)$ using $\boldsymbol{x}(T)_T$, it seems natural to posit that $y(t)$ and $\boldsymbol{x}(t)_t$ are linearly related:

$$y(t) = \boldsymbol{x}(t)_t\,\alpha + w(t), \tag{1}$$

where $w(t)$ captures both information available to the government at time t, but not to the econometrician, and information that is not available, even to the government, until after time t. From the perspective of the econometrician, $w(t)$ is a mean-zero disturbance orthogonal to $\boldsymbol{x}(t)_t$. While heteroscedasticity cannot be ruled out, there are no *a priori* reasons to expect it.

Given an estimate of α, the econometrician forecasts $y(T)$ in the obvious way, by setting

$$\hat{y}(T) \;=\; \boldsymbol{x}(T)_T\hat{\alpha}, \tag{2}$$

where hats are used to indicate forecasted or estimated values. The challenge lies in obtaining the best possible estimate of α. In practice, $y(t)$ is often observed with a substantial lag–if it is observable at all. Consequently, direct estimation of Equation 1 will typically be impossible except over a truncated sample period.[5]

To make progress on the estimation problem we must take a stand on how it is that the government goes about revising its official estimates of $y(t)$. The analysis that follows is based on the assumption that the government's initial official estimate of $y(t)$ is efficient, in the sense that future revisions are unpredictable at the time that the initial estimate is released. In particular, we assume that $\nu(t) \equiv y(t) - y(t)_t$ and $\nu(t)_s \equiv y(t)_s - y(t)_t$ (where $s \geq t$) are uncorrelated with all variables in the information set of the government at time t.[6] *A fortiori*, $\nu(t)$ and $\nu(t)_s$ are uncorrelated with all variables available to the econometrician at time t.

Although empirical tests of the efficiency assumption are not always supportive (Runkle 1998, Croushore and Stark 1999), the fact that attempts to "second guess" government statistical reports at the time of their release are rare suggests that in many applications inefficiencies in the government's estimation process are small. Hence, prescriptions based on the efficiency hypothesis will often be a good guide to econometric practice. This contention is borne out in the GDP forecasting results we present later.

III. Alternative Estimation Strategies

For the most part, we take the set of right-hand-side variables as "given" in our analysis. We focus, instead, on the vintage or vintages of data that should be used when estimating a forecasting equation to maximize real-time forecasting performance. The first strategy we consider is regressing the initial official estimate of $y(t)$ on real-time-vintage data–i.e., regressing $y(t)_t$ on $\boldsymbol{x}(t)_t$. The second strategy is similar, but replaces $y(t)_t$ with an end-of-sample-vintage official estimate of $y(t)$, $y(t)_{T-1}$. The third and final strategy we consider–used by most real-world forecasters–is to put end-of-sample-vintage data on both the left-hand side and right-hand side of the estimated forecasting equation. Hence, $y(t)_{T-1}$ is regressed on $\boldsymbol{x}(t)_T$.[7]

Strategy 1: First-Release Data on the Left-Hand Side, Real-Time-Vintage Data on the Right. Our preferred strategy is to estimate the forecasting equation with first-release data on the left-hand side and real-time-vintage data (data, at each point within the sample, that would have been available at the time) on its right-hand side. To see the rationale for this approach, note that Equation 1 implies that

$$y(t)_t = \boldsymbol{x}(t)_t \alpha + \omega(t). \tag{3}$$

where

$$\omega(t) \equiv w(t) - (y(t) - y(t)_t) = w(t) - \nu(t). \tag{4}$$

Since $\boldsymbol{x}(t)_t$ and $\nu(t)$ are uncorrelated under our efficiency assumption, ordinary least squares applied to Equation 3 yields an unbiased and consistent estimate of α. Less obviously, $\mathrm{var}(\omega(t)) \leq \mathrm{var}(w(t))$, so that even if he could somehow estimate Equation 1 directly, the econometrician should estimate Equation 3 instead.

The intuition behind this last result is straightforward. The error term in Equation 3 is the difference between the government's period-t estimate of $y(t)$ and the econometrician's estimate of $y(t)$. As such, it has a non-zero variance only to the extent that the government has more information in period t than does the econometrician. The error term in Equation 1, in contrast, reflects both the period-*t* government–private information gap and additions to the government's information set over time. If the government uses information efficiently, the government–private information gap will be uncorrelated with additions to the government's information set. Hence the error term in Equation 3 must have lower variance than the error term in Equation 1.[8]

More formally,

$$\begin{aligned}\mathrm{var}(w(t)) &= \mathrm{var}(\omega(t) + \nu(t)) \\ &= \mathrm{var}(\omega(t)) + \mathrm{var}(\nu(t)) + 2\mathrm{cov}(\mathrm{y}(t)_t - \boldsymbol{x}(t)_t\,\alpha,\ \nu(t)).\end{aligned}$$

However, both $y(t)_t$ and $\boldsymbol{x}(t)_t$ are in the government's information set at time *t* and, consequently, must be uncorrelated with $\nu(t)$. It follows that

$$\mathrm{var}(w(t)) = \mathrm{var}(\omega(t)) + \mathrm{var}(\nu(t)) \geq \mathrm{var}(\omega(t)).$$

Except in the uninteresting case in which $y(t)_t$ is never revised, the Equation 3's error term has strictly lower variance than that of Equation 1. Consequently, estimating Equation 3 will, in any finite sample, yield parameter estimates more precise than those obtained by estimating Equation 1.

There is no *a priori* reason to expect $\omega(t)$ to be heteroscedastic. Serial correlation can usually be eliminated by expanding the vector $\boldsymbol{x}(t)_t$ to include additional lags.[9]

The forecast error under Strategy 1,

$$y(T) - \hat{y}(T) = [y(T) - y(T)_T] + [y(T)_T - \hat{y}(T)_T], \tag{5}$$

has two components: the government's error in forecasting $y(T)$ based on information available to it at time T, and the econometrician's error in forecasting the government's initial estimate of $y(T)$. Each component has a zero mean. Moreover, under our assumption that the government uses available information efficiently, the two components must be uncorrelated. Hence,

$$var[y(T) - \hat{y}(T)] = var[y(T) - y(T)_T] + var[y(T)_T - \hat{y}(T)_T]$$

$$= \sigma_\nu^2 + \sigma_\omega^2 + var[\boldsymbol{x}(T)_T(\alpha - \hat{\alpha})]. \tag{6}$$

The final term in Equation 6 is the variance due to coefficient uncertainty in the estimated version of Equation 3. For any fixed $\boldsymbol{x}(T)_T$, this variance goes to zero as the sample size increases. For any given sample size, the variance due to coefficient uncertainty is increasing in the distance between $\boldsymbol{x}(T)_T$ and the sample mean of the $\boldsymbol{x}(t)_t$ ($t = 0, 1, ...T - 1$). The simplest case is that in which $\boldsymbol{x}(T)_T$ and the sample mean of the $\boldsymbol{x}(t)_t$ coincide. Equation 6 then reduces to

$$var[y(T) - \hat{y}(T)] = \sigma_\nu^2 + \sigma_\omega^2(1 + 1/T). \tag{7}$$

We will use this case as a benchmark when comparing the performance of Strategy 1 to the performance of alternative estimation strategies.[10]

Strategy 1': First-Release Data on Both the Left-Hand Side and Right-Hand Side. A special case of Strategy-1 estimation arises when the vector $\boldsymbol{x}(t)_t$ on the right-hand-side of Equation 3 is restricted to include only first-release data. By entirely eliminating the need to keep track of revisions, this artificial restriction on the set of information used by the econometrician somewhat simplifies data collection. The obvious cost is a deterioration in in-sample fit and out-of-sample forecasting precision. In Equations 3 and 7, σ_ω^2 is larger than previously. Less obviously, restricting $\boldsymbol{x}(t)_t$ to first-release data may induce serial correlation in the error term in Equation 3.[11]

Although inefficient, restricting $\boldsymbol{x}(t)_t$ to include only first-release data still produces a forecast that is unbiased, conditional on the restricted information set. Swanson (1996) and Swanson and White (1996, 1997) have used first-release data sets to illustrate how estimation using current-vintage data can distort forecasting relationships, and to compare alternative forecasting models and model-selection criteria.

Strategy 2: End-of-Sample-Vintage Data on the Left, Real-Time-Vintage Data on the Right. Under Strategy 2, the econometrician comes as close to estimating Equation 1 as possible, by using the most up-to-date available estimate of $y(t)$ on the left-hand side of the forecasting equation. This is not a *bad* method of estimating α, but it's not as efficient a method as Strategy 1. The proof of this proposition is essentially the same as that used above to establish that Equation 3's error term has a lower variance than Equation 1's error term.

Formally, under Strategy 2 the econometrician estimates

$$y(t)_{T-1} = \boldsymbol{x}(t)_t\alpha + \omega'(t) \tag{3'}$$

for $0 \leq t < T$, where $\omega'(t) \equiv w(t) - (y(t) - y(t)_{T-1}) = w(t) - (y(t) - y(t)_t) + (y(t)_{T-1} - y(t)_t) = \omega(t) + \nu(t)_{T-1}$.[12] Under our assumption that $\nu(t)_{T-1}$ is uncorrelated with information available to the government at time t, applying least squares to Equation 3' will yield an unbiased estimate of α. However, since $\omega(t) = y(t)_t - \boldsymbol{x}(t)_t\alpha$ and since both $y(t)_t$ and $\boldsymbol{x}(t)_t$ are in the government's period-t information set, $\omega(t)$ must be uncorrelated with $\nu(t)_{T-1}$. It follows that the variance of the error term in Equation 3' will be greater than the variance of the error term in Equation 3:

$$\text{var}[\omega'(t)] = \text{var}[\omega(t)] + \text{var}[\nu(t)_{T-1}] \geq \text{var}[\omega(t)], \tag{8}$$

with strict inequality holding for at least one $t < T - 1$. (The only exception occurs when the government's initial estimate of $y(t)$ is never revised, so that Strategies 1 and 2 are identical.) More generally, the variance of the error term in Equation 3' can be expected to be decreasing in t, but is always at least as great as the variance of the error term in Equation 3.[13] Hence, for any $s > t$, the relationship between $y(t)_s$ and $\boldsymbol{x}(t)_t$ is necessarily looser than that between $y(t)_t$ and $\boldsymbol{x}(t)_t$.

That the error term in Equation 3' has larger variance than the error term in Equation 3 means that the parameter vector, α, will be estimated less precisely if least squares is applied to the former equation than if it is applied to the latter.[14] The Strategy-2 counterpart to Equation 6 is

$$var[y(T) - \hat{y}'(T)] = var[y(T) - y(T)_T] + var[y(T)_T - \hat{y}(T)'_T]$$

$$= \sigma_\nu^2 + \sigma_\omega^2 + var[\boldsymbol{x}(T)_T(\alpha - \hat{\alpha}')], \tag{6'}$$

where a prime is used to distinguish the predictions and estimates associated with Strategy 2 from those associated with Strategy 1. The problem with Strategy 2 is that the third term on the right-hand side of Equation 6'–the term that measures the effects of coefficient uncertainty–is larger than the corresponding term on the right-hand side of Equation 6. In the benchmark case in which $\boldsymbol{x}(T)_T$ and the sample mean of the $\boldsymbol{x}(t)_t$ coincide, for example, Equation 6' reduces to

$$var[y(T) - \hat{y}'(T)] = \sigma_\nu^2 + \sigma_\omega^2 + (\sigma_\omega^2 + \bar{\sigma}_\nu^2)/T. \tag{7'}$$

where

$$\bar{\sigma}_\nu^2 \equiv (\frac{1}{T})\Sigma_{t=0}^{T-1} \; var[\nu(t)_{T-1}]$$

is the average variance of the revisions to the left-hand-side variable. Comparing Equations 7 and 7', the relative inefficiency of Strategy 2 is clear.[15]

Strategy 3: End-of-Sample-Vintage Data on Both Left and Right. As noted above, real-world professional forecasters generally use end-of-sample-vintage data on both the left-hand and right-hand sides of their forecasting equations. Economists who undertake conventional real-time forecasting exercises imitate this process, updating the vintage of the data used for estimating their models as they gradually extend their samples. In our notation, the conventional approach to real-time forecasting amounts to estimating

$$y(t)_{T-1} = \boldsymbol{x}(t)_T \alpha + \omega''(t) \tag{3''}$$

for $0 \leq t \leq T - 1$, where $\omega''(t) \equiv \omega(t) + \nu(t)_{T-1} - \xi(t)_T \alpha$ and where $\xi(t)_T \equiv \boldsymbol{x}(t)_T - \boldsymbol{x}(t)_t$ is a vector of revisions to the right-hand-side variables. We have already assumed that $y(t)_t$ is an efficient estimate of $y(t)_{T-1}$. Suppose, similarly, that $\boldsymbol{x}(t)_t$ is an efficient estimate of $\boldsymbol{x}(t)_T$. Then neither $\nu(t)_{T-1}$ nor $\xi(t)_T \alpha$ will be correlated with $\omega(t)$ and, so,

$$\text{var}[\omega''(t)] = \text{var}[\omega(t)] + \text{var}[\zeta(t)] \geq \text{var}[\omega(t)], \tag{9}$$

where $\zeta(t) \equiv \nu(t)_{T-1} - \xi(t)_T \alpha$. Equality between var[$\omega''(t)$] and var[$\omega(t)$] obtains only in the exceptional case in which revisions to the government's estimates of $y(t)$ and movements in $\xi(t)_T \alpha$ exactly offset one another, so that var[$\zeta(t)$] = 0. In general, though, the variance of the error term in Equation 3" will exceed the variance of the error term in Equation 3.[16]

Correlation between $\zeta(t)$ and $\boldsymbol{x}(t)_T \alpha$ in Equation 3" is likely. Only if variation in $\nu(t)_{T-1}$ completely offsets variation in $\xi(t)_T \alpha$ [so that cov($\zeta(t)$, $\xi(t)$) = 0] will such correlation be avoided. Absent a complete offset, least squares will yield an inconsistent estimate of α. The argument is similar to that which underlies textbook discussions of the classic "errors-in-variables" problem, but is complicated by the fact that some of the right-hand-side-variable errors (some of the elements of $\xi(t)_T$) may be correlated with the left-hand-side-variable error ($\nu(t)_{T-1}$).

In the classic errors-in-variables problem the limiting value of the least-squares coefficient vector is

$$plim(\hat{\alpha}^{//}) = \alpha - (\Sigma_{XX} + \Sigma_{\xi\xi})^{-1}\Sigma_{\xi\xi}\alpha, \tag{10}$$

where $\Sigma_{xx} \equiv \text{plim}[\mathbf{X}'\mathbf{X}/T]$, $\Sigma_{\xi\xi} \equiv \text{plim}[\Xi'\Xi/T]$, and $\mathbf{X}$ and Ξ are $T\times k$ matrices whose tth rows are $\boldsymbol{x}(t)_t$ and $\xi(t)_T$, respectively. Thus, the least-squares coefficient estimates are inconsistent. More generally,

$$plim(\hat{\alpha}^{//}) = \alpha + (\Sigma_{XX} + \Sigma_{\xi\xi})^{-1}[\Sigma_{\xi\nu} - \Sigma_{\xi\xi}\alpha], \tag{11}$$

where $\Sigma_{\xi\nu} \equiv \text{plim}[\Xi'\text{N}/T]$ and N is a $T \times 1$ vector whose tth element is $\nu(t)_{T-1}$. While it is theoretically possible for the term in square brackets in Equation 11 to equal zero, there are no *a priori* grounds for believing that this condition will hold. Hence, Strategy 3 will typically yield an inconsistent estimate of α and a biased forecast of $y(T)$.

As an extreme example, consider the special case in which $\alpha = 0$, so that $\boldsymbol{x}(T)_T$ is of absolutely no use in forecasting $y(T)$. Equation 11 says that estimating a relationship between y and $\boldsymbol{x}$ using end-of-sample data (Strategy 3) will, nevertheless, yield a non-zero estimate of α insofar as *revisions* to y are correlated with *revisions* to the elements of $\boldsymbol{x}$. Hence, one of the complaints that has been directed against forecasting analyses that use current-vintage data–that such analyses can lead the econometrician to include variables on the right-hand-side of his forecasting equation that, in real time, have little marginal predictive power–applies also to any real-time analysis that uses end-of-sample vintage data.

Intuitively, there really isn't that much difference between coming up with a forecast of (say) 1997:Q1 GDP growth using an equation estimated with 1997:Q1-vintage data (a Strategy-3 real-time forecasting exercise) and coming up with a forecast of 1997:Q1 GDP growth using an equation estimated through 1996:Q4 but with the latest-available data (the approach taken by analysts who don't bother to collect real-time data). Some of the data for 1995 and 1996 may differ significantly between the two estimations, but revisions are likely to be small for earlier years. If sample periods extend back very far at all, both estimations will be dominated by heavily revised data–data that may contain more information on how revisions to GDP growth are related to revisions to the right-hand-side variables than on how recently released (and largely unrevised) estimates of the right-hand-side variables are related to GDP growth.

Summary. In the typical forecasting problem, one must use today's data to try to predict a future official release of some variable of interest. There are three different ways one might go about estimating such a forecasting relationship. The most natural approach is to estimate the forecasting relationship using end-of-sample-vintage data for the left-hand-side variable and real-time-vintage data for each of the right-hand-side variables (Strategy 2). The coefficient estimates so obtained will be unbiased. We argued that in many cases one can expect to obtain more precise coefficient estimates–and, hence, better forecasts–by using real-time-vintage data on the left-hand-side of the forecasting equation as well as on the right-hand side (Strategy 1). However, real-world forecasters typically take just the opposite approach, estimating their forecasting equations using end-of-sample-vintage data on both the left-hand side and the right-hand side (Strategy 3). In general, the coefficient estimates obtained using the conventional approach are inconsistent and inefficient. An important assumption underlying our theoretical results is that future revisions to official government estimates cannot be forecasted at the time the official estimates are first released.

IV. An Example: Forecasting Current-Quarter GDP Growth

The Model. To illustrate the importance of estimating forecasting models with real-time-vintage data, we attempt to predict current-quarter real GDP growth using monthly measures of real economic activity. Following Trehan (1992), our set of monthly indicator variables includes 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).[17,18]

To obtain our forecasting models, we regress quarter-to-quarter changes in real GDP on a constant and five month-to-month changes in each of our three coincident indicators. To be precise, we estimate equations of the form:

$$\Delta y_t = \alpha_0 + \Sigma_{j=0}^{4}\beta_j e_{t,3-j} + \Sigma_{j=0}^{4}\gamma_j ip_{t,3-j} + \Sigma_{j=0}^{4}\epsilon_j rs_{t,3-j}, \qquad (12)$$

where Δy_t denotes the annualized quarterly percentage change in real GDP in quarter t, and where $e_{t,s}$, $ip_{t,s}$, and $rs_{t,s}$ are the annualized monthly percentage changes in non-farm employment, industrial production, and real retail sales, respectively, in month s of quarter t.[19] The lag length is chosen entirely on *a priori* grounds.[20]

Estimations start in the first quarter of 1980 and end on or before the third quarter of 1997. Forecasts cover the period from 1990:Q1 through 1997:Q4. Sample periods are limited by the difficulty of constructing end-of-sample-vintage data sets.

The Data. For our recursive estimation/forecasting exercises we require current-vintage, end-of-sample-vintage, real-time-vintage, and (as a special case of real-time-vintage) first-release data sets.

In our current-vintage data set, $e_{t,3-j}$, $ip_{t,3-j}$, and $rs_{t,3-j}$ (t = 1980:Q1-1997:Q4; j = 0, ..., 4) are all measured as of January 1999.

In our end-of-sample-vintage data sets, $e_{t,3-j}$, $ip_{t,3-j}$, and $rs_{t,3-j}$ (t = 1980:Q1-T; j = 0, ..., 4) are measured as of the end of the sample period (quarter T), which ranges from 1989:Q4 to 1997:Q4.[21] Data toward the beginning of a given end-of-sample-vintage data set have undergone extensive revision. Data at the very end of the data set are "first release." End-of-sample-vintage data are used by professional forecasters and in many existing real-time economic analyses.

Our real-time-vintage data set consists of a sequence of 5-month "snapshot" histories of our right-hand-side variables–one snapshot for each quarter from 1980:Q1 through 1997:Q4. Regardless of when the estimation period ends, only data that would have been available in 1980:Q1 are used to forecast 1980:Q1 GDP growth; only data that would have been available in 1980:Q2 are used to forecast 1980:Q2 GDP growth; etc. There is exactly one vintage of data for each point in the estimation period. Formally, $e_{t,3-j}$, $ip_{t,3-j}$, and $rs_{t,3-j}$ (j = 0, ..., 4) are all measured as of quarter t, for t = 1980:Q1-1997:Q4.[22]

Our first-release data set has only one employment-growth, one industrial-production-growth, and one retail-sales-growth entry for each month from November 1979 through December

1997. Each entry is the government's first estimate of the variable in question.

We variously use current-vintage, end-of-sample-vintage, and first-release measures of real GDP growth as the left-hand-side variable when estimating Equation 12.[23] Only the end-of-sample-vintage and first-release measures would have been available in real time. The switch from fixed- to chain-weight GDP is treated just like any other data revision.

Efficiency Tests. Recall that our preferred estimation technique (Strategy 1) puts first-release data on the left-hand side of the forecasting equation for estimation purposes and real-time-vintage data on the right-hand side. This approach yields unbiased coefficient estimates if, and only if, revisions to official estimates of the forecasted variable ($y(t)$) are uncorrelated with the vector of right-hand-side variables, $\boldsymbol{x}(t)_t$. In the present context, the question is whether changes in the government's GDP estimates are correlated with real-time-vintage growth in employment, industrial production and retail sales. Results from a regression of GDP revisions on the set of monthly coincident indicators–presented on Line 1 of Table 1– are unambiguous. There is no indication that real-time employment, industrial production and retail sales are related any differently to initial estimates of GDP growth than they are to latest-available (January 1999) estimates of GDP growth. The *F* statistic associated with the regression has a *P* value of 0.448.

The main competition for our preferred estimation technique is Strategy 2, which puts real-time-vintage data on the right-hand side of the forecasting equation for estimation, but end-of-sample data on the left. For Strategy 1 to yield more precise finite-sample coefficient estimates than Strategy 2 it is sufficient (not necessary) that revisions to official estimates of $y(t)$ be uncorrelated with the initial releases of that variable ($y(t)_t$) and with the vector of right-hand-side variables ($\boldsymbol{x}(t)_t$).

In the present context, we can expect Strategy 1 to yield more accurate GDP forecasts than Strategy 2 provided GDP revisions are not predictable using first-release GDP data and real-time-vintage growth in employment, industrial production and retail sales. Results from the relevant regression–displayed on Line 2 of Table 1–are encouraging: the hypothesis that GDP revisions are unpredictable cannot be rejected at the 5 percent significance level (although only just).[24]

In-Sample Performance. Table 2 compares the in-sample performance of Strategy 1 and Strategy 2 as applied to Equation 12. In addition, the table presents results for the special case of Strategy-1 estimation in which the set of real-time-vintage data is restricted to include only initial official estimates (Strategy 1') and for the most commonly used real-time estimation approach (Strategy 3), which puts end-of-sample-vintage data on both the right and left of the forecasting equation. In each case, the table displays the joint significance of the coefficients attached to the right-hand-side variables. It also displays goodness-of-fit measures, and the results of the *Q* test for serial correlation and the Goldfeld-Quandt test for heteroscedasticity.

According to the table, monthly percentage changes in employment, industrial production, and retail sales are highly statistically significant in every regression. As expected from our theoretical analysis, the Strategy-1 regression has the lowest standard error. Serial correlation is a problem only for Strategy 1', where it is expected. (See the Appendix.) Theoretically a threat also when using Strategy 3, heteroscedasticity appears to be of practical importance only for Strategy 2. (Nevertheless, all standard errors reported in the table were calculated using the Newey-West estimator, which is robust to heteroscedasticity and serial correlation.)

Recursive Forecast Comparisons. To rank our alternative estimation strategies, we conducted a series of recursive forecasting exercises. 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 1997:Q4. *The forecasts were compared with actual GDP growth, where "actual GDP growth" is GDP growth as measured in January 1999.* Presumably, current-vintage GDP growth statistics are the best available estimates of what happened to real economic activity over the period in question.

Results of these recursive forecasting exercises are displayed in Table 3 for models estimated using Strategies 1, 1', 2, and 3. Table 3 also shows how well Equation 12 appears to perform when estimated and evaluated naively, using current-vintage data throughout. This last approach yields potentially misleading results, since both the data used to obtain coefficient estimates and the data substituted into the right-hand side of the estimated equation to obtain a "forecast" would not actually have been available to an analyst in real time. Nevertheless, the naive approach to forecast evaluation is frequently used in practice.

As expected, Strategy 1 (which uses exclusively real-time-vintage data) beats out both Strategy 1' (which uses exclusively first-release data) and Strategy 2 (which uses end-of-sample-vintage data on the left-hand side of the forecasting equation and real-time-vintage data on the right-hand side). Strategy 3 performs poorly in comparison with the other methods.

Is real-time forecast evaluation worth the extra bother? If the only real-time estimation approach under consideration is Strategy 3, Table 3 suggests that the answer is "probably not." Thus, a naive evaluation methodology (which relies entirely on current-vintage data) predicts fairly accurately the root-mean-square forecast error generated by Strategy 3. In contrast, the naive approach gives a markedly too-pessimistic view of the real-time performance of Strategy 1.

The final two rows of Table 3 compare the recursive forecasting performance of our monthly-indicator model to two benchmark alternatives–a purely autoregressive model of GDP growth and the Blue Chip consensus forecast.[25] The autoregression performs poorly relative to the monthly-indicator model regardless of how the latter is estimated. The comparison with the Blue Chip consensus forecast is more interesting. It illustrates that a simple model, correctly estimated, can sometimes match the real-time performance of experienced, professional forecasters using conventional techniques. Thus, our Strategy-1 forecasts appear to do every bit as well as the Blue Chip consensus forecast over the period from 1990:Q1-1997:Q4. In contrast, when estimated using Strategy 2, Strategy 3, or naively, our model performs comparatively poorly.

Table 4 presents results of a more formal comparison of how well various versions of the monthly-indicator model perform relative to the Blue Chip consensus. In the left half of the table, recursive forecast errors from the monthly indicators model are regressed on the Blue Chip consensus forecast to see whether the Blue Chip forecast has marginal predictive power. In the right half of the table, similarly, we test whether the different versions of the monthly indicators model have predictive power beyond the Blue Chip forecast.[26] Although such encompassing tests are known to be problematic when predictions depend on estimated parameters, they are asymptotically valid when the forecasts under consideration are recursive and the models being compared are non-nested–conditions that are met here.[27]

Results displayed in the left half of the Table 4 indicate that the Blue Chip forecast does not contain information beyond that which is captured by the monthly indicators model, regardless of how that model is estimated.[28] In contrast, results vary considerably across estimation strategies when–as in the right half of Table 4–the roles of the Blue Chip and monthly indicators forecasts are reversed. An analyst using Strategy-1' or Strategy-3 estimation would conclude from these results that the monthly indicators model is of little help in forecasting GDP growth given the Blue Chip consensus. In actuality, the failure to improve on the consensus forecast is the fault of the data used to estimate the model rather than of the model *per se*. Thus, the indicators model estimated according to Strategy 1 has substantial marginal predictive power relative to the Blue Chip, and Strategy-2 forecasts are very nearly statistically significant.

Forecasting "Advance" and "Final" GDP. All of the performance statistics presented up to this point have assumed that the analyst wants the best possible prediction of "true" GDP growth. In practice, true GDP growth is approximated using GDP growth as it appeared in January 1999. Thus, Tables 3 and 4 evaluate alternative estimation strategies with reference to January 1999 GDP data. However, a case can be made that many real-world analysts are, and ought to be, more concerned with predicting GDP statistics that have undergone only a few revisions than with predicting GDP statistics released many years after the fact. It is, after all, the relatively early GDP releases that are most likely to affect the current decisions of households and businesses and, hence, the future course of the economy. Accordingly, we briefly consider the performance of alternative estimation strategies in forecasting the first ("advance") and third ("final") GDP estimates, which become available one month and three months after the end of the quarter, respectively. Results are presented in Table 5, which is similar in format to Table 3. We confine our analysis to our preferred estimation technique (Strategy 1), the most commonly used estimation technique (Strategy 3), and the Blue Chip consensus forecast.

Two main conclusions emerge from Table 5. First, early GDP releases are substantially easier to predict than are late releases. Thus, the root-mean-square errors recorded in Table 5 are uniformly lower than the corresponding root-mean-square errors in Table 3. Within Table 5, root-mean-square errors are lower when forecasting the advance GDP release than when forecasting the final GDP release. These findings are consistent with in-sample estimation results reported in Table 2, and are exactly what one would expect to see given the unpredictability of revisions to the government's GDP estimates. When revisions are unpredictable, stripping the revisions away eliminates noise from the equation. The second main conclusion to emerge from Table 5 is that the relative ranking of the alternative forecasts is unaffected by whether the goal is to predict an early or a late GDP release. If anything, the performance of the monthly indicators model appears to improve relative to that of the Blue Chip consensus forecast when predicting early releases.[29]

V. Concluding Remarks

In most economic forecasting applications, the data that are substituted into the right-hand-side of the forecasting equation to obtain an actual out-of-sample forecast have undergone few, if any, revisions. We have argued that this fact needs to be taken into account when the forecasting equation is first estimated. In particular, at each date within his sample the econometrician estimating a forecasting equation ought to use only right-hand-side data that would have been available at the time. We call these "real-time-vintage" data. Real-time-vintage data sets are more complete than "unrevised" or "first-release" data sets because at each within-sample date they include revisions that would have been known at that date. For typical lag specifications (extending back a year or less), real-time-vintage data are readily available in back issues of government publications.

Most analysts do *not* use real-time vintage data to estimate their forecasting models. Instead, they use the most up-to-date numbers available at the time of the estimation. Economists often label as "real time" forecasting exercises in which this practice is reproduced after the fact. In these exercises, the economist gradually extends the period over which a forecasting equation is estimated, each time using data as it would have appeared at the close of the sample period. The problem with using end-of-sample-vintage data to estimate a forecasting equation is that correlation between revisions to a right-hand-side variable and revisions to the left-hand-side variable can make it appear that a forecasting relationship exists when, in fact, early vintages of the first variable–the only vintages actually relevant to constructing current forecasts–have little or no marginal predictive power. In other words, the linkages between the right-hand-side and left-hand-side variables near the start of the sample period–where both are heavily revised–may be quite different from the linkages at the end of the sample period–where the available data have undergone little, if any, revision. It is only the latter linkages that are relevant for constructing an accurate current forecast.

A more subtle question than whether the right-hand-side variables in the forecasting equation ought to be real-time vintage or end-of-sample vintage is whether the *left*-hand-side variable ought to be first release or end-of-sample vintage. We have argued that as long as the government's first release fully exploits available information, superior forecasting performance can be expected if first-release data are used on the left-hand side of the equation during estimation. Intuitively, if the government's initial release is efficient, using it on the left-hand side of the equation strips away unpredictable noise from the dependent variable. Because the real-time information on the right-hand side of the forecasting equation is more closely related to the government's initial estimate of the left-hand-side variable than it is to any revised estimate, a regression with first-release data on its left-hand side will yield more precise coefficient estimates than a regression with revised data on its left-hand side. First-release data are to be preferred for estimation even if the analyst is ultimately interested in predicting revised data.

We certainly do not believe that every government estimate is fully efficient. However, we *do* believe that attempts to second guess government estimates at the time of their release are sufficiently uncommon as to suggest that efficiency may be a useful approximation in many applications.

In the particular application considered here–forecasting current-quarter GDP using monthly data on employment, industrial production, and retail sales–we find that our theoretical findings are borne out. A substantial improvement in out-of-sample performance is achieved if the forecasting equation is estimated with real-time-vintage data on its right-hand side, rather than end-of-sample vintage data. There is a further improvement if first-release GDP growth is used as a left-hand-side variable for estimation purposes. Properly estimated, our simple model appears to outperform the Blue Chip consensus GDP forecast.

APPENDIX

A. The Data Requirements of the Alternative Estimation Strategies

A simple example will help illustrate the data requirements of the alternative forecasting strategies discussed in the main text. Suppose that the econometrician is interested in forecasting GDP growth using current and two lags of employment growth. Available GDP data extend from $t = 0$ to $t = T - 1$. As shown in Table A.1, we can arrange the GDP data in a triangular array with T rows and T columns. In the upper left-hand corner is the period-0 official estimate of period-0 GDP growth ($y(0)_0$). In the upper right-hand corner is the period-T-1 official estimate of period-0 GDP growth ($y(0)_{T-1}$). And in the lower right-hand corner is the period-T-1 official estimate of period-T-1 GDP growth ($y(T-1)_{T-1}$). In general, as one moves from left to right along a row, one is looking at increasingly up-to-date estimates of a particular quarter's GDP growth. As moves from top to bottom along a column, the *vintage* of the GDP estimates stays constant, but one is looking at GDP growth in ever more recent periods.[1] Strategy 1 uses data from along the main diagonal of the data array for estimation purposes. Strategies 2 and 3 use data from the right-most column.

Under Strategy 1, as T increases (so that the sample period is extended), the econometrician simply adds a new GDP growth observation to the end of his data set. (In Table A.1, all that's needed is one new diagonal element at the lower right of the data array.) Under Strategies 2 and 3, the entire data set is discarded and replaced with a new set of GDP-growth observations of vintage T. (In Table A.1, an entire new column of data must be added to the data array.) Suppose that the econometrician wishes to conduct an *ex post*, real-time, recursive-forecast exercise–an exercise in

---

[1] The format is similar to that of Table 1 in Diebold and Rudebusch (1991).

which the model is estimated and re-estimated as the sample period is gradually extended, for each sample period using only data that would have been available at the time. Under Strategy 1, the econometrician need collect only *one* series of GDP growth estimates–a series consisting entirely of initial releases. Under Strategies 2 and 3, the econometrician must collect a data set of vintage T - 1 that covers GDP growth over the entire interval from t = 0 to t = T - 1, a data set of vintage T that extends from t = 0 to t = T, and so forth.

Table A.2 displays an array of employment-growth data of different vintages that is organized much like the GDP-growth array of Table A.1, except that there are two additional rows at the top of the array (for employment growth at t = -1 and t = -2) and one additional column (reflecting the availability of vintage-T employment data). As before, Strategy 3 uses only entries from the extreme right-hand column of the array. If the sample period is extended by one quarter, the old data set must be discarded and replaced with a new set of employment-growth observations of vintage T+1. Conducting a real-time, recursive-forecast exercise requires that the econometrician collect a sequence of long data sets, each of a different vintage.

Forecasting Strategies 1 and 2 use employment-growth data from the bottom *three* elements of each column, reflecting the fact that the employment-growth terms that appear on the right-hand side of the forecasting equation at a given date are all of the same vintage. Extending the sample period by one quarter simply requires adding three new entries at the lower right of the employment-growth array–an entirely new observation of period-T+1 employment growth, and newly revised estimates of period-T and period T-1 employment growth. Conducting a real-time, recursive-forecast exercise requires that the econometrician collect many three-element employment-growth snap shots, each of a different vintage.

Note that the employment-growth observations used under Strategies 1 and 2 in this example do *not* consist only of first-release data: at each point in the sample, current-quarter employment growth is first release, but lagged employment growth is revised. Hence, in the main text we call the right-hand-side data under Strategies 1 and 2 "real-time vintage" rather than "first release." Restricting oneself to first-release right-hand-side data (Strategy 1') amounts to taking data only from the main diagonal of the array, rather than from the main and two adjacent diagonals.

To see how serial correlation can arise under Strategy 1', when it is absent under Strategy 1, consider the difference between the right-hand-side variables under the two strategies. Under Strategy 1, the right-hand-side variables at time t are $e(t)_t$, $e(t-1)_t$, and $e(t-2)_t$. Under Strategy 1', the right-hand-side variables are $e(t)_t$, $e(t-1)_{t-1}$, and $e(t-2)_{t-2}$. The differences between the right-hand-side variables under the two strategies are thus 0, $e(t-1)_t - e(t-1)_{t-1}$, and $e(t-2)_t - e(t-2)_{t-2} = [e(t-2)_t - e(t-2)_{t-1}] + \mathbf{[}\boldsymbol{e(t-2)_{t-1} - e(t-2)_{t-2}}\mathbf{]}$. At time $t-1$, the corresponding differences are 0, $\boldsymbol{e(t-2)_{t-1} - e(t-2)_{t-2}}$, and $e(t-3)_{t-1} - e(t-3)_{t-3}$. The overlap (the terms in bold-faced type), means that the information that is left out as one goes from Strategy-1 to Strategy-1' estimation can be expected to be correlated over time.

B. More on the Gains from Strategy-1 Estimation

Strategy-1 coefficient estimates–obtained by regressing first-release data ($y(t)_t$) on real-time-vintage data ($\boldsymbol{x}(t)_t$)–are unbiased if revisions to the left-hand-side variable are uncorrelated with information available to the econometrician at the time of the variable's initial release. If $y(t)_t$ is a *fully* efficient estimate of $y(t)$, then Strategy-1 coefficient estimates can be expected to be more precise than those obtained by using some later estimate of $y(t)$ as the left-hand-side variable. This appendix documents the potential payoff, in improved forecasting accuracy, from using Strategy-1 estimation, and discusses circumstances in which the payoff is likely to be large.

We consider the case where the econometrician wishes to forecast $y(T)$ using a single right-hand-side variable and a constant. Thus, Equation 1 in the main text takes the form:

$$y(t) = \alpha_0 + \alpha_1\, x(t)_t + w(t), \tag{B.1}$$

and, similarly, Equation 3 in the main text becomes

$$y(t)_t = \alpha_0 + \alpha_1\, x(t)_t + \omega(t), \tag{B.2}$$

where $\omega(t) = w(t) - \nu(t) = w(t) - [y(t) - y(t)_t]$. For example, $y(t)$ might be quarterly real GDP growth and $x(t)$ same-quarter employment growth. We assume that the government's initial $y(t)$ estimate is efficient, so that $\nu(t)$ is uncorrelated with all variables in the government's period-t information set.

Suppose that $y(t)$ becomes available within one period of the initial release, so that direct estimation of Equation B.1 is possible–an extreme version of Strategy-2 estimation. A realistic, real-world example would be the analyst who wants to forecast the so-called "final" estimate of real GDP growth, which is published two months after GDP figures first become available. The question is whether the analyst ought to estimate his forecasting equation with "final" GDP data on its left-hand side (Equation B.1), or with first-release data (Equation B.2).

With a sample period extending from $t = 0$ to $t = T - 1$, the variance of the period-T forecast error is

$$\mathrm{var}[y(T) - \hat{y}(T)] = \sigma_w^2(1 + K) = (\sigma_\omega^2 + \sigma_\nu^2)(1 + K) \tag{B.3}$$

using Equation B.1 and

$$\begin{aligned}\mathrm{var}[y(T) - \hat{y}(T)_T] &= \mathrm{var}[y(T) - y(T)_T] + \mathrm{var}[y(T)_T - \hat{y}(T)_T] \\ &= \sigma_\nu^2 + \sigma_\omega^2(1 + K)\end{aligned} \tag{B.4}$$

using Equation B.2, where

$$K \equiv \frac{1}{T}\left[1 + \frac{T+1}{T-1}F\right],$$

$$F \equiv \frac{(x(T)_T - \bar{x})^2 / (T+1)}{[(1/T)\sum_0^{T-1}(x(t)_t - \bar{x})^2] / (T-1)},$$

and $\bar{x}$ is the sample mean of the $x(t)_t$. Note that Equation B.2 (Strategy 1) unambiguously outperforms Equation B.1 (Strategy 2). The intuition is that $\nu(t)$ is just noise from the perspective of period t under the assumption that the government's revision process is efficient. By purging the left-hand-side variable of this noise before estimation, the analyst gets more precise coefficient estimates.

Comparing Equations B.3 and B.4, the percentage improvement in forecast-error variance achieved by estimating Equation B.2 rather than Equation B.1 is

$$[\frac{\sigma_{\nu}^2}{\sigma_{\nu}^2+\sigma_{\omega}^2}][\frac{K}{1+K}] = [\frac{\sigma_{\nu}^2}{\sigma_{\nu}^2+\sigma_{\omega}^2}][\frac{1+\frac{T+1}{T-1}F}{(T+1)+\frac{T+1}{T-1}F}]. \qquad \text{(B.5)}$$

The first term in square brackets is the fraction of σ_w^2 accounted for by revisions to y. The more important are these revisions, the bigger the payoff from stripping them out of the forecasting equation before estimating its parameters. The second term in square brackets measures the fraction of the total Strategy-2-forecast-error variance accounted for by parameter uncertainty. This fraction tends to be larger the smaller is the sample size. It is also larger the more unusual is $x(T)_T$ relative to in-sample values of $x(t)_t$, where "unusualness" is measured by the squared deviation of $x(T)_T$ from $\bar{x}$ divided by the average in-sample squared deviation of $x(t)_t$ from $\bar{x}$.

To reach conclusions about the *ex ante* expected performance advantage of Strategy-1 estimation relative to Strategy-2 estimation, one must assume something about the distribution from which the $x(t)_t$ are drawn. Suppose, for example, that the $x(t)_t$ ($t = 0, 1, ...T$) are independent draws from a normal distribution, and consider the ratio of the Strategy-1forecast-error variance to the Strategy-2 forecast-error variance:

$$\frac{\text{var}[y(T)-\hat{y}(T)_T]}{\text{var}[y(T)-\hat{y}(T)]} = [\frac{\sigma_{\nu}^2}{\sigma_{\nu}^2+\sigma_{\omega}^2}][\frac{T}{T+1}][\frac{T-1}{T-1+F}] + [\frac{\sigma_{\omega}^2}{\sigma_{\nu}^2+\sigma_{\omega}^2}]. \qquad \text{(B.6)}$$

It is easily seen that F is distributed as $F_{1,T\text{-}1}$ if the $x(t)_t$ are independent normal. It follows that $(T - 1)/(T - 1 + F)$ has a beta distribution with parameters $(T - 1)/2$ and ½ (Hogg and Craig 1970, p. 134), and that the mean of the forecast-variance ratio is

$$E\{\frac{\mathrm{var}[y(T) - \hat{y}(T)_T]}{\mathrm{var}[y(T) - \hat{y}(T)]}\} = 1 - [\frac{\sigma_v^2}{\sigma_v^2 + \sigma_w^2}][\frac{2}{T+1}]. \tag{B.7}$$

Consistent with results (discussed above) for the case in which the $x(t)_t$ are known, Equation B.7 says that the expected performance advantage of Strategy 1 is greater the more important are revisions to the government's official estimates of $y(t)$ and the smaller is the sample size over which the econometrician's forecasting equation will be estimated.

For the monthly-indicators GDP-growth model developed in the main text, the in-sample estimation results reported in Table 2 suggest that $\sigma_v^2/\sigma_w^2 = 0.68$. For this value of σ_v^2/σ_w^2, Equation B.7 implies an expected forecast-variance ratio of 0.935 when $T = 20$, as compared with an expected forecast-variance ratio of 0.993 when $T = 200$. Using Equation B.6 and the fact that $(T - 1)/(T - 1 + F)$ has a beta distribution, one can further show that there is one chance in ten of observing a forecast-variance ratio of 0.880 or below when $T = 20$. When $T = 200$, the tenth percentile of the forecast-variance ratio is a substantially higher 0.987.

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.

Croushore, Dean and Tom Stark (1999) "A Real-Time Data Set for Macroeconomists," Federal Reserve Bank of Philadelphia Working Paper No. 99-4.

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

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

Fair, Ray C. and Robert J. Shiller (1990) "Comparing Information in Forecasts from Econometric Models," *American Economic Review* 80, pp. 375-89.

Faust, Jon, John H. Rogers, and Jonathan H. Wright (2000) "News and Noise in G-7 GDP Announcements," Board of Governors of the Federal Reserve System, International Finance Division Discussion Paper #690.

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.

Hogg, Robert V. and Allen T. Craig (1970) *Introduction to Mathematical Statistics*, Macmillan Publishing Co., New York.

Ingenito, Robert and Bharat Trehan (1996) "Using Monthly Data to Predict Quarterly Output," Federal Reserve Bank of San Francisco *Economic Review*, No. 3, pp. 3-11.

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 (1999) "Rational Bias in Economic Forecasts," *Quarterly Journal of Economics* 114, pp. 293-318.

Mankiw, N. Gregory and Matthew D. Shapiro (1986) "News or Noise: An Analysis of GNP Revisions," *Survey of Current Business* 66, pp. 20-5.

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.

Orphanides, Athanasios (1999) "The Quest for Prosperity Without Inflation," unpublished working paper, Board of Governors of the Federal Reserve System.

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

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

Rogers, R. Mark (1998) "A Primer on Short-Term Linkages between Key Economic Data Series," Federal Reserve Bank of Atlanta *Economic Review*, Second Quarter, pp. 40-54.

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

Swanson, Norman R. (1996) "Forecasting Using First Available Versus Fully Revised Economic Time Series Data," Working Paper No. 4-96-7, Pennsylvania State University.

Swanson, Norman R. and Halbert White (1996) "Forecasting Economic Time Series Using Flexible Versus Fixed Specification and Linear Vs. Nonlinear Econometric Models," unpublished.

Swanson, Norman R. and Halbert White (1997) "A Model Selection Approach to Real-Time Macroeconomic Forecasting Using Linear Models and Artificial Neural Networks," *Review of Economics and Statistics* 79, pp. 540-50.

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.

West, Kenneth D. and Michael W. McCracken (1998) "Regresion-Based Tests of Predictive Ability," *International Economic Review* 39, pp. 817-40.

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. Testing the Efficiency of the Government's Initial GDP Estimates: Summary statistics obtained by regressing GDP revisions, $v(t)$, on alternative information sets, 1980:Q1-1997:Q3.

Information Set	Adjusted R^2	F-Statistic	Constant Term
$\boldsymbol{x}(t)_t$	0.005	1.022 (P = 0.448)	-0.062 (P = 0.828)
$y(t)_t$ and $\boldsymbol{x}(t)_t$	0.156	1.806 (P = 0.055)	0.461 (P = 0.140)

TABLE 2. Summary of In-Sample Estimation Results, 1980:Q1-1997:Q3

	Strategy 1	Strategy 1'	Strategy 2	Strategy 3
Employment				
Joint Signif.	0.001	0.001	0.000	0.020
Sum of Coeff.	0.293	0.270	0.488	0.476
Industrial Prod.				
Joint Signif.	0.000	0.000	0.000	0.000
Sum of Coeff.	0.244	0.299	0.228	0.262
Real Retail Sales				
Joint Signif.	0.000	0.000	0.000	0.000
Sum of Coeff.	0.139	0.138	0.119	0.191
Overall				
Adjusted R^2	0.858	0.842	0.794	0.723
Std. Error of Est.	1.085	1.153	1.512	1.755
Signif. of GQ	0.097	0.179	0.030	0.189
Significance of Q	0.324	0.003	0.787	0.723

TABLE 3. Summary Statistics for Recursive Forecasting Exercise, 1990:Q1-1997:Q4

Estimation Strategy	Mean Error	Mean Absolute Error	Root Mean-Sq. Error
Strategy 1	0.19	1.14	1.40
Strategy 1'	0.17	1.29	1.59
Strategy 2	0.19	1.37	1.62
Strategy 3	0.05	1.45	1.71
Naive	-0.58	1.53	1.86
Autoregression	0.20	1.72	2.15
Blue Chip Consensus	0.37	1.22	1.48

TABLE 4. Encompassing Tests for Recursive Forecasts of Current-Vintage GDP Growth, 1990:Q1-1997:Q4

	$y(T)$ - Strategy#i = θBlueChip			$y(T)$ - BlueChip = θStrategy#i	
Strategy #	θ	**θ = 0?**		θ	**θ = 0?**
1	0.060 (0.100)	P = 0.553		0.183 (0.087)	P = 0.043
1'	0.018 (0.115)	P = 0.877		0.142 (0.086)	P = 0.109
2	0.037 (0.117)	P = 0.752		0.167 (0.082)	P = 0.051
3	-0.029 (0.123)	P = 0.818		0.134 (0.081)	P = 0.106

TABLE 5. Summary Statistics for Recursive Forecasting Exercises, 1990:Q1-1997:Q4

Estimation Strategy	Mean Error	Mean Absolute Error	Root Mean-Sq. Error
Advance GDP			
Strategy 1	0.13	0.67	0.78
Strategy 3	-0.01	0.91	1.07
Blue Chip	0.31	0.83	0.99
Final GDP			
Strategy 1	0.16	0.84	1.02
Strategy 3	0.02	1.12	1.27
Blue Chip	0.34	0.95	1.20

TABLE A.1. Available left-hand-side-variable data laid out in an array. A typical entry, $y(t)_s$, is the official estimate of $y(t)$ released at time $s \geq t$. Strategy 1 uses only entries in bold-faced type (along the main diagonal) for estimation of the forecasting equation. Strategies 2 and 3 use only entries from the right-most column.

$\boldsymbol{y(0)_0}$	$y(0)_1$	$y(0)_2$	.	.	.	$y(0)_{T-2}$	$y(0)_{T-1}$	
	$\boldsymbol{y(1)_1}$	$y(1)_2$	.	.	.	$y(1)_{T-2}$	$y(1)_{T-1}$	
		$\boldsymbol{y(2)_2}$	.	.	.	$y(2)_{T-2}$	$y(2)_{T-1}$	
						.	.	more recent
						.	.	periods
						.	.	↓
						$\boldsymbol{y(T\text{-}2)_{T-2}}$	$y(T\text{-}2)_{T-1}$	
							$\boldsymbol{y(T\text{-}1)_{T-1}}$	

more recent vintages →

TABLE A.2. Available right-hand-side-variable data laid out in an array. In the notation of the main text, $x(t)_s = [e(t)_s\ e(t - 1)_s\ e(t - 2)_s]$. A typical entry of the array, $e(t)_s$, is the official estimate of $e(t)$ released at time $s \geq t$. Strategies 1 and 2 use only the entries in bold-faced type for estimation. The bottom three entries of the right-most column are then substituted into the estimated equation to obtain a forecasted value of y(T). Strategy 3 uses only entries from the right-most column for estimation and forecasting. Strategy 1' uses only entries from the main diagonal.

$\boldsymbol{e(\text{-}2)_0}$	$e(\text{-}2)_1$	$e(\text{-}2)_2$	.	.	.	$e(\text{-}2)_{T-2}$	$e(\text{-}2)_{T-1}$	$e(\text{-}2)_T$	
$\boldsymbol{e(\text{-}1)_0}$	$\boldsymbol{e(\text{-}1)_1}$	$e(\text{-}1)_2$	.	.	.	$e(\text{-}1)_{T-2}$	$e(\text{-}1)_{T-1}$	$e(\text{-}1)_T$	
$\boldsymbol{e(0)_0}$	$\boldsymbol{e(0)_1}$	$\boldsymbol{e(0)_2}$	.	.	.	$e(0)_{T-2}$	$e(0)_{T-1}$	$e(0)_T$	
	$\boldsymbol{e(1)_1}$	$\boldsymbol{e(1)_2}$	.	.	.	$e(1)_{T-2}$	$e(1)_{T-1}$	$e(1)_T$	
		$\boldsymbol{e(2)_2}$	.	.	.	$e(2)_{T-2}$	$e(2)_{T-1}$	$e(2)_T$	
						.	.	.	more recent
						.	.	.	periods
						$\boldsymbol{e(T\text{-}4)_{T-2}}$	.	.	↓
						$\boldsymbol{e(T\text{-}3)_{T-2}}$	$\boldsymbol{e(T\text{-}3)_{T-1}}$	.	
						$\boldsymbol{e(T\text{-}2)_{T-2}}$	$\boldsymbol{e(T\text{-}2)_{T-1}}$	$e(T\text{-}2)_T$	
							$\boldsymbol{e(T\text{-}1)_{T-1}}$	$e(T\text{-}1)_T$	
								$e(T)_T$	

more recent vintages →

TABLE B.1. The distribution of $K/(1 + K)$ for alternative sample sizes, under the assumption that the $x(t)_t$ (t = 0, 1,...T) are independent draws from a normal distribution. Note that $E[K/(1 + K)] = 2/(T + 1)$.

	Percentile										
T	**0**	**10**	**20**	**30**	**40**	**50**	**60**	**70**	**80**	**90**	**Mean**
20	0.0476	0.0484	0.0509	0.0552	0.0617	0.0708	0.0833	0.1013	0.1285	0.1772	0.0952
30	0.0323	0.0328	0.0344	0.0373	0.0415	0.0475	0.0560	0.0680	0.0864	0.1197	0.0645
40	0.0244	0.0248	0.0260	0.0281	0.0313	0.0358	0.0421	0.0512	0.0651	0.0906	0.0488
50	0.0196	0.0199	0.0209	0.0226	0.0252	0.0288	0.0338	0.0411	0.0522	0.0728	0.0392
60	0.0164	0.0167	0.0175	0.0189	0.0210	0.0240	0.0282	0.0343	0.0436	0.0608	0.0328
70	0.0141	0.0143	0.0150	0.0162	0.0180	0.0206	0.0242	0.0294	0.0374	0.0522	0.0282
80	0.0123	0.0125	0.0132	0.0142	0.0158	0.0181	0.0212	0.0258	0.0328	0.0458	0.0247
90	0.0110	0.0112	0.0117	0.0126	0.0141	0.0161	0.0189	0.0229	0.0292	0.0408	0.0220
100	0.0099	0.0101	0.0105	0.0114	0.0127	0.0145	0.0170	0.0206	0.0263	0.0367	0.0198
200	0.0050	0.0051	0.0053	0.0057	0.0064	0.0072	0.0085	0.0103	0.0132	0.0184	0.0100

Notes:

1. See, for example, Fitzgerald and Miller (1989), Trehan (1989, 1992), Diebold and Rudebusch (1991), Miller and Chin (1996), and Robertson and Tallman (1998). Our "end-of-sample-vintage data" are what Swanson (1996) calls "partially revised" or "real-time" data.

2. Ehrbeck and Waldmann (1996) and Laster, Bennett, and Geoum (1999) also argue that professional forecasts are typically suboptimal. Their arguments depend on strategic behavior by forecasters, and are completely different from the arguments advanced here.

3. The list of researchers who have developed forecasting models for current-quarter GDP includes Fitzgerald and Miller (1989), Trehan (1989, 1992), Braun (1990), Ingenito and Trehan (1996), and Miller and Chin (1996). Zadrozny (1990) and Rathjens and Robins (1993) use monthly data to improve forecasts of *next* quarter's output growth.

4. Real-world employment data are available almost a month before the first official estimate of GDP growth is released.

5. Suppose that the truth is revealed after $S > 0$ periods, so that $y(t)_{t+S} = y(t)$. Then Equation 1 can be estimated over a sample period that runs from $t = 0$ to $t = T - S - 1$. We argue that even when it is possible, direct estimation of Equation 1 is likely to be suboptimal.

6. Equivalently, we assume that $y(t)_t$ is the mathematical expectation of both $y(t)$ and $y(t)_s$ given the government's period-t information set.

7. A more complete discussion of the data requirements of the different estimation strategies is provided in an appendix.

8. Consider, for example, the extreme case in which $y(t) = x(t) + z(t)$ where $x(t)$ and $z(t)$ are both white noise. Suppose that $x(t)$ is observed by both the government and the private sector at time t (so that $x(t)_t = x(t)$), but $z(t)$ is not observed until later. In the notation used above, $\alpha = 1$, $y(t)_t = x(t)\alpha$, $\omega(t) \equiv 0$ and $w(t) = \nu(t) = z(t)$. Clearly, in any finite sample the analyst will do better to "estimate" the equation $y(t)_t = x(t)\alpha$ (which fits the data perfectly) than to estimate the equation $y(t) = x(t)\alpha + w(t)$, although both strategies yield unbiased estimates of α.

9. The exception is when revisions to the government's methodology for calculating $y(t)$ are so great as to shift α. Ideally, the analyst would apply the government's latest methodology retroactively, using real-time vintage source data to obtain methodologically-consistent series for $y(t)_t$ and $\boldsymbol{x}(t)_t$. A more practical alternative is to test for structural breaks coincident with major methodological revisions and–if necessary–introduce one or more dummy variables on the right-hand side of Equation 3. Because it assumes that α is constant over time, our empirical analysis tends, if anything, to put Strategy 1 at a disadvantage relative to the usual estimation technique (Strategy 3).

10. As noted above, our primary focus is on choosing the proper vintage data, taking the set of right-hand-side variables as "given." Variable selection clearly influences the performance of a Strategy-1 forecasting model through its effects on the variance of ω in Equation 7. It follows

that the analyst using Strategy 1 will want to choose right-hand-side variables that are highly correlated with the variables used by the government in making its initial estimate of $y(t)$.

11. Basically, the difference between the Strategy 1 and Strategy 1' error terms is a moving average of lagged revisions. An example is provided in an appendix.

12. We assume that the release of $y(t)_T$ is delayed relative to the release of $\boldsymbol{x}(t)_T$ just as the release of $y(T)_T$ is delayed relative to the release of $\boldsymbol{x}(T)_T$. Hence $y(t)_{T-1}$ is the most up-to-date available estimate of $y(t)$.

13. Thus, the error term in Equation 3' can be expected to be heteroscedastic. Likely serial correlation in $\nu(t)_{T-1}$ means that $\omega'(t)$ will also typically be serially correlated.

14. The drop-off in precision will be particularly great if, in estimating Equation 3', heteroscedasticity and serial correlation are not properly taken into account. See note #13.

15. A more detailed look at Strategy-2's forecasting performance relative to that of Strategy 1 is presented in an appendix.

16. Heteroscedasticity is a likely complication. Data from the early part of the sample will have undergone more extensive revision than data near the end of the sample. Consequently, it would not be surprising to see var[$\zeta(t)$] decline as t approaches T - 1. Insofar as an upward revision to one quarter's data is also typically accompanied by upward revisions to data from neighboring

quarters, $\zeta(t)$ will also be serially correlated. See Croushore and Stark (1999) for a thorough documentation of the properties of revisions to several macroeconomic data series.

17. The final component of the coincident index–real personal income–is released substantially later than the others. For a review of the timing of various statistical releases and their relationship to GDP, see Rogers (1998).

18. Based partly on findings reported in Koenig (1996) and Fitzgerald and Miller (1989), we tried including lagged GDP growth, manufacturing capacity utilization, the aggregate hours of workers in the service-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.

19. If $x_{t,s}$ is a monthly variable, we define $x_{t,0} \equiv x_{t-1,3}$ and $x_{t,-1} \equiv x_{t-1,2}$. Note that this specification assumes that a full three months of current-quarter employment, industrial production, and retail sales data are available. Restricting the information set to one or two months of current-quarter data results in poorer forecasting performance for each alternative estimation strategy, but does not change the relative ranking of the estimation strategies. (Results available on request.)

20. 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

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

Thus, the quarter-to-quarter percentage change in output is a weighted average of five month-to-month percentage changes in the coincident indicator. In our GDP model, the exact pattern of weights suggested by this example was rejected in formal statistical tests, so we left the coefficient weights attached to the right-hand-side variables unrestricted in our regressions.

21. Ellis Tallman graciously provided industrial-production data. Other data were culled from a variety of official government sources.

22. Data are actually measured as of the middle of the month following the close of the quarter. The initial official GDP estimate is released about two weeks later.

23. The end-of-sample vintage GDP data were graciously provided by Dean Croushore. The Federal Reserve Bank of Philadelphia maintains an extensive on-line, real-time data base.

24. When the information set is restricted to a constant and the initial GDP release, coefficients are small in magnitude and statistically insignificant at conventional levels.

25. The AR results are for a model estimated using Strategy 1, with a lag length that is selected recursively by the Bayesian Information Criterion. Results change little if the lag length is held fixed as the sample period is extended, or if the AR model is estimated using Strategy 2 or Strategy 3.

26. We conducted similar formal comparisons of the forecasting performance of the monthly

indicators model with that of a real-time autoregression (Note #25). In these tests, the indicators model was found to be encompassing but not encompassed regardless of how it was estimated.

27. See West and McCracken (1998). The West-McCracken result assumes that certain technical conditions are also met. Among them is the requirement that the methodology used to estimate the forecasting model of interest yield unbiased parameter estimates (West and McCracken, Assumption 2). The unbiasedness condition is typically be violated by Strategy 3, but an analyst unfamiliar with our arguments would not be aware of this problem.

28. Including constants in the regressions does not change these results.

29. Encompassing tests similar to those reported in Table 4 indicate that the monthly-indicators forecasts are correlated with Blue Chip forecast errors, but Blue Chip forecasts are not correlated with the forecast errors of the monthly indicators model. These results obtain regardless of whether the indicators model is estimated using Strategy 1 or Strategy 3.