Board of Governors of the Federal Reserve System

International Finance Discussion Papers

Number 684

November 2000

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

Evan F. Koenig, Sheila Dolmas and Jeremy Piger

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

Evan F. Koenig, Sheila Dolmas and Jeremy Piger[*]

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

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 might be used to estimate rolling 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 rolling regressions all estimated with 1999:Q1-vintage data, one uses the prediction of a model estimated with 1990:Q1-vintage data to forecast GDP growth in 1990:Q1, the prediction of a model 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

---

[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.

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 models, we argue that analysts should generally use data of as many different vintages as there are dates in their samples. More specifically, at every point *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 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 (say) and early estimates of employment growth will typically be different from that between GDP growth and estimates of employment growth available several years after the fact. Although the latter relationship may be relevant to the analyst trying to estimate an aggregate production function or other "deep structural" model, 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). Accordingly, the coefficient estimates obtained using the conventional approach are a mix of those relevant to forecasting and those relevant to a deep structural analysis. What we call real-time-

---

[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.

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 time.

What of the choice between end-of-sample data and first-release data for the left-hand-side variable (in our example, real GDP growth)? As long as the government's initial GDP release is an efficient, unbiased estimate of the final release, a model that does a good job of predicting the initial release will also do a good job of predicting the final release. Moreover, insofar as real-time-vintage employment data are more tightly linked to the government's initial GDP release than to revised releases--a condition that we argue is likely to be met--then a forecasting equation estimated with initial-release data on its left-hand side will have lower coefficient uncertainty than an equation estimated with end-of-sample-vintage data on its left-hand side. Even if the government forecasts are not fully efficient, they may be efficient 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).

---

[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.

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. As we demonstrate, the key to our model's strong performance is the fact that we estimate it with real-time-vintage data, rather than end-of-sample-vintage data. The point is that the vintage or vintages of data one uses when estimating a model is important--sometimes as important as the choice of right-hand-side variables--in determining real-time forecasting success. A relatively simple model, correctly estimated, may match the performance of a more sophisticated model estimated conventionally.

II. The Basic Forecasting Problem

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

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

$\mathbf{x}(t) \equiv$ the "true" period-t value of $\mathbf{x}$, $0 \le t \le T$

$y_R(t) \equiv$ the official estimate of $y(t)$ released at time $t + R$, where $R \ge 0$

$v_R(t) \equiv y_R(t) - y_0(t)$

$v(t) \equiv y(t) - y_0(t)$

$\mathbf{x}_r(t) \equiv$ the official estimate of $\mathbf{x}(t)$ available at time $t + r \ge t$

$\xi_r(t) \equiv \mathbf{x}_r(t) - \mathbf{x}_0(t)$

$z(t) \equiv$ a variable observed by the government at time t, but not the econometrician

For concreteness, the reader may want to think of $\mathbf{x}_r(t)$ as an official estimate of employment growth during quarter t and $y_R(t)$ as an official estimate of GDP growth. The analyst is trying to forecast GDP growth in quarter T ($y(T)$). No official estimate of quarter-T GDP growth has yet

been released ($y_0(T)$ is, as yet, unavailable). However, an initial estimate ($\mathbf{x}_0(T)$) of quarter-T employment growth *is* available.[4] Also available are various official estimates of GDP growth ($y_R(t)$, for $0 \leq t < T$ and $0 \leq R < T - t$) and employment growth ($\mathbf{x}_r(t)$, for $0 \leq t < T$ and $0 \leq r \leq T - t$) in earlier quarters.

We put some structure on the forecasting problem by assuming that the government's initial estimate of $y(t)$ is linearly related to $\mathbf{x}_0(t)$ and $z(t)$:

$$y_0(t) = \mathbf{x}_0(t)\beta + z(t). \tag{1}$$

In general, $z(t)$ will be correlated with some of the components of $\mathbf{x}_0(t)$. Thus,

$$z(t) = \mathbf{x}_0(t)\delta + \omega(t), \tag{2}$$

where δ is a parameter vector and $\omega(t)$ is a mean-zero random disturbance orthogonal to $\mathbf{x}_0(t)$. We assume that $\omega(t)$ is i.i.d.

We also assume 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 require that $v(t)$ and $v_R(t)$ be uncorrelated with any information available to the government at time t. *A fortiori*, $v(t)$ and $v_R(t)$ are uncorrelated with any information available to private forecasters at time t. While 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. Assuming that $v(t)$ is uncorrelated with all information available to the government at time t amounts to assuming that $y_0(t)$ is the mathematical expectation of $y(t)$ given the government's information set.

---

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

III. Alternative Estimation and Forecasting Strategies

We consider three alternative estimation strategies, in turn. The first strategy is to regress $y_0(t)$ on $\mathbf{x}_0(t)$ and to obtain a forecast of $y(T)$ by evaluating the estimated equation at $\mathbf{x}_0(T)$. The second strategy is similar, but uses an end-of-sample-vintage official estimate of $y(t)$ as the left-hand-side variable instead of $y_0(t)$. The third strategy--used by most real-world forecasters--is to use end-of-sample-vintage data on both the left-hand side and right-hand side of the estimated forecasting equation.[5] Note that 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.

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 that point) on its right-hand side. To see the rationale for this approach, note that Equations 1 and 2 imply that

$$y_0(t) = \mathbf{x}_0(t)\alpha + \omega(t). \tag{3}$$

where $\alpha \equiv \beta + \delta$. Since $\mathbf{x}_0(t)$ and $\omega(t)$ are uncorrelated, ordinary least squares applied to Equation 3 yields an unbiased and consistent estimate of α. Unless $\omega(t)$ is serially correlated or heteroscedastic--and there is no *a priori* reason to expect that it will be either--OLS is also efficient.[6]

---

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

[6] Revisions to the government's methodology may sometimes be so great as to shift α. Ideally, the analyst would apply the government's end-of-sample methodology retroactively to real-time vintage source data, to obtain methodologically-consistent series for $y_0(t)$ and $\mathbf{x}_0(t)$. 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, if anything, tends to put Strategy 1 at a disadvantage relative to the usual estimation technique (Strategy 3).

The forecast of y(T) obtained under Strategy 1 is

$$\hat{y}(T) = \hat{y}_0(T) \equiv \boldsymbol{x}_0(T)\hat{\alpha}, \tag{4}$$

where a hat denotes a predicted or fitted value. The forecast error,

$$y(T) - \hat{y}(T) = [y(T) - y_0(T)] + [y_0(T) - \hat{y}_0(T)], \tag{5}$$

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 mean 0. Moreover, under our assumption that the government uses available information efficiently, the two components must be uncorrelated. Hence,

$$var[y(T) - \hat{y}(T)] = var[y(T) - y_0(T)] + var[y_0(T) - \hat{y}_0(T)]$$

$$= \sigma_v^2 + \sigma_\omega^2 + var[\boldsymbol{x}_0(T)(\alpha - \hat{\alpha})]. \tag{6}$$

The final term in Equation 6 is the variance due to coefficient uncertainty in the estimated version of Equation 3. For any fixed $\mathbf{x}_0(T)$, 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 $\mathbf{x}_0(T)$ and the sample mean of the $\mathbf{x}_0(t)$ (t = 0, 1, ...T - 1). The simplest case is that in which $\mathbf{x}_0(T)$ and the sample mean of the $\mathbf{x}_0(t)$ coincide. Equation 6 then reduces to

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

We will use this case as a benchmark when comparing the performance of Strategy 1 to the performance of alternative estimation strategies.[7]

Strategy 1': First-Release Data on Both the Left-Hand Side and Right-Hand Side. A special case of Strategy-1 estimation arrises when the vector $\mathbf{x}_0(t)$ on the right-hand-side of Equation 3 is restricted to include only current and lagged first-release data. By entirely eliminating the need to keep track of revisions, this artificial restriction on the set of information used by the analyst 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 $\mathbf{x}_0(t)$ to first-release data may induce serial correlation in the error term in Equation 3.[8]

Although inefficient, restricting $\mathbf{x}_0(t)$ 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. Given that we are ultimately interested in forecasting the *true* value of y at time T (y(T)), it might seem sensible to use an end-of-sample-vintage estimate of y on the left-hand side of the estimated forecasting equation, rather than the initial official estimate. Formally, it might seem sensible to

---

[7] 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).

[8] 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.

replace $y_0(t)$ with $y_{T-t-1}(t)$ on the left-hand side of Equation 3, under the presumption that $\text{var}[y(t) - y_R(t)]$ is decreasing in R.[9] However, this strategy is inefficient. To see why, add $v_{T-t-1}(t) = y_{T-t-1}(t) - y_0(t)$ to both sides of Equation 3 to get a relationship between $y_{T-t-1}(t)$ and $\mathbf{x}_0(t)$:

$$y_{T-t-1}(t) = \mathbf{x}_0(t)\alpha + \omega'(t) \tag{3'}$$

for $0 \le t < T$, where $\omega'(t) \equiv \omega(t) + v_{T-t-1}(t)$. Under our assumption that $v_{T-t-1}(t)$ is uncorrelated with information available to the government at time t, applying least squares to Equation 3' will certainly yield an unbiased estimate of α. However, since $\omega(t) = z(t) - \mathbf{x}_0(t)\delta$ and both $z(t)$ and $\mathbf{x}_0(t)$ are in the government's period-t information set, $\omega(t)$ must be uncorrelated with $v_{T-t-1}(t)$. 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:

$$\text{var}[\omega'(t)] = \text{var}[\omega(t)] + \text{var}[v_{T-t-1}(t)] \ge \text{var}[\omega(t)]. \tag{8}$$

Equality holds in Equation 8 for $t = T - 1$. (Recall that $v_{T-t-1}(T-1) = y_0(T-1) - y_0(T-1) = 0$.) Except in the uninteresting special case in which the government's initial estimate of $y(t)$ is never revised (so that Strategies 1 and 2 are identical), *strict* inequality will obtain for at least one $t < T - 1$. 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.[10] Hence, for any $R > 0$, the relationship between $y_R(t)$ and $\mathbf{x}_0(t)$ is necessarily looser than that between $y_0(t)$ and $\mathbf{x}_0(t)$.

---

[9] We assume that release of $y_{T-t}(t)$ is delayed relative to the release of $\mathbf{x}_{T-t}(t)$, just as the release of $y_0(T)$ is delayed relative to the release of $\mathbf{x}_0(T)$. Hence $y_{T-t-1}(t)$ is the most up-to-date available estimate of $y(t)$.

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

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 to the latter.[11] The Strategy-2 counterpart to Equation 6 is

$$var[y(T) - \hat{y}'(T)] = var[y(T) - y_0(T)] + var[y_0(T) - \hat{y}'_0(T)]$$

$$= \sigma_v^2 + \sigma_\omega^2 + var[\boldsymbol{x}_0(T)(\alpha - \hat{\alpha}')], \tag{6'}$$

where

$$\hat{y}'(T) = \hat{y}'_0(T) \equiv \boldsymbol{x}_0(T)\hat{\alpha}', \tag{4'}$$

and 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 $\mathbf{x}_0(T)$ and the sample mean of the $\mathbf{x}_0(t)$ coincide, for example, Equation 6' reduces to

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

where

---

[11] 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 #10.

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

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.

Summarizing: If $y_0(T)$ is an efficient estimate of $y(T)$, then the best forecast of $y_0(T)$ will also be the best forecast of $y(T)$. But OLS applied to Equation 3 yields the lowest-variance unbiased forecast of $y_0(T)$, conditional on $\mathbf{x}_0(T)$. Hence, Strategy 1 yields the lowest-variance unbiased forecast of $y(T)$.[12] The forecasts generated by Strategy 2 are unbiased, but inefficient.

Strategy 3: End-of-Sample-Vintage Data on Both Left and Right. As indicated 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

$$y_{T-t-1}(t) = \mathbf{x}_{T-t}(t)\alpha + \omega''(t) \tag{3''}$$

for $0 \le t \le T - 1$, where $\omega''(t) \equiv \omega(t) + v_{T-t-1}(t) - \xi_{T-t}(t)\alpha$. We have already assumed that $y_0(t)$ is an efficient estimate of $y_{T-t-1}(t)$. We now, similarly, assume that $\mathbf{x}_0(t)$ is an efficient estimate of $\mathbf{x}_{T-t}(t)$. It follows that neither $v_{T-t-1}(t)$ nor $\xi_{T-t}(t)\alpha$ will be correlated with $\omega(t)$. Hence

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

---

[12] Implicitly, we assume that the choice between Strategies 1 and 2 is "either-or." Alternatively, one might view Equations 3 and 3' as an error-component model combining cross-section and time-series data (Kmenta 1986, pp. 625-30).

where $\zeta(t) \equiv v_{T-t-1}(t) - \xi_{T-t}(t)\alpha$. 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 $\xi_{T-t}(t)\alpha$ 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.[13]

Correlation between $\zeta(t)$ and $\mathbf{x}_{T-t}(t)\alpha$ in Equation 3" is likely. Only if variation in $v_{T-t-1}(t)$ completely offsets variation in $\xi_{T-t}(t)\alpha$ 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 $\xi_{T-t}(t)$) may be correlated with the left-hand-side-variable error ($v_{T-t-1}(t)$).

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 $\Sigma_{xx} \equiv plim[\mathbf{X}_0'\mathbf{X}_0/T]$, $\Sigma_{\xi\xi} \equiv plim[\Xi'\Xi/T]$, and $\mathbf{X}_0$ and Ξ are T× k matrices whose tth rows are $\mathbf{x}_0(t)$ and $\xi_{T-t}(t)$, respectively. Thus, the least-squares coefficient estimates are inconsistent. More generally,

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

---

[13] 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 data series.

where $\Sigma_{\xi\nu} \equiv \text{plim}[\Xi'N/T]$ and N is a $T \times 1$ vector whose tth element is $\nu_{T-t-1}(t)$. While it is theoretically possible for the term in square brackets to equal zero, there are no general *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 $\mathbf{x}_0(T)$ is of absolutely no use in forecasting $y(T)$. Equation 11 says that estimating a relationship between y and **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 **x**. Hence, one of the complaints that has been voiced about 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 1999:Q1-vintage 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. The coefficient estimates so obtained will be unbiased. We argued that in many cases one can expect to obtain more precise coefficient

13

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. 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. 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 are not forecastable 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).[14,15]

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:

---

[14] The final component of the coincident index--real personal income--is released much later than the employment, retail sales, CPI, and industrial production reports. For a review of the timing of various statistical releases and their relationship to GDP, see Rogers (1998).

[15] 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 were statistically significant, and we dropped them from our analysis.

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

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.[16, 17]

      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 rolling 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

---

      [16] 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.)

      [17] Equation 12 can be motivated 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

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

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, the exact pattern of weights suggested by this simple example are rejected in formal statistical tests, so we leave the coefficient weights attached to the right-hand-side variables unrestricted in our regressions.

1997:Q4.[18] 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.[19]

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.[20] 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.

Estimation Results. As noted above, an analyst in real time must choose between three main alternative strategies when estimating a forecasting equation. Strategy 1 is to use first-release data on the left-hand side of the forecasting equation and real-time-vintage data on the right-hand side. Least-squares coefficient estimates are unbiased and efficient in this case. Strategy 2 is to use end-of-sample vintage data on the left-hand side of the equation and real-time-vintage data on the right-

---

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

[19] 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.

[20] 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.

hand side. We argued that least-squares estimation can be expected to yield unbiased but inefficient coefficient estimates under this strategy, and that heteroscedasticity may well be a problem. Strategy 3--the most commonly used approach, in practice--is to estimate the forecasting equation with end-of-sample vintage data on both sides. Least-squares coefficient estimates are inconsistent, and heteroscedasticity is again a threat.

An important element of our theoretical argument favoring Strategy 1 is the assumption that the government makes efficient use of available information when forming its initial estimate of y(t). In the case at hand, a simple test of efficiency is obtained by regressing revisions to the government's initial GDP estimate on a constant and the initial release itself. Efficiency requires that neither the constant nor the initial release be statistically significant in this regression. Results are as follows:

$$v_{T-t-1}(t) = 0.230 + 0.015y_0(t) \qquad\qquad R^2 = 0.001$$
$$(0.247) \ \ (0.067) \qquad\qquad\qquad\quad S.E. = 1.603$$

for T = 1999:Q1 and 1980:Q1 $\leq$ t $\leq$ 1997:Q3. (Standard errors of the estimated coefficients are in parentheses.) Clearly, efficiency of the government's initial estimates is not rejected over our sample period.[21] We therefore have reason to hope that our theoretical results will hold for the GDP forecasting problem considered here.

Table 1 compares the in-sample performance of the three main estimation strategies as applied to Equation 12 over the period 1980:Q1-1997:Q3. In addition, it presents results for Strategy 1'--the special case of Strategy-1 estimation in which the set of real-time-vintage data is restricted to include only initial official estimates. For each strategy, 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. Although the latter test gives a clear warning signal only for Strategy 2, all standard errors reported in the table are adjusted to correct for possible heteroscedasticity.

---

[21] Similar results are reported by Mankiw and Shapiro (1986). If the sample period is extended back into the 1960s, efficiency is rejected (Runkle 1998).

Monthly percentage changes in employment, industrial production, and retail sales are highly statistically significant in every regression. As expected, the Strategy-1 regression has the lowest standard error. That the Strategy 1 regression has a lower standard error than Strategy 2 shows that real-time-vintage monthly employment, industrial production, and retail sales data are more closely linked to first-release GDP data than to GDP data that have been extensively revised. Serial correlation is a problem only for Strategy 1', where it is expected. (See the Appendix.)

TABLE 1. 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

Forecast Comparisons: Alternative Estimation Strategies. To rank our alternative estimation strategies, we conducted a series of rolling-sample forecasting exercises. Coefficient estimates

obtained using data through 1989:Q4 were used to forecast real GDP growth in 1990:Q1.[22] 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 real-GDP growth,*

where "actual real-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 rolling forecasting exercises are displayed in Table 2 for models estimated

using Strategies 1, 1', 2, and 3. Table 2 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.

TABLE 2. Summary Statistics for Rolling-Sample 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
Blue Chip Consensus	0.37	1.22	1.48

---

[22] Formally, $\mathbf{x}_0(T)$ was substituted into the right-hand-sides of Equations 3, 3', and 3"
after these equations were estimated.

As expected, Strategy 1 (which uses first-release data on the left-hand side and real-time-vintage data on the right-hand side of the forecasting equation) beats out both Strategy 1' (which uses first-release data on both left-hand and right-hand sides) 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 (which is the most commonly used method of real-time estimation) 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 2 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 errors that are generated by estimation Strategy 3. In contrast, the naive approach gives a markedly too-pessimistic view of the real-time performance of estimation Strategy 1.

The primary objective of our empirical analysis is to demonstrate that, for a *given* set of right-hand-side variables, the vintage data one uses to estimate a forecasting equation can have a substantial impact on the equation's performance. The final row of Table 2 illustrates a related, but secondary, point: that a very simple model, correctly estimated, can sometimes match the real-time performance of experienced, professional forecasters using conventional estimation techniques. Thus, by standard measures, 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.

Are the performance differences between Strategy 1 and the alternatives significant? Table 3 presents results from tests of marginal predictive power. Current-vintage GDP growth is regressed on a constant and the forecasts generated by each of two alternative estimation strategies or models. If a forecast is statistically significant in one of these regressions, then it contains valuable information beyond that captured by the other forecast (Fair and Shiller 1990). The table shows that none of the alternative models or estimation strategies has predictive power beyond that of our simple model estimated with real-time vintage data (Strategy 1). In no case can one reject the hypothesis that the Strategy-1 forecast enters with a coefficient of 1 and that both the constant term

and the alternative forecast have coefficients that equal 0. In contrast, only in competition with the Blue Chip forecast is the Strategy-1 forecast statistically insignificant.[23]

TABLE 3. Tests of Marginal Predictive Power

Constant	Strategy 1	Strategy 1'	Strategy 2	Strategy 3	Blue Chip	Joint Test[†]
0.344 (0.406)	0.957* (0.388)	-0.026 (0.339)	---	---	---	0.850
0.333 (0.409)	0.999* (0.388)	---	-0.063 (0.338)	---	---	0.843
0.303 (0.393)	1.511** (0.422)	---	---	-0.528 (0.364)	---	0.414
0.144 (0.443)	0.595 (0.355)	---	---	---	0.464 (0.450)	0.603

† Significance level for test of the hypothesis that the Strategy-1 forecast enters with a coefficient of 1, and both the constant term and the alternative forecast have coefficients that equal 0.
* Significant at 5% level.   ** Significant at 1% level.

_____

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 2 and 3 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

_____

[23] In encompassing tests--where the coefficients of the two forecasts are restricted to sum to 1 and the constant term is restricted to equal 0 (Chong & Hendry 1986)--Strategy 1 is strictly preferred to the Blue Chip and every other alternative.

become available with one-month and three-month lags, respectively. Results are presented in Tables 4 and 5, which are similar in format to Tables 2 and 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.

TABLE 4. Summary Statistics for Rolling-Sample Forecasting Exercise, 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 5. Tests of Marginal Predictive Power

Constant	Strategy 1	Strategy 3	Blue Chip	Joint Test[†]
Advance GDP				
0.302 (0.223)	0.938** (0.239)	-0.016 (0.207)	---	0.607
0.224 (0.245)	0.786** (0.196)	---	0.185 (0.249)	0.496
Final GDP				
0.192 (0.297)	1.090** (0.319)	-0.098 (0.275)	---	0.827
0.044 (0.324)	0.721** (0.260)	---	0.361 (0.329)	0.581

† Significance level for test of the hypothesis that the Strategy-1 forecast enters with a coefficient of 1, and both the constant term and the alternative forecast have coefficients that equal 0.
* Significant at 5% level.   ** Significant at 1% level.

Two main conclusions emerge from Tables 4 and 5. First, early GDP releases are substantially easier to predict than are late releases. Thus, the root-mean-square errors recorded in Table 4 are uniformily lower than the corresponding root-mean-square errors in Table 2. Within Table 4, 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 earlier, in Table 1. The second main conclusion to emerge from the tables is that the relative ranking of the alternative forecasts is unaffected by whether the goal is to predict an early GDP release or a late GDP release. Indeed, the case for our preferred estimation strategy only becomes stronger when the goal is to achieve accurate forecasts of an early GDP release. Thus, according to Table 5, our Strategy-1 estimation methodology yields forecasts of advance and final GDP growth that unambiguously dominate the corresponding Strategy-3 and Blue-Chip-consensus forecasts.

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 will 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 point within his sample the econometrician estimating a forecasting equation ought to use only right-hand-side data that would have been available at that 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 (where lags extend 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 data 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

23

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 the first variable is useful in predicting the second when, in fact, early vintages of the first variable--the only vintages that are 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. Intuitively, 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 of the left-hand-side variable. Therefore, a regression with first-release data on its left-hand side is likely to 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 is a useful approximation in many real-world 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 is competitive with the Blue Chip consensus GDP forecast.

APPENDIX: 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 Figure 1A, 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 ($y_{T-1}(0)$). And in the lower right-hand corner is the period-T-1 official estimate of period-T-1 GDP growth ($y_0(T - 1)$). 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.[24] 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 Figure 1A, 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 Figure 1A, 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, rolling-forecast exercise--an exercise in 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.

Figure 1B displays an array of employment-growth data of different vintages that is organized exactly like the GDP-growth array of Figure 1A, except that there are two additional rows

---

[24] The format is similar to that of Table 1 in Diebold and Rudebusch (1991).
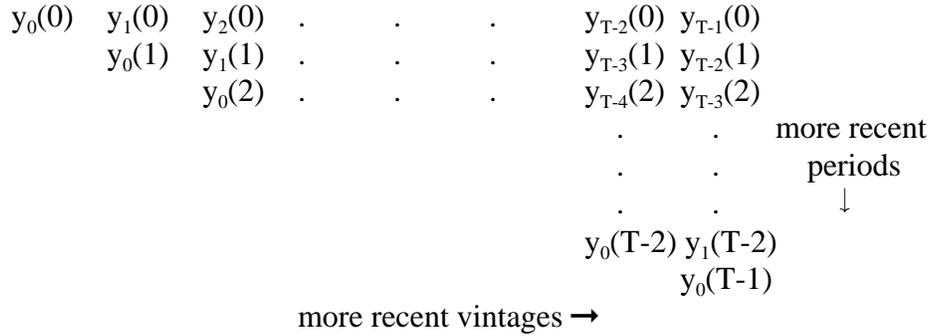
at the top of the array (for employment growth at t = -1 and t = -2). As before, Strategy 3 uses only entries from the extreme right-hand column of the array. If the sample period is extended by one quarter, the old data set must be discarded and replaced with a new set of employment-growth observations of vintage T. Conducting a real-time, rolling-forecast exercise requires that the econometrician collect a sequence of long data sets, each of a different vintage.

Strategies 1 and 2 draw employment data from the bottom *three* elements of each column of the array, reflecting the fact that the current and lagged employment growth that appear on the right-hand side of the forecasting equation are 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 employment growth, and newly revised estimates of period-T-1 and period T-2 employment growth. Conducting a real-time, rolling-forecast exercise requires that the econometrician collect many three-element employment-growth snap shots, each of a different vintage.
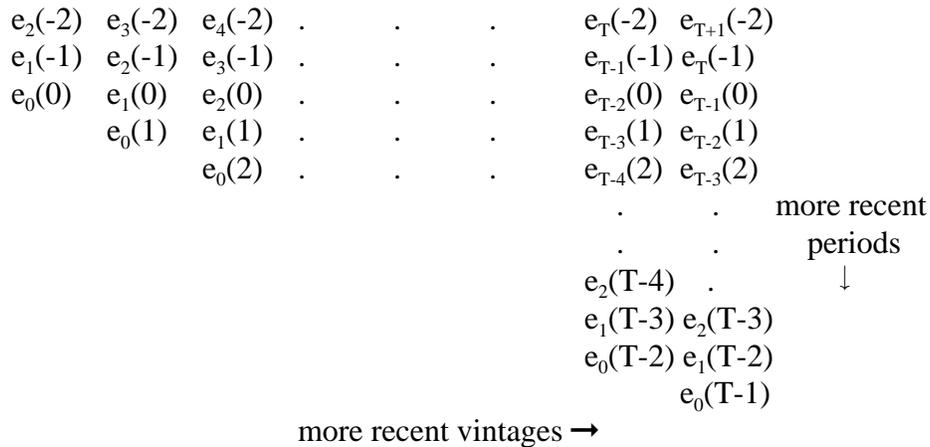
Note that the employment-growth observations used under Strategies 1 and 2 in this example do *not* consist only of first-release data: at each point in the sample, current-quarter employment growth is first release, but lagged employment growth is revised. Hence, in the main text we call the right-hand-side 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 $e_0(t)$, $e_1(t - 1)$, and $e_2(t - 2)$. Under Strategy 1', the right-hand-side variables are $e_0(t)$, $e_0(t - 1)$, and $e_0(t - 2)$. The differences between the right-hand-side variables under the two strategies are thus 0, $e_1(t - 1) - e_0(t - 1)$, and $e_2(t - 2) - e_0(t - 2) = [e_2(t - 2) - e_1(t - 2)] + \mathbf{[e_1(t - 2) - e_0(t - 2)]}$. At time t - 1, the corresponding differences are 0, $\mathbf{e_1(t - 2) - e_0(t - 2)}$, and $e_2(t - 3) - e_0(t - 3)$. The overlap (the terms in boldface 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.

27

FIGURE 1A. Available left-hand-side-variable data laid out in an array. A typical entry, $y_R(t)$, is the official estimate of $y(t)$ released at time $t + R$. Estimation Strategy 1 uses only circled entries (along the main diagonal). Strategies 2 and 3 use only entries from the right-most column.

$$
\begin{array}{llllllll}
y_0(0) & y_1(0) & y_2(0) & . & . & . & y_{T-2}(0) & y_{T-1}(0) \\
 & y_0(1) & y_1(1) & . & . & . & y_{T-3}(1) & y_{T-2}(1) \\
 & & y_0(2) & . & . & . & y_{T-4}(2) & y_{T-3}(2) \\
 & & & & & & . & . \quad \text{more recent} \\
 & & & & & & . & . \quad\quad \text{periods} \\
 & & & & & & . & . \quad\quad\quad \downarrow \\
 & & & & & & y_0(T-2) & y_1(T-2) \\
 & & & & & & & y_0(T-1)
\end{array}
$$

more recent vintages $\rightarrow$

FIGURE 1B. Available right-hand-side-variable data laid out in an array. In the notation of the main text, $x_r(t) = [e_r(t) \; e_{r+1}(t - 1) \; e_{r+2}(t - 2)]$. A typical entry of the array, $e_r(t)$, is the official estimate of $e(t)$ released at time $t + r$. Estimation Strategies 1 and 2 use only the circled entries. Estimation Strategy 3 uses only entries from the right-most column. Strategy 1' uses only entries from the main diagonal.

$$
\begin{array}{llllllll}
e_2(-2) & e_3(-2) & e_4(-2) & . & . & . & e_T(-2) & e_{T+1}(-2) \\
e_1(-1) & e_2(-1) & e_3(-1) & . & . & . & e_{T-1}(-1) & e_T(-1) \\
e_0(0) & e_1(0) & e_2(0) & . & . & . & e_{T-2}(0) & e_{T-1}(0) \\
 & e_0(1) & e_1(1) & . & . & . & e_{T-3}(1) & e_{T-2}(1) \\
 & & e_0(2) & . & . & . & e_{T-4}(2) & e_{T-3}(2) \\
 & & & & & & . & . \quad \text{more recent} \\
 & & & & & & . & . \quad\quad \text{periods} \\
 & & & & & & e_2(T-4) & . \quad\quad\quad \downarrow \\
 & & & & & & e_1(T-3) & e_2(T-3) \\
 & & & & & & e_0(T-2) & e_1(T-2) \\
 & & & & & & & e_0(T-1)
\end{array}
$$

more recent vintages $\rightarrow$

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.

Chong, Yock Y. and David F. Hendry (1986) "Econometric Evaluation of Linear Macroeconomic Models," *Review of Economic Studies* 53, pp. 671-90.

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.

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.

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.

Kmenta, Jan (1986) *Elements of Econometrics*, 2nd edition, Macmillan Publishing Co., New York.

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 Versus 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.

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.

Advance GDP

TABLE 2. Summary Statistics for Rolling-Sample Forecasting Exercise, 1990:Q1-1997:Q4

Estimation Strategy	Mean Error	Mean Absolute Error	Root Mean-Sq. Error
Strategy 1	0.13	0.67	0.78
Strategy 1'	0.10	0.95	1.14
Strategy 2			
Strategy 3	-0.01	0.91	1.07
Naive			
Blue Chip Consensus			

"Final" GDP

TABLE 2. Summary Statistics for Rolling-Sample Forecasting Exercise, 1990:Q1-1997:Q4

Estimation Strategy	Mean Error	Mean Absolute Error	Root Mean-Sq. Error
Strategy 1	0.16	0.84	1.02
Strategy 1'			
Strategy 2	0.22	0.86	1.07
Strategy 3	0.02	1.12	1.27
Naive			
Blue Chip Consensus			