Forecast Accuracy Improvement: Evidence from U.S. Nonfarm Payroll Employment

Allan W. Gregory, Hui (Julia) Zhu
Department of Economics, Queen’s University

May 2008

Abstract

The timing of data release for time series variables for the same time period of observation is often spread over weeks. For instance official government statistics are often released at different times over the quarter or month and yet cover the same time period. This paper focuses on this separation of announcement time or data release and the use of standard econometric updating methods to forecast data that has yet to be made available. The theoretical framework shows that the updating root mean squared forecast error will be accurate with higher correlation coefficients among observation innovations. There are large amounts of applications to earnings forecasts for firms and macroeconomic forecasting. With fixed timing announcement, we find that data on U.S. nonfarm payroll employment are consistent with the theoretical framework. The results highlight an important multivariate forecast accuracy measure.

*The author is grateful to and for helpful comments and suggestions. She also thanks seminar participants at Department of Economics, Queen’s University.

†Address correspondence to Hui (Julia) Zhu, Department of Economics, Queen’s University, Kingston, Ontario, Canada K7L 3N6; E-mail: hzhu@econ.queensu.ca.
1 Introduction

The timing of data release for time series variables for the same time period of observation is often spread over weeks. For instance earning announcements for firms can be spread over a two week period even though these earnings are for the same quarter or month. Further, official government statistics are often released at different times over the quarter or month and yet cover the same time period. This paper focuses on this separation of announcement time or data release and the use of standard econometric updating methods to forecast data that has yet to be made available. To the best of our knowledge this rather important aspect of forecasting which we think has direct application in financial market updating (in terms of earnings, earning per share and so on) has not been studies in the literature.

Traditional forecasting in a multivariate time series setting usually is studied in the context of vector autoregression (VAR) models. In this set-up, some common end-point, the VAR is specified, estimated and single or multiple period forecasts are conducted. Standard errors for the forecasts can be based on asymptotic normal theory or more recently use of bootstrap or some other resampling technique can be applied. The key element for this exercise is that there is a common end point of observed data and that forecasts are made for all variables over some common forecasting horizon.

The situation that we consider is different in that only some (one) of the variables comprising the system is released at a point in time with the remaining variables are release at latter dates. The latter dates may coincide for some variables or it may differ in which case there is a sequence of release times for the system as a whole.

There are two critical exogeniety assumptions in this exercise. The first assumption that the timing of the release information (say earlier or later in the release cycle) does not depend on the information that is released. That is, if firms have poor earnings they might choose to get this information out to prospective investors early (or later). From what we are able to tell about earnings announcements, the decisions as to when to announce is made long in advance of when the earnings information would be credibly known to the firm so this is unlikely to be a problem. The second exogeniety assumption is that there is no feedback on one firm’s earnings with those of another. That is, if one firm announces large earnings in say Q4 for 2007 in February 27 2008, that a related firm does not change its announced earnings for the same quarter which is released on say February 28 2008. For the present application we avoid both of these issues and start a simple bivariate example using nonfarm labor data in the United States.

Comparison with the ordinary VAR forecasts shows that the updating forecast mean squared error (MSE) are smaller than the one in the ordinary VAR forecast. We first consider one real-time variable available case in the bivariate updating forecasts. When the correlation coefficient of the innovations between two variables are lower, the updating forecast has the relative smaller MSE than the ordinary VAR forecast. When the correlation coefficient of observation innovations between two variables are higher, the updating forecast has a significant smaller MSE than the ordinary VAR forecast. In one extreme case, when correlation coefficient approaches to zero, that
is, no contemporaneous information is available to be useful for forecasting, the updating forecast has the exact same result as the ordinary VAR forecast. In another extreme case, when correlation coefficient approaches to one, that is, we have perfect linear association and there is no errors, the updating forecast is the best. Furthermore, we note that not only the correlation coefficient of observation innovation but also the coefficient parameter plays an important role in the multistep ahead forecast. Moreover, this research is the most up-to-date forecasting. To solve data revision problem in the sample, whenever data is available at the time period of forecasting, we adopt it. Finally, Application to nonfarm labor data in the United States shows that the root mean squared error in the updating VAR is less than the ordinary VAR and univariate autoregression. It shows quantitatively that the updating VAR improves 17% in one-step ahead forecast and 12% in two-step ahead forecast respectively comparing to the ordinary VAR.

Related Literature

After Sims’s (1980) influential work, VAR are widely applied in analyzing the dynamics of economic system. For over twenty years, multivariate VAR model has been proven to be powerful and reliable tools in everyday use. Stock and Watson (2001) reassess how well VARs have addressed in data description, forecasting, structural inference, and policy analysis. Working with the inflation-unemployment-interest rate VAR, they conclude that VAR either does no worse than or improves upon the univariate autoregression and that both improve upon the random walk forecast. Therefore, VAR model is now rightly in data description and forecasting. However, the standard VAR forecast does not match the facts that the timing of data release for most financial time series and macro time series for the same time period of observation is often spread over weeks.

Many economic time series are subject to revision. Revisions to measures of real economic activity may occur immediate consequent month or years after official figures are first released. Aruoba (2008) documents the empirical properties of revisions to major macroeconomic variables in the United States. He finds that there revision do not have zero mean, which indicates that the initial announcements by Statistical Agencies are biased. Croushore and Stark (2001, 2003) show how data revisions can affect forecast. They use a real-time data set to analyze data revisions. While the results of some studies are conflicting, Koenig et al. (2003) show first-release data are to be preferred for estimation even if the analyst is ultimately interest in predicting revised data. They provide three alternative strategies for estimating forecasting equations with real time data. They conclude that using first released data in the both sides of equation performs superior forecast to that obtaining final released data. This paper employs the most up-to-date estimates available at forecasting time in addition to the updating VAR forecast method.

The primary use of earnings or macroeconomic forecasts is to provide a proxy for the market expectation of a future realization. Recent work suggests that the stock market reacts to earnings or macroeconomic announcements. Some researches study how markets respond to labor data, such as Krueger (1996) and Boyd et al. (2005). The latter examines how stocks respond to unemployment news, which is measured as the surprise component. Adopting forecasts of the change in the unemployment rate to get the surprise component in the announcement of the unemployment rate,
they find that an announcement of rising unemployment is good news for stocks during economic expansions and bad news during economic contractions. Another line of research focuses on how markets respond to macro variables, such as Flannery et al. (2002), Pesaran and Wickens (1995), and Rapach et al. (2005). Faust et al. (2007) studies how U.S. macroeconomic announcements affect joint movements of exchange rates and interest rates. Using 14-year span of high-frequency data, they conclude that unexpectedly strong announcements lead either to a fall in the risk premium required for holding foreign assets or an expected net depreciation over the ensuing decade, or both.

Role of Current Research

The existing literature on accuracy of time series forecasts has focused on fixed time period. With revision in data, most estimation adopts final released data (end-of-sample). This research adds several new dimensions. First, we provide a practical updating multivariate VAR forecast method, extending the analysis into U.S. nonfarm payroll employment. Second, we construct the most up-to-date data set for estimation and forecast basis. In this paper, first released data plays an important role in estimation and forecast. Finally, we investigate how stock market reacts to current employment situation, which is the first major economic indicator released each month.

The remainder of the paper is organized as follows. Section 2 derives multistep ahead bivariate updating VAR forecasts based on one real-time variable available. In section 3 we discuss the model specification. An application to U.S. nonfarm payroll employment is illustrated in section 4. Section 5 concludes. We collect proofs in the Appendix.

2 A Theoretical Framework

Consider the $N$ multivariate stationary VAR (1) model

$$Y_t = A Y_{t-1} + \epsilon_t, \quad t = 0, \pm 1, \pm 2, \ldots,$$

where $Y_t = (y_{1t}, \ldots, y_{Nt})'$ is a $(N \times 1)$ random vector, the $A$ is fixed $(N \times N)$ coefficient matrices. Moreover, $\epsilon_t = (\epsilon_{1t}, \ldots, \epsilon_{Nt})'$ is a $N$-dimensional white noise or innovation process, that is, $E(\epsilon_t) = 0$ and $E(\epsilon_t \epsilon_t') = \Omega$, with the contemporaneous covariance $C(\epsilon_{it}, \epsilon_{jt}) = \rho_{ij} \sigma_i \sigma_j$ for $i = 1, \ldots, N$ and $j = 1, \ldots, N$. The covariance matrix $\Omega$ is assumed to be nonsingular if not otherwise stated. Since we assume stationary vector autoregressive process, the condition of correlation coefficient $|\rho_{ij}| < 1$ must hold. Finally, $\sigma_i$ is the standard deviation of the innovation $\epsilon_i$. Given the multivariate VAR model (1), the ordinary VAR one-step ahead forecast at time region $T$ is $Y_T(1) = A Y_T$ and the associated forecast error and the corresponding MSE or the covariance matrix of forecast error are $\epsilon_T(1) = \epsilon_{T+1}$ and $MSE_Y$ is in-diagonal covariance matrix of $\Omega'$ respectively, where $Y_T(1)$ denotes the forecast of $Y$ at time $T + 1$, $\epsilon_T(1)$ denotes forecast error at time $T + 1$, and $I$ is $(N \times N)$ identity matrices. The ordinary multivariate VAR forecasts are standard and can be obtained from Lütkepohl (1993) and Hamilton (1994).
Theoretically, multivariate VAR forecast is based on the certain amount of time periods of 1 through $T$. Each equation in multivariate VAR model can be estimated by ordinary least squares (OLS) regression. This OLS estimator is as efficient as maximum likelihood estimator and general least squared estimator. Therefore, the ordinary VAR forecast computed through the unbiased and consistent coefficient estimates and the variance covariance matrix estimates has the lowest MSE and is optimal. However, the fact is that most macroeconomic or financial time series we study in multivariate VAR forecast are not ending at the same time, that is, one variable is generally available for public use a couple days prior to the other variables. For example, financial time series such as firms’ earnings releases for the same quarter or the same year are available at different dates for public use; macro indicators such as inflation, employment, unemployment rate, interest rate, and etc. releases for the same month are also available at different dates for public use. Omitting the timing fact, the ordinary multivariate VAR forecast usually ignores the latest information we can obtain and adopts the same amount of certain time periods to do forecasts.

The focus of this paper is on examining how taking advantage of one or more information about one variable and matching the timing fact to improve multivariate VAR forecasts.

2.1 Bivariate VAR with Multistep Ahead Forecast

A practical method to update forecasts with one real-time variable available in advance is investigated. To simplify the discussion, consider $N = 2$, $y_{1t}$ be observable for $t = 1, \ldots T + 1$, and $y_{2t}$ be observable only up to time $T$. Then the reduced form bivariate VAR (1) model is as follows.

$$
y_{1t} = a_{11}y_{1t-1} + a_{12}y_{2t-1} + \epsilon_{1t}, \quad t = 1, \ldots, T, T + 1
$$

$$
y_{2t} = a_{21}y_{1t-1} + a_{22}y_{2t-1} + \epsilon_{2t}, \quad t = 1, \ldots, T
$$

Following the ordinary VAR forecasting method proposed by Lütkepohl (1993), the one-step ahead ordinary VAR forecast error covariance matrix (or forecast MSE matrix) is

$$
MSE[Y_T(1)] = A^0\Omega_\epsilon(A^0)'
$$

where $Y$ is a vector of $(y_{1t}, y_{2t})'$, $A$ is $2 \times 2$ dimensional coefficient matrix, and $\Omega_\epsilon$ is covariance matrix of $E(\epsilon_t, \epsilon_t')$. That is

$$
A = \begin{pmatrix}
a_{11} & a_{12} \\
a_{21} & a_{22}
\end{pmatrix},
$$

and

$$
\Omega_\epsilon = \begin{pmatrix}
\sigma_1^2 & \rho_{12}\sigma_1\sigma_2 \\
\rho_{12}\sigma_1\sigma_2 & \sigma_2^2
\end{pmatrix}.
$$

With one more real-time information available in advance, the $k = 1$ horizon forecast is performed in the following. Thereafter we denote $MSE^u$ as the updating VAR forecast MSE to distinguish with the ordinary VAR forecast MSE.
Proposition 1  Given the known full information set \( I = \{y_{11}, y_{12}, \ldots, y_{T+1}, y_{21}, \ldots, y_{2T}\} \), the mean squared error of one-step ahead forecast of variable \( y_2 \) in the updating bivariate VAR is

\[
MSE^u[y_{2T}(1)] = Var[y_{2T+1} - \hat{y}_{2T+1}] = Var[\epsilon_2 - \epsilon_{2T}(1)] = (1 - \rho_{12}^2)\sigma_2^2
\]  

(3)

Proof.  See Appendix.

This proposition shows that taking advantage of one more real-time information available in advance, the one-step ahead updating bivariate VAR forecast MSE\(^u\) is \((1 - \rho_{12}^2)\sigma_2^2\). Since \(\rho_{12}\) is coefficient correlation between \(\epsilon_1\) and \(\epsilon_2\), the condition of \(\rho_{12} < 1\) must hold by assumption of stationary vector autoregressive process. However, omitting the extra information \(y_{1T+1}\), the ordinary bivariate VAR forecast MSE gives \(MSE = \sigma_2^2\). Therefore, the one-step ahead updating bivariate VAR forecast has smaller MSE comparing with the forecast from the ordinary bivariate VAR. This implies the updating bivariate VAR forecast is more accurate than the forecast from the ordinary bivariate VAR. Additionally, the higher correlation among the error terms of the variables, the smaller MSE in the updating VAR forecast.

With one more real-time information available in advance, the \( k > 1 \) long-horizon forecast is also examined.

Proposition 2  Given the known full information set \( I = \{y_{11}, y_{12}, \ldots, y_{T+1}, y_{21}, \ldots, y_{2T}\} \), the \( k \)-step ahead forecast mean squared error matrix in the updating bivariate VAR is

\[
MSE^u[Y_T(k)] = \begin{pmatrix} 0 & 0 \\ 0 & (1 - \rho_{12}^2)\sigma_2^2 \end{pmatrix} + \sum_{i=1}^{k-1} A^1\Omega_\epsilon(A^i)'k > 1.
\]  

(4)

Proof.  See Appendix.

This proposition shows that multistep ahead updating bivariate VAR forecast takes advantage of the first step forecast derived from proposition 1. By iterating forward, the \( k \)-step ahead updating bivariate VAR forecast mean squared error matrix MSE\(^u\) in equation (4) is smaller than the recursive ordinary bivariate VAR forecast MSE matrix

\[
MSE[Y_T(k)] = \Sigma_{i=0}^{k-1} A^i\Omega_\epsilon(A^i)' + \sum_{i=1}^{k-1} A^1\Omega_\epsilon(A^i)', \quad k > 1.
\]  

To see this, the difference between the updating VAR forecast MSE\(^u\) and ordinary VAR forecasts
MSE is as follows.

\[
MSE[Y_T(k)] - MSE^u(Y_{T+k}) = \sum_{i=0}^{k-1} A^i \Omega \epsilon(A^i)' - \left( \begin{array}{cc} 0 & 0 \\ 0 & \rho_{12}^2 \sigma_2^2 \end{array} \right) - \sum_{i=1}^{k-1} A^i \Omega \epsilon(A^i)'
\]

\[
= \Omega \epsilon - \left( \begin{array}{cc} 0 & 0 \\ 0 & \rho_{12}^2 \sigma_2^2 \end{array} \right)
\]

\[
= \left( \begin{array}{cc} \sigma_1^2 & \rho_{12} \sigma_1 \sigma_2 \\ \rho_{12} \sigma_1 \sigma_2 & \sigma_2^2 \end{array} \right) - \left( \begin{array}{cc} 0 & 0 \\ 0 & (1 - \rho_{12}^2) \sigma_2^2 \end{array} \right)
\]

\[
> 0
\]

Therefore, the multistep ahead updating bivariate VAR forecast has smaller MSE comparing with the forecast from the ordinary bivariate VAR. This implies the updating bivariate VAR forecast is more accurate than the forecast from the ordinary bivariate VAR. Additionally, the higher correlation among the error terms of the variables, the smaller MSE in the updating VAR forecast.

3 Model Specification

There are alternative ways to model the relationship between two time series under the assumption in section 2. The contemporaneous regression is the traditional approach to study. In the following two subsection we show that the linear regression model is misspecified and that bivariate VAR model gains in efficiency.

To investigate misspecification of contemporaneous regression, consider a reduced form bivariate VAR (1) process,

\[
y_{1t} = a_{11}y_{1t-1} + a_{12}y_{2t-1} + \epsilon_{1t} \\
y_{2t} = a_{21}y_{1t-1} + a_{22}y_{2t-1} + \epsilon_{2t}
\]

We see from (5) that \( y_{2t-1} = (y_{1t} - a_{11}y_{1t-1} - \epsilon_{1t})/a_{12}. \) If we substitute this into (6), we find that

\[
y_{2t} = \frac{a_{22}}{a_{12}} y_{1t} + (a_{21} - \frac{a_{11} a_{22}}{a_{12}}) y_{1t-1} + \epsilon_{2t} - \frac{a_{22}}{a_{12}} \epsilon_{1t}
\]

Since \( y_{1t} \) is correlated to the error term \( \epsilon_{1t} \) by assumption, this expression (7) is equivalent to instrument variables estimation as follows:

\[
y_{2t} = \frac{a_{22}}{a_{12}} y_{1t} + (a_{21} - \frac{a_{11} a_{22}}{a_{12}}) y_{1t-1} + \epsilon_{2t}
\]

where the variables \( y_{2t-1}^u \) is not actually observed. Instead, we observe

\[
y_{1t} = y_{2t-1}^u + \epsilon_{1t}.
\]

Here \( \epsilon_{1t} \) is measurement error which is assumed to be identified, independent, and distributed with
variance $\sigma_1^2$ and to be independent of $y_{1t-1}$ and $y_{2t}$. It is also assumed that there is contemporaneous
correlation of $\epsilon_{1t}$ and $\epsilon_{2t}$. Therefore, $E(\epsilon_{1t}\epsilon_{2t}) = \rho_{12}$ for some correlation coefficient $\rho_{12}$ such that
$-1 < \rho_{12} < 1$.

If we suppose that the true DGP is the VAR (1) model, that is a special case of (8) along with
equation (9), from (7) we find that

$$y_{2t} = \frac{a_{22}}{a_{12}} y_{1t} + (a_{21} - \frac{a_{11}a_{22}}{a_{12}}) y_{1t-1} + \epsilon_{2t} - \frac{a_{22}}{a_{12}} \epsilon_{1t}$$

\[= \alpha y_{1t-1} + \beta y_{1t} + u_t \tag{10}\]

where

$$\alpha \equiv a_{21} - \frac{a_{11}a_{22}}{a_{12}}, \quad \beta \equiv \frac{a_{22}}{a_{12}}, \quad u_t \equiv \epsilon_{2t} - \beta \epsilon_{1t}$$

Thus $\text{Var}(u_t)$ is equal to $\sigma_2^2 + \beta^2 \sigma_1^2 - 2\beta \rho \sigma_1 \sigma_2$.

If we estimate simple model like (10), one effect of the measurement error in the independent
variable is to increase the variance of the error terms if $\epsilon_1$ and $\epsilon_2$ are negatively correlated. Another
severe consequence is that the OLS estimator is biased and inconsistent. Because $y_{1t} = y_{2t-1} + \epsilon_{1t}$, and $u_t$ depends on $\epsilon_{1t}$, $u_t$ must be correlated with $y_{1t}$ whenever $\beta \neq 0$. In fact, since the random
part of $y_{1t}$ is $\epsilon_{1t}$ and $\epsilon_{1t}$ is correlated to $\epsilon_{2t}$, we have that

$$E(u_t|y_{1t}) = E(u_t|\epsilon_{1t}) = \epsilon_{2t} - \beta \epsilon_{1t}.$$ 

Using the fact that $E(u_t) = 0$ unconditionally, we can see that

$$\text{Cov}(y_{1t}, u_t) = E(y_{1t}u_t) = E(y_{1t}E(u_t|y_{1t})) = E((y_{2t-1} + \epsilon_{1t})(\epsilon_{2t} - \beta \epsilon_{1t})) = \rho \sigma_1 \sigma_2 - \beta \sigma_1^2$$

This covariance is negative if $\beta > \rho \sigma_2 / \sigma_1$ and positive if $\beta < \rho \sigma_2 / \sigma_1$. Since it does not depend
on the sample size $n$, it does not go away as $n$ becomes large. Therefore the OLS assumption
that $E(u_t|X_t) = 0$ is false whenever any element of $X_t$, that is $y_{1t}$ and $y_{1t-1}$ in our special case, is
measured with error. In consequence, the OLS estimator is biased and inconsistent.

**Proposition 3** Given the true DGP of a reduced form VAR (1) in equations (5) and (6), the OLS
estimator in the linear regression model (10) is biased and inconsistent.

**Proof.** See Appendix.

This proposition shows when the true DGP is a reduced form bivariate VAR (1), linear regression
model is misspecified and OLS estimators lead serious measurement error problem.
4 Application to U.S. Nonfarm Payroll Employment

There are a large amount of applications of the updating bivariate VAR forecast. We focus on examining one of the macroeconomic indicators such as employment in the United States.

Bureau of Labor Statistics (BLS), U.S. Department of Labor, releases the employment situation each month. The announcements are usually made at 8:30 a.m. on a Friday. The employment situation is composed of household survey and establishment survey data. The household survey has a wider scope than the establishment survey since it includes the self-employed, unpaid family workers, agricultural workers, and private household workers, who are excluded by the establishment survey. However, the establishment survey employment series has a smaller margin of error on the measurement of month-to-month change than the household survey in that it has much larger sample size. The establishment survey includes several industry payroll employment information, such that the total nonfarm payroll employment, weekly and hourly earnings, weekly hours worked, and by selected industry. As a set of labor market indicators, nonfarm payroll employment counts the number of paid employees working part-time or full-time in the nation’s business and government establishments. They are used in the formulation of fiscal and economic policy. Federal officials constantly monitor this data watching for even smallest signs of potential inflationary pressure. The series are also used as an explanation of anomalous stock price behavior. By tracking the jobs data, investors can sense the degree of tightness in the job market, which affects interest rates and leads bond and stock prices fluctuate in the process.

Automatic Data Processing (ADP) contracted with Macroeconomic Advisors to compute a monthly report (ADP National Employment Report). Estimates of employment published in the ADP National Employment Report are available beginning in January of 2001. The announcements are usually made at 8:15 a.m. on a Wednesday, although very less announcements are made on other days. All announcement dates, whether Wednesday or not, are included in our study. The ADP report is a measure of nonfarm private employment, and it forms the level of employment by select industry and by size of payroll (small, medium, and large). It ultimately helps to predict monthly nonfarm payrolls from BLS employment situation. The ADP report only covers private payrolls, excluding government. However, it has been proved that nonfarm private employment released on the behalf of the ADP (hereafter called ADP’s data) are highly correlated to nonfarm payroll employment announced two days later by BLS (hereafter called BLS’s data). As a result, this study only focuses on ADP’s data and BLS’s data.

In this section, data description is illustrated firstly. Then we apply ADP’s data and BLS’s data to fit the bivariate updating VAR model and study VAR forecast performance. Comparisons to the various individual time series forecasts are also investigated. Afterall, impacts on stock market are studied.
Table 1: Descriptive Statistics

| Source | Automatic Data Processing, Inc.  
|        | Macroeconomic Advisers, LLC  
|        | Bureau of Labour Statistics |
| Frequency | Seasonal adjusted, Monthly |
| Sample period | January 2001 - April 2008 |
| Sample Size | 88 |

<table>
<thead>
<tr>
<th></th>
<th>First difference of ( adp )</th>
<th>First difference of ( bls )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>49.0814</td>
<td>63.2674</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>134.0006</td>
<td>155.1345</td>
</tr>
<tr>
<td>( \beta_1 )</td>
<td>-0.1154</td>
<td>-0.2629</td>
</tr>
</tbody>
</table>

Augmented Dickey-Fuller: -2.273 \(-3.556\)

\( H_0 : \) Nonstationary

Reject at 90\% (critical value = -1.61)

\( ^1 \) The coefficient \( \beta \) designates the autocorrelation of the series at lag \( i \). The augmented Dickey-Fuller test statistic is computed as \( \hat{\tau} = \hat{\beta}/\text{ase}(\hat{\beta}) \) in the model: \( \Delta y_t = \Sigma_{j=1}^{p} \beta_j y_{t-p} + \epsilon_t \), with \( p = 1 \) lags.

### 4.1 Data

Two time series, namely ADP’s data \( adp \) and BLS’s data \( bls \), are used to fit the bivariate VAR model in (1). \( bls \) is total nonfarm payroll employment from nonfarm payrolls establishment survey conducted by state employment security agencies in cooperation with the BLS. \( adp \) is nonfarm private employment from the ADP national employment report computed by the ADP on behalf of Macroeconomic Advisers. The 88 observations are seasonal adjusted monthly from January of 2001 to April of 2008. Table 1 reports the descriptive statistics. Dickey-Fuller nonstationary tests have been conducted, and the presence of a unit root is rejected. Both series are stationary use first differences. Since the test is known to have low power, even a slight rejection means that existence of a unit root is unlikely. The time-series plot of the data is provided in Figure 1. After the first difference of each time-series, the plot of the monthly changes of employees on nonfarm payrolls is provides in Figure 2.

The timing of two time series release is fixed and in order. The ADP national employment report is released, for public use only, two days prior to publication of the employment situation by the BLS. For year 2008 release schedule, August 2008 report is the only one which is release for public use only one day prior to publication of the employment situation by the BLS.

The BLS revises its initial monthly estimates twice, in the immediately succeeding 2 months, to incorporate additional information that was not available at the time of the initial publication of the estimates. On an annual basis, the BLS re-calculates estimates to nearly complete employment counts available from unemployment insurance tax records, usually in March. ADP revises its initial release one month later. In addition, the entire history of the report is revised once a year after the annual benchmarking of the national employment data by the BLS, usually in February.
Nonfarm payroll employment counts the number of paid employees working part-time or full-time in the nation’s business and government establishments. The ADP national employment report only covers private payrolls, excluding government.

4.2 Forecasting

A reduced form VAR expresses each variable as a linear function of its own past values, the past values of all other variables being considered. It also captures a serially correlated error term across equations. The error terms in these regressions are the co-movements in the variables after taking its past values into account. Thus, in this study the VAR involves two equations: current \( adp \) as a function of past values of the \( adp \) and the \( bls \); current \( bls \) as a function of past values of the \( adp \) and the \( bls \). Each equation can be estimated by ordinary least squares regression. This OLS estimator is as efficient as maximum likelihood estimator and general least squared estimator. The number of lagged values to include in each equation is determined by Schwarz’s Bayesian information criterion (SBIC), Akaike’s information criterion (AIC) and the Hannan and Quinn information criterion (HQIC). All latter two criterions indicate that the optimal lag selection of two combining two time series and the optimal lag selection of two for each individual time series. SBIC suggests the optimal lag selection of one for either combining two time series or individual time series. For comparison we employ one lag in all time series models such as the updating VAR, the ordinary VAR, and univariate autoregression.

The bivariate VAR (1) model estimation with standard deviation under the coefficients is as follows:

\[
\begin{align*}
\text{adp}_t &= 0.7882 \enspace \text{adp}_{t-1} + 0.0975 \enspace \text{bls}_{t-1} + \epsilon_{1t} \\
&= (0.0930) \\
\text{bls}_t &= 0.8388 \enspace \text{adp}_{t-1} + 0.1342 \enspace \text{bls}_{t-1} + \epsilon_{2t} \\
&= (0.1319) \\
\end{align*}
\]
In the ordinary VAR forecast, the coefficient estimates and the variance covariance matrix estimates are used to calculate the forecast mean squared error. In the updating VAR forecast, in addition to coefficient estimates and variance covariance matrix estimates, we also concern the correlation of residuals of VAR to compute the more efficient forecast mean squared error. The correlation of residuals of the bivariate VAR is 50.74%, which helps to use one series’ more information to forecast the other series’ unknown value.

This is real time forecast with revisions in data set. For each time period forecast, we collect all the information available at that time to do one and two-step ahead forecasts. Figure 3 shows revisions in data set and data set used for estimation in each time period. Knowing adp’s first release is always two days prior to bls’s first release, we take advantage of two days’ ahead in adp and forecast bls in the same time period. For instance, on May 30 of 2007, once ADP reports its first release of nonfarm private employment for May (2007m5), the real time information at that time is reported in column 2 and 3 in Figure 3, totally 76 observations. After the BLS reports its employment situation for May on Jun 1 of 2007, bls on 2007m4 becomes the second revisions and bls on 2007m3 becomes final revision. This can be seen from column 3 and 4 in Figure 3. Furthermore, on July 4 of 2007, once ADP reports its first release of nonfarm private employment for June (2007m6), the estimation period is from 2001m1 through 2007m5 of both adp and bls, totally 77 observations. This continues till March of 2008. On April 4 of 2008, the BLS reports the first release of employment situation for March (2007m3) while the BLS revises on annual basis. It is on the column of 25 in Figure 3.

Figure 3: Revisions in Data

<table>
<thead>
<tr>
<th>On</th>
<th>Apr</th>
<th>May</th>
<th>Jun</th>
<th>Jul</th>
<th>Aug</th>
<th>Sep</th>
<th>Oct</th>
<th>Nov</th>
<th>Dec</th>
<th>Jan</th>
<th>Feb</th>
<th>Mar</th>
</tr>
</thead>
<tbody>
<tr>
<td>2007m1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>107</td>
<td>107</td>
<td>107</td>
<td>107</td>
<td>107</td>
<td>107</td>
<td>107</td>
<td>107</td>
<td>107</td>
<td>107</td>
<td>107</td>
<td>107</td>
</tr>
<tr>
<td>2007m4</td>
<td>61</td>
<td>61</td>
<td>61</td>
<td>61</td>
<td>61</td>
<td>61</td>
<td>61</td>
<td>61</td>
<td>61</td>
<td>61</td>
<td>61</td>
<td>61</td>
</tr>
<tr>
<td>2007m5</td>
<td>97</td>
<td>98</td>
<td>98</td>
<td>98</td>
<td>98</td>
<td>98</td>
<td>98</td>
<td>98</td>
<td>98</td>
<td>98</td>
<td>98</td>
<td>98</td>
</tr>
<tr>
<td>2007m6</td>
<td>150</td>
<td>143</td>
<td>132</td>
<td>143</td>
<td>126</td>
<td>143</td>
<td>143</td>
<td>143</td>
<td>143</td>
<td>143</td>
<td>143</td>
<td>143</td>
</tr>
<tr>
<td>2007m7</td>
<td>48</td>
<td>41</td>
<td>41</td>
<td>41</td>
<td>41</td>
<td>41</td>
<td>41</td>
<td>41</td>
<td>41</td>
<td>41</td>
<td>41</td>
<td>41</td>
</tr>
<tr>
<td>2007m8</td>
<td>38</td>
<td>52</td>
<td>42</td>
<td>27</td>
<td>27</td>
<td>27</td>
<td>27</td>
<td>27</td>
<td>27</td>
<td>27</td>
<td>27</td>
<td>27</td>
</tr>
<tr>
<td>2007m9</td>
<td>58</td>
<td>61</td>
<td>61</td>
<td>61</td>
<td>61</td>
<td>61</td>
<td>61</td>
<td>61</td>
<td>61</td>
<td>61</td>
<td>61</td>
<td>61</td>
</tr>
<tr>
<td>2007m10</td>
<td>106</td>
<td>119</td>
<td>119</td>
<td>119</td>
<td>119</td>
<td>119</td>
<td>119</td>
<td>119</td>
<td>119</td>
<td>119</td>
<td>119</td>
<td>119</td>
</tr>
<tr>
<td>2007m11</td>
<td>189</td>
<td>173</td>
<td>94</td>
<td>173</td>
<td>115</td>
<td>185</td>
<td>-25</td>
<td>185</td>
<td>-325</td>
<td>185</td>
<td>-325</td>
<td>185</td>
</tr>
<tr>
<td>2007m12</td>
<td>40</td>
<td>37</td>
<td>38</td>
<td>42</td>
<td>42</td>
<td>42</td>
<td>42</td>
<td>42</td>
<td>42</td>
<td>42</td>
<td>42</td>
<td>42</td>
</tr>
<tr>
<td>2008m1</td>
<td>130</td>
<td>119</td>
<td>-27</td>
<td>119</td>
<td>-22</td>
<td>119</td>
<td>-26</td>
<td>119</td>
<td>-26</td>
<td>119</td>
<td>-26</td>
<td>119</td>
</tr>
<tr>
<td>2008m3</td>
<td>83</td>
<td>38</td>
<td>-80</td>
<td>38</td>
<td>-80</td>
<td>38</td>
<td>-80</td>
<td>38</td>
<td>-80</td>
<td>38</td>
<td>-80</td>
<td>38</td>
</tr>
<tr>
<td>2008m4</td>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The detailed estimation methodology is investigated in the following two-stage multistep VAR forecast. At the first stage, we estimate the bivariate VAR over the period from 2001m1 through 2007m4. Following the maximum likelihood estimation by VAR, one step ahead residuals prediction of both series is straightforward to be made. Then we regress the residuals of the bls on the residuals.
of the \(\text{adp}\). Since we observe the actual error term of the \(\text{adp}\) at time 2007m5, the best fitted residual of the \(\text{bls}\) at time 2007m5 is the estimated coefficients times the actual error term of the \(\text{adp}\) at time 2007m5.

At the second stage, we reestimate the bivariate VAR over the period from 2001m1 through 2007m5. One-step ahead fitted values of both series are predicted for time 2007m6. This is our best fitted values of the \(\text{bls}\) and \(\text{adp}\) at time 2007m6. If we want to do further forecasts, we reestimate the bivariate VAR over the period from 2001m1 through 2007m6. The fitted values of both series are the best forecasts at time 2007m7, and so on. This dynamic forecast is based on the real time information, that is the latest actual value of \(\text{adp}\). As well the \(\text{adp}\) time series is highly correlated with the \(\text{bls}\) time series, thus this two-stage multistep VAR forecast benefits from accuracy measures.

4.3 Forecast Accuracy Comparison

Multistep ahead forecasts, computed by iterating forward the reduced form VAR, are assessed in Table 2. The BLS releases revision of past employment announcements for the previous three months, after which the announcement is considered final. To make our multistep ahead forecasts consistent and comparable, we choose forecast horizon from 2007m5 through 2008m4. Because the ultimate test of a forecasting model is its out-of-sample performance, Table 2 focuses on out-of-sample forecasts over the period from 2001m1 through 2007m4. It examines forecast horizons of one month and two months. The dynamic forecast ‘k’ steps ahead is computed by estimating the VAR through a given month, making the forecast ‘k’ steps ahead, reestimating the VAR through the next month, making the next forecast and so on through the forecast period. The key difference between the updating VAR forecasts and the ordinary VAR forecasts is the first stage estimation. Taking advantage of one more information at the first forecasting period, the updating VAR forecast provides the best estimate we are going to predict.

<table>
<thead>
<tr>
<th>Forecast Horizon</th>
<th>Updating VAR</th>
<th>VAR</th>
<th>AR</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 month</td>
<td>71.51</td>
<td>86.14</td>
<td>71.75</td>
</tr>
<tr>
<td>2 months</td>
<td>95.64</td>
<td>108.44</td>
<td>96.44</td>
</tr>
</tbody>
</table>

† The entries are root mean squared error.

As a comparison, out-of-sample forecasts were also computed for a univariate autoregression with one lag, that is, a regression of the variable on lags of its own past values. Table 2 shows the root mean square forecast error for each of the forecasting methods. The mean squared forecast error is computed as the average squared value of the forecast error over the 2007m5 - 2008m4 out-of-sample period, and the resulting square root is the root mean squared forecast error reported
in the table. In Table 2 the entries of column 2 through 5 are the root mean squared errors of the updating VAR forecast, the ordinary VAR forecast, and univariate autoregression forecast respectively for nonfarm payroll employment \textit{blls}. The results indicate that the updating VAR forecast has lower root mean squared error than the ordinary VAR and univariate autoregressive forecast over one and two-step ahead forecast. Comparing to the ordinary VAR quantitatively, the updating VAR forecast improves 17\% in one-step ahead forecast and improves 12\% in two-steps ahead forecast.

\textbf{One-Step Ahead Time Series Forecast Performance}

One-step ahead time series forecast is based on time period of May 2007 through April 2008. Totally 12 observations are constructed. For instance, on May 30 of 2007, after ADP reported the first released nonfarm private employment, we use \textit{adp} from 2001m1 through 2007m5 and \textit{bls} from 2001m1 through 2007m4 to forecast \textit{bls} on 2007m5, in which BLS will release it two days later on June 1 of 2007.

![Figure 4: One-Step Ahead Forecast: 2007m5 - 2008m4](image)

Omitting \textit{adp} data on 2007m5, the bivariate ordinary VAR forecast is to estimate both \textit{adp} and \textit{bls} from 2001m1 through 2007m4. Then one-step ahead forecast is followed. The bivariate updating VAR forecast takes two stages to complete forecasting. In the first stage, it estimates both \textit{adp} and \textit{bls} through time period of 2001m1 to 2007m4. Since we know the true value of \textit{adp} on 2007m5, we have the true residual term of \textit{adp} on 2007m5. In the second stage, we regress residuals of \textit{bls} on residuals of \textit{adp} and obtain the coefficients of \textit{adp}. Given the true residuals of \textit{adp} on 2007m5 and the coefficients of \textit{adp}, the residual of \textit{bls} on 2007m5 is obviously the outcome of the true residual of \textit{adp} on 2007m5 multiplied by the coefficients of \textit{adp}. Then one-step ahead forecast of \textit{bls} is its fitted value adding estimated residual on 2007m5.

Consequently, standing at 2007m5, one-step ahead forecast of \textit{bls} is based on both \textit{adp} and \textit{bls} through time period of 2001m1 to 2007m5 and \textit{adp} on time period of 2007m6. Keeping the process
to continue, we obtain 12 observations of one-step ahead forecasts in Figure 4. The solid line is
the BLS first released data from 2007m5 through 2008m4. The first released data plays a key role
in markets, so that we employ it as the actual value. We find that the updating VAR forecast
outperforms the ordinary VAR and univariate autoregression forecasts.

Two-Steps Ahead Time Series Forecast Performance

Two-steps ahead time series forecast is based on time period of June 2007 through April 2008.
Totally 11 observations are constructed. The difference from one-step ahead forecast is that based
on one-step forecast of bls on 2007m5, an iterated one-step forecast of bls on 2007m6 is conducted.
Since two-steps ahead forecast is based on one-step forecast, the updating multistep VAR fore-
cast outperforms the ordinary VAR forecast. Figure 5 reports that the updating VAR forecast
outperforms the ordinary VAR and univariate autoregressive model forecasts.

Figure 5: Two-Steps Ahead Forecast: 2007m6 - 2008m4

4.4 Stock Market Reaction to Employment Announcement

Total nonfarm payroll employment by the BLS is the first major economic indicator released
each month. This is an economic report that can move the markets. Figure 6 gives the full picture
of quantitative comovement between the BLS employment situation first release and Don Jones
industrial IN. closing price. For instance, on Sept 7 of 2007 the BLS reported four thousands
decreases of employment. This led the stock market index declined dramatically by 250 points
at its closing price. On Nov 2 of 2007, the BLS reported 166 thousands increases in nonfarm
employment. This led the stock index closing price increased by 28 points at that day.
Figure 6: Impact of Report Release on Stock Market

† Don Jones Industrial IN. from Yahoo Finance at http://finance.yahoo.com/

5 Conclusion Remarks

In multivariate time series, the covariance matrix of observation innovations plays an important role in forecasting. We propose a practical method to update forecasts in multivariate VAR models. The focus and the benefit of employing the updating approach is that the true innovation of the current known observations is always to be useful to predict an innovation of the unknown observations. The theoretical framework shows that the current known observations of one variable are always going to be useful for forecasting the current unknown observations of another variables. Therefore, higher correlation among observation innovations of multi-variables implies that mean squared forecast error of the current unknown observations of the other variables will be accurate for longer.

There are large amounts of applications in real-time forecasting. This paper uses U.S. labor data to examine whether ADP estimates which is usually announced two days prior to BLS estimates are helpful to forecast the same month total nonfarm payroll employment by the BLS. Rather than use the final released data to estimate and forecast, we use data of as many different as there are dates in the sample. More specifically, at every date within a sample, variables in the model is the most up-to-date estimates at that time. We find that the predicted employment are more accurately matching the labor data by considering the real-time information than the standard time series forecast. Comparing to the ordinary VAR quantitatively, the updating VAR improves 17% in one-step ahead forecast and improves 12% in two-steps ahead forecast.

The timing of data release for time series variables for the same time period of observation is often spread over weeks. For instance earning announcements for firms can be spread over a
two week period even though these earnings are for the same quarter or month. Future research will extend the theoretical framework to more general case of two more variables available earlier by public use. Applications to earnings forecasts for firms by different industry sectors are to be studied.
Appendix

Proof of proposition 1. For the forecast horizon at time $T + 1$, we observe $\epsilon_{1T+1}$, since $\epsilon_{1T+1} = y_{1T+1} - (a_{11}y_{1T} + a_{12}y_{2T})$. So we can forecast the innovation $\epsilon_{2T}(1)$ by the relationship $\epsilon_{2T}(1) = \rho_{12}\frac{\sigma_2}{\sigma_1}\epsilon_{1T+1}$. Then forecast error becomes:

$$y_{2T+1} - \hat{y}_{2T+1} = \epsilon_{2T}(1)$$

The variance of the forecast error of $y_2$ at $T + 1$ is

$$MSE[y_{2T}(1)] = Var[y_{2T+1} - \hat{y}_{2T+1}] = Var[\epsilon_{2T} - \epsilon_{2T}(1)] = (1 - \rho_{12}^2)\sigma_2^2$$

Proof of proposition 2. Given information set $I = \{y_{11}, y_{12}, \ldots, y_{1T+1}, y_{21}, \ldots, y_{2T}\}$, where the first subscribe represents the variable and the second subscribe represents the time period. We stand at $T + 1$. We know the time series of $y_1$ from 1 till $T + 1$ while we only know the time series of $y_2$ from 1 till $T$.

At $T + 1$, we adopt all the time periods from 1 to $T + 1$ of $y_1$ and the time periods from 1 to $T$ of $y_2$. Since $\epsilon_1$ and $\epsilon_2$ are correlated, we see $\epsilon_{2T+1} = (\rho_{12}\sigma_2/\sigma_1)\epsilon_{1T+1}$. Thus the forecast error is

$$\begin{pmatrix} y_{1T+1} - \hat{y}_{1T+1} \\ y_{2T+1} - \hat{y}_{2T+1} \end{pmatrix} = \begin{pmatrix} 0 \\ \epsilon_{2T+1} \end{pmatrix} = (\rho_{12}\sigma_2/\sigma_1)\epsilon_{1T+1}$$

and the MSE or forecast error covariance matrix is

$$MSE[Y_T(1)] = Var[Y_{T+1} - \hat{Y}_{T+1}] = \begin{pmatrix} 0 & 0 \\ 0 & \rho_{12}^2\sigma_2^2 \end{pmatrix}$$

At $T + 2$,

$$MSE[Y_T(2)] = MSE[Y_T(1)] + A^1\Omega_e(A^1)'$$

At $T + 3$,

$$MSE[Y_T(3)] = MSE[Y_T(2)] + A^2\Omega_e(A^2)'$$

Recursively, our updating forecast error covariance matrix becomes

$$MSE^u[Y_T(k)] = \begin{pmatrix} 0 & 0 \\ 0 & \rho_{12}^2\sigma_2^2 \end{pmatrix} + \sum_{i=1}^{k-1} A^i\Omega_e(A^i)'$$

Proof of proposition 3. Suppose we estimate the ordinary least squares model of the form

$$y_{2t} = \beta y_{1t} + u_t$$

17
and the true DGP is as (5). OLS estimator \( \hat{\beta}_{ols} \) is

\[
\hat{\beta}_{ols} = \frac{\sum_t y_{1t} y_{2t}}{\sum y_{2t}^2} = \frac{\sum_t y_{1t} (\alpha y_{1t-1} + \beta y_{1t} + u_t)}{\sum y_{1t}^2} = \beta + \alpha \frac{\sum_t y_{1t} y_{1t-1}}{\sum y_{1t}^2} + \frac{\sum_t y_{1t} u_t}{\sum y_{1t}^2}.
\]

Since \( y_{1t} \) and \( u_t \) are correlated, the second term goes to \( \alpha \) in the limit while the third term does not go to zero in the limit and the estimator is biased, that is,

\[ E(\hat{\beta}_{ols}) = \beta + \alpha + E\left(\frac{\sum_t y_{1t} u_t}{\sum y_{1t}^2}\right). \]

Since

\[
\alpha = \frac{Var(y_{2t-1})Cov(y_{1t-1}, y_{2t}) - Cov(y_{1t-1}, y_{2t-1})Cov(y_{2t-1}, y_{2t})}{Var(y_{1t-1})Var(y_{2t-1}) - Cov(y_{1t-1}, y_{2t-1})^2} = 0
\]

by linear regression of (7), we have

\[ E(\hat{\beta}_{ols}) = \beta + E\left(\frac{\sum_t y_{1t} u_t}{\sum y_{1t}^2}\right). \]

To see the inconsistency,

\[
\text{plim}_{n \to \infty} (\hat{\beta}) = \text{plim}_{n \to \infty} \left( \frac{\sum_t y_{1t} (\beta y_{1t} + \alpha y_{1t-1} + u_t)}{\sum y_{1t}^2} \right)
\]

\[ = \beta + \text{plim}_{n \to \infty} \frac{\sum_t y_{1t} u_t}{\sum y_{1t}^2}
\]

\[ = \beta + \frac{Cov(y_{1t}, u_t)}{Var(y_{2t-1})Var(y_{1t-1}) - Cov(y_{1t-1}, y_{2t-1})^2}
\]

\[ = \beta + \frac{\rho \sigma_1 \sigma_2 - \beta \sigma_1^2}{\sigma_2^2 + \sigma_1^2}
\]

\[ = \frac{\beta}{1 + \sigma_1^2 / \sigma_2^2} + \frac{\rho \sigma_1 \sigma_2}{\sigma_1^2 + \sigma_2^2}
\]

\[ = \beta + \frac{\rho \sigma_1 \sigma_2}{1 + \sigma_1^2 / \sigma_2^2}.
\]

Thus \( \hat{\beta} \) will underestimate \( \beta \) if \( \beta > \rho \sigma_2 / \sigma_1 \) and will overestimate \( \beta \) if \( \beta < \rho \sigma_2 / \sigma_1 \). The degree of underestimation or overestimation depends on \( \sigma_1^2 / \sigma_2^2 \).
References


