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

Allan W. Gregory, Julia Hui Zhu *

Department of Economics, Queen's University†

October 2008

Abstract

The timing of data release for a specific 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 timing or data release and the use of econometric real-time methods (what we call an updated vector autoregression forecast) to forecast data that has not yet been made available. In comparison to standard time series forecasting, we find that the updated multivariate time series forecasting will be more accurate with higher correlation coefficients among observation innovations. This updated forecast has a direct application in macro and financial series. One of the macro real variables, U.S. nonfarm payroll employment, is the first of its kind in the literature. We find that the relative efficiency gain by using the updated vector autoregression forecast is 16% in the one-step-ahead forecast and 7% in the two-step-ahead forecast, respectively, in comparison to the ordinary vector autoregression forecast. The results demonstrate the usefulness of updating multivariate forecast accurate measurements.

Keywords: Timing of data release; Data revision; Nonfarm payroll employment; Forecasts

*The author is grateful to seminar participants at the CEA 2008 meetings and Queen's University for their helpful comments and discussions.

†Address correspondence to Julia Hui 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 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 month or quarter. Further, official government statistics are often released at different times over the month or quarter and yet cover the same time period. This paper focuses on this separation of announcement timing or data release and the use of standard econometric updating methods to forecast data that has not yet to be made available. To the best of our knowledge, this is an important aspect of forecasting and has a direct application in financial market updating (in terms of earnings, earning per share and so on) and has not yet been studied in the literature.

Traditional forecasting in a multivariate time series setting is usually studied in the context of vector autoregression (VAR) models. In this set-up the VAR, with the common end-point, is specified and estimated. Single or multiple period forecasts are conducted. Standard errors for the forecasts can be based on asymptotic normal theory or, more recently, the use of bootstrap or some other re-sampling technique has been applied. The key element for this research is that there is a common end point of observed data and that forecasts are made for all variables over the same forecasting horizon. The situation that we consider is different in that only some (one) of the variables comprising the system is released at a given point in time with the remaining variables being released at later dates. The later dates may coincide for some variables or they may differ, in which case there is a sequence of release times for the system as a whole.

There are two critical exogeneity assumptions in this research. The first assumption is that the timing of the release information (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 earlier (or later). From what we are able to tell about earnings announcements, the decision as to when to announce is made long in advance of the time when the earnings information would be credibly known to the firm, so this is unlikely to be a problem. The second exogeneity assumption is that the announcement of one firm's earnings on a given day will not affect the announcement of the earnings of a related firm on a later day. That is, if one firm announces large earnings in say the forth quarter of 2007 on February 27, 2008, a related firm does not change its announced earnings for the same quarter when released on, say, February 28, 2008. For the present application we avoid both of these issues and consider a simple bivariate example using nonfarm labor data in the United States.

The ordinary multivariate VAR forecast is based on the common end-point. However, most available macroeconomic variables or financial time series for a given time period are released on different announcement dates. For example, financial time series such as firms' earnings released 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 etc. released for the same month are also available at different dates for public use. In this paper, we obtain the latest available data and adopt two-stage estimation methods to forecast unknown data for the same month or quarter, what we call the updated VAR forecast hereafter.

Comparison with the ordinary VAR forecasts shows that the updated forecast mean squared error (MSE) is smaller than the one in the ordinary VAR forecast. We first consider a case with one available real-time variable in the bivariate updated forecasts. When the correlation coefficient of the innovations between two variables are lower, the updated forecast has a relatively smaller MSE than the ordinary VAR forecast. When the correlation coefficient of observation innovations between two variables is higher, the updated forecast has a significantly smaller MSE than the ordinary VAR forecast. In one extreme case, when the correlation coefficient approaches zero, that is, no contemporaneous information is available to be useful for forecasting, the updated forecast has exactly the same result as the ordinary VAR forecast. In another extreme case, when the correlation coefficient approaches one, that is, we have perfect linear association and there are no errors, the updated forecast has the best performance. Furthermore, we note that not only does the correlation coefficient of observation innovation play an important role in the multistep ahead forecast but so does the coefficient parameter.

Although there are various macroeconomic announcement series or earnings announcements for firms we could consider as an application of the research, we specifically chose one of macro real variables, nonfarm payroll employment in the United States, as a strong candidate for a forecasting variable. One reason is that total nonfarm payroll employment is the first major economic indicator released each month. This is an economic report that can move the markets. Another reason is that the employment situation has frequently been used not only in the formulation of Federal Reserve policy but also as an explanation of anomalous stock price behavior. In addition, employment announcements have tractable and fixed announcement timing and order.

We have eight years of monthly nonfarm labor data dating from January 2001 through September 2008. Therefore, this research is the first of its kind in the literature and provides the most up-to-date forecasting. One important issue dealing with macro variables is data revision. To solve data revision problems existing in most macro variables, we adopt the sample data whenever it is available at the time period of forecasting. That is, our sample indicates the first released data, the second released data, the third released data, and the final released data in turn. Rather than use final released data to estimate and forecast, we employ data of as many different release as there are dates in the sample. More specifically, at every date within a sample, both right-side and left-side variables in the bivariate VAR model should be the most up-to-date estimate variable at that time. We find that the root mean squared error in the updated VAR is less than the ordinary VAR. It shows quantitatively that the updated VAR improves 16% in a one-step ahead forecast and 7% in a two-step ahead forecast respectively compared to the ordinary VAR. We also show that in comparison to the univariate autoregression model, the updated VAR forecast is slightly better depending on the estimation sample we choose.

Related Literature

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

Many economic time series are subject to revision. Revisions to measures of real economic activity may occur immediately in the next 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 their revisions do not have a zero mean, which indicates that the initial announcements by Statistical Agencies are biased. Croushore and Stark (2001, 2003) show how data revisions can affect forecasting. 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 interested 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 both sides of the equation provides superior forecast to that obtained from final released data. Our paper employs the most up-to-date estimates available at forecasting time in addition to using the updated 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 researchers study how markets respond to labor data, such as Krueger (1996) and Boyd et al. (2005). The latter examine how stocks respond to unemployment news, which is measured as the surprise component. Adopting forecasts of the change in the unemployment rate to obtain 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. Other researchers focus 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) study how U.S. macroeconomic announcements affect joint movements of exchange rates and interest rates. Using a 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 periods. With revision in data, most estimations adopt final released data (end-of-sample). Our research contributes several new dimensions, which, to our knowledge, have not yet been addressed. First, we provide a practical updating multivariate VAR forecasting method, extending the analysis into U.S. labor data. Second, we construct the most up-to-date data set for both estimation and forecasting. In this paper, first released data plays an important role in estimation and forecasting. Finally, we investigate how the stock market reacts to the 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 updated VAR forecasts. In section 3 we discuss the model specification. An application of the updated VAR forecasts to U.S. nonfarm payroll employment data is illustrated in section 4. Section 5 concludes. Appendix provides proofs.

2 A Theoretical Framework

Consider the N multivariate stationary VAR (1) model

$$Y_t = AY_{t-1} + \epsilon_t, \qquad t = 0, \pm 1, \pm 2, ..., \tag{1}$$

where $Y_t = (y_{1t}, ..., y_{Nt})'$ is a $(N \times 1)$ random vector, and the A is fixed $(N \times N)$ coefficient matrices. The first subscribe represents the variable and the second subscribe represents the time period. Moreover, $\epsilon_t = (\epsilon_{1t}, ..., \epsilon_{Nt})$' 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 $Cov(\epsilon_{it},\ \epsilon_{jt}) = \rho_{ij}\sigma_i\sigma_j$ for $i = 1, ..., N$ and $j = 1, ..., N$. The covariance matrix Ω 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, σ_i is the standard deviation of the innovation ϵ_i. Given the multivariate VAR model (1), the ordinary multivariate VAR one-step ahead forecast at time region T is $Y_T(1) = A\ Y_T$ and the associated forecast error is $\epsilon_T(1) = \epsilon_{T+1}$, where $Y_T(1)$ denotes the forecast of Y at time $T + 1$ and $\epsilon_T(1)$ denotes forecast error at time $T + 1$. The ordinary multivariate VAR forecasts are standard and can be obtained from Lütkepohl (1993) and Hamilton (1994).

Theoretically, the multivariate VAR forecast is based on the certain amount of time periods of 1 through T. Each equation in the multivariate VAR model can be estimated by ordinary least squares (OLS) regression. This OLS estimator is as efficient as the maximum likelihood estimator and thr general least squared estimator. Therefore, the ordinary multivariate 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 do not end at the same time, that is, one variable is generally available for public use a couple of days prior to the other variables. For example, financial time series such as firms' earnings released 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 so on. released for the same month are also available at different dates for public use. Omitting the timing factor, the ordinary multivariate VAR forecast usually ignores the latest information we can obtain and adopts the same amount from certain time periods to do forecasts.

The focus of this paper is on examining how taking advantage of more data from one variable and matching the timing factor can be used to improve multivariate VAR forecasts.

2.1 Updating Bivariate VAR Forecast

A practical method to update forecasts with one real-time variable available in advance is investigated. To simplify the discussion, we consider bivariate VAR forecast, where $N = 2$, firstly. Suppose y_{1t} is observable for $t = 1, \ldots T + 1$ and y_{2t} is observable only up to time T.

Then the reduced form bivariate VAR (1) model is as follows

$$\begin{aligned} y_{1t} &= a_{11}y_{1t-1} + a_{12}y_{2t-1} + \epsilon_{1t} \qquad t = 1, ..., T, T+1 \\ y_{2t} &= a_{21}y_{1t-1} + a_{22}y_{2t-1} + \epsilon_{2t} \qquad t = 1, ..., T. \end{aligned} \tag{2}$$

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

$$MSE[Y_T(1)] = \Omega_\epsilon,$$

where Y is a vector of $(y_1, y_2)'$, and the covariance matrix of Ω_ϵ is $E(\epsilon_t, \epsilon_t')$; that is,

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

Therefore, the MSE of one-step ahead forecast of variable y_2 in the ordinary bivariate VAR is $MSE[y_{2T}(1)] = \sigma_2^2$. Hereafter we denote MSE^u as the updated multivariate VAR forecast MSE to distinguish it from the ordinary multivariate VAR forecast MSE.

With one more piece of real-time information (y_{1T+1}) available in advance, at the forecast horizon of the time period $T+1$, we observe ϵ_{1T+1}. This is due to $\epsilon_{1T+1} = y_{1T+1} - (a_{11}y_{1T} + a_{12}y_{2T})$. If we regress ϵ_2 on ϵ_1, then the best estimates of ϵ_{2T+1} is obtained by $\hat{\epsilon}_{2T+1} = E(\epsilon_{2T+1}|\epsilon_{1T+1})$. Thus, we can forecast the residual $\hat{\epsilon}_{2T+1}$ by the relationship $\hat{\epsilon}_{2T+1} = (\rho_{12}\sigma_2/\sigma_1)\epsilon_{1T+1}$. Hence, the one-step ahead forecast MSE is obtained as the following proposition:

Proposition 1 *Given the known full information set $\mathcal{I} = \{y_{11}, y_{12}, ..., y_{1T+1}, y_{21}, ..., y_{2T}\}$, the mean squared error of one-step ahead forecast of variable y_2 in the updated bivariate VAR is*

$$MSE^u[y_{2T}(1)] = (1 - \rho_{12}^2)\sigma_2^2 \tag{3}$$

Proof. *See Appendix.*

This proposition shows that taking advantage of one more piece of real-time information available in advance, the one-step ahead updated bivariate VAR forecast MSE^u is $(1-\rho_{12}^2)\sigma_2^2$. Since ρ_{12} is coefficient correlation between ϵ_1 and ϵ_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 updated bivariate VAR forecast has smaller MSE in comparison with the forecast from the ordinary bivariate VAR. This implies the updated bivariate VAR forecast is more accurate than the forecast from the ordinary bivariate VAR.

Additionally, the higher the correlation among the error terms of the variables, the smaller the MSE in the updated bivariate VAR forecast. When the correlation coefficient of the innovations between two variables is lower, the updated bivariate VAR forecast has a relatively smaller MSE than the ordinary bivariate VAR forecast. When the correlation coefficient of observation innovations between two variables is higher, the updated bivariate VAR forecast has a significantly smaller MSE than the ordinary bivariate VAR forecast. In one extreme case, when the correlation coefficient approaches zero, that is, no contemporaneous information is available to be useful for forecasting, the updated bivariate VAR forecast has exactly the same results as the ordinary bivariate VAR forecast. In another extreme case, when the correlation coefficient approaches one, that is, we have perfect linear association and there are no errors, the updated bivariate VAR forecast has the best performance.

To generalize, we also examine the $k \geq 2$ long-horizon forecast with one more piece of real-time information known in advance.

Proposition 2 *Given the known full information set $\mathcal{I} = \{y_{11}, y_{12}, ..., y_{1T+1}, y_{21}, ..., y_{2T}\}$, the k-step ahead forecast mean squared error matrix in the updated bivariate VAR is*

$$\begin{aligned} MSE^u[Y_T(k)] &= \sum_{i=0}^{k-2} A^i \Omega_\epsilon A^{i'} + A^{k-1} \begin{pmatrix} 0 & 0 \\ 0 & (1-\rho_{12}^2)\sigma_2^2 \end{pmatrix} A^{k-1'} \\ &= \Omega_\epsilon + A \quad MSE^u[Y_T(k-1)] \quad A', \qquad k \geq 2, \end{aligned} \tag{4}$$

where A is 2×2 dimensional coefficient matrix. A matrix to the power of zero is defined to be the identity matrix of the same dimensions, that is, $A^0 = I$, and

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

Proof. *See Appendix.*

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

$$
\begin{aligned}
MSE[Y_T(k)] &= \sum_{i=0}^{k-2} A^i \Omega_\epsilon (A^i)' + A^{k-1} \begin{pmatrix} \sigma_1^2 & \rho_{12}\sigma_1\sigma_2 \\ \rho_{12}\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix} A^{k-1'} \\
&= MSE[Y_T(k-1)] + A^{k-1}\Omega_\epsilon A^{k-1'} \qquad k \geq 2.
\end{aligned} \tag{5}
$$

To see the difference between the updated bivariate VAR forecast MSE^u and the ordinary bivariate VAR forecast MSE, we compare the equation (4) with (5):

$$
MSE[Y_T(k)] - MSE^u[Y_T(k)] = A^{k-1} \begin{pmatrix} \sigma_1^2 & \rho_{12}\sigma_1\sigma_2 \\ \rho_{12}\sigma_1\sigma_2 & \rho_{12}^2\sigma_2^2 \end{pmatrix} A^{k-1'}. \tag{6}
$$

In an one-step ahead forecast, where $k = 1$, equation (6) indicates that the gain of employing one available real-time information y_{1T+1} to forecast related variable y_{2T+1} depends on the correlation coefficient of the innovations, ρ_{12}. When $\rho_{12} = 0$, there is no gain at all. The updated bivariate VAR forecast has the same MSE as the ordinary bivariate VAR forecast. This implies that the updated bivariate VAR forecast is more accurate than the forecast from the ordinary bivariate VAR under short-horizon forecasting. In addition, the higher the correlation among the error terms of the variables, the smaller the MSE in the updated bivariate VAR forecast. When the correlation coefficient approaches zero, that is, no contemporaneous information is available to be useful for forecasting, the updated forecast has exactly the same results as the ordinary bivariate VAR forecast. When the correlation coefficient approaches one, that is, we have perfect linear association and there are no errors, the updated bivariate VAR forecast has the best performance. As forecasting horizon becomes longer, the MSE of the updated bivariate VAR forecast converges to the MSE of the ordinary bivariate VAR forecast.

In a two-step ahead forecast, where $k = 2$, equation (6) becomes

$$
MSE[Y_T(2)] - MSE^u[Y_T(2)] = A \begin{pmatrix} \sigma_1^2 & \rho_{12}\sigma_1\sigma_2 \\ \rho_{12}\sigma_1\sigma_2 & \rho_{12}^2\sigma_2^2 \end{pmatrix} A'
$$

$$
= \begin{pmatrix} (a_{11}\sigma_1 + a_{12}\rho_{12}\sigma_2)^2 & (a_{11}\sigma_1 + a_{12}\rho_{12}\sigma_2)(a_{21}\sigma_1 + a_{22}\rho_{12}\sigma_2) \\ (a_{11}\sigma_1 + a_{12}\rho_{12}\sigma_2)(a_{21}\sigma_1 + a_{22}\rho_{12}\sigma_2) & (a_{21}\sigma_1 + a_{22}\rho_{12}\sigma_2)^2 \end{pmatrix}.
$$

Since the difference between the forecast MSE of y_{2T+1} of the ordinary bivariate VAR and that of the updated bivariate VAR is $(a_{21}\sigma_1 + a_{22}\rho_{12}\sigma_2)^2$, this is larger or equal to zero. Therefore, the two-step ahead updated bivariate VAR forecast has smaller MSE compared with the forecast from the ordinary bivariate VAR. Even with $\rho_{12} = 0$, the gain of employing one available real-time information y_{1T+1} to forecast the related variable y_{2T+1} still exists since $a_{21}^2\sigma_1^2$ is always positive.

Furthermore, $MSE[Y_T(k)] - MSE^u[Y_T(k)]$ converges to zero rapidly as $k \to \infty$. In other words, the MSE by the updated VAR forecast converges to the MSE by the ordinary VAR forecast as the forecasting horizon is bigger enough. Under the assumption that our VAR (1) process is stationary, the polynomial $det(I_m - Az)$[1] has no roots in and on the complex unit circle. That is equivalent to say that all eigenvalues of parameter matrix A have modulus less than 1. By the properties of matrix power[2], A^{k-1} converges to 0 as $k \to \infty$.

2.2 Updating Bivariate VAR Forecast with Data from Two More Periods Known in Advance

In practice, there are some applications to the bivariate VAR with two more periods available real-time information known in advance. Many countries use their teletriage system as early warning system. For instance, in Canada Telehealth is a toll-free helpline provided by the Ontario Ministry of Health and Long-term Care's (MOHLTC)'s Telehealth program and is available to all residents of Ontario. Users are encouraged to call with any general health questions with confidential advice being given regarding any health concerns. National Ambulatory Care Reporting System (NACRS) was developed in 1997 by the Canadian Institute for Health Information (CIHI) to capture clinical, administrative and demographic information from all hospital-based and community-based ambulatory care. The concern with the two correlated sources of data is that NACRS data are one of timeliness, as data are not available in real-time, but are rather months delayed. This issue is also compounded by the fact that some hospitals have yet to complete a migration to electronic records management, making the integration of all NACRS's additionally difficult. These limitations make understanding the provincial Telehealth data and its usefulness to public health and emergency medicine essential.

In this section, we develop the updated bivariate VAR forecast with two more time periods real-time information ($s \geq 1$). Let y_{1t} be observable for $t = 1, \ldots T + s$ with $s \geq 1$

[1]Lütkepohl (1991) Appendix A.6 rule 7 on page 456: all eigenvalues of the $(m \times m)$ matrix A have modulus less than 1 if and only if $det(I_m - Az) \neq 0$ for $|z| \leq 1$, that is, the polynomial $det(I_m - Az)$ has no roots in and on the complex unit circle.

[2]Lütkepohl (1996) property 14 on page 39: A $(m \times m)$ matrix A, $A^i \to_{i\to\infty} 0 \Leftrightarrow$ all eigenvalues of A have modulus less than 1.

and y_{2t} be observable only up to time T. The simple bivariate VAR(1) model is as follows:

$$\begin{aligned} y_{1t} &= a_{11}y_{1t-1} + a_{12}y_{2t-1} + \epsilon_{1t} \qquad t = 1, ..., T, T+1, ..., T+s \\ y_{2t} &= a_{21}y_{1t-1} + a_{22}y_{2t-1} + \epsilon_{2t} \qquad t = 1, ..., T. \end{aligned} \tag{7}$$

We assume all the assumptions hold as in the section 2.1.

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

$$y_{2T+1} - \hat{y}_{2T+1} = \epsilon_{2T+1} - \rho_{12}\frac{\sigma_2}{\sigma_1}\epsilon_{1T+1}.$$

Then the variance of the forecast error is

$$Var[y_{2T+1} - \hat{y}_{2T+1}] = Var[\epsilon_{2T+1} - \hat{\epsilon}_{2T+1}] = (1 - \rho_{12}^2)\sigma_2^2. \tag{8}$$

Equation (8) shows that the results are all consistent with one-step ahead bivariate VAR forecast. The MSE^u in this updated bivariate VAR forecast given one more periods available real-time information is smaller than the ordinary bivariate VAR forecast, that is, $MSE^u = (1 - \rho_{12}^2)\sigma_2^2 < MSE = \sigma_2^2$.

For two-step ahead forecast, the known information set is $\mathcal{I} = \{y_{11}, y_{12}, ..., y_{1T+s}, y_{21}, ..., y_{2T}\}$. At $T+2$, rearrange equation (7), we obtain

$$\epsilon_{1T+2} = y_{1T+2} - (a_{11}y_{1T+1} + a_{12}y_{2T+1}).$$

Since we do not observe y_{2T+1}, we do not observe ϵ_{1T+2}. There are two ways to make a prediction on ϵ_{1T+2}. One way is to set $E(\epsilon_{1T+2}) = 0$ and to set the variance of ϵ_{1T+2} be the first element of the MSE of the ordinary bivariate VAR forecast, that is, the first element of the matrix $\Omega_\epsilon + A\Omega_\epsilon A'$. The alternative way is to predict $\hat{\epsilon}_{1T+2}$ through the residual form $\hat{\epsilon}_{1T+2} = y_{1T+2} - (a_{11}y_{1T+1} + a_{12}\hat{y}_{2T+1})$. In the latter case, the variance of the difference in error becomes

$$Var[\epsilon_{1T+2} - \hat{\epsilon}_{1T+2}] = a_{12}^2(1 - \rho_{12}^2)\sigma_2^2.$$

If the sufficient condition $\{a_{12}^2(1 - \rho_{12}^2)\sigma_2^2 <$ the first element of matrix $\Omega_\epsilon + A\Omega_\epsilon A'\}$ holds, we would use $\hat{\epsilon}_{1T+2}$ rather than $E(\epsilon_{1T+2}) = 0$. Then the variance of the forecast error is followed by

$$Var[y_{2T+2} - \hat{y}_{2T+2}] = (1 + a_{22}^2)(1 - \rho_{12}^2)\sigma_2^2.$$

We see that not only the correlation coefficient of two time series and the variance of the forecasting innovation, but also the coefficient parameter a_{22} plays a role on the multistep ahead forecast.

To generalize the standard $1 \leq k \leq s$ multi-step horizon forecast, we iterate k-step ahead updated bivariate VAR forecast.

Proposition 3 *Given the known full information set $\mathcal{I} = \{y_{11}, y_{12}, ..., y_{1T+1}, ..., y_{1T+s}, y_{21}, ..., y_{2T}\}$, and the sufficient condition $(\sum_{i=0}^{k-2} a_{22}^{2i})a_{12}^2\sigma_2^2(1-\rho_{12}^2) <$ the first element of matrix $\Sigma_{i=0}^{k-1} A^i \Omega_\epsilon A^{i'}$ holds. Then the k-step ahead forecast mean squared error matrix in the updated bivariate VAR is followed by*

$$MSE^u[y_{2T}(k)] = (\sum_{i=0}^{k-1} a_{22}^{2i})(1-\rho_{12}^2)\sigma_2^2 \qquad 1 \leq k \leq s. \tag{9}$$

Proof. *See Appendix.*

Clearly, if the absolute value of estimated a_{22} is lager than one, the MSE of the updated bivariate VAR forecast is diverged when a forecast horizon is large enough, that is $k \to \infty$. However, stationary process assumption rule out the divergence case. Therefore, when the forecast horizon is large enough, the MSE of the updated bivariate VAR forecast will approach to steady state. If the absolute value of estimated a_{22} is less or equal to one, it exists a permanent efficient gain by employing the updated bivariate VAR forecast.

3 Misspecify Updated Bivariate VAR Forecast

Given one variable y_{1t} is one time period prior to another variable y_{2t}, the natural approach researches think about in order to forecast y_{2t} is by using the contemporaneous regression model, that is, we can simply regress y_{2t} on y_{1t}. Then the forecast of y_{2t+1} is just the expectation of y_{2t} conditional on y_{1t+1}. For example, monthly total nonfarm payroll employments are released by the Bureau of Labor Statistics (BLS) of the United State while monthly private nonfarm payroll employments are reported by the Automatic Data Processing (ADP). These two source of data are highly correlated. Known one time period ADP data in advance, the common knowledge to predict the BLS data at that time period is to regress BLS on ADP. Hence, the forecast of BLS is the conditional expectation of BLS given ADP. However, we will show that this contemporaneous regression model is misspecified.

In this section we investigate model misspecification of contemporaneous regression model and show that the linear regression model is misspecified and that the bivariate VAR model gains in efficiency.

Suppose true DGP is a reduced form bivariate VAR (1) process

$$
\begin{aligned}
y_{1t} &= a_{11}y_{1t-1} + a_{12}y_{2t-1} + \epsilon_{1t} && (10)\\
y_{2t} &= a_{21}y_{1t-1} + a_{22}y_{2t-1} + \epsilon_{2t}. && (11)
\end{aligned}
$$

If we model the bivariate variable y_{1t} and y_{2t} as the contemporaneous regression

$$
y_{2t} = \beta y_{1t} + u_t, \tag{12}
$$

we see from (10) that $y_{2t-1} = (y_{1t} - a_{11}y_{1t-1} - \epsilon_{1t})/a_{12}$. If we substitute this into (11), we find that

$$
\begin{aligned}
y_{2t} &= \frac{a_{22}}{a_{12}}y_{1t} + (a_{21} - \frac{a_{11}a_{22}}{a_{12}})y_{1t-1} + \epsilon_{2t} - \frac{a_{22}}{a_{12}}\epsilon_{1t} && (13)\\
&= \beta y_{1t} + \delta y_{1t-1} + u_t,
\end{aligned}
$$

where

$$
\delta \equiv a_{21} - \frac{a_{11}a_{22}}{a_{12}}, \qquad \beta \equiv \frac{a_{22}}{a_{12}}, \qquad u_t \equiv \epsilon_{2t} - \beta\epsilon_{1t}.
$$

Thus $Var(u_t)$ is equal to $\sigma_2^2 + \beta^2\sigma_1^2 - 2\beta\rho_{12}\sigma_1\sigma_2$.

Since y_{1t} is correlated to the error term ϵ_{1t} by assumption, it is also correlated to the error term u_t. Thus this expression (13) is equivalent to instrument variables estimation, true DGP of a special case of (14) along with equation (15), as follows:

$$
y_{2t} = \frac{a_{22}}{a_{12}}y_{2t-1}^{u} + (a_{21} - \frac{a_{11}a_{22}}{a_{12}})y_{1t-1} + \epsilon_{2t}, \tag{14}
$$

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

$$
y_{1t} = y_{2t-1}^{u} + \epsilon_{1t}. \tag{15}
$$

Here ϵ_{1t} is measurement error which is assumed to be identified, independent, and distributed with variance σ_1^2 and to be independent of y_{1t-1} and y_{2t}. It is also assumed that there is contemporaneous correlation of ϵ_{1t} and ϵ_{2t}. Therefore, $E(\epsilon_{1t}\epsilon_{2t}) = \rho_{12}$ for some correlation coefficient ρ_{12} such that $-1 < \rho_{12} < 1$.

If we estimate simple model like (12), one effect of the measurement error in the independent variable is to increase the variance of the error terms if ϵ_1 and ϵ_2 are negatively correlated. Another severe consequence is that the OLS estimator is biased and inconsistent. Because $y_{1t} = y^u_{2t-1} + \epsilon_{1t}$, and u_t depends on ϵ_{1t}, u_t must be correlated with y_{1t} whenever $\beta \neq 0$. In fact, since the random part of y_{1t} is ϵ_{1t} and ϵ_{1t} is correlated to ϵ_{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

$$\begin{aligned} Cov(y_{1t}, u_t) &= E(y_{1t}u_t) \\ &= E(y_{1t}E(u_t|y_{1t})) \\ &= E((y_{2t-1} + \epsilon_{1t})(\epsilon_{2t} - \beta\epsilon_{1t})) \\ &= \rho_{12}\sigma_1\sigma_2 - \beta\sigma_1^2. \end{aligned}$$

This covariance is negative if $\beta > \rho_{12}\sigma_2/\sigma_1$ and positive if $\beta < \rho_{12}\sigma_2/\sigma_1$. Since it does not depend on the sample size T, it does not go away as T 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 4 *Given the true DGP of a reduced form VAR (1) in equations (10) and (11), the OLS estimator in the linear regression model (12) 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

Federal government agencies regularly announce the latest calculated values for economic variables. A monthly announcement reports the series' value in last month. As well, firms announce their quarterly financial report (for instance, earnings per share and net income) for public use. The schedule for these announcements is known well in advance, generally by the previous year-end. The timing of these announcement either varies or fixes. The order in which variables are announced also varies or fixes each month or quarter.

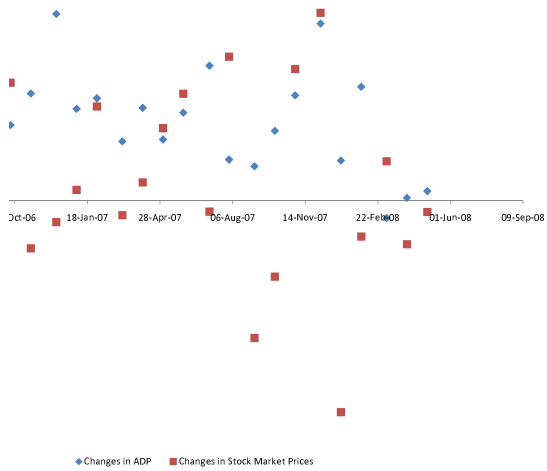
Although there are various macroeconomic announcement series or earnings announcements for firms that we could consider as an application of the research, we chose one of the real macro variables, nonfarm payroll employment in the United States, as a strong candidate for a forecasting variable. It has not only been used in the formulation of Federal Reserve policy but also used as an explanation of anomalous stock price behavior. In addition, employment announcements have tractable and fixed announcement timing and order.

One employment announcement used in the bivariate VAR model is total nonfarm payroll employment. The 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 a much larger sample size. The establishment survey includes payroll employment information, such as the total nonfarm payroll employment, weekly and hourly earnings, and weekly hours worked for several industries. 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.

Another employment announcement employed in the bivariate VAR model is private nonfarm payroll employment. 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 were made available beginning in January of 2001. Most of the announcements are made at 8:15 a.m. on a Wednesday, although a few 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 calculates 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 the BLS employment situation. The ADP report only covers private payrolls, excluding government. However, Nonfarm private employment released on the behalf of the ADP (hereafter called ADP's data) is highly correlated to nonfarm payroll employment announced two days later by BLS (hereafter called BLS's data) with a correlation of 0.82.

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 1 illustrates correlation between ADP announcements and stock market prices at the date of

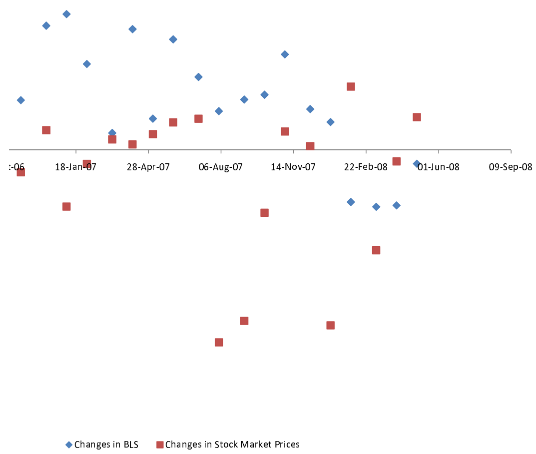
Figure 1: Effect of Private Employment Announcements on Stock Market



† The mean of changes in private employment and changes in stock price are 91.6 and 15.87 respectively. The standard deviation of changes in private employment and changes in stock price are 54.59 and 95.48 respectively. The correlation between changes in private employment and changes in stock price is 0.22.

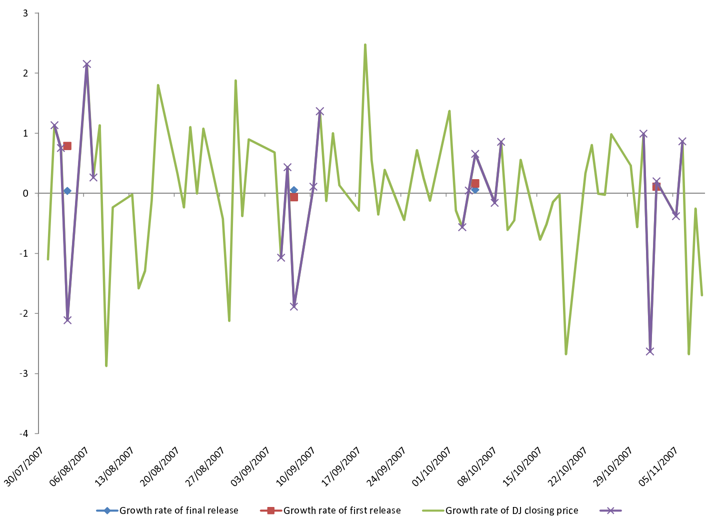
ADP announcement. The mean of changes in private employment and changes in stock price are 91.6 and 15.87 respectively. The standard deviation of changes in private employment and changes in stock price are 54.59 and 95.48 respectively. The correlation between changes in private employment and changes in stock price is 0.22. Figure 2 illustrates correlation between BLS announcements and stock market prices at the date of the BLS announcement. The mean of changes in total employment and changes in stock price are 86.84 and -31.43 respectively. The standard deviation of changes in total employment and changes in stock price are 91.34 and 108.63 respectively. The correlation between changes in total employment and changes in stock price is 0.18. Figure 3 gives the full picture of quantitative comovement between the first release of the BLS employment situation and the Dow Jones industrial index closing price. Figure 3 reports the growth rate of monthly BLS first release, final release, and daily Dow Jones Index closing price. For instance, on September 7, 2007, the BLS reported that in August, the number of nonfarm payroll employment decreased by 4,000. This led the stock market index to decline by 250 points at its closing price. On Nov 2, 2007, the BLS reported a 166 thousand increase in nonfarm employment. This led the stock index closing price to rise by 28 points at that day.

Figure 2: Effect of Total Employment Announcements on Stock Market



† The mean of changes in total employment and changes in stock price are 86.84 and -31.43 respectively. The standard deviation of changes in total employment and changes in stock price are 91.34 and 108.63 respectively. The correlation between changes in total employment and changes in stock price is 0.18.

Figure 3: Effect of the First and Final Employment Announcements on Stock Market



† The vertical axis represents the growth rate of monthly BLS first release, final release, and daily Don Jones Index closing price.

Table 1: Descriptive Statistics

Source	Automatic Data Processing, Inc. Macroeconomic Advisers, LLC.B Bureau of Labor Statistics	
Frequency	Seasonal adjusted, Monthly	
Sample period	January 2001 - September 2008	
Sample Size	93	
	First difference of *adp*	First difference of *bls*
Mean	44.53	54.22
Standard deviation	130.10	153.91
β_1	-0.121	-0.260
Augmented Dickey-Fuller	-2.421	-3.656
H_0 : Nonstationary	Reject at 95% (critical value = -1.95)	

† The coefficient β designates the autocorrelation of the series at lag i. The augmented Dickey-Fuller test is based on the following regression: $y_t - y_{t-1} = \beta_0 + \beta_1 y_{t-1} + \beta_2 \Delta y_{t-1} + \beta_3 \Delta y_{t-2} + \mu_t$. Terms added until additional lags provide no new information significant at the 5% level.

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). The 93 observations are seasonal adjusted monthly from January of 2001 to September 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 with first differences. Since the test is known to have low power, even a slight rejection means that the existence of a unit root is unlikely. The time series plot of the data is provided in Figure 4. After the first difference of each time series, the plot of the monthly changes of employees on nonfarm payrolls is provided in Figure 5.

The timing of the two time series release is fixed and in order. Table 2 shows that the ADP national employment report is released, for public use only, two days prior to publication of the employment situation by the BLS. Due to official holiday, the June and August 2008 ADP national emplyment announcements are the one which is released 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 two 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 recalculates estimates to complete employment counts available from unemployment insurance tax records, usually in

Figure 4: Nonfarm Payrolls in Level

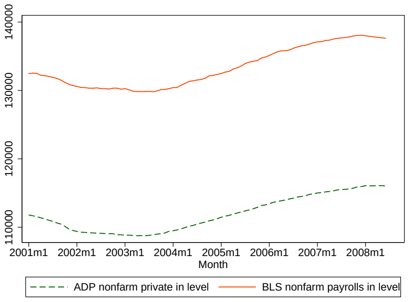
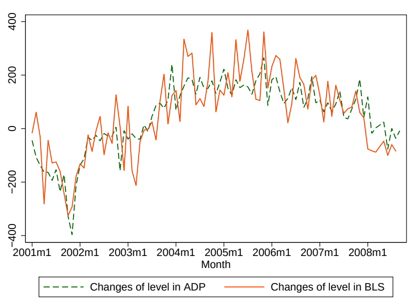


Figure 5: Changes in Nonfarm Payrolls



Source: Automatic Data Processing, Inc., Macroeconomic Advisers, LLC., Bureau of Labour Statistics. 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.

March. ADP revises its initial release every one month later. In addition to monthly revision, the entire history of the estimate of private nonfarm employment covering the period from January 2001 through January 2007 is first revised on Thursday, February 22, 2007. The first regular release of the ADP national employment report to incorporate all these revisions was published on Wednesday, March 7, 2007.

We incorporate the features of revisions existing in *bls* and *adp* time series into time series model estimation and forecasting. Rather than use final released data to estimate and forecast, we employ data of as many different releases as there are dates in the sample. More specifically, at every date within a sample, both right-side and left-side variables in the bivariate VAR model should be the most up-to-date estimate variables at that time.

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 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*, and 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 the maximum likelihood estimator and the 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). The latter two critera indicate that the optimal lag selection is two for both combining two time

Table 2: U.S. Nonfarm Payrolls Employment Announcements

	2007 Release Date		2008 Release Date	
Reference Month	ADP	BLS	ADP	BLS
January	Jan.31, Wed.	Feb.2, Fri.	Jan.30, Wed.	Feb.1, Fri.
February	Mar.7, Wed.	Mar.9, Fri.	Mar.5, Wed.	Mar.7, Fri.
March	Apr.4, Wed.	Apr.6, Fri.	Apr.2, Wed.	Apr.4, Fri.
April	May.2, Wed.	May.4, Fri.	Apr.30, Wed.	May.2, Fri.
May	May.30, Wed.	Jun.1, Fri.	Jun.4, Wed.	Jun.6, Fri.
June	Jul.5, Thu.	Jul.6, Fri.	Jul.2, Wed.	Jul.3, Thu.
July	Aug.1, Wed.	Aug.3, Fri.	Jul.30, Wed.	Aug.1, Fri.
August	Sep.5, Wed.	Sep.7, Fri.	Sep.4, Thu.	Sep.5, Fri.
September	Oct.3, Wed.	Oct.5, Fri.	Oct.1, Wed.	Oct.3, Fri.
October	Oct.31, Wed.	Nov.2, Fri.	Nov.5, Wed.	Nov.7, Fri.
November	Dec.5, Wed.	Dec.7, Fri.	Dec.3, Wed.	Dec.5, Fri.
December	Jan.2, 09, Wed.	Jan.4, 09, Fri.	Jan.7, 09, Wed.	Jan.9, 09 Fri.

series and 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 updated VAR, the ordinary VAR, and univariate autoregression.

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

$$\begin{aligned}
adp_t &= \underset{(0.0905)}{0.7812}\ adp_{t-1} + \underset{(0.0767)}{0.0999}\ bls_{t-1} + \epsilon_{1t} \\
bls_t &= \underset{(0.1282)}{0.7988}\ adp_{t-1} + \underset{(0.1087)}{0.1719}\ bls_{t-1} + \epsilon_{2t}.
\end{aligned}$$

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 updated VAR forecast, in addition to coefficient estimates and variance covariance matrix estimates, we also consider the correlation of residuals of VAR to compute the more efficient forecast mean squared error. The correlation of residuals of the bivariate VAR is 41.32%; the additional information of one series, is used to forecast the unknown value of the other series.

This is real time forecast with revisions in the data set. For each time period forecast, we collect all the information available at that time to make one- and two-step-ahead forecasts.

Table 3 shows revisions in the data set and the 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 August 1, 2007, once ADP reports its first release of nonfarm private employment for July 2007, the real time information at that time is reported in column 2 and 3 in Table 3, totalling 78 observations. After the BLS reports its employment situation for July on August 3, 2007, *bls* in June 2007 becomes the second revision and *bls* in May 2007 becomes the final revision. This can be seen from column 3 and 4 in Table 3. Furthermore, on September 5, 2007, once ADP reports its first release of nonfarm private employment for August 2007, the estimation period is from January 2001 through July 2007 of both *adp* and *bls*, for a total of 79 observations. This continues till March of 2008. On April 4, 2008, the BLS reports the first release of the employment situation for March 2007 and revises these reports on annual basis. The first release is shown in column 25 of Table 3. The last forecasting observation is on July 2, 2008, when the ADP first released its June 2008 employment report. As of July 2, 2008, we know the time series *adp* in June 2008, but we aim to forecast the time series *bls* in June 2008, which is first released on July 5, 2008.

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 January 2001 through June 2007. Following the maximum likelihood estimation by VAR, the one-step ahead residual prediction of both series is straightforward. At the second stage, we regress the residuals of the *bls* on the residuals of the *adp*. Since we observe the actual error term of the *adp* in July 2007, the best fitted residual of the *bls* in July 2007 is the estimated coefficient multiplied by the actual error term of the *adp* in July 2007.

Consequently, we reestimate the bivariate VAR over the period from January 2001 through July 2007. One-step ahead fitted values of both series are predicted for August 2007. This is our best fitted values of the *bls* and *adp* in August 2007. If we want to do further multistep forecasts, we reestimate the bivariate VAR over the period from January 2001 through August 2007. The fitted values of both series are the best forecasts in September 2007, and so on. This dynamic forecast is based on the real time information, that is, the latest actual value of *adp*. As well, the *adp* time series is highly correlated with the *bls* time series, thus this two-stage multistep VAR forecast benefits from accurate measurements.

4.3 Forecast Accuracy Comparison

Multistep ahead forecasts, computed by iterating forward the reduced form VAR, are reported in Table 4. The BLS releases revision of past employment announcements for

Table 3: Revisions in Data

On	Apr		May		Jun		Jul		Aug		Sep		Oct		Nov		Dec		Jan		Feb		Mar		April		May		June		July		Aug	
	ADP	BLS	ADP	BLS	ADP	BLS	ADP	BLS	ADP	BLS	ADP	BLS	ADP	BLS	ADP	BLS	ADP	BLS	ADP	BLS	ADP	BLS	ADP	BLS	ADP	BLS	ADP	BLS	ADP	BLS	ADP	BLS	ADP	BLS
2007m1	107	347	107	347	107	347	107	347	107	347	107	347	107	347	107	347	107	347	107	347	107	347	107	126	107	126	107	126	107	126	107	126	107	126
2007m2	62	90	62	90	62	90	62	90	62	90	62	90	62	90	62	90	62	90	62	90	62	90	62	25	62	25	62	25	62	25	62	25	62	25
2007m3	148	177	148	175	148	175	148	175	148	175	148	175	148	175	148	175	148	175	97	175	97	175	97	177	97	177	97	177	97	177	97	177	97	177
2007m4	61	88	61	80	61	122	61	122	61	122	61	122	61	122	61	122	61	122	64	122	64	122	64	46	64	46	64	46	64	46	64	46	64	46
2007m5	97		98	157	98	190	98	188	98	188	98	188	98	188	98	188	98	188	92	188	92	188	92	162	92	162	92	162	92	162	92	162	92	162
2007m6			150		143	132	143	126	143	69	143	69	143	69	143	69	143	69	141	69	141	69	141	107	141	107	141	107	141	107	141	107	141	107
2007m7					48		41	92	41	68	41	93	41	93	41	93	41	93	43	93	43	93	43	57	43	57	43	57	43	57	43	57	43	57
2007m8							38		27	-4	27	89	27	93	27	93	27	93	36	93	36	93	36	74	36	74	36	74	36	74	36	74	36	74
2007m9									58		61	110	61	96	61	44	61	44	73	44	73	44	73	81	73	81	73	81	73	81	73	81	73	81
2007m10											106		119	166	119	170	119	159	110	159	110	159	110	140	110	140	110	140	110	140	110	140	110	140
2007m11													189		173	94	173	115	185	-325	185	-325	185	60	185	60	185	60	185	60	185	60	185	60
2007m12															40		37	18	42	82	42	41	42	41	42	41	42	41	42	41	42	41	42	41
2008m1																	130		119	-17	119	-22	119	-76	119	-76	119	-76	119	-76	119	-76	119	-76
2008m2																			-23		-18	-63	-18	-76	-18	-83	-18	-83	-18	-83	-18	-83	-18	-83
2008m3																					8		3	-80	3	-71	3	-88	3	-88	3	-88	3	-88
2008m4																							10		13	-20	13	-28	13	-67	13	-67	13	-67
2008m5																									40		25	-49	25	-62	25	-47	25	-47
2008m6																											-79		-77	-62	-77	-51	-77	-100
2008m7																													9		1	-51	1	-60
2008m8																															-33		-37	-84
2008m9																																	-8	

Table 4: Root Mean Squared Errors of Out-of-Sample Forecasts

Forecast Horizon	Estimation Sample	Forecast Periods	Updating VAR	VAR	AR
1 month	2001m1-2007m4	2007m5-2008m4	71.51	86.14	71.75
	2001m1-2007m6	2007m7-2008m6	72.33	85.98	71.45
	2001m1-2007m6	2007m7-2008m9	69.12	82.39	66.80
2 months	2001m1-2007m4	2007m5-2008m4	95.64	108.44	96.44
	2001m1-2007m6	2007m7-2008m6	99.50	108.93	96.04
	2001m1-2007m6	2007m7-2008m9	94.78	101.97	88.57

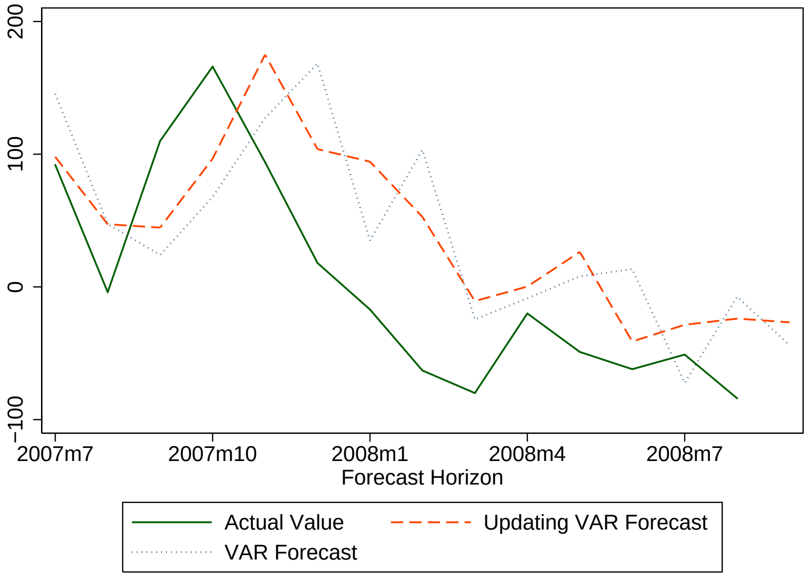
the previous three months, after which the announcement is considered final. To make our multistep-ahead forecasts consistent and comparable, we choose a forecast horizon from May 2007 through April 2008 and a forecast horizon from July 2007 through September 2008. Because the ultimate test of a forecasting model is its out-of-sample performance, Table 4 focuses on out-of-sample forecasts over the period from January 2001 through April 2007 and the period from January 2001 through June 2007. 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 and by reestimating the VAR through the next month, making the next forecast and so on through the forecast period. The key difference between the updated VAR forecasts and the ordinary VAR forecasts is the first stage estimation. Taking advantage of one more piece of information at the first forecasting period, the updated VAR forecast provides the best estimate we are going to predict.

As a comparison, out-of-sample forecasts were also computed for a univariate autoregression model with one lag, that is, a regression of the variable on lags of its own past values. Table 4 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 out-of-sample time period of May 2007 through April 2008 and of July 2007 through September 2008, and the resulting square root is the root mean squared forecast error reported in the table. In Table 4 the entries of column 3 through 6 are the root mean squared errors of the updated VAR forecast, the ordinary VAR forecast, and the univariate autoregression forecast, respectively, for nonfarm payroll employment bls. The results indicate that the updated VAR forecast has lower root mean squared error than the ordinary VAR forecast over one and two-step ahead forecast. The updated VAR forecast is slightly better or worse than univariate autoregression forecasts depending on the estimation sample we choose. Compared to the ordinary VAR quantitatively, relative efficiency gain by using the updated VAR forecast is 16% in the one-step ahead forecast and 7% in the two-step ahead forecast.

One-Step Ahead Time Series Forecast Performance

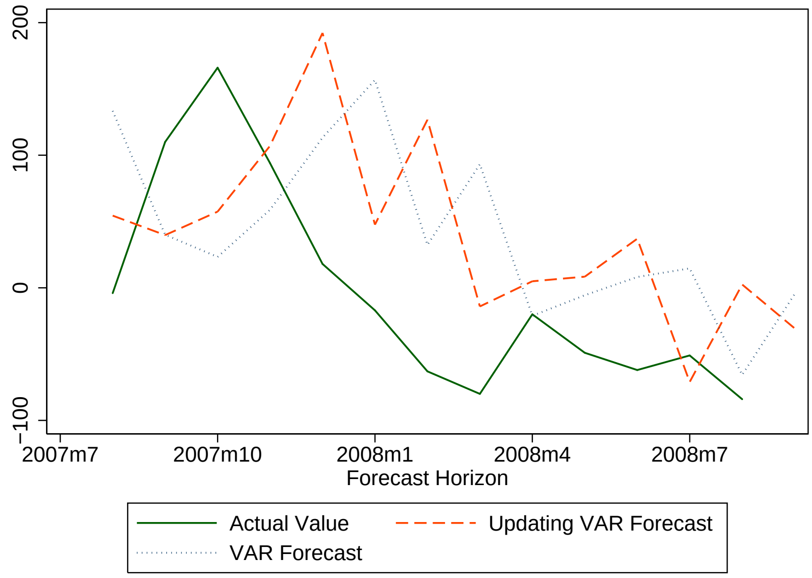
One-step ahead time series forecast is based on the time period of July 2007 through September 2008. A total 15 observations are constructed. For instance, on August 1, 2007, after ADP reported the first released nonfarm private employment, we use *adp* from January 2001 through July 2007 and *bls* from January 2001 through June 2007 to forecast *bls* in July 2007. On August 1, 2007, it was forecasted that the data for July 2007 would be released on August 3, 2007.

Figure 6: One-Step Ahead Forecast: 2007m7 - 2008m9



Omitting *adp* data in July 2007 at the time we forecasted, the ordinary bivariate VAR forecast is to estimate both *adp* and *bls* from January 2001 through June 2007. Then the one-step ahead forecast is followed. However, the updated bivariate VAR forecast takes two stages to complete. In the first stage, we estimate both *adp* and *bls* from January 2001 to June 2007. Since we know the true value of *adp* in July 2007, we have the true residual term of *adp* in July 2007. In the second stage, we regress residuals of *bls* on residuals of *adp* and obtain the coefficients of *adp*. Given the true residuals of *adp* and the coefficients of *adp*, the residual of *bls* is obviously the outcome of the true residual of *adp* multiplied by the coefficients of *adp*. Then the one-step ahead forecast of *bls* is its fitted value adding the estimated residual in July 2007.

Figure 7: Two-Steps Ahead Forecast: 2007m8 - 2008m9



As a result, standing in July 2007, the one-step ahead forecast of bls is based on both adp and bls from January 2001 to July 2007 and adp in August 2007. Continuing the process, we calculate the 15 observations of one-step ahead forecasts, which are shown in Figure 6. The solid line is the BLS first released data from July 2007 through September 2008. The first released data plays a key role in markets, so that we employ it as the actual value. We find that the updated VAR forecast outperforms the ordinary VAR forecast and is slightly better or worse than univariate autoregression forecasts depending on the estimation sample we choose.

Two-Steps Ahead Time Series Forecast Performance

The two-step ahead time series forecast is based on the time period of August 2007 through September 2008. In total, 14 observations are constructed. The difference from one-step-ahead forecast is that based on the one-step-ahead forecast of bls in July 2007, an iterated one-step ahead forecast of bls in August 2007 is conducted. Since the two-step ahead forecast is based on the one-step ahead forecast, the updated multistep VAR forecast outperforms the ordinary VAR forecast. Figure 7 reports the updated VAR forecasts, the ordinary VAR forecasts, and the univariate autoregression forecasts.

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 currently known observations will always be useful in predicting an innovation of the unknown observations. The theoretical framework shows that the currently known observations of one variable are always going to be useful for forecasting the currently unknown observations of other variables. Therefore, a higher correlation among observation innovations of multi-variables implies that the mean squared forecast error of the currently unknown observations of the other variables will be accurate for longer periods of time.

There are many applications in real-time forecasting. This paper uses U.S. labour data to examine whether *adp* estimates, which are usually announced two days prior to *bls* estimates, are helpful in forecasting the total nonfarm payroll employment number in the same month by the BLS. Rather than use the final released data to estimate and forecast, we use data of as many dates as are available in the sample. More specifically, at every date within a sample, variables in the model are the most up-to-date estimates at that time. We find that the predicted employment number is more accurate in matching the labour data when considering the real-time information than the standard time series forecast. Compared to the ordinary VAR quantitatively, the updated VAR forecast improves 16% on the one-step ahead forecast and improves 7% on the two-step ahead forecast.

The timing of data release for time series variables over 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 cases of two more available variables prior to being made public. Applications to earnings forecasts for firms by different industry sectors are yet to be developed.

Appendix

Proof of proposition 1. Given the full known information set $\mathcal{I} = \{y_{11}, y_{12}, ..., y_{1T+1}, y_{21}, ..., y_{2T}\}$, we stand at $T+1$. We know the time series of y_1 from 1 through $T+1$ while we only know the time series of y_2 from 1 through T.

For the forecast horizon at time $T+1$, we observe ϵ_{1T+1}, since $\epsilon_{1T+1} = y_{1T+1} - (a_{11}y_{1T} + a_{12}y_{2T})$. If we regress ϵ_2 on ϵ_1, the conditional expectation

$$\begin{aligned} E[\epsilon_{2T+1}|\epsilon_{1T+1}] &= \frac{Cov(\epsilon_2, \epsilon_1)}{Var(\epsilon_1)}\epsilon_{1T+1} \\ &= \frac{\rho_{12}\sigma_1\sigma_2}{\sigma_1^2}\epsilon_{1T+1}. \end{aligned}$$

So we can forecast the residual $\hat{\epsilon}_{2T+1}$ by the relationship $\hat{\epsilon}_{2T+1} = \epsilon_{2T}(1) = (\rho_{12}\sigma_2/\sigma_1)\epsilon_{1T+1}$. Then forecast error becomes:

$$\begin{aligned} y_{2T+1} - \hat{y}_{2T+1} &= \epsilon_{2T+1} - \hat{\epsilon}_{2T+1} \\ &= \epsilon_{2T+1} - \rho_{12}\frac{\sigma_2}{\sigma_1}\epsilon_{1T+1} \end{aligned}$$

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

$$\begin{aligned} MSE^u[y_{2T}(1)] &= Var[y_{2T+1} - \hat{y}_{2T+1}] \\ &= Var[\epsilon_{2T} - \epsilon_{2T}(1)] \\ &= Var[\epsilon_{2T+1} - \rho_{12}\frac{\sigma_2}{\sigma_1}\epsilon_{1T+1}] \\ &= (1 - \rho_{12}^2)\sigma_2^2 \end{aligned}$$

■

Proof of proposition 2. Given the full known information set $\mathcal{I} = \{y_{11}, y_{12}, ..., y_{1T+1}, y_{21}, ..., y_{2T}\}$, we stand at $T+1$. We know the time series of y_1 from 1 through $T+1$ while we only know the time series of y_2 from 1 through T.

The bivariate VAR (1) can be represented by the matrices in model (1). We can obtain the moving average model notation by iterating forward.

$$\begin{aligned}
Y_1 &= AY_0 + \epsilon_1 \\
Y_2 &= AY_1 + \epsilon_2 = A^2 Y_0 + A\epsilon_1 + \epsilon_2 \\
&\cdots \\
Y_t &= A^t Y_0 + \sum_{i=0}^{t-1} A^i \epsilon_{t-i}
\end{aligned}$$

The k-step ahead

$$Y_{t+k} = A^k Y_t + \sum_{i=0}^{k-1} A^i \epsilon_{t+k-i}$$

Given ϵ_{t+j}, for $j > 0$, is uncorrelated with y_{t-i}, for $i \geq 0$, the minimal forecast MSE by the ordinary VAR is

$$\begin{aligned}
MSE[Y_t(k)] &= E(\sum_{i=0}^{k-1} A^i \epsilon_{t+k-i})(\sum_{i=0}^{k-1} A^i \epsilon_{t+k-i})' \\
&= \sum_{i=0}^{k-1} A^i \Omega_\epsilon A^{i'}, \qquad (16)
\end{aligned}$$

At $T+1$, the ordinary bivariate VAR forecast MSE by equation (16) is

$$\begin{aligned}
MSE[Y_T(1)] &= \Omega_\epsilon \\
&= \begin{pmatrix} \sigma_1^2 & \rho_{12}\sigma_1\sigma_2 \\ \rho_{12}\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}.
\end{aligned}$$

For the updated VAR, we adopt all the time periods from 1 through $T+1$ for y_1 and the time periods from 1 through T for y_2. Since ϵ_1 and ϵ_2 are correlated, as well, the innovation ϵ_{1T+1} is known. By a linear regression, the best predictor of ϵ_{2T+1} based on ϵ_{1T+1} is $(\rho_{12}\sigma_2/\sigma_1)\epsilon_{1T+1}$. Thus the forecast error is

$$\begin{aligned}
Y_{T+1} - \hat{Y}_{T+1} &= \begin{pmatrix} y_{1T+1} - \hat{y}_{1T+1} \\ y_{2T+1} - \hat{y}_{2T+1} \end{pmatrix} \\
&= \begin{pmatrix} 0 \\ \epsilon_{2T+1} - \hat{\epsilon}_{2T+1} \end{pmatrix} \\
&= \begin{pmatrix} 0 \\ \epsilon_{2T+1} - \rho_{12}\frac{\sigma_2}{\sigma_1}\epsilon_{1T+1} \end{pmatrix}
\end{aligned}$$

and the MSE or forecast error covariance matrix of the updated bivariate VAR is

$$
\begin{aligned}
MSE^u[Y_T(1)] &= Var[Y_{T+1} - \hat{Y}_{T+1}] \\
&= E\begin{pmatrix} \epsilon_{1T+1} - \hat{\epsilon}_{1T+1} \\ \epsilon_{2T+1} - \hat{\epsilon}_{2T+1} \end{pmatrix} \begin{pmatrix} \epsilon_{1T+1} - \hat{\epsilon}_{1T+1} \\ \epsilon_{2T+1} - \hat{\epsilon}_{2T+1} \end{pmatrix}' \\
&= \begin{pmatrix} 0 & 0 \\ 0 & (1-\rho_{12}^2)\sigma_2^2 \end{pmatrix}.
\end{aligned}
$$

At $T+2$,

$$
\begin{aligned}
Y_{T+2} &= AY_T + \epsilon_{T+2} + A\epsilon_{T+1} \\
\hat{Y}_{T+2} &= AY_T + \hat{\epsilon}_{T+2} + A\hat{\epsilon}_{T+1}.
\end{aligned}
$$

The ordinary bivariate VAR forecast MSE followed by equation (16) is

$$
\begin{aligned}
MSE[Y_T(2)] &= \Omega_\epsilon + A\Omega_\epsilon A' \\
&= MSE[Y_T(1)] + A\Omega_\epsilon A'.
\end{aligned}
$$

The updated bivariate VAR forecast MSE is

$$
\begin{aligned}
MSE^u[Y_T(2)] &= E\left((\epsilon_{T+2} - \hat{\epsilon}_{T+2}) + A(\epsilon_{T+1} - \hat{\epsilon}_{T+1})\right)\left((\epsilon_{T+2} - \hat{\epsilon}_{T+2}) + A(\epsilon_{T+1} - \hat{\epsilon}_{T+1})\right)' \\
&= \Omega_\epsilon + A\begin{pmatrix} 0 & 0 \\ 0 & (1-\rho_{12}^2)\sigma_2^2 \end{pmatrix} A' \\
&= \Omega_\epsilon + A \;\; MSE^u[Y_T(1)] \;\; A'.
\end{aligned}
$$

At $T+3$,

$$
\begin{aligned}
Y_{T+3} &= AY_T + \epsilon_{T+3} + A\epsilon_{T+2} + A^2\epsilon_{T+1} \\
\hat{Y}_{T+3} &= AY_T + \epsilon_{T+3} + A\hat{\epsilon}_{T+2} + A^2\hat{\epsilon}_{T+1}.
\end{aligned}
$$

The ordinary bivariate VAR forecast MSE followed by equation (16) is

$$
\begin{aligned}
MSE[Y_T(3)] &= \Omega_\epsilon + A\Omega_\epsilon A' + A^2\Omega_\epsilon A^{2'} \\
&= MSE[Y_T(2)] + A^2\Omega_\epsilon A^{2'}.
\end{aligned}
$$

The updated bivariate VAR forecast MSE is

$$\begin{aligned} MSE^u[Y_T(3)] &= E\big((\epsilon_{T+3}-\hat{\epsilon}_{T+3}) + A(\epsilon_{T+2}-\hat{\epsilon}_{T+2}) + A^2(\epsilon_{T+1}-\hat{\epsilon}_{T+1})\big) \\ &\quad \big((\epsilon_{T+3}-\hat{\epsilon}_{T+3}) + A(\epsilon_{T+2}-\hat{\epsilon}_{T+2}) + A^2(\epsilon_{T+1}-\hat{\epsilon}_{T+1})\big)' \\ &= \Omega_\epsilon + A\Omega_\epsilon A' + A^2 \begin{pmatrix} 0 & 0 \\ 0 & (1-\rho_{12}^2)\sigma_2^2 \end{pmatrix} A^{2'} \\ &= \Omega_\epsilon + A \;\; MSE^u[Y_T(2)] \;\; A'. \end{aligned}$$

Recursively, the k-step ahead updated forecast error covariance matrix becomes

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

■

Proof of proposition 3. Given the full known information set $\mathcal{I} = \{y_{11}, y_{12}, ..., y_{1T+1}, ..., y_{1T+s}, y_{21}, ..., y_{2T}\}$, we know the time series of y_1 from 1 through $T+s$ while we only know the time series of y_2 from 1 through T.

At $T+2$, equation (7) gives

$$\epsilon_{1T+2} = y_{1T+2} - (a_{11}y_{1T+1} + a_{12}y_{2T+1}).$$

Since we do not observe y_{2T+1}, we do not observe ϵ_{1T+2}. There are two ways to make a prediction on ϵ_{1T+2}. One way is to set $E(\epsilon_{1T+2}) = 0$ and to set the variance of ϵ_{1T+2} be the first element of the MSE of the ordinary bivariate VAR forecast, that is, the first element of the matrix $\Omega_\epsilon + A\Omega_\epsilon A'$. The alternative way is to predict $\hat{\epsilon}_{1T+2}$ through the residual form $\hat{\epsilon}_{1T+2} = y_{1T+2} - (a_{11}y_{1T+1} + a_{12}\hat{y}_{2T+1})$. In the latter case, the variance of the difference in error becomes

$$Var[\epsilon_{1T+2} - \hat{\epsilon}_{1T+2}] = a_{12}^2(1-\rho_{12}^2)\sigma_2^2.$$

If the sufficient condition $\{a_{12}^2(1-\rho_{12}^2)\sigma_2^2 <$ the first element of matrix $\Omega_\epsilon + A\Omega_\epsilon A'\}$ holds, we would use $\hat{\epsilon}_{1T+2}$ rather than $E(\epsilon_{1T+2}) = 0$. Then notify that

$$\begin{aligned} y_{2T+2} &= a_{21}y_{1T+1} + a_{22}y_{2T+1} + \epsilon_{2T+2} \\ \hat{y}_{2T+2} &= a_{21}y_{1T+1} + a_{22}\hat{y}_{2T+1} + \hat{\epsilon}_{2T+2}. \end{aligned}$$

By equation (8) from section 2.2, the variance of the forecast error is followed by

$$\begin{aligned} Var[y_{2T+2} - \hat{y}_{2T+2}] &= a_{22}^2 Var[y_{2T+1} - \hat{y}_{2T+1}] + Var[\epsilon_{2T+2} - \hat{\epsilon}_{2T+2}] \\ &= a_{22}^2 Var[y_{2T+1} - \hat{y}_{2T+1}] + Var[\epsilon_{2T+2} - \rho_{12}\frac{\sigma_2}{\sigma_1}\hat{\epsilon}_{1T+2}] \\ &= a_{22}^2(1-\rho_{12}^2)\sigma_2^2 + (1-\rho_{12}^2)\sigma_2^2 \\ &= (1+a_{22}^2)(1-\rho_{12}^2)\sigma_2^2. \end{aligned}$$

At $T+3$, equation (7) gives

$$\epsilon_{1T+3} = y_{1T+3} - (a_{11}y_{1T+2} + a_{12}y_{2T+2}).$$

Since we do not observe y_{2T+2}, we do not observe ϵ_{1T+3}. Again, there are two ways to make a prediction on ϵ_{1T+3}. One way is to set $E(\epsilon_{1T+3}) = 0$ and to set the variance of ϵ_{1T+3} be the first element of the MSE of the ordinary bivariate VAR forecast, that is, the first element of the matrix $\Sigma_{i=0}^{2} A^i \Omega_\epsilon A^{i'}$. The alternative way is to predict $\hat{\epsilon}_{1T+3}$ through the residual form $\hat{\epsilon}_{1T+3} = y_{1T+3} - (a_{11}y_{1T+2} + a_{12}\hat{y}_{2T+2})$. In the latter case, the variance of the difference in error becomes

$$\begin{aligned} Var[\epsilon_{1T+3} - \hat{\epsilon}_{1T+3}] &= a_{12}^2 Var[y_{2T+2} - \hat{y}_{2T+2}] \\ &= a_{12}^2(1+a_{22}^2)(1-\rho_{12}^2)\sigma_2^2. \end{aligned}$$

If the sufficient condition $\{a_{12}^2(1+a_{22}^2)(1-\rho_{12}^2)\sigma_2^2 <$ the first element of matrix $\Sigma_{i=0}^{2} A^i \Omega_\epsilon A^{i'}\}$ holds, we would use $\hat{\epsilon}_{1T+3}$ rather than $E(\epsilon_{1T+3}) = 0$. Then notify that

$$\begin{aligned} y_{2T+3} &= a_{21}y_{1T+2} + a_{22}y_{2T+2} + \epsilon_{2T+3} \\ \hat{y}_{2T+3} &= a_{21}y_{1T+2} + a_{22}\hat{y}_{2T+2} + \hat{\epsilon}_{2T+3}. \end{aligned}$$

By equation (8) from section 2.2, the variance of the forecast error is followed by

$$\begin{aligned} Var[y_{2T+3} - \hat{y}_{2T+3}] &= a_{22}^2 Var[y_{2T+2} - \hat{y}_{2T+2}] + Var[\epsilon_{2T+3} - \hat{\epsilon}_{2T+3}] \\ &= a_{22}^2 Var[y_{2T+2} - \hat{y}_{2T+2}] + Var[\epsilon_{2T+3} - \rho_{12}\frac{\sigma_2}{\sigma_1}\hat{\epsilon}_{1T+3}] \\ &= a_{22}^2(1+a_{22}^2)(1-\rho_{12}^2)\sigma_2^2 + (1-\rho_{12}^2)\sigma_2^2 \\ &= (1+a_{22}^2+a_{22}^4)(1-\rho_{12}^2)\sigma_2^2. \end{aligned}$$

Iterating forward, we need to check if the sufficient condition of $\{(\Sigma_{i=0}^{k-2} a_{22}^{2i})a_{12}^2)(1-\rho_{12}^2)\sigma_2^2 <$ the first element of matrix $\Sigma_{i=0}^{k-1} A^i \Omega_\epsilon A^{i'}\}$ holds. If this sufficient condition holds, then the forecast MSE of the updated bivariate VAR given s more periods real-time information is followed by

$$MSE^u[y_{2T}(k)] = (\sum_{i=0}^{k-1} a_{22}^{2i})(1-\rho_{12}^2)\sigma_2^2, \qquad 1 \leq k \leq s.$$

■

Proof of proposition 4. Suppose the true DGP is as model (10) and (11). We estimate the ordinary least squares model of the form

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

OLS estimator ($\hat{\beta}_{ols}$) is

$$\begin{aligned}\hat{\beta}_{ols} &= \frac{\Sigma_{t=1}^T y_{1t}y_{2t}}{\Sigma_{t=1}^T y_{1t}^2} = \frac{\Sigma_{t=1}^T y_{1t}(\alpha y_{1t-1} + \beta y_{1t} + u_t)}{\Sigma_{t=1}^T y_{1t}^2} \\ &= \beta + \alpha \frac{\Sigma_{t=1}^T y_{1t}y_{1t-1}}{\Sigma_{t=1}^T y_{1t}^2} + \frac{\Sigma_{t=1}^T y_{1t}u_t}{\Sigma_{t=1}^T y_{1t}^2}.\end{aligned}$$

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

$$E(\hat{\beta}_{ols}) = \beta + \alpha + E(\frac{\Sigma_t y_{1t}u_t}{\Sigma_t y_{1t}^2}).$$

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 (12), we have

$$E(\hat{\beta}_{ols}) = \beta + E(\frac{\Sigma_t y_{1t}u_t}{\Sigma_t y_{1t}^2}).$$

This completes the proof of bias.

To see the inconsistency,

$$\begin{aligned}
plim_{t\to\infty}(\hat{\beta}_{ols}) &= plim_{t\to\infty}(\frac{\Sigma_{t=1}^{T} y_{1t}y_{2t}}{\Sigma_{t=1}^{T} y_{1t}^2}) \\
&= plim_{t\to\infty}(\frac{\sum y_{1t}(\beta y_{1t} + \alpha y_{1t-1} + u_t)}{\sum y_{1t}^2}) \\
&= \beta + plim_{t\to\infty}\frac{\sum y_{1t}u_t}{\sum y_{1t}^2} \\
&= \beta + \frac{Cov(y_{1t}, u_t)}{Var(y_{2t-1} + \varepsilon_{1t})} \\
&= \beta + \frac{\rho\sigma_1\sigma_2 - \beta\sigma_1^2}{\sigma_2^2 + \sigma_1^2} \\
&= \frac{\beta}{1 + \sigma_1^2/\sigma_2^2} + \frac{\rho\sigma_1\sigma_2}{\sigma_1^2 + \sigma_2^2} \\
&= \frac{\beta + \rho\sigma_1/\sigma_2}{1 + \sigma_1^2/\sigma_2^2}.
\end{aligned}$$

Thus $\hat{\beta}_{ols}$ will underestimate β if $\beta > \rho\sigma_2/\sigma_1$ and will overestimate β if $\beta < \rho\sigma_2/\sigma_1$. The degree of underestimation or overestimation depends on σ_1^2/σ_2^2. ■

References

Abarbanell J. S. and V. L. Bernard (1992) Tests of Analysts' Overreaction/Underreaction to Earnings Information as an Explanation for Anomalous Stock Price Behavior. *Journal of Finance* 47, 1181-1207.

Boyd, J. H., J. Hu and R. Jagannathan (2005) The Stock Markets Reaction to Unemployment News: Why Bad News Is Usually Good for Stocks. *Journal of Finance* LX, 649-672

De Bondt, Werner F. M. and Thaler, Richard H. (1985) Does the Stock Market Overreact? *Journal of Finance* 40, 793-805.

——— (1987) Further Evidence on Investor Overreaction and Stock Market Seasonality. *Journal of Finance* 42, 557-581.

Doucet, A., de Freitas N, and Gordon NJ (2001) *Sequential Monte Carlo methods in practice.* Springer: New York.

Easterwood, J. C. and S. R. Nutt (1999) Inefficiency in Analysts' Earnings Forecasts: Systematic Misreaction or Systematic Optimism? *Journal of Finance* 54, 1777-1797

Ertimur, Y., J Sunder, and S. V. Sunder (2007) Measure for Measure: The Relation between Forecast Accuracy and Recommendation Profitability of Analysts. *Journal of Accounting Research* 45, 567-606

Faust, J. et al. (2007) The high-frequency response of exchange rates and interest rates to macroeconomic announcements. *Journal of Monetary Economics* 54, 1051-1068

Flannery, M. J. and A. A. Protopapadakis (2002) Macroeconomic factors do influence aggregate stock returns. *Review of Financial Studies* 15, 751-782

Hamilton, J.D. (1994) *Time series analysis.* Princeton University Press, Princeton.

Kitagawa, G. and W. Gersch (1996) *Smoothness Priors Analysis of Time Series.* Springer: New York.

Krueger, A. B. (1996) Do markets respond more to more reliable labor market data? A test of market rationality, NBER Working paper 5769.

Lee, Charles M. C., J. Myers, and B. Swaminathan (1999) What is the intrinsic value of the Dow? *Journal of Finance* 54 (5), 1693-1741.

Lütkepohl, H. (1991) *Introduction to Multiple Time Series Analysis.* Springer: Berlin.

—— and —— (1996) *Handbook of Matrices* John Wiley & Sons: England.

Nyblom, J. and A. Harvey (2000) Tests of Common Stochastic Trends. *Econometric Theory* 16, 176-199.

Pesaran, M. H. and M. R. Wickens (1995) Handbook of Applied Econometrics: Macroeconomics. Blackwell: Oxford UK & Cambridge USA

Phillips, P.C.B. and P. Perron (1988) Testing for a Unit Root in Time Series Regressions. *Biometrika* 75, 335-346.

Rapach, D. E., M. E. Wohar, J. Rangvid (2005) Macro variables and international stock return predictability. *International Journal of Forecasting* 21, 137-166

Sarno, L. and M. Taylor (1998) Real exchange rates under the recent float: unequivocal evidence of mean reversion. *Economics Letters* 60, 131-137.

Shumway, RH. and DS. Stoffer (2000) *Time Series Analysis and its Applications.* Springer: New York.

Stock, J. H. and M. W. Watson (2001) Vector Autoregressions. *The Journal of Economic Perspectives* 15, 101-115

———— (1989) New indexes of coincident and leading economic indicators, in J. Olivier and Stanley Fisher, eds.: NBER Macroeconomics Annual 1989, Cambridge, Massachusetts.

Triantafyllopoulos, K. (2007) Covariance estimation for multivariate conditionally Gaussian dynamic linear models. *Journal of Forecasting* 26, 551-569.

Tsay, R. S. (2005) *Analysis of Financial Time Series.* John Wiley & Sons, New Jersey.