Journal of Econometrics 160 (2011) 93–101

Contents lists available at ScienceDirect

Journal of Econometrics

journal homepage: www.elsevier.com/locate/jeconom

Forecasting multivariate realized stock market volatility

Gregory H. Bauer[a,*], Keith Vorkink[b]

[a] Financial Markets Department 4E, Bank of Canada, 234 Wellington, Ottawa, Ontario, Canada K1A 0G9

[b] 667 TNRB, Provo, UT 84602, USA

ARTICLE INFO

Article history:
Available online 6 March 2010

JEL classification:
G14
C53
C32

Keywords:
HAR-RV model
Realized volatility
Covariance matrix
Factor model

ABSTRACT

We present a new matrix-logarithm model of the realized covariance matrix of stock returns. The model uses latent factors which are functions of lagged volatility, lagged returns and other forecasting variables. The model has several advantages: it is parsimonious; it does not require imposing parameter restrictions; and, it results in a positive-definite estimated covariance matrix. We apply the model to the covariance matrix of size-sorted stock returns and find that two factors are sufficient to capture most of the dynamics.

© 2010 Elsevier B.V. All rights reserved.

1. Introduction

The variances and covariances of stock returns vary over time (e.g. Andersen et al., 2005). As a result, many important financial applications require a model of the conditional covariance matrix. Three distinct categories of methods for estimating a latent conditional covariance matrix have evolved in the literature. In the first category are the various forms of the multivariate GARCH model where forecasts of future volatility depend on past volatility and shocks (e.g. Bauwens et al., 2006). In the second category, authors have modeled asset return variances and covariances as functions of a number of predetermined variables (e.g. Ferson, 1995). The third category includes multivariate stochastic volatility models (e.g. Asai et al., 2006).

In this paper, we introduce a new model of the *realized* covariance matrix.[1] We use high-frequency data to construct estimates of the daily realized variances and covariances of five size-sorted stock portfolios. By using high-frequency data we obtain an estimate of the matrix of 'quadratic variations and covariations' that differs from the true conditional covariance matrix by mean zero errors (e.g. Andersen et al. (2003) and Barndorff-Nielsen and Shephard (2004a)). This provides greater power in determining the effects of alternative forecasting variables on equity market volatility when compared to efforts based on latent volatility models.

We transform the realized covariance matrix using the matrix logarithm function to yield a series of transformed volatilities which we term the *log-space volatilities*. The matrix logarithm is a non-linear function of all of the elements of the covariance matrix and thus the log-space volatilities do not correspond one-to-one with their counterparts in the realized covariance matrix.[2] However, modeling the time variation of the log-space volatilities is straightforward and avoids the problems that plague existing estimators of the latent volatility matrix. In particular, we do not have to impose any constraints on our estimates of the log-space volatilities.

We then model the dynamics of the log-space volatility matrix using a latent factor model. The factors consist of *both* past volatilities and other variables that can help forecast future volatility. We thus are able to model the conditional covariance matrix by combining a large number of forecasting variables into a relatively small number of factors. Indeed we show that two factors can capture the volatility dynamics of the size-sorted stock portfolios.

The factor model is estimated by GMM yielding a series of filtered estimates. We then transform these fitted values, using the matrix exponential function, back into forecasts of the realized covariance matrix. Our estimated matrix is positive

* Corresponding author.
E-mail addresses: gbauer@bankofcanada.ca (G.H. Bauer), keith_vorkink@byu.edu (K. Vorkink).

1 Andersen et al. (2001) and Barndorff-Nielsen and Shephard (2002) formalized the notion of *realized volatility*.

2 The matrix logarithm has been used for estimators of latent volatility by Chiu et al. (1996) and Kawakatsu (2006) and was also suggested in Asai et al. (2006).

0304-4076/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.jeconom.2010.03.021

definite by construction and does not require any parameter restrictions to be imposed. The approach can thus be viewed as a multivariate version of standard stochastic volatility models, where the variance is an exponential function of the factors and the associated parameters.

In addition to introducing our new realized covariance matrix we also test the forecasting ability of alternative variables for time-varying equity market covariances. Our motivation is that researchers have examined a number of variables for forecasting returns but there is much less evidence that the variables forecast risks. The cross-section of small- and large-firm volatility has been examined in a number of earlier papers (e.g., Kroner and Ng (1998), Chan et al. (1999), and Moskowitz (2003)). However, these papers used models of latent volatility to capture the variation in the covariances. In contrast, we construct daily measures of the realized covariance matrix of small and large firms over the 1988 to 2002 period. Our precise measures of volatility allow a more detailed examination of the drivers of conditional covariances than prior work.

Naturally all of these advantages come at a cost. The main cost is that by performing our analysis on the log-space volatilities and then using the (non-linear) matrix exponential function, the estimated volatilities are not unbiased. However, as we show below, a simple bias correction is available that greatly reduces the problem. Another cost is that direct interpretation of the effects of an instrument on expected volatility is difficult due to the non-linear nature of the model. However, using our factor model estimates, we can obtain the derivatives of the realized covariance matrix with respect to the forecasting variables. We are able to calculate the derivatives at each point in our sample, yielding a series of conditional volatility elasticities that are functions of both the level of the volatility and the factors driving the volatility. The time series allows us to determine which variables have a large impact on time-varying expected volatility.

The paper is organized as follows. In Section 2, we present our model of matrix logarithmic realized volatility. In Section 3, we outline our method for constructing the realized volatility matrices and give the sources of the data. In Section 4, we give our results. In Section 5, we conclude.

2. Model

2.1. The matrix log transformation

In this paper, we use the matrix exponential and matrix logarithm functions to model the time-varying covariance matrix. The matrix exponential function performs a power series expansion on a square matrix A

$$V = \text{expm}\,(A) \equiv \sum_{n=0}^{\infty} \frac{1}{n!} A^n. \tag{1}$$

The matrix exponential function has a number of useful properties (Chiu et al. (1996)). The most important of these is that A is a real, symmetric matrix, if and only if V is a real, positive definite matrix. The matrix logarithm function is the inverse of the matrix exponential function. Taking the matrix logarithm of a real, positive definite matrix V results in a real, symmetric matrix A:

$$A = \text{logm}\,(V).$$

The matrix logarithm and matrix exponential functions are used in our three-step procedure to obtain forecasts of the conditional covariance matrix of stock returns. In the first step, for each day t, we use high-frequency data to construct the $P \times P$ realized conditional covariance matrix V_t.[3] The V_t matrix is positive semi-definite by construction. Applying the matrix logarithm function,

$$A_t = \text{logm}\,(V_t), \tag{2}$$

yields a real, symmetric $P \times P$ matrix A_t.

In the second step, we model the dynamics of the A_t matrix. To do this, we follow Chiu et al. (1996) and apply the vech operator to the matrix A_t

$$a_t = vech\,(A_t),$$

which stacks the elements on and below the diagonal of A_t to obtain the $p \times 1$ vector a_t, where $p = \frac{1}{2}P(P+1)$. The a_t vector forms the basis for all subsequent models. Below, we present a factor model for the a_t processes which allows both lagged values of a_t and other variables to forecast the volatility.

In the third step, we transform the fitted values from the log-volatility space into fitted values in the actual volatility space. We use the inverse of the *vech* function to form a $P \times P$ symmetric matrix $\widehat{A}_t$ of the fitted values at each time t from the vector $\widehat{a}_t$. Applying the matrix exponential function

$$\widehat{V}_t = \text{expm}\,(\widehat{A}_t), \tag{3}$$

yields the positive semi-definite matrix $\widehat{V}_t$, which is our estimate of the conditional covariance matrix for day t.

2.2. Factor models of volatility

2.2.1. Forecasting variables

We will use several different groups of variables to forecast the conditional covariance matrix. Based on the existing literature, we can separate the variables into two groups. The first are matrix-log values of realized volatility $(a_t, a_{t-1}, a_{t-2}, \ldots)$ which are used to capture the autoregressive nature of the volatility. There are three potential problems in using these variables to forecast volatility. First, the existing literature shows that capturing volatility dynamics will likely require a long lag structure. To overcome this, we adapt the Heterogeneous Autoregressive model of realized volatility (HAR-RV) of Corsi (2009) and Andersen et al. (2007) to a multivariate setting. These authors show that the aggregate market realized volatility is forecast well by a (linear) combination of lagged daily, weekly and monthly realized volatility.

The second problem is that other authors have indicated that lagged realized volatility may not be the best predictor. In particular, both Andersen et al. (2007) and Ghysels et al. (2006) find that bi-power covariation – an estimate of the continuous part of the volatility diffusion – is a good predictor of the aggregate market's realized volatility.[4] We thus construct bi-power covariation matrices aggregated over the last $d = 1, 5$ and 20 days. As in (2) above, we take the matrix logarithm of the bi-power covariation matrix over the past d days to yield $A^{BP}(d)_t$. Taking the *vech* of this matrix yields the unique elements $a^{BP}(d)_t$.

The third problem is the large number of correlated predictors. It is likely that the bi-power covariation series $a^{BP}(d)_t$ is driven by a smaller number of factors. We test this by estimating the principal components of $a^{BP}(d)_t$,

$$a^{BP}(d, i), \quad i = 1, \ldots, pc, \tag{4}$$

where $a^{BP}(d, i)$ is the ith principal component of the d-day log-space bi-power covariation matrix. We find that a small number of components captures the volatility of the daily, weekly

[3] The details of how the matrix is constructed are presented below.

[4] Barndorff-Nielsen and Shephard (2004b, 2006) develop the theory of bi-power variation, and extend their results to the multivariate case (bi-power covariation) in Barndorff-Nielsen and Shephard (2005). We construct bi-power covariation measures for our portfolios using Definition 3 of Barndorff-Nielsen and Shephard (2005).

Download English Version:

https://daneshyari.com/en/article/5096766

Download Persian Version:

https://daneshyari.com/article/5096766

Daneshyari.com