Multi-Period Forecasts of Variance: Direct, Iterated, and Mixed-Data Approaches *

Eric Ghysels† **Antonio Rubia**‡ **Rossen Valkanov**§

First Draft: May, 2006
This version: January 12, 2009

Abstract

Multi-period forecasts of stock market return variances are often used in asset pricing, portfolio allocation, risk-management and most other areas of finance where long-horizon measures of risk are necessary. Yet, very little is known about how to forecast variances several periods ahead, as most of the focus has been placed on one-period-ahead measurements and forecasts. In this paper, we compare several approaches of producing multi-period ahead forecasts of variance –iterated, direct, and mixed-data sampling (MIDAS) – as alternatives to the often-used "scaling-up" method. The comparison is conducted (pseudo) out-of-sample using returns data of the US stock market portfolio and a cross section of size and book-to-market portfolios. The comparison results are surprisingly sharp. For the market, size, and book-to-market portfolios, we obtain the same precision ordering of the variance forecasting methods. The direct approach provides the worse (in MSFE sense) forecasts; it is dominated even by the naive "scaling-up" method. Iterated forecasts are suitable for shorter horizons (5 to 10 periods ahead), but their MSFEs deteriorate fast as the horizon increases. The MIDAS forecasts perform well at long horizons: they dominate all other approaches at horizons of 10-periods ahead and higher. The MIDAS forecasting advantage becomes most apparent at horizons of 30-periods ahead and longer, where its MSFE is about 20 percent lower than that of the next best variance forecast. In sum, this study dispels the notion that variance is not forecastable at long horizons and offers an approach that delivers accurate pseudo out-of-sample predictions.

Keywords: Volatility forecasting, multi-period forecasts, mixed-data sampling
JEL: G17, C53, C52, C22

*We thank Allan Timmermann, Andrew Patton, and participants at the 2006 Jackson Hole Conference for helpful discussions. This project was initiated during Rubia's 2006 visit to UCSD. All remaining errors are our own.

†Department of Finance, Kenan-Flagler Business School and Department of Economics, University of North Carolina, Gardner Hall CB 3305, Chapel Hill, NC 27599-3305, e-mail: eghysels@unc.edu.

‡University of Alicante, Campus de San Vicente, CP 03080, Spain. E-mail: antonio.rubia@ua.es.

§Corresponding Author. Rady School of Management, UCSD, Otterson Hall, 9500 Gilman Drive, MC 0093 La Jolla, CA 92093-0093, phone: (858) 534-0898, e-mail: rvalkanov@ucsd.edu.

1 Introduction

What is the best approach to obtain multi-period forecasts of volatility? For instance, suppose we have daily stock market returns, but need weekly (5-periods ahead), monthly (20-periods ahead), or quarterly (60-periods ahead) variance forecasts for, say, portfolio allocation, risk management, or regulatory purposes. One way of obtaining these multi-period forecasts is to use a direct approach by estimating a horizon-specific forecasting model of the variance. In our example, following the seminal work of Engle (1982) and Bollerslev (1986) (surveyed in Bollerslev and Nelson (1994)), we could estimate a GARCH model with weekly, monthly, or quarterly returns and then use it to directly predict the variance over next week, month, or quarter. Alternatively, we can use an iterated approach by estimating an autoregressive forecasting model of variance (such as a GARCH) using daily returns. Weekly, monthly, or quarterly forecasts of the variance are then obtained by iterating over the daily autoregressive model for the necessary number of periods. A third method to obtain the desired multi-period forecasts is the mixed-data sampling (MIDAS) approach. A MIDAS forecasting regression, introduced by Ghysels, Santa-Clara, and Valkanov ((2005), (2006)), uses daily squared returns to produce directly multi-period volatility forecast and can be viewed as a middle ground between the direct and the iterated approaches. Yet another, perhaps less satisfactory, method is to scale-up the daily variance forecasts by k, where k is the number of trading days.[1] These approaches, with the exception of the MIDAS, have all been widely used in the empirical finance literature. Yet, little is known about their properties in the context of variance forecasting and about their ultimate out-of-sample precision.

In this paper, we compare various approaches of obtaining multi-period forecasts of stock market return variances. To our knowledge, such a comparison has never been carried out as the volatility forecasting literature has focused almost exclusively on one-period ahead forecasts. A few notable exceptions are Diebold, Hickman, Inoue, and Schuermann (1997) and Christoffersen and Diebold (2000). Perhaps a reason for the lack of papers on the subject is the theoretical difficulty of comparing multi-period forecasts. The literature on the topic hinges on the trade-off between bias and estimation variance that exists in multi-period forecasts (Findley (1983), Findley (1985), Lin and Granger (1994), Clements and Hendry (1996), Bhanzali (1999), and Chevillon and Hendry (2005)). If the one-period model is known

[1] If we make that assumption, we ignore 25 years of volatility forecasting and modeling literature, starting with Engle (1982).

(no model uncertainty) and we are only concerned with estimation uncertainty, then the iterated method produces more efficient parameter estimates and better forecasts. However, in the more realistic case of misspecification in the one-period model (model uncertainty), the direct method is more robust to biases arising from the misspecification. Because model uncertainty might lead to severe forecasting biases, the theoretical papers on the subject seem to favor direct over iterated multi-period forecasts (Bhanzali (1999), Ing (2003), and Chevillon and Hendry (2005)). Marcellino, Stock, and Watson (2006) compare direct and iterated forecasts using US macroeconomic data series. However, they don't consider variance forecasts nor do they investigate MIDAS models. It must be pointed out that, while the above papers are based on autoregressive models, none of them directly addresses multi-period variance forecasts. We make the following contributions with respect to the previous literature.

First, we investigate whether multi-horizon forecasts of the variance of US stock market returns are more accurate, in mean square error sense, than the naive but widely-used scaling approach. While it might seem obvious that the well-documented predictability of variance using one-period-ahead forecasts implies multi-period predictability, this is not necessarily the case.[2] In fact, Diebold, Hickman, Inoue, and Schuermann (1997) and Christoffersen and Diebold (2000) provide evidence that the opposite might be true in forecasting return variances. We consider variance forecasts of the US market portfolio returns as well as of five size and five book-to-market portfolio returns. The accuracy of all forecasts is assessed in a pseudo out-of-sample exercise with the mean square forecasting error (MSFE) as a criterion for forecasting accuracy. We use the MSFE because, in addition to being the most widely used loss function, it also has some desirable properties in the context of out-of-sample variance forecasts (Hansen and Lunde (2006) and Patton (2007)).

Second, we carry out an empirical comparison of the various multi-period forecasting approaches – direct, iterated, and MIDAS – using the same eleven stock return portfolios (market plus ten size and book-to-market). Because of the lack of theoretical guidance on this topic, such a comparison would not only provide some stylized facts about the long-horizon forecastability of return variances, but it would also allow us to gauge if one method produces clearly superior multi-horizon forecasts relative to the others. The results from such a data-driven comparison are ultimately conditional on the sample at hand and the design of

[2] Model uncertainty, parameter uncertainty, model instability are some of the reasons that might drive a wedge between one-period and multi-period forecasts.

the pseudo out-of-sample experiment. However, in our case, the findings (discussed below) are remarkably sharp. At the very least, they speak to the method that should be used in multi-period volatility forecasts. More generally, our results might provide guidance for future theoretical work on long-horizon variance forecasting.

As a third contribution, building on recent work by Ghysels, Santa-Clara, and Valkanov (2006), we analyze more closely the performance of the MIDAS variance forecasts because it offers a natural middle-ground between the direct and iterated approaches. Indeed in a MIDAS, daily forecasters (squared returns) are used to produce a direct forecast of the long-horizon (proxy of) variance. Similarly to an iterated approach, the MIDAS forecasts use information in the entire history of daily returns, which implies that they will be efficient. At the same time, the forecasted variable is the long-horizon (proxy of) variance, which allows us to side-step the need of aggregating the forecasts and introducing bias. We analyze five families of MIDAS forecasting models, which differ in their specification of the weights place on lagged squared returns.

Our study yields surprisingly sharp results. We carry out the pseudo out-of-sample forecasting comparison for the market and ten size and book-to-market portfolios for horizons of 5-periods ahead (weekly) to 60-periods ahead (quarterly). First, as expected, the scaling-up method performs poorly relative to the other methods. This result is consistent with the findings of Diebold, Hickman, Inoue, and Schuermann (1997) and several others who have documented the poor performance of this approach. What is surprising, however, is that the direct method is not much better. At short and medium horizons (up to 10 periods ahead), they are similar, and at long-horizons, scaling-up performs marginally better than the direct approach. Hence, if the direct method were the only alternative to the scaling-up approach, and since scaling-up is a poor forecaster of future volatility, one might come to the erroneous conclusion that the variance is hard to forecast at long horizons.

Our second result dispels that notion. We find that for the variance of the market portfolio, iterated and MIDAS forecasts perform significantly better than the scaling-up and the direct approaches. At relatively short horizons of 5- to 10-periods ahead, the iterated forecasts dominate all others. However, at horizons of 10-periods ahead and higher, MIDAS forecasts have a significantly lower MSFE relative to the other forecasts. At horizons of 30- and 60-periods ahead, the MSFE of MIDAS is more than 20 percent lower than that of the next best forecast. Hence, we find that suitable MIDAS models produce variance forecasts that are significantly better than other widely used methods.

Third, the superior performance of MIDAS in multi-period forecasts is also observed in predicting the variance of the size and book-to-market portfolios. Similarly to the market variance results, the relative precision of the MIDAS forecasts improves with the horizon. At horizons of 10-periods and higher, the MIDAS forecasts of eight out of the ten size and book-to-market portfolios dominate the iterated and direct approaches. At horizons of 30-periods and higher, the MIDAS has the smallest MSFEs amongst all forecasting methods for all ten portfolios. We observe that the variance of the size and book-to-market portfolios is significantly less predictable than that of the entire market. Also, the predictability of the variance increases with the size of the portfolio. The variance of the largest-cap stocks is the most predictable, albeit still less forecastable than the market's. We do not observe such a discernable pattern for the book-to-market portfolios.

In section two, we introduce the direct, iterated, and MIDAS multi-period forecasts. The third section discusses the loss function used to evaluate the forecasts. Section four presents the empirical results. In section five, we conclude and offer some directions for further research.

2 Multi-Period Variance Forecasts

We use the following notation. We have D daily returns indexed by $d = 1, 2, \ldots,\ D$ and $T_k = [D/k]$ long-horizon returns indexed by $t = 1_k, 2_k, \ldots,\ T_k$, where $[.]$ is the integer operator and 1_k denotes a one-step of a $k-$th horizon return, hence 1_2 is a skip-sampling scheme, etc. For instance, in our data set, we have $D = 10920$ observations from 1963 to 2005 from which we can compute $T_5 = 2184$ 5-day (or weekly) non-overlapping returns, $T_{10} = 1092$ 10-day (or bi-weekly) non-overlapping returns, and so on. To keep the notation simple, we will henceforth drop the subscript from T but will keep in mind that the number of non-overlapping returns changes with the horizon of interest, k.

The daily return is defined as $r_d = \log\left(P_{d+1}/P_d\right)$ whereas the $k-$period continuously compounded return is defined as $R_t^k = \log\left(P_{d+k}/P_d\right)$. Note that we use lower case for daily and capital for multi-period returns. All long-horizon returns are demeaned and are computed without overlap to avoid mechanical serial correlation.

We can assemble the k-period, non-overlapping, continuously compounded returns into the following information set $I_t^k = \left\{R_t^k, R_{t-1_k}^k, R_{t-2_k}^k, \ldots, R_0^k\right\}$. Similarly, the information set for the daily returns at time t is $I_t = \{r_t, r_{t-1}, r_{t-2}, \ldots, r_0\}$. We denote by I_T^k and I_T the information sets based on the history of the entire k-period and one-period returns, respectively. Note that the two information sets at the end of the sample period are different: I_T^k contains T non-overlapping k-period returns whereas I_T contains $D = Tk$ daily returns, and $I_T^k \subset I_T$.

Henceforth, we will use the following notation for the conditional variance: $V_c(a, b, i)$, where a is the starting period of the forecast, b is the forecast horizon and c is the method, and by the latter we mean *direct, iterated* and *MIDAS*. We will use the labels d, i and m, respectively. Finally, i is the information set used. Often we will drop i, or even (a, b, i) in situations where it will be unambiguous. For example, $V_I(t, 1_k,)$ versus $V_M(t, 1_k)$ are conditional forecasts, both using daily historical data I_t, to produce $k-$step ahead forecasts at time t with iterated and MIDAS methods. Finally, we will denote $V_P(t, 1_k)$ as the true - or population - conditional variance given past daily data.

2.1 Direct Variance Forecasts

The first method, perhaps the simplest to implement, is to use the multi-period returns R_t^k, and forecast the multiple horizon conditional variance directly as one step ahead. For instance, we can model $V_D(T, 1_k)$ as a GARCH(p,q) that we estimate with k-period returns in I_T^k and then forecast the next k-period variance. We call this a direct approach of forecasting and denote it by $V_D(T, 1_k, I_T^k)$, or more concisely $V_D(T, 1_k)$. One might expect this approach to yield accurate estimates on several grounds. First, the parsimony of the GARCH model makes it hard to beat in pseudo out-of-sample forecasts (Hansen and Lunde (2005)). Second, the direct approach would produce robust estimates in the sense that it does not display a bias. However, given that we use the multi-period returns R_t^k to formulate volatility forecasts, this estimator would not be as efficient as one using the information in set I_T.

In our comparison, we use a GARCH(1,1) model. We also have results from more general GARCH(p,q) models, where p and q are chosen by the Akaike Information Criterion (AIC) and Bayes Information Criterion (BIC). However, the AIC- and BIC-chosen models fail to beat the GARCH(1,1) out-of-sample. This finding confirms that the Hansen and Lunde (2005) results hold at horizons longer than one-period ahead. Henceforth, we use the GARCH(1,1) exclusively in our analysis.

2.2 Iterated Variance Forecasts

The second method is to use the daily returns r_t in I_T in estimating forecasts $V_I(T, 1_k)$. Hence, we form iterated forecasts of the daily volatility k period forward and, under the assumption that the conditional covariances are zero, we write that $V_I(T, 1_k, I_T) = \sum_{j=1}^{k} V_D(T+j, 1, I_T)$. Note that this forecast uses information at time T and the forecasts for days $T+1$, $T+2$, ..., $T+k$ would have to be iterated from the one-period daily forecasting model.

This iterated approach, seems viable because returns are serially uncorrelated (or close), but their variances are time-varying and persistent. Hence, this is an improvement over the simple scaling approach. This method has the advantage that we are using daily data to estimate the forecasting model and will hence be more efficient than the direct approach. However, since we are iterating the forecasts and summing them, then small errors due to model misspecification will be amplified. Hence, in general this method is thought, at least theoretically to be bias-prone. But it will be efficient, because the data used is high-frequency. This is particularly important in volatility models.

2.3 MIDAS Variance Forecasts

The third approach is to use the daily returns r_t and directly produce a multi-step ahead forecast using a mixed-data sampling (or MIDAS) approach. Since this approach is relatively new, we describe it more in detail. We start by formulating a MIDAS forecasting regression:

$$\tilde{V}_{t+1}^{k} = \mu_k + \phi_k \sum_{j=0}^{j^{max}} b_k(j, \theta) r_{t-j}^2 + \varepsilon_{k,t} \tag{2.1}$$

where $\tilde{V}_{t+1}^{k}$ is a measure of (future) volatility such as realized variance, e.g. $\tilde{V}_{t+1}^{k} = RV_{t+1}^{k}$ $\equiv \sum_{j=1}^{k} r_{t+j}^2$ and $b_k(j, \theta)$ is a parsimonious weighting function parameterized by a low-dimensional parameter vector θ. The intercept μ_k, slope ϕ_k, and weighting scheme parameters θ are estimated with QMLE. The regression involves data sampled at different frequencies, since in this study, the realized variance in equation (2.1) is measured at horizons ranging from one week ($k = 5$) to three months ($k = 60$), whereas the regressors are available at daily frequencies. For instance, equation (2.1) relates the realized variance over the month of, say, December (measured from the close of the market during the last trading day of November to the close of the market during the last trading day of December) with daily squared returns up to the last day of November. The weights placed on the predictive lagged squared returns are estimated in-sample and used to form an pseudo out-of-sample forecast.

As noted before, the lag coefficients $b_k(j, \theta)$ are parameterized to be a low-dimensional function of underlying parameters θ. Without this parametric restriction, the number of parameters associated with the forecasters r_{t-j}^2 would proliferate significantly, leading to in-sample overfit and poor out-of-sample forecasts. A suitable parameterization of $b_k(j, \theta)$ circumvents the problem of parameter proliferation and is one of the most important ingredients in a MIDAS regression. We consider several parameterizations of $b_k(j, \theta)$, some of which have already been suggested in previous work. Specifically we consider the following five specifications: We postulate a flexible form for the weight given to the squared return on day $t - d$:

1. Exponential:
$$b_k(j, \theta_1, \theta_2) = \frac{\exp\{\theta_1 j + \theta_2 j^2\}}{\sum_{i=0}^{\infty} \exp\{\theta_1 i + \theta_2 i^2\}}. \tag{2.2}$$
This scheme guarantees that the weights are positive (which in turn ensures that the forecasted variance is also positive) and that they add up to one. Also, the functional form in equation (2.2) can produce a wide variety of shapes for different values of the two parameters, and it is parsimonious, with only two parameters to estimate. Finally, as long as the coefficient κ_2 is negative, the weights go to zero as the lag length increases. The speed with which the weights decay controls the effective number of observations used to estimate the conditional variance.

2. Beta:
$$b_k(j, \theta_1, \theta_2) = \frac{f(\frac{j}{j^{max}}, \theta_1; \theta_2)}{\sum_{i=1}^{j^{max}} f(\frac{i}{j^{max}}, \theta_1; \theta_2)} \tag{2.3}$$
where: $f(z, a, b) = z^{a-1}(1 - z)^{b-1}/\beta(a, b)$ and $\beta(a, b)$ is based on the Gamma function, or $\beta(a, b) = \Gamma(a)\Gamma(b)/\Gamma(a + b)$. Specification (2.3) was introduced in Ghysels, Santa-Clara, and Valkanov (2002) and further explored in Ghysels, Sinko, and Valkanov (2006). One appealing feature is positivity of the coefficients, which is necessary for $a.s.$ positive definiteness of the forecasted variance. For $\theta_1 = 1$ and $\theta_2 > 1$ one has a slowly decaying pattern typical of volatility filters, which means that only one parameter is left to determine the shape, whereas in the case of $\theta_1 = \theta_2 = 1$ we obtain equal weights, which corresponds to a rolling estimator of the volatility (French, Schwert, and Stambaugh (1987), (2005)). The flexibility of the Beta function is well known and

it is often used in Bayesian econometrics to impose flexible, yet parsimonious prior distributions. The function can take many shapes, including flat weights, gradually declining weights as well as hump-shaped patterns.

3. Linear:

$$b_k(j) = 1/j^{\max} \tag{2.4}$$

where $j^{\max}$ is the truncation point specified above. This simple decay functional form has the advantage that there are no parameters to estimate in the lagged weight function and might offer good out-of-sample forecasts.

4. Hyperbolic:

$$b_k(j, \theta) = \frac{g(\frac{j}{j^{max}}, \theta)}{\sum_{i=1}^{j^{max}} g(\frac{i}{j^{max}}, \theta)} \tag{2.5}$$

where $g(j, \theta) = \Gamma(j+\theta)/\left(\Gamma(j+1)\Gamma(\theta)\right)$ which can be written equivalently as $g(0, \theta) = 1$ and $g(j, \theta) = (j + \theta - 1)\, g(j-1, \theta)/j$, for $j \geq 1$. We also impose the restriction that $\theta < 0.5$, which would ensure that the forecasted variance is stationary (Tanaka (1999)). This Gamma functional form has only one parameter to estimate. While it is not as flexible as the Beta specification about, it has been extensively used in the variance modeling literature. The weights in (2.5) decay hyperbolically rather than exponentially (Hosking (1981)). They are nothing but the impulse response of a ARFIMA model (see, Hosking (1981) and Tanaka (1999) and references therein). ARFIMA models have successfully been used by Andersen and Bollerslev (1998) and others in volatility forecasting.

5. Step-function:

$$\begin{aligned} b_k\left(j, \theta_1, \theta_2\right) &= \theta_1,\ 0 \leqslant j < s_1 \\ &= \theta_2,\ s_1 \leqslant j < s_2 \\ &= \theta_3, s_2 \leqslant j \leqslant j^{\max} \end{aligned} \tag{2.6}$$

where $\theta_1 > \theta_2 > \theta_3 > 0$ and θ_1, θ_2,and θ_3 are normalized so that the weights sum up to one as in the previous specifications. The number and location of the step is pre-specified and we will consider several specifications with two and three steps.

The MIDAS forecasts are denoted by $V_M(T, 1_k, I_T)$, or $V_M(T, 1_k)$. Whenever necessary, we will specify the weights used in the forecasts, which implicitly capture the dynamics of the conditional variance. Larger weights on distant past returns induce more persistence on the variance process. The weighting function also determines the statistical precision of the estimator by controlling the amount of data used to estimate the conditional variance. There is a tension between capturing the dynamics of variance and minimizing measurement error. Because variances change through time, we would like to use more recent observations to forecast the level of variance in the next month. However, to the extent that measuring variance precisely requires a large number of daily observations, the estimator could still place significant weight on more distant observations.

Some of the above weight specifications have been used in the previous literature, while others are new to the MIDAS framework but not to the variance forecasting literature. The exponential lag structure has been suggested by Ghysels, Santa-Clara, and Valkanov (2005) to study the risk return tradeoff, while the beta lag has been used in Ghysels, Santa-Clara, and Valkanov (2006) in comparing short-horizon forecasts using different predictors. The linear lag is a simple natural benchmark, with no parameters to estimate. Hence, it may prove robust out-of-sample. The hyperbolic weights are similar to the impulse responses of ARFIMA models which have been successfully used in the variance literature (see e.g. Andersen, Bollerslev, and Diebold (2003)). The step functions are another simple specification that does not require estimation of the parameters $b_k(j, \theta)$ but it does require a specification of the number and relative magnitude of the steps. It is the specification that presents the most data-mining problems, but is the simplest to implement. Corsi (2004) and Forsberg and Ghysels (2007) use this specification to model the variance of stock returns.

We can think of the mixed-data regression (2.1) as combining the attractive features of the iterated and direct forecasts. Notice that we can vary the forecast horizon by changing k, whereas the predictive variables remain the same and allow us to explore the richer information set I_T. This is not true for the direct approach, where the predictive variables change with the horizon and estimation and forecasts are formed using information set $I_T^k \subset I_T$. In the MIDAS forecasts, it is not the regressors that change but the estimated shape of the lag function b_k, thus changing the weights placed on the lagged daily squared returns. Moreover, we form direct forecasts of future variance at the horizon of interest without having to iterate over forecasts. This is in contrast with the iterated GARCH forecasts. Therefore, we use (2.1) to side-step the iteration and aggregation issues associated with iterated forecasts as well as the inefficient use of lagged returns that is characteristic of the direct approach.

While the MIDAS approach to formulate forecasts is quite general, we focused on regression (2.1) for several reasons. First, we could have extended the number of regression to include not only daily squared returns but other volatility forecasts such as daily absolute returns, daily range measures (high-low), and others, as done in Forsberg and Ghysels (2007) and Ghysels, Santa-Clara, and Valkanov (2006). We use daily squared returns only in order to make this forecast as directly comparable with the GARCH forecasts as possible. Moreover, the comparison of the squared daily return predictors to other predictors at shorter horizons have already been investigated extensively in Forsberg and Ghysels (2007) and Ghysels, Santa-Clara, and Valkanov (2006). MIDAS regressions typically do *not* exploit an autoregressive scheme, so that r_{t-j}^2 is not necessarily related to lags of the left hand side variable. Instead, MIDAS regressions are first and foremost regressions and therefore the selection of r_{t-j}^2 amounts to choosing the best predictor of future variance from the set of several possible measures of past fluctuations in returns. In other words, MIDAS is a reduced-form forecasting device rather than a model of conditional volatility.

3 Comparing the Forecasts

Once we have forecasts $V_D(T, 1_k)$, $V_I(T, 1_k)$ and $V_M(T, 1_k)$, we need to decide which one is the closest to the true variance. Two issues arise here. First, we don't observe the true k-period variance. Second, in evaluating the forecasts, we have to agree on a loss function. Given the first issue, we require a loss function that produces robust rankings of the forecasts even in the absence of true variance. We address these two issues below.

3.1 Proxy for Unobservable Long-Horizon Variance

The pseudo out-of-sample forecast error is

$$e_{T+1}^k \equiv V_P(T, 1_k) - V_F(T, 1_k)$$

where $V_(.)$ is the forecasted variance (either $V_D(T, 1_k)$, $V_I(T, 1_k)$, or $V_M(T, 1_k)$,) and $V_P(T, 1_k)$ is the true population variance. However, we cannot obtain e_{t+1}^k because the true variance is unobservable. Hence, we use the realized variance RV_{T+1}^k as a proxy for V_P. Andersen and Bollerslev (1998) and subsequent work show that the realized variance is a good proxy for the true variance, or at least much better than squared returns. The realized variance is computed using high-frequency returns. The consensus in that literature is that we need data that is very frequent. Unfortunately, we do not have access to high-frequency data for our sample period, nor for the size and decile portfolios in the cross-section. Hence, we use the highest frequency data that we have – daily returns – to compute the realized variance at horizon k. Given that we don't have true high frequency data, our estimated RV_{T+1}^k will be a noisy proxy of the true underlying variance. We have to keep that in mind when ranking the forecasts. Although all models are compared with the same noisy measurement.

3.2 Ranking the Forecasts: Appropriate Loss Function

Using RV_{T+1}^k, we compute the feasible pseudo out-of-sample forecast error

$$u_{T+1}^k = RV_{T+1}^k - V_F(T, 1_k, i)$$

and the sample MSFE at the k-horizon:

$$MSFE^k = \frac{1}{T_2 - T_1 + 1} \sum_{t=T_1}^{T_2} u_t^k.$$

The sample MSFE is computed for each of each forecasting horizon and for each forecasting method. We index the horizon but not the forecasting method to save on notation. For a horizon k, the empirical efficiency of the forecasts is assessed by comparing the respective MSFE. The ranking that we thus obtain will be consistent in the sense of Patton (2007). He showed that when we used the MSFE function, a forecast that dominates using the feasible errors u_{T+1}^k will also dominate using the infeasible e_{t+1}^k. In other words, the error introduced from using a proxy rather than the true volatility will not change the ranking of our forecasting methods. This robustness property is not shared by some other popular forecasting evaluation methods, such as the mean absolute forecasting error (MAFE). Hence, we focus on the MSFE as a loss function to evaluate the forecasts.

4 Data and Results

4.1 Data

We have daily CRSP log returns for the period July 1, 1963 to December 31, 2005. Using these returns, we compute k-period continuously compounded, non-overlapping returns R_t^k, for $k = 5$ (weekly), 10 (bi-weekly), 15, 20 (monthly), 30, 60 (quarterly). We also have data of five daily size and five daily book-to-market portfolio returns obtained from Kenneth French's website. In sum, we will forecast the variances of 11 portfolio returns (market plus 10 portfolios) over various horizons . The dataset is standard in empirical finance and, in the interest of conciseness, we do not provide summary statistics[3].

The log daily returns are used to compute the iterated GARCH forecasting models and as predictors in the MIDAS models. They are also used to compute the realized volatility RV_t^k for each horizon k as a proxy for the true (unobservable) variance. The long-horizon returns R_t^k are used in the direct GARCH forecasts.

4.2 Results

In Table 1, we compare the direct, iterated, and MIDAS forecasts to the scaling-up approach. Scaling the one-period variance by the horizon k is admittedly a naive approach and is not directly comparable to the other three forecasting methods. However, despite the evidence against this method (Diebold, Hickman, Inoue, and Schuermann (1997)), it is still widely used in practice.[4] We use the k-rule as a benchmark not because we think it is a particularly hard forecast to beat, but because of its wide use in the profession. The MIDAS forecasts are computed using the hyperbolic specification (2.5).

In panel A, we report the level of the MSFE of all forecasting methods discussed above for horizons of 1, 5, 10, 15, 20, 25, 30, and 60 days. Not surprisingly, the MSFEs increase approximately at rate k as the forecasting horizon increases. The one-period ahead forecasts of the iterated, direct, and integrated forecasts are the same, by construction. However, they differ as the horizon k increases. At 5-period ahead, the iterated, direct, MIDAS,

[3]They are available upon request.

[4]This method has been mentioned in the Basel II agreement, which might explain to a great extent is use by professionals.

and integrated forecasts have similar MSFEs. The k-rule forecasts are the exception with significantly higher MSFEs. As the horizon increases, the MSFE of the k-rule is similar to that of the direct and integrated forecasts. In fact, at horizons of 20 periods and higher, the k-rule forecasts are better than the direct forecasts and at par with the integrated forecasts. At horizons of 30 periods and higher, the difference is statistically significant at conventional levels. This pattern is perhaps best observed in Panel B, which reports the MSFE's of all forecasts relative to that of the k-rule. In the case of the direct forecast, the ratio of MSFE is greater than one for periods of 20 and higher. At 60 periods, the k-rule forecast has a MSFE about 30 percent lower than the direct method.

The iterated method produces the best forecasts at short horizons. In Panel B, we observe that at 5-period ahead, its relative MSFE is significantly lower than that of the k-rule. In fact, it is lower than any of the other methods, including the MIDAS. As the horizon increases, the relative forecasting performance of the iterated method quickly subsides and at 60 periods, it is similar to that of the k-rule. This result is consistent with the theoretical papers that have emphasized the bias in iterated forecasts. As the iterations increase, so does the bias. Hence, our results suggest that the iterated method is more suitable for short-horizon forecasts in the range of 5-period (one week) to 20-periods (one month).

The MIDAS method produces the best forecasts at long horizons. In Panel B, its relative MSFE is better than the k-rule at 5-period ahead forecasts and higher. More importantly, its forecasting performance only improves as the forecasting horizon increases. At 10-period ahead and higher, it produces the best forecast and its advantage increases steadily until 30-period ahead. At 60-period, its MSFE is about 23 percent lower than that of the k-rule. The long-horizon forecastability of the market variance with the MIDAS is a new result. As observed in Table 1, the rest of the forecasts are either insignificantly better or worse than the k-rule. Hence, while without the MIDAS model, one might be tempted to conclude that market variance is hard to forecast at long horizons, the MIDAS approach offers a new perspective.

It might be argued that the MIDAS approach has an unfair advantage in this pseudo out-of-sample exercise, because we have chosen the hyperbolic specification (2.5) which is known from previous work to produce good results. While the same comment can be levied against the other methods, it is interesting to see whether alternative polynomial specifications would produce vastly different forecasts. In Table 2, we compute the MSFE of various MIDAS specifications. The polynomial weights that we use in the MIDAS are the hyperbolic (2.5), the linear (2.4), the beta (2.3), the exponential (2.2), and various specifications of step functions (2.6). We also display the MSFE of the iterated forecast (as a reference) which was shown to produce the best prediction among the non-MIDAS approaches.

Panel A of Table 2 provides the MSFEs which are directly comparable with Table 1, while Panel B provides the MSFEs relative to the iterated approach. Focusing our attention on Panel B, we note that exponential weights are the worst overall. While these weights have been used successfully by Ghysels, Santa-Clara, and Valkanov (2005) to estimate the risk return tradeoff, in the context of volatility forecasting, they are dominated by the other methods. It is not surprising to find that different weights will be appropriate in different applications. The suitability of the weighting function will be determined by the stochastic properties of the predicted variable and it is not reasonable to expect one functional form of $b_k(j, \theta)$ to dominate across applications.

The hyperbolic specification produces the best multi-period forecasts. This is not surprising, given that it is very similar to an ARFIMA model (whose impulse response function is also hyperbolically decaying) and ARFIMAs produce good out-of-sample forecasts of future volatility (cite). However, it is interesting to note that the beta and most of the step specifications produce very good results, as well. With the exception of the exponential MIDAS, all other models dominate the iterated forecasts at long-horizons.

Tables 3 and 4 display results similar to those in Table 1 for the five size and five book-to-market portfolios, respectively. In the interest of conciseness, we have omitted the k-rule forecasts, as they are inferior to the alternative models for all portfolios and all horizons. First, we notice that the portfolio MSFE across forecasting methods are larger than the corresponding market MSFE in Table 1. In other words, the variances of the portofolio returns are less predictable than the variance of the market portfolio.

Looking at Panel B of Table 3, the MSFE of the direct forecasts are markedly higher than those of the iterated forecasts. For the smallest cap stocks, they are more than 4 times larger at 60-period ahead. For the largest cap stocks, they are about 17 percent larger. Hence, the direct approach is particularly inappropriate to use for volatile, small-cap stocks. By contrast, the MIDAS MSFEs are significantly lower than those of the iterated forecasts, especially at long horizons. Moreover, the large cap stocks are significantly more predictable than are smaller cap stocks. Table 4 contains similar results for the book-to-market portfolio returns. The variances of these portfolio returns seems to be more forecastable that of the size portfolios.

5 Conclusion

We consider two widely used methods to forecast volatility at long horizons: iterated and direct forecasts. In addition, we use a relatively new third approach – MIDAS. All three approaches yield multi-step ahead variance forecasts without relying on the restrictive assumption of i.i.d. returns that is implicit in the often-used “scaling-up” approach. We compare these forecasting methods in terms of their average forecasting accuracy–using the MSFE. Since no general analytic results are possible, the comparison is carried out using daily stock market returns from 1963 to 2005 for the US stock market as well as for five size and five book-to-market portfolios. All forecasts are (pseudo) out-of-sample.

The MIDAS forecasts are much more precise than those of the second best model. We conjecture that the gains in forecasting power are due to the ability of the approach to take advantage of the bias-efficiency trade-off that exists in multi-period forecasts. While the other two approaches are either efficient but biased (iterated) or unbiased but inefficient (direct), the MIDAS strikes a good balance between the two. Certainly more work is needed to establish these empirical findings on firm theoretical grounds.

References

Andersen, Torben, and Tim Bollerslev, 1998, Answering the skeptics: Yes, standard volatility models do provide accurate forecasts, *International Economic Review* 39, 885–905.

Andersen, T., T. Bollerslev, and F.X. Diebold, 2003, Parametric and nonparametric volatility measurement, in Yacine Aït-Sahalia, and Lars P. Hansen, ed.: *Handbook of Financial Econometrics*.

Bhanzali, R.J., 1999, Parameter estimation and model selection for multistep prediction of a time series: A review, in S. Ghosh, ed.: *Asymptotics, Nonparametrics, and Time Series* (Marcel Dekker).

Bollerslev, T., R. Engle, and D. Nelson, 1994, Arch models, in *Handbook of Econometrics*. Elsevier Science, Amsterdam.

Bollerslev, T., 1986, Generalized autoregressive conditional heteroskedasticity, *Journal of Econometrics* 31, 307–327.

Chevillon, Chillaume, and David Hendry, 2005, Non-parametric direct multi-step estimation for forecasting economic processes, *International Journal of Forecasting* 21, 201–218.

Christoffersen, Peter, and Francis Diebold, 2000, How relevant is volatility forecasting for financial risk management?, *Review of Economics and Statistics* 82, 12–22.

Clements, M., and David Hendry, 1996, Mutli-step estimation for forecasting, *Oxford Bulletin of Economics and Statistics* 58, 657–684.

Corsi, F., 2004, A simple long memory model of realized volatility, Tech. rep., University of Southern Switzerland.

Diebold, Francis, Andrew Hickman, Atsushi Inoue, and Til Schuermann, 1997, Converting 1-day volatility to h-day volatility: Scaling by $\sqrt{h}$ is worse than you think, mimeo, University of Pennsylvania.

Engle, R.F., 1982, Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation, *Econometrica* 50, 987–1008.

Findley, D., 1983, On the use of multiple models of multi-period forecasting, *Proceedings of the Business and Statistics Section, Americal Statistical Association* pp. 528–531.

——— , 1985, Model selection for multi-step-ahead forecasting, in S. Ghosh, and P. Young, ed.: *Proceedings of the 7th Symposium on Identification and System Parameter Estimation* . pp. 1039–1044 (Pergamon, Oxford).

Forsberg, Lars, and Eric Ghysels, 2007, Why do absolute returns predict volatility so well, *Journal of Financial Econometrics* 5, 31–67.

French, Kenneth R., William Schwert, and Robert F. Stambaugh, 1987, Expected stock returns and volatility, *Journal of Financial Economies* 19, 3–29.

Ghysels, Eric, Pedro Santa-Clara, and Rossen Valkanov, 2002, The MIDAS touch: Mixed data sampling regression models, Working paper, UNC and UCLA.

——— , 2005, There is a risk-return tradeoff after all, *Journal of Financial Economics* 76, 509–548.

——— , 2006, Predicting volatility: getting the most out of return data sampled at different frequencies, *Journal of Econometrics* 131, 59–95.

Ghysels, Eric, Arthur Sinko, and Rossen Valkanov, 2006, MIDAS Regressions: Further Results and New Directions, *Econometric Reviews* 26, 53–90.

Hansen, Peter, and Asger Lunde, 2005, A forecast comparison of volatility models: Does anything beat a garch(1,1)?, *Journal of Applied Econometrics* 20, 873–889.

——— , 2006, Consistent ranking of volatility models, *Journal of Econometrics* 131, 97–12.

Hosking, J., 1981, Fractional differencing, *Biometrika* 68, 165–176.

Ing, K.C., 2003, Multistep prediction in autoregressive processes, *Econometric Theory* 19, 254–279.

Lin, J., and Clive Granger, 1994, Model selection for multiperiod forecasts, *Journal of Forecasting* 13, 1–9.

Marcellino, Massimiliano, James Stock, and Mark Watson, 2006, A comparison of direct and iterated multistep ar methods for forecasting macroeconomic time series, *Journal of Econometrics* 135, 499–526.

Patton, Andrew, 2007, Volatility forecast comparison using imperfect volatility proxies, Working Paper, London School of Economics.

Tanaka, Katsuto, 1999, The nonstationary fractional unit root, *Econometric Theory* 15, 549–582.

Table 1: Multi-Period MSFEs of Volatility Forecasts – Market Portfolio

The table reports mean square forecasting errors (MSFE) of the market variance from the iterated V_I, direct V_D, MIDAS V_M, integrated V_{int}, and scaling-up kV_1 methods. The pseudo out-of-sample forecasts obtain by re-estimating the model at each step and using it to formulate a k-period ahead prediction. The data is from July 1, 1963 to December 31, 2004. The first 1,000 daily observations are used to estimate the first forecast. Panel A reports the level of the MSFE's whereas Panel B reports the MSFEs relative to the MSFE of the k-rule forecast.

Forecasting Method	Forecasting Horizon 1	5	10	15	20	25	30	60
				A: MSFE of All Models				
V_I	4.90E-05	1.07E-04	1.87E-04	2.69E-04	3.48E-04	4.25E-04	5.01E-04	9.54E-04
V_D	4.90E-05	1.18E-04	2.27E-04	3.07E-04	4.03E-04	5.01E-04	5.94E-04	1.26E-03
V_M	5.63E-05	1.19E-04	1.84E-04	2.47E-04	3.08E-04	3.64E-04	4.19E-04	7.41E-04
V_{int}	4.90E-05	1.17E-04	2.18E-04	3.13E-04	3.95E-04	4.66E-04	5.23E-04	9.15E-04
kV_1	5.33E-05	1.36E-04	2.29E-04	3.17E-04	3.99E-04	4.77E-04	5.52E-04	9.63E-04
				B: Relative MSFE – MSFE(Model)/MSFE(kV_1)				
V_I	0.919	0.785	0.815	0.849	0.872	0.891	0.907	0.991
V_D	0.919	0.868	0.989	0.968	1.009	1.050	1.076	1.307
V_M	1.058	0.869	0.803	0.780	0.771	0.763	0.760	0.769
V_{int}	0.919	0.857	0.948	0.988	0.989	0.978	0.947	0.950

Table 2: Multi-Period MSFEs of Various MIDAS Volatility Forecasts – Market Portfolio

The table reports mean square forecasting errors (MSFE) of the market return variance for various MIDAS specifications. The specifications are the hyperbolic denoted V_M HYPERB, and appearing in (2.5), the linear denoted V_M LINEAR and appearing in (2.4), the beta denoted V_M BETA and appearing in (2.3), the exponential denoted V_M EXP and appearing (2.2), and various specifications of step functions denoted V_M [abc] for steps os sizes a, b, and (possibly) c, as appearing in (2.6). As a reference, we also report the MSFE of the iterated method. The pseudo out-of-sample forecasts obtain by re-estimating the model at each step and using it to formulate a $k-$period ahead prediction. The data is from July 1, 1963 to December 31, 2004. The first 1,000 daily observations are used to estimate the first forecast. Panel A reports the level of the MSFE's whereas Panel B reports the MSFEs relative to the MSFE of the interated forecast.

Forecasting Method	Forecasting Horizon 1	5	10	15	20	25	30	60
				A: MSFE of All Models				
V_I	4.90E-05	1.07E-04	1.87E-04	2.69E-04	3.48E-04	4.25E-04	5.01E-04	9.54E-04
V_M HYPERB	5.63E-05	1.19E-04	1.84E-04	2.47E-04	3.08E-04	3.64E-04	4.19E-04	7.41E-04
V_M LINEAR	5.23E-05	1.20E-04	1.95E-04	2.65E-04	3.30E-04	3.92E-04	4.52E-04	7.92E-04
V_M BETA	5.85E-05	1.24E-04	1.88E-04	2.52E-04	3.16E-04	3.63E-04	4.22E-04	7.71E-04
V_M EXP	5.66E-05	1.40E-04	2.25E-04	3.09E-04	3.92E-04	4.63E-04	5.38E-04	9.51E-04
V_M [5 120]	5.90E-05	1.21E-04	1.89E-04	2.54E-04	3.18E-04	3.74E-04	4.32E-04	7.83E-04
V_M [10 120]	5.40E-05	1.12E-04	1.75E-04	2.48E-04	3.13E-04	3.66E-04	4.23E-04	7.64E-04
V_M [15 120]	6.00E-05	1.08E-04	1.73E-04	2.42E-04	3.09E-04	3.65E-04	4.19E-04	7.59E-04
V_M [20 120]	6.19E-05	1.09E-04	1.75E-04	2.39E-04	3.04E-04	3.64E-04	4.18E-04	7.63E-04
V_M [5 15 120]	5.39E-05	1.22E-04	1.91E-04	2.57E-04	3.20E-04	3.73E-04	4.34E-04	7.88E-04
V_M [5 20 120]	5.59E-05	1.22E-04	1.90E-04	2.55E-04	3.20E-04	3.68E-04	4.30E-04	7.92E-04
				B: Relative MSFE–MSFE(Model)/MSFE(V_I)				
V_M HYPERB	1.151	1.108	0.985	0.919	0.884	0.855	0.837	0.777
V_M LINEAR	1.068	1.120	1.040	0.984	0.948	0.922	0.902	0.830
V_M GAMMA	1.195	1.156	1.006	0.937	0.908	0.854	0.842	0.808
V_M EXP	1.156	1.304	1.201	1.148	1.127	1.089	1.075	0.997
V_M [5 120]	1.204	1.134	1.013	0.945	0.912	0.878	0.863	0.821
V_M [10 120]	1.102	1.047	0.936	0.924	0.897	0.861	0.845	0.801
V_M [15 120]	1.225	1.012	0.925	0.901	0.888	0.858	0.837	0.796
V_M [20 120]	1.264	1.017	0.934	0.888	0.872	0.857	0.834	0.800
V_M [5 15 120]	1.101	1.143	1.022	0.956	0.918	0.877	0.866	0.826
V_M [5 20 120]	1.142	1.139	1.015	0.950	0.920	0.866	0.858	0.830

Table 3: Multi-Period MSFEs of Volatility Forecasts – Size Portfolios

The table reports mean square forecasting errors (MSFE) of the return variance of five size-sorted portfolios. The forecasts are obtained using the V_I, direct V_D, MIDAS V_M, integrated V_{int}, and scaling-up kV_1 methods. The pseudo out-of-sample forecasts obtain by re-estimating the model at each step and using it to formulate a $k-$period ahead prediction. The data is from July 1, 1963 to December 31, 2004. The first 1,000 daily observations are used to estimate the first forecast. Panel A reports the level of the MSFE's whereas Panel B reports the MSFEs relative to the MSFE of the iterated forecast.

Size Portfolio	Forecasting Method	1	5	10	15	20	25	30	60
					Forecasting Horizon				
					A: MSFE of All Models				
1	V_I	3.13E-05	8.75E-05	1.63E-04	2.42E-04	3.19E-04	3.95E-04	4.70E-04	9.01E-04
2	V_I	4.02E-05	9.38E-05	1.68E-04	2.43E-04	3.19E-04	3.93E-04	4.65E-04	8.74E-04
3	V_I	4.12E-05	9.08E-05	1.63E-04	2.34E-04	3.06E-04	3.75E-04	4.43E-04	8.41E-04
4	V_I	4.50E-05	9.60E-05	1.71E-04	2.48E-04	3.27E-04	4.03E-04	4.78E-04	9.25E-04
5	V_I	5.88E-05	1.26E-04	2.17E-04	3.08E-04	3.95E-04	4.82E-04	5.67E-04	1.09E-03
1	V_D	3.13E-05	1.61E-04	4.42E-04	6.78E-04	1.09E-03	1.55E-03	1.87E-03	4.00E-03
2	V_D	4.02E-05	1.51E-04	3.98E-04	5.82E-04	8.23E-04	1.11E-03	1.33E-03	2.59E-03
3	V_D	4.12E-05	1.39E-04	3.36E-04	4.92E-04	6.82E-04	8.91E-04	1.05E-03	2.07E-03
4	V_D	4.50E-05	1.36E-04	2.96E-04	4.23E-04	5.68E-04	7.06E-04	8.41E-04	1.72E-03
5	V_D	5.88E-05	1.28E-04	2.30E-04	3.08E-04	4.07E-04	4.98E-04	5.94E-04	1.28E-03
1	V_M	3.39E-05	8.77E-05	1.63E-04	2.41E-04	3.14E-04	3.87E-04	4.54E-04	8.49E-04
2	V_M	4.33E-05	9.34E-05	1.68E-04	2.47E-04	3.23E-04	3.97E-04	4.65E-04	8.61E-04
3	V_M	4.31E-05	9.02E-05	1.62E-04	2.37E-04	3.08E-04	3.77E-04	4.39E-04	8.11E-04
4	V_M	4.60E-05	9.44E-05	1.70E-04	2.48E-04	3.24E-04	3.97E-04	4.62E-04	8.58E-04
5	V_M	5.72E-05	1.19E-04	2.10E-04	2.97E-04	3.89E-04	4.67E-04	5.40E-04	9.64E-04
1	V_{int}	3.13E-05	8.99E-05	1.78E-04	2.62E-04	3.39E-04	4.08E-04	4.76E-04	8.92E-04
2	V_{int}	4.02E-05	9.79E-05	1.89E-04	2.81E-04	3.61E-04	4.34E-04	4.98E-04	9.32E-04
3	V_{int}	4.12E-05	9.45E-05	1.84E-04	2.71E-04	3.45E-04	4.15E-04	4.75E-04	8.85E-04
4	V_{int}	4.50E-05	1.01E-04	1.93E-04	2.86E-04	3.66E-04	4.38E-04	5.01E-04	9.40E-04
5	V_{int}	5.88E-05	1.40E-04	2.58E-04	3.68E-04	4.63E-04	5.45E-04	6.08E-04	1.05E-03
					B: Relative MSFE–MSFE(Model)/MSFE(V_I)				
1	V_D	1.000	1.839	2.706	2.805	3.417	3.914	3.989	4.445
2	V_D	1.000	1.613	2.363	2.391	2.578	2.815	2.849	2.963
3	V_D	1.000	1.536	2.065	2.099	2.232	2.374	2.370	2.464
4	V_D	1.000	1.414	1.731	1.702	1.738	1.750	1.761	1.860
5	V_D	1.000	1.016	1.061	1.001	1.029	1.033	1.048	1.170
1	V_M	1.082	1.003	0.997	0.997	0.984	0.981	0.967	0.943
2	V_M	1.077	0.996	1.001	1.016	1.013	1.010	0.999	0.985
3	V_M	1.047	0.994	0.998	1.014	1.009	1.004	0.993	0.964
4	V_M	1.022	0.984	0.994	0.997	0.990	0.985	0.968	0.928
5	V_M	0.974	0.949	0.968	0.967	0.985	0.969	0.952	0.884
1	V_{int}	1.000	1.028	1.086	1.085	1.061	1.033	1.013	0.991
2	V_{int}	1.000	1.043	1.125	1.153	1.131	1.103	1.071	1.067
3	V_{int}	1.000	1.041	1.133	1.159	1.130	1.106	1.074	1.051
4	V_{int}	1.000	1.050	1.132	1.153	1.118	1.086	1.048	1.017
5	V_{int}	1.000	1.113	1.192	1.198	1.172	1.130	1.071	0.963

Table 4: Multi-Period MSFEs of Volatility Forecasts – BTM Portfolios

The table reports mean square forecasting errors (MSFE) of the return variance of five book-to-market-sorted (BTM) portfolios. The forecasts are obtained using the from the iterated V_I, direct V_D, MIDAS V_M, integrated V_{int}, and scaling-up kV_1 methods. The pseudo out-of-sample forecasts obtain by re-estimating the model at each step and using it to formulate a $k-$period ahead prediction. The data is from July 1, 1963 to December 31, 2004. The first 1,000 daily observations are used to estimate the first forecast. Panel A reports the level of the MSFE's whereas Panel B reports the MSFEs relative to the MSFE of the interated forecast.

BTM Portfolio	Forecasting Method	Forecasting Horizon 1	5	10	15	20	25	30	60
					A: MSFE of All Models				
1	V_I	6.44E-05	1.27E-04	2.14E-04	2.98E-04	3.78E-04	4.57E-04	5.36E-04	1.02E-03
2	V_I	5.27E-05	1.12E-04	1.94E-04	2.79E-04	3.61E-04	4.40E-04	5.17E-04	9.80E-04
3	V_I	4.47E-05	9.60E-05	1.68E-04	2.43E-04	3.16E-04	3.87E-04	4.57E-04	8.85E-04
4	V_I	4.00E-05	9.56E-05	1.68E-04	2.39E-04	3.07E-04	3.73E-04	4.35E-04	7.85E-04
5	V_I	4.17E-05	9.13E-05	1.59E-04	2.26E-04	2.91E-04	3.54E-04	4.12E-04	7.63E-04
1	V_D	6.44E-05	1.48E-04	2.83E-04	3.96E-04	5.12E-04	6.36E-04	7.65E-04	1.61E-03
2	V_D	5.27E-05	1.32E-04	2.62E-04	3.79E-04	4.75E-04	6.02E-04	7.05E-04	1.45E-03
3	V_D	4.47E-05	1.12E-04	2.21E-04	3.10E-04	3.86E-04	4.62E-04	5.62E-04	1.18E-03
4	V_D	4.00E-05	1.04E-04	1.80E-04	2.46E-04	3.21E-04	3.91E-04	4.63E-04	9.25E-04
5	V_D	4.17E-05	1.08E-04	2.20E-04	3.19E-04	4.60E-04	5.69E-04	7.07E-04	1.45E-03
1	V_M	5.35E-05	1.07E-04	1.87E-04	2.69E-04	3.51E-04	4.23E-04	4.87E-04	8.74E-04
2	V_M	6.53E-05	1.23E-04	2.10E-04	2.97E-04	3.85E-04	4.62E-04	5.38E-04	9.96E-04
3	V_M	4.55E-05	9.12E-05	1.62E-04	2.32E-04	3.01E-04	3.66E-04	4.24E-04	7.70E-04
4	V_M	3.95E-05	9.14E-05	1.60E-04	2.26E-04	2.91E-04	3.36E-04	3.90E-04	7.02E-04
5	V_M	4.20E-05	8.76E-05	1.54E-04	2.19E-04	2.84E-04	3.45E-04	3.99E-04	7.13E-04
1	V_{int}	6.44E-05	1.35E-04	2.47E-04	3.51E-04	4.41E-04	5.20E-04	5.84E-04	1.06E-03
2	V_{int}	5.27E-05	1.20E-04	2.22E-04	3.19E-04	4.04E-04	4.77E-04	5.33E-04	9.34E-04
3	V_{int}	4.47E-05	1.02E-04	1.92E-04	2.77E-04	3.54E-04	4.21E-04	4.75E-04	8.37E-04
4	V_{int}	4.00E-05	1.06E-04	1.89E-04	2.68E-04	3.32E-04	3.87E-04	4.33E-04	7.39E-04
5	V_{int}	4.17E-05	9.81E-05	1.80E-04	2.59E-04	3.25E-04	3.85E-04	4.34E-04	7.66E-04
					B: Relative MSFE–MSFE(Model)/MSFE(V_I)				
1	V_D	1.000	1.164	1.324	1.330	1.354	1.390	1.426	1.577
2	V_D	1.000	1.179	1.352	1.360	1.315	1.368	1.364	1.475
3	V_D	1.000	1.165	1.310	1.277	1.223	1.194	1.230	1.331
4	V_D	1.000	1.084	1.076	1.031	1.043	1.047	1.063	1.179
5	V_D	1.000	1.182	1.382	1.414	1.579	1.611	1.715	1.902
1	V_M	0.831	0.843	0.877	0.904	0.928	0.925	0.908	0.857
2	V_M	1.238	1.100	1.085	1.065	1.065	1.049	1.040	1.017
3	V_M	1.016	0.950	0.962	0.956	0.955	0.946	0.928	0.870
4	V_M	0.988	0.956	0.952	0.946	0.947	0.900	0.896	0.894
5	V_M	1.008	0.959	0.971	0.968	0.976	0.977	0.966	0.935
1	V_{int}	1.000	1.057	1.157	1.180	1.167	1.137	1.089	1.040
2	V_{int}	1.000	1.074	1.145	1.145	1.120	1.083	1.032	0.953
3	V_{int}	1.000	1.064	1.138	1.143	1.120	1.089	1.040	0.946
4	V_{int}	1.000	1.108	1.126	1.123	1.080	1.038	0.995	0.942
5	V_{int}	1.000	1.075	1.133	1.145	1.117	1.089	1.052	1.004