VOLATILITY MODELS AND THEIR APPLICATIONS

Luc Bauwens, Christian Hafner, Sébastien Laurent

WILEY-INTERSCIENCE

A JOHN WILEY & SONS, INC., PUBLICATION

CONTENTS

CHAPTER 1

FORECASTING VOLATILITY WITH MIDAS

Eric Ghysels (University of North Carolina), Rossen Valkanov (University of California San Diego)

1.1 INTRODUCTION

We focus on the issues pertaining to mixed frequencies - that arise typically because we would like to consider multi-step volatility forecasts while maintaining information in high frequency data. For example, when we forecast daily volatility we want to preserve the information in the intra-daily data without computing daily aggregates such as realized volatility. Likewise, when we focus on, say, weekly or monthly volatility forecasts we want to use daily returns or daily realized volatility measures.

The focus on multi-step forecasting is natural even if we do not consider the case of using intra-daily returns for the purpose of daily volatility forecasts as it features prominently in the context of Value-at-Risk (VaR) within the risk management literature. In the context of forecasting the 10-day VaR, required by the Basle accord, using daily or even intra-daily information, MIDAS models can be used to produce directly multi-step forecasts.

Volatility Models and Their Applications. By Bauwens, Hafner, Laurent

Econometric methods involving data sampled at different frequencies have been considered in recent work by [65] in a likelihood-based setting and by [64], [66] and [11] using regression-based methods. The mixed frequency setting has been labeled MIDAS, meaning Mi(xed) Da(ta) S(ampling). The original work on MIDAS focused on volatility predictions, see e.g. [4], [13], [28], [52], [25], [30], [29], [37], [39], [59], [65], [66], [67], [68], [63], [76], among others.

[14] provide a user-friendly introduction to MIDAS regressions. A Matlab Toolbox for MIDAS regressions is also available, see [85]. A topic not covered, since we deal with volatility, but noteworthy is the fact that MIDAS regressions can be related to Kalman filters and state space models, see [16].

In a first section we cover MIDAS regressions in the context of volatility forecasting. The second section covers likelihood-based models, which means we cover MIDAS as it relates to ARCH-type models. A final section covers multivariate extensions.

1.2 MIDAS REGRESSION MODELS AND VOLATILITY FORECASTING

In order to analyze the role of MIDAS in forecasting volatility let us introduce the relevant notation. Let $V_{t+1,t}$ be a measure of volatility in the next period. We focus on predicting future conditional variance, measured as increments in quadratic variation (or its log transformation), due to the large body of existing recent literature on this subject. The increments in the quadratic variation of the return process, $Q_{t+1,t}$, is not observed directly but can be measured with some discretization error. One such measure would be the sum of (future) m intra-daily squared returns, namely $\sum_{j=1}^{m} [r_{j,t}]^2$, which we will denote by $RV_{t+1,t}$. We can also consider multiple periods, which will be denoted by $RV_{t+h,t}$, for horizon h. Note that the case where no intra-daily data is available corresponds to m = 1 and RV becomes a daily squared return.

In a first subsection we cover MIDAS regressions, followed by a subsection elaborating on direct versus iterated volatility forecasting. The next subsection discusses variations on the theme of MIDAS regressions and a final subsection deals with microstructure noise and MIDAS regressions.

1.2.1 MIDAS Regressions

We start with MIDAS regressions involving daily regressors for predictions at horizon h :

$$RV_{t+h,t} = \mu + \phi \sum_{k=0}^{k^{max}} w(k,\theta) X_{t-k} + \varepsilon_t \tag{1.1}$$

The volatility specification (1.1) has a number of important features.

MIDAS regressions typically do *not* exploit an autoregressive scheme, so that X_{t-k} 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 X_{t-k} amounts to choosing the best predictor of future quadratic variation from the set of several possible measures of past fluctuations in returns. Examples of X_{t-k} are past daily squared returns (that correspond to the ARCH-type of models with some parameter restrictions, [48] and [23]), absolute daily returns (that relate to the specifications of (see e.g. [42]), realized daily volatility (e.g. [7]), realized daily power of (see [21] and [20]), and daily range (e.g. [3] and [61]). Since all of the regressors are used within a framework with the same number of parameters and the same maximum number of lags, the results from MIDAS regressions are directly comparable. Moreover, MIDAS regressions can also be extended to study the joint forecasting power of the regressors.

The weight function or the polynomial lag parameters are parameterized via Almon, Exponential Almon, Beta, linear step-functions (see below), etc., see [69], and they are especially relevant in estimating a persistent process parsimoniously, such as volatility, where distant X_{t-k} are likely to have an impact on current volatility. In addition, the parameterization allows us to compare MIDAS regressions at different frequencies as the number of parameters to estimate will be the same even though the weights on the data and the forecasting capabilities might differ across horizons. Most importantly one does not have to adjust measures of fit for the number of parameters and in most situations with one predictor one has a MIDAS model with either one or two parameters determining the pattern of the weights. Note also that in the above equation we specify a slope coefficient as the weights are normalized to add up to one. Such a restriction will not always be used in the sequel. The selection of k^{max} can be done conservatively (by taking a large value) and letting the weights die out as determined by the parameter estimation. The only cost to taking large k^{max} is the loss of initial data in the sample, which should be inconsequential in large samples.

Related to the MIDAS volatility regression is the Heterogeneous Autoregressive Realized Volatility (HAR-RV) regressions proposed by [39]. The HAR-RV model is given by:

$$RV_{t+1,t} = \mu + \beta^D RV_t^D + \beta^W RV_t^W + \beta^M RV_t^M + \varepsilon_{t+1}, \qquad (1.2)$$

which has a simple linear prediction regression using RV over heterogeneous interval sizes, daily (D), weekly (W) and monthly (M). As noted by [10] (footnote 16) and [39] (discussion on page 181) the above equation is in a sense a MIDAS regression with step-functions (in the terminology of [69]). In this regard the HAR-RV can be related to the MIDAS-RV in (1.1) of [66] and [59], using different weight functions such as the Beta, exponential Almon or step functions and different regressors not just autoregressive with mixed frequencies. Note also that both models exclude the jump component of quadratic variation. Simulation results reported in [59] also show that the difference between HAR and MIDAS models is very small for RV. For other regressors, such as the realized absolute variance, the MIDAS model performs slightly better.

It should also be noted that one can add lagged RV to the above specifications, for example for h = 1 and using intra-daily data for day t, denoted $X_{j,t}$ assuming we pick only one day of lags:

$$RV_{t+1,t} = \mu + \alpha RV_{t,t-1} + \phi \sum_{k=1}^{m} w(k,\theta) X_{j,t} + \varepsilon_t \quad (1.3)$$

The above equation is reminiscent of the ADL-MIDAS regression models used extensively in the context of macro forecasting by [12]. The above equation will also relate to the HYBRID GARCH class of models discussed later.

1.2.2 Direct versus Iterated Volatility Forecasting

The volatility measure on the left-hand side, and the predictors on the right-hand side are sampled at different frequencies. As a result the volatility in equation (1.1), can be measured at different horizons (e.g. daily, weekly, and monthly frequencies), whereas the forecasting variables X_{t-k} are available at daily or higher frequencies. Thus, this specification allows us not only to forecast volatility with data sampled at different frequencies, but also to compare such forecasts and ultimately evaluate empirically the continuous asymptotic arguments. In addition, equation (1.1) provides a method to investigate whether the use of high-frequency data necessarily leads to better volatility forecasts at various horizons.

The existent literature has placed most of the emphasis on the accuracy of one-period-ahead forecasts (see [48], [23], [5], [72]). Long-horizon volatility forecasts have received significantly less attention. Yet, financial decisions related to risk management, portfolio choice, and regulatory supervision, are often based on multi-period-ahead volatility forecasts. The preeminent long-horizon volatility forecasting approach is to scale the one-period-ahead forecasts by $\sqrt{k}$ where k is the horizon of interest. [34] and others have shown that this "scaling" approach leads to poor volatility forecasts at horizons as short as ten days. The lack of a comprehensive and rigorous treatment of multi-period volatility forecasts is linked to the more general theoretical difficulty to characterize the trade-off between bias and estimation that exists in multi-period forecasts (see [57], [58], [77], [36], [22], and [32]). The paucity of new results on this topic has lead researchers to conclude that, in general, volatility is difficult to forecast at long horizons (see [34] and [89]).

In a recent paper, [63] undertake a comprehensive empirical examination of multi-period volatility forecasting approaches, beyond the simple $\sqrt{k}$-scaling rule. They consider two alternative approaches –direct and iterative–of forming long-horizon forecasts (see [79]). The "direct" forecasting method consists of estimating a horizon-specific model of the volatility at, say, monthly or quarterly frequency, which can then be used to form direct predictions of volatility over the next month or quarter. An "iterative" forecast obtains by estimating a daily autoregressive volatility forecasting model and then iterate over the daily forecasts for the necessary number of periods to obtain monthly, or quarterly predictions of the volatility. In addition to the direct and iterated approaches, [63] consider a third, novel way of long-horizon forecasts, which is based on MIDAS regressions. A MIDAS method uses daily data to produce directly multi-period volatility forecasts and can thus be viewed as a middle ground between the direct and the iterated approaches. The results of their study suggest that long-horizon volatility is much more predictable than previously suggested at horizons as long as 60 trading days (about three months).

The direct and iterated methods [63] use are based on three volatility models: GARCH (see [48] and [23]), autoregressive models of realized volatility ([8], [6], and [9]), and integrated volatility. [63] point out that a long-horizon forecast is implicitly a joint decision of choosing the appropriate volatility model and the appropriate forecasting method. A similar distinction between a method and a model has also been made implicitly by [9] and, in a different context, by [70]. The three volatility models that [63] consider in conjunction with the iterated and direct forecasting methods give rise to six different ways to produce long-horizon forecasts. The MIDAS approach, which in essence combines the forecasting model and the long-horizon method into one step, offers a seventh way of producing multi-period-ahead forecasts of volatility.

To establish the accuracy of the seven long-term forecasts, [63] use a loss function that penalizes deviations of predictions from the ex-post realizations of the volatility (similar to [60] and [6]) and a test for predictive accuracy that allows them to compare the statistical significance of competing forecasts. They use the mean square forecasting error (MSFE) as one loss function, because of its consistency property, i.e. it delivers the same forecast ranking with the proxy as it would with the true volatility (see [82]). They use a Value-at-Risk (VAR) as an alternative metric of forecast accuracy. To test the statistical significance in predictive power, [63] use two tests. The first one, proposed by [88], takes into account parameter uncertainty, which is of particular concern in the volatility forecasting literature. The second test, proposed by [70], can be viewed as a generalization or a conditional version of the [88] test. Rather than comparing the difference in average performance, [70] consider the conditional expectation of the difference across forecasting models. This conditioning approach allows not only for parameter uncertainty (as in [88]) but also uncertainty in a number of implicit choices made by the researcher when formulating a forecast, such as what data to use, the windows of in-sample estimation period, the length of the out-of-sample forecast, among others.

Using the above setup, [63] investigate whether multi-horizon forecasts of the volatility of US stock market returns are more accurate than the naive but widely-used scaling approach. They consider volatility forecasts of the US market portfolio returns as well as of five size, five book-to-market, and ten industry portfolio returns. They find that the scaling-up method performs poorly relative to the other methods for all portfolios and horizons. This result is consistent with [41] and other papers who have documented the poor performance of this approach. More surprisingly, however, they find that the direct method does not fair much better. At horizon longer than 10 days ahead, the approach of scaling one-period-ahead forecasts performs significantly better than the direct method. Hence, if the direct method were the only alternative to the scaling approach, and since scaling is a poor forecaster of future volatility, one might come to the hasty conclusion that the volatility is hard to forecast at long horizons by any model.

[63] find that for the volatility of the market portfolio, iterated and MIDAS forecasts perform significantly better than the scaling and the direct approaches. At relatively short horizons of 5- to 10-days ahead, the iterated forecasts are quite accurate. However, at horizons of 10 days ahead and higher, MIDAS forecasts have a significantly lower MSFE relative to the other forecasts. At horizons of 30- and 60-days ahead, the MSFE of MIDAS is more than 20 percent lower than that of the next best forecast. These differences are statistically significant at the one percent level according to the [88] and [70] tests. Hence, they find that suitable MIDAS models produce multi-period volatility forecasts that are significantly better than other widely used methods.

[63] also link predictive accuracy to portfolio characteristics. They note that the superior performance of MIDAS in multi-period forecasts is also observed in predicting the volatility of the size, book-to-market, and industry portfolios. Similarly to the market volatility 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. They observe that the volatility of the size and book-to-market portfolios is significantly less predictable than that of the entire market. Also, the predictability of the volatility increases with the size of the portfolio. The volatility of the largest-cap stocks is the most predictable, albeit still less forecastable than the market's. They fail to observe such a discernible pattern for the book-to-market portfolios.

From the MSFE results, it might be tempting to generalize that the MIDAS forecasts are more accurate than the iterated forecasts which in turn dominate the direct and scaling-rule approaches. However, [63] caution that a general ranking of forecast accuracy is difficult, since it is ultimately predicated on the loss function and application at hand. As an illustration, they note that when they use the VAR as a measure of forecast accuracy, then the direct method not only dominates the iterated method, but for most portfolio returns, its coverage is close to that of the MIDAS model. Overall, however, they find that MIDAS forecasts strike a good balance between bias and estimation efficiency.

1.2.3 Variations on the Theme of MIDAS Regressions

The MIDAS approach can also be used to study various other interesting aspects of forecasting volatility. [28] provide a novel method to analyze the impact of news on forecasting volatility. The following semi-parametric regression model is proposed to predict future realized volatility (RV) with past high-frequency returns:

$$RV_{t+1,t} = \psi_0 + \sum_{j=1}^{\tau} \sum_{i=1}^{m} \psi_{i,j}(\theta) NIC(r_{j,t}) + \varepsilon_{t+1} \tag{1.4}$$

where $\psi_{i,j}(\theta)$ is a polynomial lag structure parameterized by θ, $NIC(.)$ is the news impact curve and $r_{t/m}$ is the log asset price difference (return) over some short time interval i of length m on day t. Note $i = 1, \ldots, m$ of intervals on day t.

The regression model in (1.4) shows that each intra-daily return has an impact on future volatility measured by $NIC(r^{ID}_{j,t})$ and fading away through time with weights characterized by $\psi_{i,j}(\theta)$. One can consider (1.4) as the semi-parametric (SP) model that nests a number of volatility forecasting models and in particular the benchmark realized volatility forecasting equation below:

$$RV_{t+1,t} = \psi_0 + \sum_{j=0}^{\tau} \psi_j(\theta) RV_{t-j,t-j-1} + \varepsilon_{t+1} \tag{1.5}$$

The nesting of (1.5) can be seen for $k = 1, \ldots$, when we set $\psi_{i,j} \equiv \psi_i \; \forall \; j = 1, \ldots, m$, and $NIC(r) \equiv r^2$ in equation (1.4). This nesting emphasizes the role played by both the news impact curve NIC and the lag polynomial $\psi_{i,j}$.

The reason it is possible to nest the RV AR structure is due to the multiplicative specification for $\psi_{i,j}(\theta) \equiv \psi^D_j(\theta) \times \psi^{ID}_i(\theta)$, with the parameter θ containing sub-vectors that determine the two polynomials separately. The polynomial $\psi^D_j(\theta)$ is a daily weighting scheme, similar to $\psi_i(\theta)$ in the regression model appearing in (1.5). The polynomial $\psi^{ID}_i(\theta)$ relates to the intra-daily pattern. With equal intra-daily weights one has the RV measure when NIC is quadratic. [28] adopt the following specification for the polynomials:

$$\psi^D_j(\theta)\psi^{ID}_i(\theta) = Beta(j, \tau, \theta_1, \theta_2) \times Beta(i, 1/m, \theta_3, \theta_4) \tag{1.6}$$

where τ and $1/m$ are the daily (D) and intradaily (ID) frequencies. The restriction is imposed that the intra-daily patterns wash out across the entire day, i.e. $\sum_i Beta(i, 1/m, \theta_3, \theta_4) = 1$, and also impose without loss of generality, a similar restriction on the daily polynomial, in order to identify a slope coefficient in the regressions.

The multiplicative specification (1.6) has several advantages. First, as noted before, it nests the so-called flat aggregation scheme, i.e. all intra-daily weights are equal, yields a daily model with RV when the news impact curve is quadratic. Or more formally, when $\theta_3 = \theta_4 = 1$, and $NIC(r) = r^2$ one recovers RV-based regression appearing in equation (1.5). Second, by estimating $Beta(i, 1/m, \theta_3, \theta_4)$ one lets the data decide on the proper aggregation scheme which is a generic issue pertaining in MIDAS regressions as discussed in [11]. Obviously, the intra-daily part of the polynomial will pick up how news fades away throughout the day and this - in part - depends on the well known intra-daily seasonal pattern.

Finally, the MIDAS-NIC model can also nest existing parametric specifications of news impact curves adopted in the ARCH literature, namely, the daily symmetric one when $NIC(r) = br^2$, the asymmetric GJR model when $NIC(r) = br^2 + (cr^2)\mathbf{1_{r<0}}$ (see [71]) and the asymmetric GARCH model when $NIC(r) = (b(r-c)^2)$ (see [47]).

1.2.4 Microstructure Noise and MIDAS Regressions

[68] study a regression prediction problem with volatility measures that are contaminated by microstructure noise and examine optimal sampling for the purpose of volatility prediction. The analysis is framed in the context of MIDAS regressions with regressors affected by microstructure noise. They consider univariate MIDAS regressions for the prediction performance evaluation and several realized volatility measures. Their general framework also leads us to the study of optimal sampling issues in the context of volatility prediction with microstructure noise.

The topic of their paper has been studied by a variety of authors independently and simultaneously. [62] and [67] discussed forecasting volatility and microstructure noise. [69] provided further empirical evidence expanding on [67]. [2] consider a number of stochastic volatility and jump diffusions, including the Heston and log-volatility models, and study the relative performance of the two-scales realized (henceforth TSRV) estimator versus RV estimators. They provide simulation evidence showing that TSRV largely outperforms RV.

Discussions about the impact of microstructure have mostly focused so far on *measurement* and therefore mean squared error and bias of various adjustments. [68] instead focus on prediction in a regression format, and therefore can include estimators that are suboptimal in mean square error sense, since covariation with the predictor is what matters. Previously, the optimal sampling frequency was studied in terms of *MSE of estimators* in an asymptotic setting (see [90]) and for finite samples (see [19]). They derive theoretical results for RV, TSRV, average over subsamples and [91] estimators and study theoretically optimal sampling as well.

[68] also conduct an extensive empirical study of forecasting with microstructure noise, using the same data as in [73], namely the thirty Dow Jones Industrial Average (DJIA), from January 3, 2000 to December 31, 2004. The purpose of the empirical analysis is twofold. First, they verify whether the predictions from the theory hold in actual data samples. They find that is indeed the case. Second, they also implement optimal sampling schemes empirically and check the relevance of the theoretical derivations using real data. They distinguish between “conditional” and “unconditional” optimal sampling schemes, as in [18]. They find that “conditional” optimal sampling seems to work reasonably well in practice.

1.3 LIKELIHOOD-BASED METHODS

The initial work on MIDAS and volatility involved a likelihood-based on risk-return trade-offs. In a first subsection we discuss this approach, followed by recent model specifications involving mixture of ARCH-type and MIDAS specifications. These recent extensions are covered in two subsections.

1.3.1 Risk-return Trade-off

The [80] ICAPM suggests that the conditional expected excess return on the stock market should vary positively with the market's conditional variance:

$$E_t[r_{t+1}] = \mu + \gamma Var_t[r_{t+1}], \tag{1.7}$$

where γ is the coefficient of relative risk aversion of the representative agent - which should obviously be positive and take plausible values - and, according to the model, μ should be equal to zero. The expectation and the variance of the market excess return are conditional on the information available at the beginning of the return period, time t.

[17], [60], [33], and [27] do find a positive albeit mostly insignificant relation between the conditional variance and the conditional expected return. In contrast, [26] and [81] find a significantly negative relation. [71], [74], and [86] find both a positive and a negative relation depending on the method used. The main difficulty in testing the ICAPM relation is that the conditional variance of the market is not observable and must be filtered from past returns. The conflicting findings of the above studies are mostly due to differences in the approach to modeling the conditional variance.

[65] take a different look at the risk-return tradeoff with a MIDAS forecast of the monthly variance specified as a weighted average of lagged daily squared returns and estimated via a QMLE - similar to the GARCH-in-mean approaches of and [54] and [71]. Namely, they estimate the coefficients of the conditional variance process jointly with μ and γ from the expected return equation (1.7) with quasi-maximum likelihood. Hence, this approach is very different from the MIDAS regressions discussed in the previous section. The similarity, however, is that in both MIDAS regressions and in the likelihood-based MIDAS one uses the same type of parsimoniously specified lag polynomials. In particular, [65] use an exponential Almon lag specification.

Using monthly and daily market return data from 1928 to 2000 and, with a MIDAS specification for the conditional variance, [65] find a positive and statistically significant relation between risk and return. The estimate of γ is 2.6, which lines up well with economic intuition about a reasonable level of risk aversion. The MIDAS volatility estimator explains about 40 percent of the variation of realized variance in the subsequent month and its explanatory power compares favorably to that of other models of conditional variance such as GARCH. The estimated weights on the lagged daily squared returns decay slowly, thus capturing the persistence in the conditional variance process. More impressive still is the fact that, in the ICAPM risk-return relation, the MIDAS estimator of conditional variance explains about two percent of the variation of next month's stock market returns (and five percent in the period since 1964). This is quite substantial given previous results about forecasting the stock market return. Finally, the above results are qualitatively similar when one splits the sample into two subsamples of approximately equal sizes, 1928-1963 and 1964-2000. These results are obtained when extreme outliers are winsorized.

It should be noted that the ICAPM risk-return relation has also been tested using several variations of GARCH-in-mean models. However, the evidence from that literature is inconclusive and sometimes conflicting. Using simple GARCH models,

[65] confirm the findings of [60] and [71], among others, of a positive but insignificant γ coefficient in the risk-return tradeoff. The absence of statistical significance comes both from GARCH's use of *monthly* returns in estimating the conditional variance process. The use of daily data and the flexibility of the MIDAS estimator provides the power needed to find statistical significance in the risk-return tradeoff.

A comparison of the time series of conditional variance estimated according to MIDAS, GARCH, and rolling windows reveals that while the three estimators are correlated, there are some differences that affect their ability to forecast returns in the ICAPM relation. [65] find that the MIDAS variance process is more highly correlated with both the GARCH and the rolling windows estimates than these last two are with each other. This suggests that MIDAS combines some of the unique information contained in the other two estimators. They also find that MIDAS is particularly successful at forecasting realized variance both in high and low volatility regimes. These features explain the superior performance of MIDAS in finding a positive and significant risk-return relation.

It has long been recognized that volatility tends to react more to negative returns than to positive returns. [81] and [56] show that GARCH models that incorporate this asymmetry perform better in forecasting the market variance. However, [71] show that when such asymmetric GARCH models are used in testing the risk-return tradeoff, the γ coefficient is estimated to be negative (sometimes significantly so). This stands in sharp contrast with the positive and insignificant γ obtained with symmetric GARCH models and remains a puzzle in empirical finance. To investigate this issue, [65] also extend the MIDAS approach to capture asymmetries in the dynamics of conditional variance by allowing lagged positive and negative daily squared returns to have different weights in the estimator. Contrary to the asymmetric GARCH results, they still find a large positive estimate of γ that is statistically significant. In particular, they find that what matters for the tests of the risk-return tradeoff is not so much the asymmetry in the conditional variance process but rather its persistence. In this respect, asymmetric GARCH and asymmetric MIDAS models prove to be very different. Consistent with the GARCH literature, negative shocks have a larger immediate impact on the MIDAS conditional variance estimator than do positive shocks. However, [65] find that the impact of negative returns on variance is only temporary and lasts no more than one month. Positive returns, on the other hand, have an extremely persistent impact on the variance process. In other words, while short-term fluctuations in the conditional variance are mostly due to negative shocks, the persistence of the variance process is primarily driven by positive shocks. This is an intriguing finding about the dynamics of the variance process. Although asymmetric GARCH models allow for a different response of the conditional variance to positive and negative shocks, they constrain the persistence of both types of shocks to be the same. Since the asymmetric GARCH models "load" heavily on negative shocks and these have little persistence, the estimated conditional variance process shows little to no persistence.

1.3.2 HYBRID GARCH Models

The volatility specification in [65] involves a single polynomial applied to daily data. Similar to the specification of the MIDAS regression (1.3) one could think of introducing lagged volatilities. We do not operate in a regression format, so this approach would be similar to the specification of a GARCH model.

This insight has recently been pursued by [29] and [30]. A key ingredient of conditional volatility models is that more weight is attached to the most recent returns (i.e. information). In the case of the original ARCH model (see e.g. [48]) that means the most recent (daily) squared returns have more weight when predicting future (daily) conditional volatility. While intra-daily data are used to construct RV, prediction models put more weights on recent (daily) RV, but despite the use of intra-daily data - do not differentiate among intra-daily returns. If volatility is a persistent process, it would be natural to weight intra-daily data differently, as pointed out recently by [78]. This is one example of the class of models [30] so called HYBRID GARCH models. They are a unifying framework, based on a generic GARCH-type model, that addresses the issue of volatility forecasting involving forecast horizons of a different frequency than the information set. Hence, [30] propose a class of GARCH models that can handle volatility forecasts over the next five business days and use past daily data, or tomorrow's expected volatility while using intra-daily returns. The models are called HYBRID GARCH, which stands for ***H***igh Frequenc***Y*** Data-***B***ased P***R***oject***I***on-***D***riven GARCH models as the GARCH dynamics are driven by what [30] call HYBRID processes.

Compared to [78], they go beyond linear projections - albeit in a discrete time setting. The HYBRID GARCH models do have a connecting with continuous time models as well when one restricts attention to linear projections. [30] study three broad classes of HYBRID processes: (1) parameter-free processes that are purely data-driven, (2) structural HYBRIDs where one assumes an underlying DGP for the high frequency data and finally (3) HYBRID filter processes. HYBRID-GARCH models. In case (1) the HYBRID process H_τ does not depend on parameters. The obvious case would be a simple return process such that $V_{\tau+1|\tau}$ is the conditional volatility of the next period. More recently, however, other purely data-driven examples of what we call generic HYBRID processes have been suggested. For example [51], [40], [87], [84] suggest the use of (daily) realized volatilities, high-low range or realized kernels or generic realized measures as they are called by [84]. Structural HYBRID processes appear in the context of temporal aggregation - a topic discussed extensively in the (weak) GARCH literature, see e.g. [44], [45], among others. Finally, the HYBRID process $H(\phi, \vec{r}_\tau)$ can involve parameters that are *not* explicitly related to $\tilde{a}$, $\tilde{b}$ and γ appearing in (1.8). There is no underlying high frequency data DGP that is being assumed, unlike in the structural HYBRID case. One can view this as a GARCH model driven by a filtered high frequency process - where the filter weights - (hyper-)parameterized by ϕ are estimated jointly with the volatility dynamics parameters.

A generic HYBRID GARCH model has the following dynamics for volatility:

$$V_{t+1|t} = \alpha + \beta V_{t|t-1} + \gamma H_t \qquad (1.8)$$

where H_t will be called a HYBRID process. When H_t is simply a daily squared return we have the volatility dynamics of a standard daily GARCH(1,1), or H_t a weekly squared return those of a standard weekly GARCH(1,1). However, what would happen if we want to attribute an individual weight to each of the five days in a week? In this case we might consider a process $H_t \equiv \sum_{j=0}^{4} \omega_j r_{t-j/5}^2$, where we use the notation $r_{t-j/5}$ to indicate intra-period returns - in the this case daily observations of week t (when days spill over into the previous week, we assume $r_{t-j/m} \equiv r_{t-1-(j-5)/m}$). This is an example of a parameter-driven HYBRID process $H_t \equiv H(\phi, \vec{r}_t)$ where $\vec{r}_t = (r_{t-1+1/m}, r_{t-1+2/m}, \dots, r_{t-1/m}, r_t)^T$ is $\mathbf{R^m}-$valued random vector (in this case and m = 5). In addition, the weights $(\omega_j(\phi), j = 0, \dots, m-1)$ are governed by a low-dimensional parameter vector ϕ. One can think of at least two possibilities: (1) the weights are treated as additional parameters and estimated as such (with m small this is possible, but not as m gets large), or (2) anchor the weights ω_j to an underlying daily GARCH(1,1) in which case the parameters α, β and γ and the weights in ϕ are jointly related to the assumed daily DGP. The discussion so far implicitly relates to many issues we elaborate on next.

The HYBRID process H_t may be purely data-driven and not depend on parameters. The obvious case would be a simple return process such that $V_{t+1|t}$ has the typical GARCH(1,1) dynamics. More recently, however, other purely data-driven examples of what we call generic HYBRID processes have been suggested. For example [51], [40], [87], [84], [**?**] suggest the use of (daily) realized volatilities, high-low range or realized kernels or generic realized measures. It is important to note that typically parameter-free HYBRID processes do not differentiate intra-period returns, i.e. an equal weighting scheme is supposed - although some kernel-weighting or pre-averaging may take place to eliminate micro-structure noise.

To study structural HYBRIDs consider a daily weak GARCH(1,1), as defined by [44], then the implied weekly prediction, using past daily returns is:

$$V_{t+1|t} = \alpha_m + \beta_m V_{t|t-1} + \gamma_m \sum_{j=0}^{m-1} \beta_m^{j/m} r_{t-j/m}^2, \quad t \in \mathbf{Z} \tag{1.9}$$

with m = 5, and where α_m, β_m and γ_m depend on the daily GARCH(1,1) parameters α_1, β_1 and γ_1. Note that all the parameters are driven by the daily parameters. Therefore, while the HYBRID process is parameter-driven it is in principle an integral part of the volatility dynamics and $H(\phi, \vec{r}_t)$ in (1.8) does not involve stand-alone parameters ϕ. This will have consequences when we elaborate on the estimation of HYBRID GARCH models. Indeed, the context of temporal aggregation precludes us from using, say standard QMLE methods, a topic that will be discussed later.

Finally, consider a HYBRID filtering process. Here the HYBRID process $H(\phi, \vec{r}_t)$ in (1.8) involves parameters that are *not* explicitly related to α, β and γ appearing in (1.8). There is no underlying high frequency data DGP that is being assumed, unlike in the structural HYBRID case. One can view this as a GARCH model driven by a filtered high frequency process - where the filter weights - (hyper-)parameterized by ϕ are estimated jointly with the volatility dynamics parameters. The choice of the parameterizations of is inspired by [28]. The commonly used specifications are exponential, beta, linear, hyperbolic, and geometric weights. This approach has implications too as far as estimation is concerned. Unlike the structural HYBRID case, we now can consider likelihood-based methods, although the regularity conditions required are novel and more involved as those of the usual QMLE approach to GARCH estimation for instance in [24].

So far we have done the same as [78] in terms of the formulation of HYBRID processes in the context of discrete time GARCH dynamics. At this stage, we start to deviate from the linear projection paradigm and continue the logic of GARCH modeling. In light of these finding we consider HYBRID GARCH models that feature intra-daily news impact curves - similar to the framework of [28], except that the latter use a MIDAS regression format. The HYBRID processes are of the following type:

$$H_t(\phi) = \sum_{j=0}^{m-1} \Psi_j(\phi_1) NIC(\phi_2, r_{t-j/m}), \quad \sum_{j=0}^{m-1} \Psi_j(\phi_1) = 1 \tag{1.10}$$

where $NIC(\phi_2, \cdot)$ stands for a high frequency data news impact curve discussed earlier.

Various estimation procedures can be considered - some tailored to specific cases of HYBRID processes. Let us first collect all the parameters of the model appearing in (1.8) in a parameter vector called $\theta \in \Theta$, with the (pseudo-) true parameter being denoted θ_0. One has to keep in mind that specific cases - notably involving structural HYBRID processes - may involve constraints across the parameters in (1.8) or the filtering weights of the HYBRID process may also be hyper-parameterized, so that the dimension of θ (denoted as d) depends on the specific circumstances considered. For this generic setting we have the following estimators:

$$\hat{\theta}_T^{mdrv} = \arg\min_{\theta \in \mathcal{C}} \frac{1}{T} \sum_{t=1}^{T} (RV_t - V_{t|t-1}(\theta))^2$$

where $\mathcal{C}$ is a convex compact subset of Θ such that θ_0 is in the interior of $\mathcal{C}$. This minimum distance estimator involves observations about RV, realized volatility or possibly a realized measure that corrects for microstructure effects etc. This estimator applies to volatility models involving all possible HYBRID processes, including structural ones for which a weak GARCH assumption is required. Note that this means that $V_{t|t-1}(\theta)$ in the above estimator is based on a best *linear predictor,* not the conditional variance - a technical issue that will be discussed in the next section.

A companion estimation procedure involves a single squared return process, namely:

$$\hat{\theta}_T^{mdr2} = \arg\min_{\theta \in \mathcal{C}} \frac{1}{T} \sum_{t=1}^{T} (R_t^2 - V_{t|t-1}(\theta))^2$$

The above estimator has a likelihood-based version, namely:

$$\hat{\theta}_T^{lhr2} = \arg\min_{\theta \in \mathcal{C}} \frac{1}{T} \sum_{t=1}^{T} \left(\log V_{t|t-1}(\theta) + \frac{R_t^2}{V_{t|t-1}(\theta)} \right)$$

requiring far more stringent in terms of regularity conditions, notably because $V_{t|t-1}(\theta)$ is a conditional variance, and in fact does not apply to all types of HYBRID processes - in particular structural ones. The estimator $\hat{\theta}_T^{mdr2}$ is reminiscent of QMLE estimators for semi-strong GARCH models - yet the mixed data frequencies add an extra layer of complexity discussed later in the paper. One can again replace daily squared returns by, say RV and consider the following estimator:

$$\hat{\theta}_T^{lhrv} = \arg\min_{\theta \in \mathcal{C}} \frac{1}{T} \sum_{t=1}^{T} \left(\log V_{t|t-1}(\theta) + \frac{RV_t}{V_{t|t-1}(\theta)} \right)$$

The choice of R^2 versus RV in $\hat{\theta}_T^{mdr2}$ versus $\hat{\theta}_T^{mdrv}$ and $\hat{\theta}_T^{lhr2}$ versus $\hat{\theta}_T^{lhrv}$ has efficiency implications that will be discussed as well.

Inspired by the Multiplicative Error Model (MEM) of [46] and the subsequent work by [51], [75], [35] also consider the following model

$$RV_{t+1} = \sigma_{t+1|t}^2 \eta_{t+1} \tag{1.11}$$

where conditional on $\mathcal{F}_t$, η_{t+1} is independent and identically distributed with mean 1. Suppose the cumulative distribution function of η is F. The choice of F could be a unit exponential (see [46]), or a Gamma distribution as suggested in [51], or a mixture of two gamma distributions of [75]. The resulting class of estimators will be denoted by $\tilde{\theta}_T^{mem}$.

[30] provide further detail regarding the theoretical properties of the various estimators and various HYBRID processes. They also conduct a Monte Carlo simulation study which shows that the estimator that appears to have the best finite sample properties is $\hat{\theta}_T^{lhrv}$. It is typically vastly better than the estimators based on R^2, either minimum distance or likelihood-based. It should also be noted that the MEM-type estimator - which is asymptotically equivalent to $\hat{\theta}_T^{lhrv}$ - is occasionally in small samples the most efficient for one parameter in particular, namely α. This means that the most efficient estimation of the unconditional mean of the volatility dynamic process can be achieved with the MEM principle which estimates directly the volatility process.

As far as empirical specification goes, the jury is still out. At the time this chapter was being written a thorough empirical investigation was still being conducted looking at the various types of HYBRID processes and their forecast performance at different horizons. [29] used the HYBRID GARCH class of models to predict volatility at daily horizons using intra-daily returns. The use of such returns forces one to think about how to treat intra-daily seasonality. [29] considered four approaches which we called: (1) (Unconstrained) HYBRID GARCH, (2) Periodic HYBRID GARCH, (3) (Unconstrained) SA HYBRID GARCH and (4) Constrained SA HYBRID GARCH. The former two apply to raw returns, the latter two to re-scaled returns using intra-daily unconditional volatility patterns. Overall they find that the use of seasonally adjusted returns is inferior both in-sample and out-of-sample. This means that we have essentially a relatively simple class of models that handle intra-day seasonality well.

1.3.3 GARCH-MIDAS Models

So far we did not cover component models of volatility. Empirical evidence suggests that volatility dynamics is better described by component models. [53] introduced a GARCH model with a long and short run component.[1] The volatility component model of Engle and Lee decomposed the equity conditional variance as the sum of the short-run (transitory) and long-run (trend) components.

So far we considered MIDAS filters that applied to high frequency data. Here we use the same type of filters to extract low frequency components. Hence, it is again a MIDAS setting, using different frequencies, but this time we use the polynomial specifications to extract low frequency movements in volatility.

In anticipation of the material in the next section, we consider multiple returns, although we study here still one single return series at the time. Namely, we consider a set of n assets and let the vector of returns be denoted as $\mathbf{r}_t = [r_{1,t}, \ldots, r_{n,t}]'$.

The new class of models is called GARCH-MIDAS, since it uses a mean reverting unit *daily* GARCH process, similar to [55], and a MIDAS polynomial which applies to *monthly, quarterly,* or *bi-annual* macroeconomic or financial variables. In what follows we will refer to g_i and m_i as the short and long run variance components respectively for asset i. [52] consider various specifications for g_i and we select only a specific one where the long run component is held constant across the days of the month, quarter or half-year. Alternatively, one can specify m_i based on rolling samples that change from day to day. The findings in [52] show that they yield very similar empirical fits - so we opted for the simplest to implement which involves locally constant long run components. We will denote by N_v^i the number of days that m_i is held fixed. The superscript i indicates that this may be asset-specific. The subscript v differentiates it from a similar scheme that will be introduced later for correlations. It will be convenient to introduce two time scales t and τ. In particular, while $g_{i,t}$ moves daily, $m_{i,\tau}$ changes only once every N_v^i days.

More specifically we assume that for each asset $i = 1, \ldots, n$, univariate returns follow the GARCH-MIDAS process:

$$r_{i,t} \quad = \quad \mu_i + \sqrt{m_{i,\tau} \cdot g_{i,t}}\xi_{i,t}, \quad \forall t = \tau N_v^i, \ldots, (\tau+1)N_v^i \tag{1.12}$$

where $g_{i,t}$ follows a GARCH(1,1) process:

$$g_{i,t} \quad = \quad (1 - \alpha_i - \beta_i) + \alpha_i \frac{(r_{i,t-1} - \mu_i)^2}{m_{i,\tau}} + \beta_i g_{i,t-1} \tag{1.13}$$

while the MIDAS component $m_{i,\tau}$ is a weighted sum of K_v^i lags of realized variances (RV) over a long horizon:

$$m_{i,\tau} \quad = \quad \overline{m}_i + \theta_i \sum_{l=1}^{K_v^i} \varphi_l(\omega_v^i) RV_{i,\tau-l} \tag{1.14}$$

[1] Several others have proposed related two-factor volatility models, see e.g. [43], [61], [**?**], [31] and [1] among many others.

where the realized variances involve N_v^i daily squared returns, namely:

$$RV_{i,\tau} = \sum_{j=(\tau-1)N_v^i+1}^{\tau N_v^i} (r_{i,j})^2.$$

Note that N_v^i could for example be a quarter or a month. The above specification corresponds to the block sampling scheme as defined in [52], involving so called Beta weights defined as:

$$\varphi_l(\omega_v^i) = \frac{(1-\frac{l}{K_v^i})^{\omega_v^i-1}}{\sum_{j=1}^{K_v^i}(1-\frac{j}{K_v^i})^{\omega_v^i-1}} \tag{1.15}$$

In practice we will consider cases where the parameters N_v^i and K_v^i are independent of i, i.e. the same across all series. Similarly, we can also allow for different decay patterns ω_v^i across various series, but once again we will focus on cases with common ω_v (see the next subsection for further discussion). Obviously, despite the common parameter specification, we expect that the $m_{i,\tau}$ substantially differ across series, as they are data-driven.

[52] study long historical data series of aggregate stock market volatility, starting in the 19th century, as in [83]. Their empirical findings show that for the full sample the long run component accounts for roughly 50 % of predicted volatility. During the Great Depression era even 60 % of expected volatility is due to the long run component. For the most recent period the results show roughly a 40 % contribution. Finally, they also introduce refinements of the GARCH-MIDAS model where the long run component is driven by macroeconomic series.

1.4 MULTIVARIATE MODELS

The estimation of multivariate volatility models with mixed sampling frequencies is a relatively unexplored area. In this final section we present one approach that appears promising. It was proposed by [38] and also applied by [15] to the determinants of stock and bond return co-movements.

[38] address the specification, estimation and interpretation of correlation models that distinguish short and long run components. They show that the changes in correlations are indeed very different. Dynamic correlations are a natural extension of the GARCH-MIDAS model to [49] DCC model. The idea captured by the DCC-MIDAS model is similar to that underlying GARCH-MIDAS. In the latter case, two components of volatility are extracted, one pertaining to short term fluctuations, the other pertaining to a secular component. In the GARCH-MIDAS the short run component is a GARCH component, based on daily (squared) returns, that moves around a long-run component driven by realized volatilities computed over a monthly, quarterly or bi-annual basis. The MIDAS weighting scheme helps to extract the slowly moving secular component around which daily volatility moves.

[52] explicitly link the extracted MIDAS component to macroeconomic sources. It is the same logic that is applied here to correlations. Namely, the daily dynamics obey a DCC scheme, with the correlations moving around a long run component. Short-lived effects to correlations will be captured by the autoregressive dynamic structure of DCC, with the intercept of the latter being a slowly moving process that reflects the fundamental or long-run causes of time variation in correlation.

To estimate the parameters of the DCC-MIDAS model [38] follow the two-step procedure of [49]. They start by estimating the parameters of the univariate conditional volatility models. The second step consists of estimating the DCC-MIDAS parameters with the standardized residuals. Moreover, they also discuss the regularity conditions we need to impose on the *MIDAS-filtered* long run correlation component as models of correlations are required to yield positive definite matrices.

Using the standardized residuals $\xi_{i,t}$ of the previous section it is possible to obtain a matrix $\boldsymbol{Q}_t$ whose elements are:

$$
\begin{aligned}
q_{i,j,t} &= \overline{\rho}_{i,j,t}\,(1-a-b) + a\xi_{i,t-1}\xi_{j,t-1} + bq_{i,j,t-1} \qquad (1.16)\\
\overline{\rho}_{i,j,t} &= \sum_{l=1}^{K_c^{ij}} \varphi_l\left(\omega_r^{ij}\right) c_{i,j,t-l}\\
c_{i,j,t} &= \frac{\sum_{k=t-N_c^{ij}}^{t} \xi_{i,k}\xi_{j,k}}{\sqrt{\sum_{k=t-N_c^{ij}}^{t} \xi_{i,k}^2}\sqrt{\sum_{k=t-N_c^{ij}}^{t} \xi_{j,k}^2}}
\end{aligned}
$$

where the weighting scheme is similar to that appearing in (1.14). Note that in the above formulation of $c_{i,j,t}$ one could have used simple cross-products of $\xi_{i,t}$. One can regard $q_{i,j,t}$ as the short run correlation between assets i and j, whereas $\overline{\rho}_{i,j,t}$ is a slowly moving long run correlation. Rewriting the first equation of system (1.16) as

$$
q_{i,j,t} - \overline{\rho}_{i,j,t} = a\left(\xi_{i,t-1}\xi_{j,t-1} - \overline{\rho}_{i,j,t}\right) + b\left(q_{i,j,t-1} - \overline{\rho}_{i,j,t}\right) \qquad (1.17)
$$

conveys the idea of short run fluctuations around a time varying long run relationship. The idea captured by the DCC-MIDAS model is similar to that underlying GARCH-MIDAS. In the latter case, two components of volatility are extracted, one pertaining to short term fluctuations, the other pertaining to a secular component. In the GARCH-MIDAS the short run component is a GARCH component, based on daily (squared) returns, that moves around a long-run component driven by realized volatilities computed over a monthly, quarterly or bi-annual basis. The MIDAS weighting scheme helps one to extract the slowly moving secular component around which daily volatility moves. It is the same logic that is applied here to correlations. Namely, the daily dynamics obey a DCC scheme, with the correlations moving around a long run component. Short-lived effects on correlations will be captured by the autoregressive dynamic structure of DCC, with the intercept of the latter being a slowly moving process that reflects the fundamental or secular causes of time variation in correlation.

In terms of empirical implementation [38] and [15] consider examples involving stocks and bonds. Both papers show the usefulness of the component specification in correlations and in particular the appeal of using MIDAS filters to specify long run component of correlations. Formal testing reported in both papers show that the DCC-MIDAS models outperform standard DCC models. [38] also study asset allocation with multiple international equities (five international stock markets) and a single MIDAS filter. Using the methodology proposed by [50] pertaining to minimum variance portfolio management they document the economic significance of using the DCC-MIDAS specification as well.

REFERENCES

1. T. Adrian and J. Rosenberg. Stock Returns and Volatility: Pricing the Short-Run and Long-Run Components of Market Risk. *Journal of Finance*, 63:2997–3030, 2008.
2. Y. Aït-Sahalia and L. Mancini. Out of sample forecasts of quadratic variation. *Journal of Econometrics*, 147:17–33, 2008.
3. S. Alizadeh, M. W. Brandt, and F. X. Diebold. Range-based estimation of stochastic volatility models. *Journal of Finance*, 57:1047–1091, 2002.
4. C.E. Alper, S. Fendoglu, and B. Saltoglu. Forecasting Stock Market Volatilities Using MIDAS Regressions: An Application to the Emerging Markets. Technical report, MPRA Paper No. 7460, 2008.
5. T. Andersen and T. Bollerslev. Answering the skeptics: Yes, standard volatility models do provide accurate forecasts. *International Economic Review*, 39:885–905, 1998.
6. T. Andersen and T. Bollerslev. Answering the skeptics: Yes, standard volatility models do provide accurate forecasts. *International Economic Review*, 39:885–905, 1998.
7. T. Andersen, T. Bollerslev, and F. X. Diebold. Parametric and nonparametric volatility measurement. In Yacine Aït-Sahalia and Lars P. Hansen, editors, *Handbook of Financial Econometrics*, 2010.
8. T. Andersen, T. Bollerslev, F. X. Diebold, and H. Ebens. The distribution of realized stock return volatility. *Journal of Financial Economies*, 61:43–76, 2001.
9. T. Andersen, T. Bollerslev, F. X. Diebold, and P. Labys. Modeling and forecasting realized volatility. *Econometrica*, 71:579–625, 2003.
10. T. Andersen, T. Bollerslev, and F.X. Diebold. Roughing it up: Including jump components in the measurement, modeling and forecasting of return volatility. *The Review of Economics and Statistics*, 89:701–720, 2007.
11. E. Andreou, E. Ghysels, and A. Kourtellos. Regression Models With Mixed Sampling Frequencies. *Journal of Econometrics*, 158:246–261, 2010.
12. E. Andreou, E. Ghysels, and A. Kourtellos. Should macroeconomic forecasters use daily financial data and how? Discussion Paper Universities of Cyprus and North Carolina, 2010.
13. V. Aragó and E. Salvador. Re-Examining the Risk-Return Relationship: The Influence of Financial Crisis (2007-2009). Technical report, Discussion Paper, Universitat Jaume I, Spain, 2010.
14. M. Armesto, K. Engemann, and M. Owyang. Forecasting with mixed frequencies. *Federal Reserve Bank of St. Louis Review*, (forthcoming), 2010.

15. L. Baele, G. Bekaert, and K. Inghelbrecht. The determinants of stock and bond return comovements. *Review of Financial Studies*, pages 2374–2428, 2010.

16. J. Bai, E. Ghysels, and J. Wright. State space models and midas regressions. Working paper, NY Fed, UNC and Johns Hopkins, 2009.

17. R. T. Baillie and R. P. DeGennaro. Stock returns and volatility. *Journal of Financial and Quantitative Analysis*, 25:203–214, 1990.

18. F.M. Bandi and J.R. Russell. Separating microstructure noise from volatility. *Journal of Financial Economics*, 79:655–692, 2006.

19. F.M. Bandi and J.R. Russell. Market microstructure noise, integrated variance estimators, and the accuracy of asymptotic approximations. *Journal of Econometrics* (forthcoming), 2010.

20. O. Barndorff-Nielsen, S. Graversen, and N. Shephard. Power variation and stochastic volatility: a review and some new results. *Journal of Applied Probability*, 41A:133–143, 2004.

21. O. Barndorff-Nielsen and N. Shephard. Power variation with stochastic volatility and jumps. Discussion paper, Nuffield College, 2003.

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

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

24. T. Bollerslev and J. M. Wooldridge. Quasi-maxmimum likelihood estimation and inference in dynamic models with time-varying covariances. *Econometric Reviews*, 11:143–172, 1992.

25. D. P. Brown and M. A. Ferreira. The Information in the Idiosyncratic Volatility of Small Firms. Working paper, Univesrity of Wisconsin and ISCTE, 2003.

26. J. Y. Campbell. Stock returns and the term structure. *Journal of Financial Economies*, 18:373–399, 1987.

27. J. Y. Campbell and L. Hentschel. No news is good news: An asymmetric model of changing volatility in stock returns. *Journal of Financial Economies*, 31:281–318, 1992.

28. X. Chen and E. Ghysels. News-Good or Bad-and its Impact on Volatility Forecasts over Multiple Horizons. *Review of Financial Studies*, forthcoming., 2010.

29. X. Chen, E. Ghysels, and F. Wang. On the role of Intra-Daily Seasonality in HYBRID GARCH Models. Technical report, *Journal of Time Series Econometrics*, forthcoming., 2009.

30. X. Chen, E. Ghysels, and F. Wang. The HYBRID GARCH class of models. Technical report, Working Paper, UNC, 2009.

31. M. Chernov, A. Ronald Gallant, E. Ghysels, and G. Tauchen. Alternative models for stock price dynamics. *Journal of Econometrics*, 116:225–257, 2003.

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

33. R. Chou. Volatility persistence and stock valuations: some empirical evidence using GARCH. *Journal of Applied Econometrics*, 3:279–294, 1992.

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

35. F. Cipollini, R.F. Engle, and G.M. Gallo. Vector multiplicative error models: representation and inference, 2006.

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

37. M.P. Clements, A.B. Galvão, and J.H. Kim. Quantile forecasts of daily exchange rate returns from forecasts of realized volatility. *Journal of Empirical Finance*, 15:729–750, 2008.

38. R. Colacito, R.F. Engle, and E. Ghysels. A component model for dynamic correlations. 2009.

39. F. Corsi. A simple approximate long-memory model of realized volatility. *Journal of Financial Econometrics*, 7:174–196, 2009.

40. R.G. de Vilder and M.P. Visser. Ranking and Combining Volatility Proxies for Garch and Stochastic Volatility Models. MPRA paper 11001, 2008.

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

42. Z. Ding, C. Granger, and R.F. Engle. A long memory property of stock market returns and a new models. *Journal of Empirical Finance*, 1:83–106, 1993.

43. Z. Ding and C.W.J. Granger. Modeling volatility persistence of speculative returns: A new approach. *Journal of Econometrics*, 73:185–215, 1996.

44. F.C. Drost and T.E. Nijman. Temporal aggregation of GARCH processes. *Econometrica*, pages 909–927, 1993.

45. F.C. Drost and B.J.M. Werker. Closing the GARCH gap: Continuous time GARCH modeling. *Journal of Econometrics*, 74:31–57, 1996.

46. R. Engle. New frontiers for ARCH models. *Journal of Applied Econometrics*, 17:425–446, 2002.

47. R. F. Engle. Discussion: Stock market volatility and the crash of '87. *Review of Financial Studies*, 3:103–106, 1990.

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

49. R.F. Engle. Dynamic conditional correlation - a simple class of multivariate garch models. *Journal of Business and Economic Statistics*, 20:339–350, 2002.

50. R.F. Engle and R. Colacito. Testing and valuing dynamic correlations for asset allocation. *Journal of Business and Economic Statistics*, pages 238–253, 2006.

51. R.F. Engle and G.M. Gallo. A multiple indicators model for volatility using intra-daily data. *Journal of Econometrics*, 131:3–27, 2006.

52. R.F. Engle, E. Ghysels, and B. Sohn. On the Economic Sources of Stock Market Volatility. Discussion Paper NYU and UNC, 2008.

53. R.F. Engle and G.G.J. Lee. A Permanent and Transitory Component Model of Stock Return Volatility. In *Cointegration, Causality, and Forecasting: A Festschrift in Honor of Clive W. J. Granger, Oxford University Press*, pages 475–497, 1999.

54. R.F. Engle, D.M. Lilien, and R.P. Robins. Estimating Time varying risk premia in the term structure: The ARCH-M model. *Econometrica: Journal of the Econometric Society*, 55:391–407, 1987.

55. R.F. Engle and J.G. Rangel. The spline-GARCH model for low-frequency volatility and its global macroeconomic causes. *Review of Financial Studies*, 21:1187–1222, 2008.

56. Robert F. Engle and Victor Ng. Measuring and testing the impact of news on volatility. *Journal of Finance*, 48:1749–1778, 1993.

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

58. D. Findley. Model selection for multi-step-ahead forecasting. In S. Ghosh and P. Young, editors, *Proceedings of the 7th Symposium on Identification and System Parameter Estimation*, pages 1039–1044. Pergamon, Oxford, 1985.

59. L. Forsberg and E. Ghysels. Why do absolute returns predict volatility so well? *Journal of Financial Econometrics*, 6:31–67, 2006.

60. K. R. French, W. Schwert, and R. F. Stambaugh. Expected stock returns and volatility. *Journal of Financial Economies*, 19:3–29, 1987.

61. A. Ronald Gallant, C.-T. Hsu, and G. Tauchen. Using daily range data to calibrate volatility diffusions and extract the forward integrated variance. *restat*, 81:617–631, 1999.

62. R. Garcia and N. Meddahi. Comment. *Journal of Business and Economic Statistics*, 24:184–192, 2006.

63. E. Ghysels, A. Rubia, and R. Valkanov. Multi-Period Forecasts of Volatility: Direct, Iterated, and Mixed-Data Approaches. Working Paper, UCSD and UNC, 2008.

64. E. Ghysels, P. Santa-Clara, and R. Valkanov. The MIDAS touch: Mixed data sampling regression models. Working paper, UNC and UCLA, 2002.

65. E. Ghysels, P. Santa-Clara, and R. Valkanov. There is a risk-return tradeoff after all. *Journal of Financial Economics*, 76:509–548, 2005.

66. E. Ghysels, P. Santa-Clara, and R. Valkanov. Predicting volatility: getting the most out of return data sampled at different frequencies. *Journal of Econometrics*, 131:59–95, 2006.

67. E. Ghysels and A. Sinko. Comment on realized variance and market microstructure noise by p. r. hansen and asger lunde. *Journal of Business and Economic Statistics*, 24:192–194, 2006.

68. E. Ghysels and A. Sinko. Volatility prediction and microstructure noise. *Journal of Econometrics* (forthcoming), 2006.

69. E. Ghysels, A. Sinko, and R. Valkanov. Midas regressions: Further results and new directions. *Econometric Reviews*, 26:53–90, 2006.

70. R. Giacomini and H. White. Tests of conditional predictive ability. *Econometrica*, 74:1545–1578, 2006.

71. L. R. Glosten, R. Jagannathan, and D. E. Runkle. On the relation between the expected value and the volatility of the nominal excess return on stocks. *Journal of Finance*, 48:1779–1801, 1993.

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

73. P. R. Hansen and A. Lunde. Realized variance and market microstructure noise. *Journal of Business and Economic Statistics*, 24:127–161, 2006.

74. C. R. Harvey. The specification of conditional expectations. *Journal of Empirical Finance*, 8:573–638, 2001.

75. M. Lanne. A mixture multiplicative error model for realized volatility. *Journal of Financial Econometrics*, 4:594, 2006.

76. Á. León, J.M. Nave, and G. Rubio. The relationship between risk and expected return in Europe. *Journal of Banking and Finance*, 31:495–512, 2007.

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

78. P. Malliavin and M.E. Mancino. A Fourier transform method for nonparametric estimation of multivariate volatility. *Annals of Statistics*, 37:1983–2010, 2005.

79. M. Marcellino, J. Stock, and M. Watson. A comparison of direct and iterated multistep ar methods for forecasting macroeconomic time series. *Journal of Econometrics*, 135:499–526, 2006.

80. R. C. Merton. An intertemporal capital asset pricing model. *Econometrica*, 41:867–887, 1973.

81. D. B. Nelson. Conditional heteroskedasticity in asset returns: A new approach. *Econometrica*, 59:347–370, 1991.

82. A. Patton. Volatility forecast comparison using imperfect volatility proxies. *Journal of Econometrics* (forthcoming), 2008.

83. G. W. Schwert. Why does stock market volatility change over time? *Journal of Finance*, 44:1207–1239, 1989.

84. N. Shephard and K. Sheppard. Realising the future: forecasting with high frequency based volatility (HEAVY) models. Journal of Applied Econometrics (forthcoming), 2009.

85. A. Sinko, M. Sockin, and E. Ghysels. Matlab toolbox for midas regressions. Available at `www.unc.edu/~eghysels/Software_datasets.html`, 2010.

86. C. M. Turner, R. Startz, and C. R. Nelson. A markov model of heteroskedasticity, risk, and learning in the stock market. *Journal of Financial Economies*, 25:3–22, 1989.

87. M.P. Visser. Garch parameter estimation using high-frequency data. *MPRA paper 9076*, 2008.

88. K. West. Asymptotic inference about predictive ability. *Econometrica*, 64:1067–1084, 1996.

89. K. West and D. Cho. The predictive ability of several models of exchange rate volatility. *Journal of Econometrics*, 69:367–391, 1995.

90. L. Zhang, P. A. Mykland, and Y. Aït-Sahalia. A tale of two time scales: Determining integrated volatility with noisy high frequency data. *Journal of the American Statistical Association*, 100:1394–1411, 2005.

91. B. Zhou. High-frequency data and volatility in foreign-exchange rates. *Journal of Business and Economic Statistics*, 14:45–52, 1996.