
VOLATILITY MODELS AND THEIR APPLICATIONS

Luc Bauwens, Christian Hafner, Sébastien Laurent

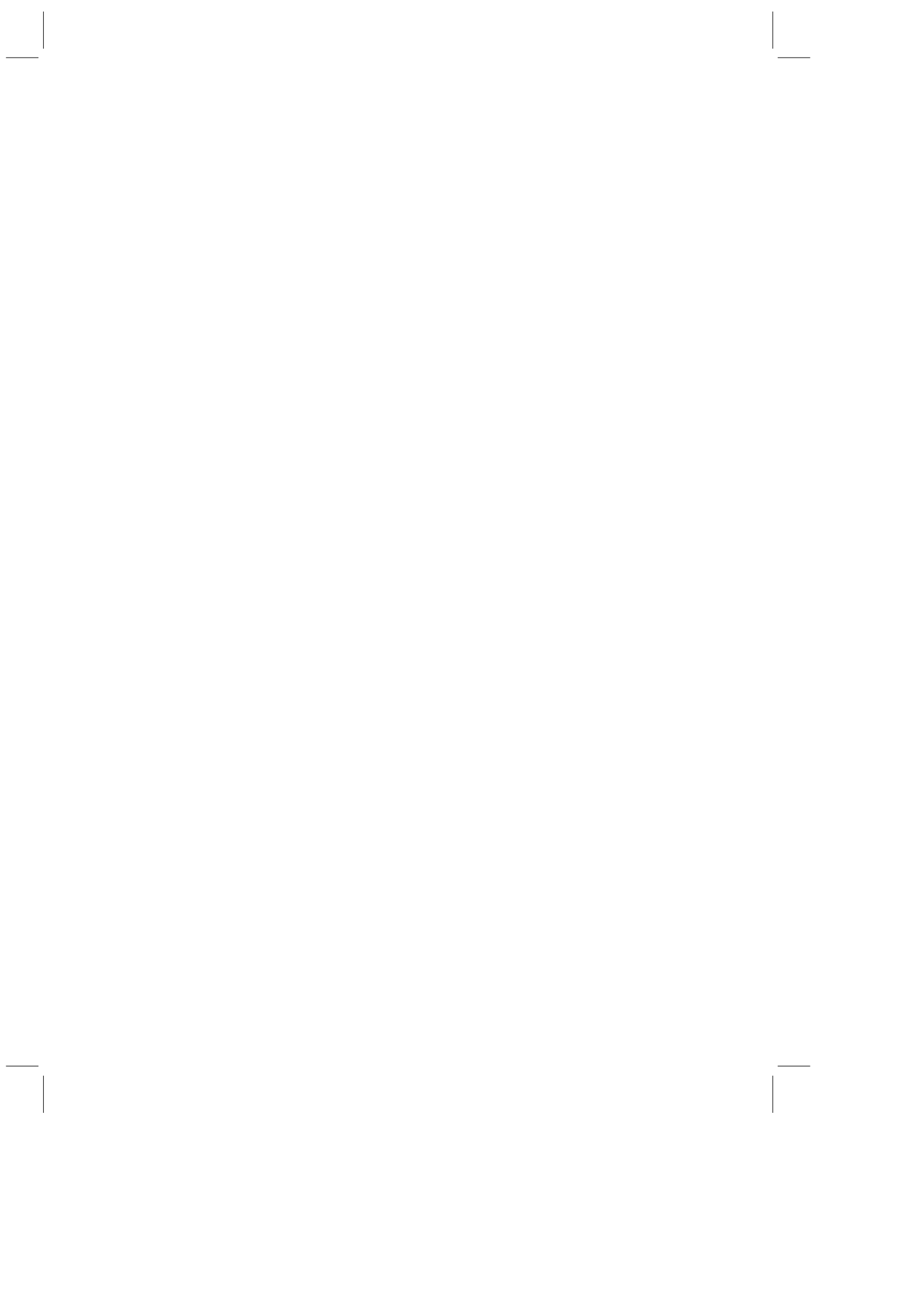
 **WILEY-
INTERSCIENCE**

A JOHN WILEY & SONS, INC., PUBLICATION



1 CONTENTS

2	1	Forecasting volatility with MIDAS	1
3	1.1	Introduction	1
4	1.2	Regressions	2
5	1.2.1	MIDAS Regressions	2
6	1.2.2	Direct versus Iterated Volatility Forecasting	4
7	1.2.3	Variations on the Theme of MIDAS Regressions	6
8	1.2.4	Microstructure Noise and MIDAS Regressions	8
9	1.3	Likelihood-based Methods	8
10	1.3.1	Risk-return Trade-off	9
11	1.3.2	HYBRID GARCH Models	11
12	1.3.3	GARCH-MIDAS Models	15
13	1.4	Multivariate Models	16
14	1.5	Concluding remarks	18
15		References	18



CHAPTER 1

FORECASTING VOLATILITY WITH MIDAS

1 Eric Ghysels (University of North Carolina), Rossen Valkanov (University of Cali-
2 fornia San Diego)

3

4 1.1 INTRODUCTION

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

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

1 Econometric methods involving data sampled at different frequencies have been
 2 considered in recent work by [64] in a likelihood-based setting and by [63], [65] and
 3 [10] using regression-based methods. The mixed frequency setting has been labeled
 4 MIDAS, meaning Mi(xed) Da(ta) S(ampling). The original work on MIDAS focused
 5 on volatility predictions, see e.g. [4], [12], [27], [51], [24], [29], [28], [36], [38], [58],
 6 [64], [65], [66], [67], [62], [76], among others.

7 [13] provide a user-friendly introduction to MIDAS regressions. A Matlab Tool-
 8 box for MIDAS regressions is also available, see [85]. A topic not covered, since we
 9 deal with volatility, but noteworthy is the fact that MIDAS regressions can be related
 10 to Kalman filters and state space models, see [15].

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

15 1.2 MIDAS REGRESSION MODELS AND VOLATILITY FORECASTING

16 In order to analyze the role of MIDAS in forecasting volatility let us introduce the
 17 relevant notation. Let $V_{t+1,t}$ be a measure of volatility in the next period. We focus on
 18 predicting future conditional variance, measured as increments in quadratic variation
 19 (or its log transformation), due to the large body of existing recent literature on this
 20 subject. The increments in the quadratic variation of the return process, $Q_{t+1,t}$,
 21 is not observed directly but can be measured with some discretization error. One
 22 such measure would be the sum of (future) m intra-daily squared returns, namely
 23 $\sum_{j=1}^m [r_{j,t}]^2$, which we will denote by $RV_{t+1,t}$. We can also consider multiple periods,
 24 which will be denoted by $RV_{t+h,t}$, for horizon h . Note that the case where no intra-
 25 daily data is available corresponds to $m = 1$ and RV becomes a daily squared return.

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

30 1.2.1 MIDAS Regressions

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

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

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

34 MIDAS regressions typically do *not* exploit an autoregressive scheme, so that
 35 X_{t-k} is not necessarily related to lags of the left hand side variable. Instead, MIDAS
 36 regressions are first and foremost regressions and therefore the selection of X_{t-k}
 37 amounts to choosing the best predictor of future quadratic variation from the set

1 of several possible measures of past fluctuations in returns. Examples of X_{t-k}
 2 are past daily squared returns (that correspond to the ARCH-type of models with
 3 some parameter restrictions, [47] and [22]), absolute daily returns (that relate to
 4 the specifications of (see e.g. [41]), realized daily volatility (e.g. [6]), realized daily
 5 power of (see [20] and [19]), and daily range (e.g. [3] and [60]). Since all of the
 6 regressors are used within a framework with the same number of parameters and the
 7 same maximum number of lags, the results from MIDAS regressions are directly
 8 comparable. Moreover, MIDAS regressions can also be extended to study the joint
 9 forecasting power of the regressors.

10 The weight function or the polynomial lag parameters are parameterized via
 11 Almon, Exponential Almon, Beta, linear step-functions (see below), etc., see [68],
 12 and they are especially relevant in estimating a persistent process parsimoniously,
 13 such as volatility, where distant X_{t-k} are likely to have an impact on current volatility.
 14 In addition, the parameterization allows us to compare MIDAS regressions at different
 15 frequencies as the number of parameters to estimate will be the same even though
 16 the weights on the data and the forecasting capabilities might differ across horizons.
 17 Most importantly one does not have to adjust measures of fit for the number of
 18 parameters and in most situations with one predictor one has a MIDAS model with
 19 either one or two parameters determining the pattern of the weights. Note also that
 20 in the above equation we specify a slope coefficient as the weights are normalized to
 21 add up to one. Such a restriction will not always be used in the sequel. The selection
 22 of k^{max} can be done conservatively (by taking a large value) and letting the weights
 23 die out as determined by the parameter estimation. The only cost to taking large
 24 k^{max} is the loss of initial data in the sample, which should be inconsequential in
 25 large samples.

26 Related to the MIDAS volatility regression is the Heterogeneous Autoregressive
 27 Realized Volatility (HAR-RV) regressions proposed by [38]. The HAR-RV model is
 28 given by:

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

29 which has a simple linear prediction regression using RV over heterogeneous interval
 30 sizes, daily (D), weekly (W) and monthly (M). As noted by [9] (footnote 16) and
 31 [38] (discussion on page 181) the above equation is in a sense a MIDAS regression
 32 with step-functions (in the terminology of [68]). In this regard the HAR-RV can be
 33 related to the MIDAS-RV in (1.1) of [65] and [58], using different weight functions
 34 such as the Beta, exponential Almon or step functions and different regressors not
 35 just autoregressive with mixed frequencies. Note also that both models exclude the
 36 jump component of quadratic variation. Simulation results reported in [58] also
 37 show that the difference between HAR and MIDAS models is very small for RV. For
 38 other regressors, such as the realized absolute variance, the MIDAS model performs
 39 slightly better.

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

1 pick only one day of lags:

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

2 The above equation is reminiscent of the ADL-MIDAS regression models used
 3 extensively in the context of macro forecasting by [11]. The above equation will also
 4 relate to the HYBRID GARCH class of models discussed later.

5 1.2.2 Direct versus Iterated Volatility Forecasting

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

15 The existent literature has placed most of the emphasis on the accuracy of one-
 16 period-ahead forecasts (see [47], [22], [5], [71]). Long-horizon volatility forecasts
 17 have received significantly less attention. Yet, financial decisions related to risk
 18 management, portfolio choice, and regulatory supervision, are often based on multi-
 19 period-ahead volatility forecasts. The preeminent long-horizon volatility forecasting
 20 approach is to scale the one-period-ahead forecasts by \sqrt{k} where k is the horizon
 21 of interest. [33] and others have shown that this “scaling” approach leads to poor
 22 volatility forecasts at horizons as short as ten days. The lack of a comprehensive and
 23 rigorous treatment of multi-period volatility forecasts is linked to the more general
 24 theoretical difficulty to characterize the trade-off between bias and estimation that
 25 exists in multi-period forecasts (see [56], [57], [77], [35], [21], and [31]). The paucity
 26 of new results on this topic has lead researchers to conclude that, in general, volatility
 27 is difficult to forecast at long horizons (see [33] and [89]).

28 In a recent paper, [62] undertake a comprehensive empirical examination of multi-
 29 period volatility forecasting approaches, beyond the simple \sqrt{k} -scaling rule. They
 30 consider two alternative approaches – direct and iterative – of forming long-horizon
 31 forecasts (see [79]). The “direct” forecasting method consists of estimating a horizon-
 32 specific model of the volatility at, say, monthly or quarterly frequency, which can
 33 then be used to form direct predictions of volatility over the next month or quarter. An
 34 “iterative” forecast obtains by estimating a daily autoregressive volatility forecasting
 35 model and then iterate over the daily forecasts for the necessary number of periods
 36 to obtain monthly, or quarterly predictions of the volatility. In addition to the direct
 37 and iterated approaches, [62] consider a third, novel way of long-horizon forecasts,
 38 which is based on MIDAS regressions. A MIDAS method uses daily data to produce
 39 directly multi-period volatility forecasts and can thus be viewed as a middle ground

1 between the direct and the iterated approaches. The results of their study suggest
2 that long-horizon volatility is much more predictable than previously suggested at
3 horizons as long as 60 trading days (about three months).

4 The direct and iterated methods [62] use are based on three volatility models:
5 GARCH (see [47] and [22]), autoregressive models of realized volatility ([7], [5], and
6 [8]), and integrated volatility. [62] point out that a long-horizon forecast is implicitly
7 a joint decision of choosing the appropriate volatility model and the appropriate
8 forecasting method. A similar distinction between a method and a model has also
9 been made implicitly by [8] and, in a different context, by [69]. The three volatility
10 models that [62] consider in conjunction with the iterated and direct forecasting
11 methods give rise to six different ways to produce long-horizon forecasts. The
12 MIDAS approach, which in essence combines the forecasting model and the long-
13 horizon method into one step, offers a seventh way of producing multi-period-ahead
14 forecasts of volatility.

15 To establish the accuracy of the seven long-term forecasts, [62] use a loss function
16 that penalizes deviations of predictions from the ex-post realizations of the volatility
17 (similar to [59] and [5]) and a test for predictive accuracy that allows them to
18 compare the statistical significance of competing forecasts. They use the mean
19 square forecasting error (MSFE) as one loss function, because of its consistency
20 property, i.e. it delivers the same forecast ranking with the proxy as it would with the
21 true volatility (see [73] and [82]). They use a Value-at-Risk (VAR) as an alternative
22 metric of forecast accuracy. To test the statistical significance in predictive power,
23 [62] use two tests. The first one, proposed by [88], takes into account parameter
24 uncertainty, which is of particular concern in the volatility forecasting literature.
25 The second test, proposed by [69], can be viewed as a generalization or a conditional
26 version of the [88] test. Rather than comparing the difference in average performance,
27 [69] consider the conditional expectation of the difference across forecasting models.
28 This conditioning approach allows not only for parameter uncertainty (as in [88])
29 but also uncertainty in a number of implicit choices made by the researcher when
30 formulating a forecast, such as what data to use, the windows of in-sample estimation
31 period, the length of the out-of-sample forecast, among others.

32 Using the above setup, [62] investigate whether multi-horizon forecasts of the
33 volatility of US stock market returns are more accurate than the naive but widely-
34 used scaling approach. They consider volatility forecasts of the US market portfolio
35 returns as well as of five size, five book-to-market, and ten industry portfolio returns.
36 They find that the scaling-up method performs poorly relative to the other methods
37 for all portfolios and horizons. This result is consistent with [40] and other papers
38 who have documented the poor performance of this approach. More surprisingly,
39 however, they find that the direct method does not fair much better. At horizon longer
40 than 10 days ahead, the approach of scaling one-period-ahead forecasts performs
41 significantly better than the direct method. Hence, if the direct method were the only
42 alternative to the scaling approach, and since scaling is a poor forecaster of future
43 volatility, one might come to the hasty conclusion that the volatility is hard to forecast
44 at long horizons by any model.

1 [62] find that for the volatility of the market portfolio, iterated and MIDAS fore-
 2 casts perform significantly better than the scaling and the direct approaches. At
 3 relatively short horizons of 5- to 10-days ahead, the iterated forecasts are quite ac-
 4 curate. However, at horizons of 10 days ahead and higher, MIDAS forecasts have
 5 a significantly lower MSFE relative to the other forecasts. At horizons of 30- and
 6 60-days ahead, the MSFE of MIDAS is more than 20 percent lower than that of the
 7 next best forecast. These differences are statistically significant at the one percent
 8 level according to the [88] and [69] tests. Hence, they find that suitable MIDAS mod-
 9 els produce multi-period volatility forecasts that are significantly better than other
 10 widely used methods.

11 [62] also link predictive accuracy to portfolio characteristics. They note that
 12 the superior performance of MIDAS in multi-period forecasts is also observed in
 13 predicting the volatility of the size, book-to-market, and industry portfolios. Similarly
 14 to the market volatility results, the relative precision of the MIDAS forecasts improves
 15 with the horizon. At horizons of 10-periods and higher, the MIDAS forecasts of
 16 eight out of the ten size and book-to-market portfolios dominate the iterated and
 17 direct approaches. At horizons of 30-periods and higher, the MIDAS has the smallest
 18 MSFEs amongst all forecasting methods for all ten portfolios. They observe that the
 19 volatility of the size and book-to-market portfolios is significantly less predictable
 20 than that of the entire market. Also, the predictability of the volatility increases
 21 with the size of the portfolio. The volatility of the largest-cap stocks is the most
 22 predictable, albeit still less forecastable than the market's. They fail to observe such
 23 a discernible pattern for the book-to-market portfolios.

24 From the MSFE results, it might be tempting to generalize that the MIDAS
 25 forecasts are more accurate than the iterated forecasts which in turn dominate the
 26 direct and scaling-rule approaches. However, [62] caution that a general ranking of
 27 forecast accuracy is difficult, since it is ultimately predicated on the loss function and
 28 application at hand. As an illustration, they note that when they use the VAR as a
 29 measure of forecast accuracy, then the direct method not only dominates the iterated
 30 method, but for most portfolio returns, its coverage is close to that of the MIDAS
 31 model. Overall, however, they find that MIDAS forecasts strike a good balance
 32 between bias and estimation efficiency.

33 1.2.3 Variations on the Theme of MIDAS Regressions

34 The MIDAS approach can also be used to study various other interesting aspects of
 35 forecasting volatility. [27] provide a novel method to analyze the impact of news on
 36 forecasting volatility. The following semi-parametric regression model is proposed
 37 to predict future realized volatility (RV) with past high-frequency returns:

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

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

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

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

9 Equation (1.4) nests (1.5) when we set $\psi_{i,j} \equiv \psi_i \forall j = 1, \dots, m$, and $NIC(r) \equiv$
 10 r^2 . This nesting emphasizes the role played by both the news impact curve NIC and
 11 the lag polynomial $\psi_{i,j}$.

12 The reason it is possible to nest the RV AR structure is due to the multiplicative
 13 specification for $\psi_{i,j}(\theta) \equiv \psi_j^D(\theta) \times \psi_i^{ID}(\theta)$, with the parameter θ containing sub-
 14 vectors that determine the two polynomials separately. The polynomial $\psi_j^D(\theta)$ is a
 15 daily weighting scheme, similar to $\psi_i(\theta)$ in the regression model appearing in (1.5).
 16 The polynomial $\psi_i^{ID}(\theta)$ relates to the intra-daily pattern. With equal intra-daily
 17 weights one has the RV measure when NIC is quadratic. [27] adopt the following
 18 specification for the polynomials:

$$\psi_j^D(\theta) \psi_i^{ID}(\theta) = Beta(j, \tau, \theta_1, \theta_2) \times Beta(i, 1/m, \theta_3, \theta_4), \quad (1.6)$$

19 where τ and $1/m$ are the daily (D) and intradaily (ID) frequencies. The restric-
 20 tion is imposed that the intra-daily patterns wash out across the entire day, i.e.
 21 $\sum_i Beta(i, 1/m, \theta_3, \theta_4) = 1$, and also impose without loss of generality, a similar
 22 restriction on the daily polynomial, in order to identify a slope coefficient in the
 23 regressions.

24 The multiplicative specification (1.6) has several advantages. First, as noted
 25 before, it nests the so-called flat aggregation scheme, i.e. all intra-daily weights are
 26 equal, yields a daily model with RV when the news impact curve is quadratic. Or
 27 more formally, when $\theta_3 = \theta_4 = 1$, and $NIC(r) = r^2$ one recovers RV -based regression
 28 appearing in equation (1.5). Second, by estimating $Beta(i, 1/m, \theta_3, \theta_4)$ one lets the
 29 data decide on the proper aggregation scheme which is a generic issue pertaining
 30 in MIDAS regressions as discussed in [10]. Obviously, the intra-daily part of the
 31 polynomial will pick up how news fades away throughout the day and this - in part -
 32 depends on the well known intra-daily seasonal pattern.

33 Finally, the MIDAS-NIC model can also nest existing parametric specifications of
 34 news impact curves adopted in the ARCH literature, namely, the daily symmetric one
 35 when $NIC(r) = br^2$, the asymmetric GJR model when $NIC(r) = br^2 + (cr^2)\mathbf{1}_{r < 0}$
 36 (see [70]) and the asymmetric GARCH model when $NIC(r) = (b(r-c)^2)$ (see [46]).

1.2.4 Microstructure Noise and MIDAS Regressions

[67] 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. [61] and [66] discussed forecasting volatility and microstructure noise. [68] provided further empirical evidence expanding on [66]. [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. [67] 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 [18]). They derive theoretical results for RV, TSRV, average over subsamples and [91] estimators and study theoretically optimal sampling as well.

[67] also conduct an extensive empirical study of forecasting with microstructure noise, using the same data as in [72], 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 [17]. 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 1.3.1 Risk-return Trade-off

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

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

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

9 [16], [59], [32], and [26] do find a positive albeit mostly insignificant relation
10 between the conditional variance and the conditional expected return. In contrast, [25]
11 and [81] find a significantly negative relation. [70], [74], and [86] find both a positive
12 and a negative relation depending on the method used. The main difficulty in testing
13 the ICAPM relation is that the conditional variance of the market is not observable
14 and must be filtered from past returns. The conflicting findings of the above studies
15 are mostly due to differences in the approach to modeling the conditional variance.

16 [64] take a different look at the risk-return tradeoff with a MIDAS forecast of the
17 monthly variance specified as a weighted average of lagged daily squared returns
18 and estimated via a QMLE - similar to the GARCH-in-mean approaches of [53] and
19 [70]. Namely, they estimate the coefficients of the conditional variance process jointly
20 with μ and γ from the expected return equation (1.7) with quasi-maximum likelihood.
21 Hence, this approach is very different from the MIDAS regressions discussed in the
22 previous section. The similarity, however, is that in both MIDAS regressions and in
23 the likelihood-based MIDAS one uses the same type of parsimoniously specified lag
24 polynomials. In particular, [64] use an exponential Almon lag specification.

25 Using monthly and daily market return data from 1928 to 2000 and, with a
26 MIDAS specification for the conditional variance, [64] find a positive and statistically
27 significant relation between risk and return. The estimate of γ is 2.6, which lines up
28 well with economic intuition about a reasonable level of risk aversion. The MIDAS
29 volatility estimator explains about 40 percent of the variation of realized variance
30 in the subsequent month and its explanatory power compares favorably to that of
31 other models of conditional variance such as GARCH. The estimated weights on
32 the lagged daily squared returns decay slowly, thus capturing the persistence in the
33 conditional variance process. More impressive still is the fact that, in the ICAPM
34 risk-return relation, the MIDAS estimator of conditional variance explains about two
35 percent of the variation of next month's stock market returns (and five percent in the
36 period since 1964). This is quite substantial given previous results about forecasting
37 the stock market return. Finally, the above results are qualitatively similar when one
38 splits the sample into two subsamples of approximately equal sizes, 1928-1963 and
39 1964-2000. These results are obtained when extreme outliers are winsorized.

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

1 [64] confirm the findings of [59] and [70], among others, of a positive but insignificant
2 γ coefficient in the risk-return tradeoff. The absence of statistical significance comes
3 both from GARCH's use of *monthly* returns in estimating the conditional variance
4 process. The use of daily data and the flexibility of the MIDAS estimator provides
5 the power needed to find statistical significance in the risk-return tradeoff.

6 A comparison of the time series of conditional variance estimated according to
7 MIDAS, GARCH, and rolling windows reveals that while the three estimators are
8 correlated, there are some differences that affect their ability to forecast returns in the
9 ICAPM relation. [64] find that the MIDAS variance process is more highly correlated
10 with both the GARCH and the rolling windows estimates than these last two are with
11 each other. This suggests that MIDAS combines some of the unique information
12 contained in the other two estimators. They also find that MIDAS is particularly
13 successful at forecasting realized variance both in high and low volatility regimes.
14 These features explain the superior performance of MIDAS in finding a positive and
15 significant risk-return relation.

16 It has long been recognized that volatility tends to react more to negative returns
17 than to positive returns. [81] and [55] show that GARCH models that incorporate
18 this asymmetry perform better in forecasting the market variance. However, [70]
19 show that when such asymmetric GARCH models are used in testing the risk-
20 return tradeoff, the γ coefficient is estimated to be negative (sometimes significantly
21 so). This stands in sharp contrast with the positive and insignificant γ obtained
22 with symmetric GARCH models and remains a puzzle in empirical finance. To
23 investigate this issue, [64] also extend the MIDAS approach to capture asymmetries
24 in the dynamics of conditional variance by allowing lagged positive and negative
25 daily squared returns to have different weights in the estimator. Contrary to the
26 asymmetric GARCH results, they still find a large positive estimate of γ that is
27 statistically significant. In particular, they find that what matters for the tests of the
28 risk-return tradeoff is not so much the asymmetry in the conditional variance process
29 but rather its persistence. In this respect, asymmetric GARCH and asymmetric
30 MIDAS models prove to be very different. Consistent with the GARCH literature,
31 negative shocks have a larger immediate impact on the MIDAS conditional variance
32 estimator than do positive shocks. However, [64] find that the impact of negative
33 returns on variance is only temporary and lasts no more than one month. Such
34 evidence is also consistent with the findings reported in [9] showing that (negative?)
35 jumps have no or little impact on future volatility. Positive returns, on the other
36 hand, have an extremely persistent impact on the variance process. In other words,
37 while short-term fluctuations in the conditional variance are mostly due to negative
38 shocks, the persistence of the variance process is primarily driven by positive shocks.
39 This is an intriguing finding about the dynamics of the variance process. Although
40 asymmetric GARCH models allow for a different response of the conditional variance
41 to positive and negative shocks, they constrain the persistence of both types of shocks
42 to be the same. Since the asymmetric GARCH models "load" heavily on negative
43 shocks and these have little persistence, the estimated conditional variance process
44 shows little to no persistence.

1 1.3.2 HYBRID GARCH Models

2 The volatility specification in [64] involves a single polynomial applied to daily
 3 data. Similar to the specification of the MIDAS regression (1.3) one could think
 4 of introducing lagged volatilities. We do not operate in a regression format, so this
 5 approach would be similar to the specification of a GARCH model.

6 This insight has recently been pursued by [28] and [29]. A key ingredient of
 7 conditional volatility models is that more weight is attached to the most recent
 8 returns (i.e. information). In the case of the original ARCH model (see e.g. [47])
 9 that means the most recent (daily) squared returns have more weight when predicting
 10 future (daily) conditional volatility. While intra-daily data are used to construct RV,
 11 prediction models put more weight on recent (daily) RV, but despite the use of intra-
 12 daily data - do not differentiate among intra-daily returns. If volatility is a persistent
 13 process, it would be natural to weight intra-daily data differently, as pointed out
 14 recently by [78]. This is one example of the class of models [29] so called HYBRID
 15 GARCH models. They are a unifying framework, based on a generic GARCH-type
 16 model, that addresses the issue of volatility forecasting involving forecast horizons of
 17 a different frequency than the information set. Hence, [29] propose a class of GARCH
 18 models that can handle volatility forecasts over the next five business days and use
 19 past daily data, or tomorrow's expected volatility while using intra-daily returns. The
 20 models are called HYBRID GARCH, which stands for **H**igh Frequent **Y** Data-**B**ased
 21 **P**roject**I**on-**D**riven GARCH models as the GARCH dynamics are driven by what
 22 [29] call HYBRID processes.

23 Compared to [78], they go beyond linear projections - albeit in a discrete time
 24 setting. The HYBRID GARCH models do have a connecting with continuous time
 25 models as well when one restricts attention to linear projections. [29] study three
 26 broad classes of HYBRID processes: (1) parameter-free processes that are purely
 27 data-driven, (2) structural HYBRIDs where one assumes some explicit data struc-
 28 ture for the high frequency data and finally (3) HYBRID filter processes. HYBRID
 29 GARCH models. In case (1) the HYBRID process H_τ does not depend on param-
 30 eters. The obvious case would be a simple return process such that $V_{\tau+1|\tau}$ is the
 31 conditional volatility of the next period. More recently, however, other purely data-
 32 driven examples of what we call generic HYBRID processes have been suggested.
 33 For example [50], [39], [87], [84] suggest the use of (daily) realized volatilities,
 34 high-low range or realized kernels or generic realized measures as they are called
 35 by [84]. Structural HYBRID processes appear in the context of temporal aggrega-
 36 tion - a topic discussed extensively in the (weak) GARCH literature, see e.g. [43],
 37 [44], among others. Finally, the HYBRID process $H_\tau \equiv H(\phi, \vec{r}_\tau)$, where $\vec{r}_\tau =$
 38 $(r_{\tau-1+1/m}, r_{\tau-1+2/m}, \dots, r_{\tau-1/m}, r_\tau)^T$ is \mathbf{R}^m -valued random vector, that can
 39 involve parameters that are *not* explicitly related to α , β and γ appearing in equation
 40 (1.8) below. There is no underlying high frequency data DGP that is being assumed,
 41 unlike in the structural HYBRID case. One can view this as a GARCH model driven
 42 by a filtered high frequency process - where the filter weights - (hyper-)parameterized
 43 by ϕ are estimated jointly with the volatility dynamics parameters.

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

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

2 where H_t will be called a HYBRID process. When H_t is simply a daily squared
 3 return we have the volatility dynamics of a standard daily GARCH(1,1), or H_t a
 4 weekly squared return those of a standard weekly GARCH(1,1). However, what
 5 would happen if we want to attribute an individual weight to each of the five days
 6 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
 7 we use the notation $r_{t-j/5}$ to indicate intra-period returns - in this case daily
 8 observations of week t (when days spill over into the previous week, we assume
 9 $r_{t-j/m} \equiv r_{t-1-(j-5)/m}$). This is an example of a parameter-driven HYBRID process
 10 $H_t \equiv H(\phi, \vec{r}_t)$ where $\vec{r}_t = (r_{t-1+1/m}, r_{t-1+2/m}, \dots, r_{t-1/m}, r_t)^T$ with in this case
 11 $m = 5$. In addition, the weights $(\omega_j(\phi), j = 0, \dots, m-1)$ are governed by a low-
 12 dimensional parameter vector ϕ . One can think of at least two possibilities: (1) the
 13 weights are treated as additional parameters and estimated as such (with m small this
 14 is possible, but not as m gets large), or (2) anchor the weights ω_j to an underlying
 15 daily GARCH(1,1) in which case the parameters α, β and γ and the weights in ϕ are
 16 jointly related to the assumed daily DGP. The discussion so far implicitly relates to
 17 many issues we elaborate on next.

18 The HYBRID process H_t may be purely data-driven and not depend on parameters.
 19 The obvious case would be a simple return process such that $V_{t+1|t}$ has the typical
 20 GARCH(1,1) dynamics. More recently, however, other purely data-driven examples
 21 of what we call generic HYBRID processes have been suggested. For example [50],
 22 [39], [87] and [84], among others, suggest the use of (daily) realized volatilities, high-
 23 low range or realized kernels or generic realized measures. It is important to note that
 24 typically parameter-free HYBRID processes do not differentiate intra-period returns,
 25 i.e. an equal weighting scheme is supposed - although some kernel-weighting or
 26 pre-averaging may take place to eliminate micro-structure noise.

27 To study structural HYBRIDs consider a daily weak GARCH(1,1), as defined by
 28 [43], then the implied weekly prediction, using past daily returns is:

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

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

36 Finally, consider a HYBRID filtering process. Here the HYBRID process $H(\phi, \vec{r}_t)$
 37 in (1.8) involves parameters that are *not* explicitly related to α, β and γ appearing
 38 in (1.8). There is no underlying high frequency data DGP that is being assumed,

1 unlike in the structural HYBRID case. One can view this as a GARCH model
 2 driven by a filtered high frequency process - where the filter weights - (hyper-
 3)parameterized by ϕ are estimated jointly with the volatility dynamics parameters.
 4 The choice of the parameterizations of the HYBRID process weighting scheme is
 5 inspired by [27]. The commonly used specifications are exponential, beta, linear,
 6 hyperbolic, and geometric weights. This approach has implications too as far as
 7 estimation is concerned. Unlike the structural HYBRID case, we now can consider
 8 likelihood-based methods, although the regularity conditions required are novel and
 9 more involved as those of the usual QMLE approach to GARCH estimation for
 10 instance in [23].

11 So far we have done the same as [78] in terms of the formulation of HYBRID
 12 processes in the context of discrete time GARCH dynamics. At this stage, we start
 13 to deviate from the linear projection paradigm and continue the logic of GARCH
 14 modeling. In light of these finding we consider HYBRID GARCH models that
 15 feature intra-daily news impact curves – similar to the framework of [27] – except
 16 that the latter use a MIDAS regression format. The HYBRID processes are of the
 17 following type:

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

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

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

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

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

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

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

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

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

4 requiring far more stringent in terms of regularity conditions, notably because
5 $V_{t|t-1}(\theta)$ is a conditional variance, and in fact does not apply to all types of HY-
6 BRID processes - in particular structural ones. The estimator $\hat{\theta}_T^{mdr2}$ is reminiscent of
7 QMLE estimators for semi-strong GARCH models - yet the mixed data frequencies
8 add an extra layer of complexity discussed later in the paper. One can again replace
9 daily squared returns by, say RV and consider the following estimator:

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

10 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
11 efficiency implications that will be discussed as well.

12 Inspired by the Multiplicative Error Model (MEM) of [45] and the subsequent
13 work by [50], [75], [34] also consider the following model

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

14 where conditional on \mathcal{F}_t , η_{t+1} is independent and identically distributed with mean
15 1. Suppose the cumulative distribution function of η is F . The choice of F could
16 be a unit exponential (see [45]), or a Gamma distribution as suggested in [50], or a
17 mixture of two gamma distributions of [75]. The resulting class of estimators will be
18 denoted by $\tilde{\theta}_T^{mem}$.

19 [29] provide further detail regarding the theoretical properties of the various esti-
20 mators and various HYBRID processes. They also conduct a Monte Carlo simulation
21 study which shows that the estimator that appears to have the best finite sample prop-
22 erties is $\hat{\theta}_T^{lhrv}$. It is typically vastly better than the estimators based on R^2 , either
23 minimum distance or likelihood-based. It should also be noted that the MEM-type
24 estimator - which is asymptotically equivalent to $\hat{\theta}_T^{lhrv}$ - is occasionally in small
25 samples the most efficient for one parameter in particular, namely α . This means that
26 the most efficient estimation of the unconditional mean of the volatility dynamic pro-
27 cess can be achieved with the MEM principle which estimates directly the volatility
28 process.

29 As far as empirical specification goes, the jury is still out. At the time this chap-
30 ter was being written a thorough empirical investigation was still being conducted
31 looking at the various types of HYBRID processes and their forecast performance at

1 different horizons. [28] used the HYBRID GARCH class of models to predict volatil-
 2 ity at daily horizons using intra-daily returns. The use of such returns forces one
 3 to think about how to treat intra-daily seasonality. [28] considered four approaches
 4 which we called: (1) (Unconstrained) HYBRID GARCH, (2) Periodic HYBRID
 5 GARCH, (3) (Unconstrained) SA HYBRID GARCH and (4) Constrained SA HY-
 6 BRID GARCH. The former two apply to raw returns, the latter two to re-scaled
 7 returns using intra-daily unconditional volatility patterns. Overall they find that
 8 the use of seasonally adjusted returns is inferior both in-sample and out-of-sample.
 9 This means that we have essentially a relatively simple class of models that handle
 10 intra-day seasonality well.

11 1.3.3 GARCH-MIDAS Models

12 So far we did not cover component models of volatility. Empirical evidence suggests
 13 that volatility dynamics is better described by component models. [52] introduced
 14 a GARCH model with a long and short run component.¹ The volatility component
 15 model of Engle and Lee decomposed the equity conditional variance as the sum of
 16 the short-run (transitory) and long-run (trend) components.

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

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

24 The new class of models is called GARCH-MIDAS, since it uses a mean reverting
 25 unit *daily* GARCH process, similar to [54], and a MIDAS polynomial which applies
 26 to *monthly*, *quarterly*, or *bi-annual* macroeconomic or financial variables. In what
 27 follows we will refer to g_i and m_i as the short and long run variance components
 28 respectively for asset i . [51] consider various specifications for g_i and we select only
 29 a specific one where the long run component is held constant across the days of
 30 the month, quarter or half-year. Alternatively, one can specify m_i based on rolling
 31 samples that change from day to day. The findings in [51] show that they yield very
 32 similar empirical fits - so we opted for the simplest to implement which involves
 33 locally constant long run components. We will denote by N_v^i the number of days
 34 that m_i is held fixed. The superscript i indicates that this may be asset-specific. The
 35 subscript v differentiates it from a similar scheme that will be introduced later for
 36 correlations. It will be convenient to introduce two time scales t and τ . In particular,
 37 while $g_{i,t}$ moves daily, $m_{i,\tau}$ changes only once every N_v^i days.

¹Several authors have proposed related two-factor volatility models, see e.g. [42], [60], [?], [30] and [1] among many others.

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

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

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

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

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

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

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

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

7 Note that N_v^i could for example be a quarter or a month. The above specification
 8 corresponds to the block sampling scheme as defined in [51], involving so called
 9 Beta weights defined as:

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

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

16 [51] study long historical data series of aggregate stock market volatility, starting
 17 in the 19th century, as in [83]. Their empirical findings show that for the full sample
 18 the long run component accounts for roughly 50 % of predicted volatility. During
 19 the Great Depression era even 60 % of expected volatility is due to the long run
 20 component. For the most recent period the results show roughly a 40 % contribution.
 21 Finally, they also introduce refinements of the GARCH-MIDAS model where the
 22 long run component is driven by macroeconomic series.

23 1.4 MULTIVARIATE MODELS

24 The estimation of multivariate volatility models with mixed sampling frequencies is a
 25 relatively unexplored area. In this final section we present one approach that appears

1 promising. It was proposed by [37] and also applied by [14] to the determinants of
2 stock and bond return co-movements.

3 [37] address the specification, estimation and interpretation of correlation models
4 that distinguish short and long run components. They show that the changes in
5 correlations are indeed very different. Dynamic correlations are a natural extension
6 of the GARCH-MIDAS model to [48] DCC model. The idea captured by the DCC-
7 MIDAS model is similar to that underlying GARCH-MIDAS. In the latter case,
8 two components of volatility are extracted, one pertaining to short term fluctuations,
9 the other pertaining to a secular component. In the GARCH-MIDAS the short
10 run component is a GARCH component, based on daily (squared) returns, that
11 moves around a long-run component driven by realized volatilities computed over
12 a monthly, quarterly or bi-annual basis. The MIDAS weighting scheme helps to
13 extract the slowly moving secular component around which daily volatility moves.
14 [51] explicitly link the extracted MIDAS component to macroeconomic sources. It
15 is the same logic that is applied here to correlations. Namely, the daily dynamics
16 obey a DCC scheme, with the correlations moving around a long run component.
17 Short-lived effects to correlations will be captured by the autoregressive dynamic
18 structure of DCC, with the intercept of the latter being a slowly moving process that
19 reflects the fundamental or long-run causes of time variation in correlation.

20 To estimate the parameters of the DCC-MIDAS model [37] follow the two-step
21 procedure of [48]. They start by estimating the parameters of the univariate condi-
22 tional volatility models. The second step consists of estimating the DCC-MIDAS
23 parameters with the standardized residuals. Moreover, they also discuss the regularity
24 conditions we need to impose on the *MIDAS-filtered* long run correlation component
25 as models of correlations are required to yield positive definite matrices.

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

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

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

$$q_{i,j,t} - \bar{\rho}_{i,j,t} = a(\xi_{i,t-1}\xi_{j,t-1} - \bar{\rho}_{i,j,t}) + b(q_{i,j,t-1} - \bar{\rho}_{i,j,t}), \quad (1.17)$$

32 conveys the idea of short run fluctuations around a time varying long run relation-
33 ship. The idea captured by the DCC-MIDAS model is similar to that underlying
34 GARCH-MIDAS. In the latter case, two components of volatility are extracted, one

1 pertaining to short term fluctuations, the other pertaining to a secular component.
2 In the GARCH-MIDAS the short run component is a GARCH component, based
3 on daily (squared) returns, that moves around a long-run component driven by real-
4 ized volatilities computed over a monthly, quarterly or bi-annual basis. The MIDAS
5 weighting scheme helps one to extract the slowly moving secular component around
6 which daily volatility moves. It is the same logic that is applied here to correlations.
7 Namely, the daily dynamics obey a DCC scheme, with the correlations moving
8 around a long run component. Short-lived effects on correlations will be captured
9 by the autoregressive dynamic structure of DCC, with the intercept of the latter be-
10 ing a slowly moving process that reflects the fundamental or secular causes of time
11 variation in correlation.

12 In terms of empirical implementation [37] and [14] consider examples involving
13 stocks and bonds. Both papers show the usefulness of the component specification
14 in correlations and in particular the appeal of using MIDAS filters to specify long
15 run component of correlations. Formal testing reported in both papers show that
16 the DCC-MIDAS models outperform standard DCC models. [37] also study asset
17 allocation with multiple international equities (five international stock markets) and a
18 single MIDAS filter. Using the methodology proposed by [49] pertaining to minimum
19 variance portfolio management they document the economic significance of using
20 the DCC-MIDAS specification as well.

21 1.5 CONCLUDING REMARKS

22 We focused on the issues pertaining to mixed frequencies - that arise typically be-
23 cause we would like to consider multi-step volatility forecasts while maintaining
24 information in high frequency data. We expect this topic to get more attention as
25 high frequency data becomes more widely available while the forecast horizons of
26 interest are mostly low frequency.

27 REFERENCES

- 28 1. T. Adrian and J. Rosenberg. Stock returns and volatility: Pricing the short-run and
29 long-run components of market risk. *Journal of Finance*, 63:2997–3030, 2008.
- 30 2. Y. Aït-Sahalia and L. Mancini. Out of sample forecasts of quadratic variation. *Journal of*
31 *Econometrics*, 147:17–33, 2008.
- 32 3. S. Alizadeh, M. W. Brandt, and F. X. Diebold. Range-based estimation of stochastic
33 volatility models. *Journal of Finance*, 57:1047–1091, 2002.
- 34 4. C.E. Alper, S. Fendoglu, and B. Saltoglu. Forecasting stock market volatilities using
35 MIDAS regressions: An application to the emerging markets. Technical report, MPRA
36 Paper No. 7460, 2008.
- 37 5. T. Andersen and T. Bollerslev. Answering the skeptics: Yes, standard volatility models
38 do provide accurate forecasts. *International Economic Review*, 39:885–905, 1998.

- 1 6. T. Andersen, T. Bollerslev, and F. X. Diebold. Parametric and nonparametric volatility
2 measurement. In Yacine Aït-Sahalia and Lars P. Hansen, editors, *Handbook of Financial*
3 *Econometrics*, 2010.
- 4 7. T. Andersen, T. Bollerslev, F. X. Diebold, and H. Ebens. The distribution of realized stock
5 return volatility. *Journal of Financial Economics*, 61:43–76, 2001.
- 6 8. T. Andersen, T. Bollerslev, F. X. Diebold, and P. Labys. Modeling and forecasting realized
7 volatility. *Econometrica*, 71:579–625, 2003.
- 8 9. T. Andersen, T. Bollerslev, and F.X. Diebold. Roughing it up: Including jump components
9 in the measurement, modeling and forecasting of return volatility. *Review of Economics*
10 *and Statistics*, 89:701–720, 2007.
- 11 10. E. Andreou, E. Ghysels, and A. Kourtellos. Regression models with mixed sampling
12 frequencies. *Journal of Econometrics*, 158:246–261, 2010.
- 13 11. E. Andreou, E. Ghysels, and A. Kourtellos. Should macroeconomic forecasters use daily
14 financial data and how? Discussion Paper Universities of Cyprus and North Carolina,
15 2010.
- 16 12. V. Aragón and E. Salvador. Re-examining the risk-return relationship: The influence of
17 financial crisis (2007-2009). Technical report, Discussion Paper, Universitat Jaume I,
18 Spain, 2010.
- 19 13. M. Armesto, K. Engemann, and M. Owyang. Forecasting with mixed frequencies.
20 *Federal Reserve Bank of St. Louis Review*, (forthcoming), 2010.
- 21 14. L. Baele, G. Bekaert, and K. Inghelbrecht. The determinants of stock and bond return
22 comovements. *Review of Financial Studies*, pages 2374–2428, 2010.
- 23 15. J. Bai, E. Ghysels, and J. Wright. State space models and midas regressions. Working
24 paper, NY Fed, UNC and Johns Hopkins, 2009.
- 25 16. R. T. Baillie and R. P. DeGennaro. Stock returns and volatility. *Journal of Financial and*
26 *Quantitative Analysis*, 25:203–214, 1990.
- 27 17. F.M. Bandi and J.R. Russell. Separating microstructure noise from volatility. *Journal of*
28 *Financial Economics*, 79:655–692, 2006.
- 29 18. F.M. Bandi and J.R. Russell. Market microstructure noise, integrated variance estimators,
30 and the accuracy of asymptotic approximations. *Journal of Econometrics* (forthcoming),
31 2010.
- 32 19. O. Barndorff-Nielsen, S. Graversen, and N. Shephard. Power variation and stochastic
33 volatility: a review and some new results. *Journal of Applied Probability*, 41A:133–143,
34 2004.
- 35 20. O. Barndorff-Nielsen and N. Shephard. Power variation with stochastic volatility and
36 jumps. Discussion paper, Nuffield College, 2003.
- 37 21. R.J. Bhanzali. Parameter estimation and model selection for multistep prediction of a time
38 series: A review. In S. Ghosh, editor, *Asymptotics, Nonparametrics, and Time Series*.
39 Marcel Dekker, 1999.
- 40 22. T. Bollerslev. Generalized autoregressive conditional heteroskedasticity. *Journal of*
41 *Econometrics*, 31:307–327, 1986.
- 42 23. T. Bollerslev and J. M. Wooldridge. Quasi-maximum likelihood estimation and infer-
43 ence in dynamic models with time-varying covariances. *Econometric Reviews*, 11:143–
44 172, 1992.

- 1 24. D. P. Brown and M. A. Ferreira. The information in the idiosyncratic volatility of small
2 firms. Working paper, Univesrity of Wisconsin and ISCTE, 2003.
- 3 25. J. Y. Campbell. Stock returns and the term structure. *Journal of Financial Economics*,
4 18:373–399, 1987.
- 5 26. J. Y. Campbell and L. Hentschel. No news is good news: An asymmetric model of
6 changing volatility in stock returns. *Journal of Financial Economics*, 31:281–318, 1992.
- 7 27. X. Chen and E. Ghysels. News-good or bad-and its impact on volatility forecasts over
8 multiple horizons. *Review of Financial Studies*, forthcoming., 2010.
- 9 28. X. Chen, E. Ghysels, and F. Wang. On the role of intra-daily seasonality in HYBRID
10 GARCH models. Technical report, *Journal of Time Series Econometrics*, forthcoming.,
11 2009.
- 12 29. X. Chen, E. Ghysels, and F. Wang. The HYBRID GARCH class of models. Technical
13 report, Working Paper, UNC, 2009.
- 14 30. M. Chernov, A. Ronald Gallant, E. Ghysels, and G. Tauchen. Alternative models for
15 stock price dynamics. *Journal of Econometrics*, 116:225–257, 2003.
- 16 31. C. Chevillon and D. Hendry. Non-parametric direct multi-step estimation for forecasting
17 economic processes. *International Journal of Forecasting*, 21:201–218, 2005.
- 18 32. R. Chou. Volatility persistence and stock valuations: some empirical evidence using
19 GARCH. *Journal of Applied Econometrics*, 3:279–294, 1992.
- 20 33. P. Christoffersen and F. Diebold. How relevant is volatility forecasting for financial risk
21 management? *Review of Economics and Statistics*, 82:12–22, 2000.
- 22 34. F. Cipollini, R.F. Engle, and G.M. Gallo. Vector multiplicative error models: representa-
23 tion and inference, 2006.
- 24 35. M. Clements and D. Hendry. Mutli-step estimation for forecasting. *Oxford Bulletin of*
25 *Economics and Statistics*, 58:657–684, 1996.
- 26 36. M.P. Clements, A.B. Galvão, and J.H. Kim. Quantile forecasts of daily exchange rate
27 returns from forecasts of realized volatility. *Journal of Empirical Finance*, 15:729–750,
28 2008.
- 29 37. R. Colacito, R.F. Engle, and E. Ghysels. A component model for dynamic correlations.
30 2011.
- 31 38. F. Corsi. A simple approximate long-memory model of realized volatility. *Journal of*
32 *Financial Econometrics*, 7:174–196, 2009.
- 33 39. R.G. de Vilder and M.P. Visser. Ranking and Combining Volatility Proxies for GARCH
34 and Stochastic Volatility Models. MPRA paper 11001, 2008.
- 35 40. F. Diebold, A. Hickman, A. Inoue, and T. Schuermann. Converting 1-day volatility to h-
36 day volatility: Scaling by \sqrt{h} is worse than you think. mimeo, University of Pennsylvania,
37 1997.
- 38 41. Z. Ding, C. Granger, and R.F. Engle. A long memory property of stock market returns
39 and a new models. *Journal of Empirical Finance*, 1:83–106, 1993.
- 40 42. Z. Ding and C.W.J. Granger. Modeling volatility persistence of speculative returns: A
41 new approach. *Journal of Econometrics*, 73:185–215, 1996.
- 42 43. F.C. Drost and T.E. Nijman. Temporal aggregation of GARCH processes. *Econometrica*,
43 pages 909–927, 1993.

- 1 44. F.C. Drost and B.J.M. Werker. Closing the GARCH gap: Continuous time GARCH
2 modeling. *Journal of Econometrics*, 74:31–57, 1996.
- 3 45. R. Engle. New frontiers for ARCH models. *Journal of Applied Econometrics*, 17:425–
4 446, 2002.
- 5 46. R. F. Engle. Discussion: Stock market volatility and the crash of '87. *Review of Financial*
6 *Studies*, 3:103–106, 1990.
- 7 47. R.F. Engle. Autoregressive conditional heteroscedasticity with estimates of the variance
8 of united kingdom inflation. *Econometrica*, 50:987–1008, 1982.
- 9 48. R.F. Engle. Dynamic conditional correlation - a simple class of multivariate garch models.
10 *Journal of Business and Economic Statistics*, 20:339–350, 2002.
- 11 49. R.F. Engle and R. Colacito. Testing and valuing dynamic correlations for asset allocation.
12 *Journal of Business and Economic Statistics*, pages 238–253, 2006.
- 13 50. R.F. Engle and G.M. Gallo. A multiple indicators model for volatility using intra-daily
14 data. *Journal of Econometrics*, 131:3–27, 2006.
- 15 51. R.F. Engle, E. Ghysels, and B. Sohn. On the economic sources of stock market volatility.
16 Discussion Paper NYU and UNC, 2008.
- 17 52. R.F. Engle and G.G.J. Lee. A permanent and transitory component model of stock return
18 volatility. In *Cointegration, Causality, and Forecasting: A Festschrift in Honor of Clive*
19 *W. J. Granger*, Oxford University Press, pages 475–497, 1999.
- 20 53. R.F. Engle, D.M. Lilien, and R.P. Robins. Estimating Time varying risk premia in the
21 term structure: The ARCH-M model. *Econometrica: Journal of the Econometric Society*,
22 55:391–407, 1987.
- 23 54. R.F. Engle and J.G. Rangel. The spline-GARCH model for low-frequency volatility and
24 its global macroeconomic causes. *Review of Financial Studies*, 21:1187–1222, 2008.
- 25 55. Robert F. Engle and Victor Ng. Measuring and testing the impact of news on volatility.
26 *Journal of Finance*, 48:1749–1778, 1993.
- 27 56. D. Findley. On the use of multiple models of multi-period forecasting. *Proceedings of the*
28 *Business and Statistics Section, American Statistical Association*, pages 528–531, 1983.
- 29 57. D. Findley. Model selection for multi-step-ahead forecasting. In S. Ghosh and P. Young,
30 editors, *Proceedings of the 7th Symposium on Identification and System Parameter Esti-*
31 *mation*, pages 1039–1044. Pergamon, Oxford, 1985.
- 32 58. L. Forsberg and E. Ghysels. Why do absolute returns predict volatility so well? *Journal*
33 *of Financial Econometrics*, 6:31–67, 2006.
- 34 59. K. R. French, W. Schwert, and R. F. Stambaugh. Expected stock returns and volatility.
35 *Journal of Financial Economics*, 19:3–29, 1987.
- 36 60. A. Ronald Gallant, C.-T. Hsu, and G. Tauchen. Using daily range data to calibrate
37 volatility diffusions and extract the forward integrated variance. *restat*, 81:617–631,
38 1999.
- 39 61. R. Garcia and N. Meddahi. Comment. *Journal of Business and Economic Statistics*,
40 24:184–192, 2006.
- 41 62. E. Ghysels, A. Rubia, and R. Valkanov. Multi-period forecasts of volatility: Direct,
42 iterated, and mixed-data approaches. Working Paper, UCSD and UNC, 2008.

- 1 63. E. Ghysels, P. Santa-Clara, and R. Valkanov. The MIDAS touch: Mixed data sampling
2 regression models. Working paper, UNC and UCLA, 2002.
- 3 64. E. Ghysels, P. Santa-Clara, and R. Valkanov. There is a risk-return tradeoff after all.
4 *Journal of Financial Economics*, 76:509–548, 2005.
- 5 65. E. Ghysels, P. Santa-Clara, and R. Valkanov. Predicting volatility: getting the most out of
6 return data sampled at different frequencies. *Journal of Econometrics*, 131:59–95, 2006.
- 7 66. E. Ghysels and A. Sinko. Comment on Realized variance and market microstructure
8 noise by Peter Hansen and Asger Lunde. *Journal of Business and Economic Statistics*,
9 24:192–194, 2006.
- 10 67. E. Ghysels and A. Sinko. Volatility forecasting and microstructure noise. *Journal of*
11 *Econometrics*, 160:257–271, 2011.
- 12 68. E. Ghysels, A. Sinko, and R. Valkanov. MIDAS regressions: Further results and new
13 directions. *Econometric Reviews*, 26:53–90, 2006.
- 14 69. R. Giacomini and H. White. Tests of conditional predictive ability. *Econometrica*,
15 74:1545–1578, 2006.
- 16 70. L. R. Glosten, R. Jagannathan, and D. E. Runkle. On the relation between the expected
17 value and the volatility of the nominal excess return on stocks. *Journal of Finance*,
18 48:1779–1801, 1993.
- 19 71. P. Hansen and A. Lunde. A Forecast Comparison of Volatility Models: Does Anything
20 Beat a GARCH(1,1)? *Journal of Applied Econometrics*, 20:873–889, 2005.
- 21 72. P. R. Hansen and A. Lunde. Realized variance and market microstructure noise. *Journal*
22 *of Business and Economic Statistics*, 24:127–161, 2006.
- 23 73. P.R. Hansen and A. Lunde. Consistent ranking of volatility models. *Journal of Econo-*
24 *metrics*, 131(1-2):97–121, 2006.
- 25 74. C. R. Harvey. The specification of conditional expectations. *Journal of Empirical*
26 *Finance*, 8:573–638, 2001.
- 27 75. M. Lanne. A mixture multiplicative error model for realized volatility. *Journal of*
28 *Financial Econometrics*, 4:594, 2006.
- 29 76. Á. León, J.M. Nave, and G. Rubio. The relationship between risk and expected return in
30 europe. *Journal of Banking and Finance*, 31:495–512, 2007.
- 31 77. J. Lin and C.W.J. Granger. Model selection for multiperiod forecasts. *Journal of Fore-*
32 *casting*, 13:1–9, 1994.
- 33 78. P. Malliavin and M.E. Mancino. A fourier transform method for nonparametric estimation
34 of multivariate volatility. *Annals of Statistics*, 37:1983–2010, 2005.
- 35 79. M. Marcellino, J. Stock, and M. Watson. A comparison of direct and iterated multistep ar
36 methods for forecasting macroeconomic time series. *Journal of Econometrics*, 135:499–
37 526, 2006.
- 38 80. R. C. Merton. An intertemporal capital asset pricing model. *Econometrica*, 41:867–887,
39 1973.
- 40 81. D. B. Nelson. Conditional heteroskedasticity in asset returns: A new approach. *Econo-*
41 *metrica*, 59:347–370, 1991.
- 42 82. A. Patton. Volatility forecast comparison using imperfect volatility proxies. *Journal of*
43 *Econometrics*, 160:246–256, 2011.

- 1 83. G. W. Schwert. Why does stock market volatility change over time? *Journal of Finance*,
2 44:1207–1239, 1989.
- 3 84. N. Shephard and K. Sheppard. Realising the future: forecasting with high frequency based
4 volatility (HEAVY) models. *Journal of Applied Econometrics* (forthcoming), 2009.
- 5 85. A. Sinko, M. Sockin, and E. Ghysels. Matlab toolbox for midas regressions. Available at
6 www.unc.edu/~eghysels/Software_datasets.html, 2010.
- 7 86. C. M. Turner, R. Startz, and C. R. Nelson. A markov model of heteroskedasticity, risk,
8 and learning in the stock market. *Journal of Financial Economies*, 25:3–22, 1989.
- 9 87. M.P. Visser. GARCH parameter estimation using high-frequency data. *MPRA paper*
10 9076, 2008.
- 11 88. K. West. Asymptotic inference about predictive ability. *Econometrica*, 64:1067–1084,
12 1996.
- 13 89. K. West and D. Cho. The predictive ability of several models of exchange rate volatility.
14 *Journal of Econometrics*, 69:367–391, 1995.
- 15 90. L. Zhang, P. A. Mykland, and Y. Aït-Sahalia. A tale of two time scales: Determining
16 integrated volatility with noisy high frequency data. *Journal of the American Statistical*
17 *Association*, 100:1394–1411, 2005.
- 18 91. B. Zhou. High-frequency data and volatility in foreign-exchange rates. *Journal of*
19 *Business and Economic Statistics*, 14:45–52, 1996.