HYBRID-GARCH

A Generic Class of Models for Volatility Predictions using Mixed Frequency Data

Xilong Chen† Eric Ghysels‡ Fangfang Wang§

First Draft: April 2009

This Draft: April 20, 2011

*We like to thank Rob Engle, Per Mykland, Eric Renault, Neil Shephard and George Tauchen for insightful comments. An early version of the paper was presented at the Stevanovich Center - CREATES conference Financial Econometrics and Statistics: Current Themes and New Directions, Skagen, Denmark 4-6 June 2009 under the title High Frequency GARCH Models. We also like to thank participants for comments at the Oxford-Mann Institute and the following conferences: 2010 FERM in Taipei, 2010 NBER-NSF Time Series conference, Duke University, 2010 New Researchers in Statistics and Probability in Vancouver, 2010 SoFiE conference in Melbourne, and the 2010 Quantitative Methods in Business Applications in Beijing.

†ETS, SAS Institute Inc., Cary, NC 27513. Email: Xilong.Chen@sas.com
‡Department of Finance, Kenan-Flagler Business School, and Department of Economics, University of North Carolina at Chapel Hill. Email: eghysels@unc.edu
§Department of Information and Decision Sciences, University of Illinois at Chicago. Email: ffwang@uic.edu
Abstract

We propose a general GARCH framework that allows the predict volatility using returns sampled at a higher frequency than the prediction horizon. We call the class of models High Frequency Data-Based Projection-Driven GARCH, or HYBRID-GARCH models, as the volatility dynamics are driven by what we call HYBRID processes. The HYBRID processes can involve data sampled at any frequency. As far as empirical specifications go, we obtain some powerful findings that deviate substantially from the existing literature. Models featuring intra-daily asymmetries (news impact curves applied to intra-daily returns) dominate symmetric models up to weekly horizons, while the reverse is true for longer horizons. Models using daily realized volatility are less preferred than HYBRID involving intra-daily weighting scheme even for longer horizons.

Keywords: HYBRID process, weak GARCH, GARCH jump diffusion, realized measure, temporal aggregation, filtering
1 Introduction

Multi-period volatility forecasts feature prominently in asset pricing, portfolio allocation, risk-management and most other areas of finance where long-horizon measures of risk are necessary. Such forecasts can be constructed in three fundamentally different ways. The first approach is to estimate a horizon-specific model of the volatility, such as a weekly or monthly GARCH which can then be used to form direct predictions of volatility over the next week, month, etc. The second approach is to estimate a daily model and then iterate forward the daily forecasts to obtain weekly or monthly predictions. The forecasting literature refers to the first approach as “direct” and the second as “iterated”. A third method is the mixed-data sampling (MIDAS) approach introduced by Ghysels, Santa-Clara, and Valkanov ((2005), (2006)). A MIDAS model uses for example daily squared returns to produce directly multi-period volatility forecasts and can be viewed as a middle ground between the direct and the iterated approaches. The MIDAS volatility literature (see Ghysels and Valkanov (2011)) has mostly focused on regressions-based models. It is the purpose of this paper to introduce ideas similar to MIDAS models in GARCH-type models. The advantages of this approach is that one focuses directly on multi-period forecasts, like in the direct approach, while one preserves the use of high frequency data.

We propose 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, we propose a class of GARCH models that can handle volatility forecasts over the next five business days and use past (intra-)daily data, or tomorrow’s expected volatility while using intra-daily returns. We call the class of models **High Frequency Data-Based Projection-Driven GARCH models** as the GARCH dynamics are driven by what we call HYBRID processes. HYBRID-
GARCH models - by their very nature - relate to many topics discussed in the large literature on volatility forecasting. These topics include - but are not limited to - iterated versus direct forecasting (as already noted), temporal aggregation, weak versus semi-strong GARCH, as well as various estimation procedures. Since there are quite a few papers written on these topics already it is hard to cite a comprehensive list here. Nevertheless, it is worth noting that we study three broad classes of HYBRID processes: (1) parameter-free processes that are purely data-driven, (2) structural HYBRIDs where one assumes an underlying data generating process (DGP) or some dynamic structure for the high frequency data and finally (3) HYBRID filter processes.

To motivate the class of models, it is worth recalling that 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. Engle (1982)) that means the most recent (daily) squared returns have more weight when predicting future (daily) conditional volatility. How does this apply to high-frequency - that is intra-daily - financial data? The foundation of so called realized volatility (RV) modeling is the theory of continuous time semi-martingale stochastic processes, more specifically stochastic volatility continuous time jump-diffusions. 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 Malliavin and Mancino (2005).1 The arguments also apply to lower frequency volatility prediction models, such as (total) weekly volatility. Here the choice is between a GARCH model - using past weekly returns, de facto putting equal weight to the daily returns within the week - and a GARCH model for weekly volatility.

1The paper was brought to our attention by George Tauchen after we wrote a first draft of our paper and presented it at the FERM meeting in Taipei, June 2010.
forecasts using daily returns. It is the latter that is novel and an example of the class of models we introduce in the paper. Compared to Malliavin and Mancino (2005), we go beyond linear projections - albeit in a discrete time setting. Our models do have a connection with continuous time models as well when we restrict our attention to linear projections.

As far as empirical specifications go, we obtain some powerful findings that deviate substantially from the existing literature. GARCH and TGARCH models with daily data are dominated at all forecast horizons by models involving intra-daily models. Models featuring intra-daily asymmetries (news impact curves applied to intra-daily returns) dominate symmetric models up to weekly horizons, while the reverse is true for longer horizons, a finding also reported in Chen and Ghysels (2011). Models using daily realized volatility (or semi-variance versions) are less preferred than HYBRID involving intra-daily weighting scheme even for longer horizons.

The paper is structured as follows. Section 2 provides an overview of the models, and the various classes of HYBRID processes involved. Section 3 is devoted to the statistical properties of the HYBRID GARCH model, and defining the various parameter estimators and studying both their asymptotic and small sample properties (via simulation). Section 4 covers the empirical model specifications, with empirical findings appearing in Section 5. Section 6 concludes the paper. Some technical details are collected in an appendix, whereas all the proofs appear in a companion document Chen, Ghysels and Wang (2011b).

2 HYBRID Processes

This section provides an overview of HYBRID GARCH models, and the various classes of HYBRID processes involved. We will start with some notation. Suppose the
underlying probability space is $(\Omega, \mathcal{F}, P)$. Let $\|X\|_p = (E|X|^p)^{1/p}$ for $X \in L^p(\Omega, \mathcal{F}, P)$ and $p < \infty$. $\|A\| = \sqrt{\text{tr}(A^T A)}$ for $A \in \mathbb{R}^{n \times n}$ or $A \in \mathbb{R}^{n \times 1}$ and $n \geq 1$. For $X \in L^2(\Omega, \mathcal{F}, P)$ and $I$ a closed subspace of $L^2(\Omega, \mathcal{F}, P)$, $P_I(X|I)$ indicates the orthogonal projection of $X$ onto $I$. We write $A > 0$ if $A$ is a positive definite matrix, and $A \geq 0$ if $A$ is positive semi-definite. If $A$ is finite element by element, then we write $A < \infty$.

Moreover, to emphasize the role of $\phi$ in the process $H(\phi, \vec{r}_t)$ we will use the notation $H_t(\phi) \equiv H(\phi, \vec{r}_t)$. The notation related to derivatives is shown in Appendix A.

### 2.1 HYBRID GARCH models: Overview

The volatility dynamics of a generic HYBRID GARCH model is as follows:

$$V_{t+1|t} = \alpha + \beta V_{t|t-1} + \gamma H_t$$  \hspace{1cm} (2.1)

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 - the daily observations of week $t$ in this case.\(^2\) This is an example of a parameter-driven HYBRID process $H_t \equiv H(\phi, \vec{r}_t)$ where $\vec{r}_t = (r_{t-1+1/m}, r_{t-1+2/m}, \ldots, r_{t-1/m}, r_t)^T$ is $\mathbb{R}^m-$valued random vector (in this case $m = 5$). In addition, the weights $(\omega_j(\phi), j = 0, \ldots, m - 1)$ are governed by a low-dimensional parameter vector $\phi$. 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 $\omega_j$ to an underlying

\(^2\)When days spill over into the previous week, we assume $r_{t-j/m} \equiv r_{t-1-(j-5)/m}$. 

daily GARCH(1,1) in which case the parameters $\alpha$, $\beta$ and $\gamma$ and the weights in $\phi$ are jointly related to the assumed daily DGP.

This example illustrates the HYBRID processes that we will define/study in this paper: parameter-free HYBRID, structural HYBRID, and HYBRID filtering processes. The parameter-free HYBRID process has the simplest structure among the three. It is purely data-driven and not depend on parameters. The parameter-free HYBRID includes, but not limited to, squared return process, (daily) realized volatilities, high-low range or realized kernels or generic realized measures (see Engle and Gallo (2006), de Vilder and Visser (2008), Visser (2011), Shephard and Sheppard (2010), Hansen, Huang, and Shek (2010)). It is important to note that parameter-free HYBRID processes do not differentiate intra-period returns, although some kernel-weighting or pre-averaging may take place to eliminate so-called microstructure noise. The $H_t$ considered in (1) is the HYBRID filtering process, while $H_t$ in (2) is structural HYBRID. We will discuss these two specifications in detail next.

2.2 Three processes of interest

Suppose the underlying (log) price process $p_s$ is a semimartingale defined on $(\Omega, \mathcal{F}, P)$. $r_s$ is the return sampled at frequency $m$, i.e., $r_s = p_s - p_{s-1/m}$. We are interested in the next-period volatility forecast based on the available discretely-sampled returns, denoted by $\sigma^2_{t+1|t}$. It is defined as the orthogonal projection of $RV_{t+1} = \sum_{j=0}^{m-1} r^2_{t+1-j/m}$ onto $\mathcal{I}_t$, which is a closed subspace of $L^2(\Omega, \mathcal{F}, P)$ and represents the information up to time $t$. In other words, $\sigma^2_{t+1|t} = P_t(RV_{t+1}|\mathcal{I}_t)$. We therefore implicitly assume that $RV_t \in L^2(\Omega, \mathcal{F}, P)$, or the return has finite fourth moment.

Denote by $[p, p]_s$ the quadratic variation of $\{p_s\}$. The predicted increment in quadratic variation is expressed as $E_t([p, p]_{t+1} - [p, p]_t) \equiv E([p, p]_{t+1} - [p, p]_t)\sigma(p_s, s \leq$
t)). We make a distinction between three objects: (1) \( V_{t+1|t} \), (2) \( \sigma^2_{t+1|t} \), and (3) \( E_t([p, p]_{t+1} - [p, p]_t) \). The latter two are population quantities, while (1) pertains to the specification of the HYBRID GARCH model. Because the first two are formulated using the available (high frequency) returns, they are not necessarily linked with an explicit continuous-time/discrete-time DGP.

Various model specifications can be considered for the HYBRID GARCH, but this is not our concern at this point. Instead, we are interested in the relation between \( \sigma^2_{t+1|t} \) and \( V_{t+1|t} \). Moreover, we will also examine how predicted increments in quadratic variation relate to the HYBRID GARCH model-based predictions \( V_{t+1|t} \). These relationships can only be well understood when (1) we impose a structure on the underlying returns and (2) we explicitly link the HYBRID GARCH to the DGP or at least some dynamic structure for high frequency returns.

At the outset it should also be noted that \( V_{t+1|t} \) inherits the properties of the HYBRID process \( H_t \) and vice versa. Hence, we will interchangeably talk about features of HYBRID process \( H_t \) and features of \( V_{t+1|t} \).

### 2.3 HYBRID Filtering Processes

Suppose that the available returns are sampled at frequency \( m \), and they are expressed as \( r_s \) where \( s \) is of the form \( t + k/m \) for some \( t \in \mathbb{Z} \) and \( k = 0, 1, 2, \ldots, m - 1 \). In general the structure of \( \sigma^2_{t+1|t} \) is not tractable due to the ignorance of \( r_s \) and \( \mathcal{I}_t \). Therefore to justify the approximation of \( \sigma^2_{t+1|t} \) with \( V_{t+1|t} \), we consider two specifications for \( \sigma^2_{t+1|t} \), and call them Scenarios 1 and 2 appearing in respectively Assumptions 2.1 and 2.2. The former views \( \sigma^2_{t+1|t} \) as the best linear predictor. Alternatively, we also consider a more general situation in Scenario 2, where \( \sigma^2_{t+1|t} \) is a conditional variance. Suppose the return process \( \{r_s\} \) satisfy Assumption A.1.\(^3\) Denote by \( \mathcal{F}^{t+m}_{t-m} \) the sigma field generated

\(^3\)For convenience we collected all the regularity conditions in Appendix A.
by the high frequency returns from \( t-m \) to \( t+m \), i.e., \( \sigma(r_s, t-m-1+1/m \leq s \leq t+m) \).

**Assumption 2.1** (Scenario 1). \( \mathcal{I}_t = \mathcal{L}_t \), the closed span of \( \{1, r_{t-k/m}, r_{t-k/m}^2; k = 0,1,2\ldots\} \) and \( P_t(r_s|\mathcal{L}_{s-1/m}) = 0 \). Therefore \( \sigma^2_{t+1|t} \) is the best linear predictor.

**Assumption 2.2** (Scenario 2). \( \mathcal{I}_t = \mathcal{F}^t_{-\infty} \), the sigma field generated by the high frequency returns up to time \( t \) and \( P_t(r_s|\mathcal{F}^t_{s-1/m}) = 0 \). The prediction equations therefore indicate that \( E(r_s|\mathcal{F}^{s-1/m}_{-\infty}) = 0 \) and \( E(RV_{t+1}|\mathcal{F}^t_{-\infty}) = E(R_{t+1}^2|\mathcal{F}^t_{-\infty}) = \sigma^2_{t+1|t} \), where \( R_{t+1} = \sum_{j=0}^{m-1} r_{t+1-j}/m \).

Since we use \( V_{t+1|t} \) driven by \( H(\phi, \vec{r}_t) \) to mimic the dynamics of \( \sigma^2_{t+1|t} \), \( H_t \) is required to satisfy Assumption A.4 in both scenarios. Particularly in Scenario 1, \( H_t \) is also assumed to be a weighted sum of 1, the intermediate returns and squared returns from period \( t-1 \) to \( t \). Assumption A.4 essentially guarantees that the HYBRID process is non-negative and satisfies measurability and identifiability when it comes to parameter estimation via of extremum estimators. A necessary condition is that the dimension of \( \phi \) is not larger than \( m \). A more detailed discussion on Assumption A.4 is available in Chen, Ghysels and Wang (2011b). The HYBRID processes that satisfy Assumption A.4 are also referred to as HYBRID filtering processes.

In both scenarios we can hyper-parameterize the filter weights, namely:

\[
H(\phi, \vec{r}_t) = \sum_{j=0}^{m-1} \Psi_j(\phi) r_{t-j/m}^2, \quad \sum_{j=0}^{m-1} \Psi_j(\phi) = 1
\]

where the weights \( (\Psi_0(\phi), \Psi_1(\phi), \Psi_2(\phi), \ldots, \Psi_{m-1}(\phi))^T \) are determined by a low-dimensional functional specification used by Chen and Ghysels (2011) which were inspired by MIDAS regression format of Ghysels, Santa-Clara, and Valkanov (2006), Ghysels, Sinko, and Valkanov (2006), Ghysels, Rubia, and Valkanov (2009). The commonly used specifications are exponential, beta, linear, hyperbolic, and geometric.
weights. Note that the weighting schemes can handle intra-daily seasonal patterns - a topic discussed in further detail by Chen, Ghysels and Wang (2011a).

Deviate from the linear projection paradigm as in (2.2), the HYBRID filtering structure allows us to consider HYBRID GARCH models that feature intra-daily news impact curves - similar to the framework of Chen and Ghysels (2011), except that the latter use a MIDAS regression format. The HYBRID processes we consider are of the following type:

$$H(\phi, \tilde{r}_t) = \sum_{j=0}^{m-1} \Psi_j(\phi) NIC(\phi, r_{t-j/m}), \quad \sum_{j=0}^{m-1} \Psi_j(\phi) = 1$$ (2.3)

where $$NIC(\phi, \cdot)$$ stands for a high frequency data news impact curve. The parameter vector $$\phi$$ determines the weights as well as the parameters that determine the news impact curve. Regarding the specification of the latter, we consider the parametric news impact curves studied in Chen and Ghysels (2011), namely:

$$NIC(\phi, r) = \gamma(r - \delta)^2$$ (2.4)

$$NIC(\phi, r) = \gamma r^2 + \delta r^2 1_{r<0}$$ (2.5)

with $$\gamma$$ and $$\delta$$ the parameters that are in the parameter vector $$\phi$$. We should note that the HYBRID process with $$NIC(\phi, \cdot)$$ given by (2.4) is specified under Scenario 1, while the HYBRID process constructed using news impact curve (2.5) is a special case of Scenario 2. To ensure the HYBRID process (2.3) meets Assumption A.4, the weight functions need to satisfy some additional conditions that appear in Assumption A.6.
2.4 Structural HYBRID Processes

When the underlying high frequency returns \( \{r_s\} \) follow a weak GARCH(1,1) of Drost and Nijman (1993), the dynamics of \( \sigma^2_{t+1|t} \) can be fully specified. In other words, with proper parameterization, the HYBRID GARCH process \( V_{t+1|t} \) and \( \sigma^2_{t+1|t} \) coincide. The HYBRID processes are therefore called structural HYBRID processes.

Suppose that \( r_{s+1/m} \) is orthogonal to \( L_s \), i.e., \( P_l(r_{s+1/m}|L_s) = 0 \), and \( \sigma^2_{s+1/m|s} \) defined as the orthogonal projection of \( r^2_{s+1/m} \) onto \( L_s \) satisfies

\[
\sigma^2_{s+1/m|s} = a + b\sigma^2_{s|s-1/m} + cr^2_s.
\]

With some algebra one obtains \( \sigma^2_{s+k/m|s} = a \left(1 - (b + c)^{k-1}\right)/(1 - (b + c)) + (b + c)^{k-1}\sigma^2_{s+1/m|s} \) for \( k \in \mathbb{Z}^+ \). Consequently, the total volatility over the period \( (t, t+1] \), denoted by \( V_{t+1|t} \equiv \sum_{k=1}^{m} \sigma^2_{t+k/m|t} \equiv \sigma^2_{t+1|t} \), can be characterized by the following GARCH-type of equation:

\[
V_{t+1|t} = \alpha_m + \beta_m V_{t|t-1} + \gamma_m \sum_{j=0}^{m-1} \beta^j_m r^2_{t-j/m}
\]

where

\[
\alpha_m = a - \frac{b^m m(1-b) - cd_m}{1-b}, \quad \beta_m = b^m, \quad \gamma_m = c d_m
\]

and \( d_m = (1 - (b + c)^m)/(1 - (b + c)) \). Clearly, (2.7) is of the form (2.1) with \( H_t = \sum_{j=0}^{m-1} \beta^j_m r^2_{t-j/m} \), and \( H_t \) is referred to as structural HYBRID process. The distinct difference between structural HYBRID and HYBRID filtering processes is that the underlying return process follows a weak GARCH(1,1) for the structural HYBRID. The HYBRID filtering process can be viewed as an extension of the structural HYBRID process by allowing for more flexible return dynamics.
It is worth noting that the structural HYBRID model allows the parameters evaluated under different sampling frequencies to be linked to each other explicitly, as is evident from (2.8). A direct implication of the relationship (2.8) is that one can use parameter estimates from say a daily model with for example 5-min returns, to formulate a weekly or lower frequency model with the same 5-min returns.

### 2.5 Diffusions, Jumps and HYBRID Processes

The HYBRID processes discussed so far pertained to discretely sampled returns at different frequencies. We turn our attention now to HYBRID processes structurally linked to continuous time processes. In addition, we also examine the possible structural HYBRID interpretation of the purely RV-driven HYBRID GARCH process, i.e., $H(\phi, \tilde{r}_t) = RV_t$. Moreover, we will characterize how the presence of jumps will have an impact on the discrete-time HYBRID process.

Inspired by Drost and Werker (1996), we consider a continuous-time GARCH model as the DGP, namely,

$$
\begin{align*}
dp_t &= \sigma_t dL_t, \quad L_t = \sqrt{1-\eta}W_t + \sqrt{\eta}N_t \\
\sigma_t^2 &= \theta(\omega - \sigma_t^2)dt + \sqrt{2\lambda\theta\sigma_t^2}dB_t
\end{align*}
$$

(2.9)

with $\theta > 0$, $\omega > 0$, $\lambda \in (0, 1)$, and $\eta \in [0, 1]$. $B_t$ and $W_t$ are standard Brownian motions. $N_t$ is a compound Poisson process with jump measure $J_N$ and Lévy measure $\nu(dy) = \zeta f(dy)$ where $f$ is the Normal density with mean 0 and variance $1/\zeta$. Moreover, $B_t$, $W_t$ and $N_t$ are independent of each other. Note that $EL_t = EL_t^3 = 0$, $EL_t^2 = t$, and $EL_t^4 = 3t^2 + 3t\eta^2/\zeta$. The discretely sampled returns $r_s = p_s - p_{s-1}/m$ from (2.9) follows a weak GARCH(1,1) appearing in equation (2.6) as discussed in Drost and Werker (1996). The relation between $(\theta, \omega, \lambda, v_L^*)$ in (2.9) and $(a, b, c, k)$ in (2.6) is stated in
Drost and Werker (1996), where \( v^*_L \equiv EL_1^4 - 3 = 3\eta^2/\zeta \) and \( k \) is the kurtosis of \( r_s \).

We turn our attention now to some prediction formulas associated with this framework. Note that the Quadratic Variation (QV) of \( p \) is \([p, p]_t = (1 - \eta) \int_0^t \sigma^2_s ds + \eta \int_0^t \int_{-\infty}^{\infty} \sigma^2_s y^2 J_N(ds, dy) \). We start with examining how \( V_{t+1|t} = \sigma^2_{t+1|t} \) relates to prediction of \( E_t([p, p]_{t+1} - [p, p]_t) \). In this subsection, we write \( V_{t+1|t} \) as \( V_m(t) \) to emphasize the role of sampling frequency \( m \).

On the one hand, using a continuous time filtration, the forecast of the increment of QV is as follows:

\[
E_t([p, p]_{t+1} - [p, p]_t) = (1 - \eta) \int_t^{t+1} E_t(\sigma^2_s) ds + \eta \int_t^{t+1} \int_{-\infty}^{\infty} E_t(\sigma^2_s) y^2 \zeta f(dy) ds \\
= \omega (1 - \theta^{-1}(1 - e^{-\theta})) + \theta^{-1}(1 - e^{-\theta}) \sigma^2_t. \tag{2.10}
\]

On the other hand, the forecast using \( \mathcal{L}_t \) yields the HYBRID GARCH equation appearing in equation (2.7). What we will show is that, although \( RV_{t+1} = \sum_{j=0}^{m-1} r_{t+1-j/m}^2 \) is a consistent estimator of \([p, p]_{t+1} - [p, p]_t\), the HYBRID GARCH process \( V_{t+1|t}^{(m)} \) may not consistently estimate \( E_t([p, p]_{t+1} - [p, p]_t) \). Using the relation between \((\theta, \omega, \lambda, v^*_L)\) and \((a, b, c, k)\) stated in Drost and Werker (1996), we have the following proposition:

**Proposition 2.1.** When the Lévy measure associated with the jump process features excess kurtosis, i.e. \( v^*_L > 0 \),

\[
\lim_{m \to \infty} \alpha_m = \omega (1 - e^{-\theta(1+\phi)}) \left(1 - \frac{\phi}{1 + \phi}\theta^{-1}(1 - e^{-\theta})\right) \\
\lim_{m \to \infty} \beta_m = e^{-\theta(1+\phi)}, \quad \lim_{m \to \infty} \gamma_m = (1 - e^{-\theta})\phi \\
\lim_{m \to \infty} \sum_{j=0}^{m-1} \beta_{m}^{\frac{j}{m}} r_{t-j/m}^2 = \int_{[t-1,t]} e^{-\theta(1+\phi)(t-s)} d[p, p]_s \quad \text{in probability} \tag{2.11}
\]

\(^4\)Note that \((a, b, c, k)\) and \( r_s \) depend on \( m \).
where $\phi = \sqrt{1 + 2\lambda/(\theta v^*_L)} - 1$. In contrast, when there are no jumps in the price process, i.e. $v^*_L = 0$, we have $\lim_{m\to\infty} \alpha_m = \omega \left(1 - \theta^{-1}(1 - e^{-\theta})\right)$, $\lim_{m\to\infty} \beta_m = 0$, and $\lim_{m\to\infty} \frac{\gamma_m}{\sqrt{m}} = \sqrt{\frac{\lambda}{\theta}}(1 - e^{-\theta})$. Moreover $\sqrt{m} \sum_{j=0}^{m-1} \beta_m^{j/m} \frac{\gamma_j}{r_{t-j/m}}$ converges to $(\theta \lambda)^{-1/2} \sigma_t^2$ in $L^2$.

The proof - as well as all subsequent ones - are omitted here; they appear in Chen, Ghysels and Wang (2011b). Comparing equations (2.7) and (2.10), we note from Proposition 2.1 that $V^{(m)}_{t+1|t}$ is a consistent estimate of $E_t([p, p]_{t+1} - [p, p]_t)$ when there are no jumps in the price process. In contrast, when jumps are present $V^{(m)}_{t+1|t}$ does not provide a consistent estimate of $E_t([p, p]_{t+1} - [p, p]_t)$ because the limit of $V^{(m)}_{t+1|t}$ involves a whole sample path of volatility up to time $t$. This is stated formally in the following corollary:

**Corollary 2.1.** Given a continuous time GARCH (2.9) as the DGP, the process $\{V^{(m)}_{t+1|t}, t\}_{m\geq 1}$ defined by equation (2.7) converges to $\{E_t([p, p]_{t+1} - [p, p]_t), t\}$ uniformly on compact sets in probability if and only if there are no jumps in the price process.

Note that without jumps, the HYBRID GARCH process still involves intra-period weighted returns, namely equation (2.7) has intra-period weights that are powers of $\beta_m$. Furthermore, it follows from the proof of Proposition 2.1 that what drives the HYBRID process as $m \to \infty$, is the instantaneous volatility $\sigma_t^2$, not the integrated process estimated by $RV$. The instantaneous volatility $\sigma_t^2$ can be consistently estimated by that very same intra-period weighted sum $mc \sum_{j=0}^{m-1} b_j^2 r_{t-j/m}^2$. Put differently, we can view the HYBRID process as a spot volatility estimator which shares some features with other data-driven spot volatility estimators considered by Foster and Nelson (1996), Zhang (2001), Andreou and Ghysels (2002), Fan, Fan, and Jiang (2007), Fan and Wang (2008), Mykland and Zhang (2008), Zhao and Wu (2008), Malliavin and Mancino (2005), among others.
To conclude it should also be noted that one could think of continuous time limits in the HYBRID filtering context and potentially link them to $E_t([p, p]_{t+1} - [p, p]_t)$.

In the above discussions we relied on the approach of Drost and Werker (1996) using exact discretization limits - which is compatible with structural HYBRID processes. We leave the broader question of diffusion limits - as in Nelson (1992), Nelson and Foster (1995), among others - and HYBRID filtering processes for future research.

3 Estimation

We study the statistical properties of the HYBRID GARCH model-based parameter estimation in this section. The structural HYBRID can be viewed as a special case of the HYBRID filtering processes. Therefore, we will focus on the latter and the results derived for the HYBRID filtering processes will carry over to the structural HYBRID accordingly. We will work exclusively with returns sampled at fixed frequency without referring to an explicit DGP and make the assumption that the returns are strictly stationary and ergodic (formally stated as Assumption A.2 in Appendix A).

Throughout this section we will assume there are no model specification errors (the analysis of potentially misspecified models is covered in Chen, Ghysels and Wang (2011b)).

3.1 Estimation under Scenario 1

We use $V_{t|t-1}(\theta)$ solved from equation (2.1) to approximate $\sigma^2_{t|t-1}$. The distance between the two time series $V(\theta) \equiv (V_{t|t-1}(\theta), t \in \mathbb{N})$ and $\sigma^2 \equiv \{\sigma^2_{t|t-1}, t \in \mathbb{N}\}$ is defined as $d(V, \sigma^2) = \sum_{t=1}^{\infty} 2^{-t} \min(\|V_{t|t-1} - \sigma^2_{t|t-1}\|_2, 1)$. Note that $\|RV_t - V_{t|t-1}(\theta)\|_2 - \|RV_t - \sigma^2_{t|t-1}\|_2 = \|V_{t|t-1}(\theta) - \sigma^2_{t|t-1}\|_2$ due to Assumption A.1. $0 \leq d(RV, V(\theta)) - d(RV, \sigma^2) \leq d(V(\theta), \sigma^2)$ where $RV = (RV_t, t \in \mathbb{N})$. $d(V(\theta), \sigma^2)$ hence measures
ignorance of the true dynamics. \( d(V(\theta), \sigma^2) \) (for a suitable choice of \( \theta \)) is 0 when equation (2.1) correctly describes the dynamics of \( \sigma^2_{t|t-1} \).

Given \( H \) that satisfies Assumption A.4, \( d(V(\theta), \sigma^2) \) has a minimum over \( C \), say at \( \theta_0 = (\alpha_0, \beta_0, \gamma_0, \phi_0) \in C \), where \( C \) is a convex compact subset of the interior of \( \Theta \equiv \{ (\alpha, \beta, \gamma, \phi) : \alpha > 0, 0 < \beta < 1, \gamma > 0, \phi \in \Phi \} \) and \( \Phi \) is a connected set which collects all the possible values of \( \phi \)'s such that \( H_t(\phi) \) meets Assumption A.4. Assume \( d(V(\theta_0), \sigma^2) = 0 \). A natural estimator of \( \theta_0 \) under Scenario 1, given \( \{ r_{1/m}, \ldots, r_1, \ldots, r_{T-1+1/m}, \ldots, r_T \} \), is the minimizer of \( \min_{\theta \in C} T^{-1} \sum_{t=1}^{T} (RV_t - \tilde{V}_t(\theta))^2 \) and \( \tilde{V}_t \) is defined recursively by

\[
\tilde{V}_t(\theta) = \alpha + \beta \tilde{V}_{t-1}(\theta) + \gamma H_{t-1}(\phi), \ t \geq 1 \quad \text{and} \quad \tilde{V}_0 = \tilde{v} \tag{3.1}
\]

where \( \tilde{v} \) is any arbitrary deterministic value. The minimizer exists due to Jennrich (1969) and Gallant and White (1988) and is denoted by \( \tilde{\theta}_T^{mdrv} \), minimum-distance RV-based estimator. Note that \( \theta_0 \) is identifiably unique.\(^5\) Namely, Chen, Ghysels and Wang (2011b) show that: letting \( \varepsilon_t(\theta) = RV_t - V_{t|t-1}(\theta) \),

**Proposition 3.1** (Scenario 1). Suppose that Assumptions A.1 and A.2 hold. \( \theta_0 = \arg \min_{\theta \in C} E \varepsilon_t(\theta)^2 \).

We therefore have the following:

**Theorem 3.1** (Scenario 1). Under Assumptions A.1 and A.2,

1. \( \tilde{\theta}_T^{mdrv} \) is identifiably unique and it is a strongly consistent estimator of \( \theta_0 \).
2. Additionally under Assumption A.3 \( \lim_{T \to \infty} \text{var} \left( \frac{1}{\sqrt{T}} \sum_{t=1}^{T} \varepsilon_t(\theta_0) \right) \) exists and is finite, denoted by \( \Omega^{mdrv} \).
3. If \( \Omega^{mdrv} > 0 \), \( \sqrt{T}(\tilde{\theta}_T^{mdrv} - \theta_0) \Longrightarrow N(0, (\Sigma^{md})^{-1}\Omega^{mdrv}(\Sigma^{md})^{-1}) \), where \( 0 < \Sigma^{md} = E \nabla V_{t|t-1}(\theta_0) (\nabla V_{t|t-1}(\theta_0))' < \infty \).

\(^5\)See Gallant and White (1988) for the definition of identifiable uniqueness.
The existence of $\Omega^{mdrv}$ and the asymptotic normality follow from the fact that $\varepsilon_t \partial_k \varepsilon_t$ is near epoch dependent (NED) on the underlying return process under suitable moment conditions, and is therefore a mixingale when the return process is $\alpha$–mixing (see Assumption A.3). It should also be noted that the size of $\alpha$–mixing is $-v_2/(v_2 + 2)$. It is weaker than the size required in Theorem 5.7 of Gallant and White (1988), i.e., $-2v_2/(v_2 + 2)$, and Goncalves and White (2004) as well which requires a size of $-\delta v_2/(v_2 + 2)$ for some $\delta > 2$, although in a different context.  

### 3.2 Estimation under Scenario 2

We now consider situations where the HYBRID GARCH model produces conditional variance predictions. In such a case we are at liberty to consider both minimum distance estimators as in Section 3.1, and quasi-maximum likelihood estimators that are standard in the GARCH literature. More specifically, let us consider the following estimators:

$$
\tilde{\theta}^{mdrv}_T = \arg\min_{\theta \in \mathcal{C}} \frac{1}{T} \sum_{t=1}^{T} \left( RV_t - \tilde{V}_t(\theta) \right)^2 
$$

(3.2)

$$
\tilde{\theta}^{mdr}_T = \arg\min_{\theta \in \mathcal{C}} \frac{1}{T} \sum_{t=1}^{T} \left( R^2_t - \tilde{V}_t(\theta) \right)^2 
$$

(3.3)

$$
\tilde{\theta}^{lhr}_T = \arg\min_{\theta \in \mathcal{C}} \frac{1}{T} \sum_{t=1}^{T} \left( \log \tilde{V}_t(\theta) + \frac{R^2_t}{\tilde{V}_t(\theta)} \right) 
$$

(3.4)

$$
\tilde{\theta}^{lhrv}_T = \arg\min_{\theta \in \mathcal{C}} \frac{1}{T} \sum_{t=1}^{T} \left( \log \tilde{V}_t(\theta) + \frac{RV_t}{\tilde{V}_t(\theta)} \right) 
$$

(3.5)

with $\tilde{V}_t$ defined in (3.1). We also consider estimator derived using Multiplicative Error Model which shares some similarities with the likelihood-RV-based estimator $\tilde{\theta}^{lhrv}_T$.

---

6See page 24 of Gallant and White (1988) for the definition of size.
3.2.1 Minimum Distance and Quasi-Likelihood Estimators

We start by extending Proposition 3.1 to the case of Scenario 2: letting $e_t(\theta) = R_t^2 - V_{t|t-1}(\theta)$,

**Proposition 3.2.** Suppose that Assumptions A.1, A.2, and A.5(2) hold. Under Scenario 2, $\theta_0 = \arg\min_{\theta \in C} E(\varepsilon_t(\theta))^2 = \arg\min_{\theta \in C} E(e_t(\theta))^2$.

Note that $\varepsilon_t(\theta_0)\partial_t\varepsilon_t(\theta_0)$ and $e_t(\theta_0)\partial_t e_t(\theta_0)$ are martingale difference sequences. $\Omega^{mdrv}$ defined in Theorem 3.1 becomes $E[(RV_t - V_{t|t-1}(\theta_0))^2\nabla V_{t|t-1}(\theta_0)\nabla V_{t|t-1}(\theta_0)']$. Define $\Omega^{mdr^2}$ as $E[(R_t^2 - V_{t|t-1}(\theta_0))^2\nabla V_{t|t-1}(\theta_0)\nabla V_{t|t-1}(\theta_0)']$. Both $\Omega^{mdrv}$ and $\Omega^{mdr^2}$ are finite and positive definite under suitable regularity conditions. Consider $\hat{\theta}^{mdrv}_T$ and $\hat{\theta}^{mdr^2}_T$ defined in (3.2) and (3.3).

**Theorem 3.2 (Scenario 2).** Under Assumptions A.1, A.2, and A.5(2),

1. $\hat{\theta}^{mdrv}_T$, $\hat{\theta}^{mdr^2}_T$ are identifiably unique and they converge to $\theta_0$ a.s.
2. Further assume that $E r^8 < \infty$ and Assumption A.5(3) holds, $\sqrt{T}(\hat{\theta}^{mdrv}_T - \theta_0) \rightarrow N(0, (\Sigma^{md})^{-1}\Omega^{mdrv}(\Sigma^{md})^{-1})$ and $\sqrt{T}(\hat{\theta}^{mdr^2}_T - \theta_0) \rightarrow N(0, (\Sigma^{md})^{-1}\Omega^{mdr^2}(\Sigma^{md})^{-1})$.

Moreover, we have the following result that ties the QMLE estimators:

**Proposition 3.3 (Scenario 2).** Suppose $E(\sup_{\phi \in \Phi} H(\phi, \hat{r}_t)) < \infty$. Under Assumptions A.1 and A.2, $\theta_0 = \arg\min_{\theta \in C} E(\log V_{t|t-1}(\theta) + RV_t/V_{t|t-1}(\theta)) = \arg\min_{\theta \in C} E(\log V_{t|t-1}(\theta) + R_t^2/V_{t|t-1}(\theta))$.

Therefore $\theta_0$ can also be estimated by $\hat{\theta}^{thrv}_T$ in (3.4) or $\hat{\theta}^{thr^2}_T$ in (3.5). Define $\Sigma^{lh}$ as $E(V_{t|t-1}(\theta)\nabla V_{t|t-1}(\theta)\nabla V_{t|t-1}(\theta)')$, $\Omega^{thrv}$ as $E(V_{t|t-1}(\theta)(RV_t - V_{t|t-1}(\theta))^2\nabla V_{t|t-1}(\theta)\nabla V_{t|t-1}(\theta)')$, and $\Omega^{thr^2}$ as $E(V_{t|t-1}(\theta)(RV_t - V_{t|t-1}(\theta))^2\nabla V_{t|t-1}(\theta)\nabla V_{t|t-1}(\theta)')$.

**Theorem 3.3 (Scenario 2).** Suppose $E(\sup_{\phi \in \Phi_0} H(\phi, \hat{r}_t)) < \infty$ and Assumptions A.1, and A.2 hold.
(1) \( \tilde{\theta}^{thr}_T, \tilde{\theta}^{hr}_T \) are identifiably unique and they converge to \( \theta_0 \) a.s.

(2) Suppose \( E(r^{4+v}) < \infty \) for some \( v > 0 \), and the following holds

\[
|\partial_\phi H(\phi, \bar{x})/H(\phi, \bar{x})| \leq g(\phi), \quad |\partial^2_\phi H(\phi, \bar{x})/H(\phi, \bar{x})| \leq g(\phi) \quad \forall \bar{x} \in \mathbb{R}^m, \phi \in \Phi
\]

for some real-valued function \( g \) which is continuous in \( \phi \). \( \sqrt{T}(\tilde{\theta}^{thr}_T - \theta_0) \Rightarrow N(0, (\Sigma^{lh})^{-1}\Omega^{thr}(\Sigma^{lh})^{-1}) \), and \( \sqrt{T}(\tilde{\theta}^{hr}_T - \theta_0) \Rightarrow N(0, (\Sigma^{lh})^{-1}\Omega^{hr}(\Sigma^{lh})^{-1}) \).

The likelihood estimation considered here is slightly different from what is discussed in the literature. First of all, \( \sigma^2_{t|t-1} \) is studied in \( L^2(\Omega, \mathcal{F}, P) \) instead of \( L^1(\Omega, \mathcal{F}, P) \). Secondly, the objective function appearing in (3.4) is not the joint quasi-log-likelihood function (modulo a constant) of \( \{R_1, R_2, R_3, \ldots, R_T\} \). Instead of conditioning on \( R_1, R_2, \ldots, R_{t-1} \), \( \log \tilde{V}_t(\theta) + R^2_t/\tilde{V}_t(\theta) \) is conditional quasi-log-likelihood w.r.t. a finer set, the sigma field generated by the high frequency returns up to time \( t - 1 \).

It should also be noted that the discussion can be extended to strictly periodically stationary and periodically ergodic time series as well.\(^7\) This is because the proofs only require \( \bar{r}_t \) to be strictly stationary ergodic.

Note that we have four estimators of \( \theta_0 \): \( \tilde{\theta}^{mdr}_T, \tilde{\theta}^{mdrv}_T, \tilde{\theta}^{hr}_T \) and \( \tilde{\theta}^{thr}_T \). The likelihood-based estimators are superior to the minimum-distance ones in terms of moment conditions. However, it is hard to compare the efficiency between \( R^2 \)-based estimator and RV-based estimator (i.e., \( \tilde{\theta}^{mdr}_T \) v.s. \( \tilde{\theta}^{mdrv}_T \), and \( \tilde{\theta}^{hr}_T \) v.s. \( \tilde{\theta}^{thr}_T \)), because the sign of \( E_{t-1}[(R^2_t - V_{t|t-1}(\theta_0))^2 - (RV_t - V_{t|t-1}(\theta_0))^2] = E_{t-1}(R^4_t - RV^2_t) \) is unclear for an arbitrary return process. Next we consider a special case.

**Corollary 3.1.** Suppose the DGP is a semi-strong GARCH(1,1) and \( E(r^3_s|F_{s-1/m}) = 0 \). Then \( E_{t-1}(R^4_t - RV^2_t) > 0 \). Under the assumptions in Theorems 3.2 and 3.3, \( \tilde{\theta}^{mdrv}_T \) (or \( \tilde{\theta}^{thr}_T \)) has a smaller asymptotic variance than \( \tilde{\theta}^{mdr}_T \) (or \( \tilde{\theta}^{hr}_T \)).

\(^{7}\)See Aknouche and Bibi (2009) for the definition.
3.2.2 Multiplicative Error Models

Inspired by the Multiplicative Error Model of Engle (2002) and the subsequent work by Engle and Gallo (2006), Lanne (2006), Cipollini, Engle, and Gallo (2006), we also consider the model

\[ RV_{t+1} = \sigma_{t+1|t}^2 \eta_{t+1}, \]

where \( \eta_{t+1} \) is independent and identically distributed with mean 1, and \( \sigma_{t+1|t}^2 \) is the conditional expectation of \( RV_{t+1} \) given information at time \( t \). Suppose the cumulative distribution function of \( \eta \) is \( F \). The choice of \( F \) could be a unit exponential (see Engle (2002)), or a Gamma distribution as suggested in Engle and Gallo (2006), or a mixture of two gamma distributions of Lanne (2006). Denote the solution to equation (2.1) as \( V_{t+1|t}(\theta) \). Suppose that \( \theta_0 \in \Theta \) is such that \( \sigma_{t+1|t}^2 = V_{t+1|t}(\theta_0) \). Therefore, we can use the marginal distribution of \( \eta \) to formulate the estimation of \( \theta_0 \). The estimator is then denoted by \( \hat{\theta}_T^{mem} \).

Suppose that the appropriate probability distribution for the error term is a Gamma distribution. In other words, the conditional density of \( RV_{t+1} \) is

\[ f(RV_{t+1}|F_{t-\infty}) = \Gamma(g)^{-1} g^{\frac{1}{2}} \exp\left(-\frac{g RV_{t+1}^2}{\sigma_{t+1|t}^2}\right). \]

Hence \( E(RV_{t+1}|F_{t-\infty}) = \sigma_{t+1|t}^2 \), and \( \text{Var}(RV_{t+1}|F_{t-\infty}) = \frac{\sigma_{t+1|t}^4}{g} \). The parameter space becomes \( \Theta \times \{ g > 0 \} \). As pointed out by Engle and Gallo (2006) and Cipollini, Engle, and Gallo (2006), the estimation of \( \theta_0 \) and \( g \) are asymptotically independent. The point estimation \( \hat{\theta}_T^{mem} \) of \( \theta_0 \) is then same as \( \hat{\theta}_T^{lhrv} \). It follows from the proof of Theorem 3.3 that \( \sqrt{T}(\hat{\theta}_T^{mem} - \theta_0) \) converges to \( N(0, (g \Sigma^{th})^{-1}) \) in distribution.

If the existence of an appropriate parametric density can not be verified, one can consider quasi-likelihood estimation which will yield the same asymptotic result as \( \hat{\theta}_T^{lhrv} \). But since the innovation \( \eta \) is independent, \( \Omega^{lhrv} \) in Theorem 3.3 becomes \( E(RV_t^2/V_{t-1}^2(\theta_0) - 1)\Sigma^{th} \) and hence we do not need the moment condition \( Er^{4+v} < \infty \) to establish the asymptotic normality of \( \hat{\theta}_T^{mem} \). Its asymptotic variance-covariance matrix becomes \( (E\eta_t^2 - 1)(\Sigma^{th})^{-1} \).
We conclude this section with a finite sample simulation study as it will become clear that asymptotic analysis is not sufficient to appraise which estimators are the most attractive for empirical work. The details of the simulation design appear in Appendix B.

The results are quite easy to summarize and therefore not reported in detail (see Chen, Ghysels and Wang (2011b) for the details). The estimator that appears to have the best finite sample properties is LHRV. The LHRV estimator is typically vastly better than the estimators based on $R^2$, either minimum distance or likelihood-based. Compared to the LHR2 estimator, we also find that MDRV - which uses also $RV$ but via a minimum distance criterion - is also less efficient, except in one case $m = 5$ and $T = 1000$. It should also be noted that the MEM estimator - which is asymptotically equivalent to LHRV - is occasionally in small samples the most efficient for one parameter in particular, namely $\alpha_m$. 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. The simulation results indicate that $\hat{\theta}_{T}^{\text{LHRV}}$ and $\hat{\theta}_{T}^{\text{MEM}}$ appear to have the most desirable finite sample properties - the reason why we select these estimators for our empirical analysis.

4 Model Specifications and Evaluations

The class of HYBRID GARCH processes we introduce allows us to address quite a few intriguing empirical modeling strategies. For example, to predict daily volatility, are we better off estimating intra-daily weighting schemes, despite the additional parameters involved compared to using realized volatility and related data-driven HYBRID processes? How should we handle asymmetries? Can we simply rely on sign-sensitive aggregates such as semi-variances, or should we rather estimate intra-
daily news impact curves, also at a cost of additional parameters? When we are interested in weekly horizon forecasts, should we keep using intra-daily data with their own weighting, or should we reply on simple daily realized volatilities? Hence, is the right sampling frequency of returns intra-daily? Or, can we come by with daily aggregates? The same question applies to longer horizons, such as bi-weekly volatility forecasts, which are most relevant for value-at-risk calculations. To address these questions we introduce various model specifications that feature weighting schemes, news impact functions etc. We will focus on cases involving intra-daily returns in a first subsection. The next subsection covers weekly and bi-weekly predictions.

4.1 Daily Volatility Forecasts

We start with data-driven processes followed by parametric specifications for the HYBRID processes. Recall that \( V_{t+1|t} = \alpha + \beta V_{t|t-1} + \gamma H_t \). In the sequel we will provide various specifications for the \( H_t \) process.

*Data-driven HYBRID processes*

A standard GARCH(1,1) model can be viewed as a special case of HYBRID-GARCH with the HYBRID process \( H_t \equiv r^2_t \) where the return in this case is the daily return. Using intra-daily data, we can also consider RV and SemiRV GARCH models, respectively. The \( RV \) GARCH is defined with \( H_t \equiv RV_t \). Likewise, we can account for asymmetries by computing semi-variances, as suggested by Barndorff-Nielsen, Kinnebrock, and Shephard (2008), \( SemiRV_t \equiv \sum_{j=1}^{m} \frac{r^2_{t-(j-1)/m} 1_{r^2_{t-(j-1)/m}<0}}{m} \) yielding the \( SemiRV \) HYBRID process \( H_t \equiv RV_t + \delta SemiRV_t \). The appeal of these HYBRID processes is that they involve only a small set of parameters. A standard GARCH(1,1) involves three parameters. The RV GARCH has the same number of parameters, but
uses intra-daily information via realized volatility. The SemiRV GARCH adds one extra slope parameter \( \delta \) that potentially captures asymmetries of negative returns.

Last but not least, we also consider so called realized volatility measures that are corrected for microstructure noise. Microstructure noise may mask the true price variation. This has prompted a substantial literature on how to correct measures of quadratic variation based on intra-daily data - which we will denote as \( RV_t^* \). See for example Aït-Sahalia, Mykland, and Zhang (2005), Bandi and Russell (2006), Barndorff-Nielsen, Hansen, Lunde, and Shephard (2006) and Hansen and Lunde (2006), and references therein. We therefore also consider data-driven HYBRID processes that correct for the presence of intra-daily microstructure noise. One may use sub-sampling (see Zhang, Mykland, and Aït-Sahalia (2005) and Aït-Sahalia, Mykland, and Zhang (2011)), or the realized kernel (see Barndorff-Nielsen, Hansen, Lunde, and Shephard (2008) and Barndorff-Nielsen, Hansen, Lunde, and Shephard (2011)), or pre-averaging (see Jacod, Li, Mykland, Podolskij, and Vetter (2009)) to construct \( RV_t^* \). In our empirical applications, we will use the realized kernel method. More specifically, we will use the Oxford-Man Institute’s realised library data described in Shephard and Sheppard (2010).\(^8\) We refrain from reporting all the details of the empirical results for two reasons: (1) to save space and (2) since our findings are similar to Patton and Sheppard (2009) - namely that the use of realized measures corrected for microstructure noise yields results that are typically similar (or often worse) to the forecast performance using \( RV \).

**News Impact Curves and HYBRID processes**

The computation of semi-variances brings us to the subject of asymmetric volatility models. In particular, it has been observed that 'good news' and 'bad news' have

\(^8\)See [http://realized.oxford-man.ox.ac.uk/](http://realized.oxford-man.ox.ac.uk/) for further detail.
different impact for the prediction of future volatility. To capture asymmetries, Engle and Ng (1993) introduced the notion of a news impact curve, both as an object of economic interest and a diagnostic tool for volatility modeling. One commonly used specification was advocated by Glosten, Jagannathan, and Runkle (1993), which involves the following HYBRID process that pertains to the so called TGARCH or threshold GARCH model: $H_t \equiv r_t^2 + \delta 1_{r_t<0}r_t^2$. Inspired by the specification Chen and Ghysels (2011) extended the notion of news impact curve applicable to a mixture of high and low frequency returns. Their approach involves MIDAS regressions with intra-daily news impact curves. Overall, they find that moderately good (intra-daily) news reduces volatility (the next day), while both very good news (unusual high intra-daily positive returns) and bad news (negative returns) increase volatility, with the latter having a more severe impact. Chen and Ghysels (2011) also find that asymmetries disappear over longer horizons. Moreover, they find that models featuring asymmetries dominate in terms of out-of-sample forecasting performance, especially during the 2007-2008 financial crisis.

We will introduce HYBRID processes involving intra-daily news impact curves - similar to Chen and Ghysels (2011) - but within the context of HYBRID-GARCH. Before we do, we need to discuss how to handle weighting schemes for intra-daily data.

*How to Better Use High-Frequency Financial Data*

The main empirical results of the paper pertain to whether we can make better use of intra-daily returns. The first approach involves a HYBRID process that allows for unequal weighting of intra-daily returns. The parametric specification of the HYBRID process is as follows:

$$H(\phi, r_t) \equiv H_t \equiv \sum_{j=1}^{m} \exp \left( \sum_{i=1}^{j-1} (\phi_0 + \phi_1 i/m + \phi_2 (i/m)^2) \right) r_{t-(j-1)/m}^2$$

(4.1)
where the use of the (quadratic) Exponential Almon MIDAS filter, with $\phi = (\phi_0, \phi_1, \phi_2)$, is inspired by a similar approach proposed in Ghysels, Santa-Clara, and Valkanov (2005).¹

Note that when $\phi_0 = \phi_1 = \phi_2 = 0$, we have the equal weighting scheme of the RV specification. Estimating the parameters amounts to allowing for different weighting schemes. Compared to the RV-driven HYBRID process, we have at most three extra parameters to estimate in order to retrieve more information from the intra-daily return series. The extra parameters come at a cost that may cause poor out-of-sample forecast performance. This is an empirical question that we will address in the next section.

While parsimony is a concern, we do want to discuss first specifications involving more parameters than we have handled so far. Indeed, following the theme of asymmetries discussed earlier, we also consider HYBRID processes involving both intra-daily weighting schemes as well as news impact curves, yielding:

$$H_t \equiv \sum_{j=1}^{m} \exp \left( \sum_{i=1}^{j-1} (\phi_0 + \phi_1 i/m + \phi_2 (i/m)^2) \right) (1 + \delta r_{t-(j-1)/m < 0}) r_{t-(j-1)/m}^2 \quad (4.2)$$

Together with the parameters $\alpha$, $\beta$, and $\gamma$ in the generic HYBRID GARCH, this is the model specification with the largest parameter space. But it involves a flexible weighting scheme as well as an intra-daily news impact curve.

The concerns for parsimony prompts us to think about putting restrictions on the specification of the HYBRID process and/or putting restrictions on the parameters $\alpha$ and $\beta$ that determine the volatility dynamics. We will make a distinction between parameter restriction on the parameters driving the HYBRID process and restrictions

¹We opted for the Exponential Almon because it yields a convenient approach to testing restrictions, as opposed to the Beta polynomial proposed by Chen and Ghysels (2011) to handle intra-daily seasonality in the context of a MIDAS regression. Note also that the Exponential Almon lags in equation (4.1) are not normalized to add up to one.
pertain to the parameters $\alpha$ and $\beta$. We will start with the latter.

Recall that with volatility being a persistent process, it is natural to weight intra-daily data differently. Recall that this is one of the motivations behind the class of HYBRID GARCH models. We can carry this logic a step further, as the persistence not only affects the intra-daily weighting scheme, but should also affect the parameter $\beta$ in equation (2.1). The connection between the parameter vector $(\phi_0, \phi_1, \phi_2)$ and $\beta$ relates to what we refer to earlier as structural HYBRID processes. One such restriction relates to what Chen, Ghysels and Wang (2011a) refer to as Periodic HYBRID GARCH models because of its connection with the periodic GARCH(1,1) model of Bollerslev and Ghysels (1996). The model can be written as follows:

$$\begin{align*}
V_{t+1|t} &= \alpha + \exp \left( \sum_{i=1}^{m} (\phi_0 + \phi_1 i/m + \phi_2 (i/m)^2) \right) V_{t|t-1} \\
&\quad + \gamma \sum_{j=1}^{m} \exp \left( \sum_{i=1}^{j-1} (\phi_0 + \phi_1 i/m + \phi_2 (i/m)^2) \right) r_{t-(j-1)/m}^2
\end{align*}$$

We will call this HYBRID slope-constrained GARCH, or HYBRID SC GARCH as the parameter $\beta$ is now replaced and linked to the parameters of the HYBRID process. Note that the above specification can also be enriched with an intra-daily news impact curve, yielding a HYBRID SC TGARCH.

Another set of restrictions pertains to the parametrization of the HYBRID process itself. In equation (4.1) we assumed a quadratic polynomial for the HYBRID process. We can set $\phi_2$ equal to zero yielding an exponentially affine weighting scheme. We will refer to such a model as FC1 HYBRID GARCH, or FC1 HYBRID TGARCH, the latter involving news impact curves. These models can also feature slope constraints, such as for example FC1 HYBRID SC GARCH. Finally, we can further restrict the HYBRID process by imposing $\phi_2 = \phi_1 = 0$. This restriction on the HYBRID process, combined with its associated slope restriction amounts to assuming a high frequency GARCH(1,1) data generating process (DGP) and is an example of a structural HYBRID GARCH.
with an underlying DGP. We will refer to this type of restriction as \textit{FC0}.

**4.2 Weekly and Bi-Weekly Volatility Forecasts**

When we turn our attention to weekly and bi-weekly forecasts, we face interesting issues which we want to highlight in this subsection. In principle, these issues also apply to intra-daily returns and daily forecasts, but they are more acute for longer term forecasts. Before we do, we would like to note that all the models discussed in the previous subsection readily apply to multiple day horizons. For example, the general HYBRID TGARCH with an \( h \) day forecast horizon can be written as:

\[
V_{t+1|t} = \alpha + \beta V_{t|t-1} + \gamma \sum_{j=1}^{mh} \exp \left( \sum_{i=1}^{j-1} (\phi_0 + \phi_1 i/mh + \phi_2 (i/mh)^2) \right) \left( 1 + \delta \mathbf{1}_{r_{t-(j-1)/m} < 0} r_{t-(j-1)/m}^2 \right).
\]

The above equation collapses to (4.2) when \( h = 1 \). Such extensions also apply to slope constrained GARCH models as well as models involving constrained HYBRID processes.

\textit{Selecting the right sampling frequency}

To motivate the discussion we start with a standard GARCH(1,1) model using weekly returns, namely:

\[
V_{t+1|t} = \alpha + \beta V_{t|t-1} + \gamma (r_{t\text{w}})^2
\]

where we emphasize the fact that returns are sampled weekly with the superscript \( w \). Likewise, we can think of a RV-GARCH(1,1) model:

\[
V_{t+1|t} = \alpha + \beta V_{t|t-1} + \gamma RV_t^{w}
\]
with the same convention for the notation. We could, however, also consider a HYBRID GARCH with daily returns to replace equation (4.5):

\[ V_{t+1\mid t} = \alpha + \beta V_{t\mid t-1} + \gamma \sum_{j=1}^{h} \exp \left( \sum_{i=1}^{j-1} (\phi_0 + \phi_1 i/h + \phi_2 (i/h)^2) \right) (r_{ht-(j-1)})^2 \] (4.7)

where the superscript \( d \) refers to daily returns and \( h = 5 \) for the weekly/daily sampling frequency combination. We will denote the above model as HYBRID GARCH \( D \).

Likewise, we can also apply the same logic to RV-driven GARCH models, namely replacing equation (4.6) with:

\[ V_{t+1\mid t} = \alpha + \beta V_{t\mid t-1} + \gamma \sum_{j=1}^{h} \exp \left( \sum_{i=1}^{j-1} (\phi_0 + \phi_1 i/h + \phi_2 (i/h)^2) \right) RV_{ht-(j-1)}^d \] (4.8)

Obviously, as before, we are facing a trade-off between parameter proliferation and the level of aggregation of returns and/or RV. In each of these cases, we can constraint the HYBRID process parameters, or we can impose a slope restriction. Likewise, we can also impose news impact curves when we consider (daily) returns. The generic specification in equation (4.4) uses all high frequency data directly and involves a general HYBRID process with intra-daily news impact curves. The key question therefore, is whether to use intra-daily data, or daily returns or daily RV, instead of weekly returns or weekly RV.

It should also be noted that the selection of the sampling frequency pertains to the issue of considering an iterated forecast for the next five days versus a direct forecast based on a coarser data set of past weekly returns. Relatively little work has been done on the comparison between direct and iterated forecasts in the context of volatility. One exception is Ghysels, Rubia, and Valkanov (2009) who consider a regression-based approach and compare several approaches of producing multi-period ahead forecasts of
volatility –iterated, direct, and mixed-data sampling (MIDAS) regressions. The latter is shown to be the best and is similar to the HYBRID GARCH class of models proposed here. While the MIDAS approach dominates, it is not clear whether we should use equation (4.4) applicable to all high frequency data directly, or daily aggregates such as $RV^d$, or daily returns, etc. These questions pertain to the selection of the proper sampling frequency in a mixed sampling frequency setting.

4.3 Data and Model Evaluation

We use the S&P 500 Futures 5-minute returns from April 1st, 1982 to December 31st, 2008. To estimate the models and perform out-of-sample evaluations, we use a rolling window sample which is moved forward monthly. There are 178 rolling windows and each window contains 120-months in-sample data and 24-month set aside for out-of-sample appraisals. We consider three horizons: one-day, one-week, and two-weeks ahead forecasts. All models, with specifications listed in Table 1 and 2, are estimated using two methods: MEM and LHRV - inspired by our Monte Carlo simulation findings. The out-of-sample forecast performance of models using the same estimation method and across different estimation methods are evaluated via the Giacomini and White (2006) test, henceforth denoted GW, which can be viewed as a generalization, or a conditional version of the Diebold and Mariano (1995) and West (1996) tests. Another appeal of using the GW test is that it can handle non-nested models, which is the case in our application. In fact, Giacomini and White (2006) stress the difference between what they call forecasting methods versus forecasting models. Loosely speaking forecasting methods are the combination of estimation sample, model specification and prediction sample. In our application this is most relevant, as we will not only compare models involving high frequency data directly, or daily aggregate measures, but we will also
include ARCH-type models involving daily returns - i.e. the original models that were used in the literature on asymmetries in volatility.

The loss function we use in the GW test is QLike, which has desirable properties and is robust to measurement error noises in volatility (see Patton (2011)). The QLike loss function is defined as \( L(h_t, RV_t) = \log h_t + RV_t/h_t - (\log RV_t + 1) \).

Finally, we construct a score for each model based on the GW tests’ p-values and ratios:

\[
s_A = \frac{\sum_{B \neq A} 1_{p_{AB} < \alpha} r_{AB} + 0.5(1 - 1_{p_{AB} < \alpha})}{n(n-1)/2} \tag{4.9}
\]

where \( s_A \) is the score of model A; \( p_{AB} \) is the p-value of GW test comparing models A and B; \( \alpha \) is the significance level and set as 10% in the paper; \( r_{AB} \) is the ratio of model A being predicted as better choice than model B; \( n \) is the number of models. The score is normalized such that the summation of scores of all models is one.\(^{10}\)

## 5 Empirical Findings

The comparison of Conditional Predictive Ability (CPA) of each model between different methods are shown in Table 3 for daily, weekly, and bi-weekly data. The results consistently show that the choice of estimation method - MEM or LHRV - is not important, except that for a few exceptions, LHRV is preferred. Hence, to compare the forecast performance of models, we only need to consider the LHRV method. The p-values of GW tests are shown in Table 4 for daily models. The details for the weekly and bi-weekly data, and the statistics of GW test are omitted in order to save space. We also report the ratio of choosing model A (the model No. shown in the row header) against model B (the model No. shown in the column header) according to the decision

\(^{10}\)Other methods for comparing large classes of models exist, see e.g. Hansen, Lunde, and Nason (2011) and references therein. These tests are unconditional, however, whereas we prefer to rely on conditional tests.
rule discussed in Giacomini and White (2006). When the p-value of GW test is below the significance level of interest, for example, 10%, the ratio is an indicator of the relative forecast performance of the two models.

Examining the results in Table 4, we note the following for daily model comparisons. GARCH and TGARCH models using daily returns, Model (1) and (2), are dominated by all other models, i.e. models using intra-daily data. Symmetric models are always dominated by their asymmetric counterparts, which implies that asymmetry does matter, despite the required extra parameters. Focusing on symmetric models, the RV GARCH model, Model (3), is dominated by all other HYBRID or slope constrained HYBRID models. Focusing on asymmetric models, the SemiRV GARCH model, Model (4), is also always less preferred than any HYBRID or slope constrained HYBRID asymmetric models. These results support two key findings for daily forecasting of volatility: (1) asymmetries matter and (2) the weighting scheme also does matter.

![Figure 1: Best Daily Model in Three Subsamples](image)

(a) Weights  
(b) News Impact Curve

Figure 1: Best Daily Model in Three Subsamples

To further document these findings we turn our attention to Figure 1(a) where we display the weights of the best daily model, i.e. the HYBRID TGARCH model, in three representative subsamples of our rolling sample scheme: P1 is for period December,
1986 to November, 1996; P2 for period August, 1989 to July, 1999; P3 for period April, 1993 to March, 2003.\textsuperscript{11} It appears from the figure that the weighting schemes are fairly stable across subsamples. In Figure 1(b) we display the news impact curve of the best daily model in the same three subsamples. We see again a stable pattern across subsamples and a pattern that is distinctly asymmetric and feature larger impact of bad news - as commonly documented in the literature.

Next we turn to the weekly models, where we find that same as daily models, GARCH and TGARCH models are dominated by all other models. Symmetric models are dominated by their asymmetric counterparts in most cases, which implies that asymmetric effects still matter at weekly horizons. The models using daily returns (having $D$ in their acronym), are always dominated by their counterparts using 5-minute returns. In fact, almost all models based on 5-minute returns show better performance than any model based on daily returns, which implies that using 5-minute returns, instead of daily returns, may result in better forecast performance for weekly models. RV GARCH models perform worse than most HYBRID or slope constrained HYBRID models. The same conclusion applies to SemiRV GARCH models. Hence, the weight scheme still matters for weekly horizons.

Figure 2 displays the weighting scheme and news impact curves of the best weekly model, i.e. FC0 HYBRID SC TGARCH model, in the aforementioned three subsamples. Arguably, the weighting schemes and impact curves vary more across subsamples. Nevertheless all display the two key features supporting our findings: (1) the weighting scheme matters and (2) asymmetries do too.

Finally we turn our attention to bi-weekly models. GARCH and TGARCH models with daily data are again dominated by other models. As to the asymmetric effect, now the conclusion is different: many asymmetric models are dominated by their

\textsuperscript{11}We will discuss later how the best model is selected - the results appear in Table 5.
symmetric counterparts. It implies that asymmetric effects have less importance as
the horizon increases, a finding also reported in Chen and Ghysels (2011). The models
using daily returns are still less preferred to their counterparts using 5-minute returns.
RV and SemiRV GARCH models are still less preferred than most HYBRID or slope
constrained HYBRID models, which means that the weight scheme still matters for
bi-weekly horizon.

Figure 2: Best Weekly Model in Three Subsamples

Figure 3: Best Bi-weekly Model in Three Subsamples
Figure 3 covers the weighting scheme and news impact curve of the best bi-weekly model, i.e. FC1 HYBRID SC GARCH model, in the same three subsamples. We clearly see that (1) the weighting schemes vary across subsamples and (2) the asymmetries have disappeared as shown in the news impact curves.

The score of a model is a good indicator of model’s forecast performance. We rank all models according to their score in Table 5 which provides an easy summary. As shown in Table 5, the top 6 models are always HYBRID or slope constrained HYBRID models, and the last two models are always GARCH and TGARCH models.

6 Conclusions

We proposed a general unifying framework that allows the use of different frequency returns to model conditional heteroskedasticity. We call the class of models HYBRID-GARCH models, as the volatility dynamics are driven by what we call HYBRID processes. The topics addressed in this paper have many applications, given the wide use of multi-period volatility forecasts in risk and portfolio analysis.

The main conclusions, based on our analysis are fairly simple: to use intra-daily weighting schemes other than purely aggregation like RV and SemiRV; to use 5-minute returns directly rather than daily returns or daily aggregates like RV even in long horizon forecast; to include asymmetric effect at short horizons but not for long horizons. These results have implications that go against most of the current practice of (1) using direct forecasting methods for long horizon forecasts, (2) using daily aggregates of intra-daily data, and (3) using daily returns or realized measures for long horizon forecasts.
References


Technical Appendices

A Regularity Conditions

The purpose of this appendix is to collect the regularity conditions used in the paper. In what follows, we use $\nabla$ to denote the vector differential operator (w.r.t. $\theta$) so that $\nabla f$ is the gradient (column vector) of scaler function $f$, and $Hess(f)$ the Hessian matrix of $f$, i.e., $ent_{i,j}Hess(f) = \partial_i \partial_j f$ where $\partial_k$ denotes the partial derivative w.r.t. the $k^{th}$ parameter in $\theta = (\alpha, \beta, \gamma, \phi)$. When $\phi$ is a vector, $\partial_{\phi}$ refers to the partial derivative w.r.t. each component of $\phi$, say $\phi_i$, and $\partial^2_{\phi}$ is treated as $\partial_{\phi_i} \partial_{\phi_j}$. $\nabla_{\phi}$ is a vector differential operator w.r.t. $\phi$ when $\phi$ is a vector.

**Assumption A.1.** $r^2_s \in L^2(\Omega, \mathcal{F}, P)$ for some probability space $(\Omega, \mathcal{F}, P)$, and $P(RV_{t+1} | I_t) = \sigma^2_t + 1 | t$, where $RV_{t+1} = \sum_{j=0}^{m-1} r^2_{t+1-j/m}$, $I_t$ is a closed subspace of $L^2(\Omega, \mathcal{F}, P)$ and it consists of the information on the high frequency returns up to time $t$. $r_s$'s are non-degenerate, and linearly independent.

**Assumption A.2.** $\{r_s\}$ is strictly stationary and ergodic.

**Assumption A.3.** $\{r_s\}$ is strictly stationary and strong mixing. The mixing coefficient $\alpha(k)$ satisfies $\sum_{k=0}^{\infty} \alpha(k)v_2/(2+v_2) < \infty$ for some $v_2 > 0$. And $E r_s^{4(2+v_2)} < \infty$.

**Assumption A.4.** Suppose that $\Phi$ is a connected set and $H$ is a mapping from $\Phi \times \mathbb{R}^m$ to $\mathbb{R}^+$ such that (1) $H(\cdot, \bar{x}) \in C^2$ for $\bar{x} \in \mathbb{R}^m$; (2) $H(\phi, \cdot), \partial_\phi H(\phi, \cdot), \partial^2_{\phi} H(\phi, \cdot)$ are $\mathcal{B}(\mathbb{R}^m)/\mathcal{B}(\mathbb{R})$ measurable for $\phi \in \Phi$; (3) for $r_t$ satisfying Assumption A.1, $1$ and $H(\phi, \bar{r}_t)$ and (each component of) $\partial_\phi(H(\phi, \bar{r}_t))$ are linearly independent for all $\phi \in \Phi$.

**Assumption A.5** (Moment conditions on $H$). For $r_t$ satisfying Assumption A.1,

1. $E \sup_{\phi \in \Phi} H(\phi, \bar{r}_t)$, $E \sup_{\phi \in \Phi} |\partial_\phi H(\phi, \bar{r}_t)|$ and $E \sup_{\phi \in \Phi} |\partial^2_{\phi} H(\phi, \bar{r}_t)|$ are finite.
2. $E(\sup_{\phi \in \Phi} H(\phi, \bar{r}_t))^2$, $E(\sup_{\phi \in \Phi} |\partial_\phi H(\phi, \bar{r}_t)|)^2$ and $E(\sup_{\phi \in \Phi} |\partial^2_{\phi} H(\phi, \bar{r}_t)|)^2$ are finite.

38
(3) $E(\sup_{\phi \in \Phi} H(\phi, \vec{r}_t))^4$ and $E(\sup_{\phi \in \Phi} |\partial_\phi H(\phi, \vec{r}_t)|^4$ are finite.

(4) $E(\sup_{\phi \in \Phi} H(\phi, \vec{r}_t)^{2(2+v_2)}$ and $E(\sup_{\phi \in \Phi} |\partial_\phi H(\phi, \vec{r}_t)|)^{2(2+v_2)}$ are finite, where $v_2$ is defined in Assumption A.3.

Assumption A.6. The rank of the matrix $(\nabla_{\phi_1} \Psi_0 \nabla_{\phi_1} \Psi_1 \ldots \nabla_{\phi_1} \Psi_{m-1})$ is same as the dimension of $\phi_1$.

Assumption A.7. (1) $\theta_\ast$ is identifiably unique in $C$. (2) $\theta_\ast \in C^0$. (3) The determinant of $E[Hess(\varepsilon_t^2)(\theta_\ast)]$ (or $E[Hess(e_t^2)(\theta_\ast)]$) is positive.

Assumption A.8. (1) $\theta_{\ast\ast}$ is identifiably unique in $C$. (2) $\theta_{\ast\ast} \in C^0$. (3) The determinant of $E[Hess(l_t)(\theta_{\ast\ast})]$ (or $E[Hess(h_t)(\theta_{\ast\ast})]$) is positive.

B Simulation study

We consider two data generating processes. The first is a discrete-time GARCH process: strong GARCH(1,1). Namely,

$$r_{s+1/m} = \sqrt{v_{s+1/m}|s} \varepsilon_{s+1/m}, v_{s+1/m}|s = a + bv_{s-1/m} + cr_s^2 \varepsilon_{s+1/m} \overset{iid}{\sim} N(0,1). \quad (B.1)$$

The second is a GARCH diffusion process, i.e., Model (2.9) with $\eta = 0$. The discretely-sampled high frequency return $r_s = p_s - p_{s-1/m}$ is therefore a weak GARCH(1,1):

$$\sigma_{s+1/m|s}^2 = a + b(\sigma_{s-1/m}^2 + cr_s^2) \sigma_{s+1/m}^2 = P_t(r_{s+1/m|L_s}), P_t(r_{s+1/m|L_s}) = 0 \quad (B.2)$$

with $a = \omega(1-e^{-\theta/m})/m, c = e^{-\theta/m} - b$, and $|b| < 1$ is the solution to $b/(1+b^2) = (\rho e^{-\theta/m} - 1)/(\rho(1+e^{-2\theta/m} - 2)$ where $\rho = [4(e^{-\theta/m} - 1+\theta/m) + 2\theta/m(1+\theta/m(1-\lambda)/\lambda)/(1-e^{-2\theta/m})].$

We then construct the HYBRID GARCH process (2.7) based on either model (B.1) or model (B.2).

The values of parameters in model (B.1) are $a = 2.8E - 06, b = 0.9770, c = 0.0225$ which are taken from Meddahi and Renault (2004). It is easy to check that $r_s$ has finite 8th moment.
(see Bollerslev (1986)) and it satisfies Assumption A.2. For the GARCH diffusion process, we consider \( \theta = 0.0350, \omega = 0.6365, \lambda = 0.2962 \) which are based on the daily DeutscheMark - US dollar exchange rate from October 1, 1987 to September 30, 1992 (See Andersen, Bollerslev, and Lange (1999)). The values of \( \alpha_m, \beta_m \) and \( \gamma_m \) are reported in table below with \( m = 5, 78, 288 \) for model (B.1) and with \( m = 24, 144, 288 \) for model (B.2).

<table>
<thead>
<tr>
<th>m</th>
<th>a</th>
<th>b</th>
<th>c</th>
<th>( \alpha_m )</th>
<th>( \beta_m )</th>
<th>( \gamma_m )</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>2.8E-06</td>
<td>0.9770</td>
<td>0.0225</td>
<td>0.0001</td>
<td>0.8902</td>
<td>0.1124</td>
</tr>
<tr>
<td>78</td>
<td>2.8E-06</td>
<td>0.9770</td>
<td>0.0225</td>
<td>0.0147</td>
<td>0.1628</td>
<td>1.7216</td>
</tr>
<tr>
<td>288</td>
<td>2.8E-06</td>
<td>0.9770</td>
<td>0.0225</td>
<td>0.1429</td>
<td>0.0012</td>
<td>6.0365</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>m</th>
<th>a</th>
<th>b</th>
<th>c</th>
<th>( \alpha_m )</th>
<th>( \beta_m )</th>
<th>( \gamma_m )</th>
</tr>
</thead>
<tbody>
<tr>
<td>24</td>
<td>3.86e-05</td>
<td>0.9794</td>
<td>0.0192</td>
<td>0.0216</td>
<td>0.6065</td>
<td>0.4523</td>
</tr>
<tr>
<td>144</td>
<td>1.07e-06</td>
<td>0.9915</td>
<td>0.0082</td>
<td>0.0204</td>
<td>0.2945</td>
<td>1.1619</td>
</tr>
<tr>
<td>288</td>
<td>2.69e-07</td>
<td>0.9940</td>
<td>0.0059</td>
<td>0.0195</td>
<td>0.1776</td>
<td>1.6590</td>
</tr>
</tbody>
</table>

The estimators considered are: \( \hat{\theta}_T^{mdr} \), defined in (3.2), and the companion estimator \( \hat{\theta}_T^{mdr^2} \), replacing \( RV \) by \( R^2 \), as well as (quasi-)likelihood-based estimators \( \hat{\theta}_T^{lhr^2} \) defined in (3.4), and \( \hat{\theta}_T^{lhrv} \), defined in (3.5). Finally, the simulation study also includes the MEM method described in Section 3.2.2. Recall that Engle and Gallo (2006) and Cipollini, Engle, and Gallo (2006) noted that the estimation of \( \theta_0 \) and \( g \) are asymptotically independent, and thus \( \hat{\theta}_T^{mem} \) is asymptotically the same as \( \hat{\theta}_T^{lhrv} \). The purpose of this section is to examine differences in small sample behavior.

In the simulation experiment, we consider 1000 replications of sample path (2.7), each having the first 1000 observations burn-in and consisting of 500 and 1000 observations left in the sample. For the continuous case, we use Euler discretization to simulate the diffusion process: take one day as a reference measure, and simulate 24 hours of trading with \( dt = 1/86400 \).
Table 1: Summary of Model Specifications (A)

Let $r_{s/m}$ denote the high-frequency return and $m$ is the frequency in one day, i.e., $r_{s/m}$ is the $i$th high-frequency return at day $j$ if $s = j-1+i/m, i=1, \ldots, m$. The return at period $t$ of interest, namely daily, weekly, and bi-weekly return, is denoted as $R_t \equiv \sum_{t=1}^{m} r_{ht-(i-1)/m}$, where $h$ is the number of days in one period, i.e., $h = 1$ for daily return, 5 for weekly return, and 10 for bi-weekly return. Especially, we use $r_t^D$ to denote daily return at day $t$. The realized volatility at period $t$ is $RV_t = \sum_{s=1}^{m} r_{ht-(i-1)/m}^2$, and the realized semivariance is $SemiRV_t = \sum_{s=1}^{m} 1_{r_{ht-(i-1)/m} < 0} r_{ht-(i-1)/m}^2$. All "*** D" models, namely the twelve models with model No. (11) to (14), (19) to (22), and (27) to (30), are only applied to weekly or bi-weekly data. "FC1 ***" models set the restriction $\phi_2 = 0$ and "FC0 ***" models $\phi_1 = \phi_2 = 0$ on the corresponding models. For example, model (27) FC0 HYBRID GARCH D is defined as the same as HYBRID GARCH D, i.e. model (11), with restriction that $\phi_1 = \phi_2 = 0$.

(1) GARCH  \quad V_{t+1|t} = \alpha + \beta V_{t|t-1} + \gamma R_t^2

(2) TGARCH  \quad V_{t+1|t} = \alpha + \beta V_{t|t-1} + \gamma R_t^2 + d1_{R_t<0} R_t^2

(3) RV GARCH  \quad V_{t+1|t} = \alpha + \beta V_{t|t-1} + \gamma RV_t

(4) SemiRV GARCH  \quad V_{t+1|t} = \alpha + \beta V_{t|t-1} + \gamma RV_t + \delta SemiRV_t

(5) RV GARCH D  \quad V_{t+1|t} = \alpha + \beta V_{t|t-1} + \gamma \sum_{j=1}^{h} r_{ht-(j-1)}^2

(6) SemiRV GARCH D  \quad V_{t+1|t} = \alpha + \beta V_{t|t-1} + \gamma \sum_{j=1}^{h} (1 + \delta 1_{1_{r_{ht-(j-1)/m}<0}}) r_{ht-(j-1)}^2

(7) HYBRID GARCH  \quad V_{t+1|t} = \alpha + \beta V_{t|t-1} + \gamma \sum_{j=1}^{h} \exp \left( \sum_{i=1}^{j-1} (\phi_0 + \phi_i/mh + \phi_2(i/mh)^2) \right) R_t^{2(j-1)/m}

(8) HYBRID TGARCH  \quad V_{t+1|t} = \alpha + \beta V_{t|t-1} + \gamma \sum_{j=1}^{h} \exp \left( \sum_{i=1}^{j-1} (\phi_0 + \phi_i/mh + \phi_2(i/mh)^2) \right) (1 + \delta 1_{1_{r_{t-1/m}<0}}) R_t^{2(j-1)/m}

(9) HYBRID SC GARCH  \quad V_{t+1|t} = \alpha + \exp \left( \sum_{i=1}^{h} (\phi_0 + \phi_i/mh + \phi_2(i/mh)^2) \right) V_{t|t-1} + \gamma \sum_{j=1}^{h} \exp \left( \sum_{i=1}^{j-1} (\phi_0 + \phi_i/mh + \phi_2(i/mh)^2) \right) R_t^{2(j-1)/m}

(10) HYBRID SC TGARCH  \quad V_{t+1|t} = \alpha + \exp \left( \sum_{i=1}^{h} (\phi_0 + \phi_i/mh + \phi_2(i/mh)^2) \right) V_{t|t-1} + \gamma \sum_{j=1}^{h} \exp \left( \sum_{i=1}^{j-1} (\phi_0 + \phi_i/mh + \phi_2(i/mh)^2) \right) (1 + \delta 1_{1_{r_{t-(j-1)/m}<0}}) R_t^{2(j-1)/m}
Table 2: Summary of Model Specifications (B)

(11) HYBRID GARCH D  
\[ V_{t+1|t} = \alpha + \beta V_{t|t-1} + \gamma \sum_{j=1}^{h} \exp \left( \sum_{i=1}^{j-1} (\phi_0 + \phi_1 i/h + \phi_2 (i/h)^2) \right) r_{ht-(j-1)}^d \]

(12) HYBRID TGARCH D  
\[ V_{t+1|t} = \alpha + \beta V_{t|t-1} + \gamma \sum_{j=1}^{h} \exp \left( \sum_{i=1}^{j-1} (\phi_0 + \phi_1 i/h + \phi_2 (i/h)^2) \right) (1 + \delta 1_{r_{ht-(j-1)} < 0}) r_{ht-(j-1)}^d \]

(13) HYBRID SC GARCH D  
\[ V_{t+1|t} = \alpha + \exp \left( \sum_{i=1}^{h} (\phi_0 + \phi_1 i/h + \phi_2 (i/h)^2) \right) V_{t|t-1} \]
\[ + \gamma \sum_{j=1}^{h} \exp \left( \sum_{i=1}^{j-1} (\phi_0 + \phi_1 i/h + \phi_2 (i/h)^2) \right) r_{ht-(j-1)}^d \]

(14) HYBRID SC TGARCH D  
\[ V_{t+1|t} = \alpha + \exp \left( \sum_{i=1}^{h} (\phi_0 + \phi_1 i/h + \phi_2 (i/h)^2) \right) V_{t|t-1} \]
\[ + \gamma \sum_{j=1}^{h} \exp \left( \sum_{i=1}^{j-1} (\phi_0 + \phi_1 i/h + \phi_2 (i/h)^2) \right) (1 + \delta 1_{r_{ht-(j-1)} < 0}) r_{ht-(j-1)}^d \]

(15) FC1 HYBRID GARCH  
(16) FC1 HYBRID TGARCH  
(17) FC1 HYBRID SC GARCH  
(18) FC1 HYBRID SC TGARCH  
(19) FC1 HYBRID GARCH D  
(20) FC1 HYBRID TGARCH D  
(21) FC1 HYBRID SC GARCH D  
(22) FC1 HYBRID SC TGARCH D  
(23) FC0 HYBRID GARCH  
(24) FC0 HYBRID TGARCH  
(25) FC0 HYBRID SC GARCH  
(26) FC0 HYBRID SC TGARCH  
(27) FC0 HYBRID GARCH D  
(28) FC0 HYBRID TGARCH D  
(29) FC0 HYBRID SC GARCH D  
(30) FC0 HYBRID SC TGARCH D
Table 3: GW Test for comparing daily models between methods
Each model is estimated through two methods: MEM and LHRV. The GW tests show that the conditional predictive ability of each model might be different due to its estimation methods. The column names “stat” stands for “Statistics” and “ratio” for “the ratio of choosing model estimated through method A other than that through method B according to the decision rule”. A ratio is only meaningful when the corresponding p-value is smaller than the given significance level, for example, 10%.

| MODEL | Daily Models | | | Weekly Models | | | Bi-Weekly Models | | |
|-------|--------------|---|---|--------------|---|---|--------------|---|
|       | LHRV | MEM | LHRV | MEM | LHRV | MEM | |
| (1)   | 1.18 | 0.55 | 0.0 | 0.54 | 0.76 | 1.0 | 2.00 | 0.37 | 1.0 |
| (2)   | 0.70 | 0.70 | 0.1 | 1.10 | 0.58 | 1.0 | 0.23 | 0.89 | 0.0 |
| (3)   | 1.04 | 0.59 | 1.0 | 0.27 | 0.87 | 0.0 | 2.09 | 0.35 | 0.0 |
| (4)   | 0.28 | 0.87 | 0.0 | 3.69 | 0.16 | 1.0 | 1.71 | 0.43 | 0.0 |
| (5)   | 1.16 | 0.56 | 0.0 | 0.83 | 0.66 | 1.0 | 2.00 | 0.37 | 1.0 |
| (6)   | 2.58 | 0.27 | 0.0 | 1.29 | 0.53 | 1.0 | 2.00 | 0.37 | 1.0 |
| (7)   | 2.98 | 0.23 | 0.9 | 5.56 | 0.06 | 1.0 | 0.92 | 0.63 | 0.9 |
| (8)   | 1.28 | 0.53 | 0.1 | 6.83 | 0.03 | 1.0 | 5.16 | 0.08 | 1.0 |
| (9)   | 0.33 | 0.85 | 0.0 | 2.58 | 0.27 | 0.9 | 1.89 | 0.39 | 0.1 |
| (10)  | 2.12 | 0.35 | 0.0 | 5.31 | 0.07 | 1.0 | 1.53 | 0.46 | 1.0 |
| (11)  | 1.17 | 0.56 | 0.1 | 1.39 | 0.50 | 0.1 | 2.00 | 0.37 | 1.0 |
| (12)  | 4.03 | 0.13 | 0.0 | 1.18 | 0.55 | 0.9 | 2.00 | 0.37 | 1.0 |
| (13)  | 0.66 | 0.72 | 0.0 | 1.93 | 0.38 | 1.0 | 2.00 | 0.37 | 1.0 |
| (14)  | 3.85 | 0.15 | 0.0 | 3.37 | 0.19 | 1.0 | 2.00 | 0.37 | 1.0 |
| (15)  | 4.99 | 0.08 | 1.0 | 3.76 | 0.15 | 1.0 | 4.21 | 0.12 | 1.0 |
| (16)  | 1.99 | 0.37 | 0.0 | 2.68 | 0.26 | 0.1 | 3.08 | 0.21 | 1.0 |
| (17)  | 0.95 | 0.62 | 0.9 | 1.63 | 0.44 | 0.1 | 2.02 | 0.36 | 1.0 |
| (18)  | 2.38 | 0.30 | 1.0 | 0.92 | 0.63 | 0.2 | 2.16 | 0.34 | 0.0 |
| (19)  | 1.89 | 0.39 | 1.0 | 1.78 | 0.41 | 1.0 | 2.00 | 0.37 | 1.0 |
| (20)  | 2.07 | 0.35 | 1.0 | 0.39 | 0.82 | 0.0 | 2.00 | 0.37 | 1.0 |
| (21)  | 6.42 | 0.04 | 0.0 | 2.27 | 0.32 | 1.0 | 2.00 | 0.37 | 1.0 |
| (22)  | 5.42 | 0.07 | 0.0 | 2.40 | 0.30 | 0.0 | 2.00 | 0.37 | 1.0 |
| (23)  | 0.68 | 0.71 | 1.0 | 0.11 | 0.95 | 1.0 | 2.96 | 0.23 | 0.0 |
| (24)  | 2.46 | 0.29 | 0.0 | 1.03 | 0.60 | 0.0 | 1.76 | 0.42 | 0.0 |
| (25)  | 2.00 | 0.37 | 0.1 | 3.99 | 0.14 | 0.9 | 3.05 | 0.22 | 1.0 |
| (26)  | 0.17 | 0.92 | 0.1 | 3.52 | 0.17 | 0.0 | 2.03 | 0.36 | 1.0 |
| (27)  | 0.79 | 0.67 | 0.9 | 1.36 | 0.51 | 1.0 | 2.00 | 0.37 | 1.0 |
| (28)  | 2.63 | 0.27 | 0.9 | 0.81 | 0.67 | 1.0 | 2.00 | 0.37 | 1.0 |
| (29)  | 1.13 | 0.57 | 0.0 | 2.05 | 0.36 | 0.0 | 2.00 | 0.37 | 1.0 |
| (30)  | 4.31 | 0.12 | 1.0 | 2.73 | 0.25 | 1.0 | 2.00 | 0.37 | 1.0 |
Table 4: GW tests for daily models estimated with LHRV method
All daily models are estimated via LHRV method. Each cell shows the p-value of GW test and in the bracket is the ratio of choosing model A other than model B according to the decision rule.

<table>
<thead>
<tr>
<th>Model B</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(7)</th>
<th>(8)</th>
<th>(9)</th>
<th>(10)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(2)</td>
<td>0.00[1.0]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(3)</td>
<td>0.00[1.0]</td>
<td>0.00[0.9]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(4)</td>
<td>0.00[1.0]</td>
<td>0.00[1.0]</td>
<td>0.00[1.0]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(7)</td>
<td>0.00[1.0]</td>
<td>0.00[1.0]</td>
<td>0.00[1.0]</td>
<td>0.00[0.3]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(8)</td>
<td>0.00[1.0]</td>
<td>0.00[1.0]</td>
<td>0.00[1.0]</td>
<td>0.00[0.9]</td>
<td>0.00[1.0]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(9)</td>
<td>0.00[1.0]</td>
<td>0.00[1.0]</td>
<td>0.00[1.0]</td>
<td>0.00[0.7]</td>
<td>0.00[0.0]</td>
<td>0.00[0.0]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(10)</td>
<td>0.00[1.0]</td>
<td>0.00[1.0]</td>
<td>0.00[1.0]</td>
<td>0.00[0.8]</td>
<td>0.00[0.8]</td>
<td>0.00[0.2]</td>
<td>0.00[1.0]</td>
<td></td>
</tr>
<tr>
<td>(15)</td>
<td>0.00[1.0]</td>
<td>0.00[1.0]</td>
<td>0.00[1.0]</td>
<td>0.00[0.9]</td>
<td>0.00[0.9]</td>
<td>0.00[0.1]</td>
<td>0.00[0.0]</td>
<td>0.00[0.1]</td>
</tr>
<tr>
<td>(16)</td>
<td>0.00[1.0]</td>
<td>0.00[1.0]</td>
<td>0.00[1.0]</td>
<td>0.00[0.9]</td>
<td>0.00[0.9]</td>
<td>0.00[0.2]</td>
<td>0.00[1.0]</td>
<td>0.00[1.0]</td>
</tr>
<tr>
<td>(17)</td>
<td>0.00[1.0]</td>
<td>0.00[1.0]</td>
<td>0.00[1.0]</td>
<td>0.00[0.9]</td>
<td>0.00[0.9]</td>
<td>0.00[0.1]</td>
<td>0.00[0.0]</td>
<td>0.00[1.0]</td>
</tr>
<tr>
<td>(18)</td>
<td>0.00[1.0]</td>
<td>0.00[1.0]</td>
<td>0.00[1.0]</td>
<td>0.00[0.9]</td>
<td>0.00[0.9]</td>
<td>0.00[0.1]</td>
<td>0.00[0.0]</td>
<td>0.00[1.0]</td>
</tr>
<tr>
<td>(23)</td>
<td>0.00[1.0]</td>
<td>0.00[1.0]</td>
<td>0.00[1.0]</td>
<td>0.00[0.8]</td>
<td>0.00[0.0]</td>
<td>0.00[0.1]</td>
<td>0.00[0.0]</td>
<td>0.00[1.0]</td>
</tr>
<tr>
<td>(24)</td>
<td>0.00[1.0]</td>
<td>0.00[1.0]</td>
<td>0.00[1.0]</td>
<td>0.00[0.6]</td>
<td>0.00[0.9]</td>
<td>0.00[0.3]</td>
<td>0.00[1.0]</td>
<td>0.00[1.0]</td>
</tr>
<tr>
<td>(25)</td>
<td>0.00[1.0]</td>
<td>0.00[1.0]</td>
<td>0.00[1.0]</td>
<td>0.00[0.9]</td>
<td>0.00[0.0]</td>
<td>0.00[0.0]</td>
<td>0.00[1.0]</td>
<td>0.00[1.0]</td>
</tr>
<tr>
<td>(26)</td>
<td>0.00[1.0]</td>
<td>0.00[1.0]</td>
<td>0.00[1.0]</td>
<td>0.00[0.9]</td>
<td>0.00[0.9]</td>
<td>0.00[0.2]</td>
<td>0.00[1.0]</td>
<td>0.00[1.0]</td>
</tr>
</tbody>
</table>

Score 0.000 0.013 0.023 0.081 0.070 0.113 0.028 0.089

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Model A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(2)</td>
<td>0.00[0.0]</td>
<td>0.00[0.0]</td>
<td>0.00[0.0]</td>
<td>0.00[0.0]</td>
<td>0.00[0.0]</td>
<td>0.00[0.0]</td>
<td>0.00[0.0]</td>
<td></td>
</tr>
<tr>
<td>(3)</td>
<td>0.00[0.1]</td>
<td>0.00[0.0]</td>
<td>0.00[0.1]</td>
<td>0.00[0.0]</td>
<td>0.00[0.1]</td>
<td>0.00[0.0]</td>
<td>0.00[0.1]</td>
<td></td>
</tr>
<tr>
<td>(4)</td>
<td>0.00[0.1]</td>
<td>0.00[0.1]</td>
<td>0.00[0.1]</td>
<td>0.00[0.1]</td>
<td>0.00[1.0]</td>
<td>0.00[0.1]</td>
<td>0.00[1.0]</td>
<td></td>
</tr>
<tr>
<td>(7)</td>
<td>0.00[0.9]</td>
<td>0.00[0.9]</td>
<td>0.00[0.9]</td>
<td>0.00[0.9]</td>
<td>0.00[0.9]</td>
<td>0.00[0.9]</td>
<td>0.00[0.9]</td>
<td></td>
</tr>
<tr>
<td>(8)</td>
<td>0.00[1.0]</td>
<td>0.00[1.0]</td>
<td>0.00[1.0]</td>
<td>0.00[1.0]</td>
<td>0.00[1.0]</td>
<td>0.00[1.0]</td>
<td>0.00[1.0]</td>
<td></td>
</tr>
<tr>
<td>(9)</td>
<td>0.04[1.0]</td>
<td>0.00[0.0]</td>
<td>0.00[0.0]</td>
<td>0.00[0.0]</td>
<td>0.00[0.3]</td>
<td>0.00[0.0]</td>
<td>0.00[0.2]</td>
<td></td>
</tr>
<tr>
<td>(10)</td>
<td>0.00[1.0]</td>
<td>0.01[0.2]</td>
<td>0.00[1.0]</td>
<td>0.00[1.0]</td>
<td>0.00[1.0]</td>
<td>0.00[1.0]</td>
<td>0.00[1.0]</td>
<td></td>
</tr>
<tr>
<td>(15)</td>
<td>0.00[1.0]</td>
<td>0.00[0.0]</td>
<td>0.06[0.2]</td>
<td>0.00[0.0]</td>
<td>0.00[0.7]</td>
<td>0.00[0.0]</td>
<td>0.06[0.5]</td>
<td></td>
</tr>
<tr>
<td>(16)</td>
<td>0.00[1.0]</td>
<td>0.00[1.0]</td>
<td>0.00[0.3]</td>
<td>0.00[1.0]</td>
<td>0.00[0.3]</td>
<td>0.00[1.0]</td>
<td>0.00[0.8]</td>
<td>0.09[0.4]</td>
</tr>
<tr>
<td>(17)</td>
<td>0.06[0.8]</td>
<td>0.00[0.0]</td>
<td>0.00[0.0]</td>
<td>0.00[0.7]</td>
<td>0.00[0.0]</td>
<td>0.00[0.0]</td>
<td>0.00[0.6]</td>
<td>0.00[0.0]</td>
</tr>
<tr>
<td>(18)</td>
<td>0.00[1.0]</td>
<td>0.00[0.7]</td>
<td>0.00[1.0]</td>
<td>0.00[1.0]</td>
<td>0.00[0.7]</td>
<td>0.00[1.0]</td>
<td>0.00[1.0]</td>
<td>0.00[0.8]</td>
</tr>
<tr>
<td>(23)</td>
<td>0.00[0.3]</td>
<td>0.00[0.0]</td>
<td>0.00[0.3]</td>
<td>0.00[0.0]</td>
<td>0.00[0.3]</td>
<td>0.00[0.0]</td>
<td>0.00[0.4]</td>
<td>0.00[0.0]</td>
</tr>
<tr>
<td>(24)</td>
<td>0.00[1.0]</td>
<td>0.00[0.2]</td>
<td>0.00[1.0]</td>
<td>0.00[0.3]</td>
<td>0.00[1.0]</td>
<td>0.00[1.0]</td>
<td>0.00[1.0]</td>
<td>0.00[4.0]</td>
</tr>
<tr>
<td>(25)</td>
<td>0.06[0.5]</td>
<td>0.00[0.0]</td>
<td>0.00[0.4]</td>
<td>0.00[0.0]</td>
<td>0.00[0.6]</td>
<td>0.00[0.0]</td>
<td>0.00[0.0]</td>
<td></td>
</tr>
<tr>
<td>(26)</td>
<td>0.00[1.0]</td>
<td>0.09[0.6]</td>
<td>0.00[1.0]</td>
<td>0.00[0.2]</td>
<td>0.00[1.0]</td>
<td>0.00[0.6]</td>
<td>0.00[1.0]</td>
<td></td>
</tr>
</tbody>
</table>

Score 0.043 0.102 0.050 0.111 0.039 0.093 0.043 0.103
Table 5: Rank of models according to scores
Score is a good indicator of how much the model is better or worse than the others. The higher the score of a model is, the better of the predictive ability the model might be. The table lists the rank of models for different horizons according to their scores.

<table>
<thead>
<tr>
<th>Rank</th>
<th>Daily Models</th>
<th>Score</th>
<th>Weekly Models</th>
<th>Score</th>
<th>Bi-weekly Models</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>HYBRID TGARCH</td>
<td>0.113</td>
<td>FC0 HYBRID SC TGARCH</td>
<td>0.059</td>
<td>FC1 HYBRID SC GARCH</td>
<td>0.053</td>
</tr>
<tr>
<td>2</td>
<td>FC1 HYBRID SC TGARCH</td>
<td>0.111</td>
<td>FC1 HYBRID SC TGARCH</td>
<td>0.058</td>
<td>FC0 HYBRID GARCH</td>
<td>0.052</td>
</tr>
<tr>
<td>3</td>
<td>FC0 HYBRID SC TGARCH</td>
<td>0.103</td>
<td>FC0 HYBRID TGARCH</td>
<td>0.057</td>
<td>FC0 HYBRID SC GARCH</td>
<td>0.052</td>
</tr>
<tr>
<td>4</td>
<td>FC1 HYBRID TGARCH</td>
<td>0.102</td>
<td>FC1 HYBRID TGARCH</td>
<td>0.056</td>
<td>FC0 HYBRID TARCH</td>
<td>0.050</td>
</tr>
<tr>
<td>5</td>
<td>FC0 HYBRID TGARCH</td>
<td>0.093</td>
<td>HYBRID SC TGARCH</td>
<td>0.055</td>
<td>HYBRID GARCH</td>
<td>0.049</td>
</tr>
<tr>
<td>6</td>
<td>HYBRID SC TGARCH</td>
<td>0.089</td>
<td>FC0 HYBRID GARCH</td>
<td>0.054</td>
<td>FC1 HYBRID TARCH</td>
<td>0.049</td>
</tr>
<tr>
<td>7</td>
<td>SemiRV GARCH</td>
<td>0.081</td>
<td>FC1 HYBRID SC GARCH</td>
<td>0.052</td>
<td>FC0 HYBRID SC TGARCH</td>
<td>0.049</td>
</tr>
<tr>
<td>8</td>
<td>HYBRID GARCH</td>
<td>0.070</td>
<td>SemiRV GARCH</td>
<td>0.051</td>
<td>FC1 HYBRID GARCH</td>
<td>0.048</td>
</tr>
<tr>
<td>9</td>
<td>FC1 HYBRID SC GARCH</td>
<td>0.050</td>
<td>FC0 HYBRID SC GARCH</td>
<td>0.051</td>
<td>FC1 HYBRID SC TARCH</td>
<td>0.048</td>
</tr>
<tr>
<td>10</td>
<td>FC1 HYBRID GARCH</td>
<td>0.043</td>
<td>SemiRV GARCH D</td>
<td>0.050</td>
<td>HYBRID TARCH</td>
<td>0.047</td>
</tr>
<tr>
<td>11</td>
<td>FC0 HYBRID SC GARCH</td>
<td>0.043</td>
<td>FC1 HYBRID GARCH</td>
<td>0.047</td>
<td>HYBRID SC TARCH</td>
<td>0.044</td>
</tr>
<tr>
<td>12</td>
<td>FC0 HYBRID GARCH</td>
<td>0.039</td>
<td>HYBRID TGARCH</td>
<td>0.044</td>
<td>SemiRV GARCH D</td>
<td>0.043</td>
</tr>
<tr>
<td>13</td>
<td>HYBRID SC GARCH</td>
<td>0.028</td>
<td>RV GARCH</td>
<td>0.043</td>
<td>HYBRID SC GARCH</td>
<td>0.043</td>
</tr>
<tr>
<td>14</td>
<td>RV GARCH</td>
<td>0.023</td>
<td>HYBRID SC GARCH</td>
<td>0.043</td>
<td>RV GARCH</td>
<td>0.042</td>
</tr>
<tr>
<td>15</td>
<td>TGARCH</td>
<td>0.013</td>
<td>HYBRID GARCH</td>
<td>0.037</td>
<td>SemiRV GARCH</td>
<td>0.040</td>
</tr>
<tr>
<td>16</td>
<td>GARCH</td>
<td>0.000</td>
<td>FC0 HYBRID TGARCH D</td>
<td>0.026</td>
<td>FC0 HYBRID SC GARCH D</td>
<td>0.034</td>
</tr>
<tr>
<td>17</td>
<td>RV GARCH D</td>
<td>0.025</td>
<td>FC1 HYBRID SC GARCH D</td>
<td>0.025</td>
<td>FC0 HYBRID SC GARCH D</td>
<td>0.026</td>
</tr>
<tr>
<td>18</td>
<td>FC0 HYBRID SC GARCH D</td>
<td>0.025</td>
<td>FC0 HYBRID TGARCH D</td>
<td>0.020</td>
<td>RV GARCH D</td>
<td>0.023</td>
</tr>
<tr>
<td>19</td>
<td>HYBRID TGARCH D</td>
<td>0.020</td>
<td>FC0 HYBRID GARCH D</td>
<td>0.019</td>
<td>FC1 HYBRID SC TARCH D</td>
<td>0.021</td>
</tr>
<tr>
<td>20</td>
<td>FC0 HYBRID GARCH D</td>
<td>0.019</td>
<td>FC0 HYBRID TARCH D</td>
<td>0.019</td>
<td>FC0 HYBRID TARCH D</td>
<td>0.021</td>
</tr>
<tr>
<td>21</td>
<td>FC1 HYBRID GARCH D</td>
<td>0.017</td>
<td>HYBRID SC TARCH D</td>
<td>0.017</td>
<td>FC1 HYBRID TARCH D</td>
<td>0.020</td>
</tr>
<tr>
<td>22</td>
<td>FC0 HYBRID SC GARCH D</td>
<td>0.016</td>
<td>FC0 HYBRID SC GARCH D</td>
<td>0.016</td>
<td>FC0 HYBRID GARCH D</td>
<td>0.020</td>
</tr>
<tr>
<td>23</td>
<td>HYBRID GARCH D</td>
<td>0.015</td>
<td>FC1 HYBRID TARCH D</td>
<td>0.015</td>
<td>FC1 HYBRID GARCH D</td>
<td>0.019</td>
</tr>
<tr>
<td>24</td>
<td>FC1 HYBRID SC TARCH D</td>
<td>0.015</td>
<td>HYBRID SC TARCH D</td>
<td>0.014</td>
<td>FC0 HYBRID SC GARCH D</td>
<td>0.019</td>
</tr>
<tr>
<td>25</td>
<td>HYBRID SC GARCH D</td>
<td>0.014</td>
<td>FC1 HYBRID GARCH D</td>
<td>0.013</td>
<td>HYBRID GARCH D</td>
<td>0.016</td>
</tr>
<tr>
<td>26</td>
<td>GARCH</td>
<td>0.002</td>
<td>GARCH</td>
<td>0.002</td>
<td>TARCH</td>
<td>0.004</td>
</tr>
</tbody>
</table>