SCHOOL OF ECONOMICS AND MANAGEMENT
FACULTY OF SOCIAL SCIENCES
AARHUS UNIVERSITY

CREATES
Center for Research in Econometric Analysis of Time Series

CREATES Research Paper 2010-13

Realized GARCH: A Complete Model of Returns and Realized Measures of Volatility

Peter Reinhard Hansen, Zhuo (Albert) Huang
and Howard Howan Shek

School of Economics and Management
Aarhus University
Bartholins Allé 10, Building 1322, DK-8000 Aarhus C
Denmark

Realized GARCH: A Complete Model of Returns and Realized Measures of Volatility*

Peter Reinhard Hansen†
Stanford University
Department of Economics

Zhuo (Albert) Huang
Stanford University
Department of Economics

Howard Howan Shek
Stanford University
iCME

March 11, 2010

Abstract

GARCH models have been successful in modeling financial returns. Still, much is to be gained by incorporating a realized measure of volatility in these models. In this paper we introduce a new framework for the joint modeling of returns and realized measures of volatility. The Realized GARCH framework nests most GARCH models as special cases and is, in many ways, a natural extension of standard GARCH models. We pay special attention to linear and log-linear Realized GARCH specifications. This class of models has several attractive features. It retains the simplicity and tractability of the classical GARCH framework; it implies an ARMA structure for the conditional variance and realized measures of volatility; and models in this class are parsimonious and simple to estimate. A key feature of the Realized GARCH framework is a measurement equation that relates the observed realized measure to latent volatility. This equation facilitates a simple modeling of the dependence between returns and future volatility that is commonly referred to as the leverage effect. An empirical application with DJIA stocks and an exchange traded index fund shows that a simple Realized GARCH structure leads to substantial improvements in the empirical fit over to the standard GARCH model. This is true in-sample as well as out-of-sample. Moreover, the point estimates are remarkably similar across the different time series.

Keywords: GARCH; High Frequency Data; Realized Variance; Leverage Effect.

JEL Classification: C10; C22; C80

*We thank Tim Bollerslev, Giampiero Gallo, and seminar participants at Humboldt University, the Einaudi Institute, and the 2009 Stanford Institute for Theoretical Economics for valuable comments.

†Peter Hansen is also affiliated with CREATES, a center at the University of Aarhus that is funded by the Danish National Research Foundation.

1 Introduction

The latent volatility process of asset returns are relevant to a wide variety of applications, such as option pricing and risk management, and GARCH models are widely used to model the dynamic features of volatility. This has sparked the development of a large number of ARCH and GARCH models since the seminal paper by Engle (1982). Within the GARCH framework, the key element is the specification for the conditional variance. GARCH models utilize daily returns (typically squared returns) to extract information about the current level of volatility, and this information is used to form expectations about the next period's volatility. A single return is unable to offer more than a weak signal about the current level of volatility. The implication is that GARCH models are poorly suited for situations where volatility changes rapidly to a new level, because the GARCH model is slow at "catching up" and it will take many periods for the conditional variance (implied by the GARCH model) to reach its new level.

High-frequency financial data are now readily available and the literature has recently introduce a number of realized measures of volatility, including the realized variance, the bipower variation, the realized kernel, and many related quantities, see Andersen and Bollerslev (1998), Andersen, Bollerslev, Diebold, and Labys (2001), Barndorff-Nielsen and Shephard (2002), Barndorff-Nielsen and Shephard (2004), Barndorff-Nielsen, Hansen, Lunde, and Shephard (2008a), Hansen and Horel (2009), and references therein. Any of these measures is far more informative about the current level of volatility than is the squared return. This makes realized measures very useful for modeling and forecasting future volatility. Estimating a GARCH-X model that includes a realized measure in the GARCH equation provides a good illustration of this point. Such models were estimated by Engle (2002) who used the realized variance and by Barndorff-Nielsen and Shephard (2007) who used both the realized variance and the bipower variation. Within the GARCH-X framework no effort is paid to explain the variation in the realized measures, so these GARCH-X models are partial (incomplete) models that have nothing to say about returns and volatility beyond a single period into the future.

Engle and Gallo (2006) introduced the first "complete" model in this context. Their model specifies a GARCH structure for each of the realized measures, so that an additional latent volatility process is introduced for each realized measure in the model. The model by Engle and Gallo (2006) is known as the Multiplicative Error Model (MEM), because it builds on the MEM structure proposed by Engle (2002). Another complete model is the HEAVY model by Shephard and Sheppard (2010) that, in terms of its mathematical structure, is nested in the MEM framework. Unlike the traditional GARCH models, these models operate with multiple latent volatility processes. For instance, the MEM by Engle and Gallo (2006) has a total of three latent volatility processes and the HEAVY model by Shephard and Sheppard (2010) has two (or more) latent volatility processes.

In this paper we introduce a new framework that combines a GARCH structure for returns with a model for realized measures of volatility. Models within our framework are called Realized GARCH models, a name that transpires both the objective of these models (similar to GARCH) and the means by which these models operate (using realized measures). A Realized GARCH model maintains the single volatility-factor structure of the traditional GARCH framework. Instead of introducing additional latent factors, we take advantage of the natural relationship between the realized measure and the conditional variance, and we will argue that there is no need for additional factors in many cases. Consider the case where the realized measure, x_t, is a consistent estimator of the integrated variance. Now write the integrated variance as a linear combination of the conditional variance and a random innovation, and we obtain the relation $x_t = \xi + \varphi h_t + \epsilon_t$. We do not impose $\varphi = 1$ so that this approach also applies when the realized measure is compute from a shorter period (e.g. 6.5 hours) than the interval that the conditional variance refers to (e.g. 24 hours). Having a measurement equation that ties x_t to h_t has several advantages. First, it induces a simple and tractable structure that is similar to that of the classical GARCH framework. For instance, the conditional variance, the realized measure, and the squared return, all have ARMA representations. Second, the measurement equation makes it simple to model the dependence between shocks to returns and shocks to volatility, that is commonly referred to as a leverage effect. Third, the measurement equation induces a structure that is convenient for prediction. Once the model is estimated it is simple to compute distributional predictions for the future path of volatilities and returns, and these predictions do not require one to introduce auxiliary future values for the realized measure.

To illustrate our framework and fix ideas, consider a canonical version of the Realized GARCH model that will be referred to as the RealGARCH(1,1) model with a linear specification. This model is given by the following three equations

$$\begin{aligned} r_t &= \sqrt{h_t} z_t, \\ h_t &= \omega + \beta h_{t-1} + \gamma x_{t-1}, \\ x_t &= \xi + \varphi h_t + \tau(z_t) + u_t, \end{aligned}$$

where r_t is the return, $z_t \sim \text{iid}(0,1)$, $u_t \sim \text{iid}(0,\sigma_u^2)$, and $h_t = \text{var}(r_t|\mathcal{F}_{t-1})$ with $\mathcal{F}_t = \sigma(r_t, x_t, r_{t-1}, x_{t-1}, \ldots)$. The last equation relates the observed realized measure to the latent volatility, and is therefore called the measurement equation. It is easy to verify that h_t is an autoregressive process of order one, $h_t = \mu + \pi h_{t-1} + w_{t-1}$, where $\mu = \omega + \gamma\xi$, $\pi = \beta + \varphi\gamma$, and $w_t = \gamma\tau(z_t) + \gamma u_t$. So it is natural to adopt the nomenclature of GARCH (generalized autoregressive conditional heteroskedasticity) models. The inclusion of the realized measure in the model and the fact that x_t has an autoregressive moving average (ARMA) representation motivate the name Realized GARCH. A simple, yet potent specification of the leverage function is $\tau(z) = \tau_1 z + \tau_2(z^2 - 1)$, which can generate an asymmetric response in volatility to return shocks. The simple structure of the model makes the model easy to estimate and interpret, and leads to a tractable analysis of the quasi maximum likelihood estimator.

We apply the Realized GARCH framework to the DJIA stocks and an exchange traded index fund, SPY. We find, in all cases, substantial improvements in the log-likelihood function when benchmarked to a standard GARCH model. Substantial improvements are found in-sample as well as out-of-sample. The empirical evidence also strongly favors inclusion of the leverage function, and the parameter estimates are remarkably similar across stocks.

The paper is organized as follows. Section 2 introduces the Realized GARCH framework as a natural extension to GARCH. We focus on linear and log-linear specification and show that squared returns, the conditional variance, and realized measures have ARMA representations in this class of Realized GARCH models. Our Realized GARCH framework is compared to MEM and some related models in Section 3. Likelihood-based inference is analyzed in Section 4, where we derive the asymptotic properties of the QMLE estimator. Our empirical analysis is given in Section 5. We estimate a range of Realized GARCH models using time series for 28 stocks and an exchange traded index fund. Additional results and some extensions are presented in Section 6. For instance, we derive the skewness and kurtosis of returns over one or more periods, and show that the Realized GARCH is capable of generating substantial skewness and kurtosis. Concluding remarks are given in Section 7, and Appendix A presents all proofs.

2 Realized GARCH

In this section we introduce the Realized GARCH model. The key variable of interest is the conditional variance, $h_t = \mathrm{var}(r_t|\mathcal{F}_{t-1})$, where $\{r_t\}$ is a time series of returns. In the GARCH(1,1) model the conditional variance, h_t, is a function of h_{t-1} and r_{t-1}^2. In the present framework, h_t will also depend on x_{t-1}, that represents a realized measure of volatility, such as the realized variance. More generally, x_t will denote a vector of realized measures, such as the realized variance, the bipower variation, the intraday range, and the squared return. A measurement equation, that ties the realized measure to the latent volatility "completes" the model. So the Realized GARCH model fully specifies the dynamic properties of both returns and the realized measure.

To simplify the exposition we will assume $\mathrm{E}(r_t|\mathcal{F}_{t-1}) = 0$. A more general specifications for the conditional mean, such as a constant or the GARCH-in-mean by Engle et al. (1987), is accommodated by reinterpreting r_t as the return less its conditional mean. The general framework for the Realized GARCH model is presented next.

2.1 The General Formulation

The general structure of the RealGARCH(p,q) model is given by

$$r_t = \sqrt{h_t} z_t, \tag{1}$$

$$h_t = v(h_{t-1}, \dots, h_{t-p}, x_{t-1}, \dots, x_{t-q}), \tag{2}$$

$$x_t = m(h_t, z_t, u_t), \tag{3}$$

where $z_t \sim iid\,(0,1)$ and $u_t \sim iid\,(0, \sigma_u^2)$, with z_t and u_t being mutually independent.

We refer to the first two equations as the *return equation* and the *GARCH equation*, and these define a class of GARCH-X models, including those that were estimated by Engle (2002), Barndorff-Nielsen and Shephard (2007), and Visser (2008). The GARCH-X acronym refers to the the fact that x_t is treated as an exogenous variable. The HYBRID GARCH framework by Chen et al. (2009) includes variants of the GARCH-X models and some related models.

We shall refer to (3) as the *measurement equation*, because the realized measure, x_t, can often be interpreted as a measurement of h_t. The simplest example of a measurement equation is: $x_t = h_t + u_t$. The measurement equation is an important component because it "completes" the model. Moreover, the measurement equation provides a simple way to model the joint dependence between r_t and x_t, which is known to be empirically important. This dependence is modeled though the presence of z_t in the measurement equation, which we find to be highly significant in our empirical analysis.

It is worth noting that most (if not all) variants of ARCH and GARCH models are nested in the Realized GARCH framework. See Bollerslev (2009) for a comprehensive list of such models. The nesting can be achieved by setting $x_t = r_t$ or $x_t = r_t^2$, and the measurement equation is redundant for such models, because it is reduced to a simple identity. This is illustrated in the following two examples.

Example 1. By setting $x_t = r_t^2$, it is easy to verify that the RealGARCH(p,q) nests the GARCH(p,q) model. For instance with $p = q = 1$ we obtain the GARCH(1,1) structure with

$$v(h_{t-1}, r_{t-1}^2) = \omega + \alpha r_{t-1}^2 + \beta h_{t-1},$$
$$m(h_t, z_t, u_t) = h_t z_t^2.$$

The measurement equation is simply an identity in this case, so that we can take $u_t = 0$ for all t.

Example 2. If we set $x_t = r_t$, then we obtain the EGARCH(1,1) model by Nelson (1991) with

$$v(h_{t-1}, r_{t-1}) = \exp\{\omega + \alpha|z_{t-1}| + \theta z_{t-1} + \beta \log h_{t-1}\}, \quad \text{since} \quad z_{t-1} = r_{t-1}/\sqrt{h_{t-1}},$$
$$m(h_t, z_t, u_t) = \sqrt{h_t} z_t.$$

Naturally, the interesting case is when x_t is a high-frequency based realized measure, or a vector containing several realized measures. Next we consider some particular variants of the Realized GARCH model.

2.2 Realized GARCH with a Log-Linear Specification

The Realized GARCH model with a simple log-linear specification is characterized by the following GARCH and measurement equations.

$$\log h_t = \omega + \sum_{i=1}^{p} \beta_i \log h_{t-i} + \sum_{j=1}^{q} \gamma_j \log x_{t-j}, \tag{4}$$

$$\log x_t = \xi + \varphi \log h_t + \tau(z_t) + u_t, \tag{5}$$

where $z_t = r_t/\sqrt{h_t} \sim iid\,(0,1)$, $u_t \sim iid\,(0,\sigma_u^2)$, and $\tau(z)$ is called the *leverage function.*

Remark 1. A logarithmic specification for the measurement equation seems natural in this context. The reason is that (1) implies that

$$\log r_t^2 = \log h_t + \log z_t^2, \tag{6}$$

and a realized measure is in many ways similar to the squared return, r_t^2, albeit a more accurate measure of h_t. It is therefore natural to explore specifications where $\log x_t$ is expressed as a function of $\log h_t$ and z_t, such as (5). Having chosen a logarithmic form for the measurement equation, makes it convenient to specify the GARCH equation with a logarithmic form, because this induces a nice ARMA structure, as we shall see below.

Remark 2. In our empirical application we adopt a quadratic specification for the leverage function, $\tau(z_t)$. The identity (6) motivated us to explore expressions that involves $\log z_t^2$, but these were inferior to the quadratic expression, and resulted in numerical issues because zero returns are occasionally observed in practice. Additional details about the leverage function is given in Section 2.2.1.

Remark 3. The conditional variance, h_t is, by definition, adapted to $\mathcal{F}_{t-1}$. Therefore, if $\gamma \neq 0$ then x_t must also be adapted to $\mathcal{F}_t$. A filtration that would satisfy this requirement is $\mathcal{F}_t = \sigma(r_t, x_t, r_{t-1}, x_{t-1}, \ldots)$, but $\mathcal{F}_t$ could in principle be an even richer σ-field.

Remark 4. Note that the measurement equation does not require x_t to be an unbiased measure of h_t. For instance, x_t could be a realized measure that is computed with high-frequency data from a period that only spans a fraction of the period that r_t is computed over. E.g. x_t could be the realized variance for a 6.5 hour long period whereas the return, r_t, is a close-to-close return that spans 24 hours. When x_t is roughly proportional to h_t, then we should expect $\varphi \approx 1$, and that is indeed what we find empirically. Both when we use open-to-close returns and close-to-close returns.

An attractive feature of the log-linear Realized GARCH model is that it preserves the ARMA structure that characterizes some of the standard GARCH models. This shows that the "ARCH" nomenclature is appropriate for the Realized GARCH model. For the sake of generality we derive the result for the case where the GARCH equation includes lagged squared returns. Thus consider the following GARCH equation,

$$\log h_t \quad = \quad \omega + \sum\nolimits_{i=1}^{p} \beta_i \log h_{t-i} + \sum\nolimits_{j=1}^{q} \gamma_j \log x_{t-j} + \sum\nolimits_{j=1}^{q} \alpha_j \log r_{t-j}^2, \tag{7}$$

where $q = \max_i\{(\alpha_i, \gamma_i) \neq (0,0)\}$.

Proposition 1. *Define* $w_t = \tau(z_t) + u_t$ *and* $v_t = \log z_t^2 - \kappa$*, where* $\kappa = \mathrm{E}\log z_t^2$*. The Realized GARCH model* *defined by (5) and (7) implies*

$$\log h_t = \mu_h + \sum_{i=1}^{p\vee q}(\alpha_i + \beta_i + \varphi\gamma_i)\log h_{t-i} + \sum_{j=1}^{q}(\gamma_j w_{t-j} + \alpha_j v_{t-j}),$$

$$\log x_t = \mu_x + \sum_{i=1}^{p\vee q}(\alpha_i + \beta_i + \varphi\gamma_i)\log x_{t-i} + w_t + \sum_{j=1}^{p\vee q}\left\{-(\alpha_j + \beta_j)w_{t-j} + \varphi\alpha_j v_{t-j}\right\},$$

$$\log r_t^2 = \mu_r + \sum_{i=1}^{p\vee q}(\alpha_i + \beta_i + \varphi\gamma_i)\log r_{t-i}^2 + v_t + \sum_{j=1}^{p\vee q}\left\{\gamma_i(w_{t-j} - \varphi v_{t-j}) - \beta_j v_{t-j}\right\},$$

where $\mu_h = \omega + \gamma_\bullet\xi + \alpha_\bullet\kappa$, $\mu_x = \varphi(\omega + \alpha_\bullet\kappa) + (1 - \alpha_\bullet - \beta_\bullet)\xi$, *and* $\mu_r = \omega + \gamma_\bullet\xi + (1 - \beta_\bullet - \varphi\gamma_\bullet)\kappa$, *with*

$$\alpha_\bullet = \sum\nolimits_{j=1}^{q}\alpha_j, \qquad \beta_\bullet = \sum\nolimits_{i=1}^{p}\beta_i, \qquad \text{and} \qquad \gamma_\bullet = \sum\nolimits_{j=1}^{q}\gamma_j,$$

using the conventions $\beta_i = \gamma_j = \alpha_j = 0$ *for* $i > p$ *and* $j > q$.

So the log-linear Realized GARCH model implies that $\log h_t$ is ARMA$(p \vee q, q-1)$, whereas $\log r_t^2$ and $\log x_t$ are ARMA$(p\vee q, p \vee q)$. If $\alpha_1 = \cdots = \alpha_q = 0$, then $\log x_t$ is ARMA$(p \vee q, p)$.

From Proposition 1 we see that the persistence of volatility is summarized by a *persistence parameter*

$$\pi = \sum_{i=1}^{p\vee q}(\alpha_i + \beta_i + \varphi\gamma_i) = \alpha_\bullet + \beta_\bullet + \varphi\gamma_\bullet.$$

Example 3. For the case $p = q = 1$ we have

$$\log h_t = \omega + \beta\log h_{t-1} + \gamma\log x_{t-1} \quad \text{and} \quad \log x_t = \xi + \varphi\log h_t + \tau(z_t) + u_t,$$

so that $\log h_t$ ∼AR(1) and $\log x_t$ ∼ARMA(1,1). Specifically

$$\log h_t = \mu_h + \pi\log h_{t-1} + \gamma w_{t-1} \quad \text{and} \quad \log x_t = \mu_x + \pi\log x_{t-1} + w_t - \beta w_{t-1},$$

where $\pi = \beta + \varphi\gamma$.

Remark 5. The measurement equation induces a GARCH structure that is similar to an EGARCH with a stochastic volatility component. Take the case in Example 3 where $\log h_t = \mu_h + \pi\log h_{t-1} + \gamma\tau(z_{t-1}) + \gamma u_{t-1}$. Note that $\gamma\tau(z_{t-1})$ captures the leverage effects whereas γu_{t-1} adds an additional stochastic component that resembles that of stochastic volatility models. So the Realized GARCH model can induce a flexible stochastic volatility structure, similar to that in Yu (2008), but does in fact have a GARCH structure because u_{t-1} is $\mathcal{F}_{t-1}$-measurable. Interestingly, for the purpose of forecasting (beyond one-step ahead predictions), the Realized GARCH is much like a stochastic volatility model since future values of u_t are unknown. This analogy does not apply to one-step ahead predictions because the lagged values, $\tau(z_{t-1})$ and u_{t-1}, are known at time $t-1$.

An obvious advantage of using a logarithmic specification is that it automatically ensures a positive variance. Here it should be noted that the GARCH model with a logarithmic specification, known as LGARCH, see Geweke (1986), Pantula (1986), and Milhøj (1987), has some practical drawbacks. These drawbacks may explain that the LGARCH is less popular in applied work than the conventional GARCH model that uses a specification for the level of volatility, see Teräsvirta (2009). One drawback is that zero returns occasionally are observed, and such will cause havoc for the log-specification unless we impose some ad-hoc censoring. Within the Realized GARCH framework, zero returns are not problematic, because $\log r_{t-1}^2$ does not appear in its GARCH equation.

2.2.1 The Leverage Function

The function $\tau(z)$ is called the leverage function because it captures the dependence between returns and future volatility, a phenomenon that is referred to as the *leverage effect*. We normalized such functions by $\mathrm{E}\tau(z_t) = 0$, and we focus on those that have the form

$$\tau(z_t) = \tau_1 a_1(z_t) + \cdots + \tau_k a_k(z_t), \qquad \text{where} \quad \mathrm{E}a_k(z_t) = 0, \quad \text{for all } k,$$

so that the function is linear in the unknown parameters. We shall see that the leverage function induces an EGARCH type structure in the GARCH equation, and we note that the functional form used in Nelson (1991), $\tau(z_t) = \tau_1 z + \tau_+(|z_t| - \mathrm{E}|z_t|)$ is within the class of leverage functions we consider. We shall mainly consider leverage functions that are constructed from Hermite polynomials

$$\tau(z) = \tau_1 z + \tau_2(z^2 - 1) + \tau_3(z^3 - 3z) + \tau_4(z^4 - 6z^2 + 3) + \cdots,$$

and our baseline choice for the leverage function is a simple quadratic form $\tau(z_t) = \tau_1 z_t + \tau_2(z_t^2 - 1)$. This choice is convenient because it ensures that $\mathrm{E}\tau(z_t) = 0$, for any distribution of z_t, so long as $\mathrm{E}z_t = 0$ and $\mathrm{var}(z_t) = 1$. The polynomial form is also convenient in our quasi likelihood analysis, and in our derivations of the kurtosis of returns generated by this model.

The leverage function $\tau(z)$ is closely related to the *news impact curve*, see Engle and Ng (1993), that maps out how positive and negative shocks to the price affect future volatility. We can define the news impact curve by

$$\nu(z) = \mathrm{E}(\log h_{t+1} | z_t = z) - \mathrm{E}(\log h_{t+1}),$$

so that $100\nu(z)$ measures the percentage impact on volatility as a function of the studentized return. From the ARMA representation in Proposition 1 it follows that $\nu(z) = \gamma_1 \tau(z)$.

2.2.2 Multiple Realized Measurements

The Realized GARCH framework makes it simple to utilize multiple realized measures. For instance, let $x_t = (r_t^2, \mathrm{R}_t, \mathrm{RV}_t)'$, where R is the intraday range (high minus low) and RV is the realized variance. Then we could consider the following model,

$$h_t = \omega + \beta h_{t-1} + \gamma' x_t, \qquad \text{and} \qquad x_t = m(h_t, z_t, u_t) = \begin{pmatrix} h_t z_t^2 \\ \exp\{\xi_{\mathrm{R}} + \varphi_{\mathrm{R}} \log h_t + u_{\mathrm{R},t}\} \\ \exp\{\xi_{\mathrm{RV}} + \varphi_{\mathrm{RV}} \log h_t + u_{\mathrm{RV},t}\} \end{pmatrix},$$

which has a structure that is similar to the MEM by Engle and Gallo (2006). A key difference between the Realized GARCH structure and the MEM is that we only have one latent variable for volatility, h_t. The MEM has one for each of the three variables in x_t. Both RV and R are naturally tied to h_t, so it is perhaps more natural to use two measurement errors instead of two additional latent volatility variables. We revisit this issue in Section 3.

2.3 Realized GARCH with a Linear Specification

In this section we adopt a linear structure that is more similar to the original GARCH model, by Bollerslev (1986). One advantage of this formulation is that the measurement equation is simple to interpret in this model. For instance if x_t is computed from intermittent high-frequency data (i.e. over 6.5 hours) whereas r_t is a close-to-close return that spans 24 hours. Then we would expect φ to reflect how much of daily volatility that occurs during the trading hours. The linear Realized GARCH model is defined by,

$$h_t = \omega + \sum\nolimits_{i=1}^{p} \beta_i h_{t-i} + \sum\nolimits_{j=1}^{q} \gamma_j x_{t-j},$$
$$x_t = \xi + \varphi h_t + \tau(z_t) + u_t,$$

As is the case for the GARCH(1,1) model the RealGARCH(1,1) model with the linear specification implies that h_t has an AR(1) representation

$$h_t = (\omega + \gamma\xi) + (\beta + \gamma\varphi) h_{t-1} + \gamma w_{t-1},$$

where $w_t = u_t + \tau(z_t)$ is an iid process.

Analogous to the properties of squared returns in a GARCH(1,1) model, the realized measure, x_t, will have an ARMA(1,1) representation,

$$x_t = \varphi\omega + (1-\beta)\xi + (\beta + \varphi\gamma) x_{t-1} + w_t - \beta w_{t-1},$$

which is consistent with the time-series properties of realized measures in this context, see Meddahi (2003).

In this linear model, it may be more appropriate to scale the leverage function by h_t, i.e. substitute $h_t\tau(z_t)$ for $\tau(z_t)$. The reason is that the asymptotic variance for many realized measures is known to be roughly proportional to h_t^2. Alternatively, we could use the return as the argument, i.e. $\tau(r_t)$. We have experimented with measurement equations of this type, but did not find a version of the linear model that we preferred to the simple log-linear specification.

3 Comparison to MEM and HEAVY

In this section we relate the Realized GARCH model to the Multiplicative Error Model (MEM) by Engle and Gallo (2006) and the HEAVY model by Shephard and Sheppard (2010).[1]

The MEM by Engle and Gallo (2006) utilizes two realized measures in addition to the squared returns. These are the intraday range (high minus low) and the realized variance, whereas the HEAVY model by Shephard and Sheppard (2010) uses the realized kernel (RK) by Barndorff-Nielsen et al. (2008a). These models introduce an additional latent volatility process for each of the realized measures. So the MEM and the HEAVY digress from the traditional GARCH models that only have a single latent volatility factor.

In comparison to the MEM by Engle and Gallo (2006) and the HEAVY model by Shephard and Sheppard (2010), the Realized GARCH has the following characteristics.

- Maintains the single factor structure of latent volatility.
- Ties the realized measure directly to the conditional variance.
- Explicit modeling of the return-volatility dependence (leverage effect).
- Implies a simple reduced-form model for $\{r_t, h_t\}$ that is useful for forecasting, see Section 6.2.
- The Realized EGARCH model, developed in Section 6.5.1, takes one step further and decomposes the contributions from the realized measure in the GARCH equation, that yields important insight about the merits of this type of models.

Key model features are given in Table 1. We have included the level specification of the Realized GARCH model because it is most similar to the GARCH, MEM, and HEAVY models. Based on our empirical analysis in Section 5 we recommend the log-linear specification in practice.

Brownless and Gallo (2010) estimates a restricted MEM model that is closely related to the Realized GARCH with the linear specification. They utilize a single realized measure, the realized kernel by Barndorff-Nielsen et al. (2008a), so that they have two latent volatility processes, $h_t = \mathrm{E}(r_t^2|\mathcal{F}_{t-1})$ and $\mu_t = \mathrm{E}(x_t|\mathcal{F}_{t-1})$. However, their model is effectively reduced to a single factor model as they introduce the constraint, $h_t = c + d\mu_t$, see Brownless and Gallo (2010, eqs.6-7). This structure is also implied by the linear version of our measurement equation. However, they do not formulate a measurement equation or relate $x_t - \mu_t$ to a leverage function. Instead they, for the purpose of simplifying the prediction problem, adopt a simple time-varying structure, $\mu_t = a_t + b_t x_{t-1}$, where a_t and b_t are defined by spline methods. Spline methods were introduced in this context by Engle and Rangel (2008) to capture the low-frequency variation in the volatility.

[1]The Realized GARCH model was conceptualized and developed concurrently and independently of Shephard and Sheppard (2010). However, in our current presentation of the model we have adopted some terminology from Shephard and Sheppard (2010).

	Latent Variables†	Observables	Distribution‡
GARCH(1,1) (Bollerslev, 1986)	$h_t = \omega + \alpha r_{t-1}^2 + \beta h_{t-1}$	$r_t = \sqrt{h_t} z_t$	$z_t \sim \text{iid}N(0,1)$
MEM (Engle & Gallo, 2006)	$h_t = \omega + \alpha r_{t-1}^2 + \beta h_{t-1} + \delta r_{t-1} + \varphi R_{t-1}^2$ $h_{R,t} = \omega_R + \alpha_R R_{t-1}^2 + \beta_R h_{R,t-1} + \delta_R r_{t-1}$ $h_{RV,t} = \omega_{RV} + \alpha_{RV} RV_{t-1} + \beta_{RV} h_{RV,t-1}$ $+\delta_{RV} r_{t-1} + \vartheta_{RV} RV_{t-1} 1_{(r_{t-1}<0)} + \varphi_{RV} r_{t-1}^2$	$r_t^2 = h_t z_t^2$ $R_t^2 = h_{R,t} z_{R,t}^2$ $RV_t = h_{RV,t} z_{RV,t}^2$	$\begin{pmatrix} z_t \\ z_{R,t} \\ z_{RV,t} \end{pmatrix} \sim \text{iid}N(0,I)$
HEAVY (Shephard & Sheppard, 2009)	$h_t = \omega + \alpha r_{t-1}^2 + \beta h_{t-1} + \gamma x_{t-1}$ $\mu_t = \omega_R + \alpha_R x_{t-1} + \beta_R \mu_{t-1}$	$r_t = \sqrt{h_t} z_t$ $x_t = \mu_t z_{RK,t}^2$	$\begin{pmatrix} z_t \\ z_{RK,t} \end{pmatrix} \sim \text{iid}N(0,I)$
Realized GARCH (linear specification)	$h_t = \omega + \beta h_{t-1} + \gamma x_{t-1}$	$r_t = \sqrt{h_t} z_t$ $x_t = \xi + \varphi h_t + \tau(z_t) + u_t$	$\begin{pmatrix} z_t \\ \frac{u_t}{\sigma_u} \end{pmatrix} \sim \text{iid}N(0,I)$
Realized GARCH (log-linear specification)	$h_t = \exp\{\omega + \beta \log h_{t-1} + \gamma \log x_{t-1}\}$	$r_t = \sqrt{h_t} z_t$ $\log x_t = \xi + \varphi \log h_t + \tau(z_t) + u_t$	$\begin{pmatrix} z_t \\ \frac{u_t}{\sigma_u} \end{pmatrix} \sim \text{iid}N(0,I)$
Realized EGARCH (Section 6.5.1)	$h_t = \exp\{\omega + \beta \log h_{t-1} + \tau(z_{t-1}) + \delta \epsilon_{t-1}\}$	$r_t = \sqrt{h_t} z_t$ $\log x_t = \xi + \log h_{t+1} + \epsilon_t$	$\begin{pmatrix} z_t \\ \frac{\epsilon_t}{\sigma_\epsilon} \end{pmatrix} \sim \text{iid}N(0,I)$

Table 1: Key model features at a glance: The realized measures, R_t, RV_t, and x_t denote the intraday range, the realized variance, and the realized kernel, respectively. In the Realized GARCH model, the dependence between returns and innovations to the volatility (leverage effect) is modeled with $\tau(z_t)$, such as $\tau(z) = \tau_1 z + \tau_2(z^2 - 1)$, so that $\mathrm{E}\tau(z_t) = 0$, when $z_t \sim (0,1)$. †The MEM specification listed here is that selected by Engle & Gallo (2006) using BIC (see their table 4). The MEM framework permits more complex specifications. ‡The distributional assumptions listed here are those used to specify the quasi log-likelihood function. (Gaussian innovations are not essential for any of the models). The Realized EGARCH is introduce in Section 6.5.1.

One of the main advantages of Realized GARCH framework is the simplicity by which dependence between return-shocks and volatility-shocks is modeled with the leverage function. The MEM is formulated with a general dependence structure for the innovations that drive the latent volatility processes. The usual MEM formulation is based on a vector of non-negative random innovations, η_t, that are required to have mean $\mathrm{E}(\eta_t) = (1, \ldots, 1)'$. The literature has explored distributions with this property such as certain multivariate Gamma distributions, and Cipollini, Engle, and Gallo (2009) use copula methods that entail a very flexible class of distributions with the required structure. Some drawbacks of this approach are that estimation is rather complex and a rigorous analysis of the asymptotic properties of these estimators seems intractable. A perhaps simpler way to achieve the structure in the multiplicative error distribution is by setting $\eta_t = Z_t \odot Z_t$, and work with the vector of random variables random variables, Z_t, instead. The required structure can be obtained with a more traditional error structure, where each element of Z_t is required to have zero mean and unit variance. This alternative formulation can be adopted without any loss of generality, since the dependence between the elements of Z_t can take any form. The estimates in Engle and Gallo (2006) and Shephard and Sheppard (2010) are based on a likelihood where the elements of η_t are independent χ^2-distributed random variables with one degree of freedom. We have used the alternative formulation in Table 1 where $(z_t^2, z_{R,t}^2, z_{RV,t}^2)'$ corresponds to η_t in the MEM by Engle and Gallo (2006).

A related framework that is currently under development is the HYBRID GARCH models by Chen et al. (2009).

3.1 Realized GARCH Nomenclature

The Realized GARCH framework can be extended to a multi-factor structure. For instance with m realized measures (including the squared return) we could specify a model with $k \leq m$ latent volatility factors. The Realized GARCH model introduced in this paper has $k = 1$ whereas the MEM has $m = k$. This hybrid framework with $1 \leq k \leq m$, provides a way to bridge the Realized GARCH models with the MEM framework.

All these models can be viewed as extensions of standard GARCH models, where the extensions are achieved by incorporating realized measures into the model in various ways.[2] For this reason we suggest that all such models be called *Realized GARCH models.* This name transpires both the objective of these models (similar to GARCH) and the means by which these models operate (using realized measures).

4 Quasi-Maximum Likelihood Analysis

In this section we discuss the asymptotic properties of the quasi-maximum likelihood estimator within the Realized GARCH(p, q) model. The structure of the QMLE analysis is very similar to that of the standard GARCH model, see Bollerslev and Wooldridge (1992), Lee and Hansen (1994), Lumsdaine (1996), Jensen and Rahbek (2004a,b), and Straumann and Mikosch (2006). Both Engle and Gallo (2006) and Shephard and

[2]A realized measure is here used a generic for statistics that are constructed from high-frequency data, such as the realized variance, the realized kernel, intraday range, the number of transactions, and volume.

Sheppard (2010) justify the standard errors they report, by referencing existing QMLE results for GARCH models. This argument hinges on the fact that the joint log-likelihood in Engle and Gallo (2006) and Shephard and Sheppard (2010) is decomposed into a sum of univariate GARCH-X models, whose likelihood can be maximized separately. The factorization of the likelihood is achieved by two facets of these models: One is that all observables (i.e. squared return and each of the realized measures) are being tied to their individual latent volatility process. The other is that the primitive innovations in these models are taken to be independent in the formulation of the likelihood function. The latter inhibits a direct modeling of leverage effect with a function such as $\tau(z_t)$, which is one of the traits of the Realized GARCH model. However, in the MEM framework leverage type dependencies can be achieved indirectly by including suitable realized measures in various GARCH equations, such as the realized semivariance, see Barndorff-Nielsen et al. (2009b), or by introducing suitable indicator functions as in Engle and Gallo (2006).

In this section we will derive the underlying QMLE structure for the log-linear Realized GARCH model. The structure of the linear Realized GARCH model is similar. We provide closed-form expressions for the first and second derivatives of the log-likelihood function. These expressions facilitate direct computation of robust standard errors, and provide insight about regularity conditions that would justify QMLE inference. For instance, the first derivative will unearth regularity conditions that enable a central limit theorem be applied to the score function.

For the purpose of estimation, we adopt a Gaussian specification, so that the log likelihood function is given by

$$\ell(r,x;\theta) = -\frac{1}{2}\sum_{t=1}^{n}[\log(h_t) + r_t^2/h_t + \log(\sigma_u^2) + u_t^2/\sigma_u^2].$$

We write the leverage function as $\tau' a_t = \tau_1 a_1(z_t) + \cdots + \tau_k' a_k(z_t)$, and denote the parameters in the model by

$$\theta = (\lambda', \psi', \sigma_u^2)', \qquad \text{where} \quad \lambda = (\omega, \beta_1, \ldots, \beta_p, \gamma_1, \ldots, \gamma_q)', \quad \psi = (\xi, \varphi, \tau')'.$$

To simplify the notation we write $\tilde{h}_t = \log h_t$ and $\tilde{x}_t = \log x_t$, and define

$$g_t = (1, \tilde{h}_{t-1}, \ldots, \tilde{h}_{t-p}, \tilde{x}_{t-1}, \ldots, \tilde{x}_{t-q})', \qquad m_t = (1, \tilde{h}_t, a_t')'.$$

So the GARCH and measurement equations can be expresses as

$$\tilde{h}_t = \lambda' g_t \qquad \text{and} \qquad \tilde{x}_t = \psi' m_t + u_t,$$

The dynamics that underlies the score and Hessian are driven by h_t and its derivatives with respect to λ. The properties of these derivatives are stated next.

Lemma 1. *Define* $\dot{h}_t = \frac{\partial \tilde{h}_t}{\partial \lambda}$ *and* $\ddot{h}_t = \frac{\partial^2 \tilde{h}_t}{\partial \lambda \partial \lambda'}$ *. Then* $\dot{h}_s = 0$ and $\ddot{h}_s = 0$ for $s \leq 0$, and

$$\dot{h}_t = \sum_{i=1}^{p} \beta_i \dot{h}_{t-i} + g_t \qquad \text{and} \qquad \ddot{h}_t = \sum_{i=1}^{p} \beta_i \ddot{h}_{t-i} + (\dot{H}_{t-1} + \dot{H}_{t-1}'),$$

where $\dot{H}_{t-1} = \left(0_{1+p+q\times 1}, \dot{h}_{t-1}, \ldots, \dot{h}_{t-p}, 0_{1+p+q\times q}\right)$ *is an* $p+q+1 \times p+q+1$ *matrix.*

(ii) When $p = q = 1$ *we have with* $\beta = \beta_1$ that

$$\dot{h}_t = \sum_{j=0}^{t-1} \beta^j g_{t-j} \qquad \text{and} \qquad \ddot{h}_t = \sum_{k=1}^{t-1} k\beta^{k-1}(G_{t-k} + G_{t-k}'),$$

where $G_t = (0_{3\times 1}, g_t, 0_{3\times 1})$.

Proposition 2. *(i) The score,* $\frac{\partial \ell}{\partial \theta} = \sum_{t=1}^{n} \frac{\partial \ell_t}{\partial \theta}$*, is given by*

$$\frac{\partial \ell_t}{\partial \theta} = -\frac{1}{2} \begin{pmatrix} (1 - z_t^2 + \frac{2u_t}{\sigma_u^2}\dot{u}_t)\dot{h}_t \\ -\frac{2u_t}{\sigma_u^2} m_t \\ \frac{\sigma_u^2 - u_t^2}{\sigma_u^4} \end{pmatrix},$$

where $\dot{u}_t = \partial u_t / \partial \log h_t = -\varphi + \frac{1}{2} z_t \tau' \dot{a}_t$ *with* $\dot{a}_t = \partial a(z_t)/\partial z_t$.

(ii) The second derivative, $\frac{\partial^2 \ell}{\partial \theta \partial \theta'} = \sum_{t=1}^{n} \frac{\partial^2 \ell_t}{\partial \theta \partial \theta'}$*, is given by*

$$\frac{\partial^2 \ell_t}{\partial \theta \partial \theta'} = \begin{pmatrix} -\frac{1}{2}\left\{z_t^2 + \frac{2(\dot{u}_t^2 + u_t \ddot{u}_t)}{\sigma_u^2}\right\}\dot{h}_t \dot{h}_t' - \frac{1}{2}\left\{1 - z_t^2 + \frac{2u_t\dot{u}_t}{\sigma_u^2}\right\}\ddot{h}_t & \bullet & \bullet \\ \frac{\dot{u}_t}{\sigma_u^2} m_t \dot{h}_t' + \frac{u_t}{\sigma_u^2} b_t \dot{h}_t' & -\frac{1}{\sigma_u^2} m_t m_t' & \bullet \\ \frac{u_t \dot{u}_t}{\sigma_u^4} \dot{h}_t' & \frac{u_t}{\sigma_u^4} m_t' & \frac{1}{2}\frac{\sigma_u^2 - 2u_t^2}{\sigma_u^6} \end{pmatrix},$$

where $b_t = (0, 1, -\frac{1}{2} z_t \dot{a}_t')'$ and $\ddot{u}_t = -\frac{1}{4}\tau' \left\{z_t \dot{a}_t + z_t^2 \ddot{a}_t\right\}$ *with* $\ddot{a}_t = \partial^2 a(z_t)/\partial z_t^2$.

An advantage of our framework is that we can draw upon results for generalized hidden Markov models. Consider the case $p = q = 1$: From Carrasco and Chen (2002, Proposition 2) it follows that $\tilde{h}_t$ has a stationary representation provided that $\pi = \beta + \varphi\gamma \in (-1, 1)$. If we assign $\tilde{h}_0$ its invariant distribution, then $\tilde{h}_t$ is strictly stationary and β-mixing with exponential decay, and $\mathrm{E}|\tilde{h}_t|^s < \infty$ if $\mathrm{E}|\tau(z_t) + u_t|^s < \infty$. Moreover, $\{(r_t, x_t), t \geq 0\}$ is a generalized hidden Markov model, with hidden chain $\{\tilde{h}_t, t \geq 0\}$, and so by Carrasco and Chen (2002, Proposition 4) it follows that also $\{(r_t, x_t)\}$ is stationary β-mixing with exponentially decay rate. See also Straumann and Mikosch (2006) who adopt a stochastic recurrence approach to analyze the QMLE properties for a broad class of GARCH models.

The robustness of the QMLE as defined by the Gaussian likelihood is, in part, reflected by the weak assumptions that make the score a martingale difference sequence. These are stated in the following Proposition.

Proposition 3. *(i) Suppose that* $\mathrm{E}(u_t | z_t, \mathcal{F}_{t-1}) = 0$, $\mathrm{E}(z_t^2 | \mathcal{F}_{t-1}) = 1$, *and* $\mathrm{E}(u_t^2 | \mathcal{F}_{t-1}) = \sigma_u^2$. *Then* $s_t(\theta) =$

$\frac{\partial \ell_t(\theta)}{\partial \theta}$ *is a martingale difference sequence.*

(ii) Suppose, in addition, that $\{(r_t, x_t, \tilde{h}_t)\}$ *is stationary and ergodic. Then*

$$\frac{1}{\sqrt{n}} \sum_{t=1}^{n} \frac{\partial \ell_t}{\partial \theta} \xrightarrow{d} N(0, \mathcal{J}_\theta) \qquad \text{and} \qquad -\frac{1}{n} \sum_{t=1}^{n} \frac{\partial^2 \ell_t}{\partial \theta \partial \theta'} \xrightarrow{p} \mathcal{I}_\theta,$$

provided that

$$\mathcal{J}_\theta = \begin{pmatrix} \frac{1}{4}\mathrm{E}(1 - z_t^2 + \frac{2u_t}{\sigma_u^2}\dot{u}_t)^2 \mathrm{E}\left(\dot{h}_t \dot{h}_t'\right) & \bullet & \bullet \\ -\frac{1}{\sigma_u^2}\mathrm{E}\left(\dot{u}_t m_t \dot{h}_t'\right) & \frac{1}{\sigma_u^2}\mathrm{E}(m_t m_t') & \bullet \\ \frac{-E(u_t^3)E(\dot{u}_t)}{2\sigma_u^6} E(\dot{h}_t') & \frac{E(u_t^3)}{2\sigma_u^6} E(m_t') & \frac{E(u_t^2/\sigma_u^2 - 1)^2}{4\sigma_u^4} \end{pmatrix},$$

and

$$\mathcal{I}_\theta = \begin{pmatrix} \left\{\frac{1}{2} + \frac{\mathrm{E}(\dot{u}_t^2)}{\sigma_u^2}\right\} \mathrm{E}(\dot{h}_t \dot{h}_t') & \bullet & 0 \\ -\frac{1}{\sigma_u^2}\mathrm{E}\left\{(\dot{u}_t m_t + u_t b_t)\, \dot{h}_t'\right\} & \frac{1}{\sigma_u^2}\mathrm{E}(m_t m_t') & 0 \\ 0 & 0 & \frac{1}{2\sigma_u^4} \end{pmatrix},$$

are finite.

Note that in the stationary case we have $\mathcal{J}_\theta = \mathrm{E}\left(\frac{\partial \ell_t}{\partial \theta}\frac{\partial \ell_t}{\partial \theta'}\right)$, *so a necessary condition for* $|\mathcal{J}_\theta| < \infty$ *is that* z_t *and* u_t *have finite forth moments. Additional moments may be required for* z_t*, depending on the complexity of the leverage function* $\tau(z)$*, because* $\dot{u}_t$ *depends on* $\tau(z_t)$*.*

Straumann and Mikosch (2006) established conventional QMLE results for a broad class of GARCH models. The mathematical structure of the Gaussian quasi log-likelihood function for the Realized GARCH model is quite similar to the structure analyzed in Straumann and Mikosch (2006). So we conjecture that Straumann and Mikosch (2006, theorem 7.1) can be adapted to the present framework, so that

$$\sqrt{n}\left(\hat{\theta}_n - \theta\right) \to N\left(0, \mathcal{I}_\theta^{-1}\mathcal{J}_\theta\mathcal{I}_\theta^{-1}\right).$$

To make this result rigorous we would need to adapt and verify conditions N.1-N.4 in Straumann and Mikosch (2006). This is not straightforward and would take up much space, so we leave this for future research. Moreover, the results in Straumann and Mikosch (2006) only applies to the stationary case, $\pi < 1$, so the non-stationary case would have to be analyzed separately using methods similar to those in Jensen and Rahbek (2004a,b).

In the context of ARCH and GARCH models, it has been shown that the QMLE estimator is consistent with an Gaussian limit distribution regardless of the process being stationary or non-stationary. The latter was established in Jensen and Rahbek (2004a,b). So unlike the case for autoregressive processes, we need not have a discontinuity of the limit distribution at the knife-edge in the parameter space that separates stationary and non-stationary processes. This is an important result for empirical applications, because the point estimates are typically found to be very close to the boundary.

Notice that the martingale difference result for the score, Proposition 3(i), does not rely on stationarity. So a suitable central limit theorem for martingale difference processes may be applicable even if the process is non-stationary, as is the case in Jensen and Rahbek (2004a,b).

Finally, we observe that the estimator of the parameters in the GARCH equation, λ, and those of the measurement equation, ψ, are not asymptotically independent. This asymptotic correlation is induced by the leverage function in our model, and the fact that we link the realized measure, x_t, to h_t with a measurement equation.

4.1 Computational issues

While standard errors for $\hat{\theta}$ may be compute from numerical derivatives, these can also be computed directly using the following expressions

$$\hat{\mathcal{J}} = \frac{1}{n}\sum_{t=1}^{n} \hat{s}_t \hat{s}_t', \qquad \text{where} \qquad \hat{s}_t = \left\{ \tfrac{1}{2}(1 - \hat{z}_t^2 + \frac{2\hat{u}_t}{\hat{\sigma}_u^2}\hat{\dot{u}}_t)\hat{\dot{h}}_t', -\frac{\hat{u}_t}{\hat{\sigma}_u^2}\hat{m}_t', \frac{\hat{\sigma}_u^2 - \hat{u}_t^2}{2\hat{\sigma}_u^4}) \right\}',$$

and

$$\begin{aligned} \hat{\mathcal{I}} &= \frac{1}{n}\sum_{t=1}^{n} \begin{pmatrix} \frac{1}{2}\left\{\hat{z}_t^2 + \frac{2(\hat{\dot{u}}_t^2 + \hat{u}_t\hat{\ddot{u}}_t)}{\hat{\sigma}_u^2}\right\}\hat{\dot{h}}_t\hat{\dot{h}}_t' + \frac{1}{2}\left\{1 - \hat{z}_t^2 + \frac{2\hat{u}_t\hat{\dot{u}}_t}{\hat{\sigma}_u^2}\right\}\hat{\ddot{h}}_t & \bullet & \bullet \\ -\hat{\sigma}_u^{-2}\left(\hat{\dot{u}}_t\hat{m}_t + \hat{u}_t\hat{b}_t\right)\hat{\dot{h}}_t' & \frac{1}{\hat{\sigma}_u^2}\hat{m}_t\hat{m}_t' & \bullet \\ -\frac{\hat{u}_t\hat{\dot{u}}_t}{\hat{\sigma}_u^4}\hat{\dot{h}}_t' & -\frac{\hat{u}_t}{\hat{\sigma}_u^4}\hat{m}_t' & \frac{1}{2}\frac{2\hat{u}_t^2 - \hat{\sigma}_u^2}{\hat{\sigma}_u^6} \end{pmatrix} \\ &= \frac{1}{n}\sum_{t=1}^{n} \begin{pmatrix} \frac{1}{2}\left\{\hat{z}_t^2 + \frac{2(\hat{\dot{u}}_t^2 + \hat{u}_t\hat{\ddot{u}}_t)}{\hat{\sigma}_u^2}\right\}\hat{\dot{h}}_t\hat{\dot{h}}_t' + \frac{1}{2}\left\{1 - \hat{z}_t^2 + \frac{2\hat{u}_t\hat{\dot{u}}_t}{\hat{\sigma}_u^2}\right\}\hat{\ddot{h}}_t & \bullet & \bullet \\ -\hat{\sigma}_u^{-2}\left(\hat{\dot{u}}_t\hat{m}_t + \hat{u}_t\hat{b}_t\right)\hat{\dot{h}}_t' & -\frac{1}{\hat{\sigma}_u^2}\hat{m}_t\hat{m}_t' & \bullet \\ -\frac{\hat{u}_t\hat{\dot{u}}_t}{\hat{\sigma}_u^4}\hat{\dot{h}}_t' & 0 & \frac{1}{\hat{\sigma}_u^4} \end{pmatrix} \end{aligned}$$

where the zero follows from the first order condition: $\sum_{t=1}^{n} \hat{u}_t \hat{m}_t' = 0$. Moreover, the first-order conditions for λ implies that $-\sum_{t=1}^{n} \frac{\hat{u}_t\hat{\dot{u}}_t}{\hat{\sigma}_u^4}\hat{\dot{h}}' = \sum_{t=1}^{n} \frac{1-\hat{z}_t^2}{2\hat{\sigma}_u^2}\hat{\dot{h}}_t$.

For our baseline leverage function, $\tau_1 z_t + \tau_2(z_t^2 - 1)$, we have

$$m_t = \begin{pmatrix} 1 \\ \log h_t \\ z_t \\ z_t^2 - 1 \end{pmatrix}, \; b_t = \begin{pmatrix} 0 \\ 1 \\ -\frac{1}{2}z_t \\ -z_t^2 \end{pmatrix}, \quad \dot{u}_t = -\varphi + \frac{1}{2}\tau_1 z_t + \tau_2 z_t^2, \quad \ddot{u}_t = -\frac{1}{4}\tau_1 z_t - \tau_2 z_t^2.$$

The results in this section is easily generalized to specifications that include the squared return (or log-squared return) in the GARCH equation. This is achieved by stacking the appropriate lags of r_t^2 (or $\log r_t^2$) to the vector g_t.

5 Empirical Analysis

In this section we present empirical results using returns and realized measures for 28 stocks and and exchange-traded index fund, SPY, that tracks the S&P 500 index. Detailed results are presented for SPY, whereas our results for the 28 other time series are less detailed to conserve space. We adopt the realized kernel, introduced by Barndorff-Nielsen et al. (2008a), as the realized measure, x_t. We estimate the realized GARCH models using both open-to-close returns and close-to-close returns. High-frequency prices are only available between "open" and "close", so the population quantity that is estimated by the realized kernel is directly related to the volatility of open-to-close returns, but only captures a fraction of the volatility of close-to-close returns.

The results for the linear and log-linear Realized GARCH models are presented in Tables 3 and 4, respectively. Our empirical results with the level specification suggest that open-to-close volatility (the period with high-frequency prices) is about 75% of daily volatility. Then we compare the linear and log-linear specifications and argue that the latter is better suited for the problem at hand. We report empirical results for all 29 assets in Table 5 and find the point estimates to be remarkable similar across the many time series. In-sample and out-of-sample likelihood ratio statistics are computed in Table 6. These results strongly favor the inclusion of the leverage function and show that the realized GARCH framework is superior to standard GARCH models, because the partial log-likelihood of any Realized GARCH models is substantially better than that of a standard GARCH(1,1). This is found to be the case in-sample, as well as out-of-sample.

5.1 Data Description

Our sample spans the period from January 1, 2002 to August 31, 2008, which we divide into an in-sample period: January 1, 2002 to December 31, 2007; leaving the eight months, 2008-01-02 and 2008-08-31, for out-of-sample analysis.[3] We adopt the realized kernel as the realized measure, x_t, using the Parzen kernel function. This estimator is similar to the well known realized variance, but is robust to market microstructure noise and is a more accurate estimator of the quadratic variation. Our implementation of the realized kernel follows Barndorff-Nielsen, Hansen, Lunde, and Shephard (2008b) that guarantees a positive estimate, which is important for our log-linear specification. The exact computation is explained in great details in Barndorff-Nielsen, Hansen, Lunde, and Shephard (2009a). When we estimate a Realized GARCH model using open-to-close returns we should expect $x_t \approx h_t$, whereas with close-to-close returns we should expect x_t to be smaller than h_t on average.

To avoid outliers that would result from half trading days, we removed days where high-frequency data spanned less than 90% of the official 6.5 hours between 9:30am and 4:00pm. This removes about three daily observation per year, such as the day after Thanksgiving and days around Christmas. When we estimate a model that involves $\log r_t^2$, we deal with zero returns by the truncation $\max(\log r_t^2, \epsilon)$ with $\epsilon = 10^{-20}$.

[3] So our sample does not include the very volatile period during the financial crisis, because these data were not available when we initiated the empirical analysis.

Symbol	$\overline{r}_{oc}$	$\min r_{oc}$	$\max r_{oc}$	$\overline{r_{oc}^2}$	$\overline{r}_{cc}$	$\min r_{cc}$	$\max r_{cc}$	$\overline{r_{cc}^2}$	$\overline{\text{RK}}$	min RK	max RK
AA	-0.12	-8.09	8.49	3.17	-0.01	-10.91	9.24	4.44	3.56	0.49	40.52
AIG	-0.08	-12.06	11.16	3.14	-0.08	-19.90	12.04	4.43	3.00	0.13	53.44
AXP	0.03	-8.50	9.42	2.63	0.01	-10.60	10.41	3.53	2.82	0.07	57.60
BA	-0.01	-6.97	9.39	2.18	0.04	-8.41	6.78	2.91	2.46	0.21	33.92
BAC	0.02	-12.50	15.87	2.45	0.01	-10.66	20.22	3.08	2.24	0.13	62.72
C	-0.09	-12.84	16.23	2.99	0.07	-15.69	8.63	3.09	3.28	0.17	89.01
CAT	0.00	-5.52	8.18	2.23	0.07	-15.69	8.63	3.09	2.32	0.30	27.93
CVX	0.00	-6.08	5.46	1.58	0.05	-6.93	5.27	1.97	1.81	0.22	18.60
DD	-0.02	-5.96	9.84	1.65	0.01	-6.78	9.42	2.13	2.03	0.28	36.99
DIS	0.05	-6.50	8.11	2.24	0.03	-9.45	13.66	3.14	2.66	0.23	45.96
GE	-0.05	-8.38	9.90	1.77	-0.01	-13.71	9.08	2.41	1.97	0.08	36.70
GM	-0.23	-12.76	13.64	4.98	-0.07	-16.33	16.64	6.64	4.49	0.22	112.63
HD	-0.03	-5.89	11.14	2.62	-0.04	-15.18	10.20	3.62	2.88	0.18	39.69
IBM	0.06	-6.39	5.95	1.52	0.00	-10.67	10.67	2.31	1.63	0.14	19.44
INTC	-0.04	-8.22	8.82	3.68	-0.02	-20.48	10.29	5.55	3.65	0.45	44.89
JNJ	0.02	-4.68	7.92	0.95	0.02	-17.25	7.91	1.37	1.29	0.07	36.63
JPM	-0.01	-16.41	25.28	3.58	0.01	-19.95	14.88	4.61	3.74	0.10	224.45
KO	0.04	-4.39	7.51	1.00	0.01	-10.62	5.33	1.35	1.29	0.04	25.21
MCD	0.09	-11.33	6.01	1.93	0.06	-13.72	8.85	2.59	2.24	0.24	37.65
MMM	-0.01	-7.14	6.86	1.20	0.02	-9.37	6.89	1.65	1.40	0.08	17.96
MRK	0.03	-11.13	9.75	1.92	-0.02	-31.15	12.22	3.51	2.30	0.14	63.78
MSFT	-0.02	-7.71	10.98	2.04	0.00	-12.07	10.55	2.85	2.14	0.14	35.54
PG	0.11	-5.94	5.17	0.83	0.04	-7.66	4.43	1.06	1.07	0.04	12.88
T	-0.03	-11.46	8.99	2.82	0.00	-10.76	8.71	2.91	2.78	0.12	54.01
UTX	-0.02	-7.99	6.88	1.68	0.05	-9.16	9.38	2.23	1.81	0.23	25.93
VZ	-0.02	-7.63	7.12	2.03	0.00	-12.57	8.87	2.58	2.41	0.16	39.50
WMT	-0.01	-4.77	7.88	1.36	0.01	-6.89	7.73	1.87	1.75	0.16	28.78
XOM	0.03	-6.83	10.62	1.64	0.05	-8.86	9.30	2.10	1.87	0.19	26.00
SPY	-0.02	-3.98	8.19	0.88	0.01	-3.98	5.80	1.09	0.80	0.06	13.14

Table 2: Summary statistics. The sample period is January 1, 2002 to August 31, 2008. Subscript-*oc* and subscript-*cc* refer to open-to-close and close-to-close returns, respectively. The realized kernel, RK, is used as our realized measure of volatility.

5.2 Some Notation related to the Likelihood and Leverage Effect

The log-likelihood function is (conditionally on $\mathcal{F}_0 = \sigma(\{r_t, x_t, h_t\}, t \leq 0)$) given by

$$\log L(\{r_t, x_t\}_{t=1}^n; \theta) = \sum_{t=1}^n \log f(r_t, x_t | \mathcal{F}_{t-1}).$$

Standard GARCH models do not model x_t, so the log-likelihood we obtain for these models cannot be compared to those of the Realized GARCH model. However, we can factorize the joint conditional density for (r_t, x_t) by

$$f(r_t, x_t | \mathcal{F}_{t-1}) = f(r_t | \mathcal{F}_{t-1}) f(x_t | r_t, \mathcal{F}_{t-1}),$$

and compare the partial log-likelihood, $\ell(r) := \sum_{t=1}^n \log f(r_t | \mathcal{F}_{t-1})$, with that of a standard GARCH model. Specifically for the Gaussian specification for z_t and u_t, we split the joint likelihood, into the sum

$$\ell(r, x) = \underbrace{-\frac{1}{2} \sum_{t=1}^n [\log(2\pi) + \log(h_t) + r_t^2/h_t]}_{=\ell(r)} + \underbrace{-\frac{1}{2} \sum_{t=1}^n [\log(2\pi) + \log(\sigma_u^2) + u_t^2/\sigma_u^2]}_{=\ell(x|r)}.$$

Asymmetries in the leverage function are summarized by the following two statistics,

$$\rho^- = \text{corr}\{\tau(z_t) + u_t, z_t | z_t < 0\} \qquad \text{and} \qquad \rho^+ = \text{corr}\{\tau(z_t) + u_t, z_t | z_t > 0\}.$$

These capture the slope of a piecewise linear news impact curve for negative and positive returns, such as that implied by the EGARCH model.

5.3 Empirical Results for the Linear Realized GARCH Model

First we consider Realized GARCH models with the linear specification. We estimate a standard GARCH(1,1) model and six Realized GARCH models using both open-to-close and close-to-close returns for SPY. We use RG(p,q) to denote the Realized GARCH model with p lags of h_t and q lags of x_t. We estimate three models with $p = q = 2$. In addition to the standard RG$(2,2)$ model we estimate a model without the leverage function (denoted RG$(2,2)^\dagger$) and an extended model, RG$(2,2)^*$, that also includes a lag of the squared return it the GARCH equation. The results for open-to-close returns are given in the left panel of Table 3, and the corresponding results for close-to-close returns are presented in the right panel of Table 3.

First we discuss the empirical results for open-to-close returns in the left half of Table 3. First we know that the empirical estimates of φ and ξ in the measurement equation are roughly $\hat{\varphi} \simeq 1$ and $\hat{\xi} \approx 0$, which shows that the realized kernel, which is used as the realized measure of volatility, x_t, is roughly unbiased as a measure of open-to-close volatility. Comparing RG$(2,2)$ with RG$(2,2)^\dagger$ shows that the leverage function is highly significant. Omitting the two τ-parameters leads to a rather large drop in the log-likelihood function. Next, if we compare the extended model RG$(2,2)^*$ with the standard model RG$(2,2)$ we see that the ARCH

Model	Open-to-Close Returns							Close-to-Close Returns						
	G(1,1)	RG(1,1)	RG(1,2)	RG(2,1)	RG(2,2)	RG(2,2)†	RG(2,2)*	G(1,1)	RG(1,1)	RG(1,2)	RG(2,1)	RG(2,2)	RG(2,2)†	RG(2,2)*
Panel A: Point Estimates and Log-Likelihood														
ω	0.05 (0.00)	0.09 (0.05)	0.02 (0.02)	0.08 (0.05)	0.02 (0.01)	0.01 (0.01)	0.01 (0.01)	0.01 (0.01)	0.07 (0.04)	0.02 (0.02)	0.07 (0.04)	0.02 (0.02)	0.01 (0.02)	0.14 (0..07)
α	0.05 (0.01)						−0.01 (0.02)	0.05 (0.01)						0.10 (0.03)
β_1	0.95 (0.01)	0.29 (0.16)	0.80 (0.13)	0.12 (0.11)	0.83 (0.24)	0.87 (0.26)	1.01 (0.18)	0.94 (0.01)	0.29 (0.15)	0.80 (0.11)	0.13 (0.11)	0.84 (0.23)	0.85 (0.25)	0.02 (0.19)
β_2				0.18 (0.12)	−0.02 (0.14)	−0.04 (0.15)	−0.12 (0.14)				0.17 (0.13)	−0.02 (0.14)	−0.03 (0.15)	0.07 (0.11)
γ_1		0.63 (0.18)	0.63 (0.13)	0.62 (0.17)	0.63 (0.14)	0.65 (0.10)	0.64 (0.17)		0.87 (0.25)	0.87 (0.17)	0.86 (0.25)	0.86 (0.17)	0.86 (0.13)	0.82 (0.19)
γ_2			−0.45 (0.18)		−0.46 (0.18)	−0.49 (0.17)	−0.52 (0.15)			−0.62 (0.23)		−0.64 (0.24)	−0.63 (0.22)	0.08 (0.22)
ξ		−0.05 (0.09)	−0.07 (0.10)	−0.06 (0.09)	−0.07 (0.10)	−0.04 (0.06)	−0.05 (0.11)		0.00 (0.08)	−0.02 (0.07)	−0.01 (0.08)	−0.02 (0.08)	−0.02 (0.06)	−0.06 (0.09)
φ		1.01 (0.19)	1.04 (0.21)	1.03 (0.20)	1.04 (0.21)	0.99 (0.11)	1.02 (0.26)		0.74 (0.14)	0.76 (0.13)	0.74 ()	0.76 (0.14)	0.75 (0.10)	0.80 (0.15)
σ_u		0.51 (0.06)	0.50 (0.05)	0.51 (0.05)	0.50 (0.05)	0.51 (0.05)	0.50 (0.05)		0.51 (0.06)	0.50 (0.05)	0.51 (0.06)	0.50 (0.05)	0.51 (0.05)	0.48 (0.04)
τ_1		−0.02 (0.02)	−0.02 (0.02)	−0.02 (0.02)	−0.02 (0.02)		−0.02 (0.02)		−0.07 (0.02)	−0.07 (0.02)	−0.07 (0.02)	−0.07 (0.02)		−0.07 (0.02)
τ_2		0.06 (0.01)	0.06 (0.01)	0.06 (0.01)	0.06 (0.01)		0.05 (0.01)		0.03 (0.01)	0.02 (0.01)	0.02 (0..01)	0.02 (0.01)		0.03 (0.01)
$\ell(r,x)$		-2827.5	-2801.4	-2816.5	-2801.3	-2829.7	-2799.0		-2998.8	-2971.7	-2988.1	-2971.6	-2994.7	-2919.8
Panel B: Auxiliary Statistics														
π	0.992	0.929	0.988	0.941	0.989	0.990	0.996	0.990	0.929	0.989	0.941	0.990	0.990	0.915
ρ		-0.02	-0.03	-0.03	-0.05		-0.05		-0.14	-0.12	-0.14	-0.15		-0.16
ρ^-		-0.16	-0.16	-0.15	-0.19		-0.15		-0.17	-0.14	-0.16	-0.16		-0.19
ρ^+		0.12	0.15	0.11	0.13		0.10		-0.01	-0.02	-0.01	-0.02		0.00
$\ell(r)$	-1737.2	-1715.8	-1713.1	-1715.0	-1713.0	-1712.2	-1707.8	-1911.4	-1881.1	-1877.6	-1879.3	-1877.5	-1877.2	-1900.7

Table 3: Results for the linear Realized GARCH model: RG(2,2)† denotes the Realized GARCH(2,2) model without the leverage function $\tau(z)$. RG(2,2)* is the (2,2) extended to include the ARCH-term αr_{t-1}^2. The standard errors (in brackets) are robust standard errors based on the sandwich estimator $\mathcal{I}^{-1}\mathcal{J}\mathcal{I}^{-1}$.

parameter is insignificant. Consider now the auxiliary statistics in Panel B. The persistence parameter π is estimated to be close to one in all models, and the models with a leverage function all suggest a rather strong asymmetry in the new impact curve, as summarized by ρ^- and ρ^+. The partial likelihood statistic $\ell(r)$ is the likelihood for the returns alone. For the case of the Realized GARCH models this amounts to the likelihood for the GARCH-X model arising from the the return and GARCH equations alone. Note that the Realized GARCH models do not maximize this term, yet the model still produces a better empirical fit than the GARCH(1,1) model.

The empirical results for the close-to-close returns in the right half of Table 3 are quite similar. Not surprisingly are the estimates of φ smaller which reflect the fact that the realized measure only measures volatility over the open-to-close period. The point estimates are $\varphi \simeq 0.75$, which suggests that volatility during the "open period" amounts to about 75% of daily volatility. Interestingly, the ARCH parameter is found to be significant in the analysis of close-to-close returns. This finding should be taken with a grain of salt, because the linear model is grossly misspecified, as we shall see in Section 5.5, and the estimate of α in the linear model is sensitive to outliers. Note that the inclusion of α causes a large decline in the partial log-likelihood for returns, $\ell(r)$. Moreover, the estimated model suggests that volatility is far less persistent than is usually found in practice, in part because the estimates of the β-parameters are unusually small.

5.4 Empirical Results for the Log-Linear Realized GARCH Model

In this section we present detailed results for Realized GARCH models with a log-linear specification of the GARCH and measurement equations. We strongly favor the log-linear specification over the linear specification for reasons that will be evident in Section 5.5 where we compare empirical aspects of the two specifications.

5.4.1 Log-Linear Models for SPY (Table 4)

Table 4 is analogous to Table 3, except that we use the log-linear specification for the Realized GARCH models. For the sake of comparison we use the logarithmic GARCH(1,1) model and the conventional benchmark, when comparing the empirical fit of the partial likelihood (for returns). Again we report results for both open-to-close returns and close-to-close returns for SPY.

From Table 4 we see that the extended model $\text{RG}(2,2)^*$, which includes the squared return in the GARCH equation, results in very marginal improvements over the standard model $\text{RG}(2,2)$, and the ARCH parameter, α, is clearly insignificant. Comparing the $\text{RG}(2,2)^\dagger$ with the standard model shows that the leverage function is highly significant. The improvement in the log-likelihood function is almost 100 units.

The robust standard errors suggest that β_2 is significant, when it is actually not the case. This is simply a manifestation of a common problem with standard errors and t-statistics in the context with collinearity. In this case, $\log h_{t-1}$ and $\log h_{t-2}$ are highly collinear which causes the likelihood surface to be almost flat along lines where $\beta_1 + \beta_2$ is constant, while there is sufficient curvature along the axis to make the standard

Model	Open-to-Close Returns G(1,1)	RG(1,1)	RG(1,2)	RG(2,1)	RG(2,2)	RG(2,2)†	RG(2,2)*	Close-to-Close Returns G(1,1)	RG(1,1)	RG(1,2)	RG(2,1)	RG(2,2)	RG(2,2)†	RG(2,2)*
Panel A: Point Estimates and Log-Likelihood														
ω	0.04 (0.01)	0.06 (0.02)	0.04 (0.02)	0.06 (0.02)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.05 (0.00)	0.18 (0.03)	0.11 (0.02)	0.19 (0.03)	0.01 (0.01)	0.01 (0.01)	0.04 (0.05)
α	0.03 (0.01)						0.00 (0.00)	0.03 (0.00)						0.00 (0.00)
β_1	0.96 (0.01)	0.55 (0.03)	0.70 (0.05)	0.40 (0.05)	1.43 (0.04)	1.42 (0.09)	1.45 (0.05)	0.96 (0.01)	0.54 (0.03)	0.72 (0.05)	0.37 (0.05)	1.40 (0.07)	1.40 (0.10)	1.35 (0.42)
β_2				0.13 (0.05)	−0.44 (0.04)	−0.44 (0.07)	−0.46 (0.04)				0.15 (0.05)	−0.42 (0.06)	−0.43 (0.08)	−0.39 (0.30)
γ_1		0.41 (0.03)	0.45 (0.04)	0.43 (0.04)	0.46 (0.04)	0.40 (0.05)	0.42 (0.04)		0.43 (0.05)	0.48 (0.06)	0.46 (0.05)	0.45 (0.06)	0.42 (0.05)	0.46 (0.07)
γ_2			−0.18 (0.06)		−0.44 (0.04)	−0.38 (0.04)	−0.41 (0.04)			−0.21 (0.07)		−0.43 (0.05)	−0.40 (0.04)	−0.42 (0.08)
ξ		−0.18 (0.05)	−0.18 (0.05)	−0.18 (0.05)	−0.23 (0.05)	−0.16 (0.05)	−0.18 (0.05)		−0.42 (0.06)	−0.42 (0.06)	−0.42 (0.06)	−0.42 (0.05)	−0.41 (0.04)	−0.42 (0.04)
φ		1.04 (0.06)	1.04 (0.07)	1.04 (0.07)	0.96 (0.08)	1.07 (0.08)	1.03 (0.07)		0.99 (0.10)	1.00 (0.10)	0.99 (0.10)	0.99 (0.10)	1.03 (0.08)	0.99 (0.08)
σ_u		0.38 (0.08)	0.38 (0.08)	0.38 (0.08)	0.38 (0.08)	0.41 (0.08)	0.38 (0.08)		0.39 (0.08)	0.38 (0.08)	0.39 (0.08)	0.38 (0.08)	0.41 (0.08)	0.38 (0.08)
τ_1		−0.07 (0.01)	−0.07 (0.01)	−0.07 (0.01)	−0.07 (0.01)		−0.07 (0.01)		−0.11 (0.01)	−0.11 (0..01)	−0.11 (0..01)	−0.11 (0..01)		−0.11 (0..01)
τ_2		0.07 (0.01)	0.07 (0.01)	0.07 (0.01)	0.07 (0.01)		0.07 (0.01)		0.04 (0.01)	0.04 (0.01)	0.04 (0.01)	0.04 (0.01)		0.04 (0.01)
$\ell(r,x)$		-2395.6	-2388.8	-2391.9	-2385.1	-2495.7	-2382.9		-2576.9	-2567.2	-2571.7	-2563.9	-2661.7	-2563.5
Panel B: Auxiliary Statistics														
π	0.988	0.975	0.986	0.976	0.999	0.999	0.999	0.988	0.974	0.987	0.975	0.999	0.999	0.999
ρ		-0.18	-0.18	-0.16	-0.19		-0.16		-0.27	-0.25	-0.25	-0.25		-0.28
ρ^{-}		-0.33	-0.32	-0.32	-0.35		-0.35		-0.31	-0.29	-0.28	-0.28		-0.31
ρ^{+}		0.12	0.12	0.13	0.13		0.14		-0.01	-0.03	-0.03	0.03		-0.02
$\ell(r)$	-1752.7	-1712.0	-1710.3	-1711.4	-1712.3	-1708.9	-1709.6	-1938.2	-1876.5	-1875.5	-1876.1	-1875.7	-1874.9	-1876.1

Table 4: Results for the log-linear specification: G(1,1) denotes the LGARCH(1,1) model that does not utilize a realized measure of volatility. RG(2,2)† denotes the Realized GARCH(2,2) model without the $\tau(z)$ function that captures the dependence between returns and innovations in volatility. RG(2,2)* is the RG(2,2) extended to include the ARCH-term $\alpha \log r_{t-1}^2$. The latter being insignificant. The standard errors (in brackets) are robust standard errors based on the sandwich estimator $\mathcal{I}^{-1}\mathcal{J}\mathcal{I}^{-1}$.

errors small.

The empirical estimates of φ are close to unity, $\hat{\varphi} \simeq 1$, for both open-to-close and close-to-close returns. This suggests that the realized measure, x_t, is roughly proportional to the conditional variance for both open-to-close returns and close-to-close returns. The fact that ξ is estimated to be smaller (more negative) for close-to-close returns that for open-to-close returns simply reflects that the realized measure is computed over an interval that spans a shorter period than close-to-close returns.

In terms of partial log-likelihood function, $\ell(r)$, the log-linear specification leads to an even better fit than the linear specifications, whereas the logarithmic GARCH(1,1) model leads to a worse fit than the GARCH(1,1) model, see Table 3. The joint log-likelihood function, $\ell(r, x)$, for the log-linear model is not directly comparable to that of the linear model (because we are modeling $\log x_t$ as oppose to x_t).

The standard errors of the extended model, $\mathrm{RG}(2,2)^*$, for close-to-close returns are rather sensitive to the truncation parameter, ϵ, we use to avoid the problem of taking the logarithm to (a squared) zero. There are about 10 days with zero returns in the sample. The problem disappears if we use a smaller truncation parameter, but the smaller truncation parameter also causes the performance of the LGARCH to deteriorate substantially.

5.4.2 Log-Linear RealGARCH(1,2) for All Stocks (Table 5)

Table 5 shows the parameter estimates for the log-linear Realized GARCH(1,2) model for all 29 assets. The empirical results are based on open-to-close returns. We observe that the estimates are remarkably similar across the stocks that span different sectors and have varying market dynamics. An interesting observation is observed from the conditional correlations, ρ^- and ρ^+. The index fund, SPY, is found to have a strong asymmetry, since $\hat{\rho}^- = -0.32$ and $\hat{\rho}^+ = 0.13$. This is consistent with the existing literature. Also consistent with the existing literature, see e.g. Yu (2008) and reference therein, is that fact that the two conditional correlations are more balanced for the individual stocks. However, two stocks, CVX and XOM have strong asymmetries of the same magnitude as the index fund, SPY. These two stocks are both are oil companies, so a possible explanation is that the valuations of CVX and XOM were strongly influenced by the volatility of oil prices in the sample period.

5.4.3 News Impact Curve (Figure 1)

The leverage function, $\tau(z)$ is closely related to the *news impact curve* that was introduced by Engle and Ng (1993). High frequency data enable are more more detailed study of the news impact curve than is possible with daily returns. A detailed study of the news impact curve that utilizes high frequency data is Ghysels and Chen (2010). Their approach is very different from ours, yet the shape of the news impact curve they estimate is very similar to ours. The news impact curve shows how volatility is impacted by a shock to the price, and our Hermite specification for the leverage function presents a very flexible framework for estimating

	ω	β	γ_1	γ_2	ξ	φ	σ_u	τ_1	τ_2	$\ell(r)$	$\ell(r,x)$	π	ρ	ρ^-	ρ^+
AA	0.03	0.77	0.33	-0.14	-0.07	1.15	0.40	-0.04	0.09	-2776.4	-3519.9	0.98	-0.08	-0.32	0.24
AIG	0.02	0.74	0.45	-0.21	-0.06	1.02	0.45	-0.02	0.04	-2403.1	-3317.2	0.98	-0.06	-0.17	0.08
AXP	0.05	0.70	0.38	-0.12	-0.16	1.08	0.43	-0.02	0.10	-2371.1	-3217.9	0.99	-0.05	-0.30	0.25
BA	0.02	0.82	0.31	-0.17	-0.13	1.22	0.39	-0.03	0.09	-2536.0	-3260.0	0.99	-0.09	-0.36	0.26
BAC	0.00	0.78	0.51	-0.29	0.00	0.99	0.42	-0.04	0.08	-2016.9	-2823.4	0.99	-0.09	-0.31	0.21
C	-0.02	0.74	0.45	-0.19	0.09	0.99	0.39	-0.03	0.09	-2260.5	-2974.0	0.99	-0.07	-0.31	0.24
CAT	0.03	0.82	0.37	-0.22	-0.14	1.07	0.38	-0.03	0.09	-2621.1	-3279.4	0.99	-0.08	-0.32	0.27
CVX	0.03	0.71	0.33	-0.14	-0.09	1.32	0.39	-0.08	0.08	-2319.1	-3021.8	0.97	-0.19	-0.35	0.14
DD	-0.01	0.77	0.37	-0.17	0.08	1.08	0.40	-0.05	0.08	-2301.2	-3067.3	0.98	-0.13	-0.35	0.20
DIS	0.01	0.85	0.39	-0.25	-0.05	1.10	0.41	-0.04	0.09	-2518.5	-3289.6	1.00	-0.09	-0.35	0.22
GE	0.00	0.81	0.38	-0.19	0.01	0.98	0.41	-0.01	0.08	-2197.8	-2988.7	0.99	-0.02	-0.26	0.25
GM	0.06	0.84	0.39	-0.24	-0.32	1.02	0.47	-0.01	0.12	-2987.9	-3967.3	0.99	-0.01	-0.33	0.31
HD	0.01	0.79	0.39	-0.20	0.00	1.01	0.41	-0.05	0.09	-2538.4	-3318.4	0.99	-0.13	-0.37	0.20
IBM	0.00	0.74	0.41	-0.15	0.01	0.94	0.39	-0.04	0.08	-2192.6	-2896.7	0.98	-0.09	-0.32	0.24
INTC	0.02	0.87	0.46	-0.33	-0.11	1.03	0.36	-0.02	0.07	-2869.1	-3481.1	1.00	-0.05	-0.24	0.22
JNJ	-0.03	0.80	0.38	-0.19	0.13	1.04	0.44	0.02	0.10	-1874.8	-2777.3	0.99	0.04	-0.25	0.30
JPM	0.01	0.81	0.49	-0.30	-0.02	0.98	0.42	-0.04	0.09	-2463.0	-3276.8	0.99	-0.10	-0.30	0.22
KO	-0.05	0.76	0.45	-0.21	0.19	0.93	0.38	-0.02	0.07	-1886.7	-2573.6	0.99	-0.06	-0.28	0.19
MCD	0.00	0.88	0.37	-0.25	-0.01	0.98	0.45	-0.05	0.11	-2461.8	-3371.9	0.99	-0.09	-0.35	0.26
MMM	0.00	0.77	0.43	-0.23	0.02	0.98	0.41	-0.02	0.07	-2140.3	-2944.8	0.97	-0.04	-0.23	0.21
MRK	0.03	0.84	0.33	-0.21	-0.19	1.23	0.47	0.01	0.07	-2479.2	-3478.5	0.98	0.04	-0.13	0.18
MSFT	-0.01	0.79	0.44	-0.22	0.08	0.92	0.38	-0.03	0.08	-2330.7	-3021.1	0.99	-0.08	-0.31	0.24
PG	-0.04	0.78	0.43	-0.25	0.18	1.04	0.41	-0.05	0.08	-1850.7	-2646.6	0.98	-0.14	-0.32	0.14
T	0.00	0.86	0.53	-0.38	0.01	0.86	0.46	-0.03	0.10	-2560.4	-3512.7	0.99	-0.07	-0.32	0.25
UTX	-0.01	0.80	0.45	-0.24	0.06	0.88	0.40	-0.01	0.10	-2302.4	-3059.2	0.99	-0.05	-0.34	0.29
VZ	-0.01	0.79	0.40	-0.20	0.07	1.01	0.43	-0.03	0.09	-2343.4	-3196.7	0.99	-0.08	-0.31	0.23
WMT	-0.02	0.80	0.37	-0.19	0.12	1.04	0.39	-0.01	0.09	-2164.9	-2893.6	0.99	-0.02	-0.29	0.30
XOM	0.03	0.71	0.34	-0.12	-0.10	1.26	0.38	-0.08	0.08	-2334.7	-2994.1	0.98	-0.20	-0.37	0.15
SPY	0.04	0.70	0.45	-0.18	-0.18	1.04	0.38	-0.07	0.07	-1710.3	-2388.8	0.99	-0.17	-0.32	0.13
Average	0.01	0.79	0.41	-0.21	-0.02	1.04	0.41	-0.03	0.09			0.99	-0.08	-0.30	0.22

Table 5: Estimates for the log-linear Realized GARCH$(1,2)$ model.

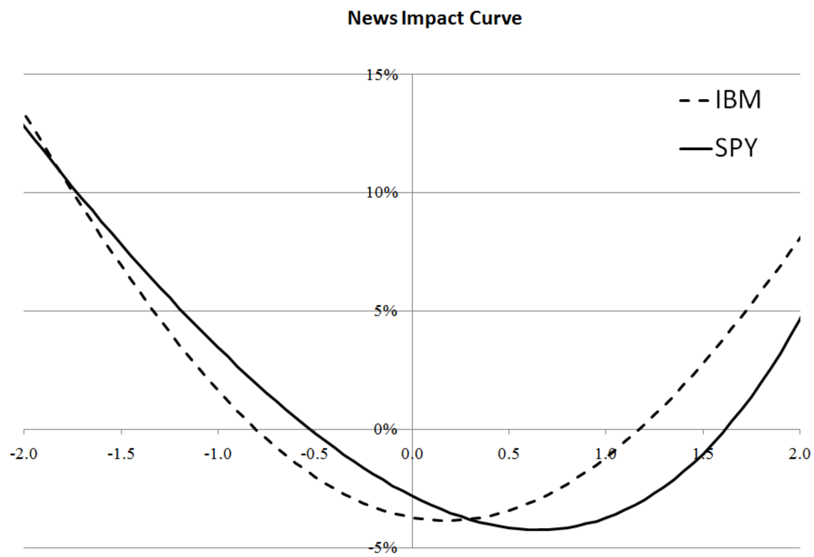

Figure 1: News impact curve for IBM and SPY

the news impact curve. In the log-linear specification we define the new impact curve by

$$\nu(z) = \mathrm{E}(\log h_{t+1}|z_t = z) - \mathrm{E}(\log h_{t+1}),$$

so that $100\nu(z)$ measures the percentage impact on volatility as a function of return-shock measures in units of standard deviations. As shown in Section 2 we have $\nu(z) = \gamma_1\tau(z)$. We have estimate the log-linear RealGARCH(1,2) model for both IBM and SPY using a flexible leverage function based on the first four Hermite polynomials. The point estimates were $(\hat{\tau}_1, \hat{\tau}_2, \hat{\tau}_3, \hat{\tau}_4) = (-0.036, 0.090, 0.001, -0.003)$ for IBM and $(\hat{\tau}_1, \hat{\tau}_2, \hat{\tau}_3, \hat{\tau}_4) = (-0.068, 0.081, 0.014, 0.002)$ for SPY. Note that the Hermite polynomials of orders three and four add little beyond the first two polynomials. The news impact curves implied by these estimates are presented in Figure 1. The fact that $\nu(z)$ is smaller than zero for some (small) values of z is an implication of its definition that implies, $\mathrm{E}[\nu(z)] = 0$.

The estimated news impact curve for IBM is more symmetric about zero than that of SPY, and this empirical result is fully consistent with the existing literature. The most common approach to model the news impact curve is to adopt a specification with a discontinuity at zero, such as that used in the EGARCH model by Nelson (1991), $\tau(z) = \tau_1 z + \tau_+(|z| - \mathrm{E}|z|)$. We also estimated the leverage functions with the piecewise linear function that leads to similar empirical results. Specifically, the implied news impact curves have the most pronounced asymmetry for the index fund, SPY, and the two oil related stocks, CVX and XOM. However, the likelihood function tends to be larger with the polynomial leverage function, $\tau(z) = \tau_1 z + \tau_2(z^2 - 1)$, and the polynomial specification simplifies aspects of the likelihood analysis.

5.4.4 In-Sample and Out-of-Sample Log-Likelihood Results

Table 6 shows the likelihood ratios for both in-sample and out-of-sample period. The statistics are based on our analysis with open-to-close returns. The statistics in Panel A are the conventional likelihood ratio statistics, where each of the five smaller models are benchmarked against the largest model. The largest model is labeled (2,2). This is the log-linear RealGARCH(2,2) model that has the squared return r_t^2 in the GARCH equation in addition to the realized measure. Thus the likelihood ratio statistics are in Panel A are defined by

$$LR_i = 2\left\{\ell_{RG(2,2)^*}(r,x) - \ell_i(r,x)\right\},$$

where i represents one of the five other Realized GARCH models. In the QMLE framework the limit distribution of likelihood ratio statistic, LR_i, is usually given as a weighted sum of χ^2-distributions. Thus comparing the LR_i to the usual critical value of a χ^2-distribution is only indicative of significance.

Comparing the RealGARCH(2,2)* to RealGARCH(2,2) leads to small LR statistics in most cases. So α tends to be insignificant in our sample. This is consistent with the existing literature that finds that squared returns adds little to the model, once a more accurate realized measure is used in the GARCH equation.

The leverage function, $\tau(z_t)$, is highly significant in all cases. The LR statistics associated with the hypothesis that $\tau_1 = \tau_2 = 0$ are well over 100 in all cases. These statistics can be computed by subtracting the statistics in the column labeled (2,2) from those in the column labeled (2,2)†. The joint hypothesis, $\beta_2 = \gamma_2 = 0$ is rejected in most cases, and so the empirical evidence does not support a simplification of the model to the RealGARCH(1,1). The results for the two hypotheses $\beta_2 = 0$ and $\gamma_2 = 0$ are less conclusive. The likelihood ratio statistics for the hypothesis, $\beta_2 = 0$ are, on average, $5.7 = 9.6 - 3.9$, which would be borderline significant when compared to conventional critical values from a $\chi^2_{(1)}$-distribution. The LR statistics for the hypothesis, $\gamma_2 = 0$, tend to be larger with an average value of 16.6. So the empirical evidence favors the RealGARCH(1,2) model over the RealGARCH(2,1) model.

Consider next the out-of-sample statistics in Panel B. These likelihood ratio statistics are computed as $\sqrt{\frac{n}{m}}\{\ell_{RG(2,2)}(r,x) - \ell_j(r,x)\}$, where n and m denote the sample sizes, in-sample and out-of-sample, respectively. The in-sample parameter estimates are simply plugged into the out-of-sample log-likelihood, and the asymptotic distribution of these statistics are non-standard because the in-sample estimates do not solve the first-order conditions out-of-sample, see Hansen (2009). The RealGARCH(2,2) model nests, or is nested in, all other models. For nested and correctly specified models where the larger model has k additional parameters that are all zero (under the null hypothesis) the out-of-sample likelihood ratio statistic is asymptotically distributed as

$$\sqrt{\frac{n}{m}}\left\{\ell_i(r,x) - \ell_j(r,x)\right\} \overset{d}{\to} Z_1' Z_2, \qquad \text{as} \quad m, n \to \infty \quad \text{with} \quad m/n \to 0,$$

where Z_1 and Z_2 are independent $Z_i \sim N_k(0, I)$. This follows from, for instance, Hansen (2009, corollary 2), and the (two-sided) critical values can be inferred from the distribution of $|Z_1' Z_2|$. For $k = 1$ the 5% and 1%

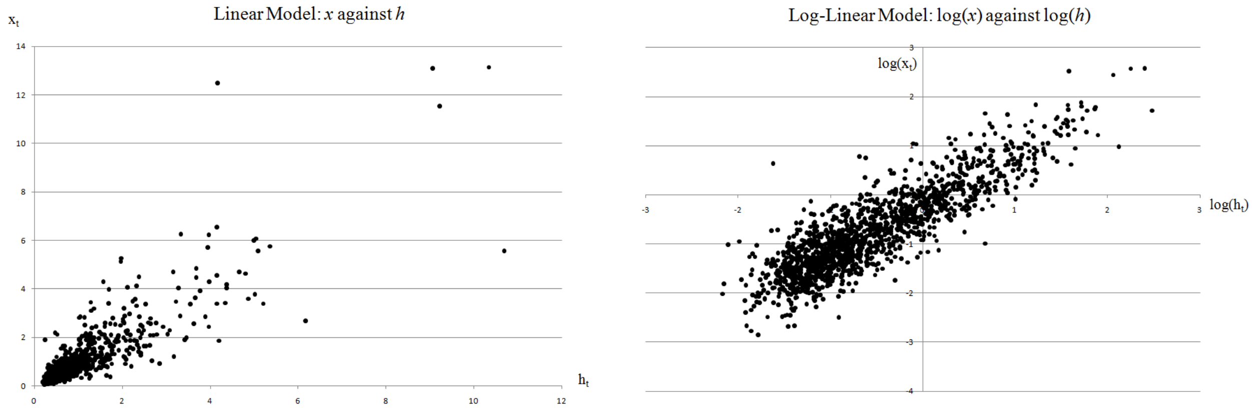

Figure 2: Heteroskedasticity in measurement equation

critical values are 2.25 and 3.67, respectively, and for two degrees of freedom ($k = 2$), these are 3.05 and 4.83, respectively. Compared to these critical values we find, on average, significant evidence in favor of a model with more lags than RealGARCH(1,1). The statistical evidence in favor of a leverage function is very strong. Adding the ARCH parameter, α, will (on average) result in a worse out-of-sample log-likelihood. As for the choice between the RealGARCH(1,2), RealGARCH(2,1), and RealGARCH(2,2) the evidence is mixed.

In Panel C, we report partial likelihood ratio statistics, that are defined by $2\{\max_i \ell_i(r|x) - \ell_j(r|x)\}$, so each model is compared with the model that had the best out-of-sample fit in term of the partial likelihood. These statistics facilitate a comparison of the Realized GARCH models with the standard GARCH(1,1) model, and we see that the Realized GARCH models also dominate the standard GARCH model in this metric. This is made more impressive by the fact that Realized GARCH models are maximizing the joint likelihood, and not the partial likelihood that is used in these comparisons.[4]

5.5 A Comparison of the Linear and Log-Linear Specifications

In this Section we focus on two empirical aspects that both favor the log-linear model over the linear model: Heteroskedasticity and the degree of misspecification of the Gaussian likelihood.

The measurement equations for the linear and log-linear model are

$$x_t = \xi + \varphi h_t + \tau(z_t) + u_t \qquad \text{and} \qquad \log x_t = \xi + \varphi \log h_t + \tau(z_t) + u_t,$$

respectively, where $\tau(z_t) + u_t$ is modeled as an iid process, hence homoskedastic.

Figure 2 displays two scatter plots. The left panel plots x_t against $\hat{h}_t$, as estimated with the linear RealGARCH(1,2) model and the right panel plots $\log x_t$ against $\log \hat{h}_t$ where the latter is estimated with a log-linear RealGARCH(1,2) model. The two models produce very similar value for h_t, however there is obviously

[4]There is not a well developed theory for the asymptotic distribution of these statistics, in part because we are comparing a model that maximizes the partial likelihood (the GARCH(1,1) model) with models that maximizes the joint likelihood (the Realized GARCH models).

	Panel A: In-Sample Likelihood Ratio						Panel B: Out-of-Sample Likelihood Ratio						Panel C: Out-of-Sample Partial Likelihood Ratio						
	(1,1)	(1,2)	(2,1)	(2,2)	$(2,2)^\dagger$	$(2,2)^*$	(1,1)	(1,2)	(2,1)	(2,2)	$(2,2)^\dagger$	$(2,2)^*$	G11	(1,1)	(1,2)	(2,1)	(2,2)	$(2,2)^\dagger$	$(2,2)^*$
AA	24.0	8.3	11.7	3.9	187.4	0	6.9	4.5	6.4	0	21.9	0.1	4.6	3.3	1.4	2.6	0.0	0.6	0.7
AIG	43.8	15.2	20.6	14.2	201.2	0	15.7	-0.2	6.0	0	25.5	12.5	56.1	9.3	5.9	7.4	6.0	0.0	7.2
AXP	30.5	25.0	26.0	0.0	197.0	0	1.2	2.7	1.5	0	13.3	0.2	24.0	0.0	0.3	0.1	1.3	1.7	1.4
BA	34.8	6.0	18.8	0.0	197.1	0	-1.4	0.4	-2.6	0	24.5	0.0	1.1	0.6	1.8	1.7	0.0	0.3	0.0
BAC	46.3	5.9	20.5	5.1	198.8	0	8.0	-0.7	-0.3	0	56.5	-2.8	147.9	4.1	0.6	0.0	1.5	19.7	0.9
C	26.8	9.5	16.0	6.3	228.4	0	1.3	-2.7	-2.5	0	-0.1	-7.2	26.9	0.3	0.9	0.5	1.1	0.0	0.4
CAT	39.8	2.3	13.8	1.6	227.3	0	1.7	-1.0	-4.2	0	9.0	2.2	47.3	0.1	0.8	0.0	1.0	1.6	1.5
CVX	17.7	0.1	3.5	0.0	194.6	0	5.3	0.0	2.2	0	20.8	-0.1	30.1	0.6	0.3	0.2	0.3	0.0	0.3
DD	31.1	10.8	15.9	6.1	167.0	0	4.6	4.1	3.0	0	-7.1	6.9	19.2	1.2	1.5	1.5	1.2	1.4	0.0
DIS	57.7	12.6	33.1	0.0	213.7	0	4.8	4.8	3.7	0	24.3	0.0	35.4	0.0	2.0	1.4	0.8	1.1	0.8
GE	37.7	12.6	20.2	12.2	200.3	0	14.1	0.6	5.9	0	9.3	0.6	41.6	0.9	0.0	0.6	0.0	0.3	0.7
GM	62.7	18.8	39.9	18.4	319.9	0	17.0	-0.2	6.8	0	7.2	14.6	57.5	1.4	0.0	0.2	0.0	0.3	0.9
HD	29.8	4.0	14.4	2.0	201.3	0	2.2	-1.0	-1.0	0	28.9	0.5	45.5	0.9	0.0	0.1	0.2	2.1	0.4
IBM	28.5	16.2	20.2	1.0	176.7	0	3.3	-0.1	1.0	0	25.0	-0.2	12.0	0.4	0.5	0.7	0.1	0.3	0.0
INTC	75.9	13.9	47.8	9.7	130.8	0	0.9	0.1	-3.4	0	36.8	19.1	133.1	1.7	0.5	2.0	0.0	0.1	1.8
JNJ	34.6	1.3	4.8	0.1	235.4	0	7.1	-0.8	2.0	0	30.7	0.2	21.8	0.4	0.0	0.1	0.1	0.1	0.0
JPM	50.7	3.5	23.6	1.9	213.5	0	0.3	-2.5	-5.4	0	-3.6	-0.6	34.9	0.0	1.3	0.1	1.6	2.0	1.6
KO	39.7	22.0	28.3	0.0	186.1	0	2.1	-2.8	-1.0	0	22.6	0.0	5.6	0.4	0.6	0.0	0.5	0.9	0.5
MCD	54.8	2.9	26.1	0.8	278.6	0	-13.0	-0.6	-9.3	0	47.4	1.4	6.8	0.0	1.5	0.2	1.8	2.0	1.7
MMM	32.0	6.7	20.4	0.1	183.5	0	-2.4	-0.2	-2.8	0	24.5	-0.1	14.5	0.0	1.9	0.9	1.4	1.0	1.4
MRK	64.6	12.0	10.6	2.9	309.0	0	-3.4	-2.5	-1.2	0	32.4	0.4	11.8	0.0	1.6	1.7	1.9	2.4	1.8
MSFT	37.3	11.5	20.5	10.2	186.1	0	8.3	4.2	7.1	0	17.0	4.1	37.1	0.1	0.5	0.4	0.0	0.2	0.1
PG	36.5	3.0	12.3	2.9	160.4	0	-1.3	0.3	-0.4	0	35.0	1.0	24.8	0.0	1.4	0.6	1.3	1.5	1.7
T	69.8	4.8	39.0	0.0	198.3	0	19.7	-0.3	13.5	0	35.1	-0.2	19.2	2.8	0.2	2.6	0.0	1.7	0.0
UTX	39.6	21.6	29.0	0.0	223.7	0	5.8	-2.0	1.2	0	33.1	-0.4	12.3	1.4	0.2	0.9	0.2	0.0	0.1
VZ	31.5	4.3	13.2	0.5	188.7	0	15.0	3.2	11.1	0	16.6	-1.3	8.7	3.5	0.7	2.8	0.3	0.0	0.1
WMT	36.2	12.0	23.0	8.3	190.0	0	-7.2	-3.8	-7.9	0	28.7	9.4	30.5	0.0	0.6	0.0	1.3	1.4	2.6
XOM	14.7	1.1	4.3	0.9	234.0	0	5.7	-0.1	1.0	0	27.9	0.6	21.6	0.0	0.5	0.3	0.5	0.5	0.5
SPY	25.3	11.6	17.9	4.2	225.6	0	6.3	-1.2	1.9	0	24.4	1.4	40.8	0.8	0.6	0.7	0.0	2.5	1.3
Average	39.8	9.6	20.5	3.9	208.8	0	4.4	0.1	1.1	0	23.0	2.1	33.5	1.2	1.0	1.0	0.8	1.6	1.0

Table 6: In-Sample and Out-of-Sample Likelihood Ratio Statistics

Standard Errors for the RealGARCH(1,2) Model						
	Linear Model			Log-linear Model		
	$\mathcal{I}^{-1}$	$\mathcal{J}^{-1}$	$\mathcal{I}^{-1}\mathcal{J}\mathcal{I}^{-1}$	$\mathcal{I}^{-1}$	$\mathcal{J}^{-1}$	$\mathcal{I}^{-1}\mathcal{J}\mathcal{I}^{-1}$
ω	0.007	0.004	0.019	0.015	0.015	0.016
β	0.034	0.017	0.125	0.040	0.031	0.053
γ_1	0.053	0.040	0.133	0.030	0.025	0.040
γ_2	0.054	0.032	0.177	0.046	0.036	0.062
ξ	0.038	0.037	0.096	0.044	0.042	0.051
φ	0.080	0.064	0.212	0.044	0.033	0.069
σ_u	0.009	0.002	0.054	0.005	0.005	0.006
τ_1	0.013	0.014	0.016	0.010	0.011	0.011
τ_2	0.008	0.013	0.011	0.006	0.008	0.006

Table 7: Conventional and robust standard errors computed for Realized GARCH(1,2) model with a quadratic leverage function. The data are open-close SPY returns.

a very pronounced degree of heteroskedasticity in the linear models. This suggests we should adopt a different leverage specification in the linear model, where $h_t\{\tau(z_t)+u_t\}$ is used in place of $\tau(z_t)+u_t$. This reinforces the value of modeling the leverage effect in this context. Homoskedastic errors are not essential for the quasi maximum likelihood estimators but causes the QMLE to be inefficient. Moreover, misspecification causes the likelihood ratio statistic to have an asymptotic distribution that is a weighted sum of $\chi^2_{(1)}$-distributed random variables, rather than a pure sum of such. Comparing likelihood ratio statistics to critical values to a standard χ^2-distribution, as an approximation, becomes very dubious when the model is highly misspecified.

In Figure 3 we present scatter plot for four variants of the Realized GARCH model. The residuals, $\{\hat{z}_t, \hat{u}_t\}_{t=1}^n$, are those for the SPY returns, obtained with the RealGARCH(1,2) model, using the four combinations of linear/log-linear specification and with/without a leverage function. The upper panels in Figure 3 are the residuals for the linear specification and the two lower panels are for the log-linear specification. The left panels are residuals obtained with out a leverage function (i.e. $\tau(z)=0$), and those on the right are the residuals obtained with a quadratic specification for $\tau(z)$. The residuals for the linear specification reveals a great deal of misspecification. The log-linear model with our a leverage function leads to residuals that strongly indicates the independence assumption between z and u is violated. The log-linear models with the quadratic leverage function offers a much better agreement with the underlying assumptions.

The fact that the log-linear model is far less misspecified than the linear model can also be illustrated by comparing robust and non-robust standard errors. In Table 7 we have computed standard errors using those of the two information matrices (the diagonal elements of $\mathcal{I}^{-1}$ and J^{-1}) and the robust standard errors computed from the diagonal of $\mathcal{I}^{-1}J\mathcal{I}^{-1}$.

5.6 The Number of Latent Volatility Factors

To illustrate that a single latent volatility factor may be sufficient in this context, we have estimated the latent volatility processes for returns, range, and the realized kernel using a simple GARCH(1,1) structure for

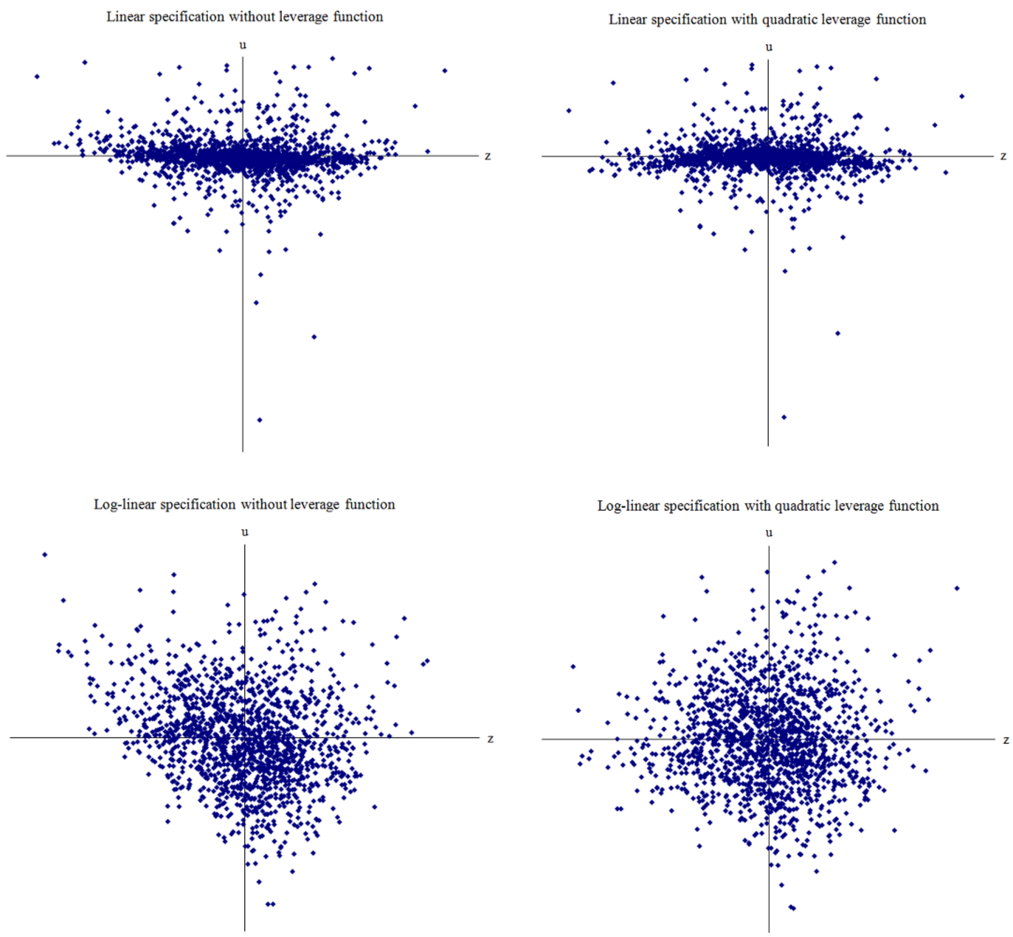

Figure 3: Scatter plots of the residuals, $(\hat{z}_t, \hat{u}_t)$, obtained with four different RealGARCH(1,2) models. The upper panels are for the linear specification and the lower panels are for the log-linear specification. The left panel are for models without a leverage function the right panels are with a quadratic leverage function. The log-linear specification with the leverage function is clearly best suited for the Gaussian structure of the quasi log-likelihood function.

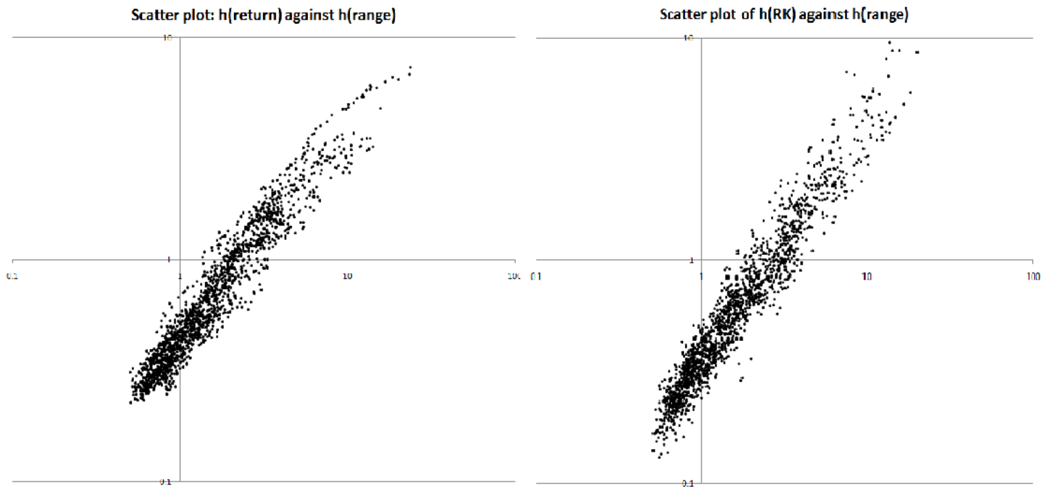

Figure 4: Scatter plots of latent volatility processes for returns, range, and the realized kernel. Each were estimated separately using a GARCH(1,1) structure. The co-linearity between these latent processes suggests that the three processes can be modeled with a single latent process.

each of them. Each of the three volatility processes were extracted by maximizing $-\frac{1}{2}\{\sum_{t=1}^{n}\log(h_t)+y_{i,t}/h_t\}$ where $h_t = \omega + \alpha y_{i,t-1} + \beta h_{t-1}$, where $y_{i,t}$ denotes either the squared return, r_t^2, the squared intraday-range, R_t^2, or the realized kernel, RK_t. For each of the three time series we maximize the quasi log-likelihood function with respect to $(\omega, \alpha, \beta, , h_0)$, so the three volatility processes are obtained separately. Figure 4 presents two scatter plots of the estimated volatility processes, and the pronounced collinearity suggests that a single latent volatility factor may be sufficient in this context.

6 Moments, Forecasting, and Insight about the Realized Measure

In this section we elaborate on the interpretations that can be deduced from the Realized GARCH model. Specifically in terms of:

- Skewness and kurtosis of cumulative returns as implied by the Realized GARCH model.
- Issues related to multi-period forecasting.
- A decomposition of the realized measure in the GARCH equation that enables us to interpret the underlying structure of the model. This decomposition motivates a modified GARCH equation that resembles that of an EGARCH model.

6.1 Properties of Cumulative Returns: Skewness and Kurtosis

We consider the skewness and kurtosis for returns generated by a Realized GARCH model. First an analytical results for a single period return for the linear Realized GARCH model.

Proposition 4. *Suppose that* $r_t = \sqrt{h_t} z_t$, *where*

$$\begin{aligned} h_t &= \omega + \alpha r_{t-1}^2 + \beta h_{t-1} + \gamma x_{t-1}, \\ x_t &= \xi + \varphi h_t + \tau(z_t) + u_t, \\ \tau(z_t) &= \tau_1 z_t + \tau_2(z_t^2 - 1) + \tau_3(z_t^3 - 3z_t) + \cdots + \tau_k H_k(z_t), \end{aligned}$$

with $z_t \sim \text{iid}N(0,1)$, $u_t \sim \text{iid}(0, \sigma_u^2)$, *and* $H_k(z_t)$ *being the* k*-th Hermite polynomial.*

Define $\pi = \alpha + \beta + \varphi\gamma$, $\mu = \omega + \gamma\xi$, $\sigma_{\tau^2}^2 = \mathrm{E}\tau(z_t)^2$, *and suppose that* $\pi^2 + 2\alpha^2 < 1$. *Then the excess kurtosis of* r_t *is given by*

$$3\frac{(1-\pi)^2}{1-\pi^2-2\alpha^2}\left(\gamma^2\frac{\sigma_u^2+\sigma_{\tau^2}^2}{\mu^2} + 4\gamma\frac{\alpha\tau_2}{\mu(1-\pi)}\right) + \frac{6\alpha^2}{1-\pi^2-2\alpha^2}.$$

In the special case were $\alpha = 0$, *the excess kurtosis is*

$$3\frac{1-\pi}{1+\pi}\gamma^2\frac{\sigma_{\tau^2}^2+\sigma_u^2}{(\omega+\gamma\xi)^2},$$

and in the special case where $\gamma = 0$ *we obtain the excess kurtosis for the GARCH(1,1) model,*

$$\frac{6\alpha^2}{1-(\alpha+\beta)^2-2\alpha^2}.$$

When $z_t \sim N(0,1)$ and the leverage function is constructed from Hermite polynomials, $\tau(z_t) = \tau_1 z_t + \tau_2(z_t^2 - 1) + \tau_3(z_t^3 - 3z_t) + \cdots$, then $\sigma_{\tau^2}^2 = \tau_1^2 + 2\tau_2^2 + 6\tau_3^2 + 4!\tau_4^2 + \cdots$.

For the log-linear Realized GARCH model we have the following results for the kurtosis of a single period return.

Proposition 5. *Consider the log-linear RealGARCH(1,1) model and define* $\pi = \beta + \varphi\gamma$ *and* $\mu = \omega + \varphi\xi$, *so that*

$$\log h_t = \pi \log h_{t-1} + \mu + \gamma w_{t-1}, \qquad \textit{where} \qquad w_t = \tau_1 z_t + \tau_2(z_t^2 - 1) + u_t,$$

with $z_t \sim \text{iid}N(0,1)$ *and* $u_t \sim \text{iid}N(0, \sigma_u^2)$. *The kurtosis of the return* $r_t = \sqrt{h_t} z_t$ *is given by*

$$\frac{\mathrm{E}(r_t^4)}{\mathrm{E}(r_t^2)^2} = 3\left(\prod_{i=0}^{\infty}\frac{1-2\pi^i\gamma\tau_2}{\sqrt{1-4\pi^i\gamma\tau_2}}\right)\exp\left\{\sum_{i=0}^{\infty}\frac{\pi^{2i}\gamma^2\tau_1^2}{1-6\pi^i\gamma\tau_2+8\pi^{2i}\gamma^2\tau_2^2}\right\}\exp\left\{\frac{\gamma^2\sigma_u^2}{1-\pi^2}\right\}. \tag{8}$$

There does not appear to be a way to further simplify the expression (8), however when $\gamma\tau_2$ is small, as we found it to be empirically, we have the approximation (see the appendix for details)

$$\frac{\mathrm{E}(r_t^4)}{\mathrm{E}(r_t^2)^2} \simeq 3\exp\left\{\frac{\gamma^2\tau_2^2}{-\log\pi} + \frac{\gamma^2(\tau_1^2+\sigma_u^2)}{1-\pi^2}\right\}. \tag{9}$$

The skewness for single period returns is non-zero, if and only if the studentized return, z_t, has non-zero

ω	=	0.04124604
β_1	=	0.70122085
γ_1	=	0.45067217
γ_2	=	-0.17604791
ξ	=	-0.17999580
φ	=	1.03749403
σ_u	=	0.38127405
τ_1	=	-0.06781023
τ_2	=	0.07015828

Table 8: Parameter estimates for the log-linear Realized GARCH(1,2) model that is used to simulate cumulative returns.

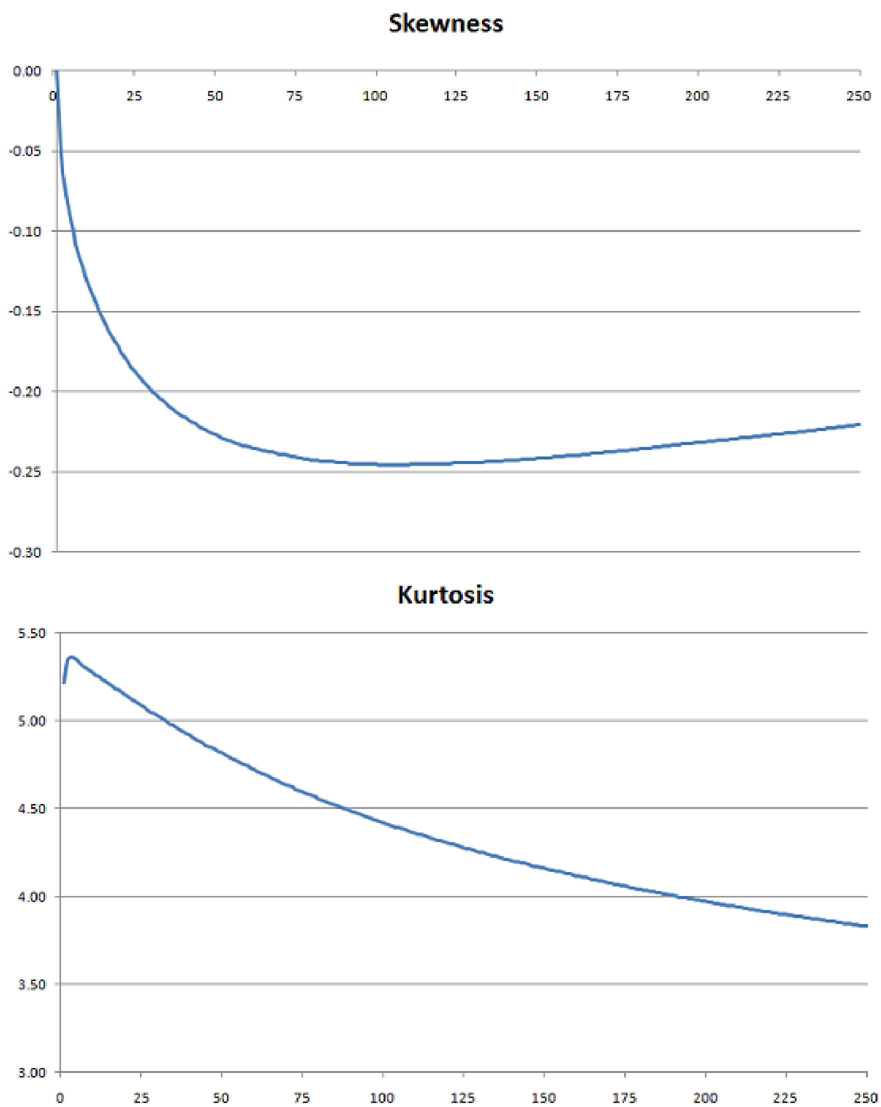

Figure 5: Skewness and kurtosis of cumulative returns from a Realized GARCH model with the log-linear specification, see Table 8. The x-axis gives the number of periods that returns are accumulated over.

skewness. This follows directly from the identity $r_t = \sqrt{h_t} z_t$, and the assumption that $z_t \perp\!\!\!\perp h_t$, that shows that,

$$\mathrm{E}(r_t^d) = \mathrm{E}(h_t^{d/2} z_t^d) = \mathrm{E}\left\{\mathrm{E}(h_t^{d/2} z_t^d | \mathcal{F}_{t-1})\right\} = \mathrm{E}(h_t^{d/2})\mathrm{E}(z_t^d),$$

and in particular that $\mathrm{E}(r_t^3) = \mathrm{E}(h_t^{3/2})\mathrm{E}(z_t^3)$. So a symmetric distribution for z_t implies that r_t has zero skewness, and this is property that is shared by standard GARCH model and Realized GARCH model alike.

For the skewness and kurtosis of cumulative returns, $r_t + \cdots + r_{t+k}$, the situation is very different, because the leverage function induces skewness. For this problem we resort to simulation methods using a design based on our empirical estimates for log-linear Realized GARCH(1,2) model that we obtained for the SPY returns. The exact configuration is given by Table 8, and the skewness and kurtosis of cumulative returns are shown in Figure 5. From Figure 5 it is evident that the Realized GARCH model can produce strong and persistent skewness and kurtosis.

6.2 Multi-Period Forecast

One of the main advantages of having a complete specification, i.e., a model that fully describes the dynamic properties of x_t is that multi-period ahead forecasting is feasible. In contrast, the GARCH-X model can only be used to make one-step ahead predictions. Multi-period ahead predictions are not possible without a model for x_t, such as the one implied by the measurement equation in the Realized GARCH model.

Multi-period ahead predictions with the Realized GARCH model is straightforward for both the linear and log-linear Realized GARCH models. Let $\tilde{h}_t$ denote either h_t or $\log h_t$, and consider first the case where $p = q = 1$. By substituting the GARCH equation into measurement equation we obtain the VARMA(1,1) structure

$$\begin{bmatrix} \tilde{h}_t \\ \tilde{x}_t \end{bmatrix} = \begin{bmatrix} \beta & \gamma \\ \varphi\beta & \varphi\gamma \end{bmatrix} \begin{bmatrix} \tilde{h}_{t-1} \\ \tilde{x}_{t-1} \end{bmatrix} + \begin{bmatrix} \omega \\ \xi + \varphi\omega \end{bmatrix} + \begin{bmatrix} 0 \\ \tau(z_t) + u_t \end{bmatrix},$$

that can be used to generate the predictive distribution of future values of $\tilde{h}_t$, $\tilde{x}_t$, as well as returns r_t, using

$$\begin{bmatrix} \tilde{h}_{t+h} \\ \tilde{x}_{t+h} \end{bmatrix} = \begin{bmatrix} \beta & \gamma \\ \varphi\beta & \varphi\gamma \end{bmatrix}^h \begin{bmatrix} \tilde{h}_t \\ \tilde{x}_t \end{bmatrix} + \sum_{j=0}^{h-1} \begin{bmatrix} \beta & \gamma \\ \varphi\beta & \varphi\gamma \end{bmatrix}^j \left\{ \begin{bmatrix} \omega \\ \xi + \varphi\omega \end{bmatrix} + \begin{bmatrix} 0 \\ \tau(z_{t+h-j}) + u_{t+h-j} \end{bmatrix} \right\}.$$

This is easily extended to the general case $(p, q \geq 1)$ where we have

$$Y_t = AY_{t-1} + b + \epsilon_t,$$

with the conventions

$$Y_t = \begin{bmatrix} \tilde{h}_t \\ \vdots \\ \tilde{h}_{t-p+1} \\ \tilde{x}_t \\ \vdots \\ \tilde{x}_{t-q+1} \end{bmatrix}, \quad A = \begin{pmatrix} (\beta_1, \dots, \beta_p) & (\gamma_1, \dots, \gamma_q) \\ (I_{p-1\times p-1}, 0_{p-1\times 1}) & 0_{p-1\times q} \\ \varphi(\beta_1, \dots, \beta_p) & \varphi(\gamma_1, \dots, \gamma_q) \\ 0_{q-1\times p} & (I_{q-1\times q-1}, 0_{q-1\times 1}) \end{pmatrix}, \quad b = \begin{pmatrix} \omega \\ 0_{p-1\times 1} \\ \xi + \varphi\omega \\ 0_{q-1\times 1} \end{pmatrix}, \quad \epsilon_t = \begin{bmatrix} 0_{p\times 1} \\ \tau(z_t) + u_t \\ 0_{q\times 1} \end{bmatrix},$$

so that

$$Y_{t+h} = A^h Y_t + \sum_{j=0}^{h-1} A^j (b + \epsilon_{t+h-j}).$$

The predictive distribution for $\tilde{h}_{t+h}$ and/or $\tilde{x}_{t+h}$, is given from the distribution of $\sum_{i=0}^{h-1} A^i \epsilon_{t+h-i}$, which also enables us to compute a predictive distribution for r_{t+h}, and cumulative returns $r_{t+1} + \cdots + r_{t+h}$.

The Realized GARCH model can be used to predict both the conditional return-variance and the realized measure. The latter has been the subject of a very active literature. See e.g. Andersen, Bollerslev, Diebold, and Labys (2003), Andersen, Bollerslev, and Meddahi (2004, 2005, 2010),Andersen, Bollerslev, and Diebold (2007).

6.2.1 A Reduced-Form Expression

If it is not an objective to predict future values of the realized measure,$\tilde{x}_t$, then the Realized GARCH implies another structure that is simpler for predicting the future path of volatility and returns. From the ARMA structure for $\tilde{h}_t$ we have that

$$\tilde{h}_t = \mu + \pi_1 \tilde{h}_t + \cdots + \pi_{\tilde{p}} \tilde{h}_{t-\tilde{p}} + \gamma_1 \{\tau(z_{t-1}) + u_{t-1}\} + \cdots + \gamma_q \{\tau(z_{t-q}) + u_{t-q}\},$$

where $\mu = \omega + \gamma_\bullet \xi$ and $\pi_j = \beta_j + \varphi\gamma_j$, for $j = 1, \dots, \tilde{p} = \max(p, q)$.

Hence, if the objective is to predict the future path of volatility and/or the predictive distribution of future returns, including multi-period returns such as, $r_{t+1} + \cdots + r_{t+h}$, then there is no need to consider axillary future values of the realized measure. Note the mixture of z_t and u_t innovations that defines the volatility path. This structure resembles the dynamic features of a stochastic volatility model with leverage effect, see e.g. Yu (2008).

6.3 Decomposing the Realized Measure in the GARCH Equation

In this section we provide a more detailed analysis of the leverage term, $\tau(z_t)$, and its dynamic effect on volatility. First we consider a hypothetical decomposition of the realized measure in the GARCH equation.

This yields valuable insight about the gains from utilizing realized measures in these models. Then we provide an alternative (but econometrically equivalent) representation of the GARCH equation. This representation suggests a simple extension of the Realized GARCH model, that offers a more flexible specification of the leverage effect. We also study a different functional form for $\tau(z)$, that induces an EGARCH structure on the GARCH equation.

6.4 A Hypothetical Decomposition

Many realized measures, such as the realized kernel used in our empirical analysis, will be consistent estimators of the quadratic variation, which is an ex-post measure of volatility.

Consider the case where x_t is an estimator of the integrated variance IV_t, such as the realized variance or the realized kernel. For realized measures of this type it is well known that the *sampling error*,

$$\eta_t = \log x_t - \log \text{IV}_t,$$

is approximately $N(0, \Sigma_{n_t})$, where $\Sigma_{n_t} \to 0$ as the number of intraday observations, $n_t \to \infty$. See e.g. Barndorff-Nielsen and Shephard (2002), Barndorff-Nielsen et al. (2008a), and Barndorff-Nielsen et al. (2008b). For instance, under weak assumptions about market microstructure noise Barndorff-Nielsen et al. (2008b) show that the "sampling error" is $\log x_t - \log \text{IV}_t = O_p(n^{-1/5})$.

The difference between the logarithmically transformed integrated variance and conditional variance is given by

$$\zeta_t = \log \text{IV}_t - \log h_t.$$

Then ζ_t captures the news about volatility that accumulated after the conditional expectation, h_t, is made at time $t-1$, and we will refer to ζ_t as the *volatility shock*. Naturally the expected value of ζ_t will depend on whether the integrated variance is measured over the same period as h_t, or a fraction thereof.

Suppose for simplicity that $\varphi = 1$ and $p = q = 1$, so that the GARCH equation can be expressed as

$$\log h_{t+1} = \mu + \pi \log h_t + \delta_1 \zeta_t + \delta_2 \eta_t.$$

Within the Realized GARCH model γ has to represent both δ_1 and δ_2, so the Realized GARCH model implicitly imposes the constraint that $\delta_1 = \delta_2$ $(= \gamma)$.

The sampling error, η_t, will be specific to the choice of realized measure (estimator of integrated variance). Since this term reflects our inability to perfectly estimate the integrated variance, we should not expect this term to be important for describing the dynamics of volatility. We should expect the volatility shock, ζ_t, to be important. Since neither ζ_t nor η_t are observed we cannot estimate a model where the realized measure is decomposed into these two terms. However, we can relate $\tau(z_t)$ and u_t to these terms, as we discuss next.

6.5 An Alternative Model Representation

An alternative representation for the RealGARCH(1,1) is

$$\begin{aligned} r_t &= \sqrt{h_t} z_t, \\ \log h_t &= \mu + \pi \log h_{t-1} + \gamma \tau(z_{t-1}) + \gamma u_{t-1}, \\ \log x_t &= \xi + \varphi \log h_t + \tau(z_t) + u_t, \end{aligned}$$

where $\pi = \beta + \varphi\gamma$ and $\mu = \omega + \gamma\xi$. The model defined by these equations will generate a process $(r_t, x_t)'$ that is observationally equivalent to log-linear RealGARCH(1,1). Note that the inclusion of x_t in the GARCH equation implicitly imposes that the coefficients associated with $\tau(z)$ and u be the same. This constraint is relaxed in the following specification,

$$\begin{aligned} \log h_t &= \mu + \pi \log h_{t-1} + \delta_1 \tau(z_{t-1}) + \delta_2 u_{t-1}, \\ \log x_t &= \xi + \varphi \log h_t + \tau(z_t) + u_t. \end{aligned} \tag{10}$$

It is natural to associate the leverage function, $\tau(z_t)$, with ζ_t, albeit there will be residual randomness in ζ_t that cannot be explained by the studentized return, z_t, alone. Consequently, u_t will be a mixture of pure sampling error, η_t, and the residual randomness $\zeta_t - \tau(z_t) - \xi$.

Since $\tau(z_t)$ is primarily related to the volatility shock, ζ_t, we should expect $\tau(z_t)$ to have a larger coefficient in the GARCH equation than u_t, and that is indeed what we find in a preliminary analysis of this particular model. Specifically we find, $\hat{\delta}_1 > \hat{\delta}_2 > 0$, where δ_2 is significant. This minor extension of the model leads to some interesting insight about the channels by which the realized measure in useful for the GARCH equation.

As discussed earlier, when x_t is included in the GARCH equation, then it does not distinguish between $\tau(z_t)$ and u_t, as it implied $\delta_1 = \delta_2$ in (10). The implication is that γ will be indicative of how accurate x_t estimates the integrated variance.

6.5.1 Realized EGARCH

The decomposition of the realized measure in the GARCH equation motivates a Realized GARCH model with the following EGARCH structure,

$$\begin{aligned} \log h_t &= \omega + \beta \log h_{t-1} + \tau(z_{t-1}) + \delta \epsilon_{t-1}, \\ \log x_t &= \xi + \varphi \log h_t + \kappa \tau(z_t) + (1+\delta)\epsilon_t. \end{aligned}$$

Here we have reparametrized the model to simplify the notation. For instance, β in this model maps into $\beta + \varphi\gamma$ in the formulation used earlier, and the leverage function has absorbed the scaling γ, and we have instead introduced the scaling κ in the measurement equation.

The Realized EGARCH model has a particularly interesting structure when $\beta = \varphi/\kappa$. In this case we can rewrite the measurement equation as

$$\log x_t = \tilde{\xi} + \kappa \log h_{t+1} + \epsilon_t, \qquad \text{where} \quad \tilde{\xi} = \xi - \kappa\omega,$$

so that the realized measure is implicitly being tied to the conditional variance for the next period.

The structure of the likelihood function for this model is different from that of our log-linear model, so we cannot utilize the QMLE results we derived in Section 4 to this model. Therefore, we leave a more detailed analysis of this model for future research.

7 Conclusion

In this paper we have proposed a complete model for returns and realized measures of volatility, x_t, where the latter is tied directly to the conditional volatility h_t. We have demonstrated that the model is straightforward to estimate and offers a substantial improvement in the empirical fit, relative to standard GARCH models. The model is informative about realized measurement, such as its accuracy.

We have shown that the realized GARCH model induces an interesting reduced-form model for $\{r_t, h_t\}$, that is similar to that of a stochastic volatility model with leverage effect.

Our empirical analysis can be extended in a number of ways. For instance, including a jump robust realized measure of volatility would be an interesting extension, because Bollerslev, Kretschmer, Pigorsch, and Tauchen (2009) found that the leverage effect primarily acts through the continuous volatility component. Another possible extension is to introduce a bivariate model of open-to-close and close-to-open returns, as an alternative to modeling close-to-close returns, see Andersen et al. (2008).

In this paper we have mainly focused on linear and log-linear specifications of the GARCH equation. In would be interesting for consider the whole family of specifications covered by the Augmented GARCH model, see Duan (1997).

The Realized GARCH framework is naturally extended to a multi-factor structure. Say m realized measures and k latent volatility variables. The Realized GARCH model discussed in this paper corresponds to the case $k = 1$, whereas the MEM framework corresponds to the case $m = k$. Such a hybrid framework would enable us to conduct inference about the number of latent factors, k. We could, in particular, test the one-factor structure, conjectured to be sufficient for the realized measure used in this paper, against the two-factor structure implied by MEM. The family of models extents traditional GARCH models by utilized the additional information provided by variables that are commonly called realized measures, still the objective of these models is essentially the same as that of GARCH models. For this reason we propose that all such models be called Realized GARCH models.

A Appendix of Proofs

Proof. [Proposition 1] The first result follows by substituting $\log x_t = \varphi \log h_t + \xi + w_t$ and $\log r_t^2 = \log h_t + \kappa + v_t$ into the GARCH equation and rearranging. Next, we substitute $\log h_t = (\log x_t - \xi - w_t)/\varphi$, $\log r_t^2 = (\log x_t - \xi - w_t)/\varphi + \kappa + v_t$, and multiply by φ, and find

$$\log x_t - \xi - w_t = \varphi\omega + \sum\nolimits_{i=1}^{p\vee q}(\beta_i + \alpha_i)(\log x_{t-i} - \xi - w_{t-i}) + \varphi \sum\nolimits_{j=1}^{q} \gamma_j \log x_{t-j} + \varphi \sum\nolimits_{j=1}^{q} \alpha_j(\kappa + v_{t-j})$$

so with $\pi_i = \alpha_i + \beta_i + \gamma_i\varphi$ we have

$$\log x_t = \xi(1 - \beta_\bullet - \alpha_\bullet) + \varphi\kappa\alpha_\bullet + \varphi\omega + \sum_{i=1}^{p\vee q} \pi_i \log x_{t-i} + w_t - \sum\nolimits_{i=1}^{p}(\alpha_i + \beta_i)w_{t-i} + \varphi \sum\nolimits_{j=1}^{q} \alpha_j v_{t-j}.$$

When $\varphi = 0$ the measurement equation shows that $\log x_t$ is an iid process. □

Proof. [Lemma 1] First note that

$$\frac{\partial g_t'}{\partial \lambda} = \left(0, \dot{h}_{t-1}, \ldots, \dot{h}_{t-p}, 0_{p+q+1\times q}\right) =: \dot{H}_{t-1},$$

Thus from the GARCH equation, $\tilde{h}_t = \lambda' g_t$, we have that

$$\dot{h}_t = \frac{\partial g_t'}{\partial \lambda}\lambda + g_t = \dot{H}_{t-1}\lambda + g_t = \sum_{i=1}^{p} \beta_i \dot{h}_{t-i} + g_t.$$

Similarly, the second order derivative, is given by

$$\ddot{h}_t = \frac{\partial(g_t + \dot{H}_{t-1}\lambda)}{\partial \lambda'} = \frac{\partial g_t}{\partial \lambda'} + \dot{H}_{t-1} + \frac{H_{t-1}}{\partial \lambda'}\lambda = \dot{H}_{t-1}' + \dot{H}_{t-1} + \sum_{i=1}^{p} \beta_i \frac{\partial \dot{h}_{t-i}}{\partial \lambda'} = \sum_{i=1}^{p} \beta_i \ddot{h}_{t-i} + \dot{H}_{t-1}' + \dot{H}_{t-1}.$$

For the starting values we observe the following: Regardless of $(h_0, \ldots, h_{p-1})$ being treated as fixed or as a vector of unknown parameters, we have $\dot{h}_s = \ddot{h}_s = 0$ for . Given the structure of $\ddot{h}_t$ this implies $\ddot{h}_1 = 0$.

When $p = q = 1$ it follows immediately that $\dot{h}_t = \sum_{j=0}^{t-1} \beta^j g_{t-j}$. Similarly we have

$$\ddot{h}_t = \sum_{j=0}^{t-1} \beta^j (\dot{H}_{t-1-j} + \dot{H}_{t-1-j}) = \sum_{j=0}^{t-2} \beta^j (\dot{H}_{t-1-j} + \dot{H}_{t-1-j})$$

where $\dot{H}_t = (0_{3\times 1}, \dot{h}_t, 0_{3\times 1})$ and where the second equality follows by $\dot{H}_0 = 0$. The results now follows by

$$\sum_{i=0}^{t-2} \beta^i \dot{h}_{t-1-i} = \sum_{i=0}^{t-2} \beta^i \sum_{j=0}^{t-1-i-1} \beta^j g_{t-1-i-j} = \sum_{i=0}^{t-2} \beta^i \sum_{k-i-1=0}^{t-i-2} \beta^{k-i-1} g_{t-k} = \sum_{i=0}^{t-2} \sum_{k=i+1}^{t-1} \beta^{k-1} g_{t-k} = \sum_{k=1}^{t-1} k\beta^{k-1} g_{t-k}.$$

□

Proof. [Proposition 2] Recall that $u_t = \tilde{x}_t - \psi' m_t$ and $\tilde{h}_t = g_t' \lambda$. So derivative with respect to $\tilde{h}_t$ are given by

$$\begin{aligned}
\frac{\partial z_t}{\partial \tilde{h}_t} &= \frac{\partial r_t \exp(-\frac{1}{2}\tilde{h}_t)}{\partial \tilde{h}_t} = -\frac{1}{2} z_t \qquad \text{so that} \qquad \frac{\partial z_t^2}{\partial \tilde{h}_t} = -z_t^2, \\
\dot{u}_t = \frac{\partial u_t}{\partial \tilde{h}_t} &= -\varphi + \frac{1}{2} z_t \tau' \dot{a}_t, \\
\ddot{u}_t = \frac{\partial \dot{u}_t}{\partial \tilde{h}_t} &= \frac{\partial \left(-\varphi + \frac{1}{2} z_t \dot{a}(z_t)' \tau\right)}{\partial \tilde{h}_t} = -\frac{1}{4} \tau' \left(z_t \dot{a}_t + z_t^2 \ddot{a}_t\right).
\end{aligned}$$

So with $\ell_t = -\frac{1}{2}\{\tilde{h}_t + z_t^2 + \log(\sigma_u^2) + u_t^2/\sigma_u^2\}$ we have

$$\frac{\partial \ell_t}{\partial u_t} = 2\frac{u_t}{\sigma_u^2} \qquad \text{and} \qquad \frac{\partial \ell_t}{\partial \tilde{h}_t} = -\frac{1}{2}\left\{1 + \frac{\partial z_t^2}{\partial \tilde{h}_t} + \frac{\partial u_t^2/\partial \tilde{h}_t}{\sigma_u^2}\right\} = -\frac{1}{2}\left\{1 - z_t^2 + \frac{2u_t \dot{u}_t}{\sigma_u^2}\right\}.$$

Derivatives with respect to λ are

$$\begin{aligned}
\frac{\partial z_t}{\partial \lambda} &= \frac{\partial z_t}{\partial \tilde{h}_t}\frac{\partial \tilde{h}_t}{\partial \lambda} = -\frac{1}{2} z_t \dot{h}_t \\
\frac{\partial u_t}{\partial \lambda} &= \frac{\partial u_t}{\partial \tilde{h}_t}\frac{\partial \tilde{h}_t}{\partial \lambda} = \dot{u}_t \dot{h}_t \\
\frac{\partial \dot{u}_t}{\partial \lambda'} &= \ddot{u}_t \dot{h}_t' \\
\frac{\partial \ell_t}{\partial \lambda} &= \frac{\partial \ell_t}{\partial \tilde{h}_t} \dot{h}_t = -\frac{1}{2}\left\{1 - z_t^2 + \frac{2u_t \dot{u}_t}{\sigma_u^2}\right\} \dot{h}_t.
\end{aligned}$$

Derivatives with respect to ψ are

$$\begin{aligned}
\frac{\partial u_t}{\partial \xi} &= -1, \qquad \frac{\partial \dot{u}_t}{\partial \xi} = 0, \qquad \text{and} \qquad \frac{\partial \ell_t}{\partial \xi} = \frac{\partial \ell_t}{\partial u_t}\frac{\partial u_t}{\partial \xi} = -2\frac{u_t}{\sigma_u^2}, \\
\frac{\partial u_t}{\partial \varphi} &= -\tilde{h}_t, \qquad \frac{\partial \dot{u}_t}{\partial \varphi} = -1, \qquad \text{and} \qquad \frac{\partial \ell_t}{\partial \varphi} = \frac{\partial \ell_t}{\partial u_t}\frac{\partial u_t}{\partial \varphi} = -2\frac{u_t}{\sigma_u^2}\tilde{h}_t, \\
\frac{\partial u_t}{\partial \tau} &= -a_t, \qquad \frac{\partial \dot{u}_t}{\partial \tau} = \frac{1}{2} z_t \dot{a}_t \qquad \text{and} \qquad \frac{\partial \ell_t}{\partial \tau} = \frac{\partial \ell_t}{\partial u_t}\frac{\partial u_t}{\partial \tau} = -2\frac{u_t}{\sigma_u^2} a_t.
\end{aligned}$$

Similarly, $\frac{\partial \ell_t}{\partial \sigma_u^2} = -\frac{1}{2}(\sigma_u^{-2} - u_t^2 \sigma_u^{-4})$. Now we turn to the second order derivatives.

$$\begin{aligned}
-2\frac{\partial^2 \ell_t}{\partial \lambda \partial \lambda'} &= \dot{h}_t \left\{-\frac{\partial z_t^2}{\partial \lambda'} + \frac{2}{\sigma_u^2}\left(\dot{u}_t \frac{\partial u_t}{\partial \lambda'} + u_t \frac{\partial \dot{u}_t}{\partial \lambda'}\right)\right\} + (1 - z_t^2 + \frac{2u_t}{\sigma_u^2}\dot{u}_t)\frac{\partial \dot{h}_t}{\partial \lambda'} \\
&= \dot{h}_t \left\{z_t^2 + \frac{2}{\sigma_u^2}\left(\dot{u}_t^2 + u_t \ddot{u}_t\right)\dot{h}_t'\right\} + (1 - z_t^2 + \frac{2u_t}{\sigma_u^2}\dot{u}_t)\ddot{h}_t.
\end{aligned}$$

Similarly, since $\frac{\partial z_t}{\partial \psi} = 0$ we have

$$
\begin{aligned}
-2\frac{\partial^2 \ell_t}{\partial \lambda \partial \xi} &= \frac{\partial(1 - z_t^2 + \frac{2u_t}{\sigma_u^2}\dot{u}_t)\dot{h}_t}{\partial \xi} = 2\dot{h}_t\left(\frac{\partial u_t}{\partial \psi'}\frac{\dot{u}_t}{\sigma_u^2} + \frac{u_t}{\sigma_u^2}\frac{\partial \dot{u}_t}{\partial \xi}\right) = 2\dot{h}_t\left(-\frac{\dot{u}_t}{\sigma_u^2} + 0\right) \\
-2\frac{\partial^2 \ell_t}{\partial \lambda \partial \varphi} &= \frac{\partial(1 - z_t^2 + \frac{2u_t}{\sigma_u^2}\dot{u}_t)\dot{h}_t}{\partial \varphi} = 2\dot{h}_t\left(\frac{\partial u_t}{\partial \varphi}\frac{\dot{u}_t}{\sigma_u^2} + \frac{u_t}{\sigma_u^2}\frac{\partial \dot{u}_t}{\partial \varphi}\right) = 2\dot{h}_t\left(-\tilde{h}_t\frac{\dot{u}_t}{\sigma_u^2} - \frac{u_t}{\sigma_u^2}\right) \\
-2\frac{\partial^2 \ell_t}{\partial \lambda \partial \tau'} &= 2\dot{h}_t\left(\frac{\partial u_t}{\partial \tau'}\frac{\dot{u}_t}{\sigma_u^2} + \frac{u_t}{\sigma_u^2}\frac{\partial \dot{u}_t}{\partial \tau'}\right) = 2\dot{h}_t\left(-a_t'\frac{\dot{u}_t}{\sigma_u^2} + \frac{u_t}{\sigma_u^2}\frac{1}{2}z_t\dot{a}_t\right),
\end{aligned}
$$

so that

$$
\frac{\partial^2 \ell_t}{\partial \lambda \partial \psi'} = \frac{\dot{u}_t}{\sigma_u^2}\dot{h}_t m_t' + \frac{u_t}{\sigma_u^2}\dot{h}_t b_t', \qquad \text{with} \quad b_t = (0, 1, -\frac{1}{2}z_t\dot{a}_t')'.
$$

$$
\frac{\partial^2 \ell_t}{\partial \lambda \partial \sigma_u^2} = -\frac{1}{2}\frac{\partial(1 - z_t^2 + \frac{2u_t}{\sigma_u^2}\dot{u}_t)\dot{h}_t}{\partial \sigma_u^2} = \frac{u_t\dot{u}_t\dot{h}_t}{\sigma_u^4}
$$

$$
\frac{\partial^2 \ell_t}{\partial \psi \partial \psi'} = -\frac{1}{\sigma_u^2}m_t m_t'
$$

$$
\frac{\partial^2 \ell_t}{\partial \psi \partial \sigma_u^2} = -\frac{1}{2}(-\frac{2u_t}{\sigma_u^4})m_t = \frac{u_t}{\sigma_u^4}m_t
$$

$$
\frac{\partial^2 \ell_t}{\partial \sigma_u^2 \partial \sigma_u^2} = -\frac{1}{2}\left(\frac{-1}{\sigma_u^4} + 2\frac{u_t^2}{\sigma_u^6}\right) = \frac{1}{2}\frac{\sigma_u^2 - 2u_t^2}{\sigma_u^6}.
$$

□

Proof. [Proposition 4] With a Gaussian specification for z_t we have $\mathrm{E}(r_t^2) = \mathrm{E}(h_t)$ and $\mathrm{E}(r_t^4) = 3\mathrm{E}(h_t^2)$. From the ARMA representation for this process we have with $\mu = \omega + \gamma\xi$ and $\pi = \alpha + \beta + \varphi\gamma \in (-1, 1)$, that

$$
h_t = \mu + \pi h_{t-1} + \gamma w_{t-1} + \alpha v_{t-1} = \sum_{i=0}^{\infty} \pi^i(\gamma w_{t-i-1} + \alpha v_{t-i-1}) + \tfrac{\mu}{1-\pi},
$$

where $w_t = \tau(z_t) + u_t$ and $v_t = h_t(z_t^2 - 1)$. So that $\mathrm{E}(h_t) = \mu/(1-\pi)$. Next we note that $\mathrm{E}(w_t^2) = \sigma_{\tau^2}^2 + \sigma_u^2$, $\mathrm{E}(v_t^2) = 2\mathrm{E}(h_t^2)$, and $\mathrm{E}(w_t v_t) = \gamma\alpha\tau_2\mathrm{E}(z_t^2 - 1)^2\mathrm{E}(h_t) = 2\gamma\alpha\tau_2\mu/(1-\pi)$, where we have used that $z_t \sim N(0,1)$ and the Hermite polynomial structure of $\tau(z)$. The second moment is given by

$$
\mathrm{E}(h_t^2) = \sum_{i=0}^{\infty} \pi^{2i}\left\{\gamma^2(\sigma_u^2 + \sigma_{\tau^2}^2) + 2\alpha^2\mathrm{E}(h_t^2) + 4\gamma\alpha\tau_2\tfrac{\mu}{1-\pi}\right\} + \tfrac{\mu^2}{(1-\pi)^2}
$$

so that $\left(1 - \frac{2\alpha^2}{1-\pi^2}\right)\mathrm{E}(h_t^2) = \frac{\gamma^2(\sigma_u^2+\sigma_{\tau^2}^2)+4\gamma\alpha\tau_2\frac{\mu}{1-\pi}}{1-\pi^2} + \frac{\mu^2}{(1-\pi)^2}$, and hence

$$
\mathrm{E}(h_t^2) = \left(\frac{1-\pi^2}{1-\pi^2-2\alpha^2}\right)\frac{\gamma^2(\sigma_u^2 + \sigma_{\tau^2}^2) + 4\gamma\alpha\tau_2\frac{\mu}{1-\pi}}{1-\pi^2} + \left(\frac{1-\pi^2}{1-\pi^2-2\alpha^2}\right)\tfrac{\mu^2}{(1-\pi)^2}
$$

Hence the excess kurtosis is given by

$$\frac{\mathrm{E}(r_t^4)}{\mathrm{E}(r_t^2)^2}-3=3\left\{\frac{\mathrm{E}(h_t^2)}{\mathrm{E}(h_t)^2}-1\right\}=3\frac{(1-\pi)^2}{1-\pi^2-2\alpha^2}\left(\gamma^2\frac{\sigma_u^2+\sigma_{\tau^2}^2}{\mu^2}+4\gamma\frac{\alpha\tau_2}{\mu(1-\pi)}\right)+\frac{6\alpha^2}{1-\pi^2-2\alpha^2},$$

and the results follow. □

Lemma 2. *Let $W=\tau_1 Z+\tau_2(Z^2-1)+U$, where $Z\sim N(0,1)$ and $U\sim N(0,\sigma_u^2)$. Then*

$$\mathrm{E}\left(\exp\left\{\pi^i\gamma W\right\}\right)=\frac{1}{\sqrt{1-2\pi^i\gamma\tau_2}}e^{\frac{\pi^{2i}\gamma^2\tau_1^2}{2(1-2\pi^i\gamma\tau_2)}-\pi^i\gamma\tau_2+\frac{\pi^{2i}\gamma^2\sigma_u^2}{2}}.$$

Proof. We have

$$\begin{aligned}\mathrm{E}\left(e^{aZ+\frac{b}{2}(Z^2-1)}\right) &= \int_{-\infty}^{\infty}e^{az+\frac{b}{2}(z^2-1)}\frac{1}{\sqrt{2\pi}}e^{\frac{z^2}{2}}\mathrm{d}z\\ &= \int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}}e^{\frac{a^2}{2(1-b)}-\frac{b}{2}-\frac{1}{2}\frac{\left(z-\frac{a}{1-b}\right)}{(1-b)^{-1}})}\mathrm{d}z\\ &= \frac{1}{\sqrt{1-b}}e^{\frac{a^2}{2(1-b)}-\frac{b}{2}},\end{aligned}$$

and from the moment generating function for the Gaussian distribution we have $\mathrm{E}\left(e^{cU}\right)=e^{\frac{c^2\sigma_u^2}{2}}$. Since Z and U are independent, we have

$$\begin{aligned}\mathrm{E}\left(\exp\left\{\pi^i\gamma W\right\}\right) &= \mathrm{E}\left(\exp\left\{\pi^i\gamma\tau_1 Z+\pi^i\gamma\tau_2(Z^2-1)\right\}\right)\mathrm{E}\left(\exp\left\{\pi^i\gamma U\right\}\right)\\ &= \frac{1}{\sqrt{1-2\pi^i\gamma\tau_2}}e^{\frac{\pi^{2i}\gamma^2\tau_1^2}{2(1-2\pi^i\gamma\tau_2)}-\pi^i\gamma\tau_2}e^{\frac{\pi^{2i}\gamma^2\sigma_u^2}{2}}.\end{aligned}$$

□

Proof. [Proposition 5] We note that

$$\begin{aligned}h_t &= \exp\left(\sum_{i=0}^{\infty}\pi^i(\mu+\gamma w_{t-1})\right)=e^{\frac{\mu}{1-\pi}}\prod_{i=0}^{\infty}\mathrm{E}\exp\left(\gamma\pi^i\tau(z_{t-i})\right)\mathrm{E}\left(\exp\left\{\pi^i\gamma u_{t-i}\right\}\right),\\ h_t^2 &= \exp\left(2\sum_{i=0}^{\infty}\pi^i(\mu+\gamma w_{t-1})\right)=e^{\frac{2\mu}{1-\pi}}\prod_{i=0}^{\infty}\mathrm{E}\exp\left(2\gamma\pi^i\tau(z_{t-i})\right)\mathrm{E}\left(\exp\left\{2\pi^i\gamma u_{t-i}\right\}\right),\end{aligned}$$

and using results, such as

$$\mathrm{E}\left(\prod_{i=0}^{\infty}\exp\left\{\pi^i\gamma u_{t-i}\right\}\right)=\prod_{i=0}^{\infty}\mathrm{E}\left(\exp\left\{\pi^i\gamma u_{t-i}\right\}\right)=\prod_{i=0}^{\infty}e^{\frac{\pi^{2i}\gamma^2\sigma_u^2}{2}}=e^{\sum_{i=0}^{\infty}\frac{\pi^{2i}\gamma^2\sigma_u^2}{2}}=e^{\frac{\gamma^2\sigma_u^2/2}{1-\pi^2}},$$

we find that

$$\begin{aligned}
\frac{\mathrm{E}h_t^2}{(\mathrm{E}h_t)^2} &= \frac{e^{\frac{2\mu}{1-\pi}} \prod_{i=0}^{\infty} \mathrm{E}\exp\left(2\gamma\pi^i w_{t-1}\right)}{e^{\frac{2\mu}{1-\pi}} \prod_{i=0}^{\infty} \left\{\mathrm{E}\exp\left(\gamma\pi^i w_{t-1}\right)\right\}^2} \\
&= \left(\prod_{i=0}^{\infty} \frac{1-2\pi^i\gamma\tau_2}{\sqrt{1-4\pi^i\gamma\tau_2}}\right) \frac{e^{\sum_{i=0}^{\infty} \frac{4\pi^{2i}\gamma^2\tau_1^2}{2(1-4\pi^i\gamma\tau_2)}}}{e^{2\sum_{i=0}^{\infty} \frac{\pi^{2i}\gamma^2\tau_1^2}{2(1-2\pi^i\gamma\tau_2)}}} \frac{e^{-\frac{2\gamma\tau_2}{1-\pi}}}{e^{-2\frac{\gamma\tau_2}{1-\pi}}} \frac{e^{\frac{2\gamma^2\sigma_u^2}{1-\pi^2}}}{e^{\frac{\gamma^2\sigma_u^2}{1-\pi^2}}} \\
&= \left(\prod_{i=0}^{\infty} \frac{1-2\pi^i\gamma\tau_2}{\sqrt{1-4\pi^i\gamma\tau_2}}\right) e^{\sum_{i=0}^{\infty} \frac{2\pi^{2i}\gamma^2\tau_1^2}{(1-4\pi^i\gamma\tau_2)} - \frac{\pi^{2i}\gamma^2\tau_1^2}{(1-2\pi^i\gamma\tau_2)}} e^{\frac{\gamma^2\sigma_u^2}{1-\pi^2}} \\
&= \left(\prod_{i=0}^{\infty} \frac{1-2\pi^i\gamma\tau_2}{\sqrt{1-4\pi^i\gamma\tau_2}}\right) e^{\sum_{i=0}^{\infty} \frac{\pi^{2i}\gamma^2\tau_1^2}{(1-6\pi^i\gamma\tau_2+8\pi^{2i}\gamma^2\tau_2^2)}} e^{\frac{\gamma^2\sigma_u^2}{1-\pi^2}}
\end{aligned}$$

where the last equality uses

$$\begin{aligned}
\frac{2\pi^{2i}\gamma^2\tau_1^2}{(1-4\pi^i\gamma\tau_2)} - \frac{\pi^{2i}\gamma^2\tau_1^2}{(1-2\pi^i\gamma\tau_2)} &= \frac{2\pi^{2i}\gamma^2\tau_1^2(1-2\pi^i\gamma\tau_2) - \pi^{2i}\gamma^2\tau_1^2(1-4\pi^i\gamma\tau_2)}{(1-4\pi^i\gamma\tau_2)(1-2\pi^i\gamma\tau_2)} \\
&= \frac{\pi^{2i}\gamma^2\tau_1^2}{(1-4\pi^i\gamma\tau_2)(1-2\pi^i\gamma\tau_2)} = \frac{\pi^{2i}\gamma^2\tau_1^2}{(1-6\pi^i\gamma\tau_2+8\pi^{2i}\gamma^2\tau_2^2)}.
\end{aligned}$$

□

A.1 Approximate Expression for the Kurtosis

In this subsection we provide a justification for the approximation (9). Recall that

$$\frac{\mathrm{E}(r_t^4)}{\mathrm{E}(r_t^2)^2} = 3\left(\prod_{i=0}^{\infty} \frac{1-2\pi^i\gamma\tau_2}{\sqrt{1-4\pi^i\gamma\tau_2}}\right) \exp\left\{\sum_{i=0}^{\infty} \frac{\pi^{2i}\gamma^2\tau_1^2}{1-6\pi^i\gamma\tau_2+8\pi^{2i}\gamma^2\tau_2^2}\right\} \exp\left\{\frac{\gamma^2\sigma_u^2}{1-\pi^2}\right\}.$$

For the first term on the right hand side, we have

$$\begin{aligned}
\log \prod_{i=0}^{\infty} \frac{1-2\pi^i\gamma\tau_2}{\sqrt{1-4\pi^i\gamma\tau_2}} &\simeq \int_0^{\infty} \log \frac{1-2\pi^x\gamma\tau_2}{\sqrt{1-4\pi^x\gamma\tau_2}} \mathrm{d}x \\
&= \frac{1}{\log\pi} \left\{\sum_{k=1}^{\infty} \frac{(2\gamma\tau_2)^k}{k^2} - \frac{1}{2}\frac{(4\gamma\tau_2)^k}{k^2}\right\} (1-2^{k-1}) \\
&= \frac{1}{\log\pi} \sum_{k=1}^{\infty} \frac{(2\gamma\tau_2)^k}{k^2}(1-2^{k-1}) \\
&= \frac{\gamma^2\tau_2^2\left\{1+\frac{8}{3}\gamma\tau_2+7(\gamma\tau_2)^2+\frac{96}{5}(\gamma\tau_2)^3+\frac{496}{9}(\gamma\tau_2)^4+\cdots\right\}}{-\log\pi}.
\end{aligned}$$

The second term can be bounded by

$$\frac{\gamma^2\tau_1^2}{1-\pi^2} \leq \sum_{i=0}^{\infty} \frac{\pi^{2i}\gamma^2\tau_1^2}{1-6\pi^i\gamma\tau_2+8\pi^{2i}\gamma^2\tau_2^2} \leq \frac{\gamma^2\tau_1^2}{1-\pi^2}\frac{1}{1-6\pi\gamma\tau_2}.$$

So the approximation error is small when $\gamma\tau_2$ is small.

References

Andersen, T., Bollerslev, T., Diebold, F. X., 2007. Roughing it up: Including jump components in the measurement, modeling and forecasting of return volatility. Review of Economics and Statistics 89, 701–720.

Andersen, T. G., Bollerslev, T., 1998. Answering the skeptics: Yes, standard volatility models do provide accurate forecasts. International Economic Review 39 (4), 885–905.

Andersen, T. G., Bollerslev, T., Diebold, F. X., Labys, P., 2001. The distribution of exchange rate volatility. Journal of the American Statistical Association 96 (453), 42–55, correction published in 2003, volume 98, page 501.

Andersen, T. G., Bollerslev, T., Diebold, F. X., Labys, P., 2003. Modeling and forecasting realized volatility. Econometrica 71 (2), 579–625.

Andersen, T. G., Bollerslev, T., Huang, X., 2008. A reduced form framework for modeling volatility of speculative prices based on realized variation measures. Working Paper.

Andersen, T. G., Bollerslev, T., Meddahi, N., 2004. Analytic evaluation of volatility forecasts. International Economic Review 45, 1079–1110.

Andersen, T. G., Bollerslev, T., Meddahi, N., 2005. Correcting the errors: Volatility forecast evaluation using high-frequency data and realized volatilities. Econometrica 73, 279–296.

Andersen, T. G., Bollerslev, T., Meddahi, N., 2010. Market microstructure noise and realized volatility forecasting. Journal of Econometrics.

Barndorff-Nielsen, O. E., Hansen, P. R., Lunde, A., Shephard, N., 2008a. Designing realised kernels to measure the ex-post variation of equity prices in the presence of noise. Econometrica 76, 1481–536.

Barndorff-Nielsen, O. E., Hansen, P. R., Lunde, A., Shephard, N., 2008b. Multivariate realised kernels: consistent positive semi-definite estimators of the covariation of equity prices with noise and non-synchronous trading. working paper.

Barndorff-Nielsen, O. E., Hansen, P. R., Lunde, A., Shephard, N., 2009a. Realised kernels in practice: Trades and quotes. Econometrics Journal 12, 1–33.

Barndorff-Nielsen, O. E., Kinnebrock, S., Shephard, N., 2009b. Measuring downside risk - realised semivariance. In: Watson, M. W., Bollerslev, T., Russell, J. (Eds.), Volatility and Time Series Econometrics: Essays in Honor of Robert Engle. Oxford University Press.

Barndorff-Nielsen, O. E., Shephard, N., 2002. Econometric analysis of realised volatility and its use in estimating stochastic volatility models. Journal of the Royal Statistical Society **B** 64, 253–280.

Barndorff-Nielsen, O. E., Shephard, N., 2004. Power and bipower variation with stochastic volatility and jumps (with discussion). Journal of Financial Econometrics 2, 1–48.

Barndorff-Nielsen, O. E., Shephard, N., 2007. Advances in Economics and Econometrics. Theory and Applications, Ninth World Congress. Econometric Society Monographs. Cambridge University Press, Ch. Variation, jumps and high frequency data in financial econometrics, pp. 328–372.

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

Bollerslev, T., 2009. Glossary to ARCH (GARCH). In: Bollerslev, T., Russell, J. R., Watson, M. (Eds.), Volatility and Time Series Econometrics: Essays in Honour of Robert F. Engle. Oxford University Press, Oxford, UK.

Bollerslev, T., Kretschmer, U., Pigorsch, C., Tauchen, G., 2009. A discrete-time model for daily s&p500 returns and realized variations: Jumps and leverage effects. Journal of Econometrics 150, 151–166.

Bollerslev, T., Wooldridge, J. M., 1992. Quasi-maximum likelihood estimation and inference in dynamic models with time-varying covariance. Econometric Reviews 11, 143–172.

Brownless, C. T., Gallo, G. M., 2010. Comparison of volatility measures: A risk management perspective. Forthcoming in Journal of Financial Econometrics.

Carrasco, M., Chen, X., 2002. Mixing and moment properties of various GARCH and stochastic volatility models. Econometric Theory 18, 17–39.

Chen, X., Ghysels, E., Wang, F., 2009. The HYBRID GARCH class of models. working paper.

Cipollini, F., Engle, R. F., Gallo, G. M., 2009. A model for multivariate non-negative valued processes in financial econometrics. working paper.

Duan, J., 1997. Augmented GARCH(p, q) process and its diffusion limit. Journal of Econometrics 79 (1), 97–127.

Engle, R. F., 1982. Autoregressive conditional heteroskedasticity with estimates of the variance of U.K. inflation. Econometrica 45, 987–1007.

Engle, R. F., 2002. New frontiers for arch models. Journal of Applied Econometrics 17, 425–446.

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

Engle, R. F., Lilien, D. V., Robins, R. P., 1987. Estimating time varying risk premia in the term structure: The ARCH-M model. Econometrica 55, 391–407.

Engle, R. F., Ng, V., 1993. Measuring and testing the impact of news on volatility. Journal of Finance 48, 1747–1778.

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

Geweke, J., 1986. Modelling persistence in conditional variances: A comment. Econometric Reviews 5, 57–61.

Ghysels, E., Chen, X., 2010. News – gode or bad – and its impact on volatility predictions over multiple horizons. Review of Financial Studies, forthcoming.

Hansen, P. R., 2009. In-sample fit and out-of-sample fit: Their joint distribution and its implications for model selection. Working Paper.

Hansen, P. R., Horel, G., 2009. Quadratic variation by markov chains. working paper `http://www.stanford.edu/people/peter.hansen`.

Jensen, S. T., Rahbek, A., 2004a. Asymptotic inference for nonstationary GARCH. Econometric Theory 20, 1203–1226.

Jensen, S. T., Rahbek, A., 2004b. Asymptotic normality of the QMLE estimator of ARCH in the nonstationary case. Econometrica 72, 641–646.

Lee, S., Hansen, B. E., 1994. Asymptotic theory for the GARCH(1,1) quasi-maximum likelihood estimator. Econometric Theory 10, 29–52.

Lumsdaine, R. L., 1996. Consistency and asymptotic normality of the quasi-maximum likelihood estimator in IGARCH(1,1) and covariance stationary GARCH(1,1) models. Econometrica 10, 29–52.

Meddahi, N., 2003. ARMA representation of integrated and realized variances. Econometrics Journal 6, 335–356.

Milhøj, A., 1987. A multiplicative parameterization of ARCH models. Tech. rep., Department of Statistics, University of Copenhagen.

Nelson, D. B., 1991. Conditional heteroskedasticity in asset returns: A new approach. Econometrica 59 (2), 347–370.

Pantula, S. G., 1986. Modelling persistence in conditional variances: A comment. Econometric Reviews 5, 71–74.

Shephard, N., Sheppard, K., 2010. Realising the future: Forecasting with high frequency based volatility (HEAVY) models. Journal of Applied Econometrics 25, 197–231.

Straumann, D., Mikosch, T., 2006. Quasi-maximum-likelihood estimation in conditionally heteroscedastic time series: A stochastic recurrence equation approach. Annals of Statistics 34, 2449–2495.

Teräsvirta, T., 2009. An introduction to univariate GARCH models. In: Mikosch, T., Kreiss, T., Davis, J.-P., Andersen, R. A., Gustav, T. (Eds.), Handbook of Financial Time Series. Springer, pp. 17–42.

Visser, M. P., 2008. Garch parameter estimation using high-frequency data. working paper.

Yu, J., 2008. A semiparametric stochastic volatility model. Working paper.

Research Papers 2010