Garch Parameter Estimation Using High-Frequency Data

Visser, Marcel P.

Korteweg-de Vries Instute for Mathematics, University of Amsterdam

10 June 2008

Garch Parameter Estimation
Using High-Frequency Data

Marcel P. Visser *

June 10, 2008

Abstract

Estimation of the parameters of Garch models for financial data is typically based on daily close-to-close returns. This paper shows that the efficiency of the parameter estimators may be greatly improved by using volatility proxies based on intraday data. The paper develops a Garch quasi maximum likelihood estimator (QMLE) based on these proxies. Examples of such proxies are the realized volatility and the intraday high-low range. Empirical analysis of the S&P 500 index tick data shows that the use of a suitable proxy may reduce the variances of the estimators of the Garch autoregression parameters by a factor 20.

JEL classification: C14, C22, C51, G1.

Key Words: volatility estimation, quasi maximum likelihood, volatility proxy, Gaussian QMLE, log-Gaussian QMLE, autoregressive conditional heteroscedasticity.

*Korteweg-de Vries Institute for Mathematics, University of Amsterdam. Plantage Muidergracht 24, 1018 TV Amsterdam, The Netherlands. Tel. +31 20 5255861. Email: m.p.visser@uva.nl.

1 Introduction

Garch models based on close-to-close daily returns do quite well in describing financial volatility, but they seem incompatible with intraday high-frequency data at first sight. The standard continuous time model for the log of asset prices is a semimartingale, and volatility is then the square root of the quadratic variation process. High-frequency data are accordingly used to estimate the daily increment in the quadratic variation. In the discrete time setting of Garch models, the day volatility is the scale factor that transforms the iid innovation Z_n into the log-return r_n.

Throughout this paper we assume that the sequence of *daily* log-returns r_n is a stationary Garch(1,1) process. We use the Garch(1,1) representation given by Drost and Klaassen (1997):

$$r_n = v_n \tau \, Z_n \tag{1}$$

$$v_n^2 = 1 + \gamma r_{n-1}^2 + \beta v_{n-1}^2, \tag{2}$$

where the innovations Z_n are iid, mean zero. For identification the second moment is standardized by $\mathbb{E}Z_n^2 = 1$. This system is equivalent to the more familiar Garch equations $r_n = \sigma_n Z_n$ and $\sigma_n^2 = \kappa + \alpha r_{n-1}^2 + \beta \sigma_{n-1}^2$ by writing $\sigma_n = v_n \tau$, and $\kappa = \tau^2$, $\alpha = \gamma \tau^2$. The system given by (1) and (2) has the advantage that the standardization of Z_n affects only the norming parameter τ. The focus on Garch(1,1) is for simplicity of exposition only. The principle below allows one to improve estimation of the parameters of any scale process v_n.

Let us say a few words on parameter estimation in this model. The returns r_n, $n = 1, \ldots, N$, are observable, the volatilities v_n are not. One may estimate the parameter $\theta = (\tau, \gamma, \beta)$ in (1) and (2) by maximizing the log-likelihood of the observations r_n. If the Z_n are standard Gaussian, one obtains the likelihood by using that the returns r_n are conditionally Gaussian distributed with mean zero and variance $v_n^2 \tau^2$. If the distribution of the random variables Z_n is unknown, one may still proceed as if the Z_n were standard Gaussian. The estimator is then called a quasi-ML estimator (QMLE).

Our goal is to improve estimation of the autoregression parameters γ and β by making use of high-frequency data. There have been attempts to make use of high-frequency data for parameter estimation. One could derive the parameters of the daily Garch process by estimation of the Garch process with a five-minute time unit using the time aggregation results of Drost and Nijman (1993). Such an approach runs into problems since it does not take into account the daily volatility cycle observed in five-minute returns, see Andersen and Bollerslev

(1997). One may also start from a continuous time diffusion. The discretized process is then a stochastic volatility model and one may use the high-low range for parameter estimation, see Alizadeh, Brandt and Diebold (2002). If the diffusion coefficient is an Ornstein-Uhlenbeck process, or a CEV process, then the daily integrated volatility is an ARMA(1,1) process. The ARMA parameters may then be estimated by state space methods, see Barndorff-Nielsen and Shephard (2002).

The present paper takes a different approach. We start out from the Garch system (1) and (2) for the daily close-to-close returns r_n. For each day n we observe the entire intraday log-return process $R_n(\cdot)$. To distill the day volatility from R_n one may use the empirical realized quadratic variation RQV_n based on five-minute intervals (also called realized variance). One obtains RQV_n by summing the squared five-minute increments over the n-th trading day. The realized volatility $H_n = \sqrt{RQV_n}$ is generally seen as a good proxy for volatility. Now, the parameters γ and β play a role in the likelihood for the H_n. If one could construct this likelihood, one hopes to find an efficiency gain compared with estimation based on the likelihood for the returns r_n. To obtain the likelihood for the proxies H_n one needs to embed the close-to-close return r_n in a model for the intraday return process R_n. As a model we shall propose a simple extension of the daily Garch process to a continuous time intraday log-return process R_n. This intraday extension yields the following system for the volatility proxy H_n:

$$
\begin{aligned}
H_n &= v_n \tau_H \, Z_{H,n} && (3)\\
v_n^2 &= 1 + \gamma r_{n-1}^2 + \beta v_{n-1}^2, && (4)
\end{aligned}
$$

where the innovations $Z_{H,n} \geq 0$ are iid and have standardization $\mathbb{E} Z_{H,n}^2 = 1$. The system given by (3) and (4) has the property that the parameters γ, β in (4) have the same value as in equation (2). So H_n and r_n share the daily factor v_n. We derive the likelihood for the observations H_n and show how one may estimate the parameter $\theta = (\tau_H, \gamma, \beta)$ by quasi maximum likelihood. More generally, we shall show that one may replace H_n by other proxies than the realized volatility; for example the intraday high-low range, or the absolute value of the maximal decrease of R_n over a fifteen minute interval.

The theory developed in the paper gives exact relationships for the asymptotic relative efficiency of QML estimators for γ and β using alternative proxies H_n. The quality of the estimator is determined by the innovation Z_H. If the variance of Z_H^2 is smaller than $\mathrm{var}(Z^2)$, then the QML estimator for γ, β based on the proxies H_n is sharper than the one based on the returns r_n. Theorem 3.1 gives conditions for the asymptotic normality of the QMLE based on H_n. Its proof is based on the likelihood theory in Straumann and Mikosch (2006). A similar estimation theory may be developed using the log proxies, $\log(H_n)$.

The estimators are applied to the four years 1992–1995 of the S&P 500 index. We motivate this choice of time period in Section 4. We emphasize that it is not our aim to identify an optimal volatility model. For using volatility proxies as predictors of future volatility we refer to Engle and Gallo (2006) and Ghysels, Santa-Clara, and Valkanov (2006). The purpose of the present analysis is to judge the potential benefits of using volatility proxies based on intraday data for parameter estimation. Figure 1 gives an impression of the empirical efficiency gains. It shows four 95% confidence ellipses for estimates of (γ, β), based on $|r_n|$

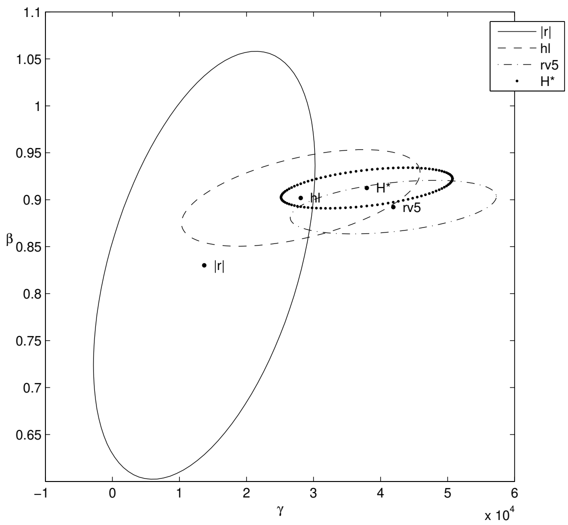


Figure 1: Confidence regions for estimators of (γ, β), based on alternative volatility proxies H. The data are the S&P 500 index futures over 1992–01–01 to 1995–12–31 (1001 days). In the figure $|r|$ depicts the Gaussian QMLE applied to the absolute returns $|r_n|$ (the usual Garch(1,1) QMLE). The other estimates are based on the log-Gaussian QMLE applied to *hl* (high-low range), RV5 (five-minute realized volatility) and H^* (a proxy determined in de Vilder and Visser (2007)).

and on three other proxies. The confidence regions are computed using Bollerslev-Wooldridge (1992) robust covariances. The main point of this paper is made in the figure: *one may greatly improve the parameter estimation for Garch processes by the use of suitable proxies based on high-frequency data.*

The remainder of the paper is organized as follows. Section 2 introduces the model and discusses volatility proxies. Section 3 presents the theory on estimation by intraday volatility measures. Building on the principle of quasi maximum likelihood it provides the theory for parameter estimation by the multiplicative equation (3), which we shall refer to as Gaussian QML. It also provides results for estimation based on the log proxies, log-Gaussian QML.

Section 4 applies the QMLE's to the S&P 500 index data. Section 5 compares simulations for estimators of (γ, β) based on realized variance with the standard QMLE based on close-to-close returns. Our conclusions are presented in Section 6. Appendices A, B, C, and D give a description of the data, background on QML estimation, proofs, and background on simulations.

2 Preliminaries

Section 2.1 introduces the model for the intraday return process $R_n(\cdot)$. Section 2.2 characterizes the volatility proxies that may be used for QML estimation. For a more detailed account of the model for R_n, and of proxies, see de Vilder and Visser (2007).

2.1 Intraday Return Process

To deal with high-frequency data in the daily Garch system given by (1) and (2) one needs to embed the sequence of daily close-to-close returns (r_n) in a continuous time process. Each day we observe the continuous time, intraday log-return process $R_n(\cdot)$: observable information is the filtration $(\mathcal{F}_n)$, given by $\sigma(R_k,\ k \leq n)$. The process R_n represents within day n the log-return with respect to the previous day's close. Describe $R_n(\cdot)$ as the product of the scale factor $v_n\tau$ and a cadlag[1] process $\Psi_n(\cdot)$ on the time interval $[0, 1]$, the trading day:

$$R_n(u) = v_n\tau\ \Psi_n(u), \qquad 0 \leq u \leq 1,$$

where the processes $\Psi_n(\cdot)$ are iid over different days, have standardization $\mathbb{E}\Psi_n^2(1) = 1$, and intraday time u advances from zero to one. So $R_n(0)$ gives the overnight return and $R_n(1)$ equals the close-to-close return r_n. The scale factor $v_n\tau$ is the same as in the discrete time model (1), and is constant within the day. The process Ψ_n may be any process representing the intraday price pattern. This continuous time model is simple enough to allow for analysis, and it takes into account the diversity in the behaviour of the market on successive trading days. One may recover the close-to-close returns r_n by setting $Z_n = \Psi_n(1)$:

$$r_n = R_n(1) = v_n\tau\, Z_n.$$

[1]The sample paths are right-continuous and have left limits.

2.2 Volatility Proxies

Let us introduce proxies for the volatility $v_n\tau$. In general we call the random variable $H_n = H(R_n)$ (or the functional H) a *proxy* whenever H is positive and is positively homogeneous in R_n. Positive homogeneity means:

$$H(sR_n) = sH(R_n), \qquad s \geq 0. \tag{5}$$

The absolute return $|r_n|$ is a proxy. Other examples are the intraday high-low range and the realized volatility.

We assume that the random variable $H(\Psi)$ is not identically zero,

$$\mu_2^H = \sqrt{\mathbb{E}H^2(\Psi)} > 0.$$

Let us introduce the normalized innovation Z_H by setting

$$Z_H = H(\Psi)/\mu_2^H,$$

so $\mathbb{E}Z_H^2 = 1$. By homogeneity $H_n = H(R_n) = v_n\tau H(\Psi_n)$, which gives (*cf.* (3))

$$H_n = v_n\tau_H\, Z_{H,n},$$

where the positive, iid innovations $Z_{H,n} \geq 0$ have $\mathbb{E}Z_H^2 = 1$, and $\tau_H = \tau\,\mu_2^H$. Replacing H by $3H$ only adds a factor 3 to the norming parameter τ_H. A good proxy H distills the factor $v_n\tau_H$ from R_n without much error.

3 QML Estimators Based on a General Volatility Measure

This section develops the theory for estimation of the parameters γ and β using the proxy H_n, as sketched in the introduction. We first treat the Gaussian QML estimator, which is based on the multiplicative equation $H_n = v_n\tau_H Z_{H,n}$. We then discuss the log-Gaussian QMLE, which is a Gaussian QMLE applied to the additive equation $\log(H_n) = \log(v_n) + \log(\tau_H) + \log(Z_{H,n})$.

Let us address one important issue first. Why should one bother with likelihood methods if one can simply obtain $v_n^2\tau^2$ from the intraday return process $R_n(\cdot)$? Consider for example the quadratic variation. The quadratic variation (QV) is the limit of the sum of squared intraday returns, as the length of the sampling intervals approach zero. If the process $\Psi(\cdot)$ of Section 2.1 is a Brownian motion, then $QV(\Psi) = 1$, so $QV(R_n) = v_n^2\tau^2$. In general we do not have this exact relationship. Under fairly mild conditions the quadratic variation of R_n is an unbiased estimator of the conditional variance of the daily return,

$$\mathbb{E}(QV_n|\mathcal{F}_{n-1}) = \text{var}(r_n|\mathcal{F}_{n-1}) = v_n^2\tau^2,$$

see for instance Andersen, Bollerslev, Diebold, and Labys (2003). Generally $QV_n \neq \text{var}(r_n|\mathcal{F}_{n-1})$ so the conditional variance $v_n^2\tau^2$ remains unobservable. If one happens to be in the fortunate circumstance of having a perfect proxy, $H_n = v_n\tau_H$, then the QML estimation below yields perfect estimates. A second reason for considering likelihood methods is that one may want to study the dynamics of a sequence of proxies (H_n). These dynamics are determined by the volatilities (v_n). So the (v_n) are central to understanding the time series behaviour of, for example, the realized volatilities RV_n.

3.1 Gaussian QMLE

This section extends the usual Garch QMLE based on close-to-close returns to a QMLE based on the proxies H_n. For a brief review of the Garch(1,1) QMLE based on close-to-close returns, see Appendix B.3.

Recall that the intraday return process $R_n(\cdot) = v_n\tau\,\Psi_n(\cdot)$ yields close-to-close returns $r_n = v_n\tau Z_n$. From Section 2.2 we know that the volatility proxy H_n satisfies

$$H_n = v_n\tau_H\, Z_{H,n}. \tag{6}$$

Similarly to the case of squared returns one has the relation $\mathbb{E}(H_n^2|\mathcal{F}_{n-1}) = v_n^2\tau_H^2$. The volatility dynamics (v_n) and the autoregression parameters (γ, β) are the same as those for r_n. The norming parameter τ_H^0 is related to τ^0 for the returns r_n by

$$\tau_H^0 = \tau^0\mu_2^H, \tag{7}$$

reflecting that the overall scale of H_n may differ from the overall scale of the absolute returns $|r_n|$. The principle of quasi maximum likelihood may be applied to the multiplicative equation (6). First consider the absolute returns $|r_n|$. Treating these as absolute values of mean zero Gaussian random variables gives the same likelihood as simply treating the returns r_n as mean zero Gaussian random variables. Using the QML-notation of Appendix B.1 one may set the observation $y_n = |r_n|$, the conditional mean $\mu_n = 0$, and the conditional variance $h_n = v_n^2\tau^2$, since the Gaussian log-likelihood needs the value for $y_n^2 = r_n^2$ only, and not the value of r_n:

$$L_N(\theta;\ y_1,\dots,y_N) = -\frac{1}{2}\sum_{n=1}^{N}\left(\log(v_n^2(\gamma,\beta)\,\tau^2) + \frac{y_n^2}{v_n^2(\gamma,\beta)\,\tau^2}\right), \tag{8}$$

modulo an unimportant constant.

Similarly, treating H_n as if it were the absolute value of a mean zero Gaussian random variable yields a QML estimator for (τ_H,γ,β). So one may set $y_n = H_n$, $\mu_n = 0$, and $h_n = v_n^2\tau_H^2$, to obtain the QMLE $\hat{\theta}_N$. We refer to this QMLE as the *Gaussian QMLE* (based on H_n).

For notational convenience we write

$$\sigma_{H,n} = v_n\tau_H. \tag{9}$$

Equation (9) suppresses the parameter θ in $\sigma_{H,n} = \sigma_{H,n}(\theta)$ for $\theta = (\tau_H,\gamma,\beta)$. Define the matrix G_H by

$$G_H(\theta)_{i,j} = \mathbb{E}\big[\frac{1}{\sigma_{H,0}^4(\theta)}\big(\frac{\partial\ \sigma_{H,0}^2(\theta)}{\partial\theta_i}\big)\big(\frac{\partial\ \sigma_{H,0}^2(\theta)}{\partial\theta_j}\big)\big]. \tag{10}$$

The QML covariance matrix V_0, Appendix B.1 equation (28), now simplifies to the matrix given in (12). One obtains the regularity conditions for the Gaussian QMLE by adjusting the six conditions of Appendix B.3 for the QMLE based on close-to-close returns. One has to adjust the condition $\mathbb{E}Z^4 < \infty$ to $\mathbb{E}Z_H^4 < \infty$, and replace τ by τ_H in condition (2). One has to keep τ^0 in condition (3). This yields the following assumptions:

A1. (Z_n) is an iid sequence with $\mathbb{E}Z^2 = 1$,

A2. $\tau_H > 0,\ \gamma > 0,\ \beta \in [0,1)$,

A3. $\mathbb{E}\log(\gamma^0(\tau^0)^2Z^2 + \beta^0) < 0$,

A4. Z^2 is non-degenerate,

A5. $\mathbb{E}Z_H^4 < \infty$,

A6. $\mathbb{P}(|Z| \le z) = o(z^\mu)$ as $z \downarrow 0$, for some $\mu > 0$.

The only condition that concerns Z_H is (A5). Most conditions concern the innovation Z of the close-to-close returns r_n. This is because Z_n appears in the volatility process v_n, which is driven by the close-to-close returns. For more background on the conditions (A1) to (A6), see Appendix B.3.

Theorem 3.1. *Let $\theta^0 = (\tau_H^0, \gamma^0, \beta^0)$ and $\tau_H^0 = \tau^0 \mu_2^H$, see equation (7). Assume conditions (A1) to (A6). Then the Gaussian QMLE $\hat{\theta}_N$ is asymptotically normal:*

$$\sqrt{N}(\hat{\theta}_N - \theta^0) \xrightarrow{d} \mathcal{N}(0, V_0), \qquad N \to \infty, \tag{11}$$

with

$$V_0 = \mathrm{var}(Z_H^2)\ G_H^{-1}(\theta^0). \tag{12}$$

The proof of Theorem 3.1 consists of an adjustment of the proof of Straumann and Mikosch (2006) for the QMLE based on the returns $y_n = r_n$ to the case that $y_n = H_n$. One may find it in Appendix C.

Let us recall the notion of asymptotic relative efficiency. If two competing estimators $\hat{\phi}_N^{(1)}$ and $\hat{\phi}_N^{(2)}$ are consistent and asymptotically normal estimators of a parameter ϕ with asymptotic variances $(\sigma_\phi^{(1)})^2$ and $(\sigma_\phi^{(2)})^2$, then the *asymptotic relative efficiency* (ARE) is given by

$$ARE = (\sigma_\phi^{(1)})^2 / (\sigma_\phi^{(2)})^2.$$

The following lemma enables the comparison of the QML covariance matrices V_0 for estimators of γ and β based on alternative proxies H. The proof may be found in Appendix C.

Lemma 3.2. *The (γ, β)-block of $G_H^{-1}(\theta^0)$ in Theorem 3.1 does not depend on the particular proxy H.*

Corollary 3.3 below follows from Theorem 3.1 and Lemma 3.2.

Corollary 3.3. *Consider two Gaussian QMLE's for γ and β from Theorem 3.1, the first based on proxies H_n' and the other based on H_n. These estimators have asymptotic relative* *efficiency*

$$ARE_{Gaussian}(H', H) = \frac{\text{var}(Z^2_{H'})}{\text{var}(Z^2_H)}. \quad (13)$$

As a final remark, suppose that the volatilities v_n are a scale process other than Garch(1,1). One may then still extend the daily returns r_n to $R_n(\cdot) = v_n \tau \, \Psi_n(\cdot)$, and obtain results analogous to the results in the present section.

3.2 Log-Gaussian QMLE

One may also estimate the parameters (γ, β) of the Garch system given by (1) and (2) by a log-Gaussian QMLE. This section develops the log-Gaussian QMLE, similarly to the Gaussian QMLE. Readers may prefer to skip Sections 3.2 to 3.4 upon first reading, and proceed directly to the empirical results of Section 4.

The log-Gaussian QMLE consists of applying Gaussian quasi maximum likelihood to the log proxies $\log(H_n)$. Applying logarithms to H_n yields the equation $\log(H_n) = \log(v_n) + \log(\tau_H) + \log(Z_{H,n})$. Define $\tilde{\tau}_H = \tau_H \exp(\mathbb{E}\log(Z_{H,n}))$, and

$$U_{H,n} = \frac{\log(Z_{H,n}) - \mathbb{E}\log(Z_{H,n})}{\sqrt{\text{var}(\log(Z_{H,n}))}}.$$

We may now write the *additive equation*

$$\log(H_n) = \log(v_n) + \log(\tilde{\tau}_H) + \lambda U_{H,n}, \quad (14)$$

where the errors $U_{H,n}$ are iid(0,1). The system (14) yields $\mathbb{E}(\log(H_n)|\mathcal{F}_{n-1}) = \log(v_n) + \log(\tilde{\tau}_H)$, and $\text{var}(\log(H_n)|\mathcal{F}_{n-1}) = \lambda^2$. The parameter λ^2 represents the measurement variance of $\log(H_n)$, a proxy for log volatility, with

$$(\lambda^0)^2 = \text{var}(\log(Z_H)).$$

Define $\tilde{\theta} = (\tilde{\tau}_H, \gamma, \beta)$ and define the extended parameter

$$\eta = (\tilde{\theta}, \lambda).$$

The parameters γ, β in $\tilde{\theta}$ are the same as the γ, β in the parameter θ for the Gaussian

QMLE of Section 3.1. The additive equation (14) fits into the framework of quasi maximum likelihood estimation (see Appendix B.1), setting $y_n = \log(H_n)$, $\mu_n(\eta) = \log(\sigma_{H,n}(\tilde{\theta}))$ and $h_n(\eta) = \lambda^2$. We refer to the maximizer $\hat{\eta}_N$ as the *log-Gaussian* QMLE. Let us determine the QML covariance matrix V_0 of Appendix B.1. The matrix A_0 is block diagonal since the mean and variance functions do not share parameters. Applying

$$\frac{\partial \mu_n(\eta)}{\partial \eta_i} = \frac{1}{2\sigma^2_{H,n}(\tilde{\theta})} \frac{\partial \sigma^2_{H,n}(\tilde{\theta})}{\partial \eta_i},$$

one finds that the $\tilde{\theta}$-block and the diagonal element for λ of A_0 satisfy

$$(A_0)_{\tilde{\theta}} = \frac{1}{4(\lambda^0)^2} G_H(\tilde{\theta}^0), \quad (A_0)_\lambda = \frac{2}{(\lambda^0)^2},$$

with G_H given by equation (10). The $\tilde{\theta}$-block of B_0 equals the $\tilde{\theta}$-block of A_0, the diagonal element for λ equals $(B_0)_\lambda = \frac{1}{(\lambda^0)^2}\mathrm{var}(U_H^2)$. The off-diagonal $(\tilde{\theta}, \lambda)$-column of B_0 equals

$$(B_0)_{\tilde{\theta},\lambda} = \frac{1}{(\lambda^0)^2} \mathbb{E} U_H^3 \ \mathbb{E} \frac{\partial \mu_n}{\partial \tilde{\theta}} (\tilde{\theta}^0)',$$

making use of $\mu_n(\eta) = \mu_n(\tilde{\theta})$. The covariance matrix $V_0 = A_0^{-1} B_0 A_0^{-1}$ divided into $(\tilde{\theta}, \lambda)$-blocks now reads

$$V_0 = 4(\lambda^0)^2 \begin{pmatrix} G_H^{-1}(\tilde{\theta}^0) & \frac{1}{2}\mathbb{E}U_H^3 \ \mathbb{E}\frac{\partial \mu_n}{\partial \tilde{\theta}}(\tilde{\theta}^0)' \\ \frac{1}{2}\mathbb{E}U_H^3 \ \mathbb{E}\frac{\partial \mu_n}{\partial \tilde{\theta}}(\tilde{\theta}^0) & \frac{1}{16}\mathrm{var}(U_H^2) \end{pmatrix}. \tag{15}$$

Assume conditions (A1) to (A6) and replace condition (A5) by

A5'. $\mathbb{E}(\log(Z_H))^4 < \infty$.

The QML theory of Appendix B.1 suggests that the log-Gaussian QMLE $\hat{\eta}_N$ is asymptotically normal,

$$\sqrt{N}(\hat{\eta}_N - \eta_0) \xrightarrow{d} \mathcal{N}(0, V_0), \qquad N \to \infty, \tag{16}$$

with V_0 the covariance matrix given by (15), though we do not produce a formal proof like the proof of Theorem 3.1. The covariance matrix (15) makes clear that the smaller $(\lambda^0)^2 = \mathrm{var}(\log(Z_H))$ the more efficient the QMLE for γ and for β. Similarly to Corollary 3.3 the asymptotic relative efficiency of two log-Gaussian QMLE's for γ and β based on two different proxies H_n' and H_n is given by

$$ARE_{\text{log-Gaussian}}(H', H) = \frac{\text{var}(\log(Z_{H'}))}{\text{var}(\log(Z_H))}. \tag{17}$$

De Vilder and Visser (2007) define an optimal proxy H^* as a proxy with minimal variance of the logarithm,

$$\text{var}(\log(H^*(\Psi))) = \inf_H \text{var}(\log(H(\Psi))).$$

Such an optimal proxy also yields the most efficient log-Gaussian QMLE for γ and β.

We end this section with a remark that is relevant to practical implementation of the log-Gaussian QMLE. The numerical value of $\hat{\lambda}$ does not influence the numerical values of the parameters in $\tilde{\theta}$. This is due to the usual effect that the value of the variance parameter does not influence the value of the mean parameter for Gaussian QML (this is true if the variance function and the mean function do not share parameters). Moreover, the usual 'sandwich' QML covariance matrix $\hat{V}$ estimated by plugging in $\hat{A}$ and $\hat{B}$ in equation (28), also does not depend on the numerical value of $\hat{\lambda}$ as far as the $\tilde{\theta}$-parameters are concerned. So the value of $\hat{\lambda}$ is irrelevant to inference on $\tilde{\theta}$.

3.3 Efficiency of log-Gaussian QMLE versus Gaussian QMLE

Let us briefly compare the asymptotic efficiency of $\hat{\gamma}, \hat{\beta}$ for the log-Gaussian and Gaussian QMLE. Comparing the (γ, β)-blocks of V_0 in equations (12) and (15), one finds that the asymptotic relative efficiency of the log-Gaussian and Gaussian QMLE's for γ and β, based on the same proxy H_n is given by

$$ARE(\text{log-Gaussian}, \text{Gaussian}) = \frac{4\text{var}(\log(Z_H))}{\text{var}(Z_H^2)}. \tag{18}$$

So, the log-Gaussian QMLE is more efficient if $4(\lambda^0)^2 = \text{var}(\log(Z_H^2)) \leq \text{var}(Z_H^2)$, where $\mathbb{E}Z_H^2 = 1$. This inequality does not always hold: $\text{var}(\log(Z_H))$ may be large if Z_H has values close to zero, while $\text{var}(Z_H^2)$ may be large if Z_H has heavy tails. The following example considers the case that Z_H has a lognormal distribution.

Example 3.3.1. *Let Z_H have a lognormal$(-\sigma^2, \sigma^2)$ distribution. Then $\log(Z_H) \sim \mathcal{N}(-\sigma^2, \sigma^2)$. The j-th moment of a lognormal(μ, σ^2) equals $e^{j\mu + j^2\sigma^2/2}$, so $\mathbb{E}Z_H^2 = 1$ and $\text{var}(Z_H^2) = e^{4\sigma^2} - 1$.*

Apply relation (18) to find

$$ARE(\textit{log-Gaussian}, \textit{Gaussian}) = \frac{4\sigma^2}{e^{4\sigma^2} - 1}.$$

Since $4\sigma^4 \le e^{4\sigma^2} - 1$ *the log-Gaussian QMLE is more efficient for all values of* σ^2*. In this example the log-Gaussian QMLE is the exact maximum likelihood estimator.*

3.4 Relative Error of Volatility Extraction

One may also be interested in the quality of the estimator of the scale factor $\sigma_{H,n} = v_n \tau_H$, for some fixed n. The volatility extraction $\theta \to \hat{\sigma}_{H,n}(\theta)$, with initialization $\hat{v}_0$ is a function of θ. To simplify the notation we omit the hat on $\sigma_{H,n}$ in this section. If we plug in the estimator $\hat{\theta}_N$, we obtain the estimated volatility extraction $\sigma_{H,n}(\hat{\theta}_N)$. The asymptotic distribution of $\sigma_{H,n}(\hat{\theta}_N)$ for $N \to \infty$ may be found by the Delta method. Let the row vector $\dot{\sigma}_{H,n}$ denote the derivative of $\sigma_{H,n}$ with respect to θ. Let V_0 denote the asymptotic covariance matrix of θ. The Delta method gives

$$\sqrt{N}(\sigma_{H,n}(\hat{\theta}_N) - \sigma_{H,n}(\theta^0)) \to \mathcal{N}(\,0\,,\ \dot{\sigma}_{H,n}(\theta^0) V_0\, \dot{\sigma}_{H,n}(\theta^0)'), \qquad N \to \infty, \tag{19}$$

for fixed n. It is natural to look at the *relative error* of $\sigma_{H,n}$,

$$\mathrm{re}(\sigma_{H,n}) = \frac{\sigma_{H,n}(\hat{\theta}_N)}{\sigma_{H,n}(\theta^0)} - 1.$$

The relative error itself is not observed. One may estimate its variance by

$$\frac{1}{\sigma^2_{H,n}} \widehat{\mathrm{var}}(\sigma_{H,n}), \tag{20}$$

where $\widehat{\mathrm{var}}(\sigma_{H,n})$ is the empirical counterpart of the variance in equation (19). The estimate (20) does not depend on $\hat{\tau}_H$, see formula (33) in Appendix C. So the asymptotic variance of the relative error is proportional to $\mathrm{var}(Z_H^2)$ and $\mathrm{var}(\log(Z_H))$, for Gaussian and log-Gaussian estimation.

For practical implementation one needs the derivatives $\dot{\sigma}_{H,n}(\hat{\theta}_N)$. Let $h_n(\theta) = \sigma^2_{H,n}(\theta)$. The analytical derivatives $\dot{h}_n$ in $\theta = \hat{\theta}_N$ are available from the optimization procedure, so one may estimate the variance $\widehat{\mathrm{var}}(\sigma_{H,n})$ in equation (20) by a straightforward application of the Delta method, making use of the chain rule:

$$\dot{\sigma}_{H,n}(\hat{\theta}_N) = \frac{1}{2\sigma_{H,n}(\hat{\theta}_N)}\dot{h}_n(\hat{\theta}_N). \tag{21}$$

Of course, if one wishes to construct a confidence interval for $v_n\tau$, instead of $v_n\tau_H$, one has to carry out estimation based on the returns r_n.

4 Empirical Efficiency Gain for the S&P 500 Index

This section examines empirically the differences in efficiency of using alternative volatility proxies for the estimation of the Garch(1,1) parameters γ and β. The analysis is carried out for both the Gaussian and the log-Gaussian QMLE. The estimates in this section are based on 1001 days of S&P 500 index tick data over the period 1992–1995. For a description of the data, see Appendix A. We use this time period, since it is a fairly stable period without clear structural breaks in the level of volatility, see Figure 3 in Appendix A. We take care in avoiding structural breaks, since it is well known that Garch parameter estimation may break down in the presence of such breaks. Parameter estimators are no longer consistent, and the persistence of volatility tends to be overestimated if the level of volatility has a change-point, see Mikosch and Starica (2004), and Hillebrand (2005).

The efficiency of the QMLE's based on alternative proxies H is determined by the variance of Z_H^2 or the variance of its logarithm. For each proxy H we estimate the parameters by both the Gaussian and the log-Gaussian QMLE. We then use the standardized residuals, $\hat{Z}_{H,n}$, to compare the quality of the estimators. Table 1 provides an efficiency factor that expresses the efficiency gain with respect to the standard Garch(1,1) QMLE (as 1/ARE). The proxy H^* is constructed in de Vilder and Visser (2007). Moving down from absolute returns to H^* reveals an efficiency gain by a factor 15 for the Gaussian QMLE. The log-Gaussian QMLE yields an efficiency gain by a factor 20. This means that *estimation of* (γ, β) *based on* $\log(H^*)$ *needs roughly 20 times fewer days of observations than the usual QMLE based on squared close-to-close returns to obtain the same precision for the parameter estimates.* There are no entries for $H = |r|$ for the log-Gaussian QMLE since these would involve taking the log of zeros. The table reflects the differences in the confidence regions in Figure 1. Notice that in this figure the estimate based on $|r_n|$ is situated below and to the left of the other estimates. In the simulations below we observe a similar effect. This effect seems to be due to finite-sample bias.

The log-Gaussian QMLE outperforms the Gaussian QMLE for the proxies hl, $RV^{(81)}$, and H^*. One possible interpretation is that these proxies are closer to having the distribution of a lognormal random variable than to the absolute value of a Gaussian random variable. In empirical research it has been found that log realized volatility and the log high-low range may have a distribution that is nearly symmetrical and nearly Gaussian, see for instance Andersen, Bollerslev, Diebold and Ebens (2001), and Alizadeh, Brandt, and Diebold (2002).

We apply the Delta method of Section 3.4 to obtain the standard errors of the relative error in the volatility extraction. Table 2 lists these standard errors for the final scale factors $\sigma_{H,n}$, $n = N = 1001$. The first entry, 3.8%, suggests that the interval $\hat{\sigma}_{H,1001} \pm 7.6\%$ encloses the true $\sigma_{H,1001}$ with probability 95%. The log-Gaussian QMLE based on H^* gives a more than 4 times tighter interval. One should not interpret these percentages as typical for this Garch(1,1) process: they depend on the path of the process before $n = 1001$.

We also checked what Tables 1 and 2 would look like if they are based on the full sample over the years 1988–2006, $n = 1, \ldots, 4575$, (ignoring possible structural breaks). We briefly mention these results without providing the tables. For the full sample the patterns in both tables are similar to the patterns in Tables 1 and 2, though the efficiency gains in Table 1 become more pronounced: instead of a factor 20 for the log-Gaussian QMLE based on H^*, we find a gain by a factor more than 40.

	Gaussian		log-Gaussian	
H	$\widehat{\text{var}}(Z_H^2)$	eff. factor	$\widehat{\text{var}}(\log(Z_H^2))$	eff. factor
$\|r\|$	3.34	1	—	—
hl	1.41	2.4	0.68	4.9
$RV^{(81)}$	0.48	7.0	0.25	13.2
H^*	0.23	14.8	0.17	20.1

Table 1: Empirical QMLE efficiency for the volatility proxies: absolute return, high-low, realized volatility based on 81 five-minute intervals, and H^*. The table reports $\widehat{\text{var}}(Z_H^2)$ and $\widehat{\text{var}}(\log(Z_H^2))$, see Sections 3.1 and 3.2. The numbers are based on residuals of Garch(1,1) estimation of the S&P 500 over 1992–01–01 to 1995–12–31, or 1001 observations. The efficiency factor is the gain with respect to the usual Garch(1,1) QMLE, expressed as 1/ARE, so 2.4=3.34/1.41.

5 Finite-Sample Properties

The estimates for the S&P 500 in Section 4 are based on one sample path only. To explore the finite-sample properties of the QML estimators we perform simulations. Other places that provide simulations of the QMLE's for the Garch(1,1) parameters include Bollerslev

H	Gaussian $\widehat{s.e.}(\text{re}_N)$ %	log-Gaussian $\widehat{s.e.}(\text{re}_N)$ %
$\|r\|$	3.8	—
hl	2.2	1.7
$RV^{(81)}$	1.2	1.0
H^*	0.9	0.8

Table 2: Estimates of the standard error of the relative error in $\hat{\sigma}_{H,1001}$. The quantities reported are $100 \times \widehat{\text{s.e.}}(\hat{\sigma}_{H,N})/\hat{\sigma}_{H,N}$, see also equation (20). Numbers are based on the same volatility proxies and data as in Table 1.

and Wooldridge (1992), Lumsdaine (1995), Fiorentini, Calzolari, and Panattoni (1996), and Straumann (2005). The simulations in the present paper focus on the difference between the inference based on the close-to-close returns $H_n = |r_n|$ and inference by the square root of realized variance

$$H_n = RV_n^{(m)} = \sqrt{RQV_n^{(m)}}.$$

To generate the realized variance one has to simulate the process $\Psi(\cdot)$ at $(m+1)$ equidistant points in $[0,1]$. A Brownian motion will not do, since the realized volatility based on 81 intervals then has $\text{var}(\log(Z_H^2)) \approx 0.025$, which would yield unrealistically precise parameter estimates, *cf.* $RV^{(81)}$ in Table 1, which has $\text{var}(\log(Z_H^2)) \approx 0.25$ where $Z_H = RV^{(81)}(\Psi)$.

We consider an intraday diffusion, with an Ornstein-Uhlenbeck process for the log of the diffusion coefficient:

$$d\Psi(u) = \exp(Y(u))\, dB^{(1)}(u), \qquad u \in [0,1], \tag{22}$$

where $Y(u)$ is Ornstein-Uhlenbeck:

$$dY(u) = -\delta(Y(u) - \mu)du + \sigma_Y dB^{(2)}(u). \tag{23}$$

The Brownian motions $B^{(1)}$ and $B^{(2)}$ are uncorrelated, $\Psi(0) = 0$, $Y(0) = Y_0$. We sample Y_0 from its stationary distribution. For $\mu = -\sigma_Y^2/(2\delta)$, the realized variance $RQV^{(m)}(\Psi)$ for all m, as well as the quadratic variation over the unit interval have expectation 1, see Appendix D. Choose

$$\delta = \frac{1}{2}, \qquad \sigma_Y = \frac{1}{4}, \qquad \mu = -\frac{1}{8}.$$

Then the return innovations Z_n satisfy

$$\mathbb{E}Z^2 = 1, \qquad \mathrm{var}(Z^2) \approx 2.77.$$

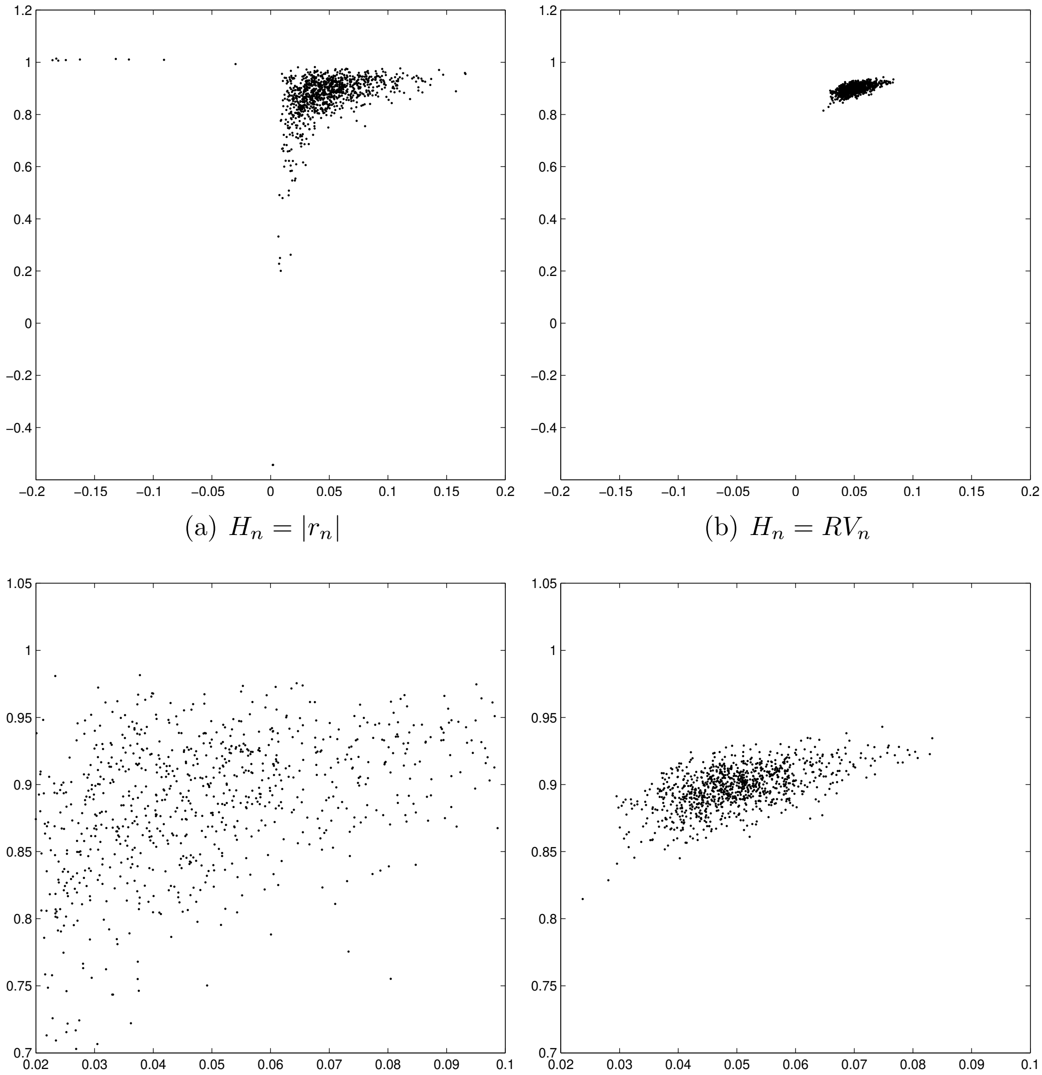
For the realized volatility we take $m = 81$ intervals, yielding innovations Z_H that satisfy

$$\mathbb{E}Z_H^2 = 1, \qquad \mathrm{var}(Z_H^2) \approx 0.27, \qquad \mathrm{var}(\log(Z_H^2)) \approx 0.24. \tag{24}$$

The simulations below consist of 10000 replications. First generate 10000 sets of 2500 days of realizations of Ψ. For each sequence (Ψ_n), $n = 1, \ldots, 2500$, we generate the paths $(v_n \tau)$ for five different configurations (γ, β), fixing $\tau = 1$. One may now examine the finite-sample properties of the Garch(1,1) QMLE's $(\hat{\gamma}, \hat{\beta})$ for sample lengths $N = 250,\ 500,\ 1000,\ 2500$. Figure 2 shows the estimates for 1000 of such paths for $(\gamma, \beta) = (0.05,\ 0.9)$ and sample length $N = 1000$ days. The left figures are based on absolute returns as a volatility proxy, the right figures are based on the realized volatility $RV_n^{(81)}$. The estimates based on $RV_n^{(81)}$ are more concentrated around the true parameter value, and have no outliers.

Table 3 provides a more complete overview of the finite-sample properties than Figure 2. The first two rows list 100 $\times$ the bias and 100 $\times$ the root mean square error (RMSE) of $\hat{\gamma}$ for $(\tau, \gamma, \beta) = (1,\ 0.05,\ 0.9)$. The first four columns in the first row contain the biases for the return based Garch(1,1) QMLE for increasing sample sizes. The next eight columns give this bias using the volatility proxy $RV_n^{(81)}$, for the Gaussian and the log-Gaussian QMLE.

While the small-sample biases of $\hat{\gamma}, \hat{\beta}$ tend to be substantial for the return based QMLE, they are moderate to negligible for the realized volatility based QMLE. The asymptotic relative efficiencies with respect to the usual Garch(1,1) QMLE may be deduced from equation (24) and equations (13) and (17). For the square root of realized variance this yields an efficiency factor $2.77/0.27 \approx 10$ for the Gaussian QMLE and efficiency factor 11 for the log-Gaussian QMLE. So the RMSE for $H_n = RV_n$ is more than a factor three smaller for large samples. This factor reflects the difference in RMSE between using returns or realized volatility, for $N = 2500$. For smaller sample sizes the efficiency gain is larger, suggesting that return based estimation suffers more from small-sample effects. The quality of the parameter estimates using 250 observations of realized volatility resembles using somewhere between 1000–2500 close-to-close returns. As predicted by the asymptotic efficiency factors for RV_n computed above (11 versus 10), the log-Gaussian QMLE does slightly better than the Gaussian QMLE.



(a) $H_n = |r_n|$

(b) $H_n = RV_n$

(c) $H_n = |r_n|$, zoomed in

(d) $H_n = RV_n$, zoomed in

Figure 2: Scatters for $(\hat{\gamma}, \hat{\beta})$ plane, 1000 sample paths ($\tau = 1, \gamma = 0.05, \beta = 0.9$). The intraday process $\Psi_n(\cdot)$ is given by equations (22) and (23) with ($\delta = 0.5$, $\sigma_Y = 0.25$, $\mu = -0.125$). Upper and lower left: estimates based on absolute returns (Gaussian QMLE). Upper and lower right: realized volatility (Gaussian QMLE). Figure (a) leaves out four points where $\hat{\gamma} > 0.2$. Figures (b) and (d) contain all points.

$\hat{\gamma}, \hat{\beta}$ Sampling Distributions

$100 \times$ bias, $100 \times$ RMSE

param	true		$H_n = \lvert r_n \rvert$; Gaussian QMLE				$H_n = RV_n^{(m=81)}$; Gaussian QMLE				$H_n = RV_n^{(m=81)}$; log-Gaussian QMLE			
			N=250	N=500	N=1000	N=2500	N=250	N=500	N=1000	N=2500	N=250	N=500	N=1000	N=2500
γ	0.05	bias	-4.8	-1.4	-0.2	-0.1	0.3	0.0	-0.0	-0.0	0.2	0.0	-0.0	-0.0
		RMSE	13.6	8.3	3.8	1.7	3.0	1.4	0.9	0.5	2.7	1.3	0.9	0.5
β	0.9	bias	-4.0	-4.7	-2.7	-0.9	-1.1	-0.4	-0.2	-0.1	-1.0	-0.4	-0.2	-0.1
		RMSE	22.5	17.3	10.3	4.0	5.2	2.5	1.6	0.9	4.8	2.3	1.5	0.9
γ	0.15	bias	-1.7	-0.7	-0.3	-0.2	0.0	-0.1	-0.1	-0.0	-0.0	-0.1	-0.0	-0.0
		RMSE	13.6	8.1	5.3	3.3	4.1	2.5	1.7	1.0	3.9	2.3	1.6	1.0
β	0.8	bias	-5.1	-2.2	-1.0	-0.4	-0.4	-0.2	-0.1	-0.0	-0.3	-0.1	-0.1	-0.0
		RMSE	17.4	8.4	4.7	2.7	3.1	2.0	1.3	0.8	2.9	1.9	1.3	0.8
γ	0.35	bias	-0.2	-0.2	-0.1	-0.1	-0.0	-0.1	-0.0	-0.0	-0.1	-0.1	-0.0	-0.0
		RMSE	21.3	13.3	9.2	5.7	6.4	4.2	2.8	1.8	6.0	3.9	2.7	1.7
β	0.6	bias	-3.4	-1.3	-0.7	-0.3	-0.2	-0.1	-0.1	-0.0	-0.2	-0.1	-0.1	-0.0
		RMSE	13.5	8.0	5.2	3.2	3.5	2.3	1.6	1.0	3.3	2.2	1.5	0.9
γ	0.25	bias	0.3	0.1	0.0	0.0	0.0	-0.0	-0.0	-0.0	0.0	-0.0	-0.0	-0.0
		RMSE	18.3	10.4	7.0	4.3	4.8	3.2	2.2	1.3	4.5	3.0	2.0	1.3
β	0.6	bias	-5.9	-2.5	-1.2	-0.5	-0.4	-0.2	-0.1	-0.0	-0.4	-0.2	-0.1	-0.0
		RMSE	20.8	12.3	7.7	4.7	5.1	3.3	2.3	1.4	4.8	3.2	2.2	1.3
γ	0.05	bias	-6.8	-1.9	0.4	0.4	0.8	0.2	0.0	0.0	0.7	0.2	0.1	0.0
		RMSE	26.5	18.8	9.1	2.2	3.5	1.7	1.1	0.6	3.4	1.5	1.0	0.6
β	0.8	bias	-10.0	-10.3	-6.7	-3.0	-3.0	-1.3	-0.6	-0.2	-2.8	-1.1	-0.5	-0.2
		RMSE	37.9	33.5	24.1	13.2	15.0	7.9	4.8	2.8	14.2	7.3	4.5	2.6
γ	0.05	bias	-5.4	-2.2	-0.9	-0.4	-0.6	-0.1	-0.1	-0.1	-0.6	-0.2	-0.1	-0.1
		RMSE	7.7	4.7	3.1	1.9	2.9	2.0	1.2	0.6	2.7	1.9	1.1	0.6
β	0.94	bias	-0.6	-1.8	-0.9	-0.2	-0.6	-0.2	-0.1	-0.0	-0.6	-0.2	-0.1	-0.0
		RMSE	14.5	8.5	3.5	1.2	2.1	1.0	0.6	0.4	2.0	0.9	0.6	0.3

Table 3: Sampling distributions of Garch(1,1) QMLE, based on 10000 replications. The intraday process $\Psi_n(\cdot)$ is given by equations (22) and (23) with ($\delta = 0.5,\ \sigma_Y = 0.25,\ \mu = -0.125$). All simulations use $\tau = 1$. From top to bottom there are six panels of different parameters (γ, β). For each parameter setting the table gives $100 \times$ the bias and $100 \times$ the root mean squared error of $\hat{\gamma}$ and $\hat{\beta}$, for different lengths of the time series: 250, 500, 1000, 2500. The case $H_n = \lvert r_n \rvert$, based on the Gaussian QML is the usual Garch(1,1) QMLE. The cases $H_n = RV_n$ give the results for the realized volatility based on 81 intraday returns as a volatility proxy, using the Gaussian and the log-Gaussian QMLE, see Sections 3.1 and 3.2.

6 Conclusions

This paper develops Garch quasi maximum likelihood estimation based on intraday volatility proxies. One may achieve a substantial efficiency gain by using a suitable volatility proxy other than the absolute or squared close-to-close return. The paper starts out from the Garch(1,1) system

$$\begin{aligned} r_n &= v_n \tau \, Z_n \\ v_n^2 &= 1 + \gamma r_{n-1}^2 + \beta v_{n-1}^2, \end{aligned}$$

and makes use of the extension of the returns r_n to the intraday return process $R_n(u) = v_n \tau \Psi_n(u), \quad u \in [0,1]$, where the processes $\Psi_n(\cdot)$ are iid over different days. The setup does not make particular assumptions for the process Ψ_n. One obtains sharp estimators $\hat{\gamma}, \ \hat{\beta}$ by making use of a suitable volatility proxy $H(R_n)$. Here, H is positive and positively homogeneous. For the S&P 500 index data the estimated variances of the estimators decrease by a factor 20. The QMLE has the additional advantage that it does not require the usual condition that the conditional fourth moment of the close-to-close returns is finite. The QMLE works provided that the proxy H has a finite conditional fourth moment.

A good parameter estimation for financial processes is important for several reasons. It gives better predictions for future market behaviour. A sharp estimation procedure may also clear up fundamental questions around the stationarity of certain financial processes. Do parameters change over time? Is this change slow or abrupt? We hope that the results in this paper help to find answers to such questions in the future.

The intraday extension employed in this paper and the resulting QML theory apply equally well to other volatility models. It would be interesting to apply the methods of this paper to asymmetric Garch models, or to models where the volatility v_n is driven by statistics different from the squared return r_{n-1}^2. For instance, from Andersen *et al.* (2003) we know that a log-ARFIMA model for realized volatilities fits well. One may expect that realized volatilities could also enhance the latent volatilities v_n. We leave this to future research.

7 Acknowledgment

The author thanks Guus Balkema, Peter Boswijk, Chris Klaassen, Remco Peters, and Robin de Vilder for detailed comments and suggestions.

Appendices

A Data

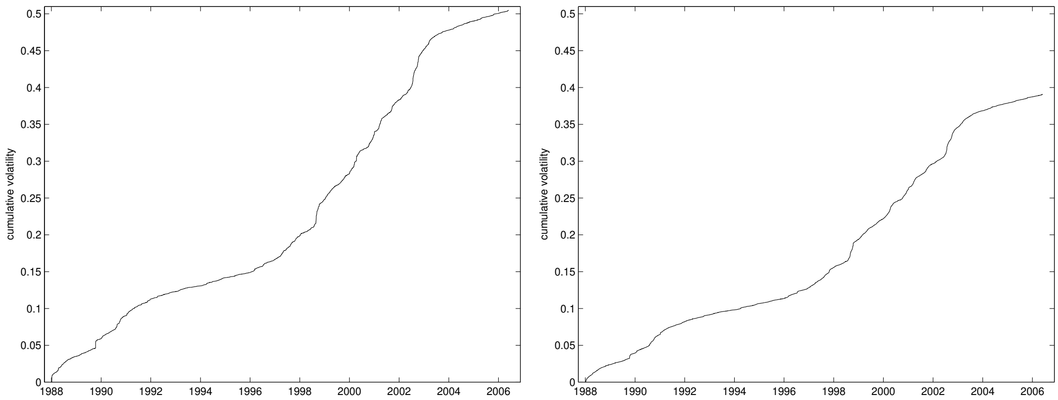
Our data set is the U.S. Standard & Poor's 500 stock index future, traded on the Chicago Mercantile Exchange (CME), for the period 1st of January, 1988 until May 31st, 2006. The data were obtained from Nexa Technologies Inc. (www.tickdata.com). The futures trade from 8:30 A.M. until 15:15 P.M. Central Standard Time. Each record in the set contains a timestamp (with one second precision) and a transaction price. The tick size is \$0.05 for the first part of the data and \$0.10 from 1997–11–01. The data set consists of 4655 trading days. We removed sixty four days for which the closing hour was 12:15 P.M. (early closing hours occur on days before a holiday). Sixteen more days were removed, either because of too late first ticks, too early last ticks, or a suspiciously long intraday no-tick period. These removals leave us with a data set of 4575 days with nearly 14 million price ticks, on average more than 3 thousand price ticks per day, or 7.5 price ticks per minute.

There are four expiration months: March, June, September, and December. We use the most actively-traded contract: we roll to a next expiration as soon as the tick volume for the next expiration is larger than for the current expiration.

Figure 3 gives an impression of the course of volatility over the years 1988–2006. It depicts the cumulative of volatility. The left figure is based on squared daily close-to-close returns, the right one on the daily realized variance based on five-minute returns. The slope in the figure based on realized variance is smaller, since it does not take into account the overnight return. The growth of cumulative volatility is low in certain periods and high in other periods. The years 1992–1995 form a period without clear qualitative changes in the level of volatility. The empirical analysis in Section 4 is based on these four years.

B Quasi Maximum Likelihood

This section contains background for the QML theory presented in Section 3. Sections B.1 and B.2 discuss QML estimation and the regularity conditions. Section B.3 briefly discusses the standard Garch(1,1) QMLE.



(a) Cumulative of r_n^2

(b) Cumulative of RQV_n

Figure 3: S&P 500 cumulative volatility over the years 1988–2006. Figure (a) estimates cumulative volatility by the sum of squared daily close-to-close returns. Figure (b) shows the cumulative of the daily realized variance, RQV_n, based on 81 five-minute returns.

B.1 Principle of QML

The estimation method used in this paper is quasi maximum likelihood (QML). Let us briefly describe the principle of Gaussian quasi maximum likelihood estimation, as discussed in Bollerslev and Wooldridge (1992). Let (y_n) be a stationary sequence adapted to the filtration $(\mathcal{F}_n)$. The conditional mean and variance functions $\mu_n(\theta), h_n(\theta)$ are parameterized by a finite dimensional parameter θ and there is a true value $\theta^0 \in \Theta$ in the sense that

$$\mu_n(\theta^0) = \mathbb{E}(y_n|\mathcal{F}_{n-1}), \quad h_n(\theta^0) = \text{var}(y_n|\mathcal{F}_{n-1}), \tag{25}$$

for all n. The likelihood of the sample $(y_1, \dots, y_N)$ is a function of θ. The parameter θ may be estimated by maximizing the Gaussian likelihood, even if the true conditional probability distribution of y_n is not Gaussian. The likelihood is constructed then as if y_n is $\mathcal{N}(\mu_n, h_n)$, and is called quasi-likelihood. Let the residual function $\varepsilon_n(\theta) = \varepsilon_n(y_n, \theta)$ denote the standardized y_n,

$$\varepsilon_n(\theta) = \frac{y_n - \mu_n(\theta)}{\sqrt{h_n(\theta)}}.$$

This leads to a log-likelihood

$$L_N(\theta) = \sum_{n=1}^{N} l_n(\theta), \tag{26}$$

where, by the Gaussian likelihood,

$$l_n(\theta) = -\frac{1}{2}\,[\log(2\pi) + \log(h_n(\theta)) + \varepsilon_n(\theta)^2].$$

Let the QMLE $\hat{\theta}_N$ denote the maximizer of the log-likelihood. Under regularity (see Appendix B.2) the QMLE is asymptotically normal,

$$\sqrt{N}(\hat{\theta}_N - \theta^0) \xrightarrow{d} \mathcal{N}(0, V_0), \qquad N \to \infty, \tag{27}$$

where

$$V_0 = A_0^{-1} B_0 A_0^{-1}. \tag{28}$$

The matrices A_0 and B_0 are given by the expected Hessian and the expectation of the outer product of the scores (which is the covariance matrix of the scores):

$$(A_0)_{i,j} = -\mathbb{E}\frac{\partial^2 l_0(\theta^0)}{\partial\theta_i \partial\theta_j}, \quad (B_0)_{i,j} = \mathbb{E}\; s_{0,i}(\theta^0) s_{0,j}(\theta^0),$$

where, using stationarity, the expectation is taken at time $n = 0$. The scores $s_{n,i}(\theta)$ are given by

$$s_{n,i}(\theta) \;\;=\;\; \frac{\partial l_n(\theta)}{\partial\theta_i} = \frac{\varepsilon_n(\theta)}{\sqrt{h_n(\theta)}}\Big(\frac{\partial\mu_n(\theta)}{\partial\theta_i}\Big) + \frac{\varepsilon_n^2(\theta) - 1}{2h_n(\theta)}\Big(\frac{\partial h_n(\theta)}{\partial\theta_i}\Big).$$

The expected Hessian A_0 may be expressed as

$$(A_0)_{i,j} = \mathbb{E}\Big[\frac{1}{h_0(\theta^0)}\Big(\frac{\partial\mu_0(\theta^0)}{\partial\theta_i}\Big)\Big(\frac{\partial\mu_0(\theta^0)}{\partial\theta_j}\Big) + \frac{1}{2h_0^2(\theta^0)}\Big(\frac{\partial h_0(\theta^0)}{\partial\theta_i}\Big)\Big(\frac{\partial h_0(\theta^0)}{\partial\theta_j}\Big)\Big].$$

If the true conditional probability distribution is Gaussian, the QMLE reduces to the Gaussian maximum likelihood estimator and the information matrix equality $A_0 = B_0$ holds, so V_0 reduces to A_0^{-1}, and the QMLE is efficient.

B.2 QML Regularity Conditions

Bollerslev and Wooldridge (1992) provide abstract regularity conditions allowing for additional regressors (x_n), and without assuming stationarity for (y_n). We restate these conditions below, assuming stationarity, and leaving out x_n. The scores s_n are row vectors. Let $\ddot{l}_n$ denote the Hessian of $l_n(\theta)$, so $\ddot{l}_n = \dot{s}_n$. We first state the definition of the Uniform Weak Law of Large Numbers, as given by Wooldridge (1990, Definition A.1). A sequence of random functions $q_n(y_n, \theta)$ satisfies the UWLLN if

$$\sup_{\theta \in \Theta} |N^{-1} \sum_{n=1}^{N} q_n(y_n, \theta) - \mathbb{E}q_n(y_n, \theta)| \xrightarrow{P} 0, \qquad N \to \infty.$$

The QML regularity conditions are:

1. Θ is compact, has nonempty interior and $\theta^0 \in \text{int } \Theta$.
2. The mean and variance functions μ_n, h_n are measurable functions of the data for all $\theta \in \Theta$, are twice continuously differentiable with respect to θ on int Θ, and the variance is nonsingular (with probability one), for all $\theta \in \Theta$.
3. (a) $(l_n(\theta))$ satisfies the UWLLN.
   (b) θ^0 is the identifiably unique maximizer of $\mathbb{E}l_n(\theta)$.
4. (a) The Hessians $(\ddot{l}_n(\theta))$ satisfy the UWLLN.
   (b) The expected Hessian $A_0 = \mathbb{E}\ddot{l}_n(\theta^0)$ is positive definite.
5. (a) The expected outer product $B_0 = \mathbb{E}s_n' s_n(\theta^0)$ is positive definite.
   (b) $\frac{1}{\sqrt{N}} B_0^{-1/2} \sum s_n'(\theta^0) \xrightarrow{d} \mathcal{N}(0, I_p), \qquad N \to \infty.$
6. The outer product of the scores $(s_n' s_n(\theta))$ satisfies the UWLLN.

B.3 QML Regularity Conditions for Garch(1,1)

The verification of the conditions for asymptotic normality of quasi maximum likelihood given in Appendix B.2, has to be carried out on a case-by-case basis. The Garch(1,1) system (1) and (2) corresponds to $y_n = r_n$, $\mu_n(\theta) = 0$, $h_n(\theta) = v_n^2(\gamma, \beta)\tau^2$, with $\theta = (\tau, \gamma, \beta)$. In the case of a Garch type process a problem is that one cannot evaluate the exact likelihood for a given parameter θ, since the unobservable volatilities $v_n(\theta)$ have to be replaced by approximations

$\hat{v}_n(\theta)$. The unobserved volatility is approximated by the volatility recursion, with initialization $\hat{v}_0^2 > 0$. There are several papers on the Gaussian QMLE for Garch(1,1) including Lee and Hansen (1994), Lumsdaine (1996), Berkes, Horvath, and Kokoszka (2003), and Francq and Zakoïan (2004). The QMLE $\hat{\theta}_N$ satisfies the asymptotic normality of equation (27) and one may consistently estimate the covariance matrix V_0 by using the empirical counterparts of A_0 and B_0; we refer to Straumann and Mikosch (2006) for the following regularity conditions, see also the monograph of Straumann (2005). The observations $y_1, \ldots, y_N$ are part of a stationary sequence (y_n) that satisfies (*cf.* (1) and (2))

$$y_n = v_n \tau Z_n \tag{29}$$

$$v_n^2 = 1 + \gamma\tau^2 v_{n-1}^2 Z_{n-1}^2 + \beta v_{n-1}^2, \tag{30}$$

where

1. (Z_n) is an iid sequence with $\mathbb{E}Z^2 = 1$,
2. $\tau > 0,\ \gamma > 0,\ \beta \in [0, 1)$,
3. $\mathbb{E}\log(\gamma^0(\tau^0)^2 Z^2 + \beta^0) < 0$,
4. Z^2 is non-degenerate,
5. $\mathbb{E}Z^4 < \infty$,
6. $\mathbb{P}(|Z| \le z) = o(z^\mu)$ as $z \downarrow 0$, for some $\mu > 0$.

Condition (6) is fulfilled if Z has a density that is bounded in a neighbourhood of zero. Straumann and Mikosch (2006) also require $\mathbb{E}Z = 0$ in condition (1) to ensure that y_n has mean zero. As we observe in Section 3.1, the requirement $\mathbb{E}Z = 0$ is not needed, see also Francq and Zakoïan (2004). To be precise, one should read condition (2) as: Θ is a compact subset of the space given by condition (2), and $\theta^0 \in \operatorname{int}\Theta$. Condition (3) is the usual condition for strict stationarity and ergodicity of the Garch process. If $\gamma^0(\tau^0)^2 + \beta^0 < 1$ then condition (3) is fulfilled by Jensen's inequality, and in addition the process is weakly stationary. Condition (4) is needed for the identifiability of θ. For consistency it suffices that $\mathbb{E}Z^2 < \infty$, but condition (5) is necessary for asymptotic normality of the Gaussian QMLE. Instead of (Z_n) iid in condition (1), Lee and Hansen (1994) use the weaker constraint that (Z_n) is strictly stationary, ergodic. They require that $\mathbb{E}(Z_n^4|\mathcal{F}_{n-1})$ is uniformly bounded, and that $\sup_n \mathbb{E}(\log(\gamma^0(\tau^0)^2 Z_n^2 + \beta^0)|\mathcal{F}_{n-1}) < 0$.

C Proofs

Proof of Theorem 3.1. The proof of Theorem 3.1 in the present paper applies the likelihood theory of Straumann and Mikosch (2006). The asymptotic normality of the usual Garch(1,1) QMLE follows from Theorem 8.1 of Straumann and Mikosch (2006). The proof of that theorem relies on their more general Theorem 7.1. We extend the assumptions needed to invoke Theorem 8.1 in Straumann and Mikosch, check that this set of assumptions establishes asymptotic normality of the Gaussian QMLE in the present paper, and then remove the redundant assumptions. We collected the conditions for the usual Gaussian QMLE based on close-to-close returns as conditions (1) to (6) in our Appendix B.3. Let us extend these assumptions by duplication: copy the conditions for τ and Z to τ_H and Z_H : assume $\tau_H > 0$ and add to each condition for Z the same condition for Z_H. We now have a set of (temporary) conditions (D1) to (D6), concerning both Z and Z_H.

Under conditions (D1) to (D4) the usual Garch model satisfies the consistency conditions (C1) to (C4) of Straumann and Mikosch, pp. 2473 (for a verification, see their Section 5.2). Let us first verify that the Gaussian QMLE in the present paper is consistent. Let $L_{H,N}(\theta) = \sum_{n=1}^{N} l_{H,n}(\theta)$ denote the log-likelihood (modulo a constant), where

$$\begin{aligned} l_{H,n}(\theta) &= -\frac{1}{2}\left(\log(h_n(\theta)) + H_n^2/h_n(\theta)\right) \\ &= -\frac{1}{2}\left(\log(h_n(\theta)) + \frac{v_n^2(\gamma^0,\beta^0)(\tau_H^0)^2 Z_{H,n}^2}{h_n(\theta)}\right), \end{aligned}$$

and $h_n(\theta) = v_n^2(\gamma,\beta)\tau_H^2$. It is important to note that the innovation $Z_{H,n}$ is independent of $h_n(\theta)$ and v_n and satisfies $\mathbb{E}Z_{H,n}^2 = 1$. The function $L(\theta) = \mathbb{E}l_{H,0}(\theta)$ equals

$$L(\theta) = -\frac{1}{2}\mathbb{E}\left(\log(h_0(\theta)) + \frac{v_0^2(\gamma^0,\beta^0)(\tau_H^0)^2}{h_0(\theta)}\right).$$

One may now follow the proof of Theorem 4.1 of Straumann and Mikosch, pp. 2473 part 1.i, to obtain that $L_{H,N}/N$ converges to L uniformly. The rest of the proof of Theorem 4.1 needs no adjustment and shows that the QMLE converges almost surely to $(\tau_H^0, \gamma^0, \beta^0)$.

Straumann and Mikosch, Section 7, treat the asymptotic normality of their general QMLE under their assumptions (N1) to (N4), see Theorem 7.1. One may follow their exposition, replacing X by H, until the second display on pp. 2488, for which we may write

$$\dot{L}_{H,n}(\theta^0) = \sum_{n=1}^{N} \dot{l}_{H,n}(\theta^0) = \frac{1}{2}\sum_{n=1}^{N} \frac{\dot{h}_n(\theta^0)}{h_n(\theta^0)}(Z_{H,n}^2 - 1),$$

where $\dot{l}_{H,n}(\theta^0)$ is a martingale difference sequence since $Z_{H,n}$ is independent of $\mathcal{F}_{n-1}$ and $\mathbb{E}Z_{H,n}^2 = 1$. Accordingly one may apply the central limit theorem for martingale differences, assuming $\mathbb{E}Z_H^4 < \infty$. So, an application of Theorem 7.1 to the Gaussian QMLE in the present paper needs $\mathbb{E}Z_H^4 < \infty$, which is satisfied by (D5).

Under conditions (1) to (6) of Appendix B.3, the standard Garch model satisfies condition (N1) to (N4) of Straumann and Mikosch, see also their Theorem 8.1. For the Gaussian QMLE of the present paper we have to establish (N1) to (N4) under our duplicated conditions (D1) to (D6). The only conditions that are left for reexamination are conditions N3.iii and N3.iv: $\mathbb{E}||\dot{l}_0||_\Theta < \infty$, and $\mathbb{E}||\ddot{l}_0||_\Theta < \infty$. Let us follow the lines of Section 8 of Straumann and Mikosch. We may write

$$\mathbb{E}||H_0^2/h_0||_\Theta^\nu \quad = \quad \mathbb{E}||h_0(\theta^0)/h_0(\theta)||_\Theta^\nu \, \mathbb{E}Z_{H,0}^{2\nu}.$$

By (D1), $\mathbb{E}Z^2 < \infty$, and by (D6): $\mathbb{P}(|Z| \le z) = o(z^\mu)$ as $z \downarrow 0$. So by Lemma 5.1 of Berkes *et al.* (2003) one has $\mathbb{E}||h_0(\theta^0)/h_0(\theta)||_\Theta^\nu < \infty$, for $0 \le \nu < 1$. Therefore

$$\mathbb{E}||H_0^2/h_0||_\Theta^\nu < \infty, \qquad 0 \le \nu < 1.$$

One may now follow the arguments of Straumann and Mikosch to establish their conditions N.3.iii and N.3.iv. This establishes the asymptotic normality of the Gaussian QMLE of Theorem 3.1 in the present paper.

Let us finally remove the redundant conditions from (D1) to (D6), and establish conditions (A1) to (A6) of Section 3.1. The assumption $\mathbb{E}Z_H^4 < \infty$ and equation (6) already imply that $(Z_{H,n})$ is an iid sequence with $\mathbb{E}Z_H^2 = 1$, yielding (A1). One should read condition (2) of Appendix B.3 as a description of the parameter space. This does not need $\tau > 0$, since we optimize L_H over τ_H, not τ. Furthermore $\tau^0 > 0$ is equivalent to $\tau_H^0 > 0$ by equation (7), hence (A2). Condition (D3) is used for establishing stationarity, ergodicity, and invertibility of (v_n). These properties do not rely on the innovations $Z_{H,n}$, yielding (A3). Condition (D4) helps to establish that v_n is uniquely determined by θ, again a property that does not depend on Z_H, hence (A4). Conditions (D5) and (D6) are used to establish asymptotic normality. Condition (D5) is needed to obtain a finite variance in the application of the martingale difference central limit theorem to the derivative of $L_{H,N}$, which only requires $\mathbb{E}Z_H^4 < \infty$, and not $\mathbb{E}Z^4 < \infty$, see the arguments above. Consider assumption (D6): $\mathbb{P}(|Z| \leq z) = o(z^\mu) \quad \text{as } z \downarrow 0$, for some $\mu > 0$. This assumption helps to establish $||h_0(\theta^0)/h_0(\theta)||_\Theta^\nu < \infty$, for all $0 \leq \nu < 1$, see above. This does not depend on Z_H, hence (A6). □

Proof of Lemma 3.2. Differentiation yields $\frac{\partial \sigma^2_{H,0}(\theta)}{\partial \tau_H} = 2\tau_H v_0^2(\theta)$, $\frac{\partial \sigma^2_{H,0}(\theta)}{\partial \gamma} = \tau_H^2 \frac{\partial v_0^2(\theta)}{\partial \gamma}$, and $\frac{\partial \sigma^2_{H,0}(\theta)}{\partial \beta} = \tau_H^2 \frac{\partial v_0^2(\theta)}{\partial \beta}$, so

$$G_H(\theta) = \mathbb{E}\begin{pmatrix} \frac{4}{\tau_H^2} & \frac{2}{\tau_H v_0^2}\frac{\partial v_0^2}{\partial \gamma} & \frac{2}{\tau_H v_0^2}\frac{\partial v_0^2}{\partial \beta} \\ \frac{2}{\tau_H v_0^2}\frac{\partial v_0^2}{\partial \gamma} & \frac{1}{v_0^4}(\frac{\partial v_0^2}{\partial \gamma})^2 & \frac{1}{v_0^4}\frac{\partial v_0^2}{\partial \gamma}\frac{\partial v_0^2}{\partial \beta} \\ \frac{2}{\tau_H v_0^2}\frac{\partial v_0^2}{\partial \beta} & \frac{1}{v_0^4}\frac{\partial v_0^2}{\partial \gamma}\frac{\partial v_0^2}{\partial \beta} & \frac{1}{v_0^4}(\frac{\partial v_0^2}{\partial \beta})^2 \end{pmatrix}_\theta . \tag{31}$$

The lower right block of the inverse of a matrix

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix},$$

equals $C^{-1} = (A_{22} - A_{21}A_{11}^{-1}A_{12})^{-1}$. So the $\nu = (\gamma, \beta)$ block of G^{-1} equals the inverse of the 2×2 matrix given by

$$(C)_{i,j} = \text{cov}(\frac{1}{v_0^2}\frac{\partial v_0^2}{\partial \nu_i}, \frac{1}{v_0^2}\frac{\partial v_0^2}{\partial \nu_j}). \tag{32}$$

Formula (32) does not depend on H. □

On the relative error $re(\sigma_{H,n})$ in Section 3.4. Let $h_n(\theta) = \sigma^2_{H,n}(\theta) = v_n^2\tau_H^2$. The derivative of h_n is given by

$$\dot{h}_n = \begin{pmatrix} 2\tau_H\, v_n^2 & \tau_H^2(r_{n-1}^2 + \beta\frac{\partial v_{n-1}^2}{\partial \gamma}) & \tau_H^2(v_{n-1}^2 + \beta\frac{\partial v_{n-1}^2}{\partial \beta}) \end{pmatrix}.$$

The Gaussian QMLE $\hat{\theta}_N$ has asymptotic variance $V_0 = \text{var}(Z_H^2)G_H^{-1}(\theta^0)$. The asymptotic variance ($N \to \infty$) of $h_n(\hat{\theta}_N)$ is given by $V_{h_n} = \dot{h}_n(\theta^0)V_0\dot{h}_n(\theta^0)'$. Partition the matrix G in (31) into τ_H and (γ, β) blocks. Using partitioned inverses one finds that the asymptotic variance of h_n (for fixed n) equals

$$V_{h_n} = c\,\text{var}(Z_H^2)\tau_H^4,$$

where c is a constant that does not depend on H. The asymptotic variance of $\sigma_{H,n} = \sqrt{h_n}$

may be obtained by the Delta method using formula (21):

$$V_{\sigma_{H,n}} = \frac{1}{4\sigma_{H,n}^4(\theta^0)} V_{h_n}. \tag{33}$$

One sees that the parameter τ_H drops out. □

D Realized Variance of Ornstein-Uhlenbeck Log-Volatility

Consider the intraday process $\Psi(\cdot)$ of Section 5. The accompanying volatility process $Y(\cdot)$ satisfies

$$Y(u) = \exp(-\delta u)Y(0) + \mu(1 - \exp(-\delta u)) + \sigma_Y \int_{s=0}^{u} \exp(-\delta(u-s))\ dB^{(2)}(s).$$

Simulation of the process Y is straightforward since $Y(u+\Delta)|Y(u)$ has a normal distribution with mean

$$\exp(-\delta\Delta)Y(u) + (1 - \exp(-\delta\Delta))\mu,$$

and variance

$$\sigma_Y^2 \frac{1}{2\delta}(1 - \exp(-2\delta\Delta)),$$

see for instance Glasserman (2003). The process Y has a stationary version which is normally distributed with mean μ and variance $\frac{\sigma_Y^2}{2\delta}$. We sample Y_0 from this stationary distribution: this yields a simple expression for the expectation of the realized quadratic variation using step size Δ. We shall use that the expectation of the squared increment in Ψ equals the expectation of the increment in the quadratic variation. The expected increment in QV equals

$$\begin{aligned}
\mathbb{E}QV[u, u+\Delta] &= \int_{s=u}^{t+\Delta} \mathbb{E}\exp(2Y(s))ds \\
&= \int_{s=u}^{u+\Delta} \mathbb{E}\exp(2Y_0)ds \\
&= \exp(2\mu + 2\frac{\sigma_Y^2}{2\delta})\ \Delta.
\end{aligned}$$

So, for $\mu = -\sigma_Y^2/(2\delta)$, the quadratic variation over the unit interval has expectation 1. This implies that the realized variance $RQV^{(m)}$ has expectation 1 for all m. We simulate a realized variance based on $m = 81$ intervals. Each of those intervals is divided into 10 subintervals using equally spaced grid points. The simulation of the process Y on all grid points is exact. The value of Ψ on each grid point is obtained by Euler discretization.

References

Alizadeh, S., Brandt, M.W. and Diebold, F.X. (2002). Range-based estimation of stochastic volatility models. *The Journal of Finance*, **57,** number 3, 1047–1091.

Andersen, T.G. and Bollerslev, T. (1997). Intraday periodicity and volatility persistence in financial markets. *Journal of Empirical Finance*, **4,** 115–158.

Andersen, T.G., Bollerslev, T., Diebold, F.X. and Ebens, H. (2001). The distribution of realized stock return volatility. *Journal of Financial Economics*, **61,** 43–76.

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

Barndorff-Nielsen, O.E. and Shephard, N. (2002). Econometric analysis of realized volatility and its use in estimating stochastic volatility models. *Journal of the Royal Statistical Society, Series* B, **64,** number 2, 253–280.

Berkes, I., Horvath, L. and Kokoszka, P. (2003). Garch processes: structure and estimation. *Bernoulli*, **9,** number 2, 201–227.

Bollerslev, T. and Wooldridge, J.M. (1992). Quasi-maximum likelihood estimation and inference in dynamic models with time-varying covariances. *Econometric Reviews*, **11,** number 2, 143–172.

de Vilder, R.G. and Visser, M.P. (2007). Volatility proxies for discrete time models. MPRA Working paper no. 4917.

Drost, F.C. and Klaassen, C.A.J. (1997). Efficient estimation in semiparametric Garch models. *Journal of Econometrics*, **81,** 193–221.

Drost, F.C. and Nijman, T.E. (1993). Temporal aggregation of Garch processes. *Econometrica*, **61,** number 4, 909–927.

Engle, R.F. and Gallo, G.M. (2006). A multiple indicators model for volatility using intra-daily data. *Journal of Econometrics*, **131,** number 1-2, 2–27.

Fiorentini, G., Calzolari, G. and Panattoni, L. (1996). Analytic derivatives and the computation of Garch estimates. *Journal of Applied Econometrics*.

Francq, C. and Zakoïan, J-M. (2004). Maximum likelihood estimation of pure GARCH and ARMA-GARCH processes. *Bernoulli*, **10,** number 4, 605–637.

Ghysels, E., Santa-Clara, P. and Valkanov, R. (2006). Predicting volatility: getting the most out of return data sampled at different frequencies. *Journal of Econometrics*, **131,** number 1-2, 59–95.

Glasserman, P. (2003). *Monte Carlo Methods in Financial Engineering*. Applications of Mathematics 53. Springer.

Hillebrand, E. (2005). Neglecting parameter changes in GARCH models. *Journal of Econometrics*, **129,** number 1-2, 121–138.

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

Lumsdaine, R.L. (1995). Finite-sample properties of the maximum likelihood estimator in Garch(1,1) and IGarch(1,1) models: A monte carlo investigation. *Journal of Business & Economic Statistics*, **13,** number 1, 1–10.

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*, **64,** number 3, 575–596.

Mikosch, T. and Starica, C. (2004). Nonstationarities in Financial Time Series, the Long-Range Dependence, and the IGARCH Effects. *The Review of Economics and Statistics*, **86,** number 1, 378–390.

Straumann, D. (2005). *Estimation in Conditionally Heteroscedastic Time Series Models*. Lecture Notes in Statistics 181. Springer.

Straumann, D. and T., Mikosch (2006). Quasi-maximum-likelihood estimation in conditionally heteroscedastic time series: a stochastic recurrence equations approach. *The Annals of Statistics*, **34,** number 5, 2449–2495.

Wooldridge, J.M. (1990). A unified approach to robust, regression-based specification tests. *Econometric Theory*, **6,** 17–43.