Web Appendix to "Realized GARCH: A Joint Model for Returns and Realized Measures of Volatility"

Peter Reinhard Hansen[*] Zhuo Huang[†] Howard Howan Shek[‡]

November 3, 2010

Supplement to Section 4: Computing Standard Errors

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 $-\sum_{t=1}^n \frac{\hat{u}_t\hat{\dot{u}}_t}{\hat{\sigma}_u^4}\hat{\dot{h}}' = \sum_{t=1}^n \frac{1-\hat{z}_t^2}{2\hat{\sigma}_u^2}\hat{\dot{h}}_t$.

[*]Stanford University, Department of Economics & CREATES.
[†]China Center for Economic Research, National School of Development, Peking University.
[‡]Stanford University, iCME.

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.

Supplement to Section 5: Summary Statistics and Additional Empirical Results

Summary statistics for the data are given in Table 1.

Results for the Linear Specification

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 $\mathrm{RG}(2,2)$ model we estimate a model without the leverage function (denoted $\mathrm{RG}(2,2)^{\dagger}$) and an extended model, $\mathrm{RG}(2,2)^{*}$, that also includes a lag of the squared return in the GARCH equation. The results for open-to-close returns are given in the left panel of Table 2, and the corresponding results for close-to-close returns are presented in the right panel of Table 2.

First we discuss the empirical results for open-to-close returns in the left half of Table 2. 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 $\mathrm{RG}(2,2)$ with $\mathrm{RG}(2,2)^{\dagger}$ 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 $\mathrm{RG}(2,2)^{*}$ with the standard model $\mathrm{RG}(2,2)$ we see that the ARCH 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

Symbol	$\overline{r}_{oc}$	$\min r_{oc}$	$\max r_{oc}$	$\overline{r_{oc}^2}$	$\overline{r}_{cc}$	$\min r_{cc}$	$\max r_{cc}$	$\overline{r_{cc}^2}$	$\overline{\mathrm{RK}}$	$\min \mathrm{RK}$	$\max \mathrm{RK}$
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 1: 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.

Model	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)*
	Open-to-Close Returns							Close-to-Close Returns						
	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 (0.14)	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 2: 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 $\mathcal{I}^{-1}\mathcal{J}\mathcal{I}^{-1}$.

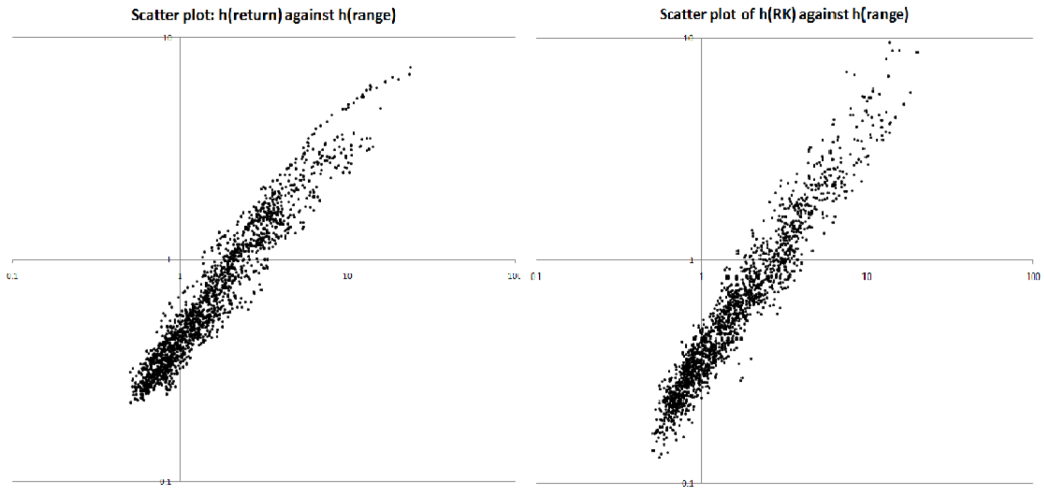

Figure 1: Scatter plots of latent volatility processes for returns, range, and the realized kernel. Each were estimated separately using a GARCH(1,1) structure. The co-linearity between these latent processes suggests that the three processes can be modeled with a single latent process.

the GARCH(1,1) model.

The empirical results for the close-to-close returns in the right half of Table 2 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.

Empirical Results Concerning the 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 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 1 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.

Supplement to Section 6

Skewness and Kurtosis with the Linear Specification

Analytical results for a single period return for the linear Realized GARCH model.

Proposition 1. *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 $z_t \sim \mathrm{iid}N(0,1)$, $u_t \sim \mathrm{iid}(0, \sigma_u^2)$, *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*

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

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 $\sigma_{\tau^2}^2 = \tau_1^2 + 2!\tau_2^2 + 3!\tau_3^2 + 4!\tau_4^2 + \cdots$.

Proof of Proposition 1. 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

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

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

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

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

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. □

Skewness and Kurtosis with Log-Linear Specification

The kurtosis and skewness in Figure 4 was simulated using a RealGARCH(1,2) model with the parameter configuration given in Table 3.

The Approximate Expression for the Kurtosis

Here we provide a justification for the approximation using in Section 6.1. 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.

A Decomposition of 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.

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*,

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

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. (2008), and Barndorff-Nielsen et al. (2010). For instance, under weak assumptions about market microstructure noise Barndorff-Nielsen et al. (2010) show that the "sampling error" is $\log x_t - \log \mathrm{IV}_t = O_p(n^{-1/5})$.

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

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

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.

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,

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

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 (1). The implication is that γ will be indicative of how accurate x_t estimates the integrated variance.

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 the paper to this model. Therefore, we leave a more detailed analysis of this model for future research.

References

Barndorff-Nielsen, O. E., Hansen, P. R., Lunde, A., Shephard, N., 2008. 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., 2010. Multivariate realised kernels: consistent positive semi-definite estimators of the covariation of equity prices with noise and non-synchronous trading. Jounal of Econometrics forthcoming.

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.