GMM Estimation of a Stochastic Volatility Model with Realized Volatility: A Monte Carlo Study

Pierre Chaussé* and Dinghai Xú†

This Version: Oct., 2011

Abstract

This paper investigates alternative generalized method of moments (GMM) estimation procedures of a stochastic volatility model with realized volatility measures. The extended model can accommodate a more general correlation structure. General closed form moment conditions are derived to examine the model properties and to evaluate the performance of various GMM estimation procedures under Monte Carlo environment, including standard GMM, principal component GMM, robust GMM and regularized GMM. An application to five company stocks and one stock index is also provided for an empirical demonstration.

Keywords: Generalized Method of Moments; Principal Component; Robust GMM; Regularized GMM; Heteroscedasticity and Autocorrelation Consistent; Monte Carlo Simulation; Stochastic Volatility Model; Realized Volatility Measure.

Preliminary Draft

---

*Department of Economics, University of Waterloo, Waterloo, Ontario, N2L 3G1, Canada. Email: pchausse@uwaterloo.ca ; Tel: 001-519-888-4567 ext. 32422.

†Department of Economics, University of Waterloo, Waterloo, Ontario, N2L 3G1, Canada. Email: dhxu@uwaterloo.ca; Tel: 001-519-888-4567 ext. 32047.

1 Introduction

Since the seminal works by Engle (1982) and Taylor (1986), there has been considerable interest in modeling the dynamics of the latent financial return volatility. Under the Generalized Autoregressive Conditional Heteroscedasticity (ARCH/GARCH) and Stochastic Volatility (SV) frameworks, the conditional variance is typically specified as certain function of the past information on squared returns and volatilities. Despite that the ARCH/GARCH and SV models have been extensively used in the literature, as Andersen, Bollerslev, Diebold and Labys (2003) and Hansen, Huang and Shek (2010) argued, these traditional models are slow at updating the information especially when the volatility changes rapidly to a new level. This naturally sparks interest in developing and building up alternative volatility models to accommodate these empirical characteristics.

On the other hand, rapid development in computer technology in the past two decades has made the financial transaction data "visible" at the highest granularity. There is an expanding literature on constructing volatility proxy using realized volatility measures based on these high frequency trading data, see Andersen and Bollerslev (1998), Andersen, Bollerslev, Diebold and Labys (2003), Barndorff-Nielsen and Shephard (2004), Zhang, Mykland and Ait-Sahalia (2005), Hansen and Lunde (2005), Barndorff-Nielsen, Hansen, Lunde and Shephard (2008, 2010) and references therein. Although these realized measures are more or less contaminated by microstructure noises and construction biases, they reveal some important information about the current level of the volatility and are consequently useful for explaining the dynamic features of the volatility. Therefore, incorporating the realized proxy element into the traditional volatility models seems to be a natural extension, which helps for modeling and forecasting the volatility movement.

Engle and Gallo (2006) introduce a GARCH process with realized measures, known as Multiplicative Error Model (MEM). Hansen, Huang and Shek (2010) extend the model to a more generalized structure by allowing a more flexible functional form on both latent volatility and realized volatility equations, known as the realized GARCH model. Alternatively, within the SV framework, Takahashi, Omori and Watanabe (2009) develop an extended SV structure by jointly modeling return, latent volatility and the corresponding realized volatility measure. In this paper, we refer to this model as realized SV. The common characteristic of both the realized GARCH and realized SV models is the link equation (or measurement equation), which specifies a potential connection between the latent volatility and the corresponding realized measure (or proxy). The realized SV and realized GARCH models have many attractive features and perform better than the conventional volatility models.

The realized SV and realized GARCH are comparable specifications in the literature. However, the SV estimation has been demonstrated difficult. In the traditional SV framework, as is well known, the likelihood function has no closed form expression. The problem essentially comes from the latent volatility sequence. In other words, the latent conditional volatility at time t has to be integrated out in order to construct the objective likelihood function. Consequently, the standard likelihood function for an SV model involves an integral with a dimension of sample size. This high dimensional integral is, if not impossible, very difficult to solve. This estimation problem is also embedded in the realized SV model structure. Various procedures for estimating the traditional SV parameters have been proposed in the literature, such as Simulated Maximum Likelihood (SML) by Denielsson and Richard (1993), Quasi Maximum Likelihood (QML) by Harvey, Ruiz and Shephard (1994), Markov Chain Monte Carlo (MCMC) by Jacquier, Polson and Rossi (1994), Efficient Method of Moments (EMM) by Gallant and Tauchen (1996), Generalized Method of Moments (GMM) by Andersen and Sorensen (1996), Characteristic Function (CF) by Knight, Satchell and Yu (2002) etc. In particular, Takahashi, Omori and Watanabe (2009) apply a simulation-based MCMC estimation for the realized SV model. As Andersen and Sorensen (1996) argued, the simulation based estimation strategies (such as MCMC and SML) would possibly suffer from the expensive computational cost as the SV setting is getting more and more complicated. For this particular reason, Takahashi, Omori and Watanabe (2009) assume a simple correlation structure in their realized SV model to simplify the MCMC procedure. In this paper, we examine the alternative GMM estimators under the realized SV structure by Monte Carlo methods. We first (slightly) extend the realized SV model proposed by Takahashi, Omori and Watanabe (2009) by accommodating a more flexible correlation structure. Furthermore, we investigate the finite sample properties of the different GMM estimation procedures of the extended realized SV model.

In this paper, we focus on analyzing the properties of different types of GMM estimations. Because returns may have fat tail distributions, the standard GMM procedure could produce bad estimates due to the unboundedness of its influence function. An outlier-robust version of GMM estimator is therefore proposed by Ronchetti and Trojani (2001). Another important issue is the selection of the moment conditions. We can derive a very large number of moment conditions from our model, but it is not obvious which one should be selected. As one can see in Section 4, arbitrary selections may deteriorate the quality of the estimates. This problem has been raised in the context of the instrumental variable estimation by Dominguez and Lobato (2004) but not in the case of moment conditions that are not derived from orthogonality conditions. One approach, which was proposed by Carrasco (2010), is to regularize the weighting matrix of the set of moment conditions. Alternatively, Doran and Schmidt (2006) propose to select the most influential conditions using the principal component approach. We investigate these alternative GMM procedures via Monte Carlo study.

The remainder of the paper is organized as follows. Section 2 presents the model specification with the associated moment conditions. Section 3 discusses the GMM estimation procedures. Section 4 conducts the Monte Carlo experiments and provides an empirical illustration. Section 5 concludes the paper. All the proofs are presented in Appendix A and some results' tables and figures are collected in the Appendix B.

2 Model Specification and Theoretical Moment Conditions

Following Taylor (1986), Ghysels, Harvey and Renault (1996), a standard discrete-time SV model structure is presented as follows,

$$\begin{aligned} x_t &= \exp(h_t/2)\epsilon_t \\ h_{t+1} &= \lambda + \alpha h_t + \eta_t \end{aligned}$$

There are two stochastic processes describing the dynamics of the returns and latent volatilities. In the above set-up, x_t is the continuously compounded return time series, which can be constructed using the logarithmic price differences. Assuming unit variance on the innovation (ϵ_t) of the return process, $\exp(h_t)$ characterizes the conditional variance at time t. The log-volatility, h_t, is normally assumed to follow a stationary AR(1) process. In general, to capture the leverage effect, we allow a certain correlation structure between the innovations from the return and volatility processes. In particular, following Harvey and Shephard (1996) and Yu (2005), the bivariate structure is assumed to be as follows:

$$\begin{pmatrix} \epsilon_t \\ \eta_t \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho\sigma_\eta \\ \rho\sigma_v & \sigma_\eta^2 \end{pmatrix} \right)$$

The asymmetric relationship between the return and future volatility can be captured in the correlation coefficient parameter, ρ. Empirically, this correlation is found to be significantly negative. In the literature, the above normally is referred as the asymmetric stochastic volatility (ASV) model.

Takahashi, Omori and Watanabe (2009) extend the classical ASV model by incorporating realized volatility measures into the above setting. Consequently, they propose a more general model, asymmetric SV with realized volatility (ASV-RV) , which is specified as follows,

$$\begin{aligned} x_t &= \exp(h_t/2)\epsilon_t & (1) \\ y_t &= \beta + h_t + u_t & (2) \\ h_{t+1} &= \lambda + \alpha h_t + \eta_t & (3) \end{aligned}$$

where the residuals follow the tri-variate Gaussian,

$$\begin{pmatrix} \epsilon_t \\ u_t \\ \eta_t \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 & \rho\sigma_\eta \\ 0 & \sigma_u^2 & 0 \\ \rho\sigma_\eta & 0 & \sigma_\eta^2 \end{pmatrix} \right) \tag{4}$$

and y_t denotes the logarithm of realized volatility at time t. Due to the microstructure and non-trading hours noise, the realized volatility, constructed from the intra-daily high frequency trading prices, may be a contaminated measure of the true latent volatility. Therefore, (2) builds up a link (or measurement) function between the constructed realized measure and the true volatility.

In the ASV-RV model, one can see that there is no correlation assumed either between u_t and ϵ_t or u_t and η_t. As Diebold and Strasser (2010) point out, the zero-correlation assumption between the return (or price) and microstructure noise in the literature is perhaps *erroneous*. Interestingly, they detect a negative contemporaneous correlation between the return (or price) and the microstructure noise. Therefore, this paper further extends the ASV-RV model by accommodating the correlation between the residuals from the measurement equation (u_t) and the return process (ϵ_t). To generalize the correlation structure, we also allow for the correlation between u_t and η_t for additional statistical flexibility.[1] Furthermore, in the proposed model, we allow for some scale effects between the realized measure and latent volatility. Consequently, we define the generalized ASV-RV (GASV-RV) as follows,

$$\begin{aligned} x_t &= \exp(h_t/2)\epsilon_t && (5)\\ y_t &= \beta_1 + \beta_2 h_t + u_t && (6)\\ h_{t+1} &= \lambda + \alpha h_t + \eta_t && (7) \end{aligned}$$

where the residuals follow the tri-variate Gaussian distribution,

$$\begin{pmatrix} \epsilon_t \\ u_t \\ \eta_t \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho_1\sigma_u & \rho_2\sigma_\eta \\ \rho_1\sigma_u & \sigma_u^2 & \rho_3\sigma_u\sigma_\eta \\ \rho_2\sigma_\eta & \rho_3\sigma_u\sigma_\eta & \sigma_\eta^2 \end{pmatrix}\right) \tag{8}$$

In the GASV-RV model, the unknown parameter vector to be estimated is defined as $\theta = (\beta_1, \beta_2, \lambda, \alpha, \sigma_u, \sigma_\eta, \rho_1, \rho_2, \rho_3)$. To further examine the model properties, three sets of moment conditions are derived and used as the inputs in the subsequent GMM estimation. These three sets of moment conditions include moments and cross-moments of the return series, moments and cross-moments of the log realized volatility series and cross-moments of both the return and log realized volatility series. We provide closed-form expressions in the following three propositions.

Proposition 1 Given the GASV-RV model specified in Equations (5) to (8), for m, n, $k \geq 0$, the closed form cross-moment expression for x_t and x_{t+k} is given as follows,[2]

$$\begin{aligned} E\left(x_t^n x_{t+k}^m\right) &= \exp\left(\frac{m\lambda}{2}\sum_{j=1}^{k}\alpha^{j-1}\right)\exp\left(\frac{m^2\sigma_\eta^2}{8}\sum_{j=2}^{k}\alpha^{2(k-j)}\right) \\ &\times \frac{\partial M_2^{(n)}}{\partial r^{(n)}}\|_{r=(n+m\alpha^k)/2} \times \frac{\partial M_1^{(m)}}{\partial r_1^{(m)}}\|_{r_1=0,r_2=0,r_3=0} \\ &\times \frac{\partial M_1^{(n)}}{\partial r_1^{(n)}}\|_{r_1=0,r_2=0,r_3=m\alpha^{k-1}/2} \end{aligned} \tag{9}$$

where M_1 and M_2 are defined as two moment generating functions (MGF) specified in the proof.

[1] We have not found any paper investigating this correlation. The addition of this correlation is purely for a more flexible statistical model structure. If there exists no such correlation in practice, the parameter estimate of this correlation should be expected statistically insignificant, and vice versa.

[2] We use the convention that $\sum_{j=a}^{b} f_j = 0$ for $b < a$, where f_j is the functional form indexed by j.

Proof: *see Appendix A*.

Proposition 2 Given the GASV-RV model specified in Equations (5) to (8), for $k \geq 0$, the first two order moments and the cross-moment expressions for y_t and y_{t+k} are given as follows,

$$\begin{aligned} Ey_t &= \beta_1 + \frac{\lambda\beta_2}{1-\alpha} && (10) \\ Ey_t^2 &= \beta_1^2 + 2\beta_1\beta_2\frac{\lambda}{1-\alpha} + \sigma_u^2 + \frac{\beta_2^2\lambda^2}{(1-\alpha)^2} + \frac{\beta_2^2\sigma_\eta^2}{1-\alpha^2} && (11) \end{aligned}$$

and

$$\begin{aligned} E(y_t y_{t+k}) &= \beta_1^2 + \frac{2\beta_1\beta_2\lambda}{1-\alpha} + \frac{\beta_2^2\lambda^2}{(1-\alpha)}\sum_{j=1}^{k}\alpha^{j-1} \\ &+ \beta_2^2\alpha^k\left(\frac{\sigma_\eta^2}{1-\alpha^2} + \frac{\lambda^2}{(1-\alpha)^2}\right) + \beta_2\rho_3\sigma_u\sigma_\eta && (12) \end{aligned}$$

Proof: *see Appendix A*.

Proposition 3 Given the GASV-RV model specified in Equations (5) to (8), for m, n, $k > 0$, the closed form cross-moment expressions for x_t and y_{t+k} are given as follows,

$$\begin{aligned} E(x_t^n y_{t+k}) &= \beta_1 E(x_t^n) + \beta_2\lambda\sum_{j=1}^{k}\alpha^{j-1}E(x_t^n) + \beta_2\alpha^k\frac{\partial M_2}{\partial r}\|_{r=n/2} \times \frac{\partial M_1^{(n)}}{\partial r_1^{(n)}}\|_{r_1=0,r_2=0,r_3=0} \\ &+ \beta_2\alpha^{k-1}M_2\left(r=\frac{n}{2}\right) \times \frac{\partial M_1^{(n+1)}}{\partial r_1^{(n)}\partial r_3}\|_{r_1=0,r_2=0,r_3=0} && (13) \end{aligned}$$

$$\begin{aligned} E(x_t^n y_t^2) &= \beta_1^2 E(x_t^n) + \beta_2^2\frac{\partial M_2^{(2)}}{\partial r^{(2)}}\|_{r=n/2} \times \frac{\partial M_1^{(n)}}{\partial r_1^{(n)}}\|_{r_1=0,r_2=0,r_3=0} && (14) \\ &+ M_2\left(r=\frac{n}{2}\right)\frac{\partial M_1^{(n+2)}}{\partial r_1^{(n)}\partial r_2^{(2)}}\|_{r_1=0,r_2=0,r_3=0} \\ &+ 2\beta_1\beta_2\frac{\partial M_2}{\partial r}\|_{r=n/2}\ \frac{\partial M_1^{(n)}}{\partial r_1^{(n)}}\|_{r_1=0,r_2=0,r_3=0} \\ &+ 2\beta_1 M_2\left(r=\frac{n}{2}\right)\frac{\partial M_1^{(n+1)}}{\partial r_1^{(n)}\partial r_2^{(1)}}\|_{r_1=0,r_2=0,r_3=0} \\ &+ 2\beta_2\frac{\partial M_2}{\partial r}\|_{r=n/2}\ \frac{\partial M_1^{(n+1)}}{\partial r_1^{(n)}\partial r_2^{(1)}}\|_{r_1=0,r_2=0,r_3=0} \end{aligned}$$

and,

$$
\begin{aligned}
E(x_t^n y_{t+1}^2) &= \beta_1^2 E(x_t^n) + \beta_2^2\lambda^2 E(x_t^n) + \beta_2^2\alpha^2 \frac{\partial M_2^{(2)}}{\partial r^{(2)}} \|_{r=n/2} \frac{\partial M_1^{(n)}}{\partial r_1^{(n)}} \|_{r_1=0,r_2=0,r_3=0} \qquad (15)\\
&+ \beta_2^2 M_2\left(r=\frac{n}{2}\right) \frac{\partial M_1^{(n+2)}}{\partial r_1^{(n)}\partial r_3^{(2)}} \|_{r_1=0,r_2=0,r_3=0} + 2\lambda\alpha\beta_2 \frac{\partial M_2}{\partial r} \|_{r=n/2} \frac{\partial M_1^{(n)}}{\partial r_1^{(n)}} \|_{r_1=0,r_2=0,r_3=0}\\
&+ 2\lambda\beta_2^2 M_2\left(r=\frac{n}{2}\right) \frac{\partial M_1^{(n+1)}}{\partial r_1^{(n)}\partial r_3} \|_{r_1=0,r_2=0,r_3=0} + 2\alpha\beta_2^2 \frac{\partial M_2}{\partial r} \|_{r=n/2} \frac{\partial M_1^{(n+1)}}{\partial r_1^{(n)}\partial r_3} \|_{r_1=0,r_2=0,r_3=0}\\
&+ M_2\left(r=\frac{n}{2}\right) \frac{\partial M_1^{(n)}}{\partial r_1^{(n)}} \|_{r_1=0,r_2=0,r_3=0} \frac{\partial M_1^{(2)}}{\partial r_2^{(2)}} \|_{r_1=0,r_2=0,r_3=0} + 2\beta_1\beta_2\lambda E(x_t^n)\\
&+ 2\beta_1\beta_2\alpha \frac{\partial M_2}{\partial r} \|_{r=n/2} \frac{\partial M_1^{(n)}}{\partial r_1^{(n)}} \|_{r_1=0,r_2=0,r_3=0} + 2\beta_1\beta_2 M_2\left(r=\frac{n}{2}\right) \frac{\partial M_1^{(n+1)}}{\partial r_1^{(n)}\partial r_3} \|_{r_1=0,r_2=0,r_3=0}
\end{aligned}
$$

Proof: *see Appendix A.*

Based on the formulas provided in Propositions 1 to 3, some specific moments of interests can be easily backed out from the proposed model. We will investigate these moment conditions further in the subsequent sections.

3 GMM Estimation on GASV-RV Model

Let ψ_t be a $q \times 1$ vector with typical element $(x_t^n x_{t+k}^m)$, $(y_t^n y_{t+k}^m)$ or $(x_t^n y_{t+k}^m)$, for some m, k, and $n \in \{0, 1, 2, 3, ...\}$, and $\psi(\theta_0) = E(\psi_t(\theta_0))$ be the theoretical moments of the GASV-RV model defined by the equations (5) to (8). Let $g_t(\theta) = [\psi_t - \psi(\theta)]$, then the GMM estimator $\hat{\theta}$ of the true vector of coefficients θ_0 is based on the following moment conditions:

$$E[g_t(\theta_0)] = 0, \qquad (16)$$

and is the solution to:

$$\min_{\theta\in\Theta} \bar{g}(\theta)'\hat{\Omega}^{-1}\bar{g}(\theta), \qquad (17)$$

where Θ is the admissible parameter space implied by the model, $\bar{g}(\theta) = [\sum_{t=1}^T \psi_t/T - \psi(\theta)]$ and $\hat{\Omega}$ is a consistent estimate of the asymptotic covariance matrix of $\sqrt{n}\bar{g}(\theta_0)$. In most cases, the estimation of the covariance matrix requires a first step estimate of θ_0 in which the weighting matrix is set to the identity matrix, unless the continuous updated GMM is used. However, because the theoretical moment condition is not a function of the data, there is no need to obtain a first step estimate as long as we compute the covariance matrix from the centered sample moment conditions. Indeed, if $\tilde{\theta}$ is a first step estimate, then $(\psi_t - \psi(\tilde{\theta}) - \bar{g}(\tilde{\theta})) = (\psi_t - \bar{\psi})$, where $\bar{\psi} = \sum_t \psi_t/T$, for any $\tilde{\theta}$. Hall (2000) argues that the power of the GMM over-identifying restrictions test may be improved because the centered covariance matrix estimator is consistent under both the null that the model is correctly specified and under the alternative. Therefore, the estimator defined by equation (17) is a one-step GMM with the estimate of the covariance matrix given by:

$$\hat{\Omega} = \sum_{t=-T+1}^{T-1} \omega(t/s)\hat{\Gamma}_t,$$

where ω is the kernel function. s is the bandwidth and $\hat{\Gamma}_i = \sum_t(\psi_t - \bar{\psi})(\psi_{t-i} - \bar{\psi})'/T$. For the choice of kernel and bandwidth, we use the quadratic spectral kernel and the bandwidth selection procedure proposed by Andrews (1991). This choice produces a more efficient estimate of the covariance matrix but is more computationally intensive since the weights converge to zero at a slower rate than the ones associated with the Parzen or Bartlett kernel. Furthermore, Andersen and Sorensen(1996) show that such automatic bandwidth selection improves the quality of the estimator using a Monte Carlo study in a simple SV model estimation. Because the one-step GMM does not require a first step estimate to compute the optimal weighting matrix, the properties of the estimator should be comparable to the continuous updated GMM. The latter is shown to be less biased than the two-step GMM because the presence of a first step estimate in the objective function adds an extra term in the second order bias (see Newey and Smith (2004)).

It is well-known in the literature on robust statistics that sample moments are not the most efficient methods to estimate the population moments especially when the data comes from fat-tail distributions. Outliers can not only increase the volatility of sample moments, but also bias the statistics based on them. For example, population mean is sometimes more efficiently estimated by the sample median or weighted sum rather than by sample mean. In the context of GMM, Ronchetti and Trojani (2001) show how the fat-tail distributed data may affect the properties of the estimators by looking at its influence function defined as:

$$IF_t(\theta_0) = -\left[G'\Omega^{-1}G\right]^{-1} G'\Omega^{-1}g_t(\theta_0),$$

where $G = E[\partial g_t(\theta_0)/\partial\theta]$. The condition of robustness for any statistical method is the boundness of its influence function. Clearly, if we have observations from heavy tail distributions, the influence function of GMM may not be bounded because it is linearly related to the moment function which is itself unbounded. Several methods have been proposed to increase the robustness of GMM estimators. One approach is to choose a norm that is less sensitive to outliers. A natural choice is the L_1 norm, as suggested by Jong and Han (2002), because the objective function is the sum of absolute values instead of squared values like the usual L_2 norm. However, the objective function based on L_1 is not differentiable everywhere which increases the difficulty of obtaining a solution especially when $\bar{g}(\theta)$ is highly non-linear. Another approach, which is similar to the M-estimates in the context of least square regressions, is to transform $g_t(\theta)$ by a function that is less sensitive to outliers. Ronchetti and Trojani (2001) propose to bound the moment function by applying the following Huber function:

$$\mathcal{H}(a;C) = \left\{ \begin{array}{cl} a & \text{if} \quad |a| \leq C \\ \text{sign}(a)C & \text{if} \quad |a| > C \end{array} \right. ,$$

where C is the robustness parameter and $\text{sign}(a) = 1$ if $a \geq 0$ and -1 if $a < 0$. Lee and Halverson (2004), who analyze the properties of a robust estimator of the variance based on the Huber function for fat tail generalized Gaussian random variables, show that the robustness parameter should increase with the sample size. In other words a larger sample size requires less robustness. Park (2009) applies the same robust GMM (RGMM) to a GARCH-M model and finds that RGMM estimators out-perform the standard GMM in terms of the mean square errors. We will analyze different level of robustness in RGMM in the Monte Carlo section.

Since $\bar{g}(\theta) = [\bar{\psi} - \psi(\theta)]$, we only need to replace $\bar{\psi}$ by a robust estimator of $E(\psi_t)$. Because our moment conditions have different magnitudes, we standardize them before selecting which ones to be truncated. We can then choose the same C for all moment conditions. In order to take into account the correlations between different sample moments, the standardized vector ψ_t^s is defined as:

$$\psi_t^s = \hat{\Omega}^{-1/2}\psi_t.$$

We then define the following weights:

$$w(\psi_{it}^s; C) = \begin{cases} 1 & \text{if} \quad |\psi_{it}^s| \leq C \\ C/|\psi_{it}^s| & \text{if} \quad |\psi_{it}^s| > C \end{cases},$$

for $i = 1, ..., q$. Then, the robust estimator of $E(\psi_{it})$ is defined as:

$$\bar{\psi}_{ri} = \frac{1}{n}\sum_{t=1}^{T} w(\psi_{it}; C)\psi_{it}$$

One can see that the larger is the outlier the smaller is $w(\psi_{it}; C)$, which reduces its impact on the estimated moment. If we were to have unbounded observation, the weight would go to zero. The influence function of GMM based on these robust estimates of the population moments is therefore bounded. Hence, the RGMM estimator is defined as:

$$\hat{\theta}_r = \arg\min_{\Theta} \bar{g}_r(\theta)'\hat{\Omega}_r^{-1}\bar{g}_r(\theta),$$

where $\bar{g}_r(\theta) = [\bar{\psi}_r - \psi(\theta)]$ and $\hat{\Omega}_r$ is an heteroscedasticity and autocorrelation consistent (HAC) estimate of the covariance matrix of $\sqrt{T}(\bar{\psi}_r)$. Inference on $\hat{\theta}$ can be based on the truncated moment conditions. We approximate the distribution of $\hat{\theta}$ by its asymptotic distribution, which is normal, with variance:

$$[\hat{G_r}'\hat{\Omega}_r^{-1}\hat{G}_r]^{-1}/T,$$

where $\hat{G}_r = \partial\bar{g}_r(\hat{\theta}_r)/\partial\theta = \partial\psi(\hat{\theta}_r)/\partial\theta$. In section 4, we examine the properties of RGMM with different robustness parameters.

When we have a large number of possible moment conditions, it is not obvious which one should be selected. If we select a small number of conditions, the objective function may become locally flat and prevent the algorithm from converging within the parameter space, especially if some conditions weakly identify the parameters. If we choose too many conditions, the covariance matrix of the moment conditions may become badly conditioned. One approach, proposed by Carrasco (2010) in the context of instrumental variable estimation, is to select a large set of moment conditions and then regularize the weighting matrix. The method was proposed for GMM based on conditional moment conditions for which the instruments are related to the regressors through an unknown function. In that case, the number of possible moment conditions may be higher than the sample size. It is therefore required to use some regularization scheme. Carrasco (2010) also argues that bad performance of GMM in some applications, such as the estimation of the return to education in which the number of instruments are over 200, may come from the large number of instruments more than their weakness. However, what is considered to be a large number is not clearly established. It is therefore worth exploring in our case even if the number of conditions is (at most) 36. Propositions 1 to 3 show that, in theory, the number could be increased without limit. Using the singular value decomposition of the covariance matrix $\hat{\Omega}$, we can write the GMM objective function as follows:

$$\bar{g}(\theta)'\hat{\Omega}^{-1}\bar{g}(\theta) = \sum_{i=1}^{q}\frac{1}{\hat{\mu}_i} < \bar{g}, \hat{\psi}_i >^2, \tag{18}$$

where $\hat{\mu}_i$ is the i^{th} eigenvalue of $\hat{\Omega}$ in decreasing order and $\hat{\phi}_i$ its associated orthonormalized eigenvector. The objective function can become very unstable if the number of conditions is large and many eigenvalues are close to zero. The regularized GMM (RLGMM) is defined as:

$$\hat{\theta}_r = \arg\min_{\Theta} \sum_{i=1}^{q} \frac{\hat{\mu}_i}{\hat{\mu}_i^2 + \nu} < \bar{g}, \hat{\psi}_i >^2, \tag{19}$$

where ν is the regularization parameter that prevents the objective function from being unstable when the eigenvalues are close to zero. Since $\mu_i/(\mu_i^2 + \nu)$ is negligible when the eigenvalue is close to zero, the method selects automatically the most influential moment conditions. Carrasco (2010) proposes a data driven method for selecting ν. But the approach can be only applied to linear class of models with conditional moment conditions. It remains uncertain how it affects the properties of the estimator when the conditions are highly nonlinear. In the linear case, increasing ν is equivalent to decreasing the number of conditions. Following Newey and Smith (2004), it should make the estimator less efficient but less biased. The effect on the RMSE is uncertain [3]. In the simulation bellow, the value of ν depends on the singular values. In particular, we analyze $\nu = \hat{\mu}_V$ for different V values. The distribution of the J-test cannot be easily derived in that case because ν is a form of nuisance parameter that contaminates the Chi-square distribution. There are ways to deal with that problem such as using the approach proposed by Imhof (1961), but this is beyond the scope of this paper.

Alternatively, we can truncate the summation of the objective function (18):

$$\bar{g}(\theta)'\hat{\Omega}_{pc}^{-1}\bar{g}(\theta) = \sum_{i=1}^{\nu} \frac{1}{\hat{\mu}_i} < \bar{g}, \hat{\psi}_i >^2, \tag{20}$$

where $\nu < q$ and $\hat{\Omega}_{pc}^{-1}$ is the singular value decomposition of $\hat{\Omega}^{-1}$ in which the inverse of the $(q-\nu)$ smallest singular values of $\hat{\Omega}$ in the diagonal matrix are set to zero. Since by definition, $\hat{\psi}$ associated with the largest singular value is obtained by solving $\max_\psi \widehat{Var(\psi'[\sqrt{T}\bar{g}])}$ subject to $\psi'\psi = 1$, this method selects the most influential orthogonal combination of moment conditions. This principal component GMM (PCGMM) approach is proposed by Doran and Schmidt (2006) as a way to improve the finite sample properties of GMM estimator when the number of conditions is large. The J-test of PCGMM is asymptotically distributed as a Chi-squared distribution with $(\nu - \dim(\theta))$ degrees of freedom. Increasing ν is therefore like adding moment conditions accordingly.

4 Monte Carlo Experiments and Empirical Illustration

4.1 Monte Carlo Design

In this section, we carry out several Monte Carlo experiments to investigate the finite sample properties of different GMM estimators presented in the previous section under certain controlled environment. In particular, the following four sets of simulations are conducted: (1) Sensitivity analysis on choice of moment conditions; (2) Analysis of the efficiency with respect to variations of the model parameters and sample sizes; (3) Comparison across/within

[3]See Chaussé (2011) for a Monte Carlo study based on nonlinear moment conditions.

alternative GMM estimators, including standard GMM, RGMM, RLGMM and PCGMM; (4) Examination of non-nested model mis-specifications. Each simulation case is replicated 1000 times using R ("GMM" estimation based on Chaussé (2010)). To test the convergence, we use random number generator to initialize the starting parameter values.[4]

First, in order to make fair comparisons across different simulation cases, we set up a benchmark model with parameter values, $\theta \equiv (\beta_1, \beta_2, \lambda, \alpha, \sigma_u, \sigma_\eta, \rho_1, \rho_2, \rho_3) =$ (0.10, 0.90, -0.01, 0.95, 0.30, 0.20, -0.10, -0.30, 0.00). Most of the parameters' values in the benchmark are chosen to be close to some parameter estimates based on Takahashi, Omori and Watanabe (2009). But given that Takahashi, Omori and Watanabe (2009)'s model does not have ρ_1 and ρ_3, we choose some common empirical values for these two parameters based on the data.[5]

Based on Propositions 1, 2 and 3, we can produce as many closed-form theoretical moment conditions as needed. Following Andersen and Sorensen (1996), the general guide to our initial moment selection is to focus on relatively lower-order moments with small lags (no more than 5). Therefore, our simulation study relies on subsets of 36 moments. More specifically, the 36 moments consist of (a) 4 marginal moments for x_t and y_t, i.e., $E(x_t^i)$ with $(i = 2, 4)$ and $E(y_t^j)$ with $(j = 1, 2)$; (b) 15 auto-correlation moments for x_t and y_t, i.e., $E(x_t^i x_{t+k}^j)$ with ($i = 1, 2$; $j = 1, 2$ and $k = 1, 2, 3, 4, 5$) and $E(y_t y_{t+k})$ with $(k = 1, 2, 3, 4, 5)$; (c) 17 cross moments of x_t and y_t, i.e., $E(x_t^i y_{t+k}^j)$ with ($i = 1, 2$; $j = 1, 2$ and $k = 1, 2, 3, 4, 5$). We believe that these 36 moments are sufficient for the GMM estimation of 9 unknown parameters and are sufficient for practical purposes.[6]

In the first group of experiments, we examine GMM estimation on various combinations of the above set (or certain subsets) of moments. The experimental design is presented in Table 1. More specifically, case 1a uses the full set (36) of moments in the estimation, and in cases 1b to 1d, we pick some subsets of the moment conditions which are commonly used in the literature, such as first four marginal moments, autocorrelations (with different orders and lags) and cross-moments (with different orders and lags) etc. In the last set ($1x^*$), the moment selection is determined by the corresponding principal components (PC), which would effectively drop some "less important" moments automatically according to the rank of the eigenvalues of the weighting matrix, see Doran and Schmidt (2006). Since there is no prior information about the optimal number of non-zero eigenvalues, which is still one of the open questions in the GMM literature, we analyze all possibilities from 23 to 35 (denoted as 23momPC to 35momPC). So in total, there are 13 sub-simulations in $1x^*$.

In the second group of simulations, we investigate the performance of the GMM estimator under different parameter configurations and under different sample sizes. The experiment

[4]We compare unbiased and biased starting values to verify the robustness of the estimation procedures. The results show that the convergence is pretty stable.

[5]There are no typical values for β_1 and β_2 (depending on the quality of the realized volatility measures) in practice. In our case, we intentionally set β_1 and β_2 to be 0.10 and 0.90 to create some bias and scale effects between the true volatility and RV proxy in the simulation. Theoretically, if the realized measure is a good approximation for the true volatility, β_1 is normally close to 0 and β_2 is expected to be close to 1. In addition, we found that ρ_3 is statistically insignificant according to our empirical results. Therefore, in the simulations, we set ρ_3 to be zero. We have also experimented with many alternative sets of the parameters' values, including all the scenarios in the empirical section and some other cases. Those results are available upon request.

[6]We have also experimented with some other larger sets of the moment conditions, such as extending the lags up to 10 and increasing the power to higher orders. We found that the results are very similar as those presented in this paper.

Table 1: Monte Carlo Design #1.

	Moments Selection	# of Moments
(1a)	Full Set	36
(1b)	$E(x_t^i)$ (i=2,4), $E(y_t^i)$ (i=1,2), $E(x_t^2 x_{t+1}^2)$, $E(y_t y_{t+1})$, $E(x_t^n y_{t+k}^m)$(n=1,2; m=1,2; k=0,1)	13
(1c)	13 moments in (1b), $E(x_t^2 x_{t+2}^2)$, $E(y_t y_{t+2})$, $E(x_t^n y_{t+k}^m)$(n=1,2; m=1,2; k=2)	19
(1d)	Full set without the moments with lag of $k = 5$	32
($1x$*)	Auto-Selection via Principle Component	{23, 24, ..., 35}

*Note: * denotes that the moments selection process has been done using Principal Component (PC) approach proposed by Doran and Schmidt (2006).*

characteristics are displayed in Table 2. Case 2a is set to be the benchmark case in this group for comparisons. Case 2b decreases the sample sizes to 500, while case 2c increases the sample sizes to 3000. Case 2d sets $\beta_2 = 1$, ρ_1 and ρ_3 equal to zero and other parameters remain the same as the benchmark. With this particular parameter setting, our model is reduced to the model presented in Takahashi, Omori and Watanabe (2009). Lastly, case 2e only increases the variance (σ_u^2) of the measurement equation to see the model performance when the realized volatility measure is a noisy estimator of the latent volatility.[7]

Table 2: Monte Carlo Design #2.

	β_1	β_2	λ	α	σ_u	σ_η	ρ_1	ρ_2	ρ_3	n
(2a)	0.10	0.90	-0.01	0.95	0.30	0.20	-0.10	-0.30	0.00	1500
(2b)	0.10	0.90	-0.01	0.95	0.30	0.20	-0.10	-0.30	0.00	**500**
(2c)	0.10	0.90	-0.01	0.95	0.30	0.20	-0.10	-0.30	0.00	**3000**
(2d)	0.10	**1.00**	-0.01	0.95	0.30	0.20	**0.00**	-0.30	0.00	1500
(2e)	0.10	0.90	-0.01	0.95	**1.20**	0.20	-0.10	-0.30	0.00	1500

Note: The bold numbers highlight the differences for each case comparing to the benchmark case, (2a).

In the third group of simulations, we investigate alternative GMM estimation procedures discussed in section 3. We compare the performance of these GMM procedures on the benchmark case. In particular, for standard GMM and PCGMM, we use the results from simulation groups #1 and #2. For RGMM, we analyze the estimator with different levels of the robustness parameter C. For RLGMM, we examine the performance with respect to different regularization coefficients, v.

In the last group of experiments, we examine the potential mis-specification effects on both the model and GMM procedures. In particular, two sub-experiments (nested and non-nested mis-specifications) are conducted in this group. First, we generate the data

[7]We have also done simulations by changing other parameter values. To save space, we do not report those results in this paper. However, the results are available upon request.

following the benchmark model (in case 2a) and estimate the model with the restrictions $\rho_1 = 0$ and $\rho_3 = 0$, which are typically assumed in the literature. Second, we misspecify the data generating process (DGP) by using another popular model, the realized GARCH, recently proposed by Hansen, Huang and Shek (2010). More specifically, we simulate the data from a log-linear realized GARCH (1,2) process with the parameter values taken from S&P 500 empirical estimates (see Table 3 in Hansen, Huang and Shek (2010)). The DGP is as follows,

$$\begin{aligned}
x_t &= \sqrt{h_t}\epsilon_t \qquad \epsilon_t \sim N(0,1) \\
\log(y_t) &= -0.18 + 1.04\log(h_t) - 0.07\epsilon_t + 0.07(\epsilon_t^2 - 1) + u_t \qquad u_t \sim N(0, 0.38^2) \\
\log(h_t) &= 0.04 + 0.70\log(h_{t-1}) + 0.45\log(y_{t-1}) - 0.18\log(y_{t-2})
\end{aligned}$$

4.2 Monte Carlo Results

In the first group of experiments, various sets of moment conditions (see moments selection specification in Table 1) are used for estimating the model parameters. Since there are as many as 19 sets of simulations in this group, to save space, we only report some representative results in the paper. These results are presented in Table 3. According to the results, there is no "best" (or "optimal") moments' combination which uniformly dominates other cases in terms of the bias and RMSE measures. However, selecting moments using the PC technique in general produces slightly smaller bias and RMSE comparing to the corresponding case without the PC-selection process (for example, 25mom versus 25momPC, 32mom versus 32momPC and etc). This is generally consistent with the findings established in Doran and Schmidt (2006). We also find that 23momPC to 35momPC performs similarly, in other words, there is no significant difference regarding the bias and RMSE. Theoretically, given a large sample size, the empirical standard deviations from the simulation of GMM estimators should be close to the standard deviations based on the asymptotic distribution. But, we observe some differences between these two measures. For example, in the 13mom case, almost for all the parameters, we observe significant gaps between the sample and asymptotic components. The Kolmogorov-Smirnov (K-S) test is conducted to investigate the asymptotic normality property of the estimator. The normality is rejected for almost all the parameter estimates from 13mom. This suggests that the estimation based on 13mom is not reliable and the distribution of estimates is not well approximated by the asymptotic distribution. The inference based on the results may provide some misleading information. A possible reason is that the number of moments is not enough for the estimation, which creates an identification problem. In this group of experiments, we also conduct the J-test and record the non-convergence rate for each case. The results are presented in Table 4. As expected, as the number of moments increases, the empirical size of J-test deviates from the nominal level. However, 23momPC to 33momPC produce reasonable J-test statistics. Lastly, as the number of moments increases, the non-convergence rate drops. In general, the non-convergence rate is low except for the 13mom case where the failure rate is as high as 16%, which is certainly not recommended for practical implementations.

Sensitivity experimental (group # 2.) results are summarized in Table 5.[8] Consistent

[8]To save space, we only report the simulation results based on the full set of 36 moments. The results

with our expectation, when the sample size decreases/increases (see case 2b/2c), both the bias and RMSE increase/ decrease uniformly. The difference between the sample standard error and the asymptotic standard error becomes smaller in general as the sample size increases. This result indicates that the distribution of the estimates converges to the corresponding asymptotic (limiting) distribution as the sample size increases. Comparing between case 2d (Takahashi, Omori and Watanabe (2009)'s model) and case 2a (benchmark), we find very similar simulation results in terms of bias and RMSE. This result implies that even in the true DGP where there are no correlations assumed on ρ_1 and ρ_3, the proposed methodology can still capture the characteristics and all parameter estimates exhibit good finite sample properties. Lastly, in case 2e, we increase the variance on the measurement equation (i.e, the constructed realized volatility estimator is contaminated with some large unexpected measurement errors). We find that the quality of the estimates in general become a little worse (RMSE increases for all parameters). The K-S test statistics (K-S Stats) and p-value (K-S p-value) are presented in the last two columns of Tables 5. Overall, increasing the sample size improves the quality of the K-S measure. For example, when the sample size is 500 (case 2b), the Normality is rejected for four parameter estimates, while when sample size increases to 3000 (case 2c), the Normality cannot be rejected for all parameter estimates. Figure 1 presents the distributions of the estimates over the 1000 replications via the QQ-plots for case 2a and 2c. As shown, most of the estimates fit well with the 45-degree quantile line against the normal distribution. This reinforces the K-S test results reported in Table 5. Overall, the estimator produces good finite sample properties.

As mentioned, the third group simulations compare across/within alternative GMM estimators on the GASV-RV model. The comparisons are based on the estimation of the benchmark case (2a). In particular, three GMM estimators are investigated via simulations. We provide the results of PCGMM with different moment truncations, RGMM with different robust parameters and RLGMM with different regularization coefficients. These results are reported in Tables 6-8, respectively. We have done experiments of PCGMM for various moment truncations. To save space, we only report the results for the truncations of {23, 26, 29, 32, 33, 35}. In general, the RMSE decreases as the number of moment conditions increases. But the RMSE stabilizes around 32-33 truncation level. As for RGMM, we have also done experiments for a wide range of C values, but to save space, we present the RMSE measures from $C = 1$ to $C = 6.5$ in Table 7. Similar as PCGMM, as C increases, the RMSE becomes more and more stabilized around $C = 5$ for all parameter estimates. As for the RLGMM, we have run the simulations from $v = 0.001$ to $v = 0.1$. To save space, we report the RMSE results only for $V \in \{0.001, 0.005, 0.01, 0.04, 0.07, 0.1\}$ in Table 8. As one can see, the RLGMM performs very similarly in terms of RMSE when v is in the interval of [0.001, 0.1]. But the difference is not very significant. RLGMM does not smooth the inverse of the weighting matrix as drastically as PCGMM. The difference among the ν's is therefore not significantly big. Comparing across all these GMM estimations, we find that the PCGMM performs generally better than the alternatives. All three GMM estimation procedures will be applied in the empirical section.

for other combinations of the moment conditions are available upon request.

In the last group of the simulations, we examine the performance under certain misspecified environment. In particular, we first simulate the data following the benchmark case and estimate the model with the restrictions $\rho_1 = \rho_3 = 0$ using 33momPC and 36mom, respectively. The results are presented in Table 9. In general, 33momPC and 36mom produce very similar estimates for all parameters. The most interesting finding is that under this misspecified case, the bias and RMSE of ρ_2 become significantly larger comparing to the benchmark case (see model 2a in Table 5). In particular, the bias of ρ_2 increases from around 0.02 to 0.10 and the RMSE of ρ_2 increases from around 0.086 to 0.120. This result implies that in practice ignoring ρ_1 in the model will produce positive bias on the leverage coefficient. This result also consistently supports our earlier argument that the assumption of ρ_1 being zero may not be a reasonable one in practice. In the second sub-experiment in this group, we simulate the data following a completely different DGP (realized LGARCH). Four different GMM procedures are applied to estimate the simulated data, namely 33momPC, RLGMM(v=0.01), RGMM(C=5.5) and 36mom. Since these two models do not share any common parameters, the standard measures (such as bias and RMSE) can not be constructed for evaluation. Table 9 reports the mean, median and sample standard deviation over 1000 replications for each GMM estimator respectively. From the Table 9, one can see that these four estimators produce similar estimates for all parameters. It is worth mentioning that the true persistence in the realized LGARCH is around 0.99, but, the estimated persistence is only around 0.93 under the GASV-RV. This is consistent with the result in Carnero, Pena and Ruiz (2004) that the GARCH-like models tend to generate larger persistence than SV-type models. To further evaluate the performance of the proposed model under this mis-specified environment, we present the moment comparisons in Table 11.[9] Following the true DGP, we approximate the true moments by using large sample simulations. Specifically, we take the mean values of each moment condition over 10,000 replications with sample size of 50,000. These moment values are used for the benchmark. Then, we plug the mean parameter estimates from 33momPC, RLGMM(v=0.01), RGMM(C=5.5) and 36mom into the theoretical moment expressions to get the corresponding moment values. As we can see, GASV-RV model can still produce moments quite close to the true ones. This indicates that even under this mis-specified process, the GASV-RV model could capture many characteristics from the data in terms of moment measures.

4.3 An Empirical Illustration

In this section, we provide an empirical illustration of the proposed methodology using five company stocks and one index data, including AIG, CVX, JPM, PG, T and S&P 500. The data set consists of intra-daily high frequency (tick-by-tick) transaction prices over the period roughly from 2003 to 2008. This data set has also been examined by Hansen, Huang and Shek (2010). Therefore, for comparisons, we also estimate the competing model (realized GARCH) proposed by Hansen, Huang and Shek (2010). Two inputs for the model are used for the estimation. First, daily return is calculated as the logarithmic price differences.[10] The second input of the model is the realized volatility

[9] As a note, only a subset of representative moments are reported in Table 11.

[10] In this paper, we adopt the open-to-close return definition to capture the market open activity, see Hansen, Huang and Shek (2010). In addition, we have also used the close-to-close return in the empirical

measure. In this paper, following Hansen, Huang and Shek (2010), we use the realized kernel proposed by Barndorff-Nielsen, Hansen, Lunde and Shephard (2008, 2010). As demonstrated in the aforementioned papers, the realized kernel estimator is robust to the market microstructure noise and provides a better proxy for the latent volatility. Figure 2 presents the plots for the return and realized kernel estimates for S&P 500 index.

We apply PCGMM, RGMM and RLGMM estimation procedures to the GASV-RV model. The empirical estimation results are reported in Table 12.[11] In general, the empirical estimates are consistent with some well-established findings in the literature. The latent volatility process is highly persistent, i.e., α values are generally close to one, and are all statistically significant. As expected, the persistence under the SV structure is slightly lower than the GARCH-like specification, see the results from Monte Carlo results in section 4.2 (simulation group # 4) and Carnero, Pena and Ruiz (2004). The estimates on β_1 and β_2 in the link equation are similar to those reported in Hansen, Huang and Shek (2010), which reenforces the fact that the realized measure of the volatility based on the kernel estimator is a fairly good proxy for the conditional variance. Especially, β_2 is close to one, which suggests the realized kernel is roughly proportional (scale) to the latent volatility. To capture the leverage effect, the realized GARCH model specifies a leverage function (a Hermite polynomial in Hansen, Huang and Shek (2010)), while the GASV-RV model assumes a general tri-variate correlation structure. The leverage coefficient ρ_2 is found to be negative and significant, which is consistent with SV literature. Moreover, we find that estimates on ρ_1 are negative and most of them are significant. This empirically supports the argument in Diebold and Strasser (2010) that there may exist a negative contemporaneous correlation between the return and the realized volatility noise.[12] The J-test statistics in the last column show that the proposed model cannot be rejected at 1% significance level for most of the empirical data (except SPY). Furthermore, for comparison, we also estimate the GASV-RV model by restricting ρ_1 to be zero. The empirical results are provided in Table 13. We find that all the parameter estimates behave similar expect the leverage coefficient (ρ_2). Specifically, ρ_2 becomes uniformly smaller (in absolute terms) and sometimes this correlation becomes positive (although not significant), which is not consistent with the literature findings. This phenomenon can be explained from the results based on our nested mis-specification simulation (see Experiment #.4). In other words, if ρ_1 is dropped out from the model, the leverage effect is reduced. In addition, from the J-statistics, we observe that the model without ρ_1 is rejected for most of the cases.

Next, we want to make empirical comparisons on the performance of the realized GARCH and GASV-RV models based on the sample data. It is difficult to conduct a direct comparison between the two specifications although both models assume great similarity of the structure. Therefore, we construct empirical moment evaluation for four model specifications including GASV-RV, Restricted GASV-RV (RGASV-RV), LGARCH(1,2)-

estimation. These results are available upon request.

[11] We generally found that ρ_3 is insignificant for these empirical data. Therefore, in Table 12, we report the estimates from different GMM procedures based on the model without ρ_3.

[12] As a note, we should be careful with these results when the J-test is rejected. As shown by Hall and Inoue (2003), the asymptotic distribution of the GMM estimators under misspecified models can be very different from the one under correctly specified models.

RV and LGARCH(1,1)-RV in Table 14. More specifically, we use the data moments as the benchmark. The corresponding model moments from GASV-RV,[13] LGARCH(1,1) and LGARCH (1,2) are calculated based on the empirical estimates from each model.[14] In general, GASV-RV behaves very similarly as the other two realized LGARCH models in terms of moments, which indicates that both specifications are competitive alternatives for practical applications. One interesting observation from Table 14 is that both GASV-RV and LGARCH-RV underestimate the empirical kurtosis of the returns. For instance, the empirical return kurtosis of JPM is 28.50, but the implied kurtosis from either GASV-RV or LGARCH-RV is much smaller. One possible reason is that under both model specifications, the innovation in the return process is assumed to be Gaussian, which is not enough to accommodate the empirical heavy-tail characteristics. This suggests that an alternative thick-tail distribution (such as t distribution or mixture of normal distribution) would be more appropriate to capture the extra kurtosis.[15] We also find that in general, the proposed GASV-RV performs slightly better than alternative models. Especially, the GASV-RV can capture the correlation between the return and future realized volatility better.

5 Conclusion

This paper provides a good extension of Monte Carlo study in Andersen and Sorensen (1996) by further examining the GMM estimation of an extended SV model with realized volatility measures. General closed form moment conditions are achieved and used in alternative GMM procedures. Given a (relatively large) set of moments, different moment selection schemes with respect to the weighting matrix are investigated. The Monte Carlo results show that selecting moments automatically via PCGMM and RGMM procedures improves the efficiency of the GMM estimator (in terms of RMSE) than the arbitrary moment selections. In the case of badly conditioned weighting matrix of the moments (weak identification), RLGMM provides an efficient way to solve the estimation problem. It is as expected that the PCGMM, RGMM and RLGMM procedures improve the quality of the GMM estimator than the standard GMM. Empirical applications to five stocks and one stock index are also provided for illustration. PCGMM, RGMM and RLGMM produce similar empirical parameter estimates. Besides the common findings in the literature (such as significant leverage effect) are detected, we do find some negative correlations between the measurement equation (realized volatility) and return process. Empirical results also show that the GASV-RV and realized GARCH models are comparable specifications and behave similarly in terms of moments.

Lastly, we want to summarize several issues which remain of interest in this paper. The estimated return kurtosis is found to be not enough to explain the empirical heavy-

[13]Noticing that the empirical estimates across PCGMM, RGMM and RLGMM are similar, we only construct the empirical moments based on PCGMM for demonstration.

[14]The realized LGARCH model is estimated by using the Quasi maximum likelihood (QML) method proposed in Hansen, Huang and Shek (2010).

[15]The theoretical moments would be very different and complicated if one distribution is not Gaussian in the tri-variate structure. One possible solution would be to use Copula-based method to accommodate general dependence with specified marginals. We will leave this for future research.

tail characteristics from the data. How do we accommodate a heavy tail distribution into the model specification without complicating the estimation procedure? Can we find a more efficient way to choose moments and number of moment conditions? What are the asymptotic comparisons between the realized GARCH and realized GASV specifications? How do we achieve robust and reliable inference given the model is misspecified? We will leave these for future research.

Appendix

A Proofs

Proof of Proposition 1.
Based on the tri-variate Gaussian specification in (8), we first define the following two MGFs:

$$\begin{aligned} M_1 &= E\ \exp(r_1\epsilon_t + r_2u_t + r_3\eta_t) \\ &= \int_{\epsilon_t}\int_{u_t}\int_{\eta_t} \exp(r_1\epsilon_t + r_2u_t + r_3\eta_t) f(\epsilon_t, u_t, \eta_t) d\eta_t du_t d\epsilon_t \\ &= \exp\left(\frac{1}{2}r_1^2 + \frac{1}{2}r_2^2\sigma_u^2 + \frac{1}{2}r_3^2\sigma_\eta^2 + \rho_1 r_1 r_2 \sigma_u + \rho_2 r_1 r_3 \sigma_\eta + \rho_3 r_2 r_3 \sigma_u \sigma_\eta\right) \end{aligned}$$

where $f(\epsilon_t, u_t, \eta_t)$ is the tri-variate Gaussian density from (8). Similarly we have,

$$\begin{aligned} M_2 &= E\ \exp(rh_t) \\ &= \exp\left(\frac{\lambda r}{1-\alpha} + \frac{\sigma_\eta^2 r^2}{2(1-\alpha^2)}\right) \end{aligned}$$

Then,

$$E\left(x_t^n x_{t+k}^m\right) = E\left[\exp\left(\frac{nh_t}{2}\right)\epsilon_t^n \exp\left(\frac{mh_{t+k}}{2}\right)\epsilon_{t+k}^m\right]$$

Given that h_t follows an AR(1) process specified in (7), by recursive substitutions, we can easily achieve h_{t+k} as follows,

$$h_{t+k} = \lambda\sum_{j=1}^{k}\alpha^{k-1} + \alpha^k h_t + \sum_{j=1}^{k}\alpha^{k-j}\eta_{t+j-1}$$

By substituting h_{t+k} into the above expectation, we have,

$$\begin{aligned} E\left(x_t^n x_{t+k}^m\right) &= E\left[\exp\left(\frac{n}{2}h_t\right)\exp\left(\frac{m\lambda}{2}\sum_{j=1}^{k}\alpha^{k-1} + \frac{m\alpha^k}{2}h_t + \frac{m}{2}\sum_{j=1}^{k}\alpha^{k-j}\eta_{t+j-1}\right)\epsilon_t^n\epsilon_{t+k}^m\right] \\ &= \exp\left(\frac{m\lambda}{2}\sum_{j=1}^{k}\alpha^{k-1}\right) \times E\left[\exp\left(\frac{n+m\alpha^k}{2}h_t\right)\right] \times E\left[\exp\left(\frac{m}{2}\alpha^{k-1}\eta_t\right)\epsilon_t^n\right] \\ &\times\ E\left(\epsilon_{t+k}^m\right) \times E\left(\frac{m}{2}\sum_{j=2}^{k}\alpha^{k-j}\eta_{t+j-1}\right) \end{aligned}$$

To work out the expectations in closed forms, we need to use the properties of the joint MGFs defined above. Define $\frac{\partial M^{(n)}}{\partial r^{(n)}}||_{\underline{r}=\underline{a}}$ as taking the nth partial derivative of the moment generating function M with respect to the corresponding variable r and evaluating the whole resulting function at ($\underline{r} = \underline{a}$). Then, we have,

$$E\left[\exp\left(\frac{n + m\alpha^k}{2}h_t\right)\right] = \frac{\partial M_2^{(n)}}{\partial r^{(n)}} \|_{r=(n+m\alpha^k)/2}$$

$$E\left[\exp\left(\frac{m}{2}\alpha^{k-1}\eta_t\right)\epsilon_t^n\right] = \frac{\partial M_1^{(n)}}{\partial r_1^{(n)}} \|_{r_1=0,r_2=0,r_3=m\alpha^{k-1}/2}$$

$$E\left(\epsilon_{t+k}^m\right) = \frac{\partial M_1^{(m)}}{\partial r_1^{(m)}} \|_{r_1=0,r_2=0,r_3=0}$$

$$E\left(\frac{m}{2}\sum_{j=2}^{k}\alpha^{k-j}\eta_{t+j-1}\right) = \exp\left(\frac{m^2\sigma_\eta^2}{8}\sum_{j=2}^{k}\alpha^{2(k-j)}\right)$$

Combining all the above closed-form expressions, we complete the proof of the Proposition 1.

Proof of Proposition 2.
Given y_t and h_t specifications in (6) and (7),

$$E(y_t) = \beta_1 + \beta_2 E(h_t) = \beta_1 + \beta_2\lambda/(1-\alpha)$$

$$\begin{aligned} E(y_t^2) &= \beta_1^2 + \beta_2^2 E(h_t^2) + E(u_t^2) + 2\beta_1\beta_2\lambda/(1-\alpha) + 2\beta_2 E(h_t u_t) \\ &= \beta_1^2 + \frac{\beta_2^2\lambda^2}{(1-\alpha)^2} + \frac{\beta_2^2\sigma_\eta^2}{1-\alpha^2} + \sigma_u^2 + 2\beta_1\beta_2\frac{\lambda}{1-\alpha} \end{aligned}$$

$$E(y_t y_{t+k}) = E(\beta_1 + \beta_2 h_t + u_t)(\beta_1 + \beta_2 h_{t+k} + u_{t+k})$$

From (7), we have $h_{t+k} = \lambda\sum_{j=1}^{k}\alpha^{k-1} + \alpha^k h_t + \sum_{j=1}^{k}\alpha^{k-j}\eta_{t+j-1}$. We substitute the h_{t+k} into the $E(y_t y_{t+k})$ expression and get,

$$\begin{aligned} E(y_t y_{t+k}) &= \beta_1^2 + 2\beta_1\beta_2\lambda/(1-\alpha) + \beta_2^2 E\left[h_t\left(\lambda\sum_{j=1}^{k}\alpha^{k-1} + \alpha^k h_t + \sum_{j=1}^{k}\alpha^{k-j}\eta_{t+j-1}\right)\right] \\ &= \beta_1^2 + \frac{2\beta_1\beta_2\lambda}{1-\alpha} + \frac{\beta_2^2\lambda^2}{(1-\alpha)}\sum_{j=1}^{k}\alpha^{j-1} + \beta_2^2\alpha^k\left(\frac{\sigma_\eta^2}{1-\alpha^2} + \frac{\lambda^2}{(1-\alpha)^2}\right) + \beta_2\rho_3\sigma_u\sigma_\eta \end{aligned}$$

Proof of Proposition 3.
Based on the tri-variate Gaussian specification in (8), we define two MGFs, M_1 and M_2,

given in the Proof of Proposition 1.

$$
\begin{aligned}
E(x_t^n y_{t+k}) &= E\left[\exp(nh_t/2)\epsilon_t^n(\beta_1 + \beta_2 h_{t+k} + u_{t+k})\right] \\
&= \beta_1 E[\exp(nh_t/2)\epsilon_t^n] + \beta_2 E\left[\exp(nh_t/2)\epsilon_t^n\left(\lambda\sum_{j=1}^{k}\alpha^{k-1} + \alpha^k h_t + \sum_{j=1}^{k}\alpha^{k-j}\eta_{t+j-1}\right)\right] \\
&= \beta_1 E(x_t^n) + \beta_2\lambda\sum_{j=1}^{k}\alpha^{j-1}E(x_t^n) + \beta_2\alpha^k E\left[\exp\left(\frac{n}{2}h_t\right)h_t\right]E(\epsilon_t^n) \\
&+ \beta_2\alpha^{k-1}E\left[\exp\left(\frac{n}{2}h_t\right)\right]E(\epsilon_t^n\eta_t)
\end{aligned}
$$

Based on the property of the MGF, we can work out each expectation in the above expression as follows,

$$
E\left[\exp\left(\frac{n}{2}h_t\right)h_t\right] = \frac{\partial M_2}{\partial r}\,\|_{r=n/2}
$$

$$
E\left(\epsilon_t^n\right) = \frac{\partial M_1^{(n)}}{\partial r_1^{(n)}}\,\|_{r_1=0,r_2=0,r_3=0}
$$

$$
E\left[\exp\left(\frac{n}{2}h_t\right)\right] = M_2\left(r = \frac{n}{2}\right)
$$

$$
E(\epsilon_t^n\eta_t) = \frac{\partial M_1^{(n+1)}}{\partial r_1^{(n)}\partial r_3}\,\|_{r_1=0,r_2=0,r_3=0}
$$

Combining all the above closed-form expressions, we complete the proof of $E(x_t^n y_{t+k})$ in the Proposition 3.

Similarly, we have,

$$
\begin{aligned}
E(x_t^n y_t^2) &= E\left[\exp(nh_t/2)\epsilon_t^n(\beta_1 + \beta_2 h_t + u_t)^2\right] \\
&= E\left[\exp(nh_t/2)\epsilon_t^n(\beta_1^2 + \beta_2^2 h_t^2 + u_t^2 + 2\beta_1\beta_2 h_t + 2\beta_1 u_t + 2\beta_2 h_t u_t)\right] \\
&= \beta_1^2 E(x_t^n) + \beta_2^2\frac{\partial M_2^{(2)}}{\partial r^{(2)}}\,\|_{r=n/2} \times \frac{\partial M_1^{(n)}}{\partial r_1^{(n)}}\,\|_{r_1=0,r_2=0,r_3=0} \\
&+ M_2\left(r = \frac{n}{2}\right)\frac{\partial M_1^{(n+2)}}{\partial r_1^{(n)}\partial r_2^{(2)}}\,\|_{r_1=0,r_2=0,r_3=0} + 2\beta_1\beta_2\frac{\partial M_2}{\partial r}\,\|_{r=n/2}\;\frac{\partial M_1^{(n)}}{\partial r_1^{(n)}}\,\|_{r_1=0,r_2=0,r_3=0} \\
&+ 2\beta_1 M_2\left(r = \frac{n}{2}\right)\frac{\partial M_1^{(n+1)}}{\partial r_1^{(n)}\partial r_2^{(1)}}\,\|_{r_1=0,r_2=0,r_3=0} + 2\beta_2\frac{\partial M_2}{\partial r}\,\|_{r=n/2}\;\frac{\partial M_1^{(n+1)}}{\partial r_1^{(n)}\partial r_2^{(1)}}\,\|_{r_1=0,r_2=0,r_3=0}
\end{aligned}
$$

and,

$$
\begin{aligned}
E(x_t^n y_{t+1}^2) &= E\left[\exp(nh_t/2)\epsilon_t^n(\beta_1+\beta_2 h_{t+1}+u_{t+1})^2\right] \\
&= E\left[\exp(nh_t/2)\epsilon_t^n(\beta_1^2+\beta_2^2 h_{t+1}^2+u_{t+1}^2+2\beta_1\beta_2 h_{t+1}+2\beta_1 u_{t+1}+2\beta_2 h_{t+1}u_{t+1})\right] \\
&= \beta_1^2 E(x_t^n)+\beta_2^2\lambda^2 E(x_t^n)+\beta_2^2\alpha^2\frac{\partial M_2^{(2)}}{\partial r^{(2)}}\|_{r=n/2}\ \frac{\partial M_1^{(n)}}{\partial r_1^{(n)}}\|_{r_1=0,r_2=0,r_3=0} \\
&+ \beta_2^2 M_2\left(r=\frac{n}{2}\right)\frac{\partial M_1^{(n+2)}}{\partial r_1^{(n)}\partial r_3^{(2)}}\|_{r_1=0,r_2=0,r_3=0}+2\lambda\alpha\beta_2\frac{\partial M_2}{\partial r}\|_{r=n/2}\ \frac{\partial M_1^{(n)}}{\partial r_1^{(n)}}\|_{r_1=0,r_2=0,r_3=0} \\
&+ 2\lambda\beta_2^2 M_2\left(r=\frac{n}{2}\right)\frac{\partial M_1^{(n+1)}}{\partial r_1^{(n)}\partial r_3}\|_{r_1=0,r_2=0,r_3=0}+2\alpha\beta_2^2\frac{\partial M_2}{\partial r}\|_{r=n/2}\ \frac{\partial M_1^{(n+1)}}{\partial r_1^{(n)}\partial r_3}\|_{r_1=0,r_2=0,r_3=0} \\
&+ M_2\left(r=\frac{n}{2}\right)\frac{\partial M_1^{(n)}}{\partial r_1^{(n)}}\|_{r_1=0,r_2=0,r_3=0}\ \frac{\partial M_1^{(2)}}{\partial r_2^{(2)}}\|_{r_1=0,r_2=0,r_3=0}+2\beta_1\beta_2\lambda E(x_t^n) \\
&+ 2\beta_1\beta_2\alpha\frac{\partial M_2}{\partial r}\|_{r=n/2}\ \frac{\partial M_1^{(n)}}{\partial r_1^{(n)}}\|_{r_1=0,r_2=0,r_3=0}+2\beta_1\beta_2 M_2\left(r=\frac{n}{2}\right)\frac{\partial M_1^{(n+1)}}{\partial r_1^{(n)}\partial r_3}\|_{r_1=0,r_2=0,r_3=0}
\end{aligned}
$$

As a note, the $E(x_t^n)$ expression can be directly taken from the results in Proposition 1.

B Results

B.1 Monte Carlo Simulations

Table 3: Different Moment Combinations (Selective) (Non-convergence rate for 13 to 36 moments are: 0.16, 0.002, 0.002, 0.002, 0.077, 0.0089, 0.097, 0.006,), (J-test size for 13 to 36 moments are: 0.17, 0.13, 0.098, 0.08, 0.082, 0.19, 0.11, 0.17,)

		Bias	RMSE	sample S-E	K-S stats	P-value (K-S)
β_1	13mom	0.0227	0.0481	0.0424	0.0225	0.6920
	17mom	0.0285	0.0513	0.0427	0.0279	0.4181
	21mom	0.0323	0.0538	0.0431	0.0274	0.4417
	25mom	0.0359	0.0566	0.0438	0.0319	0.2597
	25momPC	0.0324	0.0544	0.0428	0.0339	0.2006
	32mom	0.0630	0.0820	0.0525	0.0366	0.1362
	32momPC	0.0571	0.0765	0.0508	0.0380	0.1106
	36mom	0.0625	0.0821	0.0533	0.0510	0.0111
β_2	13mom	0.0636	0.1100	0.0898	0.0562	0.0036
	17mom	0.0509	0.0953	0.0806	0.0490	0.0165
	21mom	0.0562	0.0983	0.0806	0.0477	0.0210
	25mom	0.0615	0.1016	0.0809	0.0528	0.0076
	25momPC	0.0527	0.0953	0.0719	0.0494	0.0151
	32mom	0.1078	0.1425	0.0932	0.0616	0.0010
	32momPC	0.1176	0.1593	0.1076	0.0478	0.0206
	36mom	0.1083	0.1436	0.0943	0.0558	0.0040
λ	13mom	-0.0090	0.0170	0.0144	0.0926	0.0000
	17mom	-0.0040	0.0085	0.0075	0.0437	0.0436
	21mom	-0.0045	0.0085	0.0072	0.0516	0.0098
	25mom	-0.0050	0.0086	0.0071	0.0477	0.0213
	25momPC	-0.0052	0.0075	0.0065	0.0400	0.0816
	32mom	-0.0079	0.0193	0.0176	0.2083	0.0000
	32momPC	-0.0069	0.0099	0.0072	0.0471	0.0237
	36mom	-0.0080	0.0181	0.0162	0.1964	0.0000
α	13mom	-0.0345	0.0599	0.0490	0.0604	0.0014
	17mom	-0.0085	0.0191	0.0172	0.0341	0.1958
	21mom	-0.0085	0.0156	0.0131	0.0468	0.0251
	25mom	-0.0085	0.0147	0.0120	0.0358	0.1547
	25momPC	-0.0046	0.0127	0.0093	0.0357	0.1556
	32mom	-0.0123	0.0628	0.0616	0.3216	0.0000
	32momPC	-0.0127	0.0207	0.0163	0.0269	0.4622
	36mom	-0.0121	0.0595	0.0582	0.3245	0.0000

MC Experiment #1. Results (Different Moment Combinations) [Table 3 cont'ed]

		Bias	RMSE	sample S-E	K-S stats	P-value (K-S)
σ_u	13mom	-0.0157	0.0326	0.0286	0.0729	0.0000
	17mom	0.0004	0.0122	0.0122	0.0195	0.8429
	21mom	0.0002	0.0101	0.0101	0.0346	0.1828
	25mom	-0.0003	0.0098	0.0098	0.0236	0.6344
	25momPC	-0.0002	0.0060	0.0061	0.0664	0.0003
	32mom	0.0011	0.0101	0.0101	0.0169	0.9389
	32momPC	-0.0145	0.0489	0.0467	0.0692	0.0001
	36mom	0.0004	0.0097	0.0097	0.0177	0.9137
σ_η	13mom	0.0132	0.0642	0.0629	0.0286	0.3847
	17mom	-0.0169	0.0331	0.0284	0.0322	0.2525
	21mom	-0.0179	0.0278	0.0213	0.0192	0.8561
	25mom	-0.0190	0.0273	0.0196	0.0234	0.6441
	25momPC	-0.0110	0.0237	0.0107	0.0261	0.5048
	32mom	-0.0339	0.0395	0.0203	0.0178	0.9109
	32momPC	-0.0301	0.0371	0.0217	0.0233	0.6506
	36mom	-0.0340	0.0393	0.0197	0.0241	0.6091
ρ_1	13mom	0.0069	0.0590	0.0586	0.0232	0.6535
	17mom	0.0129	0.0546	0.0531	0.0189	0.8662
	21mom	0.0137	0.0551	0.0534	0.0187	0.8736
	25mom	0.0135	0.0556	0.0540	0.0192	0.8535
	25momPC	0.0023	0.0474	0.0474	0.0450	0.0348
	32mom	0.0194	0.0545	0.0510	0.0210	0.7691
	32momPC	0.0123	0.0520	0.0508	0.0301	0.3251
	36mom	0.0188	0.0548	0.0515	0.0211	0.7665
ρ_2	13mom	0.0234	0.1534	0.1517	0.1341	0.0000
	17mom	0.0091	0.0989	0.0985	0.0449	0.0353
	21mom	0.0158	0.0875	0.0861	0.0241	0.6092
	25mom	0.0166	0.0858	0.0842	0.0178	0.9080
	25momPC	0.0163	0.1059	0.1027	0.0574	0.0027
	32mom	0.0190	0.0854	0.0833	0.0180	0.9023
	32momPC	0.0229	0.0811	0.0882	0.0347	0.1788
	36mom	0.0193	0.0859	0.0837	0.0181	0.8988

Table 4: J-test and Non-Convergence Rate

	Size = 0.01	Size = 0.05	Size = 0.1	Non-Convergence(%)
13mom	0.0670	0.1660	0.2610	0.1597
17mom	0.0380	0.1270	0.2010	0.0020
21mom	0.0280	0.0980	0.1690	0.0020
25mom	0.0230	0.0800	0.1490	0.0020
25momPC	0.0240	0.0820	0.1300	0.0775
32mom	0.0800	0.1910	0.2760	0.0089
32momPC	0.0410	0.1100	0.1840	0.0967
36mom	0.0730	0.1720	0.2620	0.0060

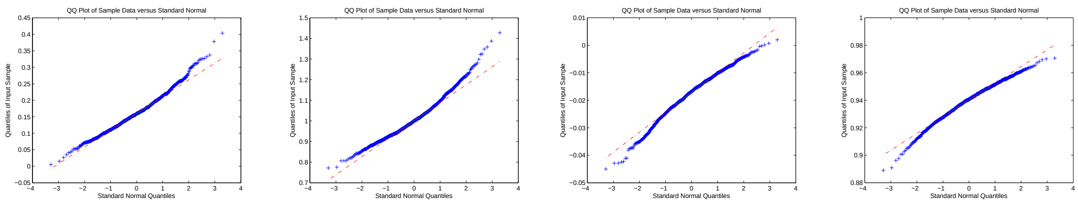
Table 5: Monte Carlo Experiment #2. Results

		Bias	RMSE	sample S-E	K-S stats	P-value (K-S)
β_1	model2a	0.0626	0.0821	0.0531	0.0514	0.0101
	model2b	0.1643	0.2078	0.1274	0.0677	0.0003
	model2c	0.0370	0.0495	0.0328	0.0336	0.2094
	model2d	0.0689	0.0901	0.0580	0.0474	0.0226
	model2e	0.0651	0.0990	0.0746	0.0424	0.0549
β_2	model2a	0.1090	0.1427	0.0922	0.0574	0.0028
	model2b	0.2259	0.3436	0.2591	0.0877	0.0000
	model2c	0.0694	0.0886	0.0551	0.0293	0.3568
	model2d	0.1182	0.1549	0.1002	0.0567	0.0032
	model2e	0.1933	0.2559	0.1678	0.0476	0.0215
λ	model2a	-0.0075	0.0106	0.0075	0.0457	0.0309
	model2b	-0.0198	0.0302	0.0228	0.0947	0.0000
	model2c	-0.0043	0.0064	0.0048	0.0267	0.4727
	model2d	-0.0082	0.0112	0.0077	0.0459	0.0296
	model2e	-0.0035	0.0124	0.0118	0.1230	0.0000
α	model2a	-0.0103	0.0160	0.0122	0.0466	0.0263
	model2b	-0.0245	0.0431	0.0355	0.0955	0.0000
	model2c	-0.0061	0.0101	0.0081	0.0199	0.8228
	model2d	-0.0125	0.0175	0.0123	0.0554	0.0043
	model2e	0.0000	0.0334	0.0334	0.0807	0.0000
σ_u	model2a	0.0004	0.0097	0.0097	0.0183	0.8927
	model2b	-0.0074	0.0199	0.0185	0.0273	0.4708
	model2c	0.0023	0.0074	0.0070	0.0167	0.9442
	model2d	0.0002	0.0102	0.0102	0.0255	0.5329
	model2e	-0.0233	0.0360	0.0275	0.0207	0.7862
σ_η	model2a	-0.0339	0.0392	0.0196	0.0259	0.5147
	model2b	-0.0493	0.0631	0.0395	0.0357	0.1730
	model2c	-0.0255	0.0289	0.0138	0.0293	0.3551
	model2d	-0.0310	0.0362	0.0186	0.0264	0.4889
	model2e	-0.0565	0.0757	0.0504	0.0264	0.4892
ρ_1	model2a	0.0187	0.0548	0.0515	0.0205	0.7947
	model2b	0.0201	0.0966	0.0945	0.0217	0.7541
	model2c	0.0179	0.0412	0.0371	0.0192	0.8556
	model2d	0.0139	0.0573	0.0556	0.0222	0.7089
	model2e	0.0070	0.0304	0.0297	0.0156	0.9675
ρ_2	model2a	0.0193	0.0859	0.0837	0.0177	0.9118
	model2b	0.0265	0.1685	0.1665	0.0426	0.0613
	model2c	0.0210	0.0620	0.0584	0.0393	0.0912
	model2d	0.0181	0.0841	0.0821	0.0208	0.7777
	model2e	-0.0462	0.1981	0.1927	0.0728	0.0001

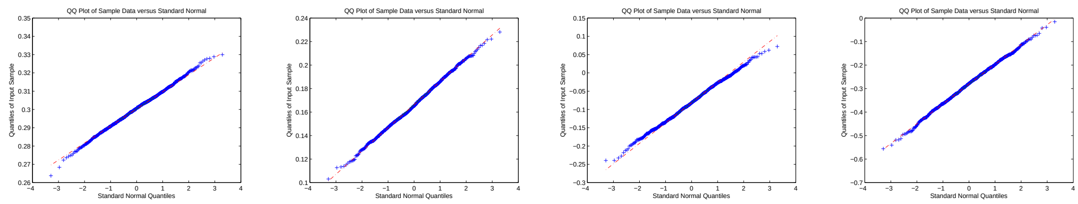
Figure 1: QQ-Plots of the Estimates

Case 2a

β_1 β_2 λ α

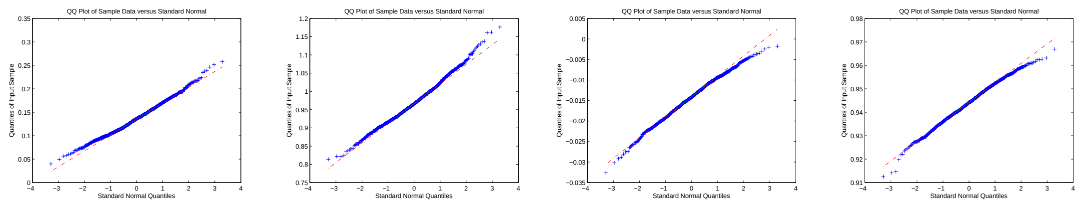


σ_u σ_η ρ_1 ρ_2



Case 2c

β_1 β_2 λ α



σ_u σ_η ρ_1 ρ_2

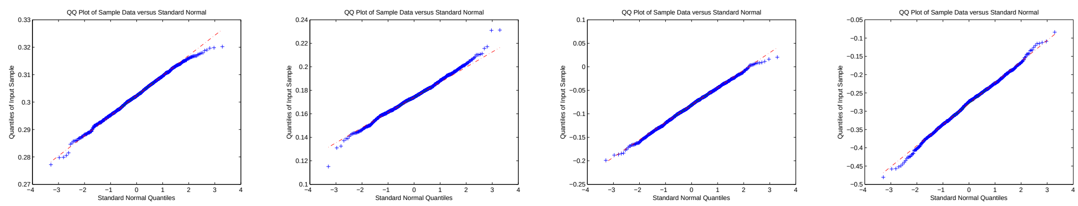


Table 6: RMSE for the PC-GMM

	Trunc=23	Trunc=26	Trunc=29	Trunc=32	Trunc=33	Trunc=35
β_1	0.0756	0.0746	0.0746	0.0765	0.0750	0.0748
β_2	0.2180	0.1891	0.1685	0.1593	0.1331	0.1311
λ	0.0100	0.0102	0.0101	0.0099	0.0094	0.0092
α	0.0281	0.0253	0.0214	0.0207	0.0160	0.0153
σ_u	0.0665	0.0589	0.0524	0.0489	0.0134	0.0100
σ_η	0.0479	0.0410	0.0373	0.0371	0.0375	0.0378
ρ_1	0.0881	0.0756	0.0672	0.0520	0.0553	0.0552
ρ_2	0.1125	0.0993	0.0930	0.0811	0.0858	0.0859

Table 7: RMSE for the Robust-GMM

	C = 2	C = 3	C = 3.5	C = 4	C = 5.5	C = 6
β_1	0.1883	0.1236	0.1127	0.1054	0.0934	0.0908
β_2	0.2572	0.2029	0.1929	0.1831	0.1616	0.1566
λ	0.0782	0.0113	0.0107	0.0105	0.0101	0.0100
α	0.0200	0.0160	0.0161	0.0167	0.0165	0.0166
σ_u	0.0429	0.0128	0.0116	0.0109	0.0101	0.0100
σ_η	0.0671	0.0479	0.0458	0.0442	0.0412	0.0405
ρ_1	0.0604	0.0544	0.0543	0.0542	0.0541	0.0539
ρ_2	0.0925	0.0858	0.0853	0.0853	0.0845	0.0842

Table 8: RMSE for the Regularized-GMM

	ν =0.001	ν =0.005	ν =0.01	ν =0.04	ν =0.07	ν =0.1
β_1	0.0760	0.0741	0.0734	0.0723	0.0718	0.0716
β_2	0.1389	0.1515	0.1615	0.1907	0.2082	0.2223
λ	0.0095	0.0092	0.0091	0.0088	0.0086	0.0084
α	0.0158	0.0170	0.0180	0.0207	0.0220	0.0228
σ_u	0.0113	0.0202	0.0272	0.0457	0.0569	0.0674
σ_η	0.0377	0.0375	0.0379	0.0409	0.0431	0.0448
ρ_1	0.0547	0.0558	0.0577	0.0655	0.0712	0.0784
ρ_2	0.0865	0.0872	0.0884	0.0933	0.0965	0.0990

Table 9: Monte Carlo Experiment #4. Results (Nested mis-specification)

		Bias	RMSE	sample S-E	K-S stats	P-value (K-S)
β_1	33momPC	0.05660	0.07515	0.04946	0.04288	0.05114
	36mom	0.05709	0.07548	0.04939	0.03838	0.10601
β_2	33momPC	0.08743	0.12506	0.08947	0.05096	0.01127
	36mom	0.08862	0.12494	0.08811	0.05198	0.00914
λ	33momPC	-0.00596	0.00910	0.00688	0.04791	0.02058
	36mom	-0.00604	0.00918	0.00691	0.04431	0.03991
α	33momPC	-0.00905	0.01587	0.01305	0.05452	0.00533
	36mom	-0.00917	0.01578	0.01286	0.04988	0.01401
σ_u	33momPC	0.00302	0.01365	0.01332	0.02371	0.62925
	36mom	0.00188	0.00996	0.00978	0.02151	0.74566
σ_η	33momPC	-0.03480	0.04019	0.02011	0.01856	0.88214
	36mom	-0.03445	0.03984	0.02003	0.01314	0.99533
ρ_2	33momPC	0.09927	0.11904	0.06573	0.02048	0.79694
	36mom	0.10071	0.11963	0.06461	0.02409	0.60928

Table 10: Monte Carlo Experiment #4. Results (Non-nested mis-specification)

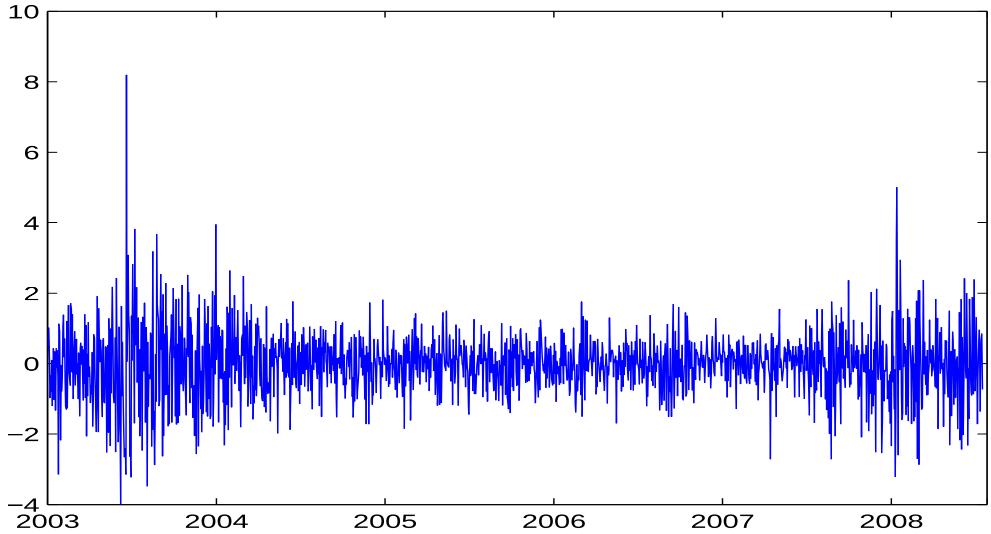
		Mean	Median	sample S-E
β_1	Full Set of Moments (36)	-0.01730	-0.01612	0.08289
	PCGMM (33)	-0.01094	-0.01336	0.08105
	RLGMM ($\nu = 0.01$)	-0.04295	0.01740	0.60583
	RGMM (C=5.5)	0.00528	0.00581	0.07730
β_2	Full Set of Moments (36)	1.10467	1.09871	0.08602
	PCGMM (33)	1.10109	1.09556	0.08380
	RLGMM ($\nu = 0.01$)	1.14533	1.18676	0.42474
	RGMM (C=5.5)	1.12299	1.12039	0.08678
λ	Full Set of Moments (36)	-0.05565	-0.05004	0.03426
	PCGMM (33)	-0.04759	-0.04310	0.02883
	RLGMM ($\nu = 0.01$)	-0.04516	-0.04277	0.02165
	RGMM (C=5.5)	-0.05885	-0.05306	0.03117
α	Full Set of Moments (36)	0.93075	0.93669	0.02791
	PCGMM (33)	0.94056	0.94567	0.02599
	RLGMM ($\nu = 0.01$)	0.93692	0.93998	0.02131
	RGMM (C=5.5)	0.92890	0.93522	0.02840
σ_u	Full Set of Moments (36)	0.26152	0.26213	0.01366
	PCGMM (33)	0.28700	0.28957	0.02254
	RLGMM ($\nu = 0.01$)	0.22090	0.22510	0.04912
	RGMM (C=5.5)	0.25749	0.25871	0.01481
σ_η	Full Set of Moments (36)	0.20595	0.20371	0.02886
	PCGMM (33)	0.19140	0.18801	0.02956
	RLGMM ($\nu = 0.01$)	0.18900	0.19066	0.02860
	RGMM (C=5.5)	0.20243	0.19910	0.02889
ρ_1	Full Set of Moments (36)	-0.24703	-0.24663	0.09035
	PCGMM (33)	-0.22666	-0.22759	0.08608
	RLGMM ($\nu = 0.01$)	-0.30677	-0.28985	0.13602
	RGMM (C=5.5)	-0.25277	-0.24982	0.09074
ρ_2	Full Set of Moments (36)	-0.15144	-0.15085	0.11247
	PCGMM (33)	-0.16571	-0.16362	0.12421
	RLGMM ($\nu = 0.01$)	-0.14187	-0.14446	0.15199
	RGMM (C=5.5)	-0.15505	-0.15286	0.11204

Table 11: Monte Carlo Experiment #4. Results (Selective Moments Comparison)

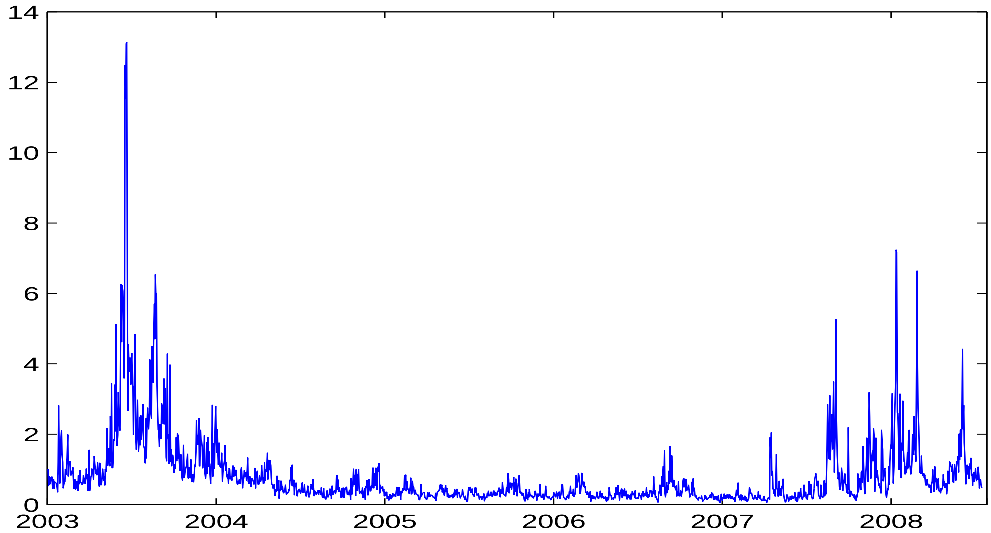
	Full	PCGMM	RLGMM	RGMM	Benchmark
Var(x_t)	0.5247	0.5260	0.5610	0.5068	0.7721
Kurt(x_t)	4.1197	4.1211	4.0189	4.2516	4.5033
Mean(y_t)	-0.9050	-0.8926	-0.8629	-0.9273	-0.6505
Var(y_t)	0.4654	0.4675	0.4323	0.4323	0.5815
Kurt(y_t)	2.9837	2.9765	2.9521	2.9056	2.8742
E($y_t y_{t+1}$)	1.1793	1.1589	1.1039	1.2043	0.9005
E($y_t y_{t+3}$)	1.1312	1.1171	1.0600	1.1563	0.8580
E($x_t y_t$)	-0.0450	-0.0454	-0.0491	-0.0447	-0.0600
E($x_t y_{t+1}$)	-0.0240	-0.0243	-0.0223	-0.0225	-0.0286
E($x_t y_{t+3}$)	-0.0208	-0.0215	-0.0195	-0.0209	-0.0204
E($x_t y_{t+5}$)	-0.0180	-0.0191	-0.0172	-0.0180	-0.0189
E($x_t^2 y_t^2$)	0.4047	0.4056	0.4075	0.4050	0.4897
E($x_t y_{t+1}^2$)	0.0356	0.0355	0.0314	0.0372	0.0213
E($x_t^2 y_{t+1}^2$)	0.4130	0.4120	0.4128	0.4125	0.4912

B.2 Empirical Results

Figure 2: SPY Daily Return and Realized Kernel



Daily Return



Realized Kernel

Table 12: Empirical Estimates

		β_1	β_2	λ	α	σ_u	σ_η	ρ_1	ρ_2	J-test
AIG	Full	**0.30**	**0.74**	**-0.01**	**0.96**	**0.34**	**0.27**	-0.09	**-0.32**	***46.48***
		(0.03)	(0.04)	(0.00)	(0.01)	(0.01)	(0.03)	(0.06)	(0.11)	(0.02)
	PC	**0.31**	**0.74**	**-4.5e-03**	**0.97**	**0.36**	**0.26**	-0.07	**-0.32**	***36.59***
		(0.03)	(0.04)	(0.00)	(0.01)	(0.01)	(0.03)	(0.06)	(0.12)	(0.06)
	RG	**0.29**	**0.75**	-4.1e-03	**0.96**	**0.33**	**0.28**	-0.07	**-0.31**	***33.84***
		(0.03)	(0.04)	(0.00)	(0.01)	(0.03)	(0.03)	(0.07)	(0.11)	(0.21)
	R	**0.30**	**0.74**	**-0.01**	**0.97**	**0.33**	**0.25**	**-0.11**	**-0.36**	56.76
		(0.02)	(0.04)	(0.00)	(0.01)	(0.01)	(0.03)	(0.06)	(0.12)	(0.00)
CVX	All	**0.14**	**1.15**	1.8e-03	**0.93**	**0.29**	**0.16**	**-0.26**	**-0.26**	51.38
		(0.03)	(0.09)	(0.00)	(0.01)	(0.01)	(0.02)	(0.05)	(0.07)	(0.00)
	PC33	**0.15**	**1.13**	1.1e-03	**0.94**	**0.32**	**0.15**	**-0.22**	**-0.26**	***44.70***
		(0.03)	(0.09)	(0.00)	(0.01)	(0.01)	(0.02)	(0.04)	(0.08)	(0.01)
	Reg(0.01)	**0.12**	**1.22**	3.5e-03	**0.92**	**0.25**	**0.17**	**-0.26**	**-0.21**	***36.82***
		(0.04)	(0.13)	(0.00)	(0.02)	(0.05)	(0.02)	(0.09)	(0.08)	(0.12)
	Rob(5.5)	**0.14**	**1.19**	2.3e-04	**0.93**	**0.29**	**0.15**	**-0.26**	**-0.26**	58.38
		(0.03)	(0.10)	(0.00)	(0.01)	(0.01)	(0.02)	(0.04)	(0.07)	(0.00)
JPM	All	**0.19**	**0.82**	-0.01	**0.95**	**0.31**	**0.27**	**-0.18**	**-0.27**	***43.60***
		(0.03)	(0.06)	(0.00)	(0.01)	(0.01)	(0.03)	(0.06)	(0.09)	(0.03)
	PC33	**0.20**	**0.80**	-4.9e-03	**0.96**	**0.35**	**0.25**	**-0.14**	**-0.29**	***30.18***
		(0.03)	(0.06)	(0.00)	(0.01)	(0.02)	(0.03)	(0.05)	(0.10)	(0.22)
	Reg(0.01)	**0.19**	**0.82**	-4.1e-03	**0.95**	**0.30**	**0.27**	**-0.16**	**-0.31**	***27.20***
		(0.03)	(0.06)	(0.00)	(0.01)	(0.04)	(0.03)	(0.07)	(0.10)	(0.51)
	Rob(5.5)	**0.19**	**0.83**	-4.9e-03	**0.95**	**0.31**	**0.27**	**-0.18**	**-0.26**	***44.51***
		(0.03)	(0.06)	(0.00)	(0.01)	(0.01)	(0.03)	(0.06)	(0.09)	(0.02)
PG	All	**0.30**	**0.94**	**-0.06**	**0.89**	**0.29**	**0.25**	**-0.22**	**-0.37**	60.57
		(0.04)	(0.07)	(0.01)	(0.02)	(0.01)	(0.02)	(0.05)	(0.06)	(0.00)
	PC33	**0.30**	**0.92**	**-0.05**	**0.90**	**0.33**	**0.23**	**-0.20**	**-0.41**	***45.38***
		(0.04)	(0.07)	(0.01)	(0.01)	(0.01)	(0.02)	(0.04)	(0.07)	(0.01)
	Reg(0.01)	**0.29**	**0.95**	**-0.06**	**0.90**	**0.32**	**0.23**	**-0.15**	**-0.38**	***34.36***
		(0.05)	(0.09)	(0.01)	(0.02)	(0.03)	(0.03)	(0.05)	(0.07)	(0.19)
	Rob(5.5)	**0.30**	**0.91**	**-0.06**	**0.89**	**0.29**	**0.24**	**-0.18**	**-0.33**	66.21
		(0.04)	(0.07)	(0.01)	(0.02)	(0.01)	(0.02)	(0.05)	(0.06)	(0.00)
SPY	All	-0.09	**0.88**	**-0.12**	**0.90**	**0.25**	**0.30**	**-0.42**	**-0.42**	124.72
		(0.07)	(0.06)	(0.02)	(0.01)	(0.01)	(0.02)	(0.06)	(0.06)	(0.00)
	PC33	-0.06	**0.90**	**-0.08**	**0.92**	**0.29**	**0.26**	**-0.34**	**-0.46**	107.64
		(0.06)	(0.06)	(0.02)	(0.01)	(0.01)	(0.03)	(0.05)	(0.07)	(0.00)
	Reg(0.01)	0.01	**1.03**	**-0.04**	**0.95**	**0.29**	**0.20**	**-0.26**	**-0.40**	58.41
		(0.04)	(0.05)	(0.01)	(0.01)	(0.04)	(0.03)	(0.06)	(0.09)	(0.00)
	Rob(5.5)	-0.03	**0.94**	**-0.09**	**0.91**	**0.25**	**0.26**	**-0.38**	**-0.41**	116.17
		(0.07)	(0.07)	(0.02)	(0.01)	(0.01)	(0.02)	(0.06)	(0.06)	(0.00)
T	All	**0.24**	**0.81**	**0.01**	**0.93**	**0.34**	**0.28**	**-0.15**	**-0.35**	***38.21***
		(0.03)	(0.05)	(0.00)	(0.01)	(0.01)	(0.03)	(0.06)	(0.09)	(0.09)
	PC33	**0.25**	**0.79**	**0.01**	**0.94**	**0.37**	**0.26**	**-0.13**	**-0.37**	***33.89***
		(0.03)	(0.05)	(0.00)	(0.01)	(0.02)	(0.03)	(0.06)	(0.10)	(0.11)
	Reg(0.01)	**0.24**	**0.80**	**0.01**	**0.93**	**0.33**	**0.28**	**-0.16**	**-0.35**	***31.30***
		(0.03)	(0.05)	(0.00)	(0.01)	(0.03)	(0.03)	(0.07)	(0.10)	(0.30)
	Rob(5.5)	**0.25**	**0.80**	**0.01**	**0.93**	**0.34**	**0.28**	**-0.15**	**-0.35**	***36.33***
		(0.03)	(0.06)	(0.00)	(0.01)	(0.01)	(0.03)	(0.06)	(0.10)	(0.13)

Note: Bold numbers represent the significance above 10% level (for example, including 1% and 5%). The Bold and Italic numbers on the last column represent that the J-statistics cannot reject the model at 1% significance level.

Table 13: Empirical Estimates ($\rho_1 = 0$)

GMM		β_1	β_2	λ	α	σ_u	σ_η	ρ_2	J-test
AIG	All	**0.293**	**0.734**	**-0.006**	**0.962**	**0.336**	**0.278**	**-0.191**	***48.842***
		(0.025)	(0.035)	(0.003)	(0.008)	(0.010)	(0.030)	(0.058)	(0.012)
	PC33	**0.304**	**0.733**	**-0.005**	**0.967**	**0.365**	**0.261**	**-0.210**	***38.223***
		(0.025)	(0.035)	(0.003)	(0.008)	(0.013)	(0.032)	(0.064)	(0.058)
	Reg(0.01)	**0.291**	**0.748**	-0.004	**0.961**	**0.330**	**0.280**	**-0.239**	***34.872***
		(0.027)	(0.037)	(0.003)	(0.009)	(0.028)	(0.034)	(0.081)	(0.209)
	Rob(5.5)	**0.296**	**0.725**	**-0.008**	**0.964**	**0.332**	**0.256**	**-0.179**	**59.892**
		(0.024)	(0.040)	(0.003)	(0.009)	(0.010)	(0.033)	(0.063)	(0.001)
CVX	All	**0.148**	**1.090**	0.003	**0.928**	**0.292**	**0.167**	0.027	85.830
		(0.033)	(0.087)	(0.003)	(0.014)	(0.009)	(0.018)	(0.054)	(0.000)
	PC33	**0.163**	**1.066**	0.002	**0.947**	**0.334**	**0.148**	0.014	73.119
		(0.033)	(0.083)	(0.002)	(0.013)	(0.014)	(0.020)	(0.062)	(0.000)
	Reg(0.01)	**0.123**	**1.213**	0.003	**0.923**	**0.258**	**0.159**	-0.048	51.241
		(0.040)	(0.134)	(0.003)	(0.021)	(0.046)	(0.023)	(0.063)	(0.007)
	Rob(5.5)	**0.156**	**1.145**	0.001	**0.933**	**0.294**	**0.148**	0.042	95.711
		(0.034)	(0.102)	(0.002)	(0.015)	(0.009)	(0.019)	(0.058)	(0.000)
JPM	All	**0.180**	**0.813**	-0.005	**0.950**	**0.308**	**0.263**	-0.057	53.402
		(0.028)	(0.061)	(0.004)	(0.012)	(0.009)	(0.031)	(0.050)	(0.004)
	PC33	**0.193**	**0.792**	-0.005	**0.961**	**0.353**	**0.236**	-0.090	***37.630***
		(0.028)	(0.061)	(0.004)	(0.012)	(0.015)	(0.033)	(0.060)	(0.066)
	Reg(0.01)	**0.189**	**0.815**	-0.004	**0.953**	**0.308**	**0.259**	**-0.181**	***32.638***
		(0.029)	(0.061)	(0.004)	(0.013)	(0.035)	(0.034)	(0.081)	(0.293)
	Rob(5.5)	**0.180**	**0.820**	-0.005	**0.948**	**0.306**	**0.267**	-0.046	53.224
		(0.027)	(0.065)	(0.004)	(0.013)	(0.010)	(0.031)	(0.049)	(0.004)
PG	All	**0.276**	**0.880**	**-0.055**	**0.891**	**0.293**	**0.255**	**-0.167**	83.158
		(0.033)	(0.059)	(0.009)	(0.015)	(0.012)	(0.022)	(0.044)	(0.000)
	PC33	**0.280**	**0.869**	**-0.046**	**0.911**	**0.331**	**0.235**	**-0.195**	68.309
		(0.033)	(0.056)	(0.009)	(0.015)	(0.014)	(0.023)	(0.049)	(0.000)
	Reg(0.01)	**0.276**	**0.927**	**-0.050**	**0.901**	**0.321**	**0.236**	**-0.291**	***43.212***
		(0.038)	(0.076)	(0.011)	(0.019)	(0.027)	(0.028)	(0.062)	(0.044)
	Rob(5.5)	**0.282**	**0.867**	**-0.057**	**0.892**	**0.293**	**0.252**	**-0.159**	82.064
		(0.034)	(0.061)	(0.010)	(0.015)	(0.010)	(0.023)	(0.043)	(0.000)
SPY	All	-0.065	**0.914**	**-0.058**	**0.934**	**0.263**	**0.234**	-0.056	169.611
		(0.046)	(0.048)	(0.011)	(0.010)	(0.010)	(0.023)	(0.044)	(0.000)
	PC33	-0.050	**0.923**	**-0.047**	**0.946**	**0.300**	**0.209**	-0.078	148.879
		(0.046)	(0.049)	(0.010)	(0.010)	(0.013)	(0.024)	(0.049)	(0.000)
	Reg(0.01)	0.004	**1.056**	**-0.026**	**0.961**	**0.289**	**0.181**	**-0.135**	78.963
		(0.036)	(0.052)	(0.007)	(0.009)	(0.035)	(0.024)	(0.062)	(0.000)
	Rob(5.5)	-0.014	**0.960**	**-0.057**	**0.937**	**0.264**	**0.223**	-0.061	158.296
		(0.056)	(0.056)	(0.011)	(0.010)	(0.011)	(0.023)	(0.045)	(0.000)
T	All	**0.242**	**0.805**	**0.008**	**0.931**	**0.339**	**0.279**	**-0.161**	***45.602***
		(0.029)	(0.054)	(0.004)	(0.012)	(0.011)	(0.026)	(0.052)	(0.026)
	PC33	**0.254**	**0.789**	**0.007**	**0.943**	**0.372**	**0.257**	**-0.181**	***40.361***
		(0.029)	(0.053)	(0.004)	(0.012)	(0.017)	(0.028)	(0.059)	(0.036)
	Reg(0.01)	**0.241**	**0.801**	**0.009**	**0.935**	**0.337**	**0.274**	**-0.208**	***36.932***
		(0.031)	(0.054)	(0.004)	(0.013)	(0.027)	(0.029)	(0.077)	(0.148)
	Rob(5.5)	**0.253**	**0.808**	**0.008**	**0.937**	**0.337**	**0.273**	**-0.157**	***42.478***
		(0.030)	(0.060)	(0.004)	(0.012)	(0.011)	(0.028)	(0.054)	(0.051)

Note: Bold numbers represent the significance above 10% level (for example, including 1% and 5%). The Bold and Italic numbers on the last column represent that the J-statistics cannot reject the model at 1% significance level.

Table 14: Empirical Evaluation between GASV-RV versus RVGARCH

Moments		Data	GASV-RV	RGASV-RV	LGARCH (1,2)	LGARCH (1,1)
	AIG	3.1346	1.2390	1.4019	1.8787	1.6636
	CVX	1.5760	1.1426	1.1517	1.4670	1.5278
Var(x_t)	JPM	3.5840	1.4562	1.2982	9.9298	0.8106
	PG	0.8210	0.6304	0.6947	0.8525	0.8354
	SPY	0.8833	0.5517	0.6355	0.7774	0.8110
	T	2.8238	1.5404	1.5416	1.4044	2.5776
	AIG	9.8994	7.5016	7.8777	5.0188	6.4528
	CVX	4.3621	3.6221	3.6668	3.3380	3.6125
Kurt(x_t)	JPM	28.5051	6.3402	5.9266	10.1135	5.3895
	PG	5.8355	3.9596	3.9656	3.4715	3.6782
	SPY	8.0488	4.5231	4.5727	4.4107	5.3376
	T	7.3966	5.3548	5.3171	4.4916	5.3364
	AIG	2.8758	2.9925	2.9888	2.9956	2.9943
	CVX	3.1848	2.9986	3.0011	3.0345	3.0161
Kurt(y_t)	JPM	2.7852	3.0037	2.9943	2.9534	2.9979
	PG	4.0287	2.9967	2.9967	3.0216	3.0120
	SPY	3.0837	3.0022	2.9909	2.9997	2.9977
	T	3.1894	2.9963	2.9964	3.0145	3.0050
	AIG	-0.0948	-0.0722	-0.0423	-0.0101	-0.0120
	CVX	-0.1514	-0.0734	0.0081	-0.0577	-0.0195
Corr(x_t, y_{t+1})	JPM	-0.0440	-0.0405	-0.0465	-0.0134	-0.0287
	PG	-0.0387	-0.1351	-0.1016	-0.0376	-0.0288
	SPY	-0.1192	-0.1087	-0.0333	-0.0411	-0.0285
	T	-0.0625	-0.0646	-0.0481	-0.0182	-0.0131
	AIG	0.4529	0.3417	0.3437	0.3518	0.3696
	CVX	0.4010	0.2426	0.2358	0.3299	0.3518
Corr(x_t^2, y_t)	JPM	0.2900	0.3451	0.3406	0.4040	0.4059
	PG	0.3739	0.2583	0.2564	0.3392	0.3609
	SPY	0.4001	0.3122	0.3166	0.3868	0.4068
	T	0.4556	0.3195	0.3198	0.4023	0.4147

References

[1] Arellano, M., Hansen, L.P. and Sentana, E., 2009. Underidentification?, working paper

[2] Andersen, T and Sorensen, B., 1996. GMM Estimation of a Stochastic Volatility Model: A Monte Carlo Study, Journal of Business and Economic Statistics, 14, 329-352.

[3] Andersen, T.G. and Bollerslev, T., 1998. Answering the Skeptics: Yes, Standard Volatility Models Do Provide Accurate Forecasts, International Economic Reviews, Vol. 39, 115-158.

[4] Andersen, T., Bollerslev, T., Diebold, F.X. and Labys, P., 2003. Modelling and Forecasting Realized Volatility, Econometrica, Vol. 71, 529-626.

[5] Andrews, D. W. K. 1991. Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation, Econometrica, 59, 817-858.

[6] 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.

[7] 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.

[8] Barndorff-Nielsen, O. E. and Shephard, N., 2002. Econometric Analysis of Realized Volatility and its Use in Estimating Stochastic Volatility Models, Jounal of the Royal Statistical Society B, Vol. 64, 253-280.

[9] Bollerslev, T., 1986. Generalized Autoregressive Conditional Heteroskedasticity, Journal of Econometrics, 31, 307-327.

[10] Bollerslev, T., Litvinova, J., and Tauchen, G., 2006. Leverage and Volatility Feedback Effects in High-Frequency Data, Journal of Financial Econometrics, Vol. 4, No. 3, 353-384.

[11] Bollerslev, T. and Zhou, H., 2006. A Simple Framework for Gauging Return-Volatility Regressions, Journal of Econometrics, Vol. 131, 123-150.

[12] Broto, C. and E. Ruiz, 2004. Estimation Methods for Stochastic Volatility Models: A Survey, Journal of Economic Surveys, Vol. 18, No. 5, 613-649.

[13] Carnero, A., Pena, D. and Ruiz, E, 2004. Persistence and Kurtosis in GARCH and Stochastic Volatility Models, Journal of Financial Econometrics, 2, 319-342.

[14] Carrasco, M., 2010. A Regularization Approach to the Many Instruments Problem, Working Paper, University de Montreal.

[15] Chaussé, P., (2010), Computing Generalized Method of Moments and Generalized Empirical Likelihood with R. Journal of Statistical Software, 34(11), 1-35 (http://www.jstatsoft.org/v34/i11/).

[16] Chaussé, P., (2011), The Generalized Empirical Likelihood for a Continuum of Moment Conditions. Working Paper, University of Waterloo.

[17] Danielsson, J. and Richard, J. F. 1993, Accelerated Gaussian Importance Sampler with Application to Dynamic Latent Variable Models, Journal of Applied Econometrics, 8, 153-173.

[18] De Jong, R. and Han, C. 2002, The Properties of L_p-GMM Estimators, Econometric Theory, 18, 491-504.

[19] Diebold, F. X. and Strasser, G. H. , 2010. On the Correlation Structure of Microstructure Noise: A Financial Economic Approach, NBER Working Paper.

[20] Dominguez, M. and Lobato, I., 2004, Consistent Estimation of Models Defined by Conditional Moment Restrictions, Econometrica, 72, 1601 - 1615.

[21] Doran, H.E. and Schmidt, P., 2006, GMM Estimators with Improved Finite Sample Properties using Principal Components of the Weighting Matrix, wit an Application to the Dynamic Panel Data Model, Journal of Econometrics, 133, 387 - 409.

[22] Engle, R. F., 1982. Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of United Kingdom inflation, Econometrica, 50, 987-1007.

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

[24] Gallant, A. R. and Tauchen, G. 1996, Which Moments to Match, Econometric Theory, 12, 657-681.

[25] Ghysels, E., Harvey, A. C. and Renault, E. 1996, "Stochastic Volatility", In G.S. Maddala and C.R.Rao (Eds) Statistical Methods in Finance. 119-191.

[26] Hall, A. R. 2000, Covariance Matrix Estimation and the Power of the Overidentifying Restrictions Test, Econometrica, 68, 1517-1527.

[27] Hall, A. R. and Inoue, A., 2003, The Large Sample Behaviour of the GMM Estimator in Misspecified Models, Journal of Econometrics, 114, 361-394

[28] Hansen, P.R., Huang, Z. and Shek, H. H., 2010. Realized GARCH: A Joint Model for Returns and Realized Measures of Volatility, Journal of Applied Econometrics, Forthcoming.

[29] Hansen, P.R. and A. Lunde, 2005. A Realized Variance for the Whole Day Based on Intermittent High-Frequency Data, Journal of Financial Econometrics, 3, 525-554.

[30] Harvey, A. C., Ruiz, E. and Shepard, N.G. 1994, Multivariate Stochastic Variance Models, Review of Economic Studies, 61, 247-264.

[31] Harvey, A. C. and Shephard, N. G. 1996. Estimation of an asymmetric stochastic volatility model for asset returns. Journal of Business and Economic Statistics 14: 429-434.

[32] Imhof, G.H. 1961, Computing the Distribution of Quadratic Forms of Normal Variables, Biometrika, 48: 419-426.

[33] Jacquier, E., Polson, N. G. and Rossi, P.E. 1994, Bayesian Analysis of Stochastic Volatility Models (with discussion). Journal of Business and Economic Statistics, 12: 371-417.

[34] Knight, J. L. and Yu, J. 2002, The Empirical Characteristic Function in Time Series Estimation, Econometric Theory, 18, 691-721.

[35] Lee, H-C. and Halverson, D.R. 2004, Measuring performance of robust estimator for the variance of generalized Gaussian distribution, Eighth International Symposium on Spread Spectrum Techniques and Applications.

[36] McAleer, M. and Medeiros, M., 2008. Realized Volatility: A Review, Econometric Reviews, Vol 27, 10-45.

[37] Martens, M., 2002. Measuring and Forecasting S&P 500 Index Futures Volatility Using High-Frequency Data, Journal of Furture Markets, Vol. 22, 497-518.

[38] Martens, M. and van Dijk, 2007. Measuring Volatility with the Realized Range, Journal of Econometrics, Vol. 138, 181-207.

[39] Meddahi, N., 2002. A Theoretical Comparison Between Integrated and Realized Volatility, Journal of Applied Econometrics, 2002, 17, 479-508.

[40] Newey, W. K., and Smith, R. J., 2004. Higher Order Properties of GMM and Generalized Empirical Likelihood Estimators, Econometrica, Vol. 72, No. 1, 219-255.

[41] Park, B-J., 2009. Risk-Return Relationship in Equity Markets: using a Robust GMM Estimator for GARCH-M models, Quantitative Finance, Vol. 9, No. 1, 93-104.

[42] Ronchetti, E. and Trojani, F. 2001, Robust Inference with GMM Estimators, Journal of Econometrics, Vol. 101, 37-69.

[43] Takahashi, M., Omori, Y. and Watanabe. T., 2009. Estimating Stochastic Volatility Models using Daily Returns and Realized Volatility Simultaneously, Computational Statistics & Data Analysis, Vol. 53, 2404-2426.

[44] Taylor, S. J., 1986. Modelling Financial Time Series, Wiley: Chichester, UK.

[45] Yu, J., 2005. On Leverage in a Stochastic Volatility Model, Journal of Econometrics, 127, 165-178.

[46] Zhang, L. Mykland, P. A., and Ait-Sahalia, Y., 2005. A Tale of Two Scales: Determining Integrated Volatility with Noisy High Frequency Data, Journal of the American Statistical Association, Vol. 100, 1394-1411.