Journal of Economic Theory and Econometrics, Vol. 25, No. 3, Sep. 2014, 81–100

Semiparametric ARCH-X Model for Leverage Effect and Long Memory in Stock Return Volatility

Shen Zhang* Heejoon Han†

Abstract This paper investigates a new semiparametric ARCH-X model to account for both leverage effect and long memory property in volatility. It is a partial linear model combining a nonparametric ARCH component and an exogenous covariate that is persistent in memory. This model can allow for a flexible functional form of the asymmetric relationship between stock return and volatility and generate the long memory property in volatility. We adopt a realized volatility measure as the covariate in our model. For the daily FTSE 100 Index return series, the nonparametric component of our model captures the leverage effect and is estimated to be a complex nonlinear function. It is shown that our model outperforms other parametric or nonparametric volatility models both in within-sample and out-of-sample forecasts.

Keywords nonparametric ARCH, semiparametric volatility model, local maximum log-likelihood, leverage effect, long memory property

JEL Classification C14, C22, G12

*SCAPE, Department of Economics, National University of Singapore. E-mail: ecszhang@nus.edu.

†Corresponding author. Department of Economics, Sungkyunkwan University. E-mail: heejoonhan@skku.edu.

Received March 1, 2014, Revised June 30, 2014, Accepted Semptember 19, 2014

1. INTRODUCTION

Financial time series such as stock or exchange rate return series commonly exhibit the *long memory property in volatility*; the autocorrelation of squared return series decays very slowly. Ding et al. (1993) found earlier that it is possible to characterize the power transformation of stock return series to be long memory. For stock return series, it is also well known that the relationship between stock return and volatility is asymmetric; volatility is higher when stock return is negative than when it is positive. Such an asymmetric relationship between stock return and volatility is called the *leverage effect*.

In the literature of parametric ARCH type models, there have been active attempts to accommodate the long memory property in modeling volatility. See Baillie et al. (1996), Ding and Granger (1996), Bollerslev and Mikkelsen (1996), Engle and Lee (1999), Diebold and Inoue (2001), Mikosch and Starica (2004), Granger and Hyung (2004), Park (2002) and Han and Park (2008). However, there has been less attention on this property in the literature of nonparametric or semiparametric volatility models. Han and Zhang (2012) only recently focused on the long memory property in the framework of a nonparametric volatility model. However, their model does not explicitly allow for the leverage effect.

To capture the leverage effect, a simple nonlinear function is typically adopted in the literature of parametric ARCH type models; for example, the GJR-GARCH model by Glosten et al. (1993) and the EGARCH model by Nelson (1991). In the mean time, there has been active research on nonparametric or semiparametric volatility models that allow for a general shape to the asymmetric relationship between stock return and volatility. See Linton (2009) and Linton and Yan (2011) for excellent reviews. The nonparametric ARCH literature begins with Pagan and Schwert (1990) and Pagan and Hong (1991), where the conditional variance $\left(\sigma_t^2\right)$ of a demeaned return series (y_t) is defined as

$$\sigma_t^2 = m(y_{t-1}) \quad \text{or} \quad \sigma_t^2 = m(y_{t-1}, y_{t-2}, \cdots, y_{t-d})$$

for a smooth but unknown function $m(\cdot)$.

To the best of our knowledge, there has been no model that accommodates the long memory property while it allows for a flexible functional form of the asymmetric relationship between stock return and volatility. The aim of this paper is to fill this gap. We propose a new model that captures the two common properties observed in stock return series and also allows for a general shape to the leverage effect. Our model is a simple semiparametric model, defined as

$$\sigma_t^2 = m(y_{t-1}) + \pi x_{t-1}^2,$$

where $m(\cdot)$ is a smooth but unknown function and the covariate (x_t) is chosen among persistent economic or financial indicators. We refer to this model as the *semiparametric ARCH-X* model. The model captures the asymmetric relationship between stock return and volatility in a flexible way by $m(y_{t-1})$. It can also generate the long memory property in volatility, according to asymptotic results in Han (2014), if the added covariate x_{t-1}^2 is persistent in memory.

As the covariate (x_t) , we choose a realized volatility measure constructed from high frequency data in the empirical application of the model. Recently, various realized volatility measures have been adopted as covariates in the GARCH-X models with the rapid development seen in the field of realized volatility; see Barndorff-Nielsen and Shephard (2007), Engle (2002), Engle and Gallo (2006), Hansen et al. (2012), and Shephard and Sheppard (2010). It is shown that using a realized volatility measure as a covariate is useful in within-sample fitting and out-of-sample forecasting. Moreover, realized volatility measures are known to be persistent. Therefore, a realized volatility measure is a reasonable choice as the covariate in our model.

It is of our interest to investigate whether the features of our model will be helpful in forecasting volatility. We consider the daily FTSE 100 Index return series from 21 October 1997 to 27 February 2009 (2844 trading days) and investigate within-sample and out-of-sample predictive power of our model. The forecast evaluation is based on the QLIKE loss function and we use the realized kernel, introduced by Barndorff-Nielsen et al. (2008), as the proxy for actual volatility. Our model performs the best providing the smallest QLIKE loss in both within-sample and out-of-sample forecasts.

The estimation result shows that the nonparametric component of our model captures the asymmetric relationship between stock return and volatility. Moreover, the estimated nonlinear function of the lagged stock return is complex and some parametric counterparts to our model has difficulty in capturing such a complex nonlinear relationship. This could be one of the main reasons why our model outperforms the benchmark models.

The rest of the paper is organized as follows. Section 2 introduces the model with required assumptions and discusses the estimation method. Section 3 provides an empirical application of the model, which includes data description, estimation results, and within-sample and out-of sample forecast evaluation results. Section 4 concludes the paper.

2. MODEL AND ESTIMATION METHOD

2.1. THE MODEL

Our new semiparametric volatility model is introduced in the following assumptions. We observe $\{y_t, x_t\}$ at time t. We write the time series (y_t) to be modeled as

$$y_t = \sigma_t \varepsilon_t$$

and let $(\mathscr{F}_t)$ be a filtration with $\mathscr{F}_t$ for each t denoting information available at time t.

Assumption 2.1 Assume that
(a) (ε_t) is iid (0,1) with $\mathbb{E}[\varepsilon_t|\mathscr{F}_{t-1}] = 0$ and $\mathbb{E}\left[\varepsilon_t^2|\mathscr{F}_{t-1}\right] = 1$,
(b)

$$\sigma_t^2 = m(y_{t-1}) + \pi x_{t-1}^2 \tag{1}$$

for $\pi \geq 0$ and a smooth but unknown function $m(\cdot)$ such that $m(z) > 0$ for all $z \in \mathbb{R}$.

Under Assumption 2.1, we have

$$\mathbb{E}(y_t|\mathscr{F}_{t-1}) = 0 \quad \text{and} \quad \mathbb{E}(y_t^2|\mathscr{F}_{t-1}) = \sigma_t^2.$$

The time series (y_t) has conditional mean zero with respect to the filtration $(\mathscr{F}_t)$, and therefore, $(y_t, \mathscr{F}_t)$ is a martingale difference sequence. However, it is conditionally heteroskedastic with conditional variance (σ_t^2).

Assumption 2.2A Assume that
(a) for $1/4 < d < 1/2$,

$$(1-L)^d x_t = v_t,$$

(b) (v_t) is iid $N(0, \sigma_v^2)$.

Assumptions 2.1 and 2.2A define our semiparametric ARCH-X model. Our model includes the nonparametric ARCH component $m(y_{t-1})$ introduced by Pagan and Schwert (1990). As shown by Pagan and Schwert (1990), $m(y_{t-1})$ captures the asymmetric relationship between stock return and volatility in a flexible way. While the nonparametric ARCH component explains the short-term movement of volatility, the other parametric component πx_{t-1}^2 explains the long-term movement of volatility because the covariate x_t^2 is persistent in memory.

We employ x_{t-1}^2 instead of x_{t-1} in (1) to guarantee that the conditional variance is positive. We can consider (x_t^2) to represent a nonnegative covariate that we choose among economic or financial indicators. If the chosen covariate contains useful information on the volatility of time series, it will improve the performance of the model in within-sample fitting or out-of-sample forecasting.

Most importantly, under Assumption 2.2A, our model can generate the long memory property in volatility. The parametric counterpart to our model is the ARCH-X model defined as

$$\sigma_t^2 = \omega + \alpha y_{t-1}^2 + \pi x_{t-1}^2.$$

This model is investigated as the GARCH-X model with $\beta = 0$ in Han (2014) and it is shown that the ARCH-X process can generate the long memory property as long as the covariate is persistent. Han (2014) showed that, for $1/4 < d < 1/2$, the autocorrelation function of y_t^2 in the ARCH-X process decreases hyperbolically. This is because the covariate x_t^2 is a long memory process LM($2d - 1/2$) under Assumption 2.2A as shown by Dittmann and Granger (2002).[1] See the remarks 2.2 and 2.4 in Han (2014) for details. Since the persistence in the covariate x_t^2 plays a key role in generating the slow decaying autocorrelation of y_t^2 in the ARCH-X model, we expect that our model can also explain the long memory property in volatility. For example, if $m(y_{t-1})$ in (1) includes αy_{t-1}^2 for $\alpha < 1$, our model can exhibit the slow decaying autocorrelation of y_t^2 as the ARCH-X model can.

While this paper considers only the semiparametric ARCH-X model defined as Assumptions 2.1 and 2.2A, it is also possible to consider the following cases.

Assumption 2.2B Assume that (x_t) satisfies Assumption 2.2A with $-1/2 < d \leq 1/4$.

Assumption 2.2C Assume that (x_t) is a fractionally integrated process $I(d)$ for $1/2 < d < 3/2$ satisfying Assumption 2B of Han (2014).

Dittmann and Granger (2002) showed that (x_t^2) is a short memory process LM(0) under Assumption 2.2B. Han (2014) showed that the ARCH-X process cannot generate the long memory property in this case. Hence, a short memory case can be covered by the model under Assumption 2.2B. Meanwhile,

[1] We use the notation of LM(d) following Dittmann and Granger (2002). If $(x_t) \sim$ LM(d) with $0 < d < 1/2$, (x_t) is a long memory process with a hyperbolically decreasing autocovariance function of the form $\gamma_x(k) \sim k^{2d-1}$.

in the nonstationary case under Assumption 2.2C, the model can generate the long memory property. Han (2014) showed that the autocorrelation of y_t^2 in the ARCH-X process under Assumption 2.2C decreases exponentially at first and converges to a positive random limit. See the remarks 2.3 and 2.4 in Han (2014) for details. However, we leave the nonstationary case under Assumption 2.2C as future work. Since our model is a semiparametric partial linear model, we adopt the two-step estimation method using a feasible least squares estimator. For the nonstationary case, it is hard to establish the asymptotic distribution of the estimator. We provide more discussions on this issue in the next subsection.

Our model can be also extended to the following nonstationary semiparametric model:

$$\sigma_t^2 = m(y_{t-1}) + f(x_{t-1}) \tag{2}$$

where (x_t) is integrated or near-integrated and $f(\cdot) : \mathbb{R} \to \mathbb{R}_+$ is a parametric nonlinear function. $f(\cdot)$ can be either *integrable* ($f \in \mathbb{I}$) or *asymptotically homogeneous* ($f \in \mathbb{H}$). The reader is referred to Park and Phillips (1999) for more details on these function classes. The parametric counterpart to this model is the *ARCH-NNH* model by Han and Park (2008) given as

$$\sigma_t^2 = \alpha y_{t-1}^2 + f(x_{t-1}).$$

They provided asymptotic theories showing that the ARCH-NNH model captures the long memory property in volatility due to the persistent covariate x_{t-1}. The model in (2) is not partial linear and therefore one cannot use the two-step estimation method using a feasible least squares estimator. Alternatively, we can use the local likelihood estimation method, but it is challenging to establish the asymptotic distribution because of nonstationarity. We also leave this nonstationary model as future work.

2.2. ESTIMATION METHOD

To estimate our model, we use the following two-step estimation method. We arrange the model as

$$y_t^2 = m(y_{t-1}) + \pi x_{t-1}^2 + u_t$$

where $u_t = \sigma_t^2(\varepsilon_t^2 - 1)$. Here (u_t) is a martingale difference sequence. We first obtain the estimate of π from the least squares regression

$$\left[y_t^2 - \mathbb{E}\left(y_t^2 | y_{t-1}\right)\right] = \pi \left[x_{t-1}^2 - \mathbb{E}\left(x_{t-1}^2 | y_{t-1}\right)\right] + u_t,$$

which is

$$\hat{\pi} = \frac{\sum_{t=1}^{n}\left(x_{t-1}^{2} - \widehat{\mathbb{E}}\left(x_{t-1}^{2}|y_{t-1}\right)\right)\left(y_{t}^{2} - \widehat{\mathbb{E}}\left(y_{t}^{2}|y_{t-1}\right)\right)}{\sum_{t=1}^{n}\left(x_{t-1}^{2} - \widehat{\mathbb{E}}\left(x_{t-1}^{2}|y_{t-1}\right)\right)^{2}} \tag{3}$$

where $\widehat{\mathbb{E}}\left(x_{t-1}^{2}|y_{t-1}\right)$ and $\widehat{\mathbb{E}}\left(y_{t}^{2}|y_{t-1}\right)$ are the kernel-based estimators of $\mathbb{E}\left(x_{t-1}^{2}|y_{t-1}\right)$ and $\mathbb{E}\left(y_{t}^{2}|y_{t-1}\right)$, respectively. The kernel estimate of $\mathbb{E}\left(w_t|z_t = z\right)$ is

$$\hat{m}(z) = \sum_{j=1}^{n}\left(K_h\left(z - z_j\right) \Big/ \sum_{k=1}^{n} K_h\left(z_k - z\right)\right) w_j.$$

We let $K(s)$ be a nonnegative real function and set $K_h(s) = h^{-1}K(s/h)$. The Gaussian kernel is used for nonparametric estimation in all cases throughout the paper.

Robinson (1988) and Speckman (1988) showed that $\hat{\pi}$ in (3) is $\sqrt{n}$-consistent and asymptotically normal under some regularity conditions. Andrews (1994) provided a general framework for asymptotics for semiparametric models using results concerning the stochastic equicontinuity of stochastic processes. Under Assumptions 2.2A, (x_t^2) is strictly stationary and ergodic as shown in the proof of Lemma 1 in Han (2014). This implies that (y_t) is in general strictly stationary if we assume that $m(y_{t-1})$ is stationary. For example, if $m(y_{t-1}) = \omega + \alpha y_{t-1}^2$ for $\alpha < 1$ as in the ARCH-X model, (y_t) is strictly stationary under Assumptions 2.2A. Therefore, the asymptotic result established by Andrews (1994) can be applied for $\hat{\pi}$ in our model.

In the second step, the nonparametric component $m(y_{t-1})$ can be estimated from

$$y_t^2 - \hat{\pi} x_{t-1}^2 = m\left(y_{t-1}\right) + u_t. \tag{4}$$

However, in this step we do not apply the Nadaraya-Watson kernel method because we have to guarantee that $\hat{m}\left(y_{t-1}\right) > 0$. Given the least squares estimate $\hat{\pi}$, it is possible that some values of $\left(y_t^2 - \hat{\pi} x_{t-1}^2\right)$ are negative, which could result in some negative values in $\hat{m}\left(y_{t-1}\right)$ if we use the Nadaraya-Watson kernel method in the second step. This actually happens in our empirical application in Section 3.

Instead, we choose the local maximum likelihood estimation method of Fan and Yao (1998). To overcome the negativity problem, we take the exponential of the linear form to approximate the nonparametric component $m(y_{t-1})$ in the model. The weighted conditional local log-likelihood function is

$$-\frac{1}{2}\sum_{t=1}^{n} L\left\{\exp(\alpha+\beta(z_t-z))+\widehat{\pi}x_{t-1}^2, y_t\right\} K_h(z_t-z)$$

where $L(\lambda, y)=\lambda^{-1}y^2+\log\lambda$ and z_t is y_{t-1}. We maximize this log-likelihood function and obtain the local maximum likelihood estimator of $m(y_{t-1})$ in (4), which is

$$\hat{m}(z)=\exp(\widehat{\alpha}). \tag{5}$$

This completes the two-step estimation of our model.

Finally, we comment on some difficulties in establishing the asymptotic distribution of $\hat{\pi}$ for the nonstationary case under Assumption 2.2C. To establish the asymptotic distribution of $\hat{\pi}$, we also need asymptotic results for the nonparametric estimators of $\mathbb{E}\left(x_{t-1}^2|y_{t-1}\right)$ and $\mathbb{E}\left(y_t^2|y_{t-1}\right)$. However, there exists no available asymptotic result for such cases considering that (x_{t-1}) is a nonstationary fractionally integrated process and (y_{t-1}) is neither stationary nor integrated. Nonparametric estimation has been conducted mostly in a stationary mixing framework. See, for example, Robinson (1988), Marsry and Tjϕstheim (1995), and references therein. For most nonparametric or semiparametric volatility models, it is assumed that (y_t) is strictly or covariance stationary with some mixing conditions when necessary. Recently some works introduced nonparametric estimation theory in a nonstationary situation. See Karlsen and Tjϕstheim (2001), Karlsen et al. (2007) and Wang and Phillips (2009a, 2009b). However, these are still not suitable for the nonstationary case of our model because (y_{t-1}) is neither integrated nor fractionally integrated in the model. Hence, without development in the asymptotic theory of a nonparametric estimator for such variables, it is hard to establish the asymptotics for the nonstationary case of the semiparametric ARCH-X model.

3. EMPIRICAL APPLICATION

3.1. THE DATA, MODELS AND ESTIMATION METHODS

We consider the daily FTSE 100 Index return series from 21 October 1997 to 27 February 2009 (2844 trading days). We demean the return series by subtracting its sample mean which is close to zero. We use the demeaned return series as (y_t). As the covariate for our semiparametric ARCH-X model, we choose a realized volatility measure constructed from high frequency data. Realized volatility measures are known to be persistent and recently realized volatility measures have been included as exogenous covariates in the framework of the GARCH-X model. See Engle (2002), Engle and Gallo (2006), Barndorff-Nielsen and Shephard (2007), Cipollini *et al.* (2007), Shephard and Sheppard (2010), and Hansen *et al.* (2012). These works motivate us to adopt a realized volatility measure as the covariate in our model. In particular, we choose the realized kernel, introduced by Barndorff-Nielsen *et al.* (2008), because it has some robustness to market microstructure effects. The realized kernel of the FTSE 100 Index return is persistent; the log periodogram estimate of the memory parameter d is 0.410 and its standard error is 0.027.[2]

In this section, we estimate the following models and compare their within-sample and out-of-sample predictive ability:

$$\begin{array}{ll}
\sigma_t^2 = \omega + \alpha y_{t-1}^2 + \beta\sigma_{t-1}^2 & \text{GARCH(1,1)}, \\
\sigma_t^2 = \omega + (\alpha + \gamma s_{t-1}) y_{t-1}^2 + \beta\sigma_{t-1}^2 & \text{GJR-GARCH(1,1)}, \\
\sigma_t^2 = \omega + \alpha y_{t-1}^2 + \pi x_{t-1}^2 & \text{ARCH-X}, \\
\sigma_t^2 = \omega + (\alpha + \gamma s_{t-1}) y_{t-1}^2 + \pi x_{t-1}^2 & \text{GJR-ARCH-X}, \\
\sigma_t^2 = m(x_{t-1}^2, y_{t-1}) & \text{nonparametric model}, \\
\sigma_t^2 = m(y_{t-1}) + \pi x_{t-1}^2 & \text{semiparametric ARCH-X},
\end{array}$$

where (y_t) and $\left(x_t^2\right)$ are the demeaned stock return series and the realized kernel, respectively. The first benchmark model is the GARCH(1,1) model and the second one is the GJR-GARCH model with a dummy variable s_{t-1}, where s_t is 1 when $y_t < 0$ and zero otherwise. While the volatility at time t is a function of all past values of y_t in these two models, the rest models use only y_{t-1} and x_{t-1}^2. The third benchmark model is the ARCH-X model. The fourth one includes an asymmetric term as the GJR-GARCH model and therefore we name it as the GJR-ARCH-X model. The fifth benchmark model is a two-dimensional nonparametric model. The last model is our semiparametric ARCH-X model, which is a semiparametric counterpart to the ARCH-X, GJR-ARCH-X and the nonparametric model.

For the four parametric ARCH type models, we use the quasi-maximum likelihood estimation method which is the standard estimation method. For the nonparametric model, we use the Nadaraya-Watson kernel estimation method.

For the bandwidth selection in the estimation of our model, we use the cross-validation bandwidth obtained from the following iterating procedure. Given an initial value for bandwidth h (we use the Silverman's bandwidth as the initial value), we first obtain the least squares estimate $\hat{\pi}$ in (3) and the nonparametric

[2]The realized kernel of the daily FTSE 100 index return series is available at the database 'Oxford-Man Institute's realised library' produced by Heber *et al.* (2009).

estimate $\hat{m}(y_{t-1})$ in (5). Then we minimize the following weighted local log-likelihood criterion to obtain the cross-validation bandwidth h_{cv}:

$$h_{cv} = \arg\min_h \frac{1}{n} \sum_{j=1}^{n} \sum_{t=1, t\neq j}^{n} \frac{K_h(z_j - z_t)}{\sum_{k=1}^{n} K_h(z_k - z_j)} \left\{ \frac{y_t^2}{\widehat{m}(z_t) + \widehat{\pi} x_{t-1}^2} + \log(\widehat{m}(z_t) + \widehat{\pi} x_{t-1}^2) \right\}$$

where z_t is y_{t-1}. Next, we use this h_{cv} as a new initial bandwidth and repeat the same procedure. This procedure is repeated until the convergence of bandwidth is reached. For our data, we need only a few iterations for the convergence. We also tried different bandwidths including the Silverman's bandwidth. It is shown that the estimation results of our model are relatively robust to the choice of bandwidth.

3.2. ESTIMATION RESULTS

Table 1 reports the within-sample estimation results of the models. The estimation results of the GARCH(1,1) and the GJR-GARCH models are similar to typical results reported before in the literature. For the GARCH(1,1) model, the ARCH effects are close to unity ($\hat{\alpha} + \hat{\beta} = 0.996$). For the GJR-GARCH model, the coefficient of the asymmetric term $s_{t-1} y_{t-1}^2$ is estimated to be 0.123 and it is significant. This implies that the model captures the leverage effect.

For the ARCH-X model, the coefficient of the covariate x_{t-1}^2 is estimated to be significant. This implies that the ARCH-X model captures the slow decaying autocorrelation of squared returns because the covariate, realized kernel, is persistent. However, the coefficient of y_{t-1}^2 is estimated to be insignificant. When we include the asymmetric term, the coefficients of y_{t-1}^2 and $s_{t-1} y_{t-1}^2$ are still insignificant in the GJR-ARCH-X model. These results may indicate that it is unnecessary to include y_{t-1}^2 or $s_{t-1} y_{t-1}^2$ in the ARCH-X model when the realized kernel is added as the covariate. Alternatively, these results may be due to misspecification of the functional form of y_{t-1}.

Compared to the ARCH-X model, our semiparametric ARCH-X model let the functional form of the lagged stock return y_{t-1} be flexible. For our model, the coefficient of the covariate x_{t-1}^2 is still estimated to be significant. However, when the functional form of y_{t-1} is allowed to be flexible, the magnitude of $\hat{\pi}$ becomes smaller than that of the ARCH-X or GJR-ARCH-X model: $\hat{\pi}$ is 1.268 in our model while it is 1.769 and 1.786 in the ARCH-X and the GJR-ARCH-X models, respectively.

The estimated nonparametric component $\widehat{m}(y_{t-1})$ in our model is plotted on the grid of values $\{y_{t-1} = -0.09 + 0.002k;\ k = 0, 1, \cdots, 90\}$ in Figure 1. The stock return series, used as (y_t), is ranged from -0.093 to 0.088 in our sample. The shape of the nonparametric component $\widehat{m}(y_{t-1})$ is clearly nonlinear. In general, $\widehat{m}(y_{t-1})$ is larger for a negative stock return y_{t-1}. However, more importantly, the simple asymmetric term $s_{t-1}y_{t-1}^2$ adopted in the GJR-ARCH-X model cannot adequately capture the nonlinearity in $\widehat{m}(y_{t-1})$. In Figure 1, as y_{t-1} moves from 0 to -0.09, $\widehat{m}(y_{t-1})$ exhibits a (bell-shaped) hump before it rapidly increases. The estimation result of our model shows that the nonlinear function of y_{t-1} is complex and a simple parametric function cannot capture it.

3.3. FORECAST EVALUATION RESULTS

Evaluation Criterion

Following Pagan and Schwert (1990), we evaluate the performance of our model by comparing predictive power of volatility models. The fundamental problem with the evaluation of volatility forecasts is that volatility is unobservable and therefore actual values with which to compare the forecasts do not exist. To overcome this problem, there has been recent developments in the literature of volatility forecast evaluation.

First, recent works support using a realized volatility measure as the proxy for actual volatility. See Hansen and Lunde (2006) and Patton and Sheppard (2009). It is because realized volatility measures are better estimates of actual volatility than squared return series. See Barndorff-Nielsen and Shephard (2002) and Andersen *et al.* (2003). Therefore, we use the realized kernel as the proxy for actual volatility as in Shephard and Sheppard (2010).

Second, there has been research on loss functions that are robust to the use of a noisy volatility proxy. See Hansen and Lunde (2006), Patton (2011) and Patton and Sheppard (2009). Even if realized volatility measures are better measures, they are still imperfect and noisy proxies for actual volatility. Therefore, it is important to use a robust loss function to prevent any spurious ranking of various volatility forecasts. Hansen and Lunde (2006) and Patton (2011) showed that the MSE and QLIKE are robust. In particular, Patton and Sheppard (2009) show that the QLIKE loss function has the highest power among robust loss functions. Therefore, we use the QLIKE loss function as the loss function, which is defined as

$$L(\hat{\sigma}_t^2, \sigma_t^2) = \frac{\sigma_t^2}{\hat{\sigma}_t^2} - \log\frac{\sigma_t^2}{\hat{\sigma}_t^2} - 1 \tag{6}$$

where $\left(\sigma_t^2\right)$ is the proxy for actual volatility and $\left(\hat{\sigma}_t^2\right)$ is the within-sample or out-of-sample forecast.

To test for equal forecasting accuracy of two competing models, we apply the Diebold-Mariano and West (henceforth DMW) test (see Diebold-Mariano(1995) and West (1996)). A DMW statistic is computed using the difference in the losses of two models

$$
\begin{aligned}
d_t &= L(\hat{\sigma}_{t,1}^2, \sigma_t^2) - L(\hat{\sigma}_{t,2}^2, \sigma_t^2), \\
DMW_T &= \frac{\sqrt{T}\bar{d}_T}{\sqrt{\widehat{avar}\left(\sqrt{T}\bar{d}_T\right)}} \sim N(0,1),
\end{aligned} \tag{7}
$$

where $\bar{d}_T$ is the sample mean of d_t and T is the number of forecasts. The asymptotic variance of the average is computed using a Newey-West variance estimator with the number of lags set to $\left[T^{1/3}\right]$. The null and alternative hypotheses in this case are

$$H_0 : d_t = 0 \text{ versus } H_1 : d_t \neq 0.$$

Within-sample forecast comparison

Table 2 contains the within-sample forecast evaluation result based on the QLIKE loss function.[3] In this case, $\hat{\sigma}_t^2$ in (6) denotes the fitted values of the volatility models for the entire sample period and $T = 2843$ in (7). Our model does not encompass the GARCH(1,1) and GJR-GARCH models and, as a consequence, there is no reason to expect our model to outperform these two models even for the within-sample forecast.

In Table 2, our model shows the smallest QLIKE loss of 0.257.[4] The second best model is the GJR-GARCH model and its QLIKE is 0.279. The ARCH-X and GARCH(1,1) models exhibit similar QLIKE losses of 0.292 and 0.294, respectively. The two-dimensional nonparametric model performs very poorly with the largest QLIKE of 0.719. We test the null hypothesis of equal loss by the DMW test procedure, and the test results show that the null hypotheses of equal loss between our model and five benchmark models are all rejected at either 5% or 1% significance level. This means that, in terms of the within-sample fitting, our semiparametric ARCH-X model provides a better explanation of the stock return volatility than the rest models.

[3]We also considered the nonparametric ARCH model $\sigma_t^2 = m(y_{t-1})$ by Pagan and Schwert (1990). However, we do not repot its result because it performs very poorly.

[4]We also estimated our model using the Silverman's bandwidth $\hat{\sigma}_y n^{-1/5}$ where $\hat{\sigma}_y$ is the sample standard deviation of (y_t). For $\hat{\sigma}_y n^{-1/5}$, $0.5\hat{\sigma}_y n^{-1/5}$ and $2\hat{\sigma}_y n^{-1/5}$, the QLIKE of our model is 0.256, 0.257 and 0.262, respectively. These are similar to the result obtained by the cross-validation bandwidth.

Out-of-sample forecast comparison

To check the possibility of over-fitting, we evaluate out-of-sample forecasts. If over-fitting is a serious problem, the QLIKE statistics for out-of-sample forecasts should be much larger than the QLIKE's for the within-sample forecasts. We adopt the rolling window forecast procedure with moving windows of 8 years (2016 trading days). This means that we obtain one-step ahead forecasts of the models for the period from 21 November 2005 to 27 February 2009. In this case, $\hat{\sigma}_t^2$ in (6) now denotes one-period ahead volatility forecasts at time $t-1$ and $T = 827$ in (7). For our model, we use the cross validation bandwidth chosen in the within-sample case. Table 3 reports the QLIKE's of the models and the DMW test statistics.

As in the within-sample case, our model shows the smallest QLIKE loss of 0.236. It shows that over-fitting is not a serious problem for our model because the out-of-sample QLIKE is even smaller than the within-sample counterpart. The second best model is still the GJR-GARCH model and its QLIKE is 0.248. The GARCH(1,1) model has the QLIKE of 0.281, which is smaller than that of the ARCH-X model (0.360). The worst model is still the two-dimensional nonparametric model. According to the DMW test result, the null hypothesis of equal loss between our model and each benchmark model is rejected except for the GJR-GARCH model.

4. CONCLUSION

This paper proposes and investigates a new semiparametric volatility model, which is a partial linear model combining a nonparametric ARCH component and a persistent exogenous covariate. The nonparametric component allows for a flexible functional form of the asymmetric relationship between stock return and volatility and a persistent covariate is included to generate the long memory property in volatility. We show that the features of our model is indeed helpful in forecasting stock return volatility.

For the daily FTSE UK 100 Index return series for the period from 21 October 1997 to 27 February 2009 (2844 trading days), we evaluate the within-sample fitting and the out-of-sample forecasting of the model. We use the daily realized kernel as the covariate in our model. Realized volatility measures are known to be persistent. Moreover, recently many works included them as covariates in the framework of the GARCH-X model and showed they are useful in forecasting volatility. The estimation result shows that the nonparametric component captures the leverage effect and is estimated to be a complex nonlinear function that parametric counterparts to our model cannot properly represent. This could be one of main reasons why our model outperforms the benchmark models in both within-sample and out-of-sample forecasts.

Table 1. Within-sample estimation results of the models

Parameter	GARCH(1,1)	GJR-GARCH	ARCH-X	GJR-ARCH-X	Semi-ARCH-X
α	0.105^{***}	0.012	-0.017	-0.001	
	(0.015)	(0.015)	(0.017)	(0.019)	
β	0.891^{***}	0.917^{***}			
	(0.013)	(0.010)			
γ		0.123^{***}		-0.047	
		(0.025)		(0.030)	
π			1.769^{***}	1.786^{***}	1.268^{***}
			(0.081)	(0.084)	(0.057)

Notes: The table reports the estimation results of the models described in Section 3.1. For the parametric models, it reports the quasi-maximum likelihood estimates. For the semiparametric ARCH-X model, it reports the least squares estimate $\hat{\pi}$ described in (3). The standard errors are reported in parentheses. *** indicates that the coefficient is significant at 1% level.

Table 2. Comparison of within-sample predictive power for the stock return volatility
(1997.10.21-2009.02.27)

models		QLIKE	DMW
GARCH(1,1)	$\omega+\alpha y_{t-1}^2+\beta\sigma_{t-1}^2$	0.294	3.520^{***}
GJR-GARCH	$\omega+(\alpha+\gamma s_{t-1})y_{t-1}^2+\beta\sigma_{t-1}^2$	0.279	2.007^{**}
ARCH-X	$\omega+\alpha y_{t-1}^2+\pi x_{t-1}^2$	0.292	5.612^{***}
GJR-ARCH-X	$\omega+(\alpha+\gamma s_{t-1})y_{t-1}^2+\pi x_{t-1}^2$	0.288	6.128^{***}
nonparametric model	$m(x_{t-1}^2, y_{t-1})$	0.719	13.381^{***}
Semi-ARCH-X	$m(y_{t-1})+\pi x_{t-1}^2$	**0.257**	

Notes: The QLIKE loss is defined in (6) and the DMW test statistic is defined in (7). *,** and *** signify rejecting the null hypothesis of equal loss for 10%, 5% and 1% tests, respectively.

Table 3. Comparison of out-of-sample predictive power for the stock return volatility
(2005.11.21-2009.02.27)

models		QLIKE	DMW
GARCH(1,1)	$\omega+\alpha y_{t-1}^2+\beta\sigma_{t-1}^2$	0.281	2.484^{**}
GJR-GARCH	$\omega+(\alpha+\gamma s_{t-1})y_{t-1}^2+\beta\sigma_{t-1}^2$	0.248	0.608
ARCH-X	$\omega+\alpha y_{t-1}^2+\pi x_{t-1}^2$	0.360	1.863^{*}
GJR-ARCH-X	$\omega+(\alpha+\gamma s_{t-1})y_{t-1}^2+\pi x_{t-1}^2$	0.300	5.260^{***}
nonparametric model	$m(x_{t-1}^2, y_{t-1})$	0.738	6.039^{***}
Semi-ARCH-X	$m(y_{t-1})+\pi x_{t-1}^2$	**0.236**	

Notes: The QLIKE loss is defined in (6) and the DMW test statistic is defined in (7). *,** and *** signify rejecting the null hypothesis of equal loss for 10%, 5% and 1% tests, respectively.

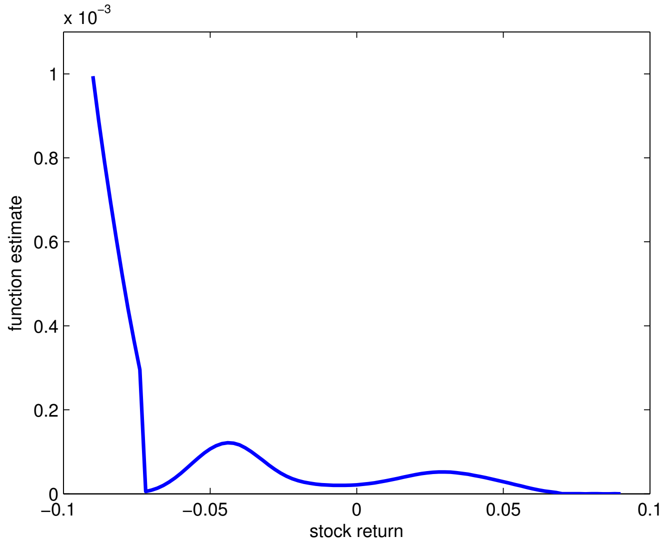

Figure 1. Estimate of $m(y_{t-1})$ in the semiparametric ARCH-X model

REFERENCES

Andersen, T.G., Bollerslev, T., Diebold, F.X. and Labys, P. (2003), Modeling and forecasting realized volatility, *Econometrica,* 71, 529-626.

Andrews, D.W.K. (1994), Asymptotics for semiparametric econometric models via stochastic equicontinuity, *Econometrica*, 62, 43-72.

Baillie, R.T., Bollerslev, T. and Mikkelsen, H.O. (1996), Fractionally integrated generalized autoregressive conditional heteroskedasticity, *Journal of Econometrics,* 74, 3-30.

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

Barndorff-Nielsen, O.E. and Shephard, N. (2007), ”Variation, jumps and high frequency data in financial econometrics” in R. Blundell, T. Persson and W.K. Newey (eds.), *Advances in Economics and Econometrics. Theory and Applications, Ninth World Congress*, Econometric Society Monographs, 328-372. Cambridge: Cambridge University Press.

Barndorff-Nielsen, O.E., Hansen, P.R., Lunde, A. and Shephard, N. (2008), Designing realized kernels to measure the ex-post variation of equity prices in the presence of noise. *Econometrica,* 76, 1481-1536.

Bollerslev, T. and Mikkelsen, H.O. (1996), Modeling and pricing long-memory in stock market volatility, *Journal of Econometrics,* 73, 151-184.

Cipollini, F., Engle, R.F. and Gallo, G. (2007), A model for multivariate non-negative valued processes in financial econometrics, mimeograph, Stern School of Business, New York University.

Diebold, F.X. and Inoue, A. (2001), Long memory and regime switching, *Journal of Econometrics,* 105, 131-159.

Diebold, F.X. and Mariano, R.S. (1995), Comparing predictive accuracy, *Journal of Business and Economic Statistics,* 13, 253-263.

Ding, Z. and Granger, C.W.J. (1996), Modeling volatility persistence of speculative returns: A new approach, *Journal of Econometrics,* 73, 185-215.

Ding, Z., Granger, C.W.J., and Engle, R.F. (1993), A long memory property of stock market returns and a new model, *Journal of Empirical Finance,* 1, 83-106.

Dittmann, I. and Granger, C.W.J. (2002), Properties of nonlinear transformations of fractionally integrated processes. *Journal of Econometrics,* 110, 113-133.

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

Engle, R.F. and Lee, G.J. (1999), A long-run and short-run component model of stock return volatility, in Engle R. and H. White ed. *Cointegration, Causality, and Forecasting: A Festschrift in Honour of Clive W.J. Granger*, Chapter 10, 475-497, Oxford University Press.

Engle, R.F. (2002), New frontiers for ARCH models, *Journal of Applied Econometrics* 17, 425-446.

Fan, J. and Yao, Q. (1998), Efficient estimation of conditional variance functions in stochastic regression. *Biometrika*, 85, 645-660.

Glosten, L.R., Jagannathan, R. and Runkle, D. (1993), On the relation between the expected value and the volatility of nominal excess returns on stocks. *Journal of Finance,* 48, 1779-1801.

Granger, C.W.J. and Hyung, N. (2004), Occasional structural breaks and long memory with an application to the S&P 500 absolute stock returns, *Journal of Empirical Finance,* 11, 399-421.

Han, H. (2014), Asymptotic Properties of GARCH-X processes, *Journal of Financial Econometrics*, forthcoming.

Han, H. and Park, J.Y. (2008), Time series properties of ARCH processes with persistent covariates, *Journal of Econometrics*, 146, 275-292.

Han, H. and Zhang, S. (2012), Non-stationary non-parametric volatility model, *The Econometrics Journal*, 15, 204-225.

Hansen, P.R. and Lunde, A. (2006), Consistent ranking of volatility models, *Journal of Econometrics*, 131, 97-121.

Hansen, P.R., Huang, Z. and Shek, H.H. (2012), Realized GARCH: a joint model for returns and realized measures of volatility. *Journal of Applied Econometrics,* 27, 877-906.

Heber, G., Lunde, A., Shephard, N. and Sheppard, K. (2009), *OMI's realised library*, Version 0.1, Oxford-Man Institute, University of Oxford.

Karlesn, H.A. and Tjϕstheim, D. (2001), Nonparametric estimation in null recurrent time series, *Annals of Statistics*, 29, 372-416.

Karlsen, H.A., Myklebust, T. and Tjϕstheim, D. (2007), Nonparametric estimation in a nonlinear cointegration type model, *Annals of Statistics*, 35, 252-299.

Linton, O.B. (2009), Semiparametric and nonparametric ARCH modeling, in Andersen, T.G., Davis, R.A., Kreiss, J.P. and Mikosch, Th. (eds.), *Handbook of Financial Time Series*, 157-168, Springer.

Linton, O.B. and Yan, Y. (2011), Semi- and nonparametric ARCH processes, *Journal of Probability & Statistics,* Volume 2011, 1-17.

Masry, E. and Tjϕstheim, D. (1995), Nonparametric estimation and identification of nonlinear ARCH time series: strong convergence and asymptotic normality, *Econometric Theory*, 11, 258-289.

Mikosch, T. and Starica, C. (2004), Nonstationarities in financial time series, the long-range dependence, and the IGARCH effects, *The Review of Economics and Statistics,* 86, 378-390.

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

Pagan, C.R. and Hong, Y.S. (1991), Nonparametric estimation and the risk premium, in Barnett, W., Powell, J, and Tauchen, G.E. (eds.), *Nonparametric and Semiparametric Methods in Econometrics and Statistics*, Cambridge University Press.

Pagan, A.R. and Schwert, G.W. (1990), Alternative models for conditional stock volatility, *Journal of Econometrics*, 45, 267-290.

Park, J.Y. (2002), Nonstationary nonlinear heteroskedasticity, *Journal of Econometrics*, 110, 383-415.

Park, J.Y. and Phillips, P.C.B. (1999), Asymptotics for nonlinear transformations of integrated time series, *Econometric Theory,* 15, 269-298.

Patton, A. J. (2011). Volatility forecast comparison using imperfect volatility proxies. *Journal of Econometrics*, 160, 246-256.

Patton, A.J. and Sheppard, K. (2009), Evaluating volatility and correlation forecasts, in Andersen, T.G., Davis, R.A., Kreiss, J.P. and Mikosch, Th. (eds.), *Handbook of Financial Time Series*, 801-838, Springer.

Robinson, P.M. (1988), Root-n-consistent semiparametric regression, *Econometrica*, 56, 931-954.

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

Speckman, P. (1988), Kernel smoothing in partial linear models, *Journal of the Royal Statistical Society B*, 50, 413-446.

Wang, Q. and Phillips, P.C.B. (2009a), Asymptotic theory for local time density estimation and nonparametric cointegrating regression, *Econometric Theory,* 25, 1-29.

Wang, Q. and Phillips, P.C.B. (2009b), Structural nonparametric cointegrating regression, *Econometrica*, 77, 1901-1948.

West, K.D. (1996), Asymptotic inference about predictive ability, *Econometrica*, 64, 1067-1084.