Available online at www.sciencedirect.com
SCIENCE DIRECT®

Journal of Empirical Finance 12 (2005) 650–665

Journal of EMPIRICAL FINANCE

www.elsevier.com/locate/econbase

# The relationship between stock returns and volatility in international stock markets

Qi Li [a,*], Jian Yang [b], Cheng Hsiao [c], Young-Jae Chang [d]

[a] Department of Economics, Texas A&M University, College Station, TX 77843-4228, USA
[b] Department of Accounting, Finance and MIS, Prairie View A&M University, Prairie View, TX 77446-0638, USA
[c] Department of Economics, University of Southern California Los Angeles, CA 90089, USA
[d] Department of Business Administration, Inje University, Gimhae 621-749, Korea

Accepted 7 March 2005
Available online 30 September 2005

## Abstract

This study examines the relationship between expected stock returns and volatility in the 12 largest international stock markets during January 1980 to December 2001. Consistent with most previous studies, we find a positive but insignificant relationship during the sample period for the majority of the markets based on parametric EGARCH-M models. However, using a flexible semiparametric specification of conditional variance, we find evidence of a significant negative relationship between expected returns and volatility in 6 out of the 12 markets. The results lend some support to the recent claim [Bekaert, G., Wu, G., 2000. Asymmetric volatility and risk in equity markets. Review of Financial Studies 13, 1–42; Whitelaw, R., 2000. Stock market risk and return: an empirical equilibrium approach. Review of Financial Studies 13, 521–547] that stock market returns are negatively correlated with stock market volatility.


*JEL classification:* G12; G15; C14
*Keywords:* Risk–return tradeoff; GARCH; Semiparametric estimation

* Corresponding author. Tel.: +1 979 845 9954; fax: +1 979 847 8757.
*E-mail address:* qi@econmail.tamu.edu (Q. Li).

0927-5398/$ - see front matter 
doi:10.1016/j.jempfin.2005.03.001

## 1. Introduction

The relationship between the return on an asset and its variance (or volatility) as a proxy for risk has been an important topic in financial research. The theoretical asset-pricing models (e.g., Sharpe, 1964; Linter, 1965; Mossin, 1966; Merton, 1973, 1980) typically link the return (or the price change) of an asset to its own return variance, or to the covariance between its return and the return on the market portfolio. However, whether such a relationship is positive or negative has been controversial. As summarized in Baillie and DeGennarro (1990), most asset-pricing models (e.g., Sharpe, 1964; Linter, 1965; Mossin, 1966; Merton, 1973) postulate a positive relationship between a stock portfolio's expected returns and volatility. On the other hand, there is also a long tradition in finance that models stock return volatility as negatively correlated with stock returns (Black, 1976; Cox and Ross, 1976; Bekaert and Wu, 2000; Whitelaw, 2000). For example, Bekaert and Wu (2000, p. 1) recently claim that "it appears that volatility in equity markets is asymmetric: returns and conditional volatility are negatively correlated." Although their paper is critically motivated by such a claim, the empirical evidence for such a negative relationship between expected returns and volatility is mixed in the US stock markets and has not yet been reported in international stock markets other than the US. In this context, our study substantially complements Bekaert and Wu (2000). Furthermore, Glosten et al. (1993) and Nelson (1991) argue that across time there is no theoretical agreement about the relationship between returns and volatility within a given period of time and that either a positive or a negative relationship between current stock returns and current volatility is possible.

Numerous empirical studies have been conducted to investigate the relationship between stock market returns and volatility. The findings of early studies are mixed (e.g., Pindyck, 1984; Poterba and Summers, 1986). As pointed out by Bollerslev et al. (1992, pp. 17–18), inference from early studies may not be reliable because variance modeling in these studies does not make efficient use of the data. More recent studies have typically used (G)ARCH-in-Mean models (Engle et al., 1987) to allow for time-varying behavior of volatility. Surprisingly, most find an insignificant relationship between returns and conditional variance (as defined by the parametric GARCH process) in international stock markets. Although French et al. (1987) document a significant positive relationship between US stock market returns and the conditional variance of these returns, Baillie and DeGennarro (1990) report that such a positive relationship is weak and almost nonexistent in the US stock market. Similarly, Theodossiou and Lee (1995) and Lee et al. (2001) also find a positive but insignificant relationship between stock market returns and the conditional variance in many other international stock markets. In contrast, Nelson (1991) documents a negative but insignificant relationship between expected returns and the conditional variance of the US stock market. Glosten et al. (1993) show evidence that such a negative relationship is significant in the US market. Obviously, the empirical findings remain inconclusive.

The finding of an insignificant relationship appears puzzling. Though a significant impact of volatility on the stock prices can take place only if shocks to volatility persist over a long period of time (Poterba and Summers, 1986), it is well documented that stock market volatility is persistent. Hence, many of the previous studies, e.g., Baillie and DeGennarro (1990), Theodossiou and Lee (1995), and Choudhry (1996), challenge the appropriateness of using the conditional variance (as modeled by a parametric GARCH process) to proxy for risk and attribute the finding of the weak relationship to the lack of a proper measure of risk. In view of the above mixed results, this study uses a flexible semiparametric specification of conditional variance to examine the relationship between expected returns and volatility in 12 major stock markets. The use of a flexible functional form for conditional variance is appealing because estimation of a parametric GARCH-M model is sensitive to model misspecification. Consistent estimation in the (G)ARCH-M model requires that the full model be correctly specified (Bollerslev et al., 1992, p. 14). Indeed, the problem that inferences drawn on the basis of GARCH-M models may be highly susceptible to model misspecification is well known to applied researchers. For example, Jones et al. (1998) choose not to estimate a GARCH-M model to measure a possible change in the risk premium, simply due to the concern that a potential misspecification problem may contaminate the estimation of conditional variance parameters. Nelson (1991, p.347) also argues that parameter restrictions imposed by GARCH models may unduly restrict the dynamics of the conditional variance process. In contrast, a semiparametric specification of the conditional variance allows flexible functional forms, and therefore can lead to more reliable estimation and inference. In this paper, we propose a semiparametric test for testing the null hypothesis of zero GARCH-M effect. The simulation results show that the proposed test has good finite sample performance compared with a parametric test based on EGARCH specification. We then apply the proposed test to the empirical data of 12 largest international stock markets and show some evidence that a significant negative relationship between (current) stock market returns and (current) market volatility prevails in most major stock markets, which has not yet been reported in the literature. The rest of this paper is organized as follows. Section 2 discusses the empirical methodology, Section 3 first presents a small-scale Monte Carlo simulations to examine the finite sample performance of the proposed semiparametric test for GARCH-M effect, and then reports the empirical findings, and finally, Section 4 concludes the paper.

## 2. Empirical methodology

This section presents a brief review of empirical methodology used in this study. To examine the relationship between stock market returns and volatility, we use both a parametric and a semiparametric method. The time-varying pattern of stock market volatility has been widely recognized and modeled as a conditional variance in the parametric GARCH framework, as originally developed by Engle (1982) and generalized by Bollerslev (1986).

### *2.1. A parametric GARCH-M specification*

The parametric method in this study is based on an AR($k$)-EGARCH ($p$,$q$)-M model specified as follows:

$$y_t = \mu + \sum_{s=1}^{k} \ell_s y_{t-s} + \delta\sigma_t^2 + \varepsilon_t \tag{1}$$

$$\varepsilon_t|\Omega_{t-1} \sim \mathrm{GED}\left(0, \sigma_t^2, v\right) \tag{2}$$

$$\ln\sigma_t^2 = \omega + \sum_{i=1}^{p} \alpha_i \left\{ \phi \frac{\varepsilon_{t-i}}{\sigma_{t-i}} + \gamma \left[ \left| \frac{\varepsilon_{t-i}}{\sigma_{t-i}} \right| - E\left| \frac{\varepsilon_{t-i}}{\sigma_{t-i}} \right| \right] \right\} + \sum_{j=1}^{q} \beta_j \ln\sigma_{t-j}^2 \tag{3}$$

where $y_t$ is the stock market return,[1] $\varepsilon_t$ is the innovation distributed as a generalized exponential distribution (GED) with zero mean and time-varying conditional variance $\sigma_v^2$, $v$ represents the scale parameter for the tail-thickness of the GED distribution that takes particular values for some particular density functions, $\mu$, $\ell$, $\delta$, $v$, $\omega$, $\alpha\phi$, $\alpha\gamma$, and $\beta$ are parameters to be estimated. Eq. (1) represents dynamic changes in the mean returns, while Eq. (3) describes time variation in the conditional variance. Among all the parameters to be estimated, the most relevant one for this study is the parameter $\delta$, because the sign and significance of the parameter $\delta$ directly shed light on the nature of the relationship between stock market returns and its volatility. Also of interest is the product of $\alpha_i\phi$. The EGARCH method allows the conditional variance process to respond asymmetrically to positive and negative shocks in stock returns, which may be reflected in the value of the parameter $\alpha_i\phi$. If $\alpha_i\phi<0$ ($>0$), such an asymmetry exists and the variance tends to rise (fall) when the shock is negative (positive).

Estimating the above model requires the adoption of some density function for the innovation vector $\varepsilon_t$. The most commonly used density function is the Normal distribution (see Engle, 1982). However, the Normal density function may not be appropriate here because it fails to capture the "fat-tail" which is a common feature of stock return distributions. Following Nelson (1991), we employ the GED distribution to model stock return innovations. The density function of the GED takes the following form:

$$f(\varepsilon_t) = v\left\{\exp\left[-0.5\left|^{(\varepsilon t/\sigma t)}/\lambda\right|^{v}\right]\right\}\left\{\lambda 2^{(1+1/v)}\Gamma(1/v)\right\}^{-1},$$

where $\Gamma(\cdot)$ is the gamma function, and $\lambda = \sqrt{\left[2^{-2/v)}\Gamma(1/v)\right]/\Gamma(3/v}$. The GED is quite general. It encompasses many different density functions, depending on various values of the parameter $v$. For example, when $v=2$ we have the standard normal density, for $v=1$ we get the double exponential distribution, and if $v\rightarrow\infty$ we obtain the uniform distribution.

### 2.2. A semiparametric GARCH-M specification

The semiparametric GARCH-M model we will estimate is given by:

$$y_t = \alpha_0 + \alpha_1 y_{t-1} + \delta\sigma_t^2 + u_t \equiv x_t\alpha + u_t \tag{4}$$

where $y_t$ is the stock market returns, $x_t=(1, y_{t-1}, \sigma_t^2)$, $\alpha=(\alpha_0, \alpha_1, \delta)'$, $\sigma_t^2=\mathrm{var}(y_t|\Omega_{t-1})$ is the conditional variance of $y_t$ conditional on $\Omega_{t-1}$, $\Omega_{t-1}$ is the information set available at time $t-1$. The error term is a martingale difference process, i.e., $E(u_t|\Omega_{t-1})=0$. We are

[1] The use of the stock return is consistent with Baillie and DeGennarro (1990), Nelson (1991), Choudhry (1996), and Lee et al. (2001), among others. Also see more related discussion on this issue in the data section below.

interested in testing the null hypothesis of $H_0$: $\delta=0$ versus $H_1$: $\delta\neq 0$. The null hypothesis says that $\sigma_t^2$ does not affect return $y_t$. If $H_0$ is rejected, a positive $\delta$ implies that the expected stock return and volatility are positively related, while a negative $\delta$ implies that they are negatively related.

Before we discuss our semiparametric specification of $\sigma_t^2$, we first briefly review the nonparametric specification considered by Pagan and Ullah (1988). Pagan and Ullah (1988) suggest to use a truncated fixed $r$-lag specification to approximate $\sigma_t^2$, i.e., using $\mathrm{var}(y_t|y_{t-1},\ldots, y_{t-r})$ to approximate $\mathrm{var}(y_t|\Omega_{t-1})$, and they suggested to estimate $\mathrm{var}(y_t|y_{t-1},\ldots,y_{t-r})$ by the nonparametric kernel method. This approach can only allow a small number of lags (say $r=2$ or 3) to be used in practice because it suffers the 'curse of dimensionality' problem if $r$ is large. Therefore, this approach is difficult to capture the highly persistent nature of the variance process.

The semiparametric GARCH model we consider below does not suffer from the 'curse of dimensionality' problem as in Pagan–Ullah's specification. We first consider a simple semiparametric GARCH model ($\sigma_t^2=\mathrm{var}(y_t|I_{t-1})=\mathrm{var}(u_t|I_{t-1})$):

$$\sigma_t^2 = m(u_{t-1}) + \gamma\sigma_{t-1}^2 \tag{5}$$

where the functional form of $m(\cdot)$ is unspecified. If $m(u_{t-1})=\alpha+\beta u_{t-1}^2$, (5) reduces to a standard GARCH(1,1) model. Under the null ($H_0$) of $\delta=0$ and from (4), we have $u_{t-1}=y_{t-1}-\alpha_0-\alpha_1 y_{t-2}$. A more general form of (5) would be to let

$$\mathrm{var}(y_t|I_{t-1}) = \sigma_t^2 = g(y_{t-1},y_{t-2}) + \gamma\sigma_{t-1}^2. \tag{6}$$

When $g(y_{t-1}, y_{t-2})=m(y_{t-1}-\alpha_0-\alpha_1 y_{t-2})=m(u_{t-1})$, (6) reduces back to (5). (6) allows the conditional variance to have general interactions between $y_{t-s}$ and $y_{t-s-1}$ ($s=1,\ldots,\infty$).

Denoting $z_{t-1}=(y_{t-1}, y_{t-2})$ and substituting (6) recursively yields

$$\sigma_t^2 = g(z_{t-1}) + \gamma g(z_{t-2}) + \gamma^2 g(z_{t-3}) + \ldots + \gamma^{d-1} g(z_{t-d}) + \ldots\ldots \tag{7}$$

Given that $0<\gamma<1$, we may approximate (7) by a finite lag model if $d$ is sufficiently large:

$$\sigma_t^2 \cong g(z_{t-1}) + \gamma g(z_{t-2}) + \gamma^2 g(z_{t-3}) + \ldots + \gamma^{d-1} g(z_{t-d}). \tag{8}$$

Eq. (8) is a restricted additive model with the restriction that the different additive functions are proportional to each other. Therefore, for a fixed value of $d$, (8) is one-dimensional nonparametric model because there is only one univariate $g(\cdot)$ function that needs to be estimated. This model can allow many lagged $y_{t-s}$'s to be included at the right-hand side of (8). Unlike a purely nonparametric model with $d$-lagged valued regressors (e.g., Pagan and Ullah, 1988), the additive model (8) does not suffer from the 'curse of dimensionality' problem (e.g., Newey, 1994; Li, 2000).

Yang (2002) considers a similar model but with $y_t=u_t$ (hence $z_{t-1}$ becomes the univariate $y_{t-1}$ in Yang, 2002). Yang suggests a kernel-based method to estimate model (8). Although Eq. (8) is only a two-dimensional nonparametric model, it can be difficult to estimate by the popular kernel method when $d$ is large. This is because at the initial estimation stage, the kernel method requires one to estimate a $d$-dimensional nonparametric regression model: $E(y_t|y_{t-1},\ldots,y_{t-d})$. Then at the second stage, one uses the additive proportional model structure of (8) to obtain an estimate of the two-dimensional $g(y_{t-1}, y_{t-2})$ function. In finite sample applications and when $d$ is large, the kernel method can give quite unreliable estimation results due to its failure to impose the additive model structure at the initial estimation stage.

We suggest to estimate model (8) by the nonparametric series method (say, spline or power series). The advantage of using the series method is that the additive proportional model structure is imposed directly and the estimation is done in one step. To see this, let $\{\phi_l(y)\}_{l=0}^{\infty}$ denote a series-based function that can be used to approximate any univariate function $m(y)$, we can use a linear combination of the product base function to approximate $g(y_{t-1}, y_{t-2})$, i.e., we approximate $g(y_{t-s}, y_{t-s-1})$ by

$$\begin{aligned}\sum_{l=0}^{q}\sum_{l'=0}^{q} a_{ll'}\phi_l(y_{t-s})\phi_{l'}(y_{t-s-1}) = {} & a_{00}\phi_0(y_{t-s})\phi_0(y_{t-s-1}) + \ldots + a_{0q}\phi_0(y_{t-s})\phi_q(y_{t-s-1}) \\ & + a_{10}\phi_1(y_{t-s})\phi_0(y_{t-s-1}) + \ldots + a_{1q}\phi_1(y_{t-s})\phi_q(y_{t-s-1}) \\ & + \ldots + a_{q0}\phi_q(y_{t-s})\phi_0(y_{t-s-1}) + \ldots \\ & + a_{qq}\phi_q(y_{t-s})\phi_q(y_{t-s-1}),\end{aligned}$$

for $s=1,\ldots, d$. The approximating function, after re-arranging terms, becomes:

$$\begin{aligned}\sigma_t^2 \cong {} & a_{00}\sum_{s=1}^{d}\gamma^{s-1}\phi_0(y_{t-s})\phi_0(y_{t-s-1}) + \ldots + a_{0q}\sum_{s=1}^{d}\gamma^{s-1}\phi_0(y_{t-s})\phi_q(y_{t-s-1}) \\ & + a_{10}\sum_{s=1}^{d}\gamma^{s-1}\phi_1(y_{t-s})\phi_0(y_{t-s-1}) + \ldots + a_{1q}\left(\sum_{s=1}^{d}\gamma^{s-1}\phi_1(y_{t-s})\phi_q(y_{t-s-1}) + \ldots\right. \\ & + a_{q0}\sum_{s=1}^{d}\gamma^{s-1}\phi_q(y_{t-s})\phi_0(y_{t-s-1}) + \ldots + a_{qq}\sum_{s=1}^{d}\gamma^{s-1}\phi_q(y_{t-s})\phi_q(y_{t-s-1}). \quad (9)\end{aligned}$$

There are $(q+1)^2+1$ parameters: $\gamma$ and $a_{ij}$ ($i$, $j=0,\ldots, q$). Note that the number of parameters in model (9) does not depend on $d$, the number of lags included in the model. For example, if $q$ is fixed, then the number of parameters is also fixed, it does not change as $d$ increases. Therefore, we can let $d\to\infty$ as $T\to\infty$ (with $d/T\to0$). Asymptotically, it allows an infinite lag structure without having the curse of dimensionality problem (since $q$ is independent of $d$).

The estimation procedure is as follows. Under $H_0$, $\delta=0$ so that $y_t=\alpha_0+\alpha_1 y_{t-1}+u_t$. Denotes $\tilde{y}_t=y_t-\alpha_0-\alpha_1 y_{t-1}$, then we have $E(\tilde{y}_t^2|\Omega_{t-1})=\sigma_t^2$, or

$$\tilde{y}_t^2 = \sigma_t^2 + v_t, \quad (10)$$

with $E(v_t|\Omega_{t-1})=0$. Using the series approximation of $\sigma_t^2$ given in (9) to replace $\sigma_t^2$ in (10), and replacing $\alpha_0$ and $\alpha_1$ by the least squares estimators of $\alpha_0$ and $\alpha_1$, we can estimate $\gamma$ and $a_{ij}$'s by the nonlinear least squares method. Or we can fix a value of $\gamma$, and estimate $a_{ij}$'s ($i$, $j=0, 1,\ldots q$) by the least squares regression of regressing $\tilde{y}_t^2$ on the series approximating base functions, and then search over $\gamma\in[0,1]$.

In order to ensure that the above procedure leads to a consistent estimate of the $g(\cdot)$ function, we need to let $q\to\infty$ and $q/T\to0$ as $T\to\infty$. The condition $q\to\infty$ ensures that the series approximation error (the asymptotic bias) goes to zero. For example, if one uses power series as the base function, then it is well known that the approximation error of using a $q$th-order polynomial to approximate a smooth function $g(\cdot)$ goes to zero as $q \to \infty$. The condition $q/T \to 0$ ensures that the estimation variance goes to zero as sample increases. See Newey (1997) and de Jong (2002) for more details on the rate of convergence of series estimation.

Let $\hat{\sigma}_t^2$ denote the resulting nonparametric series estimator of $\sigma_t^2$, replacing $\sigma_t^2$ by $\hat{\sigma}_t^2$ in Eq. (4), we get

$$y_t = \alpha_0 + \alpha_1 y_{t-1} + \delta \hat{\sigma}_t^2 + \varepsilon_t, \tag{11}$$

where $\varepsilon_t = u_t + \delta(\sigma_t^2 - \hat{\sigma}_t^2)$. We estimate $\alpha = (\alpha_0, \alpha_1, \delta)'$ by the least squares method of regressing $y_t$ on the vector $(1, y_{t-1}, \hat{\sigma}_t^2)$. Eq. (11) contains a nonparametrically generated regressor $\hat{\sigma}_t^2$. Let $\hat{\alpha} = (\hat{\alpha}_0, \hat{\alpha}_1, \hat{\delta})'$ denote the resulting estimator of $\alpha$. If one ignores the additive structure of $\sigma_t^2$, and estimate $\sigma_t^2 = \mathrm{Var}(y_t|\Omega_{t-1}) \cong \mathrm{Var}(y_t|y_{t-1}, \ldots, y_{t-d})$ by the nonparametric kernel-based method (e.g., Pagan and Ullah, 1988) and the lag number $d$ is finite ($d$ can be an arbitrarily large fixed positive integer), the asymptotic distribution of $\hat{\alpha}$ is derived in Baltagi and Li (2001) who demonstrate that

$$\sqrt{n}(\hat{\alpha} - \alpha) \to N(0, \Sigma) \text{ in distribution} \tag{12}$$

where $\Sigma$ is the asymptotic variance of $\sqrt{n}(\hat{\alpha} - \alpha)$. Baltagi and Li (2001) show that $\Sigma$ consists of two parts. One part corresponds to the case when $\sigma_t^2$ is observable. The other part comes from the fact that $\sigma_t^2$ is in fact not observable but has to be estimated. Newey (1994) has shown that, for a general semiparametric model, the asymptotic variance of the semiparametric estimator (of a finite dimensional parameter) is invariant to the nonparametric estimation techniques used to estimate the model. Thus, if one ignores the additive structure of $\mathrm{var}(y_t|y_{t-1}, \ldots, y_{t-d})$ and estimate it by the nonparametric series method, Newey's result implies that whether one uses the nonparametric series method or kernel method to estimate $\sigma_t^2$, the asymptotic variance of $\hat{\alpha}$ will be the same as given in Baltagi and Li (2001). However, we used the additive structure when estimating $\mathrm{var}(y_t|y_{t-1}, \ldots, y_{t-d})$, intuitively, the asymptotic variance of $\delta$ should be smaller (at least no larger) than the one given in Baltagi and Li (2001) because using the additive structure will lead to more accurate estimation of the conditional variance. The asymptotic distribution of our $\hat{\delta}$ is quite complicated and we leave it as a future research topic. In the empirical application below, we will assume that $d$ is a fixed positive integer and we will use the result of Baltagi and Li (2001) to compute the asymptotic variance of $\hat{\delta}$ which should provide an upper bound for the asymptotic variance of $\hat{\delta}$.

Our nonparametric estimator $\hat{\sigma}_t^2$ based on the least squares estimation of (9) is a consistent estimator for $\sigma_t^2$ under the null hypothesis of $\delta = 0$. However, if the null hypothesis is false, then $\hat{\sigma}_t^2$ is, in general, not a consistent estimator for $\sigma_t^2$ because when $\delta \neq 0$, $y_t$ follows a very complex nonlinear process. Following the approach of Carrasco and Chen (2002), it might be possible to show that a general GARCH-M process (with $\delta \neq 0$) is a stationary $\beta$-mixing process, and one may be able to establish the asymptotic theory for $\hat{\delta}$ when $\delta \neq 0$. We leave this challenging task as a future research topic. Note that the lack of related asymptotic theory does not affect the *null* distribution of our test statistic since under $H_0$, $\delta = 0$, any proxy for $\sigma_t^2$ would suffice to produce a consistent test for $\delta = 0$

provided that it is correlated with $\sigma_t^2$. So the question is whether the nonparametric proxy can lead to a better finite sample power of the test compared with some conventional parametric tests, say a test based on the EGARCH specification of the conditional variance. In Section 3, we report a small-scale Monte Carlo simulations to examine the finite sample performance of the test based on nonparametric estimation of $\sigma_t^2$ and the test based on an EGARCH specification. The simulation reported there show that our proposed test compare favorably with the parametric test based on an EGARCH specification.

Next, we discuss how to select the number of lag $d$ and the order of series approximation $q$, in finite sample applications. For a fixed value of $q$ (for a fixed order of series approximation), the number of parameters to be estimated is fixed and does not depend on $d$. In particular, choosing a large value of $d$ does not lead to over fitting because the number of parameters (that need to be estimated) does not vary as $d$ increases. Therefore, it makes sense to select the value of $d$ that minimizes the sum of squares of residuals. In this manner, we can pick up all the necessary lags to capture the persistent dynamics without overfitting the model. Therefore, we recommend to select the value of $d$ that minimizes the sum of squares of residuals.

The series approximating terms $q$ is selected as follows. Again we use a linear combination of a product base function to approximate $g(y_{t-1}, y_{t-2})$. If we use up to the $q$th univariate base functions for each component of $x_t$, to approximate $g(x_t)=g(y_{t-1}, y_{t-2})$, the number of approximating base function is $k=(q+1)^2(\phi_{l_1}(y_{t-1})\phi_{l_2}(y_{t-2})$ for $0\leq l_1$, $l_2\leq q)$. For a fixed value of $T$ (the sample size), the larger the $k$, the smaller the series approximation error (smaller bias), but the larger the estimation variance. One should choose $k$ optimally in balancing the bias square term and the variance term, i.e., minimizing the mean square error. It is known that the leave-one-out least squares cross-validation method can lead to optimal selection of the smoothing parameter $k$ in the sense that it minimizes the asymptotic estimation mean squares errors (e.g., Li, 1987). We will use the least squares cross-validation method to select $k$ in the empirical applications reported in the next section. One problem with the least squares cross-validation method is that it is computationally quite costly as one has to estimate the model at each observation point. Alternatively, one may select $k$ by minimizing some kind of modified AIC criteria which can be computationally simple. In nonparametric kernel estimations, Hurvich et al. (1998), Li and Racine (2004), and Racine and Li (2004) show that an improved (a modified) AIC criterion performs well in selecting smoothing parameters. We leave the exploration for alternative computationally simple methods as a future research topic.

## 3. Monte Carlo and empirical results

### 3.1. Monte Carlo results

In this section, we conduct a small-scale Monte Carlo simulations to examine the finite sample performance of our proposed test. The true data generating process is that of a GARCH(1,1)-in-mean process:

$$y_t = a_0 + a_1 y_{t-1} + \delta\sigma_{t-1} + u_t$$

$$\sigma_t^2 = \beta_0 + \beta_1 u_{t-1}^2 + \beta_2 \sigma_{t-1}^2$$

$$u_t \sim \epsilon_t \sqrt{\sigma_t^2}$$

where $\epsilon_t$ is i.i.d. $N(0,1)$. Following Engle and Ng (1993), we choose $\beta_0=0.01$, $\beta_1=0.09$, $\beta_2=0.9$ for the GARCH(1,1) process, and we choose $a_0=0.002$, $a_1=-0.004$ for the main equation. $\delta=0$ when $H_0$ is true and $\delta\neq 0$ when the null hypothesis is false. The sample sizes we consider are $n=200$ and $n=400$, and the number of replications are 2000 for all cases.

We use two different estimation methods to estimate the above model. (i) The proposed semiparametric estimation method, and (ii) estimation based on an EGARCH(1,1) model.

First, we examine the size estimation. Under $H_0$ of $\delta=0$, the estimated sizes is reported in Table 1. From Table 1, we observe that both tests are somewhat oversized for sample sizes we consider. However, size distortion for $n=400$ is much smaller than that for $n=200$, and its performance seems to be acceptable for $n=400$.

To examine the power of the proposed test, we set $\delta=(0.5, 1, 2, -0.5, -1, -2)$. In order to get a fair comparison for power, we compute size-adjusted power and the results are given in Table 2. From Table 2, we observe that for most cases, the semiparametric-based test is more powerful than the EGARCH-based test. Also, it is interesting to observe that the power of the semiparametric test is monotonically increasing as the model moves further away from the null model (i.e., as $|\delta|$ increases), while the parametric EGARCH-specification-based test does not have this monotonic power property. For example, for the case of $n=400$, as $\delta$ changes from 1 ($-1$) to 2 ($-2$), the power of the EGARCH-based test decreases. Similar phenomenon is also observed by Juhl and Xiao (2005) who show, via Monte Carlo simulations, that a parametric test for testing a structural break does not exhibit monotonic power while a nonparametric test enjoys the monotonic power property. We also computed the mean (over 2000 replications) values of the test statistics under $H_1$, for example, for $n=400$, and for $\delta=0.5$, 1 and 2, the mean statistics for the semiparametric test are 2.21, 4.36 and 6.21, respectively, and for $\delta=-.5$, $-1$ and $-2$, the mean statistics are $-2.17$, $-4.30$ and $-6.14$, respectively. Hence, as $|\delta|$ increases, the absolute value of the semiparametric test statistic also increases, leading to a monotonic power for our semiparametric test. In contrast, for $n=400$ and for $\delta=0.5$, 1 and 2, the mean statistics for the EGARCH-specification-based test are 2.19, 3.84 and 3.54, respectively (a similar pattern is observed for the $\delta<0$ case), which does not increase monotonically as $\delta$ increases.

Table 1
Estimated sizes

| | EGARCH | | | Semiparametric | | |
|---|---|---|---|---|---|---|
| | 1% | 5% | 10% | 1% | 5% | 10% |
| $n=200$ | 1.72 | 6.85 | 13.1 | 2.16 | 7.23 | 13.8 |
| $n=400$ | 1.65 | 6.20 | 11.6 | 1.86 | 6.43 | 12.3 |

Table 2
Estimated power

| | | $\delta=0.5$ | $\delta=1$ | $\delta=2$ | $\delta=-.5$ | $\delta=-1$ | $\delta=-2$ |
|---|---|---|---|---|---|---|---|
| *n=200* | | | | | | | |
| EGARCH | 1% | 0.198 | 0.341 | 0.393 | 0.074 | 0.011 | 0.169 |
| | 5% | 0.327 | 0.447 | 0.495 | 0.303 | 0.443 | 0.425 |
| | 10% | 0.376 | 0.486 | 0.529 | 0.384 | 0.514 | 0.537 |
| | 1% | 0.156 | 0.383 | 0.680 | 0.109 | 0.377 | 0.662 |
| Semiparametric | 5% | 0.272 | 0.504 | 0.775 | 0.280 | 0.485 | 0.752 |
| | 10% | 0.323 | 0.567 | 0.811 | 0.343 | 0.544 | 0.801 |
| *n=400* | | | | | | | |
| EGARCH | 1% | 0.530 | 0.699 | 0.668 | 0.017 | 0.067 | 0.30 |
| | 5% | 0.674 | 0.783 | 0.747 | 0.680 | 0.769 | 0.706 |
| | 10% | 0.734 | 0.811 | 0.778 | 0.739 | 0.801 | 0.740 |
| | 1% | 0.340 | 0.756 | 0.952 | 0.345 | 0.748 | 0.934 |
| Semiparametric | 5% | 0.468 | 0.832 | 0.973 | 0.435 | 0.820 | 0.955 |
| | 10% | 0.557 | 0.861 | 0.976 | 0.517 | 0.854 | 0.964 |

Another advantage of the semiparametric-based test is its computational simplicity, because the nonparametric series estimator for $\sigma_t^2$ only involves a simply least square regression, the semiparametric-based test is much faster to compute than the EGARCH-based test. EGARCH-based test requires that one estimates a highly nonlinear (misspecified) EGARCH model and it can be computationally quite costly. Sometimes in the simulation iterations, the EGARCH model does not lead to a convergent estimation result, and we have to throw away the non-convergence results. The unstable EGARCH performance is also reported by Engle and Ng (1993, p.1771).

Although one cannot draw general conclusions regarding the relative performance of the semiparametric-based test and the EGARCH-based test from a small-scale Monte Carlo simulations, the simulation results reported here at least show that the proposed semiparametric test is a worthy alternative to the popular EGARCH-based test in testing the GARCH-in-mean effect.

### 3.2. Empirical results

In this subsection, we report the result of an empirical application. Following Lo and Mackinlay (1988), we use weekly data from January 1980 to December 2001 for the 12 largest stock markets in the world in terms of market capitalization. These indexes include the United States (US), Canada (CA), Japan (JP), Australia (AU), Hong Kong (HK), Singapore (SG), the United Kingdom (UK), Germany (GM), France (FR), Italy (IT), Netherlands (NT), and Switzerland (SW). All the data are in local currency terms and obtained from Morgan Stanley Capital International (MSCI). The MSCI indexes represents approximately 60% of the aggregate market value of the stock markets in these countries. All the indexes are market-value weighted.

The stock returns used in this study are defined as the first difference of the logarithm of stock index prices including dividends ($y_t=\ln(p_t/p_{t-1})$). Table 3 provides some basic statistics. Because risk-free interest rate data are not available to all the markets under

Table 3
Summary statistics

| Country | Mean | Standard deviation | Skewness | Kurtosis |
|---|---|---|---|---|
| United States | 0.002*** | 0.023 | −1.737 | 20.584*** |
| Canada | 0.001* | 0.024 | −0.936*** | 9.70*** |
| Japan | 0.001 | 0.026 | −0.333*** | 4.318*** |
| Australia | 0.002** | 0.027 | −1.781*** | 17.986*** |
| Hong Kong | 0.002*** | 0.045 | −0.98*** | 6.866*** |
| Singapore | 0.001 | 0.034 | −2.016*** | 23.272*** |
| United Kingdom | 0.002*** | 0.023 | −1.597*** | 16.370*** |
| Germany | 0.002** | 0.027 | −1.224*** | 7.196*** |
| France | 0.002*** | 0.027 | −0.912*** | 5.115*** |
| Italy | 0.003*** | 0.034 | −0.339*** | 2.522*** |
| Netherlands | 0.002*** | 0.024 | −1.457*** | 12.915*** |
| Switzerland | 0.002*** | 0.023 | −1.370*** | 8.355*** |

* Denotes significance at the 10% level.
** Denotes significance at the 5% level.
*** Denotes significance at the 1% level.

consideration during the whole sample period, stock market volatility is measured based on stock returns instead of excess stock returns (which is equal to stock returns minus the risk-free interest rate). Many researchers (Baillie and DeGennarro, 1990; Nelson, 1991; Choudhry, 1996; Lee et al., 2001) argue that such a practice produces little difference in estimation and inference in this line of research. In particular, Baillie and DeGennarro (1990) compare two models based on stock returns and excess stock returns, with the former ignoring and the latter allowing for the risk-free interest rate. They find that the results of the two models are very similar (p. 207). Similarly, Nelson (1991) fits two EGARCH-M models with one including and the other excluding both dividends and the risk-free interest rate. He concludes that there is “virtually no difference in either the estimated parameters or the fitted variances” (p. 356).

The preliminary analysis is conducted on AR($k$)-EGARCH ($p$,$q$)-M specifications for $p$,$q$ = 1, 2, 3, 4. An AR(1)-EGARCH (1,1)-M process generally provides a good approximation of the data generating process for stock returns under consideration, with the exception that AR(2)-EGARCH (1,1)-M appears most appropriate for Hong Kong. To conserve space, only the results for the two most relevant parameters, $\delta$ and $\alpha_1\phi$, are reported in Table 4 (the complete results are available upon request). The time-varying pattern of stock market volatility is confirmed because the coefficients of GARCH effects ($\alpha_1\gamma$ and $\beta_1$ in Eq. (3), not reported here) are significant at any conventional significance level in all markets. As discussed previously, in this study it is particularly important to check the sign and significance of the coefficient estimate $\delta$. Table 4 shows that the coefficient $\delta$ is positive for all markets with the exception of Japan and Singapore, where there exists a negative but insignificant estimate of the parameter $\delta$ for these two countries. Furthermore, none of the $\delta$ parameter estimates is statistically significant at any conventional significance level, with the exception of the UK at the 10% significance level. This may be interpreted as lack of evidence for a significant relationship between expected return and volatility in international stock markets. The results are generally consistent with many previous works, including Baillie and DeGennarro (1990), Nelson

Table 4
EGARCH (1,1)-M estimation results

| Coefficient ($t$-ratio) | US | CA | JP | AU | HK | SG |
|---|---|---|---|---|---|---|
| $\delta$ | 2.30 (0.87) | 0.68 (0.33) | −0.30 (−0.14) | 2.82 (1.08) | 0.11 (0.10) | −0.07 (−0.05) |
| $\alpha_1\phi$ | 0.12 (0.68) | 0.22 (1.56) | 0.28** (2.76) | 0.25 (1.35) | 0.49*** (3.99) | 0.29** (2.25) |
| **Coefficient ($t$-ratio)** | **UK** | **GM** | **FR** | **IT** | **NT** | **SW** |
| $\delta$ | 7.73* (1.78) | 0.32 (0.17) | 0.96 (0.37) | 2.28 (1.23) | 1.66 (0.71) | 0.59 (0.25) |
| $\alpha_1\phi$ | 0.58 (1.56) | 0.15 (1.52) | 0.38** (2.62) | −0.07 (−0.64) | 0.17 (1.41) | 0.22** (2.04) |

See the note in Table 3.

(1991), Theodossiou and Lee (1995), Choudhry (1996), and Lee et al. (2001). In addition, the parameter estimate of $\alpha_1\phi$ is significant at the 5% significance level for five of the markets (Japan, Hong Kong, Singapore, France and Switzerland). This suggests that there is an asymmetric response of conditional variance to negative and positive stock return innovations in these markets. However, the EGARCH-M(1,1) estimation results should be interpreted with caution, because the following semiparametric analysis shows that (at least part of) EGARCH-M(1,1) model may be misspecified and its estimation results may not be reliable.

Next, we conduct the semiparametric analysis. We estimate model (8) and we use B-spline (Schumaker, 1981) as the approximating base function. The estimation results for the whole sample period are reported in Table 5. From Table 5, we observe that the estimated coefficients of $\delta$ are negative in 11 out of 12 markets. Moreover, the parameter estimate is significantly different from zero at 5% level for four markets (Australia, France, Italy, Netherlands) and in two more markets (US and Switzerland) at the 10% significant level. This provides some evidence of a negative relationship between stock market return and volatility in international stock markets. Such a finding implies that high/low volatility precedes low/high returns.

It has been argued that the 1987 international stock market crash may have had a substantial impact on international stock market behavior. In particular, Choudhry (1996) provides evidence of changes in the ARCH parameters, the risk premium and volatility

Table 5
Semiparametric GARCH-M estimation results (01/80–12/01)

| Coefficient ($t$-ratio) | US | CA | JP | AU | HK | SG |
|---|---|---|---|---|---|---|
| $\delta$ ($t$-ratio) | −11.99* (−1.88) | −1.83 (−0.25) | −11.86 (−1.06) | −9.89** (−2.32) | −0.04 (−0.03) | −4.86 (−0.96) |
| **Coefficient ($t$-ratio)** | **UK** | **GM** | **FR** | **IT** | **NT** | **SW** |
| $\delta$ ($t$-ratio) | 3.37 (0.40) | −3.07 (−0.70) | −7.29** (−2.17) | −10.24** (−1.99) | −14.32** (−2.23) | −6.96* (−1.81) |

See the note in Table 3.

Table 6
Semiparametric GARCH-M estimation results (01/80–09/87)

| Coefficient ($t$-ratio) | US | CA | JP | AU | HK | SG |
|---|---|---|---|---|---|---|
| $\delta$ ($t$-ratio) | −7.00<br>(−0.17) | 4.00<br>(0.18) | 12.19<br>(0.82) | 9.23<br>(0.43) | −1.09<br>(−0.23) | 1.22<br>(0.19) |
| **Coefficient ($t$-ratio)** | **UK** | **GM** | **FR** | **IT** | **NT** | **SW** |
| $\delta$ ($t$-ratio) | −25.52**<br>(−2.29) | 5.40<br>(0.24) | −10.47***<br>(−3.11) | −12.41**<br>(−2.77) | −21.97<br>(−1.15) | −18.39<br>(−0.83) |

See the note in Table 3.

persistence before and after the 1987 crash in several developing markets. Ignoring such a potential structural change due to the 1987 crash may bias the estimates based on the whole sample period. To address the impact of the 1987 crash, we further divide the data into two subperiods: the pre-crash period (January 1980–September 1987) and the post-crash period (November 1987–December 2001). The observations for October 1987 are excluded. The semiparametric analysis is conducted on the two periods and the results are reported in Tables 6 and 7. As shown in Table 6, the coefficient estimates of $\delta$ are negative in 7 out of 12 markets and significant for 3 markets (Italy at the 1% level, UK and France at the 5% level). It is also interesting to note that the coefficient estimates of $\delta$ are positive in 5 out of 12 markets in the pre-crash period, although none of them are statistically significant at any conventional significance levels.

In contrast, the results for the post-crash period show that coefficient estimates of $\delta$ are negative for 11 out of 12 markets (with the exception of Canada) and are significant for 7 markets. These 7 markets include Australia (1% level), Hong Kong, UK, Germany, France, Switzerland (5% level), and Singapore (10% level). Also noteworthy, comparing the results in Tables 6 and 7, the coefficient estimates of $\delta$ are larger during the post-crash period than during the pre-crash period for some markets, including Australia, Hong Kong, Singapore, Germany, and France. In contrast, the coefficient estimates of $\delta$ are somewhat smaller in the post-crash period than in the pre-crash period for several European markets (UK, Italy and Switzerland). The evidence shows that risk premiums

Table 7
Semiparametric GARCH-M estimation results (11/87–12/01)

| Coefficient ($t$-ratio) | US | CA | JP | AU | HK | SG |
|---|---|---|---|---|---|---|
| $\delta$ ($t$-ratio) | −20.23<br>(−1.41) | 5.41<br>(0.42) | −4.92<br>(−0.53) | −20.04***<br>(−3.82) | −7.06**<br>(−2.68) | −8.75*<br>(−1.74) |
| **Coefficient ($t$-ratio)** | **UK** | **GM** | **FR** | **IT** | **NT** | **SW** |
| $\delta$ ($t$-ratio) | −19.19**<br>(−2.63) | −18.91**<br>(−2.13) | −17.32**<br>(−2.27) | −5.19<br>(−1.08) | −15.4<br>(−1.30) | −16.67**<br>(−1.97) |

See the note in Table 3.

have increased after the 1987 crash in some stock markets, though not for some other markets. Consistent with Choudhry (1996), the fact that the change in risk premium is not uniform and depends on the individual market, suggests that factors other than the 1987 crash may also be responsible for the change.

## 4. Conclusions

This study examines the relationship between expected stock returns and volatility in the 12 largest international stock markets. We show that the estimated relationships between return and volatility are sensitive to the way volatilities are estimated. When parametric EGARCH-M models are estimated, we obtain results that are similar to previous findings. Ten out of 12 markets have positive but statistically insignificant relationship (with the only possible exception at the 10% significance level). On the other hand, using a flexible semiparametric specification of conditional variance, we show that negative relationships between returns and volatility prevail in most of these markets. Moreover, the negative relationships are significant in six markets based on the whole sample period and seven markets after the 1987 international stock market crash.

Given the fact that the semiparametric specification is more robust than a parametric conditional variance specification, the result of this study lends some support to the claim that stock return volatility is negatively correlated with stock returns (Black, 1976; Cox and Ross, 1976; Bekaert and Wu, 2000). One explanation of such a negative relationship is based on leverage (Black, 1976). A drop in the value of the firm's stock (negative return) increases financial leverage used by the firm and its debt-to-equity ratio, which makes the stock riskier and increases its volatility. Another explanation based on volatility feedback (Pindyck, 1984; French et al., 1987) suggests that if volatility is priced, an anticipated increase in volatility raises the required return on equity, leading to an immediate stock price decline (negative return). More formally, Whitelaw (2000) theoretically shows that a general equilibrium exchange economy characterized by a regime-switching consumption process generates a negative unconditional relationship between expected returns and volatility at the market level. Our results are also consistent with the empirical findings of Glosten et al. (1993) and Whitelaw (2000). However, it contradicts the prediction of a positive relation made by many asset pricing models (e.g., Sharpe, 1964; Linter, 1965; Mossin, 1966; Merton, 1973) and the empirical finding of an insignificant relationship consistently reported in the previous literature (Baillie and DeGennarro, 1990; Nelson, 1991; Theodossiou and Lee, 1995; Choudhry, 1996; Lee et al., 2001).

The findings of this study have some important implications. For example, as pointed out in Bekaert and Wu (2000, p. 2), the negative relationship between market volatility and expected market return immediately implies that the time-varying risk premium theory cannot be valid to explain the stock market behavior. Further investigation may be conducted on whether such a negative relationship is time-varying, as suggested in the model of Whitelaw (2000), and also prevails in emerging stock markets. It is also of interest to examine whether different explanations exist for such a negative relationship in international stock markets following the work of Bekaert and Wu (2000). For example, although both explanations based on leverage effects and volatility feedback may explain the negative (contemporaneous) relationship between stock returns and volatility, they carry different implications for causality between the returns and volatility (Bekaert and Wu, 2000). Future research may also employ the semiparametric specification of conditional variance used in this study to explore other topics concerning the relationship between mean and conditional variance, such as that between inflation rate and inflation volatility.

## Acknowledgements

We are grateful to a referee and a co-editor for their insightful comments that have greatly improved the paper. We would also like to thank Xiaohong Chen and Thomas Fetherston for their helpful discussions and comments, and David Drukker, Tae-Hwy Lee, Insik Min, and Byeongseon Seo for their help in programming. An earlier version of the paper was presented at 2003 Financial Management Association annual meetings. Li's research is partially supported by the Private Enterprise Research Center, Texas A&M University.

## References

Baillie, R.T., DeGennarro, R.P., 1990. Stock returns and volatility. Journal of Financial and Quantitative Analysis 25, 203–214.

Baltagi, B.H., Li, Q., 2001. Estimation of econometric models with nonparametrically specified risk terms: with applications to foreign exchange markets. Econometric Reviews 20, 445–460.

Bekaert, G., Wu, G., 2000. Asymmetric volatility and risk in equity markets. Review of Financial Studies 13, 1–42.

Black, F., 1976. Studies of stock price volatility changes. Proceedings of the 1976 Meeting of Business and Economics Statistics Section of the American Statistical Association, vol. 27, pp. 399–418.

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

Bollerslev, T., Chou, R.Y., Kroner, K.F., 1992. ARCH modeling in finance. Journal of Econometrics 52, 5–59.

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

Choudhry, T., 1996. Stock market volatility and the crash of 1987: evidence from six emerging markets. Journal of International Money and Finance 15, 969–981.

Cox, J., Ross, S., 1976. The valuation of options for alternative stochastic process. Journal of Financial Economics 3, 145–166.

de Jong, R., 2002. Convergence rates and asymptotic normality for series estimators: uniform convergence rates. Journal of Econometrics 111, 1–9.

Engle, R.F., 1982. Autoregressive conditional heteroscedasticity with estimates of the variance of U.K. inflation. Econometrica 50, 987–1008.

Engle, R.F., Ng, V.K., 1993. Measuring and testing the impact of news on volatility. Journal of Finance 48, 1749–1778.

Engle, R.F., Lillian, D.M., Robins, R.P., 1987. Estimating time-varying risk premium in the term structure: ARCH-M Model. Econometrica 55, 391–407.

French, K.R., Schwert, G.W., Stambaugh, R.F., 1987. Expected stock returns and volatility. Journal of Financial Economics 19, 3–30.

Glosten, L.R., Jagannathan, R., Runkle, D.E., 1993. On the relation between the expected value and the volatility of nominal excess return on stocks. Journal of Finance 48, 1779–1801.

Hurvich, C.M., Simonoff, J.S., Tsai, C.L., 1998. Smoothing parameter selection in non-parametric regression using an improved Akaike information criterion. Journal of the Royal Statistical Society. Series B 60, 271–293.

Jones, C., Lamont, O., Lumsdaine, R., 1998. Macroeconomic news and bond market volatility. Journal of Financial Economics 47, 315–337.

Juhl, T., Xiao, Z., 2005. A nonparametric test for changing trends. Journal of Econometrics 127, 179–199.

Lee, C.F., Chen, G., Rui, O., 2001. Stock returns and volatility on China's stock markets. Journal of Financial Research 26, 523–543.

Li, K.C., 1987. Asymptotic optimality for $C_p$, $C_L$, cross-validation, and generalized cross-validation: discrete index. Annals of Statistics 15, 975–985.

Li, Q., 2000. Efficient estimation of additive partially linear models. International Economic Review 41, 1073–1092.

Li, Q., Racine, J., 2004. Cross-validated local linear nonparametric regression. Statistica Sinica 14, 485–512.

Linter, J., 1965. The valuation of risky assets and the selection of risky investments in stock portfolios and capital budgets. Review of Economics and Statistics 47, 13–37.

Lo, A.W., Mackinlay, A.C., 1988. Stock market prices do not follow random walks: evidence from a simple specification test. Review of Financial Studies 1, 41–66.

Merton, R.C., 1973. An intertemporal capital asset pricing model. Econometrica 41, 867–887.

Merton, R.C., 1980. On estimating the expected return on the market: an exploratory investigation. Journal of Financial Economics 8, 323–361.

Mossin, J., 1966. Equilibrium in a capital asset market. Econometrica 41, 768–783.

Nelson, D., 1991. Conditional heteroscedasticity in asset returns: a new approach. Econometrica 59, 347–370.

Newey, W., 1994. Asymptotic variance of semiparametric estimators. Econometrica 62, 1349–1382.

Newey, W., 1997. Convergence rates and asymptotic normality for series estimators. Journal of Econometrics 79, 147–168.

Pagan, A.R., Ullah, A., 1988. The econometric analysis of models with risk terms. Journal of Applied Econometrics 3, 87–105.

Pindyck, R., 1984. Risk, inflation and the stock market. American Economic Review 74, 335–351.

Poterba, J., Summers, L., 1986. The persistence of volatility and stock market fluctuations. American Economic Review 76, 1142–1151.

Racine, J., Li, Q., 2004. Nonparametric estimation of regressions with both categorical and continuous data. Journal of Econometrics 119, 99–130.

Schumaker, L.L., 1981. Spline Functions: Basic Theory. Wiley, New York.

Sharpe, W.F., 1964. Capital asset prices: a theory of market equilibrium under conditions of market risk. Journal of Finance 19, 425–442.

Theodossiou, P., Lee, U., 1995. Relationship between volatility and expected returns across international stock markets. Journal of Business Finance and Accounting 22, 289–300.

Whitelaw, R., 2000. Stock market risk and return: an empirical equilibrium approach. Review of Financial Studies 13, 521–547.

Yang, L., 2002. Direct estimation in an additive model when the components are proportional. Statistica Sinica 12, 801–821.