No. 1501 [EN]

IMI Working Paper

Risk Perception and Equity ReturnsEvidence from the SPX and VIX

Gang Jianhua and Li Xiang

INTERNATIONAL MONETARY INSTITUTE

IMI Working Paper No. 1501 [EN]

Risk Perception and Equity Returns: Evidence from the SPX and VIX

***By* GANG JIANHUA[*] and LI XIANG[**]**

January 2015

Abstract

We use the semi-nonparametric (SNP) model to study the relationship between the innovation of the Volatility Index (VIX) and the expected S&P 500 Index (SPX) returns. We estimate the one-step-ahead contemporaneous relation subject to leverage GARCH effect. Results agree with a body of newly established literature arguing non-linearity, and asymmetries. In addition, the risk-return behaviour depends on the signs as well as magnitudes of the perceived risk. We conclude that influence of fear or exuberance on the conditional market return is non-monotonic and hump-shaped. Very deep fear does not necessarily mean huge losses, instead, the loss may not be as bad as fears of normal levels. Results pass the robustness tests.

JEL Classification: G17, C01, C02, C14, C15

Keywords: Conditional Joint Density, GARCH Models, Semi-nonparametric, SPX, VIX

[*] Research fellow of IMI, School of Finance, Renmin University of China. Corresponding author, email: jhgang@ruc.edu.cn.

[**] Department of Finance, School of Economics, Shanghai University, China

1. Introduction

The relationship between risk and return is crucial in the finance world. Whether or not asset prices and hence expected return should reflect investors' willingness to bear risk attracts tremendous attention and there has been a huge body of literature contributed to this topic. It is important in asset pricing, hedging, derivative pricing and risk management. But there is an ongoing debate on the nature of the risk-return relationship. Standard finance theory, e.g., Merton's (1973, 1980) Intertemporal Capital Asset Pricing Model (ICAPM) implies that the cross-section of stock returns should be affected by systematic volatility and this relation turns out to be linear and positive. Similar results are also indicated by Ghysels *et al.* (2005) using weighted rolling sample windows in the variance measurements. For a survey of these and related studies, see Lettau and Ludvigson (2009).

Motivated by Merton's work, many other studies, e.g. Campbell (1987) and Glosten *et al.* (1993), find evidence of a negative relation using US datasets. Brandt and Kang (2004) model the conditional mean and volatility of stock returns as a latent VAR process without relying on exogenous predictors. They also find a negative correlation between the innovations to the conditional moments leading to pronounced counter-cyclical variation in the Sharpe ratio. Harvey (2001), on the other hand, uses exogenous predictors and concludes that the correlation between the moments generally depends on the model and the information set used in conditional moments.

Recent work in asset pricing on the question of volatility innovations studies the cross- sectional risk premia induced by covariance between volatility changes and stock returns and finds negative premia, e.g., Ang *et al.* (2006). A sensible explanation resides in a scenario where a volatile time of period (such as a recession) comes, co-varing stocks (in term of volatility) pay off and hence are required for less premia.

While Ang *et al.* show that the cross-sectional pricing of sensitivity to innovations in implied market volatility is robust, it does not account for the asymmetric return responses to positive and negative changes in expected systematic volatility, as found by Dennis *et al.* (2006). Thus, the relation between sensitivity to market volatility innovations and returns may not yet be fully understood. We hence propose a new fundamental work to reveal the relation between market returns and expected volatility changes while allowing for asymmetric volatility responses on a time-series dimension. We find that the asymmetric volatility phenomenon is an important element in the return process. The coexistence of negative and positive relationship is unearthed and is subject to non-monotonic trend.

Our paper implies the possibility that the expected risk-return trade-off documented in literature might be subject to misspecified models and most importantly, the sign as well as magnitude of market sentiment changes. As clearly stated in Christensen *et al.* (2010) that restrictive linear conditions are unreasonable and inconsistent to empirical fact, we therefore use semi-nonparametric (SNP) models to nest explicit expected risk metrics (proxied by the innovation in the Volatility Index (VIX)) and the total S&P 500 Index (SPX) returns. The choice of the risk measurement is justified by Chen (2003) who demonstrates that changes in the expectation of future market volatility are a source of risk. Our bivariate system is led by a flexible GARCH term which allows leverage effect or complex non-parametric characteristics.[1]The SNP model in our paper enables us to exploit potential asymmetries and non-monotonic nonlinearity which was pioneered by Pagan and Hong (1991). They argue that the risk premium, μ_t, and the conditional variance, σ_t^2, are highly non-linear functions of the past whose form is not captured by standard parametric GARCH-M models.

Recent scholars also generalize GARCH-family models to explain the non-linear risk-return relationship. Following the conjecture of inadequacy in non-linearity by Das and Sarkar (2000) using an ARCH-in-Nonlinear-Mean (ARCH-NM) model,[2] Linton and Perron (2003) step forward by suggesting an algorithm to allow a satisfactory non-linear property. Linton and Perron's (2003) model is semiparametric in the sense that it is parametric in the conditional variance function while at the same time allowing for an arbitrary functional form to describe the relationship between risk and returns at market level. An application of their theory to the exponential GARCH-M model uncovers a non-linear and non-monotonic relationship. Similar work but revised estimators in the conditional variance function based on the GARCH-M model also finds its way in Conrad and Mammen (2008), and a simpler version (without testing the parametric specification of the mean function) of the model is derived by Christensen *et al.* (2010).

We also propose to use a semi-nonparametric tool to study the risk-return relation, but in a quite different way.

First, we do not add in any GARCH-family model unconditionally, and even when the model selecting procedure does imply the GARCH effect, it is still subject to diagnostic tests and potential expansion in the conditional mean term. More specifically, the conditional one-step ahead joint density starts from a standard normal distribution and is optimally expanded using Hermite polynomials and estimated by maximum likelihood (ML) method. This SNP procedure has been developed recently by Gallant and Tauchen (2006) to nest a generalized asymmetric VAR-GARCH model as its leading density together with the leverage effect. The advantage of adopting such a strategy lies in the fact that a specific predetermined functional form of the risk-return correlation is sometimes redundant and an adequate but parsimonious model which nests necessary observed information might be a better choice.

Second, on the aspect of choosing the metrics of risk, we follow the work by Chen (2003), who demonstrates that changes in the expectation of future market volatility are a source of risk. This prediction is later on verified by Ang *et al.* (2006), in which they find that sensitivities to changes in implied market volatility have a cross-sectional effect on firm-level returns. In our paper, we choose the daily innovation in the implied volatility (CBOE VIX) as a proxy of change in the expected market volatility on the time series dimension.

---

[1] Models are selected based on Baysian information criteria and author-designed diagnostics. For detailed procedures of Hermite expansion whilst alowing different leading terms, readers are advised to refer to the work by Gallant and Tauchen (2006).

[2] Das and Sarkar (2000) defines the risk premium as a Box-Cox power transformation of the conditional variance.

We find insufficient literature that contributes to direct estimations of the relation between the innovations of VIX and market returns. A critical question that needs to be answered is to what extent and how the sentiment shifts (fear or exuberance) change expected market returns. In other words, should Chen's (2003) argument hold, risk perception would be reflected in market returns through shifts in sentiment. Significant literature, however, shows the asymmetry where positive returns are associated with smaller changes in implied volatilities than negative returns of the same magnitude. Specifically, Dennis *et al.* (2006) examine the relation between stock returns and *VIX* allowing for stock returns to react asymmetrically to volatility shocks. Their goal, however, is to determine if the asymmetric volatility phenomenon stems from systematic or idiosyncratic effects but not directly test for a risk-return relation.

However, we conclude that the contemporaneous risk-return behaviour depends not only on the sign of risk metrics (sentiment shifts), but also on the magnitudes of the change. In other words, fear or exuberance (extreme innovation of VIX) does correlate to conditional return, but the correlation is non-monotonic and hump-shaped. On the one hand, very deep fear does not necessarily mean huge losses, instead, the loss may not be as bad as fears of normal levels, while on the other, exuberance does not correlated to big returns.

Finally, the present paper is different from a similar work by Christensen and Nielsen (2007), in which a negative but monotonic relation is found and hence is consistent with the cross-sectional case by considering aggregated returns and volatility innovations in both realized and implied volatility. Our results partly agree on their conclusion. We argue that the negativity is subject to the magnitude of innovations of the expected volatility and that a positive relation also exists given dramatic changes of sentiment, so that a non-monotonic relationship seems to be the full story. Empirical threshold is also given.

Our method is also different from an SNP estimation by Linton and Perron (2003), who find a non-monotonic, hump-shaped risk premium-variance relation using a non-parametric EGARCH model. We, however, do not estimate the conditional variance of returns, but instead examine the marginal density and moments of conditional returns against innovations of the implied volatility. Therefore, our metrics of risk is proxied by the fear gauge or change of sentiment/expectation.

Nonetheless, our framework is not intended to be predictive or keen to solve some causal link of anything, but merely to directly illustrate the conditional contemporaneous risk-return relation. And the relation is also allowed to change conditional on the contemporaneous market condition and on how much the VIX increases or decreases. The paper can be thought as an extension of Christensen *et al.* (2010) on the dimension of nesting not only the signs but also magnitudes of innovations as the information set. Our results partly agree with Linton and Perron (2003), in which we also show a hump-shaped pattern of the total returns relative to risk perceptions but it turns out to be more smooth. And the robustness check indicates our model works reasonably well.

This paper is organized as follows: Section II explains the dataset and adjustment; Section III exhibits the estimation strategy and model specifications; Section IV and V discuss the empirical models and results; Section VI shows the robustness tests; and finally, Section VII delivers the conclusion.

2. Data Description

2.1 Risk metrics: innovation of VIX

Lots of study on volatility dynamics relies on volatility estimated from historical data (namely, the realized volatility), but statistical estimation may produce sampling and model specification errors. Its lack of prediction power also limits the connection to real trading. We tend to care more about the expectation of the future realized volatility, especially the downside volatility, namely the risk. There hence exist incentives to use some *ex ante* measurement which directly targets an estimation of such risks.

The use of such a proxy of risk as changes in the implied volatility is justified thanks to a pioneer work by Chen (2003) which demonstrates that changes in the expectation of future market volatility are a source of risk. This prediction is further verified by Ang *et al.* (2006), where they find that sensitivities to changes in implied market volatility have a cross-sectional effect on firm-level returns. There is also a body of literature that uses the innovation of VIX as a measurement of risk, for example, a recent work by Dennis *et al.* (2006) examines the relation between stock returns and *VIX* allowing for stock returns to react asymmetrically to volatility shocks. Their goal, however, is to determine if the asymmetric volatility phenomenon stems from systematic or idiosyncratic effects while not directly testing for a risk-return relation. A paper by Christensen and Nielsen (2007) also uses the same innovation of VIX. Their conclusion suggests a monotonic negative relation by considering aggregated returns and volatility innovations in both realized and implied volatility.

In the present paper, we choose the innovation of the CBOE VIX as a risk metrics. Specifically, we calculate the daily innovation as a proportion of the previous-day VIX levels (namely the percentage change of VIX, denoted by %VIX) to be a proxy of change in the expected market volatility. Financial literature finds conclusive evidence of the relation between the implied volatility and the realized volatility. Such literature can date back to the work by Feinstein (1989) which demonstrates that implied volatility from ATM and near expiration option provides the closest approximation to the average volatility over the life of the option. Poon and Granger (2003) review 93 papers regarding the forecasting performance of various volatility models (historical, stochastic and implied volatility). Their key conclusion is that the performance of option implied standard deviation outstands that of the alternative methods.

More recent articles dealing with the US equity markets suggest that, in general, implied volatility is a superior predictor of future volatility. Three representative articles are by Giot, Jiang and Tian, and Core and Miller. Giot (2005) evaluates the information content of VIX and VXN as the predictors of the realized volatility and finds meaningful forecasting results. Similarly, Corrado and Miller (2005) report that the CBOE implied volatility indices (VIX, VXO and VXN) act as an outperforming estimators of the future realized volatility compared to the forecast from the historical volatility. Jiang and Tian (2005) investigate the characteristics of the model-free approximation. They discover that the model-free implied variance, represented by the new VIX, subsumes all information contained in the Black-Scholes implied volatility.

As for the performance of the implied volatility outside the US, empirical findings are mainly consistent with those dealing with the implied volatility in the US. Although Dowling and Muthuswamy (2005) report that their Australian implied volatility index is a poor predictor of the future realized volatility, Bluhm and Yu (2001), Skiadopoulos (2004), Nishina *et al.* (2006) and Areal (2008) all find their implied volatility indices are superior estimators of the future standard deviation.

The CBOE VIX, introduced in 1993, quickly became the benchmark for stock market volatility. It is widely followed and cited in hundreds of new articles in the *Wall Street Journal*, *Barron's* and other leading financial publications. The VIX measures market expectations of near term volatility conveyed by stock index option prices. Since the VIX signifies financial turmoil, it is commonly referred to as the 'investor fear gauge' by market practitioners and academics as well. The VIX is based on weighted averages of Black-Scholes put and call implied volatility and it is designed to be a forward-looking measure of volatility which predicts the volatility of the following 30 calendar days (22 trading days for estimation purpose) for the S&P 500 index (CBOE ticker: SPX). Since the depth of the index option market ensures that transacted prices are representative of the aggregate consensus, the VIX index is often regarded as market participants' best guess of the volatility associated with the SPX index.

Since the VIX is not a statistical estimated volatility, it does not induce traditional estimation errors. Moreover, model misspecification error is relatively small compared to statistical volatility metrics, insofar as the underlying option-pricing model based on the work of Black and Scholes (1973) and Merton (1973) is robust and widely used in the market.

2.2 Dataset

Return dataset is the total daily returns of the S&P500 index (CBOE ticker: SPX) and the daily innovation of CBOE implied volatility (CBOE ticker: VIX) in terms of percentage change (denoted by %VIX).

The data period is between 2 January 1990 and 29 December 2006, where 2 January 1990 is the introduction date of the VIX when it was officially traded in the market. The VIX is based on forming portfolios of European options and measures the market's expectation of te next 30 calendar days (22 trading days) forward S&P500 index volatility implicit in the index option prices.

Plots of the SPX index and VIX are shown in Figure 1, in which the dotted line is the raw *VIX* level and the solid line is the SPX index. Intuitively, when SPX goes up, it tends to calm the market, so that the *VIX* displays a decline and vice versa, but this is not always true. Figure 2 shows a mean-reverting process for the return series, but outliers seem to play a role.[3]

2.3 Seasonal effects

To adjust seasonalities, we use dummy-variable models to identify seasonal effects. The following dummies are considered:

- Day-of-the-week dummies (one for each day, Tuesday through Friday);

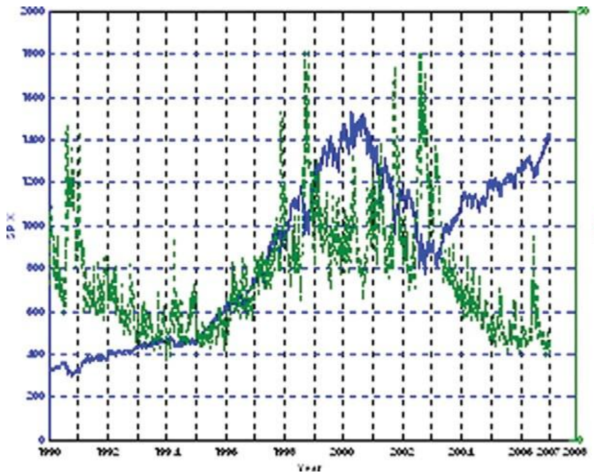

Figure 1. VIX level and SPX level.

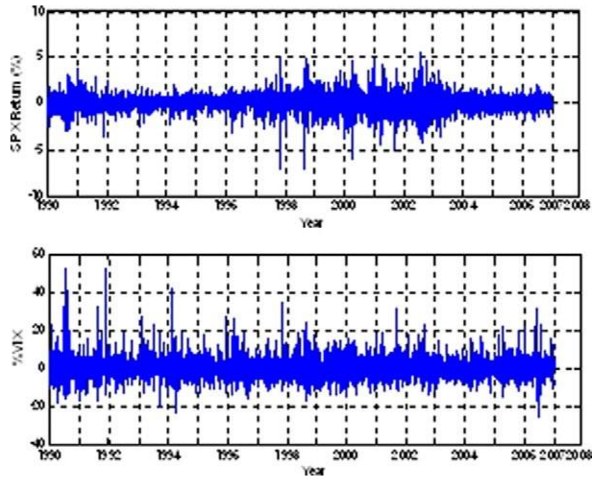

Figure 2. SPX return and %*VIX*.

• Dummies for each number of non-trading days preceding the current trading day(one non-trading day: 39, two non-trading days: 781, and three non-trading days: 105);

• Dummies for months of March, April, May, June, July, August, September, October, and November;

• Dummies for each week of December and January;

---

[3] Adjustment is introduced separately using a spline function to the original dataset to eliminate the outlier effect as can be found in later sections. Similar adjustment is also used by Gallant *et al.* (1992).

• t, $t2$ trend variables (not included in the mean regressions for the price change).

We first regress the SPX total returns and the $\%VIX$ on these dummies respectively to do a mean adjustment (location adjustment) and then regress the residuals from previous models on these dummies to do a variance adjustment.

More specifically, the location adjustment is:

$$w = x'\beta + u \qquad \text{(mean equation)}$$

where w is the series to be adjusted (dataset) and x contains the adjustment regressors (dummies). The least squares residuals (u) are taken from the mean equation to construct a variance equation, and the variance equation is used to standardize the residuals from the mean equation.

$$\log(\widehat{u}^2) = x'\gamma + \epsilon \qquad \text{(variance equation)}$$

Finally, a final linear transformation is performed to calculate the adjusted w:

$$w_{adj} = a + b\left[\frac{\widehat{u}}{\exp(x'\gamma/2)}\right]$$

where a and b are chosen so that the sample means and variances of w and w_{adj} are the same. The linear transformation makes the units of measurement of adjusted and unadjusted data the same, which facilitates interpretation of our empirical results.

In our research, however, we find that all coefficients for seasonal dummies of both conditional mean and variance equations are generally insignificant. As a result, the fitted conditional mean and variance curves should not be very smooth and there should be local maxima in the process of estimation. Therefore, if we adjust the seasonal effects unconditionally, the procedure itself will make the adjusted data series spurious simply because datasets of SPX returns and VIX innovations suggest little evidence of seasonal effects. The exact SAS procedure can be obtained by request. Results are listed in Appendix I.

2.4 Outliers

For occasional outliers in our dataset we introduce a spline treatment as follows to the raw dataset to eliminate such effects:

$$\widehat{x}_i = \begin{cases} \frac{1}{2}\left\{x_i + \frac{4}{\pi}\arctan\left[\frac{\pi}{4}(x_i + \sigma_{tr})\right] - \sigma_{tr}\right\} & x_i < -\sigma_{tr} \\ x_i & -\sigma_{tr} \le x_i \le \sigma_{tr} \\ \frac{1}{2}\left\{x_i + \frac{4}{\pi}\arctan\left[\frac{\pi}{4}(x_i + \sigma_{tr})\right] + \sigma_{tr}\right\} & \sigma_{tr} < x_i \end{cases}$$

where x_i denotes an element of $x_t{-}1$ (lagged raw datasets). This is a trigonometric spline transformation that has no effect on values of x_i within $[-\sigma_{tr}, \sigma_{tr}]$, but progressively compresses values that exceed $\pm\sigma_{tr}$.

Because it affects only $y_t{-}1, \ldots, y_t{-}L$ (the conditioning set) and not y_t (data), so the asymptotic properties of SNP estimators are unaltered.

For data from financial markets, Gallant and Tauchen (2006) cited a huge amount of empirical evidence that suggests a long simulation from a fitted model has unconditional variance and kurtosis much larger than the variance and kurtosis of the sample. When the spline transform is imposed, this anomaly is effectively eliminated while the estimated coefficients and the value of likelihood are not much affected. Thanks to their contribution we can apply spline transformations to the x_{t-1} that enters $P(z, x)$, μ_x, and $_x$ with $\sigma_{tr} = 2$. The order the transformations are implemented is as follows:

$\tilde{y}_t$	$\rightarrow$	y_t	$\rightarrow$	x_{t-1}	$\rightarrow$	$\hat{x}_{t-1}$	$\rightarrow$	μ_x, R_x
raw data		centred, scaled		lagged		spline data		location-scale transformed

where all inputs and outputs are in the units of the raw data $\tilde{y}_t$.

3. Model of Density

The present work begins with an examination of the characteristics of the law of motion itself, with the primary objective to determine the extent to which it deviates from the Gaussian vector autoregressive (VAR) model. Elaborating on a paper by Phillips (1983) and Gallant and Nychka (1987), in which they propose to approximate the unknown density in a model by Hermite series. This approach finally develops to a parsimonious but sufficient procedure of empirical density function by Gallant and Tauchen (1999). This estimation strategy is called the SemiNonParametric (SNP) methodology, which is an approach that applies conventional estimation and testing to models derived from series expansions.

The method is based on the notion that a Hermite expansion can be used as a general approximation to an empirical density function. Let z denote a vector of a dimension M, the probability density of this vector can be written as an approximation by Hermite polynomials (namely the Hermite density) which is of the form,

$$h(z) = \frac{[\rho(z)]^2 \phi(z)}{\int [\rho(s)]^2 \phi(s) ds} \tag{1}$$

where $\rho(z)$ denotes a multivariate polynomial of degree K_z and $\varphi(z)$ denotes the density function of a multivariate Gaussian distribution with mean zero and the identity as its variance-covariance matrix. Denote the coefficient vector of ρ (z) by a whose length depends on K_z and M. The constant factor $\frac{1}{\int [\rho(s)]^2 \phi(s) ds}$ makes sure h (z) to integrate to one.

Given the Hermite density above in Equation (1), we can easily expand the density even further by allowing conditional heteroskedasticities, heterogeneities and potential interactions in the multivariate case. The variables within y (the dataset) are standardized using the location scale transformation $y = Rz + \mu$, where R is an upper triangular matrix and μ is an M-vector, gives

$$f(y|\theta) = \frac{\{\rho\,[R^{-1}(y-\mu)]\}^2 \{\phi\,[R^{-1}(y-\mu)] / |\det(R)|\}}{\int [\rho(s)]^2 \phi(s) ds} \tag{2}$$

Because $\varphi[R^{-1}(y - \mu)]/|\det(R)|$ is the density function of the M-dimensional, multivariate, Gaussian distribution with mean μ and variance-covariance matrix $\Sigma = RR'$, and because the leading term of the polynomial part is one, then the leading term of the entire expansion is proportional to the multivariate, Gaussian density function. Denote the Gaussian density of dimension M with mean vector μ and variance-covariance matrix Σ by $n_M(y|\mu, \Sigma)$ and write

$$f(y|\theta) = \frac{[\rho(z)]^2 n_M(y|\mu, \Sigma)}{\int [\rho(s)]^2 \phi(s) ds} \tag{3}$$

where $z = R^{-1}(y - \mu)$. And the parameter set θ is made up of the coefficients a of the polynomial $\rho(z)$ together with μ and R and they are estimated by maximum likelihood. When K_z is put to zero, one gets $f(y|\theta) = n_M(y|\mu, \Sigma)$ exactly because the leading term in the Hermite expansion of $[\rho(z)]^2$ is one. When K_z is positive, one gets a Gaussian density whose shape is modified due to multiplication by a polynomial $[\rho(z)]^2$. The shape modifications thus can be arbitrarily rich and hence give increasing precision of the density approximation as K_z becomes large.

It is also possible that some heterogeneity property exists (the distribution of z_t depends on x_t−1). In this case, each coefficient of the polynomial $\rho(z)$ is a polynomial of degree K_x in x (same as x_t−1). Denote this polynomial by $\rho(z, x)$. Denote the mapping from x to the coefficients a of $\rho(z)$ such that $\rho(z|a_x) = \rho(z, x)$ by a_x and the number of lags on which it depends by L_p. The form of the density with this modification is

$$f(y|x, \theta) = \frac{[\rho(z, x)]^2 n_M(y|\mu_x, \Sigma)}{\int [\rho(s)]^2 \phi(s) ds} \tag{4}$$

where $y_t = Rz_t + \mu_x$, and μ_x is a linear function that depends on L_u lags,

$$\mu_x = b_0 + Bx_{t-1} \tag{5}$$

So if K_x is put to a positive integer, the shape of the density will depend upon x. Hence, all moments can depend upon x and the density can, in principle, approximate any form of conditional heterogeneity (Gallant and Tauchen, 1999; Gallant *et al.*, 1991).

It is obvious that the leading term of the expansion is $n_M(y|\mu_x, \Sigma)$ which is called the Gaussian vector auto regression or Gaussian VAR. When K_z is put to non-zero, one gets a semiparametric VAR density that can approximate well over a large class of densities whose first moment depends linearly on x according to Equation (5) and whose shape is constant with respect to variation in x.

For the case of $M > 1$(multivariate estimation), a number of interactions (cross product terms) for even modest settings of degree K_z would exist. Accordingly, additional tuning parameters, I_z and I_x are introduced in estimations to control higher order interactions.

In practice, the leading term $n_M(y|\mu_x, \Sigma)$ can be put to a Gaussian GARCH rather than a Gaussian VAR. The form is:

$$\Sigma_{x_{t-1}} = R_0 R_0' \tag{6}$$

$$+\sum_{i=1}^{Lg} Q_i \Sigma_{x_{t-1-i}} Q_i' \tag{7}$$

$$+\sum_{i=1}^{Lr} P_i(y_{t-i} - \mu_{x_{t-1-i}})(y_{t-i} - \mu_{x_{t-1-i}})' P_i' \tag{8}$$

$$+\sum_{i=1}^{Lv} \max[0, V_i(y_{t-i} - \mu_{x_{t-1-i}})] \max[0, V_i(y_{t-i} - \mu_{x_{t-1-i}})]' \tag{9}$$

$$+\sum_{i=1}^{Lw} W_i x_{(1),t-i} x'_{(1),t-i} W_i' \tag{10}$$

where R_0 is a factorized upper triangular matrix, the matrices P_i, Q_i, V_i and W_i can be scalar, diagonal or full M by M matrices, the notation $x_{(1),t}-i$ indicates that only the first column of $x_{(1),t}-i$ enters the computation, and the $max(0, x)$ function is applied elementwise. Accordingly we call these four types of heteroscedasticities *Q-type*, *P-type*, *V-type*, and *W-type*, and the lags of x_t-1 on which they depend are denoted as L_g, L_r, L_v and L_w, respectively. Clearly if $L_r > 0$, an ARCH effect is present, which makes the leading term a Gaussian ARCH, and when $L_g > 0$, and $L_r > 0$ at the same time a GARCH effect is entering the leading term, which is of the form of Gaussian GARCH.

Therefore a much general leading term is achieved by allowing conditional heteroskedasticity, so with $\Sigma_{x_{t-1}}$ specified as either an ARCH or GARCH as above, the form of the conditional density becomes:

$$f(y|x,\theta) = \frac{[\rho(z,x)]^2 n_M \left(y|\mu_x, \Sigma_{x_{t-1}}\right)}{\int [\rho(s)]^2 \phi(s) ds} \tag{11}$$

where $z_t = R_x^{-1}(y - \mu_x)$. And all the coefficients are arranged inside the vector of θ,

$$\theta = vec\left[a_0|A|b_0|B|R_0|P_1 \ldots P_p|Q_1 \ldots Q_q|V_1 \ldots V_q|W_1 \ldots W_q\right]$$

in which a_0 is the subset of a that does not depend on x, and A controls the mapping from x to the subset of a that does depend on x. The parameters of the location function are $[b_0|B]$ whose length is controlled by L_u. The other variables in θ are all stated in above texts. And all these parameters are estimated by maximum likelihood estimation.

To sum up, as seen in Table 1, the SNP method employs an expansion in Hermite functions to approximate the conditional density of a multivariate process by setting principal tuning parameters and hence different model modifications away from a Gaussian distribution can be achieved.

An appealing feature of this expansion is that it is a non-linear non-parametric model that directly nests the Gaussian VAR model, the semi-parametric VAR model, the Gaussian ARCH model, the semi-parametric ARCH model, the Gaussian GARCH model, and the semiparametric GARCH model. The SNP model is fitted using conventional maximum likelihood together with a model selection strategy determining the appropriate order of expansion.

TABLE 1

Model Types and Principal Tuning Parameters

		Parameter setting				*Characterization of* $\{y_t\}$
Lu = 0	Lg = 0	Lr = 0	Lp ≥ 0	Kz = 0	Kx = 0	iid Gaussian Gaussian
Lu > 0	Lg = 0	Lr = 0	Lp ≥ 0	Kz = 0	Kx = 0	VAR semi-parametric
Lu > 0	Lg = 0	Lr = 0	Lp ≥ 0	Kz ≥ 0	Kx = 0	VAR Gaussian ARCH
Lu ≥ 0	Lg = 0	Lr >0	Lp ≥ 0	Kz = 0	Kx = 0	semi-parametric ARCH
Lu ≥ 0	Lg = 0	Lr >0	Lp ≥ 0	Kz >0	Kx = 0	Gaussian GARCH
Lu ≥ 0	Lg > 0	Lr >0	Lp ≥ 0	Kz = 0	Kx = 0	semi-parametric
Lu ≥ 0	Lg > 0	Lr >0	Lp ≥ 0	Kz >0	Kx = 0	GARCH non-linear
Lu ≥ 0	Lg ≥ 0	Lr ≥ 0	Lp > 0	Kz >0	Kx > 0	non-parametric

4. Empirical Evidence

Results of the SNP model specifications and diagnostic tests are summarized in Table 2 for both bivariate models (Panel A and B) and univariate models (Panel C), in which all the values are comparable and they are listed in a way along the excessively rich parameterization (increasing P_θ), and the 'Obj.' column contains minimized negative log-likelihoods in each set of SNP estimation, and the column labelled 'BIC' indicates the Schwarz information criterion. All the other labels in Table 2 have already been defined in the previous section.

Among the models in Panel A of Table 2 which specify the ARCH-leading models in the bivariate Hermite density expansion, the Schwarz-preferred model has $L_u = 3$, $L_r = 13$, $L_p = 1$, $I_z = I_x = 0$,and $K_z = 4$ with $P_\theta = 51$. Under this specification the short-term diagnostic Wilks's lambda is significant for the conditional mean regression. This indicates that there exists short-term conditional heterogeneity that is not accounted for by the current model specification. Separately, the short-term conditional variance is adequately approximated. Therefore a single Schwarz criterion turns out to be too aggressive to cut down further extension of non-linear property (the case of $K_x > 0$) from entering the Hermite density specification. This fact can be found by comparing the 8th and the 9th row in Panel A where the BIC shows a jump in value when K_x moves from 0 to 1, whereas the specification test based on Wilks's lambda on conditional mean moves from significant to highly insignificant in the mean time.

The GARCH-leading term SNP bivariate model shown in Panel B of Table 2 indicates that the Schwarz-preferred model has $L_u = 3$, $L_g = 1$, $L_v = 1$, $L_r = 1$, $L_p = 1$, $I_z = I_x = 0$,and $K_z = 4$ with $P_\theta = 29$, while the short-term diagnostic Wilks's lambda is significant for the conditional mean. It suggests an insufficiency to fully explain the heterogeneity. We hence increase K_x (the degree of polynomials in the coefficients of Hermite expansion $\rho(z, x)$ in both ARCH-leading and GARCH-leading models) and find the optimized value is $K_x = 1$ under both leading terms, but the BICs increase by less than 0.002 compared with the models with K_x =0. However, by doing so, the short-term diagnostic tests are highly insignificant which suggests the Hermite expansions deviate little from the truth. In other words, the customized diagnostics can be seen as a remedy to compensate the conservative and aggressive properties of the BIC. Therefore, the BIC-preferred fittings further justified by diagnostics can then be trusted to be parsimonious enough as well as reflecting the true data generating process with little lack of explanatory variables. It is clear a parsimonious GARCH-leading SNP model dominates a large ARCH-leading term model because of smaller BIC and the diagnostics from our results in Table 2. Diagnostics apart from the information criteria are also suggestions by Gallant and Tauchen (2006) but exact procedures vary across literature.

In past decades, the risk-return relationship has been examined by lots of researchers but it is still open to further study. Some issues pertain to the predictability of price changes, the nature of the relationship between price changes and volatility, and the shape characteristics of the probability density of price changes. Others concern asymmetry of the conditional variance function (leverage effect) and the relationship between the risk and conditional price volatility. Our methods look into these issues by an SNP estimation of the one step-ahead, bivariate, conditional density. The estimation itself embodies sufficient sample information without unreasonable presumptions (property of non-parametric approaches), and the customized diagnostics/information criteria keep the model parsimonious in the meanwhile.

The optimized conditional density is a function of 31 variables, it is hard to describe directly about all the economic meanings of Hermite expansion coefficients and conditional heteroscedasticity coefficients directly. The final fit of GARCH-leading SNP model is the one with $L_u = 3$, $L_g = 1$, $L_v = 1$, $L_r = 1$, $L_p = 1$, $I_z = I_x = 1$, $K_x = 1$ and $K_z = 4$ (bottom row in Panel B of Table 2). Due to the length of this paper, only the final fit is reported in Table 3, others can be distributed upon request.

TABLE 2

SNP Return and Change of VIX Series: Optimized Models

													Specification Tests on 20 lag cubic			
													Mean		Variance	
Lu	Lg	Lr	Lp	Lv	Lw	Kz	Iz	Kx	Ix	P_θ	*Obj.*	*BIC*	*Wilks*	*pv*	*Wilks*	*pv*
Panel A: Bivariate ARCH Leading Term SNP Estimation																
1	0	0	1	0	0	0	0	0	0	9	2.80	2.81*	0.85	<.00	0.68	<.00
3	0	0	1	0	0	0	0	0	0	17	2.79	2.80*	0.87	<.00	0.68	<.00
3	0	13	1	0	0	0	0	0	0	43	2.59	2.63*	0.94	0.14	0.96	0.99
3	0	13	1	0	0	4	0	0	0	51	2.51	2.56*	0.93	0.01	0.95	0.80
3	0	13	1	0	0	4	0	1	0	53	2.51	2.57*	0.94	0.11	0.95	0.91
3	0	13	1	0	0	4	1	1	0	53	2.51	2.57*	0.94	0.11	0.95	0.91
3	0	13	1	0	0	4	1	1	1	53	2.51	2.57*	0.94	0.54	0.95	0.91
Panel B: Bivariate GARCH Leading Term SNP Estimation																
1	0	0	1	0	0	0	0	0	0	9	2.80	2.81*	0.85	<.00	0.68	<.00
3	0	0	1	0	0	0	0	0	0	17	2.79	2.80*	0.87	<.00	0.68	<.00
3	1	1	1	0	0	0	0	0	0	20	2.58	2.60*	0.94	0.10	0.93	0.01
3	1	1	1	1	0	0	0	0	0	21	2.57	2.59*	0.93	0.04	0.94	0.12
3	1	1	1	1	0	4	0	0	0	29	2.50	2.52*	0.93	0.01	0.94	0.08
3	1	1	1	1	0	4	0	1	0	31	2.50	2.53*	0.94	0.06	0.94	0.20
3	1	1	1	1	0	4	1	1	0	31	2.50	2.53*	0.93	0.01	0.94	0.20
3	1	1	1	1	0	4	1	1	1	31	2.50	2.53*	0.94	0.06	0.94	0.20
Panel C: Univariate GARCH Leading Term SNP Estimation																
1	0	0	1	0	0	0	0	0	0	3	1.40	1.41*	3.64	<.00	16.59	<.00
1	1	1	1	0	0	0	0	0	0	5	1.27	1.27*	1.21	0.13	0.80	0.86
1	1	1	1	1	0	0	0	0	0	6	1.26	1.26*	1.05	0.37	0.61	0.99
1	1	1	1	1	0	4	0	0	0	10	1.24	1.25*	1.03	0.41	0.59	0.99
1	1	1	1	1	0	4	0	1	0	11	1.24	1.25*	1.02	0.43	0.59	0.99

Note: Superscript '*' means the relevant model is optimized.

TABLE 3
Optimized Bivariate GARCH-leading SNP Model

Index	*Theta (θ)*	*Standard error*	*t-statistic*	*Descriptor*
1	0.01306	0.01414	0.92361	a0[1]
2	−0.09052	0.00993	−9.11647	a0[2]
3	−0.05483	0.00914	−5.99613	a0[3]
4	0.08526	0.00811	10.51614	a0[4]
5	0.09326	0.0274	3.40326	a0[5]
6	−0.18137	0.01081	−16.7801	a0[6]
7	0.06177	0.01871	3.30162	a0[7]
8	0.11363	0.01435	7.91636	a0[8]
9	1	0	0	A(1,1)
10	0.05053	0.02573	1.96348	A(1,2)
11	0.04123	0.02582	1.59703	A(1,3)
12	−0.13	0.04646	−2.79795	b0[1]
13	−0.02814	0.02096	−1.34273	b0[2]
14	−0.05728	0.01412	−4.0564	B(1,1)
15	0.07749	0.01134	6.83398	B(2,1)
16	0.00653	0.01314	0.49723	B(1,2)
17	−0.05578	0.01495	−3.7313	B(2,2)
18	−0.06344	0.0136	−4.66633	B(1,3)
19	0.04396	0.0108	4.0695	B(2,3)
20	−0.0399	0.01315	−3.03481	B(1,4)
21	−0.00578	0.01422	−0.40682	B(2,4)
22	−0.06009	0.0138	−4.35524	B(1,5)
23	0.04558	0.01165	3.9138	B(2,5)
24	−0.02747	0.01207	−2.27641	B(1,6)
25	−0.01661	0.0141	−1.17845	B(2,6)
26	0.16132	0.01117	14.44261	R0[1]
27	−0.02577	0.004	−6.44176	R0[2]
28	0.07895	0.0092	8.58245	R0[3]
29	0.20366	0.01051	19.36872	P(1,1)d
30	0.2336	0.0108	21.62118	P(2,1)d
31	−0.9707	0.00178	−545.342	Q(1,1)s
32	0.16568	0.01575	10.52097	V(1,1)s
SUMMARY STATISTICS				
Length of θ $(P_\theta + 1)$		32		
s_n		2.49596383		
AIC		2.50320344		
HQ		2.51133872		
BIC		2.52623326		

Note: All coefficients and tests are consistent with the final fit with a model specification of $L_u = 3$, $L_g = 1$, $L_v = 1$, $L_r = 1$, $L_p = 1$, $I_z = I_x = 1$, $K_x = 1$ and $K_z = 4$.
AIC, HQ, BIC denote the three information criteria: Akaike Information Criterion (Akaike, 1973), Hannan and Quinn Information Criterion (Hannan and Quinn, 1979), and Schwarz Information Criterion (Schwarz, 1978), respectively.
s_n denotes the objective value.

It is interesting that there is some strong evidence of a V-type heteroscedasticity (shown in Equation (9)) within the GARCH-leading term, and it is positive (0.166) as shown in Table 3 (index number 32). The significance means there exists a leverage effect rather than a symmetric GARCH. We also document that an efficient procedure to test this effect should be treating the leverage as an add-on within the GARCH discipline, not some pre-determined asymmetric function, otherwise the leverage is highly likely to be washed away by high-order ARCH terms. We do not find any W-type heteroskedasticity.

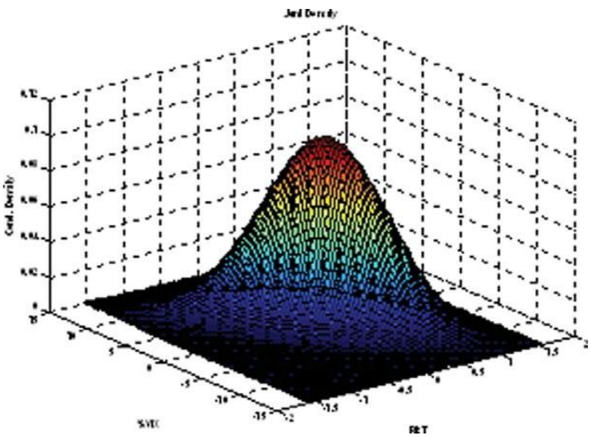

Figure 3. Surface of bivariate one-step-ahead joint density.

5. Results

5.1 Contemporaneous relationship

The empirical one-step-ahead, bivariate, conditional density at the mean can be written as:

$$f\left(ret_t, \%VIX_t | ret_{t-1:4282} = \overline{ret}, \%VIX_{t-1:4282} = \overline{\%VIX}\right) \tag{12}$$

where ret_t defines the SPX daily return, $\%VIX_t$ defines the daily innovation (in terms of percentage change) of CBOE VIX; the $\overline{ret}$ and $\overline{\%VIX}$ calculate their unconditional means respectively. Function (12) contains an information set defined as of lagged returns and volatility innovations: $\Omega = \{ret_{t-1:4282} = \overline{ret}, \%VIX_{t-1:4282} = \overline{\%VIX}\}$. Therefore, conditioning on this particular information set in the function gives an in-sample one-step-ahead prediction of the bivariate conditional density. It embodies the sample information pertaining to the predictability of the price changes, the nature of the possible relationship between returns and volatility index, and the shape characteristics of the probability density of returns. Because the fitted conditional density is a generalized Hermite expansion of 32 variables (including intercept) thanks to the GARCH-leading non-parametric optimization, it is difficult to describe all the parameters directly. Our strategy is therefore to examine features of the density by looking at marginal, low-order moments, and conditional moment functions and to interpret these features in view of the economic issues.

Figure 3 shows the three-dimensional surface of the one-step-ahead joint density as shown in Function (12). It suggests the fitted density is quite smooth over the dataset. No hump-shaped feature exists in the density function. Figure 4 depicts the contour plots of it demonstrating a downward-sloping shape; the downward tendency also gives us the intuitive impression that there is a rough negative risk-return correlation.

Given the joint density estimation as before, the marginal conditional density of ret_t is computed as:

$$f_{ret}\left(ret_t | ret_{t-1:4282} = \overline{ret}, \%VIX_{t-1:4282} = \overline{\%VIX}\right)$$
$$= \int f\left(ret_t, \%VIX_t | ret_{t-1:4282} = \overline{ret}, \%VIX_{t-1:4282} = \overline{\%VIX}\right) d\%VIX_t$$

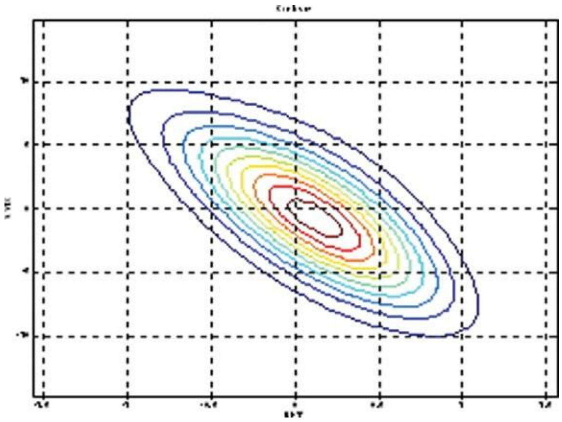

Figure 4. Contours of bivariate joint density.

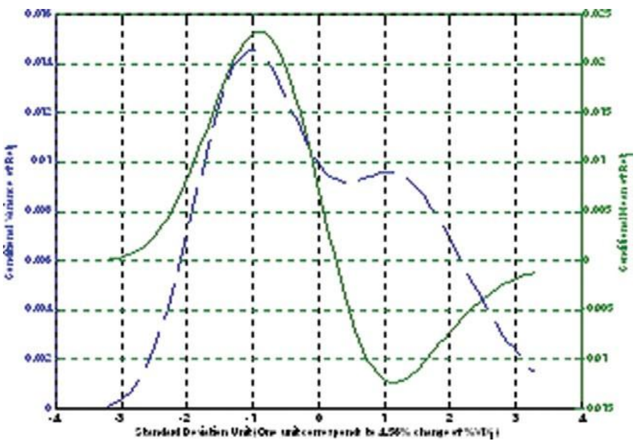

Figure 5. Contemporaneous relationships.

Since the in-sample prediction nests the unconditional means of SPX return and the %*VIX* as its information set, so it is an intuitive guess that the one-step-ahead prediction of SPX return might not deviate from the unconditional tendency too much. And, provided the joint density, it should also be easy to calculate all the conditional moments accordingly.

The contemporaneous relationship between SPX return (denoted as ret_t) and the percentage change of VIX (denoted as $\%VIX_t$) is revealed by looking at the conditional mean and variance of ret_t given $\%VIX_t$ (along slices of the bivariate (ret_t, $\%VIX_t$) density). Figure 5 shows the first two moments of ret_t conditional on $\%VIX_t$. These are the mean and variance of ret_t univariate marginal density obtained by slicing the bivariate density show in Figure 3 along a line through (−15,15) on the $\%VIX_t$ axis parallel to the ret_t axis. The horizontal axis in Figure 5 is in standardized units (divided by the standard deviation) of the marginal conditional density of

$\%VIX_t$,

$$f_{\%VIX}\left(\%VIX_t | ret_{t-1:4282} = \overline{ret}, \%VIX_{t-1:4282} = \overline{\%VIX}\right)$$

The range of the horizontal axis in Figure 5 extends four standard deviations on either side (each unit corresponds to 4.56 change in values of %*VIX*). We focus particularly on curves within three standard deviation units because the moment functions become oscillatory outside this bound. The reason is that, for the data of %*VIX*, it is too rare for %*VIX* to exceed over three standard deviations change (12 or more change in value of %*VIX* happens less than 2 per cent of the time). Although we already impose trimmings to deal with outliers, yet the practical market situation should not be ignored after all. We choose a relatively less conservative bound of three standard deviation units because our dataset is moderately large enough. A more conservative limit, two standard deviations, could be chosen because %*VIX* already rarely exceeds this limit (less than 8 per cent of the time). Fortunately, the degree of freedom in daily data is still reasonable for a bigger bound without spurious fittings and most importantly, we can also look into the tail behaviour.

Generally speaking, Figure 5 depicts an intuitive and straightforward result that a decrease in %*VIX*(negative %*VIX*) means a positive mean return while an increase in *VIX*(positive %*VIX*) implies conditional losses. In addition, there are also some untraditional features that can be observed as follows.

First, it is very interesting to note that the direction of the conditional daily return (solid line) is related to the contemporaneous innovation of the implied volatility in a non-linear way, and there seems to be some threshold to define this non-linearity. Specifically, within the range of one standard deviation change of *VIX* (in absolute value), the conditional return is negatively and monotonically related to the %*VIX*. The relationship is very close to a linear correlation and the slope is −0.11, which means one *unit* change in *VIX* implies a 0.11 decrease in the conditional total return. But this relationship, however, reverts itself and becomes positive outside the bound of one standard deviation unit. So a hump-shaped curve exists and shows non-monotonicity. Our findings might provide some explanations for the conflicting results put forward by early literature in which different signs of linear relations are found using various methodologies. The reason is that a risk-return relationship may not be as simple as we might think: a linear model has a tendency to smooth the non-linearity by simply ignoring higher-order terms; pure nonlinear parametric models may not be flexible enough to accommodate heterogeneities or even to distort the truth due to unreasonable assumptions (for example, we must assume distributions for the errors), and ultimately prevent the model from revealing the true DGP. Our SNP model is a distribution-free non-parametric model which starts from a standard normal distribution and then adjusts the density function by higher moments or leading terms determined by the dataset, while balancing over-fittings by penalties and customized diagnostics. Advantages as such make the SNP discipline dominate parametric models by allowing the real dataset to play a role, and in the mean time, circumvent over-fitting problems typically found in non-parametric methodologies.

Our result also partly agrees with Linton and Perron (2003) which shows a hump-shaped pattern of the risk premium, although our result turns out to be more smooth and less complex. It is possible that we could model short-term dynamics artificially well by simply expanding the conditional density function using higher-order Hermite polynomials. By doing so, a much complicated relationship might be derived and the non-monotonicity could be far more sensitive to innovations of the VIX. However, a complex relationship might not have much meanings for real trading strategies. Statistical relationships may not always apply in practice, especially when they are highly complicated. Traders may be more interested in summarizing a meaningful and aggregated relationship such that benchmarks of long/short positions could be made. Therefore, a simple but useful result might be preferred. In our results, we penalize some of the meaningless dynamics and focus more on thresholds at which the relationship reverts to different signs. We conclude that (empirically) the threshold is one standard deviation of the VIX innovation (in percentage terms). If the VIX innovation exceeds this bound (positive or negative), the correlation itself reverts to a positive one. Therefore, the contemporaneous risk-return behavior depends not only on the sign of risk metrics (sentiment shifts), but also on the magnitudes of the change. In other words, fear or exuberance (extreme innovation of VIX) does affect the conditional return, but the influence is non-monotonic and hump-shaped. On the one hand, very deep fear does not necessarily mean huge losses, instead, the loss may not be as bad as fears of normal levels, while on the other, exuberance does not correlate to big returns.

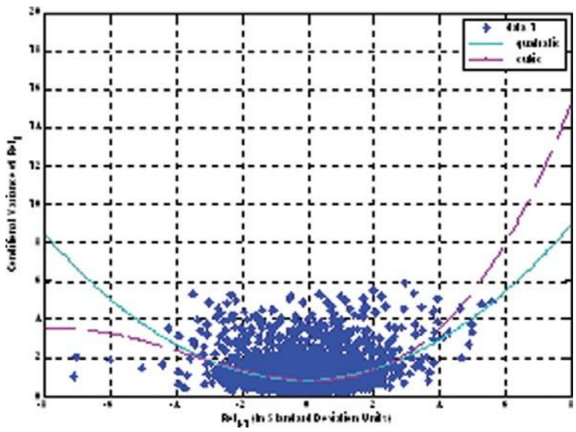

Figure 6. Bivariate conditional variance.

Secondly, in Figure 5, there also exists a generally bigger positive expected return in the negative %*VIX* region, whilst on the positive region, a generally smaller expected loss can be clearly observed. This asymmetry suggests that information pertaining to %*VIX* can actually help investors to reduce their losses and increase their total returns as well. More technically, when the daily %*VIX* enters the bivariate SNP system, it effectively carries information that can be used to hedge contemporaneous exposure to conditional loss (the so-called loss aversion), so that it mitigates the expected loss and more importantly, it even contains information about profiting over small (but frequent) positive %*VIX*.

Thirdly, it is also interesting to observe the second-order conditional moment of the total return (broken line in Figure 5). We find that positive %*VIX* correlates to a volatility moderation (to the right) relative to negative *VIX* innovations (to the left). The result is consistent to intuitive explanations of how traders treat different signs of VIX innovation. Less fear (or even exuberance) eases motives to stabilize total return fluctuations simply because the volatility in the upside conditional returns is only treated as sweeteners, but higher expected risk (measured by fear gauge) will intensify market turmoil and hence increase trader's motivation to foil conditional volatility of returns. In other words, investors do not consider a decrease in the implied volatility as a true market risk. The VIX's contemporaneous effect seems to mitigate the conditional variance of return only when %*VIX* is positive.

5.2 Asymmetries of the conditional variance

Figure 6 depicts the marginal conditional variance function of total return against lagged return itself (also in terms of standard deviation unit). It is clear that the conditional variance function displays the conditional heteroscedasticity captured by traditional GARCH-family applications. As before, on the horizontal axis, one unit stands for one standard deviation corresponding to 0.95 change in lagged SPX returns.

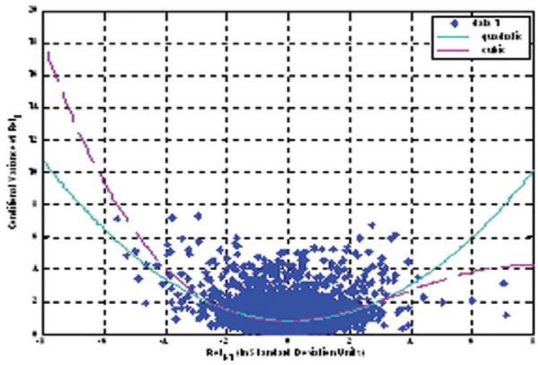

Figure 7. Univariate conditional variance.

It should be noticed that even though the SNP estimation does not impose any form of symmetry towards the marginal conditional variance function of return, the plots generated from which shows a symmetric feature, as shown by the quadratic fitting.

However, the symmetry conflicts with the findings by Nelson (1989, 1991), Pagan and Schwert (1990), and others who find evidence of asymmetric properties in the conditional variance function. Literature names asymmetries of the sort the leverage effect after early studies by Black (1976) and Christie (1982). They both provide evidence that changes in the equity value of a firm affect the riskiness of the firm's equity. Recent evidence, for instance, Low (2004) uses a simple linear model to prove the significance of this effect. But tests of the leverage hypothesis by French *et al.* (1987) and by Schwert (1989, 1990) suggest that financial leverage could not be responsible for asymmetries of the magnitude reported in literature. Nevertheless, the consensus is that leverage effect is due to the asymmetry in conditional variance of returns.

The reason why our SNP model fails to capture this asymmetry is because our SNP constructs a joint density function in which the SPX daily return and the implied volatility change are considered simultaneously. Hence, the conditioning information set includes both past returns and %*VIX* as well. When we exclude the %*VIX* out of the system and use a univariate SNP model to estimate the return series, we get a univariate conditional volatility plot shown in Figure 7. And by doing so, evidence of asymmetry is also uncovered. Therefore, Figure 7 still confirms a rather mild leverage effect by noticing the quadratic fitting curve is higher on the left than on the right, which is consistent to previous findings (negative past return causes higher conditional volatility). However, almost all the other articles regarding leverage effect examine a marginal price process instead of considering multivariate cases. This difference suggests that introducing VIX in to the analysis is responsible for producing the symmetry seen in Figure 6.

We also note from Figure 6 that although the introduction of VIX contributes to the symmetry, this symmetry is not a real one in the sense that the %*VIX* finally makes the conditional variance tilting to the left which can be fit using a cubic fitting.

6. Monte Carlo Simulations and Robustness Tests

6.1 Monte Carlo description

We use Monte Carlo simulations to do the robustness check. The upper panel in Figure 8 plots the real SPX return series, while bottom panels show simulations (of the same length) from bivariate ARCH-leading and GARCH-leading SNP models, respectively.

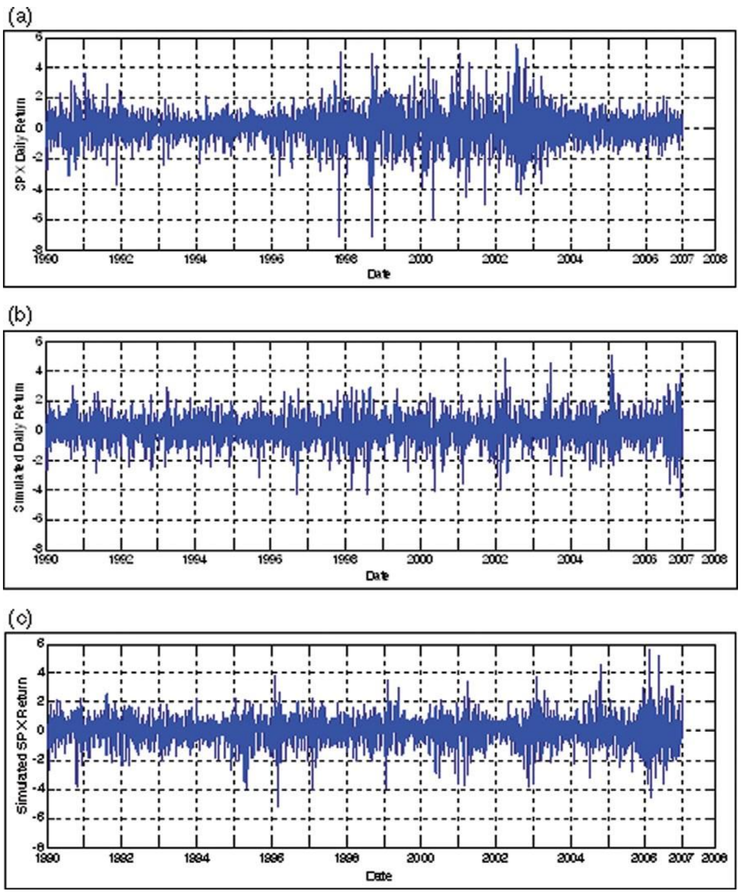

Figure 8. (a) Real SPX return series; (b) ARCH simulated daily returns; (c) GARCH simulated daily return.

Simulations capture features as in the real return series, but are less volatile. This is due to the spline transformation for outliers. We can see the GARCH-leading model carries the volatility moderation between 1992 and 1996, a volatile period from middle 2002 to middle 2003 and around the year 2001.

Table 4 shows the first four moments of all series. They all show excess kurtosis and negative skewness, which are typical to financial time series. The GARCH-leading model gives a much higher kurtosis which is closer to the kurtosis from the real series.

However, we should also notice all kurtosis of simulated series cannot explain the real one fully as observed in the financial market. A fundamental explanation is because the BIC is still conservative even when diagnostics kick in to determine a richer model. Therefore some temporary and abrupt fluctuations may be erased by the Schwarz penalty.

TABLE4

First four moments

	Real SPX daily return	*ARCH generator*	*GARCH generator*
Mean	0.032	0.037	0.031
Variance	0.995	0.808	0.781
Skewness	-0.101	-0.148	-0.199
Kurtosis	6.908	4.986	5.627

Despite its aggressiveness, researchers still prefer the Schwarz criterion, because evidence shows it does a good job in finding abrupt drops in integrated squared error which is the point at which one would like to truncate in Efficient Methods of Moments applications (Gallant and Tauchen, 1999; Coppejans and Gallant, 2002)

6.2 Robustness check

In order to check the validity of the SNP model, we construct the following model:

$$|RET_t| = \beta_0 + \beta_1|RET_{t-1}| + \beta_2\,[RET_{t-1}\, I\, (RET_{t-1} <0)] + u_t \qquad (13)$$

This regression effectively passes a V-shaped line through the central cloud. The parameter β_2 is the asymmetry coefficient: the more negative is β_2, the steeper is the slope on the left half of the 'V'. If the hypothesis of the legitimacy of the SNP mode can be accepted, then the model itself must generate similar series in terms of identical conditional properties. In order to ensure that findings from this regression are also features of the fitted (GARCH-leading) SNP model, I fit the same regressions to the simulations generated from the bivariate SNP data generating process (DGP) and from the univariate SNP DGP. The first 50 actual observations from January 1990 to March 1990, are used as the initial conditions for both simulations, and simulations are of the same length as the original data.

Table 5 shows three regressions of SPX return allowing asymmetric leverage effects. All coefficients are statistically significant at the 1 per cent significance level. And in particular, all the simulations based on the GARCH-leading SNP models (bivariate and univariate models) yield to similar features as the real SPX returns by observing the negativity of β_2(the dummy). These results support the GARCH-leading SNP model to be a powerful and precise estimate of the real data generating process. Comparing Panel B with Panel C, we find that the coefficient of the dummy from the univariate model is more negative than the one from the bivariate model. This is consistent with my previous results on the conditional variance that the introduction of VIX contributes to the symmetry. And when we eliminate the VIX from the system and estimate using the univariate SNP model the asymmetry is back and becomes more significant.

7. Conclusion

We use the bivariate semi-nonparametric (SNP) model developed recently by Gallant and Tauchen (2006) to study the contemporaneous relationship between the innovation of VIX and the expected SPX returns. We estimate the bivariate conditional joint density function using optimal Hermite expansion. The conditional density function is also subject to a possible leverage GARCH effect. We use the expectation of future market volatility as a source of risk (Chen, 2003, and Ang *et al.*, 2006). In our paper, this particular metrics of risk is calculated by the daily innovation in the implied volatility (the CBOE VIX).

We conclude that the contemporaneous risk-return behaviour depends not only on the sign of the risk metrics (sentiment shifts), but also on the magnitude of the change. In other words, fear or exuberance (extreme innovation of VIX) does correlate to conditional return, but the correlation is non-monotonic and hump-shaped. On the one hand, very deep fear does not necessarily mean huge losses, instead, the loss may not be as bad as fears of normal levels, while on the other, exuberance does not necessarily correlate with big returns.

Our result partly agrees with Christensen and Nielsen (2007) on negative and monotonic correlation but we argue that the negativity is subject to the magnitude of innovations of the expected volatility, and that positive relation also exists given dramatic changes of sentiment, so that a non-monotonic relationship seems to be a full story.

TABLE5

Estimates of coefficients for linear models

Variable	*Parameter estimate*	*Standard error*	*t-value*	*Pr* > | t |
Panel A: Estimates from SPX Daily Returns				
Intercept (β_0)	0.58344	0.01488	39.22	<0.0001
$\|RET_t{-}1\|$(β_1)	0.10501	0.01820	5.77	<0.0001
Dummy (β_2)	0.15470	0.02094	−7.39	<0.0001
Panel B: Estimates from Bivariate SNP DGP				
Intercept (β_0)	0.57527	0.01341	42.91	<0.0001
$\|RET_t{-}1\|$(β_1)	0.10312	0.01853	5.56	<0.0001
Dummy (β_2)	0.05489	0.02016	−2.72	0.0065
Panel C: Estimates from Univariate SNP DGP				
Intercept (β_0)	0.56957	0.01352	42.14	<0.0001
$\|RET_t{-}1\|$(β_1)	0.09517	0.01787	5.33	<0.0001
Dummy (β_2)	0.09572	0.02038	4.70	<0.0001

Note: Panel A regresses the real dataset of SPX returns whilst datasets in Panel B and Panel C are Monte Carlo simulations based on Bivariate SNP and Univariate SNP data generating processes respectively.

However, our framework is not intended to be predictive or keen to show any form of causality, but merely to directly illustrate the conditional contemporaneous risk-return relation. And the relationship is also allowed to change conditional on the contemporaneous market condition and on how much the VIX increases or decreases. The paper can be thought of as an extension of Christensen *et al.* (2010) on the dimension of nesting not only the signs but also magnitudes of innovations as the information set. Our results partly agree with Linton and Perron (2003) but differ in the risk measures. We also show a hump-shaped pattern of the total returns relative to risk perceptions but it turns out to be more smooth.

References:

Akaike, H. (1973). 'Information theory and an extension of the maximum likelihood principle', in Petrov, B. N. and Csaki, F. (eds), *Second International Symposium on Information Theory*. Budapest: Akademiai Kiado, pp. 267–81.

Ang, A., Hodrick, R. J., Xing, Y. and Zhang, X. (2006). 'The cross-section of volatility and expected returns', *Journal of Finance*, 61, pp. 259–99.

Areal, N. (2008). 'FTSE-100 Implied Volatility Index', Working paper series, available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1102135.

Black, F. (1976). 'Studies of stock price volatility changes', Proceedings of the 1976 Meetings of the Business and Economics Statistics Section, pp. 177–81.

Black, F. and Scholes, M. (1973). 'The pricing of options and corporate liabilities',

The Journal of Political Economy, 81, pp. 637–54.
Bluhm, H. H. W. and Yu, J. (2001). 'Forecasting volatility: evidence from the German stock market', *Working Paper*, University of Auckland (presented at the International Conference on Modelling and Forecasting Financial Volatility, September).
Brandt, M. W. and Qiang Kang (2004). 'On the relationship between the conditional mean and volatility of stock returns: a latent VAR approach', *Journal of Financial Economics*, 72, pp. 217–57.
Campbell, J. Y. (1987). 'Stock returns and the term structure', *Journal of Financial Economics*, 18(2), pp. 373–99.
Chen, J. (2003). 'Intertemporal CAPM and the cross-section of stock returns', *Working Paper*, University of Southern California.
Christensen, B. J., Dahl, C. M. and Iglesias, E. M. (2010). 'A class of semiparametric GARCH in mean models and its semiparametric efficiency bound', *CREATES Research Papers*, School of Economics and Management, University of Aarhus.
Christensen, B. J. and Nielsen, M. (2007). 'The effect of long memory in volatility on stock market fluctuations', *Review of Economics and Statistics*, 89 (4), pp. 684–700.
Christie, A. A. (1982). 'The stochastic behavior of common stock variances: value, leverage and interest rate effects', *Journal of Financial Economics*, 10, pp. 407–32.
Conrad, C. and Mammen, E. (2008). 'Nonparametric regression on latent covariates with an application to semiparametric GARCH-in-mean models', *Working Paper*, University of Mannheim.
Coppejans, M. and Gallant, A. R. (2002). 'Cross-validated SNP density estimates', *Journal of Econometrics*, 110, pp. 27–65.
Corrado, C. and Miller, T., Jr (2005). 'The forecast quality of CBOE implied volatility indexes', *Journal of Futures Markets*, 25, pp. 339–73.
Das, S. and Sarkar, N. (2000). 'An ARCH in the nonlinear mean model', *The Indian Journal of Statistics*, 62, pp. 327–44.
Dennis, P., Mayhew, S. and Stivers, C. (2006). 'Stock returns, implied volatility innovations, and the asymmetric volatility phenomenon', *Journal of Financial and Quantitative Analysis*, 41, pp. 381–406.
Dowling, S. and Muthuswamy, J. (2005). 'The implied volatility of Australian index options', *The Review of Futures Markets*, p. 14.
Feinstein, S. (1989). 'A theoretical and empirical investigation of the Black-Scholes implied volatility', PhD dissertation (Yale University, New Haven, CT).
French, K., Schwert, G. W. and Stambaugh, R. (1987). 'Expected stock returns and volatility', *Journal of Financial Economics*, 19, pp. 3–29.
Gallant, A. R., Hsieh, D. A. and Tauchen, G. E. (1991). 'On fitting a recalcitrant series: the pound/dollar exchange rate, 1974–83', in Barnett, W. A., Powell, J. and Tauchen, G. E. (eds), *Nonparametric and Semiparametric Methods in Econometrics and Statistics*, Proceedings of the Fifth International Symposium in Economic Theory and Econometrics, Cambridge: Cambridge University Press, Chapter 8, pp. 199–240.
Gallant, A. R. and Nychka, D. W. (1987). 'Seminonparametric maximum likelihood

estimation', *Econometrica*, 55, pp. 363–90.
Gallant, A. R., Rossi, P. E. and Tauchen, G. (1992). 'Stock prices and volume', *The Review of Financial Studies*, 5, pp. 199–242.
Gallant, A. R. and Tauchen, G. (1999). 'The relative efficiency of method of moments estimators', *Journal of Econometrics*, 92, pp. 149–72.
Gallant, A. R. and Tauchen, G. (2006). 'SNP: a program for nonparametric time series analysis Version 9.0 User's Guide', available at http://econ.duke.edu/web_les/arg/snp.
Ghysels, E., Santa-Clara, P. and Valkanov, R. (2005). 'There is a risk-return tradeoff after all', *Journal of Financial Economics*, 76, pp. 509–48.
Giot, P. (2005). 'Implied volatility indexes and daily value-at-risk models', *Journal of Derivatives*, 12, pp. 54–64.
Glosten, L., Jagannathan, R. and Runkle, D. E. (1993). 'On the relation between the expected value and the volatility of the nominal excess return on stocks', *Journal of Finance*, 48(5), pp. 1779–1801.
Hannan, E. J. and Quinn, B. G. (1979). 'The determination of the order of an autoregression', *Journal of the Royal Statistical Society, Series B (Methodological)*, 41(2), pp. 190–5.
Harvey, C. (2001). 'The specification of conditional expectations', *Journal of Empirical Finance*, 8, pp. 573–638.
Jiang, G. J., and Tian, Y. S. (2005). 'The model-free implied volatility and its information content', *The Review of Financial Studies*, 18, pp. 1305–42.
Lettau, M. and Ludvigson, S. C. (2009), 'Measuring and modeling variation in the risk-return tradeoff', in Ait-Sahalia, Y. and Hansen, L. P. (eds), *Handbook of Financial Econometrics*, Amsterdam: North-Holland.
Linton, O. and Perron, B. (2003). 'The shape of the risk premium: evidence from a semiparametric GARCH model', *Journal of Business and Economic Statistics*, 21(3), pp. 354–67.
Low, C. (2004). 'The fear and exuberance from implied volatility of S&P100 index options', *Journal of Business*, 77(3), pp. 527–46.
Merton, R. (1973). 'An intertemporal capital asset pricing model', *Econometrica*, 41, pp. 867–87.
Merton, R. (1980). 'On estimating the expected return on the market', *Journal of Financial Economics*, 8, pp. 323–61.
Nelson, D. (1989). 'Modeling stock market volatility changes', *Proceedings of the American Statistical Association, Business and Economics Section*, pp. 93–8.
Nelson, D. (1991). 'Conditional heteroskedasticity in asset returns: a new approach', *Econometrica*, 59, pp. 347–70.
Nishina K., Maghrebi, N. and Kim, M. S. (2006). 'Stock market volatility and the forecasting accuracy of implied volatility indices', *Discussion Papers in Economics and Business No. 06–09*, Graduate School of Economics and Osaka School of International Public Policy.
Pagan, A. and Hong, Y. (1991). 'Nonparametric estimation and the risk premium', in

Barnett, W., Powell, J. and Tauchen, G. (eds), *Nonparametric and Seminonparametric Methods in Econometrics and Statistics*, Cambridge: Cambridge University Press.

Pagan, A. and Schwert, G. W. (1990). 'Alternative models for conditional stock volatility', *Journal of Econometrics*, 45(1–2), pp. 267–90.

Phillips, P. C. B. (1983). 'ERAs: A new approach to small sample theory', *Econometrica*, 51, pp. 1505–25.

Poon, S.-H. and Granger, C. W. J. (2003). 'Forecasting volatility in financial markets', *Journal of Economic Literature*, 41, pp. 478–539.

Schwarz, G. (1978). 'Estimating the dimension of a model', *The Annals of Statistics*, 6(2), pp. 461–4.

Schwert, G. W. (1989). 'Indexes of United States stock prices from 1802–1987', *Papers 89–04*, Rochester, Business – General.

Schwert, G. W. (1990). 'Indexes of US stock prices from 1802 to 1987', *The Journal of Business*, 63(3), pp. 399–426.

Skiadopoulos, G. (2004). 'The Greek implied volatility index: construction and properties', *Applied Financial Economics*, 14(16), pp. 1187–96.

Appendix I: Seasonal Effect Adjustments

Adjustment Regressions for SPX Returns

	Location		Variance	
	Coeff.	S.D.	Coeff.	S.D.
Day of the week				
Monday	–	–	–	–
Tuesday	0.371	0.206	0.536	0.497
Wednesday	0.394	0.212	0.389	0.510
Thursday	0.358	0.212	0.622	0.510
Friday	0.345	0.212	0.572	0.509
No. of days since the preceding trading day				
GAP1	0.071	0.165	−0.302	0.396
GAP2	0.437	0.212	0.509	0.509
GAP3	0.750	0.121	0.672	0.292
Month or week				
Jan 1-7	0.028	0.132	−0.096	0.318
Jan 8-14	−0.529	0.121	0.100	0.292
Jan 15-21	0.022	0.127	0.131	0.305
Jan 22-31	0.057	0.106	0.080	0.255
Feb		–	–	
Mar	0.026	0.076	−0.030	0.183
April	0.052	0.077	−0.000	0.184
May	0.065	0.076	−0.125	0.181
June	−0.005	0.076	−0.238	0.181
July	−0.002	0.076	0.154	0.181
August	−0.064	0.075	−0.112	0.182
Sept	−0.052	0.077	−0.147	0.186
Oct	0.084	0.075	0.111	0.183
Nov	0.091	0.077	−0.122	0.187
Dec 1-7	0.144	0.121	0.247	0.292
Dec 8-14	−0.105	0.121	−0.610	0.292
Dec 15-21	0.129	0.121	0.022	0.292
Dec 22-31	0.160	0.113	−1.088	0.272
Trend				
Intercept	−0.373	0.223	−3.214	2.535
t	–	–	4.618	3.507
t^2	–	–	−3.897	2.490

Adjustment Regressions for Percentage Change of VIX

	Location		Variance	
	Coeff.	S.D.	Coeff.	S.D.
Day of the week				
Monday	–	–	–	–
Tuesday	−1.090	1.154	1.159	0.482
Wednesday	−0.923	1.184	1.137	0.494
Thursday	−0.484	1.184	1.269	0.494
Friday	−1.218	1.184	1.410	0.494
No. of days since the preceding trading day				
GAP1	1.315	0.921	−0.238	0.384
GAP2	1.720	1.184	1.329	0.494
GAP3	5.722	0.678	0.675	0.283
Month or week				
Jan 1-7	−0.561	0.739	0.128	0.309
Jan 8-14	0.293	0.679	0.661	0.283
Jan 15-21	−0.843	0.710	0.486	0.296
Jan 22-31	−0.126	0.593	0.120	0.248
Feb	–	–	–	–
Mar	0.106	0.424	0.362	0.177
April	−0.334	0.432	0.507	0.180
May	−0.201	0.426	0.097	0.178
June	−0.258	0.426	−0.074	0.178
July	0.310	0.427	0.200	0.178
August	0.268	0.423	0.178	0.177
Sept	0.186	0.431	0.281	0.180
Oct	0.100	0.423	0.328	0.177
Nov	−0.331	0.433	−0.148	0.180
Dec 1-7	−0.345	0.679	−0.179	0.283
Dec 8-14	−0.263	0.679	0.405	0.283
Dec 15-21	−0.729	0.679	0.191	0.283
Dec 22-31	0.625	0.632	0.332	0.264
Trend				
Intercept	0.479	1.222	0.377	0.519
t	–	–	0.305	0.492
t^2	–	–	−0.504	0.476