Australian School of Business

Working Paper

Never Stand Still

Australian School of Business

Australian School of Business Research Paper No. 2012 ECON 31

On the Risk Return Relationship

Jianxin Wang
Minxian Yang

This paper can be downloaded without charge from
The Social Science Research Network Electronic Paper Collection:
http://ssrn.com/abstract=2049649

ON THE RISK RETURN RELATIONSHIP

Jianxin Wang[1] and Minxian Yang[2]

Abstract

While the risk return trade-off theory suggests a positive relationship between the expected return and the conditional volatility, the volatility feedback theory implies a channel that allows the conditional volatility to negatively affect the expected return. We examine the effects of the risk return trade-off and the volatility feedback in a model where both the return and its volatility are influenced by news arrivals. Our empirical analysis shows that the two effects have approximately the same size with opposite signs for the daily excess returns of seven major developed markets. For the same data set, we also find that a linear relationship between the expected return and the conditional standard deviation is preferable to polynomial-type nonlinear specifications.

Key words: Risk premium, volatility feedback, GARCH-in-mean, Maximum likelihood, Mixture distributions, Time series

*We thank the participants at the 2011 NCER Econometrics Workshop at the Queensland University of Technology and a seminar at Macquarie University for their comments. The support from the Australian Council Linkage Grant LP0882468 is gratefully acknowledged.

[1] UTS Business School, the University of Technology, Sydney, email: Jianxin.Wang@uts.edu.au
[2] Corresponding author, School of Economics, the University of New South Wales, email: M.Yang@unsw.edu.au

1. INTRODUCTION

The risk return trade-off is a fundamental relationship in finance, which suggests that the expected excess return is positively related to the conditional variance. However the empirical evidence based on index return series has been mixed in the context of GARCH-in-mean models. A positive relationship between the expected excess return and the conditional variance is documented by French, Schwert and Stambaugh (1987), Chou (1988), Campbell and Hentschel (1993), and Lundblad (2007) among others, whereas a negative relationship is found by Nelson (1991), Glosten, Jagannathan and Runkle (1993) and Jensen and Lunde (2001) among others. Outside the GARCH-in-mean framework, the results from alternative models are not clear-cut either, see Campbell (1987), Pagan and Hong (1991), Harrison and Zhang (1999), Bandt and Kang (2004) and Ghysels, Santa-Clara and Valkanov (2005) among others.

Various explanations for the lack of a consensus in the empirical results are offered. Lundblad (2007) reasons that the sample sizes used in the literature are typically too small to accurately estimate the risk return relationship. Ghysels et al (2005) suggest that exploring data at various frequencies can sharpen the estimates of the conditional variance and its effect on the expected return. Rossi and Timmermann (2010) point out that the relationship between the expected return and the conditional volatility may be non-monotonic that cannot be correctly revealed by the models used in the literature. Yang (2011) argues that the volatility feedback effect proposed by French *et al* (1987) should be accounted for when quantifying the risk return trade-off in the GARCH-in-mean framework.

In the current paper, we consider two questions in the framework of GARCH-in-mean to further our understanding of the risk return relationship. First, when the volatility feedback is accounted for, does the conditional volatility still have a significant effect on the expected return? Second, is the relationship between the expected return and the conditional volatility nonlinear? The first question is interesting because, once the volatility feedback is taken into account, the relationship between the expected return and the conditional volatility is the sum of the risk return trade-off effect and the volatility feedback effect. This sum can no longer be interpreted as a risk premium and there is no reason to expect the conditional volatility to positively influence the expected return. The answer to this question quantifies the predictive power of the conditional volatility for the returns, shedding some light on the sources of the return predictability. The second question is interesting because its answer addresses, to some extent, the issue of non-monotonicity in the relationship between the expected return and the conditional volatility. As Rossi and Timmermann (2010) point out, theoretical pricing models do not generally imply that this relationship is linear. Indeed, in the GARCH-in-mean literature, some authors specify that the expected return is linearly related to the conditional variance (see French et al (1987), Nelsen (1991), Campbell and Hentschel (1992), Glosten et al (1993) and Lundblad (2007). Others specify that it is linearly related to the conditional standard deviation (French eta al (1987), Jensen and Lunde (2001) and Yang (2011)). Adding to the previous studies, we examine this issue in a model where the volatility feedback effect is accounted for and the polynomial-type nonlinearity is allowed.

Our model is an extension of Yang (2011), where there are two effects with opposite signs, risk return trade-off and volatility feedback, for the conditional volatility to influence the expected return. The contemporaneous correlation (CC) between the return and volatility is interpreted as the effect of volatility feedback and the reasoning of volatility feedback implies that CC should be negative. We examine the empirical validity of this interpretation by exploring the fact that the volatility feedback relies on the risk return trade-off. The interpretation will be invalid if the CC is observed without the presence of the risk return trade-off. In the event that this interpretation is incorrect, the "volatility feedback effect" mentioned elsewhere in this paper is understood to be the effect of the CC.

Our model is built on a general premise that both the return and its volatility are contemporaneously influenced by news arrivals. We show that it leads to a reduced-form GARCH-in-mean model where the conditional distribution is a mixture of the distributions of the return shock and the volatility shock. Our benchmark model follows the normal inverse Gaussian (NIG) mixture of Barndorff-Nielsen (1997) with the conditional variance being the asymmetric power ARCH of Ding, Granger and Engle (1993). We demonstrate that it is the mixture structure (not necessarily NIG mixture) that facilitates the separate identification of the risk return trade-off and the volatility feedback. Our data set consists of the daily excess return series computed from the MSCI price indices of seven developed markets. From the estimation results, we find little evidence against the hypothesis that the effects of risk return trade-off and volatility feedback sum up to zero in the expected return and the hypothesis that the conditional mean is linearly related to the conditional volatility (standard deviation). For all the markets, except Japan, the risk return trade-off effect and the volatility feedback effect are individually significant with opposite signs. Further, in the empirical results, there is little evidence against the interpretation of the CC as the volatility feedback.

In the rest of this paper, Section 2 gives the details of our model. Section 3, with three subsections, contains data descriptions, estimation results and robustness checks. Concluding remarks are given in Section 4.

2. MODEL

The model below is an extension of Yang (2011). It covers a class of conditional distributions for the return and a class of nonlinear specifications for the conditional mean. Let x_t be the excess return (the change of log asset price minus the risk-free rate) of an asset at the end of date t and $\mathcal{F}_t$ be the information set generated by $\{x_t, x_{t-1}, \ldots\}$. The risk premium of the asset for given $\mathcal{F}_{t-1}$ is specified as

(1) $\mu_t = m_0 + m_1 h_t, \qquad h_t^2 = \mathrm{var}(x_t | \mathcal{F}_{t-1}).$

Here the coefficient m_1 is interpreted as the risk return trade-off effect, i.e., the effect in the absence of the contemporaneous correlation between the return shock and volatility shock (see below). This is our benchmark specification and its nonlinear extension will be considered at the end of this section. We allow news or information arrivals to influence both the return and its volatility. Let $s_t > 0$ be the innovation[3] to the volatility and be independent of $\mathcal{F}_{t-1}$. Define the instantaneous variance as $\sigma_t^2 = \mathrm{var}(x_t | \mathcal{F}_{t-1}, s_t) = c_2^2 h_t^2 s_t^2$, where the positive constant c_2 depends of the mean and variance of s_t^2 and will be determined shortly. The innovation s_t is labelled as the *volatility shock* since it controls the magnitude (large or small) of the unexpected price change. For example, for given $\mathcal{F}_{t-1}$, a large instantaneous volatility σ_t results from a large volatility shock s_t. On the other hand, the standardised score $\varepsilon_t = (x_t - \mu_t)/(c_2 h_t s_t)$ is called the *return shock* as it sets the direction (up or down) of the unexpected price change. In terms the volatility and return shocks, the excess return may be expressed as

(2) $x_t = \mu_t + c_2 h_t s_t \varepsilon_t,$

where μ_t and h_t are functions of the information set $\mathcal{F}_{t-1}$. As the return and volatility shocks are a manifestation of news arrivals, we assume that ε_t and s_t are independent of $\mathcal{F}_{t-1}$. To capture the possible correlation between ε_t and s_t, the return shock is decomposed as

(3) $\varepsilon_t = \xi_t + \beta s_t,$

[3] Alternatively, one may think of $\eta_t = \log(s_t) - E[\log(s_t)]$ as a "deeper level" shock that is centered at zero. In that case, the innovation s_t can be expressed as a constant multiple of $\exp(\eta_t)$.

where ξ_t is independent of s_t with $E(\xi_t) = 0$ and $\text{var}(\xi_t) = 1$. The constant parameter β dictates the sign of the covariance $\text{cov}(\varepsilon_t, s_t) = \beta\text{var}(s_t)$. The shocks ε_t and s_t, driven by unobserved news arrivals, represent the impacts of the news on the direction and the magnitude of the unexpected prices change. In this framework, the hypothesis that bad news is associated with high volatility can be formulated as $\beta < 0$.

By (2) and (3), the constant c_2 must satisfy

$$h_t^2 = \text{var}(x_t|\mathcal{F}_{t-1}) = \text{var}(c_2 h_t s_t \varepsilon_t|\mathcal{F}_{t-1}) = c_2^2 h_t^2 \text{var}(s_t \varepsilon_t)$$

$$= c_2^2 h_t^2 [\beta^2 \text{var}(s_t^2) + E(s_t^2)].$$

It follows that

(4) $$c_2 = 1/\sqrt{\text{var}(s_t \varepsilon_t)} = 1/\sqrt{\beta^2 \text{var}(s_t^2) + E(s_t^2)}.$$

Further, (2) and (3) imply

(5) $$x_t = \mu_t + \beta c_2 h_t s_t^2 + c_2 h_t s_t \xi_t,$$

which, conditional on $\mathcal{F}_{t-1}$, follows the mixture distribution defined by ξ_t and s_t^2. In particular, $x_t|(\mathcal{F}_{t-1}, c_2^2 s_t^2) \sim \text{N}[\mu_t + (\beta/c_2/h_t) c_2^2 h_t^2 s_t^2,\ c_2^2 h_t^2 s_t^2]$ when $\xi_t \sim \text{N}(0,1)$. From (5), the contemporaneous correlation between x_t and $\sigma_t^2 = c_2^2 h_t^2 s_t^2$ is found to be

(6) $$\text{corr}(x_t, \sigma_t^2|\mathcal{F}_{t-1}) = \beta/\sqrt{\beta^2 + E(s_t^2)/\text{var}(s_t^2)}$$

and its sign is determined by β. To be consistent with the usual ARCH-type model, we define the overall shock in (2) as the standardised $\varepsilon_t s_t$

(7) $$v_t = \left(\varepsilon_t s_t - E(\varepsilon_t s_t)\right)/\sqrt{var(\varepsilon_t s_t)} = c_2\left(\varepsilon_t s_t - \beta E(s_t^2)\right),$$

which obviously satisfies $E(v_t) = 0$ and $\text{var}(v_t) = 1$. It is the mixture of the distributions of ξ_t and s_t^2. By (2) and (7), the excess return in (2) may also be written as

(8) $$x_t = \mu_t + \beta c_2 E(s_t^2) h_t + h_t v_t$$

which is clearly a GARCH-in-mean model when h_t follows a GARCH process.

It can be seen that (3) and (5) nest some frequently-used conditional distributions. In particular, when $\beta = 0$, the conditional distribution of x_t in (5) is determined by the distribution of $c_2 s_t \xi_t$. When $\beta = 0$ and ξ_t is normal, $x_t|\mathcal{F}_{t-1}$ is normal if $s_t \equiv 1$; and $x_t|\mathcal{F}_{t-1}$is Student's t if s_t^2 follows the inverted Gamma-2 distribution. When $\beta \neq 0$, the setup here not only covers the mixture distributions used by Jensen and Lund (2001) and

Yang (2011) but also other types of mixtures. For instance, when $\ln(s_t^2) \sim N(0,\gamma)$ with a positive parameter γ and $\xi_t \sim N(0,1)$, the conditional distribution of x_t is a normal log normal mixture. These mixture distributions all have the capacity to capture the skewness and excess kurtosis in the data.

While there is a variety of choices for the distribution of (ξ_t, s_t^2), our benchmark specification in this paper is

(9) $\xi_t \sim iid\ \mathrm{N}(0,1), \qquad s_t^2 \sim iid\ \mathrm{IG}(1,\gamma),$

where IG is the inverse Gaussian distribution[4] with γ being a positive parameter. Under (9), it is known that $E(s_t^2) = 1/\gamma$, $\mathrm{var}(s_t^2) = 1/\gamma^3$ and $c_2 = \gamma^{3/2}/\sqrt{\gamma^2+\beta^2}$ (see Jensen and Lunde (2001) and Yang (2011) among others). For (9), the overall shock in (7) specialises to

(10) $v_t = c_2(s_t\varepsilon_t - \beta/\gamma) = -\beta c_1 + (\beta/c_2)c_2^2 s_t^2 + c_2 s_t \xi_t \quad \text{with} \quad c_1 = c_2/\gamma.$

The distribution of v_t is known as the normal inverse Gaussian (NIG) mixture with the *invariant* parameterisation[5]. We denote this distribution by $\mathrm{NIG}_I(\gamma,\ \beta,\ -\beta c_1,\ c_2)$, which is the mix of $v_t|c_2^2 s_t^2 \sim \mathrm{N}(-\beta c_1 + (\beta/c_2)c_2^2 s_t^2,\ c_2^2 s_t^2)$ and $c_2^2 s_t^2 \sim \mathrm{IG}(c_2,\ \gamma/c_2)$. The NIG mixture is a member of the generalised hyperbolic (GH) distribution of Barndorff-Nielsen (1997) and possesses desirable properties for fitting speculative return series. In particular, the skewness and kurtosis of v_t are given by (see Jensen and Lunde (2001)),

(11) $\mathrm{Skew}(v_t) = 3\beta/\sqrt{(\beta^2+\gamma^2)\gamma}, \qquad \mathrm{Kurt}(v_t) = 3 + 3(1 + 4\beta^2/(\beta^2+\gamma^2))/\gamma.$

For any finite γ and non-zero β, the NIG is thick-tailed and skewed, where the skewness direction is determined by the sign of β. Its capacity to accommodate skewness and thick tails in data is a major advantage over the normal and other symmetric distributions. Indeed, the normal distribution N(0,1) is the special case of the NIG when $\beta = 0$ and $\gamma \to \infty$. Further, the NIG is computationally easier to implement than other mixture distributions.

For the specification (9), the excess return in (8) becomes

(12) $$x_t = \mu_t + \beta c_1 h_t + h_t v_t$$
$$= m_0 + (m_1 + \beta c_1)h_t + h_t v_t, \qquad v_t \sim iid\ \mathrm{NIG}_I(\gamma, \beta, -\beta c_1, c_2),$$

[4] For $Y \sim \mathrm{IG}(\delta,\gamma)$, where δ and γ are positive parameters, its probability density function is $\mathrm{pdf}_Y(y) = (2\pi)^{-1/2}\delta^{3/2}y^{-3/2}\exp\{\delta\gamma - .5(\delta^2 y^{-1} + \gamma^2 y)\}$ with $E(Y) = \delta/\gamma$ and $\mathrm{var}(Y) = \delta^2/\gamma^3$. If $Y \sim \mathrm{IG}(\delta,\gamma)$, then $c^2 Y \sim IG(\delta c, \gamma/c)$ for any positive constant c.

[5] If $X|Y \sim \mathrm{N}(\mu + (\beta/\delta)Y, Y)$ and $Y \sim \mathrm{IG}(\delta, \gamma/\delta)$ for positive parameters δ and γ, then the marginal distribution of X is $\mathrm{NIG}_I(\gamma,\beta,\mu,\delta)$. Note that the parameters (γ,β) here are equivalent to the *invariant* parameters $(\bar{\gamma},\bar{\beta})$ of Barndorff-Nielsen (1997).

where $c_1 = c_2/\gamma$. While (12) is a GARCH-in-mean process when h_t is a GARCH process, the risk return trad-off effect (m_1) and the contemporaneous correlation effect (βc_1) are clearly separated in the conditional mean. To complete the model, we specify h_t as the asymmetric power ARCH (APARCH) of Ding *et al* (1993),

$$(13) \qquad h_t^{\vartheta} = \omega + a(|u_{t-1}| - \tau u_{t-1})^{\vartheta} + b h_{t-1}^{\vartheta}, \qquad u_t = h_t v_t,$$

where $(\omega, \tau, a, b, \vartheta)$ are parameters satisfying $\omega > 0$, $|\tau| < 1$, $a > 0$, $b \geq 0$ and $\vartheta > 0$. The parameter τ captures the asymmetric effect of u_{t-1} on h_t (negative shocks predict higher volatility than positive ones when $\tau > 0$). The merits of the APARCH include the capacities of: (a) capturing asymmetric responses of the conditional variance to u_{t-1}; (b) allowing a data-determined flexible functional form via the Box-Cox transformation parameter ϑ; (c) nesting a number of popular GARCH formulations such as that of Glosten et al (1993) (see Ding *et al* (1993), Jensen and Lunde (2001) and Yang (2011) among others). It can be verified from (5) and (9) that $x_t|\mathcal{F}_{t-1} \sim \mathrm{NIG}_I(\gamma, \beta, \mu_t, c_2 h_t)$ with the conditional density

$$(14) \qquad \mathrm{pdf}(x_t|\mathcal{F}_{t-1}) = \frac{\sqrt{\beta^2+\gamma^2}}{\pi c_2 h_t} q\left(\frac{x_t-\mu_t}{c_2 h_t}\right)^{-1} K_1\left(\sqrt{\beta^2+\gamma^2}\, q\left(\frac{x_t-\mu_t}{c_2 h_t}\right)\right) \exp\left(\gamma + \beta \frac{x_t-\mu_t}{c_2 h_t}\right),$$

where $q(\cdot) = \sqrt{1 + (\cdot)^2}$ and $K_1(\cdot)$ is the modified Bessel function of third kind with index 1. The conditional density (14), together with (13) and (1) is the basis for the maximum likelihood (ML) estimation of the parameters $(m_0, m_1, \omega, \tau, a, b, \gamma, \beta, \vartheta)$. Note that $c_1 = c_2/\gamma$ and $c_2 = \gamma^{3/2}/\sqrt{\gamma^2 + \beta^2}$ are not free parameters.

In (12), the total GARCH-in-mean effect is $(m_1 + \beta c_1)$, where m_1 represents the risk return trade-off and is expected to be positive as a high expected volatility should be compensated by a high expected return. The term βc_1 is induced by the contemporaneous correlation (CC) between the volatility shock and the return shock. The two parameters in $(m_1 + \beta c_1)$ are jointly identified by the conditional (on $\mathcal{F}_{t-1}$) mean of x_t and the shape of the conditional distribution of x_t. In general, restricting either of m_1 and β to be zero without justification may lead to an inconsistent estimate of the other. Hypotheses about m_1 and βc_1 can be easily tested in (12). In particular, we are interested in the hypothesis that the CC cancels out the risk return trade-off in the conditional mean: $m_1 + \beta c_1 = 0$. It is of interest because the sum $(m_1 + \beta c_1)$ cannot be regarded as the risk return trade-off effect and there is no reason to expect it to be positive. The test provides an insight on the contribution of the conditional volatility to the predictability of the return. To control for the predictability induced by other factors, we include the lag of x_t in the conditional mean for our empirical analysis.

Yang (2011) interprets a negative CC as the volatility feedback effect suggested by French *et al* (1987). The volatility feedback may be summarised as follows: (i) a large volatility shock induces an upward revision of the expected volatility, which in turn leads to an increase in the expected return and the discount rate for future cash flows; (ii) the increase in the discount rate reduces the present value of future cash flows and causes the current price to fall; (iii) consequently, price falls (bad news) tend to be contemporaneously associated with high volatility. Clearly, in (i), the volatility feedback relies on the return risk trade-off and the volatility clustering (i.e., a high/low volatility is likely followed by a high/low volatility). It follows that the volatility feedback cannot stand alone without the presence of the return risk trade-off. Hence, the validity of the interpretation of βc_1 as the volatility feedback effect can be empirically examined: the interpretation would be invalid for cases where $\beta c_1 < 0$ but $m_1 = 0$. In such cases, the phrase "volatility feedback effect" mentioned elsewhere should be understood to be the CC effect.

Recently, Rossi and Timmermann (2010) suggest that the relationship between the conditional volatility and the expected return may be non-monotonic such that the conditional volatility may have a positive (negative) impact on the expected return when the volatility is low (high). In our framework, non-monotonicity amounts to μ_t being a nonlinear function of h_t. To address this issue, we also extend (1) to the cases where μ_t is a polynomial of h_t

(15) $$\mu_t = m_0 + \sum_{i=1}^{l} m_i h_t^i, \qquad l = 2,3,$$

which can accommodate many different non-monotonic relationships between μ_t and h_t. To investigate the monotonicity issue and provide further evidence on the functional form between the expected return and the conditional volatility, we also test the hypothesis that $m_2 = m_3 = 0$.

3. EMPIRICAL RESULTS

3.1 DATA

We consider the returns of seven major developed markets including Canada, France, Germany, Italy, Japan, UK and US. The returns are based on the MSCI price indices of these markets. While treasury bill rates are usually used to proxy the risk free interest rates in the literature, they are not available for some markets. Table 1 lists the rates used. The lengths of these series are determined by the availability of interest rates as well as genuine daily observations on the MISC indices. All data are obtained from Datastream.

The summary statistics of the excess returns are listed in Table 2. The usual features of index return series are apparent: near-zero means, large standard deviations, large kurtosis, negative skewness, and fairly strong autocorrelations (large Ljung-Box Q-statistics). While the mean excess returns should be positive theoretically, that of Japan turns out to be negative for the periods considered (restricted by data availability). The magnitude of Japan's skewness is much smaller than other markets.

3.2 MAIN RESULTS

The model outlined in Section 2 is estimated for the seven markets. In this implementation, the lag of x_t is also included in the conditional mean to account for the autocorrelation that is not induced by the conditional volatility. The ML estimation results are presented in Table 3, where the standard errors are computed using the robust "sandwich" formula (see Theorems 3.2 and 3.4 of White (1982) and Bollerslev and Wooldridge (1992)). The standard errors of derived parameters βc_1 and $m_1 + \beta c_1$ are obtained using the "delta" method (see section 5.2.4 of Greene (2003)).

The usual characteristics in ARCH-type models for market indices are evident: clustering (a being positive), persistence (b being close to unity), asymmetric news curves (τ being positive), and fat-tails (γ being finite). Interestingly, the estimates of ϑ suggest that the original GARCH formulation $\vartheta = 2$ is strongly rejected by all the markets. In particular, the data does not reject the hypothesis of $\vartheta = 1$ for the US market. The Ljung-Box Q_{30}statistics for the standardised residuals and their squares in Table 3 are low in comparison to those of the returns in Table 1, indicating that the model is able to capture the main dynamic features of the data.

Regarding the risk return trade-off effect m_1, all the markets, except Japan, exhibit significantly positive estimates of m_1 at the 2% level of significance. These estimates are also economically significant in comparison against the sample means and standard deviations in Table 2. For Japan, the p-value of the estimate of m_1 is about 0.085. Regarding the contemporaneous correlation parameter β and the volatility feedback effect βc_1, the estimates from all the markets, except Japan, are significantly negative at the 1% level. For

Japan, the p-value of the estimate of βc_1 is about 0.112. Notice that for each of these markets, the estimated m_1 and βc_1 have opposite signs with similar magnitudes.

The null hypothesis $H_0: m_1 + \beta c_1 = 0$ cannot be rejected for all the markets. The magnitudes of $m_1 + \beta c_1$ are much smaller than those of m_1 or βc_1 for all the markets but Japan. The conditional volatility appears to have little predictive power for the expected return, as the risk return trade-off and the volatility feedback cancel out each other in the conditional mean. Further, as mentioned earlier, the cases with significant βc_1 and insignificant m_1 represent evidence against the interpretation of the CC as the volatility feedback. In Table 3, no estimates fall into this category and there is no evidence against this interpretation.

The effects of restricting either $\beta = 0$ (no volatility feedback) or $m_1 = 0$ (no risk return trade-off) are presented in Table 4. While these restrictions are rejected (except Japan) according to Table 3, their full impacts can only be seen in Table 4. The likelihood ratio (LR) statistics strongly reject $\beta = 0$ (or $m_1 = 0$) for all the markets but Japan and the restricted estimates of m_1 (or βc_1) are either insignificant or with much smaller magnitudes than the unrestricted estimates in Table 3. Further, the last column of Table 4 is about restricting the conditional distribution to be normal, where the huge LR statistics convincingly reject the normality restriction for all the markets. The main difference between our model and the models in the literature is that our model allows the contemporaneous correlation (interpreted as volatility feedback) between the volatility shock and return shock. The results in Table 4 demonstrate that omitting this allowance can have serious impact on the estimates of the return risk relationship.

The extended specification (15) is used to check whether or not the expected return is linearly related to the conditional volatility (standard deviation). The linearity is rejected if a large improvement in the log likelihood is observed for the extended specification. The estimation results for the case with $l = 2$ are presented in Table 5. Clearly, the main conclusions drawn from Table 4 also hold for Table 5. In fact, the estimates of m_2 are insignificant for all the markets and the LR tests on the restriction $m_2 = 0$ provide no support for the extended specification with $l = 2$. For the extension (15) with $l = 3$, the LR statistics for the restrictions $m_2 = m_3 = 0$ are presented in the last row of Table 5, which again do not support for the extended specification. This exercise confirms that the linear specification (1) is preferred for our data set and there is little evidence to support the nonlinear relationship specified in (15).

3.3 ROBUSTNESS CHECKS

In this subsection, we check the robustness of the results documented in Section 3.2 by estimating the models for the following variations on the benchmark model in Table 3, ceteris paribus.

Choice of Risk-free Rates

There are many possible but similar proxies for the risk-free rates. However, because the magnitudes of variations in the raw returns are much larger, the impact of the choice of risk-free rates on the main results is negligible. To confirm this claim, we estimate the model with the raw returns of the seven market indices and present the results in Table 6. Apparently, there are no major differences between Tables 3 and 6.

GARCH Forms

Among possible GARCH specifications, APARCH in (13) is used for the main results in Section 3.2 because of its flexibility. We also estimate the models with the GJR formulation (Glosten et al (1993)) and the original GARCH formulation, which amount to restricting $\vartheta = 2$ and $(\vartheta, \tau) = (2,0)$ respectively. The estimates for the key parameters $(m_1, \beta c_1)$ do not materially differ from those in Table 3 and therefore are not reported.

Distributional Specifications

The model described in Section 2 covers various specifications for the return shock ε_t and volatility shock s_t in (3), in addition to our choice (9) for the main results in Section 3.2. To check whether the main results in Table 3 are sensitive to the distributional specification, we consider the alternative specification $\xi_t \sim iid\ \mathrm{N}(0,1)$ and $\ln(s_t^2) \sim iid\ \mathrm{N}(0,\gamma)$. For this case, according to (4), $c_2 = 1/\sqrt{\beta^2 e^{\gamma}(e^{\gamma} - 1) + e^{\gamma/2}}$. The overall shock v_t in (7) follows the normal log normal (NLN) mixture $\mathrm{NLN}(\gamma, \beta, -\beta c_1, c_2)$ defined by

(16) $$v_t | c_2^2 s_t^2 \sim iid\ \mathrm{N}(-\beta c_1 + (\beta/c_2) c_2^2 s_t^2\ ,\ c_2^2 s_t^2), \qquad \ln(c_2^2 s_t^2) \sim iid\ \mathrm{N}(\ln(c_2^2)\,, \gamma),$$

where $c_1 = c_2 e^{\gamma/2}$. The expression $x_t = m_0 + (m_1 + \beta c_1) h_t + h_t v_t$ in (12) is valid for the c_2 and c_1 defined here. The estimation results for the NLN mixture is presented in Table 7, where the estimates for the key parameters $(m_1, \beta c_1)$ are qualitatively the same as those of Table 3. This exercise shows that it is the mixture structure (not necessarily the NIG mixture specification) that is important for characterising the impact of news arrivals on the return and for separating the volatility feedback effect from the risk return trade-off effect.

Time Aggregation

Although our focus is on the short-term risk return relationships, it is of interest to check whether or not the main short-term results obtained in Section 3.2 remain true for weekly data. The results for weekly excess return series (end-of-Friday return) are presented in Table 8. The main conclusion drawn from Table 3 largely hold for the weekly returns series except UK. For UK, the estimate of m_1 for the weekly series becomes less sharp with a p-value being about 0.083. Nonetheless, it appears that a moderate change in the data frequency does not alter our main conclusions in Section 3.2.

In summary, the results in Section 3.2 are insensitive to the variations considered here. The computation of all the empirical results is carried out in the R Environment (2011).

4. Conclusion

We analyse the relationship between the market index return and its volatility in a GARCH-in-mean model with the normal inverse Gaussian mixture distribution, taking into account the volatility feedback effect. The hypothesis that the volatility feedback effect cancels out the risk premium effect in the expected return cannot be rejected for the market index returns of seven major developed economies. The hypothesis that the relationship between the expected return and the conditional standard deviation is linear cannot be rejected either. We also consider a set of variations on the conditional mean, conditional variance, and conditional distribution. The main results are robust to these variations. For our data set, the results indicate that the conditional volatility has little predictive power for the expected return.

Our results are obtained primarily for daily return series and largely hold for weekly return series. Ghysels et al (2005) and Lundblad (2007) find positive relationships between the expected return and the conditional variance at the monthly frequency. It is of interest to investigate whether or not the volatility feedback effect exists in monthly return series. As there are ten parameters in our model, reasonably-accurate estimates require long time series (see Lundblad (2007)). For our data set, however, the length of the monthly return series is at most 371. For this reason, monthly return series are not analysed in the current paper.

5. REFERENCES

Barndorff-Nielsen, O.E. (1997), Normal inverse Gaussian distributions and stochastic volatility modelling, *Scandinavian Journal of Statistics*, 24(1), 1-12.

Brandt, M.W. and Q. Kang (2004), On the relationship between the conditional mean and volatility of stock returns: a latent VAR approach, *Journal of Financial Economics*, 72, 217-257.

Bollerslev, T. and J.M. Wooldridge (1992), Quasi-maximum likelihood estimation and inference in dynamic models with time varying covariances, *Econometric Reviews*, 11(2), 143-172

Campbell, J.Y. (1987), Stock returns and the term structure, *Journal of Financial Economics*, 18, 373-399.

Campbell, J.Y. and L. Hentschel (1992), No news is good news: an asymmetric model of changing volatility in stock returns, *Journal of Financial Economics*, 31, 281-318.

Chou, R. (1988), Volatility persistence and stock evaluations: some empirical evidence using GARCH, *Journal of Applied Econometrics*, 3, 279-294

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

French, K.R., G.W. Schwert and R.F Stambaugh (1987), Expected stock returns and volatility," *Journal of Financial Economics*, 19, 3-29.

Ghysels, E., P. Santa-Clara and R. Valkanov (2005), There is a risk return tradeoff after all, *Journal of Financial Economics*, 76, 509-548.

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

Greene, W. H. (2003), Econometric Analysis, 5th edition, Pearson Education, Inc, New Jersey

Harrison, P. and H. H. Zhang (1999), An investigation of the risk and return relation at long horizons, *Review of Economics and Statistics*, 81, 399-408

Jensen, M.B. and A. Lunde (2001), The NIG-S&ARCH model: a fat-tailed, stochastic and autoregressive conditional heteroskedastic volatility model, *Econometrics Journal*, 4, 319-342.

Lundblad, C.(2007), The risk return tradeoff in the long run: 1836-2003, *Journal of Financial Economics*, 85, 123-150.

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

Pagan, A. and Y. Hong (1991), nonparametric estimation and the risk premium, in Barnett, W., J. Powell and G. Tauchen (eds), *Nonparametric and Seminonparametric Methods in Econometrics and Statistics*, Cambridge University Press, Cambridge, UK.

R Development Core Team (2011), R: A language and environment for statistical computing, R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org

Rossi, A. and A. Timmermann (2010), What is the shape of the risk return relation? Working Paper

White, H. (1982), Maximum likelihood estimation of misspecified models, *Econometrica*, 50(1), 1-25

Yang, M. (2011), Volatility feedback and risk premium in GARCH models with generalized hyperbolic distributions, *Studies in Nonlinear Dynamics & Econometrics*, 15(3), 124-142, Article 6 (http://www.bepress.com/snde/vol15/iss3/art6).

Table 1. Sources of Interest Rates

Market	Interest Rate
Canada	CANADA TREASURY BILL 3 MONTH - MIDDLE RATE
France	FRANCE EU-FRANC 3M (FT/ICAP/TR) - MIDDLE RATE
Germany	GERMANY EU-MARK 3M (FT/ICAP/TR) - MIDDLE RATE
Italy	ITALY EURO-LIRE 3M (FT/ICAP/TR) - MIDDLE RATE
Japan	JAPAN AVG.TIME DEP. 10+ MIL.YEN 3MTH - MIDDLE RATE
United Kingdom	UK TREASURY BILL TENDER 3M - MIDDLE RATE
United States	US T-BILL SEC MARKET 3 MONTH (D) - MIDDLE RATE

Table 2. Summary Statistics

Market	Mean	StdDev	Skewness	Kurtosis	$Q_{30}(x)$	no.Obs	Start Date
Canada	0.009	1.055	-0.695	14.823	96.105	8173	1980-01-02
France	0.012	1.265	-0.286	9.742	94.009	8173	1980-01-02
Germany	0.013	1.316	-0.361	10.625	72.864	8173	1980-01-02
Italy	0.010	1.386	-0.241	8.434	157.323	8173	1980-01-02
Japan	-0.016	1.359	-0.110	9.298	72.615	5197	1991-05-30
UK	0.007	1.100	-0.496	13.066	131.352	6866	1985-01-04
USA	0.017	1.124	-1.218	31.875	71.677	8173	1980-01-02

All series end at the end of April 2011. $Q_{30}(x)$ is the Ljung-Box Q-statistics at lag 30 for the excess return.

Table 3. Results for Benchmark Model

$$x_t - \varphi x_{t-1} = m_0 + (m_1 + \beta c_1)h_t + h_t v_t, \qquad v_t \sim iid\ \mathrm{NIG}_I(\gamma, \beta, -\beta c_1, c_2),$$

$$h_t^{\vartheta} = \omega + a(|u_{t-1}| - \tau u_{t-1})^{\vartheta} + b h_{t-1}^{\vartheta}, \qquad u_t = h_t v_t,$$

$$c_2 = \gamma^{3/2}/\sqrt{\gamma^2 + \beta^2}, \qquad c_1 = c_2/\gamma.$$

	Canada	France	Germany	Italy	Japan	UK	USA
m_1	0.199	0.134	0.271	0.116	0.100	0.328	0.146
	0.045	0.056	0.055	0.047	0.058	0.084	0.043
βc_1	-0.150	-0.124	-0.240	-0.095	-0.060	-0.301	-0.113
	0.032	0.042	0.046	0.035	0.037	0.074	0.027
$m_1 + \beta c_1$	0.050	0.010	0.032	0.021	0.040	0.027	0.033
	0.033	0.038	0.032	0.032	0.049	0.041	0.035
γ	1.911	2.700	2.655	2.100	1.786	3.988	1.601
	0.206	0.390	0.458	0.257	0.246	0.787	0.187
β	-0.208	-0.205	-0.395	-0.139	-0.080	-0.609	-0.144
	0.049	0.073	0.096	0.053	0.051	0.190	0.037
ϑ	1.641	1.471	1.324	1.391	1.431	1.520	1.187
	0.154	0.142	0.108	0.117	0.152	0.194	0.103
τ	0.197	0.404	0.318	0.230	0.484	0.382	0.639
	0.047	0.055	0.051	0.039	0.085	0.071	0.077
a	0.074	0.082	0.085	0.094	0.082	0.074	0.059
	0.009	0.008	0.007	0.009	0.009	0.008	0.006
b	0.924	0.906	0.916	0.910	0.904	0.915	0.939
	0.010	0.010	0.008	0.009	0.012	0.010	0.007
ω	0.010	0.027	0.018	0.020	0.037	0.017	0.013
	0.002	0.005	0.003	0.005	0.008	0.003	0.003
m_0	-0.028	0.009	-0.009	-0.011	-0.063	-0.011	-0.011
	0.024	0.037	0.030	0.033	0.053	0.034	0.029
φ	0.128	0.066	0.019	0.085	0.031	0.021	-0.001
	0.011	0.011	0.011	0.011	0.013	0.012	0.010
logL	-9885.15	-12019.49	-12010.09	-12844.27	-8196.83	-9006.93	-10658.62
$\mathrm{Q}_{30}(v)$	41.68	47.86	56.49	65.87	23.07	37.74	39.34
$\mathrm{Q}_{30}(v^2)$	34.17	14.79	6.54	16.00	33.40	36.45	15.42

Here x_t is the excess return at the end of day t. The black rows are ML estimates of parameters. The blue rows are the standard errors computed with the robust "sandwich" formula (see White (1982) and Bollerslev and Wooldridge (1992)). The standard errors for βc_1 and $m_1 + \beta c_1$ are obtained by using the delta method (see section 5.2.4 of Greene (2003)). The "logL" row contains the log likelihood values. $\mathrm{Q}_{30}(v)$ and $\mathrm{Q}_{30}(v^2)$ are the Ljung-Box Q-statistics at lag 30 for the standardized residual (from the benchmark model) and squared standardized residual respectively.

Table 4. Restricted Estimation Results

	Restricted Benchmark Model: $\beta = 0$		Restricted Benchmark Model: $m_1 = 0$		Restricted Benchmark Model: $\beta = 0,\ \gamma = \infty$	
	m_1	LR-stat	βc_1	LR-stat	m_1	LR-stat
Canada	0.064	22.39	-0.052	21.34	0.035	434.82
	0.033		0.023		0.036	
France	0.017	9.73	-0.051	6.42	0.019	304.03
	0.038		0.028		0.048	
Germany	0.047	36.53	-0.073	30.46	0.034	510.79
	0.031		0.028		0.033	
Italy	0.028	7.95	-0.033	6.78	0.029	404.48
	0.031		0.024		0.035	
Japan	0.046	2.59	-0.023	2.97	0.031	224.82
	0.049		0.030		0.053	
UK	0.037	27.45	-0.082	22.30	0.027	213.06
	0.041		0.035		0.042	
USA	0.046	15.89	-0.052	12.29	0.051	517.74
	0.034		0.023		0.040	

The blue rows are the standard errors computed with the robust "sandwich" formula (see White (1982) and Bollerslev and Wooldridge (1992)). The likelihood ratio, LR-stat, under the restriction, is asymptotically distributed as $\chi^2(1)$ for the second and third column and $\chi^2(2)$ for the last column. The 5% and 1% critical values are 3.84 and 6.63 for $\chi^2(1)$ respectively. The 5% and 1% critical values are 5.59 and 9.21 for $\chi^2(2)$ respectively.

Table 5. Results for Extended Model

$$x_t - \varphi x_{t-1} = m_0 + (m_1 + \beta c_1)h_t + m_2 h_t^2 + h_t v_t, \qquad v_t \sim iid\ \mathrm{NIG}_I(\gamma, \beta, -\beta c_1, c_2),$$

$$h_t^{\vartheta} = \omega + a(|u_{t-1}| - \tau u_{t-1})^{\vartheta} + b h_{t-1}^{\vartheta}, \qquad u_t = h_t v_t,$$

$$c_2 = \gamma^{3/2}/\sqrt{\gamma^2 + \beta^2}, \qquad c_1 = c_2/\gamma.$$

	Canada	France	Germany	Italy	Japan	UK	USA
m_1	0.243	0.125	0.392	0.107	0.015	0.376	0.163
	0.092	0.123	0.118	0.094	0.101	0.127	0.072
βc_1	-0.151	-0.124	-0.238	-0.096	-0.061	-0.302	-0.113
	0.032	0.042	0.044	0.035	0.038	0.074	0.027
$m_1 + \beta c_1$	0.092	0.001	0.154	0.012	-0.046	0.074	0.050
	0.086	0.115	0.110	0.086	0.096	0.103	0.067
m_2	-0.019	0.003	-0.046	0.003	0.030	-0.020	-0.007
	0.035	0.041	0.041	0.031	0.029	0.042	0.027
γ	1.913	2.701	2.639	2.101	1.776	3.992	1.606
	0.208	0.391	0.442	0.258	0.245	0.794	0.186
β	-0.210	-0.205	-0.391	-0.139	-0.081	-0.610	-0.144
	0.049	0.073	0.090	0.052	0.052	0.192	0.037
ϑ	1.645	1.468	1.324	1.391	1.423	1.540	1.194
	0.154	0.149	0.106	0.117	0.154	0.204	0.106
τ	0.196	0.405	0.313	0.230	0.489	0.379	0.645
	0.047	0.057	0.050	0.039	0.087	0.072	0.077
a	0.073	0.082	0.085	0.094	0.082	0.073	0.059
	0.009	0.008	0.007	0.009	0.009	0.008	0.006
b	0.924	0.906	0.916	0.910	0.903	0.915	0.939
	0.010	0.010	0.008	0.009	0.012	0.010	0.007
ω	0.010	0.027	0.017	0.020	0.037	0.017	0.013
	0.002	0.005	0.003	0.005	0.009	0.003	0.003
m_0	-0.047	0.015	-0.077	-0.006	-0.008	-0.035	-0.019
	0.044	0.073	0.065	0.057	0.074	0.058	0.039
φ	0.128	0.066	0.019	0.085	0.030	0.021	-0.001
	0.011	0.011	0.011	0.011	0.013	0.012	0.010
logL	-9884.99	-12019.48	-12009.15	-12844.26	-8196.50	-9006.80	-10658.59
LR($m_2 = 0$)	0.33	0.01	1.89	0.01	0.67	0.26	0.06
LR($m_2 = m_3 = 0$)	0.33	1.10	4.42	0.31	2.42	3.34	0.07

Here x_t is the excess return at the end of day t. The black rows are ML estimates of parameters. The blue rows are the standard errors computed with the robust "sandwich" formula (see White (1982) and Bollerslev and Wooldridge (1992)). The standard errors for βc_1 and $m_1 + \beta c_1$ are obtained by using the delta method (see section 5.2.4 of Greene (2003)). The "logL" row contains the log likelihood values. The LR($\cdot$) rows are the likelihood ratio statistics for testing the restriction specified in the brackets. The LR($m_2 = 0$) statistic is asymptotically $\chi^2(1)$ distribution under the restriction. The LR($m_2 = 0, m_3 = 0$) statistic is asymptotically $\chi^2(2)$ distribution under the restriction.

Table 6. Results for Benchmark Model: Raw Return Series

Market	m_1	βc_1	$m_1+\beta c_1$	γ	β	ϑ	τ	a	b	ω
Canada	0.198	-0.152	0.046	1.909	-0.211	1.632	0.202	0.074	0.924	0.010
	0.045	0.032	0.033	0.207	0.049	0.153	0.048	0.009	0.010	0.002
France	0.129	-0.124	0.005	2.682	-0.204	1.471	0.404	0.082	0.906	0.027
	0.056	0.041	0.038	0.385	0.072	0.141	0.055	0.008	0.010	0.005
Germany	0.265	-0.240	0.025	2.644	-0.395	1.311	0.327	0.085	0.916	0.018
	0.055	0.045	0.031	0.450	0.094	0.109	0.052	0.007	0.008	0.003
Italy	0.123	-0.095	0.027	2.090	-0.138	1.408	0.222	0.094	0.909	0.020
	0.046	0.035	0.032	0.256	0.052	0.117	0.039	0.009	0.010	0.005
Japan	0.100	-0.061	0.039	1.779	-0.082	1.431	0.485	0.082	0.904	0.037
	0.059	0.038	0.049	0.245	0.052	0.154	0.085	0.009	0.012	0.008
UK	0.330	-0.302	0.028	4.000	-0.611	1.524	0.373	0.074	0.914	0.018
	0.085	0.075	0.041	0.796	0.194	0.194	0.070	0.008	0.010	0.003
USA	0.150	-0.112	0.038	1.596	-0.142	1.174	0.644	0.060	0.939	0.014
	0.041	0.027	0.033	0.184	0.036	0.104	0.078	0.006	0.007	0.003

Table 7. Results for Benchmark Model: Normal Log Normal Mixture

$$x_t - \varphi x_{t-1} = m_0 + (m_1 + \beta c_1)h_t + m_2 h_t^2 + h_t v_t, \qquad v_t \sim iid\ \mathrm{NLN}(\gamma, \beta, -\beta c_1, c_2),$$

$$h_t^{\vartheta} = \omega + a(|u_{t-1}| - \tau u_{t-1})^{\vartheta} + b h_{t-1}^{\vartheta}, \qquad u_t = h_t v_t,$$

$$c_2 = 1/\sqrt{\beta^2 e^{\gamma}(e^{\gamma} - 1) + e^{\gamma/2}}, \qquad c_1 = c_2 e^{\gamma/2}.$$

Market	m_1	βc_1	$m_1+\beta c_1$	γ	β	ϑ	τ	a	b	ω
Canada	0.197	-0.148	0.050	0.433	-0.133	1.638	0.195	0.073	0.924	0.010
	0.044	0.032	0.033	0.037	0.031	0.160	0.046	0.007	0.007	0.002
France	0.130	-0.121	0.009	0.323	-0.112	1.460	0.405	0.082	0.906	0.027
	0.054	0.041	0.038	0.033	0.039	0.137	0.055	0.007	0.008	0.004
Germany	0.269	-0.237	0.032	0.324	-0.221	1.335	0.313	0.085	0.916	0.017
	0.052	0.046	0.031	0.032	0.043	0.113	0.046	0.007	0.007	0.003
Italy	0.113	-0.093	0.020	0.401	-0.084	1.394	0.230	0.094	0.909	0.020
	0.045	0.035	0.032	0.035	0.031	0.124	0.038	0.008	0.008	0.004
Japan	0.094	-0.056	0.038	0.475	-0.050	1.432	0.483	0.082	0.904	0.037
	0.057	0.036	0.049	0.052	0.033	0.159	0.076	0.009	0.010	0.007
UK	0.329	-0.301	0.028	0.225	-0.288	1.528	0.379	0.073	0.915	0.017
	0.077	0.075	0.041	0.029	0.066	0.164	0.063	0.007	0.008	0.003
USA	0.141	-0.108	0.032	0.515	-0.096	1.187	0.640	0.058	0.940	0.013
	0.042	0.026	0.035	0.043	0.026	0.105	0.080	0.005	0.005	0.002

Table 8. Results for Benchmark Model: Weekly Excess Return Series

Market	m_1	βc_1	$m_1+\beta c_1$	γ	β	ϑ	τ	a	b	ω
Canada	0.354	-0.337	0.017	2.575	-0.553	1.499	0.267	0.097	0.880	0.117
	0.143	0.102	0.106	0.622	0.218	0.387	0.115	0.020	0.030	0.066
France	0.511	-0.556	-0.044	4.305	-1.197	1.305	0.341	0.100	0.874	0.151
	0.129	0.136	0.088	1.656	0.487	0.327	0.119	0.017	0.023	0.071
Germany	0.951	-0.959	-0.008	6.276	-2.602	1.550	0.294	0.121	0.846	0.202
	0.339	0.323	0.093	2.618	1.434	0.342	0.100	0.019	0.024	0.077
Italy	0.274	-0.266	0.008	2.920	-0.460	1.504	0.119	0.130	0.862	0.163
	0.130	0.105	0.087	0.781	0.211	0.340	0.070	0.021	0.026	0.092
Japan	0.397	-0.403	-0.006	4.605	-0.880	1.225	0.488	0.087	0.846	0.279
	0.235	0.151	0.206	1.709	0.418	0.358	0.235	0.050	0.118	0.318
UK	0.516	-0.527	-0.011	3.557	-1.035	1.365	0.631	0.085	0.875	0.144
	0.298	0.220	0.125	1.619	0.651	0.218	0.168	0.018	0.024	0.051
USA	0.694	-0.681	0.013	4.355	-1.504	0.856	0.586	0.100	0.883	0.079
	0.084	0.053	0.054	0.638	0.219	0.201	0.122	0.020	0.029	0.033