The Impact of Volatility Long Memory on Option Valuation: Component GARCH versus FIGARCH

Yintian Wang
Tsinghua University

January 10, 2008

Abstract

This paper aims to investigate the impact of volatility long memory on European option valuation. We compare two groups of GARCH models that allow for long memory: the component Heston-Nandi GARCH model developed in the first chapter, in which the volatility of returns consists of a long-run and a short-run component; and a fractionally integrated Heston-Nandi GARCH model based on Baillie, Bollerslev and Mikkelsen (1996). We empirically investigate the models using S&P 500 index returns and cross-sectional European options data. The component GARCH model slightly outperforms the FIHNGARCH in fitting return data but significantly dominates the FIHNGARCH in capturing option prices. This is due to the shorter memory of the FIHNGARCH model, which, in turn, is attributable to the artificially prolonged leverage effect resulting from fractional integration and limitations of the affine structure.

Keywords: option valuation, long memory, long-run component, short-run component

1 Introduction

It has been widely reported that many financial and macroeconomic time series have a highly persistent volatility. See, for example, Briedt, Crato and de Lima (1998), Ding, Granger, and Engle (1993), and Harvey (1993). Andersen, Bollerslev, Diebold and Labys (2003) confirmed this finding using realized volatility. One approach to model persistent volatility, proposed by Baillie, Bollerslev and Mikkelsen (1996) and Bollerslev and Mikkelsen (1996) is to incorporate long-memory fractional differencing into the GARCH model. The ensuing model is called the fractionally integrated GARCH model or the FIGARCH model. Comte and Renault (1998) developed a fractionally integrated stochastic volatility model. The main characteristic of a FIGARCH model is that conditional variances exhibit not only short-run dynamics of the ARMA type, as in the standard GARCH model, but also long-run persistence that decays slowly at hyperbolic rates.

The literature on GARCH variance component models is rapidly expanding. Component GARCH models, which where first proposed by Engle and Lee (1993), constitute a convenient method of incorporating long-memory-like features into a short-memory model, at least for the horizons relevant for option valuation. Maheu (2002) presented Monte Carlo evidence that a component model can capture long-range volatility dynamics. Adrian and Rosenberg (2005) demonstrated the relevance of the component volatility structure for cross-sectional asset pricing. The fact that GARCH component variance models are also related to stochastic volatility component models has received empirical support; see Alizadeh, Brandt and Diebold (2002), Chernov, Gallant, Ghysels and Tauchen (2003), and Taylor and Xu (1994) for examples.

Given the empirical support for these volatility long-memory models in fitting S&P 500 index returns, it is natural to apply them to derivative pricing. Bollerslev and Mikkelsen (1996, 1999) and Comte, Coutin and Renault (2001) investigated and discussed the implications of fractionally integrated volatility for option valuation. While they use Monte Carlo simulation to illustrate the differences in European option prices for five alternative volatility dynamics, no empirical evidence was presented regarding the performance of a FIGARCH model in fitting option prices. Christoffersen, Jacobs and Wang's working paper, which is refered as CJW(2007) in the following context, found the component models significantly superior to the GARCH(1,1) model in capturing European option prices even if the latter model turns in a very solid empirical performance. Since both the FIGARCH model and the component GARCH model are designed to capture the long memory of volatility, it is of interest to compare both models theoretically and empirically.

In this paper, we develop a fractionally integrated Heston-Nandi GARCH model which allows for easier valuation of European options. We derive an approximate closed form option valuation formula and investigate the impact of long memory for option pricing. In addition, we characterize key properties of the model, including the conditional term structure across maturities, and conditional leverage and variance of variance paths. We discern important differences between the fractionally integrated Heston-Nandi GARCH model and the component Heston-Nandi GARCH model developed in CJW(2007). Please note that we refer to the fractionally integrated Heston-Nandi GARCH model as FIHNGARCH, and refer to the component Heston-Nandi GARCH as component GARCH. Both models are estimated using maximum likelihood estimation on S&P 500 returns, and their empirical performance is compared in terms of fitting historical returns and cross-sectional option data. Specifically, we compare two structures that capture the long memory of volatility: hyperbolic decay and exponential decay. Our results show that both the likelihood criterion and the option pricing errors strongly favor the component models.

The remainder of the paper is structured as follows. In Sections 2 and 3 we introduce the new FIHNGARCH model as well as the GARCH component models. Section 2 gives a brief review of the component model in CJW(2007) and its related properties. Section 3 introduces the fractionally integrated Heston-Nandi GARCH model, derives a number of its properties, and discusses option valuation for this component dynamic. Section 4 presents empirical model comparisons based on both the maximum likelihood estimation of returns and the root mean squared errors from valuing options on the S&P 500 index. Finally, Section 5 concludes.

2 The Component Heston-Nandi GARCH Model

2.1 Return Dynamics

The component GARCH model is an extension of a Heston-Nandi GARCH (1,1) model. The Heston-Nandi (2000) model is designed with option valuation in mind. Like the Heston (1993) model, it contains a leverage effect, allows for volatility clustering, and leads to a closed-form solution due to its affine structure. Heston and Nandi (2000) demonstrated how their model performs satisfactorily relative to ad-hoc benchmarks for the purpose of option valuation. This paper uses their model as an initial starting point. The model is

$$\begin{aligned} R_{t+1} &\equiv \ln\frac{S_{t+1}}{S_t} = r + \lambda h_{t+1} + \sqrt{h_{t+1}} z_{t+1} \qquad (1) \\ h_{t+1} &= w + b h_t + a\left(z_t - c\sqrt{h_t}\right)^2 \end{aligned}$$

where S_{t+1} denotes the underlying asset price, r the risk-free rate, λ the price of risk, z_t the *i.i.d.* return innovation with zero mean and unit variance, and h_{t+1} the daily variance on day $t+1$ which is known at the end of day t.

The unconditional variance is

$$\sigma^2 \equiv E\left(h_{t+1}\right) = \frac{w+a}{1-b-ac^2}$$

We can rewrite the conditional variance as

$$h_{t+1} = \sigma^2 + b\left(h_t - \sigma^2\right) + a\left((z_t - c\sqrt{h_t})^2 - \left(1 + c^2\sigma^2\right)\right) \tag{2}$$

The component GARCH model is obtained by replacing the constant σ^2 with a time-varying long-run component q_{t+1}. The conditional variance h_{t+1} now varies around a long-run component which is, itself, autoregressive of the first order. Using Greek letters for component model parameters, we write

$$\begin{aligned} h_{t+1} &= q_{t+1} + \widetilde{\beta}(h_t - q_t) + \alpha h_t v_{1,t} \\ q_{t+1} &= \omega + \rho q_t + \varphi h_t v_{2,t} \end{aligned} \tag{3}$$

where $v_{i,t} = (z_t^2 - 1) - 2\gamma_i z_t \sqrt{h_t}$ for $i = 1, 2$ can be viewed as zero-mean innovations to the volatility components.

We will assume that the *i.i.d.* return innovation z_t follows the standard normal distribution. We also derive a number of properties; these are key for understanding both Heston-Nandi GARCH(1,1) and the component counterpart's ability to capture the salient features of speculative returns and to fit option prices. To save space, we only illustrate the properties for the component GARCH model. Please see CJW(2007) for more details.

2.2 Variance Term Structures

Following CJW(2007), we define two measures of the variance term structure. One convenient measure denotes a cumulative k-days ahead forecast of variances divided by the unconditional variance.

$$\frac{h_{t+1:t+K}}{\sigma^2} \equiv \frac{\frac{1}{K}\sum_{k=1}^{K} E_t\left[h_{t+k}\right]}{\sigma^2} = 1 + \frac{\frac{q_{t+1}}{\sigma^2} - 1}{K}\frac{1-\rho^K}{1-\rho} + \frac{\frac{h_{t+1}-q_{t+1}}{\sigma^2}}{K}\frac{1-\tilde{\beta}^K}{1-\tilde{\beta}} \tag{4}$$

where σ^2 is the unconditional variance. This measure succinctly captures important information about the model's potential to explain the variation of option values across maturities. We can also learn about the dynamics of the variance term structure through impulse response functions, which are defined as

$$\partial E_t\left[h_{t:t+K}\right]/\partial z_t^2 = \frac{\alpha\left(1 - \gamma_1\sqrt{h_t}/z_t\right)}{K}\frac{1-\tilde{\beta}^K}{1-\tilde{\beta}} + \frac{\phi\left(1 - \gamma_2\sqrt{h_t}/z_t\right)}{K}\frac{1-\rho^K}{1-\rho} \tag{5}$$

The latter equation measures the effect of a shock at time t, z_t on the expected k-days ahead variance. Both measures estimate the persistence of variances.

2.3 Conditional Leverage and Variance of Variance

To assess the asymmetric response of volatility to positive versus negative return shocks, we derive the conditional covariance, $Cov_t\left(R_{t+1}, h_{t+2}\right)$, and refer to it as the conditional leverage effect. For the component model, the conditional leverage effect is given by

$$Cov_t\left(R_{t+1}, h_{t+2}\right) = -2\left(\alpha\gamma_1 + \varphi\gamma_2\right) h_{t+1} \tag{6}$$

We define the conditional variance of variance as $Var_t\left(h_{t+2}\right)$, which is given by

$$Var_t\left(h_{t+2}\right) = 2\left(\alpha + \varphi\right)^2 + 4\left(\gamma_1\alpha + \gamma_2\varphi\right)^2 h_{t+1} \tag{7}$$

Given the simple structure of the component model, it is easy to see that the magnitudes of both the conditional leverage and variance of variance are positively related to the leverage parameters γ_1, γ_2 and c. The relationship suggests that the leverage effect built in the model not only introduces negative skewness but also a more volatile variance dynamic.

3 An Affine FIGARCH Model

3.1 Return Dynamics

Just like fractionally integrated ARFIMA models generalize the standard ARIMA models, Baillie, Bollerslev and Mikkelsen (1996) introduced a new class of fractionally integrated GARCH models that generalize GARCH models. Analogous to the ARFIMA class of models for the conditional mean, a shock to the conditional variance in the FIGARCH model is transitory, in the sense that the influence on the forecast of the future conditional variance recedes at a slow hyperbolic rate of decay. The authors further extended the basic FIGARCH model to FIEGARCH to allow for the leverage effect. However, neither of the two models yields an analytical form for European option prices. To simplify the valuation of European options, we develop a new FIHNGARCH model based on the Heston-Nandi structure, which accommodates approximate closed formulae for European options.

First, we rewrite the Heston-Nandi GARCH(1,1)

$$\begin{aligned} R_{t+1} &\equiv \ln\frac{S_{t+1}}{S_t} = r + \eta h_{t+1} + \sqrt{h_{t+1}} z_{t+1} \\ h_{t+1} &= \omega_1 + \beta_1 h_t + \alpha_1\left(z_t - \gamma_1\sqrt{h_t}\right)^2 \end{aligned}$$

into

$$(z_t - \gamma_1\sqrt{h_t})^2\left(1 - \varphi_1 L\right) = 1 + \gamma_1^2\omega_1 - \beta_1 + \left(1 - \beta_1 L\right) v_t \tag{8}$$

where $v_t = \left(z_t - \gamma_1\sqrt{h_t}\right)^2 - \left(1 + \gamma_1^2 h_t\right)$ and $\varphi_1 = \beta_1 + \gamma_1^2\alpha_1$. Please note that, to avoid notational confusion, we use η to represent the risk price in the fractional integration GARCH model. Equation (8) is readily interpreted as an ARMA model for $(z_t - \gamma_1\sqrt{h_t})^2$. Analogously to the ARFIMA(k,d,l) process, a FIHNGARCH(p,d,q) process is naturally defined by

$$(z_t - \gamma_1\sqrt{h_t})^2\,(1-\varphi_1 L)\,(1-L)^d = 1 + \gamma_1^2\omega_1 - \beta_1 + (1-\beta_1 L)\,v_t \tag{9}$$

An alternative representation is

$$\begin{aligned}(1-\beta_1 L)\left(1+\gamma_1^2 h_t\right) &= \left(1+\gamma_1^2\omega_1 - \beta_1\right) + ... \\ &\quad + \left(1 - \beta_1 L - (1-\varphi_1 L)\,(1-L)^d\right)(z_t - \gamma_1\sqrt{h_t})^2\end{aligned} \tag{10}$$

The fractional differencing operator is defined by its Maclaurin series expansion. In order to better comprehend the statistical properties of this model, we rewrite the FIHNGARCH(p,d,q) model in terms of the observationally equivalent infinite ARCH representation,

$$\begin{aligned}h_t &= \frac{\omega_1}{1-\beta_1} + \left(\frac{1 - \frac{(1-\varphi_1 L)(1-L)^d}{(1-\beta_1 L)}}{\gamma_1^2}\right)(z_t - \gamma_1\sqrt{h_t})^2 \\ &= \varpi + \frac{\lambda(L)}{\gamma_1^2}(z_t - \gamma_1\sqrt{h_t})^2\end{aligned} \tag{11}$$

where $\varpi = \frac{\omega_1}{1-\beta_1}$, $\lambda(L) = \lambda_1 L + \lambda_2 L^2 + ...$ Please note that, in this model, γ_1 is not only the leverage parameter, but also appears in the denominator of the infinite ARCH coefficients that adjusts the magnitude of innovations impacting on conditional variance. The ARCH parameters in the lag polynomial $\lambda(L)$ can be written as

$$\begin{aligned}\lambda_1 &= (\varphi_1 - \beta_1 + d) = \alpha_1\gamma_1^2 + d \\ \lambda_k &= \beta_1\lambda_{k-1} + \left((k-1-d)\,k^{-1} - \varphi_1\right)\delta_{d,k-1} \text{ for } k \geq 2\end{aligned} \tag{12}$$

where

$$\begin{aligned}\delta_{d,1} &= d \\ \delta_{d,k} &= \delta_{d,k-1}k^{-1}\,(k-1-d) \quad \text{for } k \geq 2\end{aligned} \tag{13}$$

$\left(1 - \frac{(1-\varphi_1 L)(1-L)^d}{(1-\beta_1 L)}\right)$ evaluated at $L = 1$ equals zero, so that $\sum_{i=1}^{\infty}\lambda_i = 1$. The second moment of the unconditional distribution in the FIHNGARCH(p,d,q) model, therefore does not exist in the case of a positive ϖ, and R_{t+1} is not covariance-stationary. This feature is shared by an integrated GARCH (IGARCH) model when $d = 1$. Neither (11) nor an IGARCH model satisfy the sufficient conditions developed by Giraitis, Kokoszka and Leipus (2000) for covariance stationarity. However, Nelson (1990) showed that the IGARCH(1,1), which was extended to the general IGARCH(p,q) by Bougerol and Picard (1992), is strictly stationary and ergodic. Baillie, Bollerslev and Mikkelsen (1996) posited that the high-order lag coefficients in the infinite ARCH representation of any FIGARCH model may be dominated in an absolute value sense by the corresponding IGARCH coefficients. Therefore, a direct extension of the proofs for the IGARCH case can reveals that the FIGARCH(p,d,q) and FIHNGARCH models in our case are strictly stationary and ergodic for $0 \leq d \leq 1$. Please see Nelson (1990) for more details.

In the ARFIMA class of models, the short-run behavior of the time series is captured by the conventional ARMA parameters, while the long-run dependence is conveniently modeled through the fractional differencing parameter d. A similar result may well hold when modeling conditional variances. A shock to the optimal forecast of the future conditional variance decays at an exponential rate for the covariance-stationary GARCH(p,q) model, and remains important for forecasts of all horizons for the IGARCH(p,q) model. In contrast, in the FIGARCH(p,d,q) model, the effect of a shock to the forecast of the future conditional variance will die out at a slow hyperbolic rate. The fractional differencing parameter is therefore identifiable by the decay rate of a shock to the conditional variance, and not by the ultimate impact on the forecast for the long-run conditional variance.

3.2 Variance Term Structures

We again define the variance term structure

$$\frac{h_{t+1:t+k}}{\sigma^2} = \frac{1}{K}\frac{1}{\sigma^2}\sum_{k=1}^{K} E_t\left(h_{t+k}\right) \tag{14}$$

where

$$E_t h_{t+k} = \overline{\omega} + \sum_{i=1}^{k-1}\frac{\lambda_i}{\gamma_1^2}\left(1+\gamma_1^2 E_t h_{t+k-i}\right) + \sum_{i=k}^{\infty}\frac{\lambda_i}{\gamma_1^2}\left(z_{t-i+k} - \gamma_1\sqrt{h_{t-i+k}}\right)^2$$

and the impulse response functions are

$$\begin{aligned} \partial E_t\left[h_{t:t+K}\right]/\partial z_t^2 &= \sum_{i=1}^{K}\frac{\partial E_t h_{t+i}}{\partial z_t^2} \\ \frac{\partial E_t h_{t+k}}{\partial z_t^2} &= \sum_{i=1}^{k-1}\frac{\lambda_i}{\gamma_1^2}\frac{\partial E_t h_{t+k-i}}{\partial z_t^2} + \frac{\lambda_k}{\gamma_1^2}\left(1-\gamma_1\frac{\sqrt{h_t}}{z_t}\right) \\ \frac{\partial E_t h_{t+1}}{\partial z_t^2} &= \frac{\lambda_1}{\gamma_1^2}\left(1-\gamma_1\frac{\sqrt{h_t}}{z_t}\right) \end{aligned} \tag{15}$$

3.3 Conditional Leverage and Variance of Variance

For the FIHNGARCH model, the conditional variance of variance and the conditional leverage effect are given by

$$\begin{aligned} Var_t(h_{t+2}) &= E_t\left[h_{t+2} - E_t\left[h_{t+2}\right]\right]^2 \\ &= \left(2 + 4\gamma_1^2 h_{t+1}\right)\frac{\lambda_1^2}{\gamma_1^4} \end{aligned} \tag{16}$$

$$\begin{aligned} Cov_t(\ln\left(S_{t+1}\right), h_{t+2}) &= E_t\left[\left(\ln\left(S_{t+1}\right) - E_t\left[\ln\left(S_{t+1}\right)\right]\right)\left(h_{t+2} - E_t\left[h_{t+2}\right]\right)\right] \\ &= E_t\left[\sqrt{h_{t+1}}z_{t+1}\frac{\lambda_1}{\gamma_1^2}\left(z_{t+1}^2 - 2\gamma_1 z_{t+1}\sqrt{h_{t+1}} - 1\right)\right] \\ &= -2\frac{\lambda_1}{\gamma_1}h_{t+1} \end{aligned} \tag{17}$$

In contrast to the component GARCH model, the magnitudes of the conditional leverage and the variance of variance are both nonlinear in the leverage parameter γ_1.

3.4 The Autocorrelation Function for the Squared Innovation

We also provide the ACF of squared innovations for the FIHNGARCH model. In essence, this measure recounts the same story as the variance term structures about volatility persistence.

$$Corr_t(\varepsilon_{t+1}^2, \varepsilon_{t+k}^2) = \frac{Cov_t(\varepsilon_{t+1}^2, h_{t+k})}{\sqrt{Var_t\left[\varepsilon_{t+1}^2\right]}\sqrt{Var_t\left[\varepsilon_{t+k}^2\right]}} \tag{18}$$

where

$$\begin{aligned} Cov_t(\varepsilon_{t+1}^2, h_{t+k}) &= \frac{2\lambda_{k-1}}{\gamma_1^2}h_{t+1} \\ Var_t\left[\varepsilon_{t+1}^2\right] &= 2h_{t+1}^2 \\ Var_t\left[\varepsilon_{t+k}\right] &= 2\begin{pmatrix} \Delta^2 + 2\Delta\sum_{i=1}^{k-1}\frac{\lambda_i}{\gamma_1^2}\left(1 + \gamma_1^2 E_t h_{t+k-i}\right) + ... \\ \sum_{i=1}^{k-1}\frac{\lambda_i^2}{\gamma_1^4}\left(4 + 8\gamma_1^2 E_t h_{t+k-i} + \gamma_1^4 E_t h_{t+k-1}^2\right) \end{pmatrix} \end{aligned} \tag{19}$$

and

$$\Delta = \overline{\omega} + \frac{1}{\gamma_1^2}\left(\lambda_k L^k + \lambda_{k+1}L^{k+1} + ...\right)\left(z_{t+k} - \gamma_1\sqrt{h_{t+k}}\right)^2$$

3.5 Risk Neutralization and Option Valuation

As in CJW(2007), we assume Duan (1995)'s Locally Risk-Neutral Valuation Relationship assumption. In the risk-neutral world, the asset price S_t follows

$$\begin{aligned} \ln(S_{t+1}) &= \ln(S_t) + r - 0.5h_{t+1} + \sqrt{h_{t+1}} z_{t+1}^* \\ h_t &= \overline{\omega} + \frac{\lambda(L)}{\gamma_1^2}(z_t^* - \gamma_1^* \sqrt{h_t})^2 \end{aligned} \tag{20}$$

where z_t^* is standard normally distributed in a risk-neutral world, and $\gamma_1^* = \gamma_1 + 0.5 + \eta$. Given the risk-neutral dynamics, option valuation is straightforward. A European call option with strike price K that expires at time T, is worth

$$\text{Call Price} = e^{-r(T-t)} E_t^*[Max(S_T - K, 0)]$$

$$\begin{aligned} &= \frac{1}{2}S_t + \frac{e^{-r(T-t)}}{\pi}\int_0^\infty \text{Re}\left[\frac{K^{-i\phi} f^*(t,T;i\phi+1)}{i\phi}\right] d\phi... \\ &\quad -Ke^{-r(T-t)}\left(\frac{1}{2} + \frac{1}{\pi}\int_0^\infty \text{Re}\left[\frac{K^{-i\phi} f^*(t,T;i\phi)}{i\phi}\right] d\phi\right) \end{aligned} \tag{21}$$

where $f_t(\phi) = E_t\left[S_T^\phi\right]$ is the generating function, which is also the moment-generating function of the logarithm of S_T. Let $f_t^*(\phi)$ denote the conditional-generating function of the asset price in the risk-neutral world, In the Appendix, we show that the generating function takes the form

$$\begin{aligned} f_t &= E_t \exp(\phi \ln(S_T)) \\ &= \exp\left(\phi x_t + A_t + B_t h_{t+1} + \Lambda_t(L)\left(z_{t+1} - \gamma_1\sqrt{h_{t+1}}\right)^2\right) \end{aligned} \tag{22}$$

where $x_t = \ln(S_t)$. The coefficients $\{A_t, B_t, \Lambda_{t,1}, \Lambda_{t,2}, \Lambda_{t,3}, ...\}$ depend on the parameters of the model. Appendix A displays that the coefficients in the moment-generating function are

$$\begin{aligned} A_t &= \phi r + A_{t+1} + B_{t+1}\overline{\omega} - \frac{1}{2}\ln\left(1 - 2\left(B_{t+1}\overline{\lambda}_1 + \Lambda_{t+1,1}\right)\right) \\ B_t &= -0.5 + \gamma_1 + \frac{\left(B_{t+1}\overline{\lambda}_1 + \Lambda_{t+1,1}\right)\gamma_1^2 + \frac{\phi^2}{2} - \phi\gamma_1}{1 - 2\left(B_{t+1}\overline{\lambda}_1 + \Lambda_{t+1,1}\right)} \\ \Lambda_{t,1} &= B_{t+1}\overline{\lambda}_2 + \Lambda_{t+1,2} \\ \Lambda_{t,2} &= B_{t+1}\overline{\lambda}_3 + \Lambda_{t+1,3} \\ \Lambda_{t,n} &= B_{t+1}\overline{\lambda}_{n+1} + \Lambda_{t+1,n+1} \end{aligned} \tag{23}$$

where $\overline{\lambda}_i = \frac{\lambda_i}{\gamma_1^2}$, and n goes to infinity. The terminal conditions are

$$A_T = B_T = \Lambda_{T,1} = 0$$

One important feature is that the MGF can also be written as an infinite-weighted combination of shocks. In the evaluation of European options, a truncation of 1000 is employed as in the maximum likelihood estimation.

4 Empirical Results

This section presents the empirical results. While a formal proof of consistency and asymptotic normality of the MLE estimates of the FIGARCH process remains an outstanding issue, Baillie, Bollerslev and Mikkelsen (1996) assessed the practical applicability and small sample performance of the MLE procedure for the estimation of FIGARCH processes through a detailed simulation study. The simulations indicate that MLE is reasonably accurate.[1] Although no numerical or analytical investigation has been undertaken on FIGARCH models with leverage effects, it is still worthwhile attempting maximum likelihood estimation in our settings. To better understand the performance of FIHNGARCH, we add one additional benchmark, the Heston Nandi GARCH(1,1) model, for purpose of comparison.[2] We carry out maximum likelihood estimation for the three models on a long time series of S&P 500 return data. Then, we discuss the parameter estimates and their implications for the salient properties of the models.

4.1 Parameter Estimates from Daily Return Data

Panel A of Table 2 presents the maximum likelihood estimation results obtained using daily return data from June 1966 through December 31, 2001. The return data are obtained from CRSP. Standard errors are calculated from the outer product of the gradient and are given in parentheses. Since the fractional differencing operator is designed to capture the long-memory features of the process, truncating at too low a lag may destroy important long-run dependencies. For the estimation results reported here, we fixed the truncation lag at 1000, about four years' observations.

First, almost all parameters are estimated significantly different from zero at conventional significance levels. In terms of fit, the log likelihood values indicate that the fit of the component model is slightly superior to that of the FIHNGARCH model, which in turn fits better than the GARCH(1,1) model. We compute the test statistics in Vuong (1989), which

[1] The accuracy is evaluated through the simulated bias, root mean squared error, average estimated standard error of the QMLE, and the simulated rejection frequencies for the t-tests across 500 replications.

[2] For related properties of the Heston-Nandi GARCH(1,1) model, please see Chapter 1.

are designed to compare the goodness of fit of models when neither competing model is nested into the other. In our case, the standard normal statistic of 0.522 suggests that the component GARCH does not significantly dominate FIHNGARCH.

In the FIHNGARCH model, the estimate of β_1 is 0.664, lower than the 0.766 measured in the component model. This lower β_1, in turn, induces a lower short-run persistence $\varphi_1 = \beta_1 + \alpha_1 \gamma_1^2 = 0.5355$. We know that the short-run parameter ϕ_1 measures the persistence of the shocks over a relative short horizon, while the parameter d governs the long memory of shocks. Therefore, it is intuitive that with the introduction of the fractional differencing parameter d, volatility persistence is mostly governed by the long-run persistence parameter and, hence, the short-run persistence need not be as high as before. This finding is consistent with the previous literature. In contrast, the d value given by the model is lower than the estimates obtained in earlier research which are usually over 0.4.

Another interesting feature is that the estimate of α_1 is $-1.240E-05$. Although a positive α_1 is sufficient to guarantee the non-negativity of the conditional variance, this is not necessary the case when the parameters λ_i are positive for all i.

Panel A of Table 2 also presents unconditional summary statistics for different models. For the component model, the unconditional variance of variance is computed using the estimate for the unconditional variance in the expressions for the conditional moments (7). For the FIHNGARCH model, the unconditional volatility and the unconditional volatility of variance are undefined. To facilitate a comparison, we take the average of the conditional variance and then compute the standard deviation of variance based on the conditional moment in (16). To allow a comparison of the unconditional leverage for models, we report the moments in (6) and (17) divided by h_{t+1}. Overall, the leverage and the volatility of variance of the component GARCH model are greater in absolute value than those of the GARCH(1,1) model, while the FIHNGARCH model generates more leverage and more volatile variance than the component model.

4.2 Dynamic Model Properties

Figure 1 plots the conditional variances for the period 1990-1996. This period includes the dates for the option valuation exercise that are presented in Section 4.3. Notice that the conditional variance patterns across the three GARCH models display numerous similarities; the models all capture the low variances during the equity market run-up in 1993-1996, preceded by higher volatility during the first Gulf War and the 1990-1991 recession. However, Figure 1 also reveals differences between the models. The FIHNGARCH model appears to display slightly more variation in the conditional variance in the more recent past. We plot the conditional variance of variance path, $Var_t(h_{t+2})$ for each model. Figure 2 confirms the findings in Figures 1. The FIHNGARCH model displays a larger variance of variance than the component GARCH and the GARCH(1,1). Figure 3 plots the conditional leverage path, $Cov_t(R_{t+1}, h_{t+2})$ for each model under consideration. Note that the FIGARCH model has a larger (more negative) and more volatile leverage effect than the other two models. This is consistent with the higher unconditional levels presented in Table 2.

Figures 4a and Figure 4b plot the impulse responses to the term structure of variance for $h_t = \sigma^2$ and $z_t = 2$ and $z_t = -2$, respectively, as defined in (5). The figures present the variance term structure for up to 250 days, which corresponds approximately to the number of trading days in a year, and, therefore, captures the empirically relevant term structure for option valuation. In both figures, the effects of shocks prove significantly more persistent in the component model than in either the FIHNGARCH model or the GARCH(1,1) model. However, although a negative shock in the FIHNGARCH model persists longer than in the GARCH(1,1) model, the FIHNGARCH model does not sufficiently distinguish itself from the GARCH(1,1) following a positive shock. Comparing across Figures 4a and Figure 4b, it is also clear that the term structure of the leverage of the component model is more flexible. As a result, current shocks and the current state of the economy potentially have a more profound impact on the pricing of options across maturities in the component model than in the FIHNGARCH and the GARCH(1,1). To save space, we do not plot the autocorrelation functions of squared innovations which confirm the patterns in Figure 4.

These findings differ somewhat from those contained in the existing literature. Maheu (2002) found that the simple FIGARCH model generates a decay pattern for the autocorrelation function of the absolute value of return series for S&P 500 data, similar to that of a component model. Both autocorrelation functions diminish to zero around 2000. This shorter memory is reflected by a relative low fractional differencing parameter d. As documented by Bollerslev and Mikkelsen (1996), d is estimated at 0.447 for S&P 500 index returns from January 2, 1953 through December 31, 1990. We get $\widehat{d} = 0.2032$. To see how closely the d value relates to memory, we present Figure 5, which is an altered Figure 4a with d varying from 0.1 to 0.4, while keeping all other estimates fixed. It is evident that the impulse response of the variance term structure to a positive shock tends to decay slowly with an increasing d, while it tends to decay fast with a decreasing d. The same thing is true for a negative shock.

One possible explanation is the leverage effect, as imposed to the long lags, in this model. Fractional integration imposes hyperbolic decay pattern for shocks while, at the same time, it extends the memory for the leverage effect. Moreover, the squared innovations tend to put higher weights on large negative shocks, hence enhancing the leverage effect. It is widely documented that the leverage effect introduced by Black (1976) and Christie (1982) merely comprises temporary behavior for the S&P 500 index.[3] From an economic point of view, the debt-equity ratio may be hard to adjust in the short run, but there is no reason that a firm will not be able to adjust its capital structure over time in order to correct the overly strong leverage effect. Generally, this side effect is inevitable for many fractionally integrated models that allows for leverage effect, such as the fractionally integrated EGARCH, fraction-

[3]Engle and Lee (1992), Gallant, Rossi, and Tauchen (1993), and Giraitis, Leipus, and Robinson (2003).

ally integrated TGARCH or fractionally integrated NGARCH. In contrast, the component GARCH separates the variance into two components: long-run and short-run, each of which has its own leverage effect governed by the level of γ_1 and γ_2 respectively. That the leverage effect is modeled more flexibly as two parts, helps to avoid the dilemma of fractional integration.

Overall, understanding all implication of the affine structure turns out to be more complicated than expected. Affine models are convenient because they lead to closed-form solutions for prices of European options. CJW(2007) and Christoffersen, Jacobs, and Mimouni (2005) documented the limitations of the affine structure in terms of fitting returns as well as fitting European options. In order to address the limitations of the affine structure, the Heston (1993) model, which is a continuous-time limit of the Heston-Nandi GARCH(1,1) model, is often combined with models of jumps in returns and volatility. However, relatively little is known about the empirical biases that result from imposing the affine structure. However, the fairness of the comparison in our context is not compromised as long as the affine structure is also employed for the component model.

To shed more light on the driving forces behind the shorter memory or lower d value, we estimate another two models by maximum likelihood. One is a simple FIGARCH model which is free from leverage effects

$$\begin{aligned} h_t &= \frac{\omega}{1-\beta} + \left(1 - \frac{(1-\varphi L)(1-L)^d}{(1-\beta L)}\right)\varepsilon_t^2 \\ h_t &= \overline{\omega} + \lambda(L)\,\varepsilon_t^2 \end{aligned}$$

Our aim is to ascertain whether in the absence of a leverage effect, we obtain longer memory than that obtained in the benchmark FIGARCH model. The other model that we develop is a fractionally integrated nonlinear GARCH model with leverage effect (NGARCH)

$$\begin{aligned} h_t &= \frac{\omega}{1-\beta} + \frac{\left(1 - \frac{(1-\varphi L)(1-L)^d}{(1-\beta L)}\right)}{1+\gamma^2} h_t(z_t-\gamma)^2 \\ h_t &= \overline{\omega} + \frac{\lambda(L)}{1+\gamma^2} h_t(z_t-\gamma)^2 \end{aligned}$$

By switching to a nonlinear structure with leverage effect, we wish to establish the impact of the affine structure on memory. In both cases, $\lambda(L)$ has the same structure as in the FIHNGARCH model. Table 3 presents the MLE and the log likelihood function values for these two models. d is the parameter most directly related to memory. For the FIGARCH model, $d = 0.442$; in the case of the FIHNGARCH model, $d = 0.480$. Figure 6 illustrates impulse responses for a positive shock 2 for all three models. Consistent with the estimated values of d, the non-affine model yields the slowest decay or the highest memory, while the affine model yields the fastest decay. The simple FIGARCH model lies somewhere in between. To some extent, this confirms our conjecture that both the leverage effect and the affine structure reduce the model's memory and that the affine structure constitutes the dominant determinant.

The properties illustrated in the above section are interesting. They suggest that, on the one hand, the leverage effect in the model restrains the long memory, which mitigates the model's ability in fitting derivatives prices; On the other hand, incorporating a leverage parameter γ_1 helps to generate more volatile higher moments. It is undeniable that under our settings, a lower γ_1 generates more negative skewness as well as higher variance of variance by taking the derivatives of (17) and (16) with respect to γ_1. We know that higher moments such as skewness and kurtosis play important roles in determining option prices. Consequently, the model's ability to capture higher moments determines the ability of the FIHNGARCH model in fitting European option data.

4.3 Out-of-Sample Performance with Option Data

We use six years of S&P 500 call option data covering the period 1990-1995. Starting from the raw data from the Berkeley Option data base, we apply standard filters following Bakshi, Cao and Chen (1997). We only use options with more than seven days to maturity. Also, we only use Wednesday options data because Wednesday is the day of the week least likely to be a holiday. It is also less likely than other days (such as Monday and Friday) to be affected by day-of-the-week effects. If Wednesday is a holiday, we use the next trading day. Using only Wednesday data allows us to study a fairly long time series, which is useful in considering the highly persistent volatility processes.

Table 1 presents descriptive statistics for the options data for 1990-1995 by moneyness and maturity. Panel A reports the number of contracts available after filtering. Our sample consists of 21,752 options that span a wide range of moneyness and maturity. Panel B shows the average call price in each of the bins in Panel A. Quite predictably, the average price increases significantly as the moneyness increases (moving down the rows) and as maturity increases (moving from left to right). The average overall price is $27.91.

In Panel C of Table 1, we report the average Black-Scholes implied volatility for the option contracts in each bin. Panel C clearly documents the volatility smirk evident in quoted equity index option prices. The average implied volatility tends to increase as we move down the rows in each column of Panel C, the effect being most dramatic for the short maturities in the left-hand columns. This empirical regularity illustrates that the Black-Scholes option valuation formula, which assumes a constant per-period volatility across time, maturity and strike prices, will generate systematic pricing errors. This motivates the use of stochastic volatility and GARCH models for option valuation.

When calculating option prices, we risk neutralize the MLE estimates in Table 1. The risk-neutral parameters are used to compute the conditional variance based on the structure of (20). Variances on Wednesday are then selected, together with other inputs such as strike, maturity, interest rate, and equity price, to compute the European option prices. As illustrated in the previous section, the variance has the analytical form of (21).

Panel B of Table 2 reports the RMSEs for the two GARCH models from 1990 to 1995. The RMSE is computed as

$$RMSE = \sqrt{\frac{1}{N^T} \sum_{i,t} \left(C_{i,t}^{MKT} - C_{i,t}^{GARCH} \right)^2} \tag{24}$$

where $C_{i,t}^{MKT}$ is the market price of option i at time t, $C_{i,t}^{GARCH}$ is the model price, and $N^T = \sum_{t=1}^{T} N_t$. T is the total number of days included in the sample, and N_t the number of options included in the sample at date t.

We present the absolute values of the RMSEs as well as the normalized RMSEs, defined as the ratio of RMSEs of the component GARCH and the FIHNGARCH model, devided by the GARCH(1,1) RMSEs. It is discernible that the FIHNGARCH model yields the highest RMSEs ranging from 1.801 to 3.583. While the component model generates the lowest RMSEs ranging from 1.263 to 2.559, the GARCH(1,1) model lies in between with RMSEs ranging from 1.608 to 3.239. We also display the RMSEs by moneyness and maturity in Table 4. In general, the component GARCH model performs the best across moneyness and maturity, but especially for options with maturities between 20 days and 180 days. In addition, slightly longer memory for the FIHNGARCH model cannot guarantee the superiority of its out-of-sample performance over that of the GARCH(1,1) model. In fact, the FIHNGARCH framework may boost the likelihood function for daily returns without improving much the conditional density function for returns that are relevant for option valuation. To confirm this, we compute option prices of the FIHNGARCH model by Mont Carlo simulation and derive similar RMSEs.

Figure 7 presents the average weekly biases from 1990 to 1995. The biases seem to be highly related across the three models: all give negative biases from 1990 through 1991, and positive biases from 1992 through 1995. We plot the CBOE volatility index (VIX) in Figure 8b. Since the VIX shows the expected market volatility for a 30 day horizon in Figure 8a, we plot the cumulative 30-day ahead forecasted conditional variance for all three models as defined in 4. When comparing Figures 8a and 8b, we observe that, during the entire period of 1990 to 1995, the variances from the three models are much flatter than that of the VIX.[4] For the 1990 and 1991 recessions, the modeled variances are considerably lower than the implied variances and, therefore, all models generate much lower option prices than the real prices. On the other hand, since 1992, the market started to recover and became increasingly less volatile through 1992 to 1995. Although Figure 8a illustrates that the models can capture this

[4] Please note that, under Duan's Locally Risk-Neutral Valuation Relationship assumption, the risk-neutralized variance is supposed to be identical to the physical variance.

trend in sample, the out-of-sample performances are poorer; the models cannot fully forecast the downward trend of volatility, and, hence, generate higher option prices. Nevertheless, the component GARCH yields better forecasts of future volatility than do the GARCH(1,1) and the FIHNGARCH, and, consequently, achieves the best out-of-sample performance. We also plot the average weekly RMSE over the same period in Figure 9. One important conclusion which may be drawn from Figure 9 is that the improved performance of the component GARCH does not stem from any particular subsample.

Another point worth of mention is that the RMSEs are computed from the maximum likelihood estimates. So far, the theoretical property of the maximum likelihood estimations of any FIGARCH model have not been established. Baillie, Bollerslev and Mikkelsen (1996) justified the usage of the approximate maximum likelihood procedure for a simple FIGARCH model by Mont Carlo simulations. The consistency and other asymptotic properties of the MLE estimates of other fractionally integrated GARCH models including FIEGARCH remain unverified. In Figure 10, we simulate the log-likelihood function and the RMSEs by varying γ and d in reasonable ranges, while leaving other parameters unchanged as MLEs. It appears that RMSE reaches its minimum when $\gamma = 120$ and $d = 0.35$, compared to the maximum of the likelihood function at $\gamma = 100$ and $d = 0.20$. The change of γ is trivial, while the larger d from the minimum of RMSEs confirms that a longer memory will enrich the volatility dynamic and, therefore, better capture the option prices. The goodness-of-fit of the FIHNGARCH model could clearly be improved by using NLS to yield a larger d. The discrepancy existing in the optimal estimates between MLE and NLS sheds light on the latent inconsistency between the MLE estimates and nonlinear least square estimates.

5 Conclusion

Bollerslev and Mikkelsen (1996, 1999) and Comte, Coutin and Renault (2001) investigated and discussed some of the implications of long memory for option valuation. However, their work merely illustrated the implication of long memory on European option prices through Monte Carlo simulations, and little empirical work in fitting options data has been done.

This paper compares two groups of GARCH models that allow for long memory in volatility: the component Heston-Nandi GARCH model developed by CJW(2007), and the fractionally integrated Heston-Nandi GARCH model based on Baillie, Bollerslev and Mikkelsen (1996). We investigate the models using S&P 500 index returns and cross-sectional European options data. The component GARCH model is slightly better than FIHNGARCH in fitting S&P 500 returns, and significantly outperforms FIHNGARCH in fitting the option prices. In return, the FIHNGARCH model dominates the GARCH(1,1) in terms of log-likelihood function while yielding higher option price RMSEs than does the GARCH(1,1) model. This superiority is mainly due to the shorter memory of the FIHNGARCH model, which, in turn, can be attributed to either an artificially prolonged leverage effect created during the procedure of fractional integration or an undesired property of the affine structure. Although FIGARCH models are not quite uncommon in the literature, our findings are novel.

Our paper inspires many directions for further research. To avoid the affine structure, we could develop a fractionally integrated nonlinear GARCH model (NGARCH), introduced by Engle and Ng (1993), and compare it to a component NGARCH model. The better performance of the NGARCH model is reported widely in the existing literature, such as Christoffersen, Jacobs, and Mimouni (2005), and Duan (1995). The downside is that no analytical form of option pricing formula exists and one has to use Monte Carlo simulations.

Figure 10 shows potential to improve the memory of FIHNGARCH by doing NLS estimation. Accordingly, we compare models using information contained in options data. Moreover, we avoid the latent inconsistency between approximate MLE estimates and NLS estimates.

This paper focuses on discrete-time models. Another approach would be to use continuous-time models that allow for long memory, such as the model proposed by Comte, Coutin and Renault (2001), and the continuous-time variance component model of Duffie, Pan and Singleton (1999). It would be an interesting experiment to investigate and compare the abilities of this model to generate long memory with that of the component GARCH model.

References

[1] Ait-Sahalia, Y., Lo, A., 1998. Nonparametric estimation of state-price densities implicit in financial asset prices. Journal of Finance 53, 499-547.

[2] Alizadeh, S., Brandt, M., Diebold, F., 2002. Range-based estimation of stochastic volatility models. Journal of Finance 57, 1047-1091.

[3] Andersen, T., Benzoni, L., Lund, J., 2002. An empirical investigation of continuous-time equity return models. Journal of Finance 57, 1239-1284.

[4] Baillie, R., Bollerslev, T., Mikkelsen, H., 1996. Fractionally integrated generalized autoregressive conditional heteroskedasticity. Journal of Econometrics 74, 3-30.

[5] Bakshi, C., Cao, C., Chen, Z., 1997. Empirical performance of alternative option pricing models. Journal of Finance 52, 2003-2049.

[6] Bates, D., 1996. Jumps and stochastic volatility: exchange rate processes implicit in deutsche mark options. Review of Financial Studies 9, 69-107.

[7] Bates, D., 2000. Post-87 crash fears in S&P500 futures options. Journal of Econometrics 94, 181-238.

[8] Black, F., 1976. Studies of stock price volatility changes. In: Proceedings of the 1976 meetings of the business and economic statistics section, American Statistical Association, pp. 177-181.

[9] Black, F., Scholes, M., 1973. The pricing of options and corporate liabilities. Journal of Political Economy 81, 637-659.

[10] Bollerslev, T., Mikkelsen, H., 1996. Modeling and pricing long memory in stock market volatility. Journal of Econometrics 73, 151-184.

[11] Bollerslev, T., Mikkelsen, H., 1999. Long-term equity anticipation securities and stock market volatility dynamics. Journal of Econometrics 92, 75-99.

[12] Bollerslev, T., Zhou, H., 2002. Estimating stochastic volatility using conditional moments of integrated volatility. Journal of Econometrics 109, 33-65.

[13] Brandt, M., Jones, C., 2006. Forecasting volatility with range-based EGARCH models. Journal of Business and Economic Statistics 24, 470-486.

[14] Broadie, M., Chernov, M., Johannes, M., 2004. Model specification and risk premiums: the evidence from the futures options. Journal of Finance, forthcoming.

[15] Carr, P., Wu, L., 2004. Time-changed Levy processes and option pricing. Journal of Financial Economics 17, 113–141.

[16] Christoffersen, P., Heston, S., Jacobs, K., 2006. Option valuation with conditional skewness. Journal of Econometrics 131, 253-284.

[17] Christoffersen, P., Jacobs, K., 2004. Which GARCH model for option valuation? Management Science 50, 1204-1221.

[18] Christoffersen, P., Jacobs, K., Mimouni, K., 2005. An empirical comparison of affine and non-affine models for equity index options. Unpublished working paper, McGill University.

[19] Comte, F., Coutin, L., Renault, E., 2001. Affine fractional stochastic volatility models. Unpublished working paper, University of Montreal.

[20] Dai, Q., Singleton, K., 2000. Specification analysis of affine term structure models. Journal of Finance 55, 1943-1978.

[21] Ding, Z., Granger, C., Engle, R., 1993. A long memory property of stock market returns and a new model. Journal of Empirical Finance 83–106.

[22] Duan, J.-C., 1995. The GARCH option pricing model. Mathematical Finance 5, 13-32.

[23] Duan, J.-C., Ritchken, P., Sun, Z., 2005. Jump starting GARCH: pricing and hedging options with jumps in returns and volatilities. Unpublished working paper, Rotman School, University of Toronto.

[24] Duffie, D., Pan, J., Singleton, K., 2000. Transform analysis and asset pricing for affine jump-diffusions. Econometrica 68, 1343-1376.

[25] Engle, R., Lee, G., 1999. A permanent and transitory component model of stock return volatility. In: Engle, R., White, H. (Eds.), Cointegration, causality, and forecasting: a festschrift in honor of Clive W.J. Granger. Oxford University Press, New York, pp. 475-497.

[26] Eraker, B., 2004. Do stock prices and volatility jump? Reconciling evidence from spot and option prices. Journal of Finance 59, 1367-1403.

[27] Eraker, B., Johannes, M., Polson, N., 2003. The impact of jumps in volatility and returns. Journal of Finance 58, 1269-1300.

[28] Fama, E., French, K., 1988. Permanent and temporary components of stock prices. Journal of Political Economy 96, 246-273.

[29] Hansen, P. and A. Lunde (2005), "A Forecast Comparison of Volatility Models: Does Anything Beat a GARCH(1,1)?" forthcoming in Journal of Applied Econometrics. 71-104.

[30] Heston, S., 1993. A closed-form solution for options with stochastic volatility with applications to bond and currency options. Review of Financial Studies 6, 327-343.

[31] Heston, S., Nandi, S., 2000. A closed-form GARCH option pricing model. Review of Financial Studies 13, 585-626.

[32] Hsieh, K., Ritchken, P., 2000. An empirical comparison of GARCH option pricing models. Unpublished working paper, Case Western Reserve University.

[33] Huang, J.-Z., Wu, L., 2004. Specification analysis of option pricing models based on time-changed Levy processes. Journal of Finance 59, 1405–1439.

[34] Jones, C., 2003. The dynamics of stochastic volatility: evidence from underlying and options markets. Journal of Econometrics 116, 181-224.

[35] Maheu, J., McCurdy, T., 2004. News arrival, jump dynamics and volatility components for individual stock returns. Journal of Finance 59, 755–793.

[36] Nandi, S., 1998. How important is the correlation between returns and volatility in a stochastic volatility model? Empirical evidence from pricing and hedging in the S&P 500 index options market. Journal of Banking and Finance 22, 589-610.

[37] Pan, J., 2002. The jump-risk premia implicit in options: evidence from an integrated time-series study. Journal of Financial Economics 63, 3–50.

[38] Poterba, J., Summers, L., 1988. Mean reversion in stock returns: evidence and implications. Journal of Financial Economics 22, 27-60.

[39] Ritchken, P., Trevor, R., 1999. Pricing options under generalized GARCH and stochastic volatility processes. Journal of Finance 54, 377-402.

[40] Schroder, M., 2004. Risk-neutral parameter shifts and derivative pricing in discrete time. Journal of Finance 59, 2375-2401.

[41] Scott, L., 1987. Option pricing when the variance changes randomly: theory, estimators and applications. Journal of Financial and Quantitative Analysis 22, 419-438.

[42] Vuong, Q. (1989), "Likelihood Ratio Tests for Model Selection and Non-Nested Hypotheses," Econometrica, 57, 307-333

6 Appendix

6.1 The FIHNGARCH MGF

Define $\overline{\lambda}(L) = \overline{\lambda}_1 L + \overline{\lambda}_2 L^2 + ...$and $\overline{\lambda}_i = \frac{\lambda_i}{\gamma_1^2}$, we guess that the moment-generating function has the log-linear form[5]

$$f_t = E_t \exp(\phi \ln(S_T)) = \exp\left(\phi x_t + A_t + B_t h_{t+1} + \Lambda_t(L)\left(z_{t+1} - \gamma_1\sqrt{h_{t+1}}\right)^2\right)$$

and $\Lambda_t(L) = \Lambda_{t,1}L + \Lambda_{t,2}L^2 + ... + \Lambda_{t,n}L^n$

We have the terminal condition $A_T = B_T = \Lambda_{Ti} = 0$, $i = 1, 2, 3...1000$. Applying the law of iterated expectations to $f_{t;T,\phi}$,we obtain

$$f_t = E_t[f_{t+1}] = E_t \exp\left(\phi x_{t+1} + A_{t+1} + B_{t+1}h_{t+2} + \Lambda_{t+1}(L)\left(z_{t+2} - \gamma_1\sqrt{h_{t+2}}\right)^2\right)$$

Substituting the dynamics of x_t gives

$$f_t = E_t \exp\begin{pmatrix} \phi(r + x_t) - 0.5\phi h_{t+1} + \phi\sqrt{h_{t+1}}z_{t+1} + A_{t+1} + B_{t+1}h_{t+2} + \\ \Lambda_{t+1}(L)\left(z_{t+2} - \gamma_1\sqrt{h_{t+2}}\right)^2 \end{pmatrix}$$

$$= E_t \exp\begin{pmatrix} \phi(x_t + r) - 0.5\phi h_{t+1} + \phi\sqrt{h_{t+1}}z_{t+1} + A_{t+1} + \\ B_{t+1}\left(\overline{\omega} + \lambda(L)\left(z_{t+2} - \gamma_1\sqrt{h_{t+2}}\right)^2\right) + \\ \Lambda_{t+1}(L)\left(z_{t+2} - \gamma_1\sqrt{h_{t+2}}\right)^2 \end{pmatrix}$$

$$= E_t \exp\begin{pmatrix} \phi(x_t + r) - 0.5\phi h_{t+1} + \phi\sqrt{h_{t+1}}z_{t+1} + A_{t+1} + \\ B_{t+1}\left(\overline{\omega} + \overline{\lambda}_1\left(z_{t+1} - \gamma_1\sqrt{h_{t+1}}\right)^2 + \overline{\lambda}_2 L\left(z_{t+1} - \gamma_1\sqrt{h_{t+1}}\right)^2 + ...\right) \\ \Lambda_{t+1,1}\left(z_{t+1} - \gamma_1\sqrt{h_{t+1}}\right)^2 + \Lambda_{t+1,2}L\left(z_{t+1} - \gamma_1\sqrt{h_{t+1}}\right)^2 + ... \end{pmatrix}$$

$$= E_t \exp\begin{pmatrix} \phi(x_t + r) - 0.5\phi h_{t+1} + A_{t+1} + B_{t+1}\overline{\omega} \\ \left(B_{t+1}\overline{\lambda}_1 + \Lambda_{t+1,1}\right)\left(z_{t+1} - \left(\gamma_1 - \frac{\phi}{2\left(B_{t+1}\overline{\lambda}_1 + \Lambda_{t+1,1}\right)}\right)\sqrt{h_{t+1}}\right)^2 + ... \\ \left(\phi\gamma_1 - \frac{\phi^2}{4\left(B_{t+1}\overline{\lambda}_1 + \Lambda_{t+1,1}\right)}\right)h_{t+1} + ... \\ \left(B_{t+1}\overline{\lambda}_2 + \Lambda_{t+1,2}\right)L\left(z_{t+1} - \gamma_1\sqrt{h_{t+1}}\right)^2 + ... \\ \left(B_{t+1}\overline{\lambda}_3 + \Lambda_{t+1,3}\right)L^2\left(z_{t+1} - \gamma_1\sqrt{h_{t+1}}\right)^2 + ... \end{pmatrix}$$

[5]Please note that the MGF developed here is for the physical process. A risk neutralized MGF can be developed in a similar way by risk neutralizing correspondent parameters first.

$$= E_t \exp \left(\begin{array}{c} \phi(x_t + r) + A_{t+1} + B_{t+1}\overline{\omega} - \frac{1}{2}\ln\left(1 - 2\left(B_{t+1}\overline{\lambda}_1 + \Lambda_{t+1,1}\right)\right) \\ \left(\phi\left(-0.5 + \gamma_1\right) + \frac{\left(B_{t+1}\overline{\lambda}_1 + \Lambda_{t+1,1}\right)\left(\gamma_1 - \frac{\phi}{2\left(B_{t+1}\overline{\lambda}_1 + \Lambda_{t+1,1}\right)}\right)^2}{1 - 2\left(B_{t+1}\overline{\lambda}_1 + \Lambda_{t+1,1}\right)} - \frac{\phi^2}{4\left(B_{t+1}\overline{\lambda}_1 + \Lambda_{t+1,1}\right)} \right) h_{t+1} \\ + \left(B_{t+1}\overline{\lambda}_2 + \Lambda_{t+1,2}\right) L \left(z_{t+1} - \gamma_1\sqrt{h_{t+1}}\right)^2 + ... \\ \left(B_{t+1}\overline{\lambda}_3 + \Lambda_{t+1,3}\right) L^2 \left(z_{t+1} - \gamma_1\sqrt{h_{t+1}}\right)^2 + ... \end{array} \right)$$

where we apply

$$E\left[\exp\left(a\left(z+b\right)^2\right)\right] = \exp\left(-\frac{1}{2}\ln\left(1-2a\right) + \frac{ab^2}{1-2a}\right) \tag{A2}$$

Therefore, equating two sides of (A2), we have

$$A_t = \phi r + A_{t+1} + B_{t+1}\overline{\omega} - \frac{1}{2}\ln\left(1 - 2\left(B_{t+1}\overline{\lambda}_1 + \Lambda_{t+1,1}\right)\right)$$

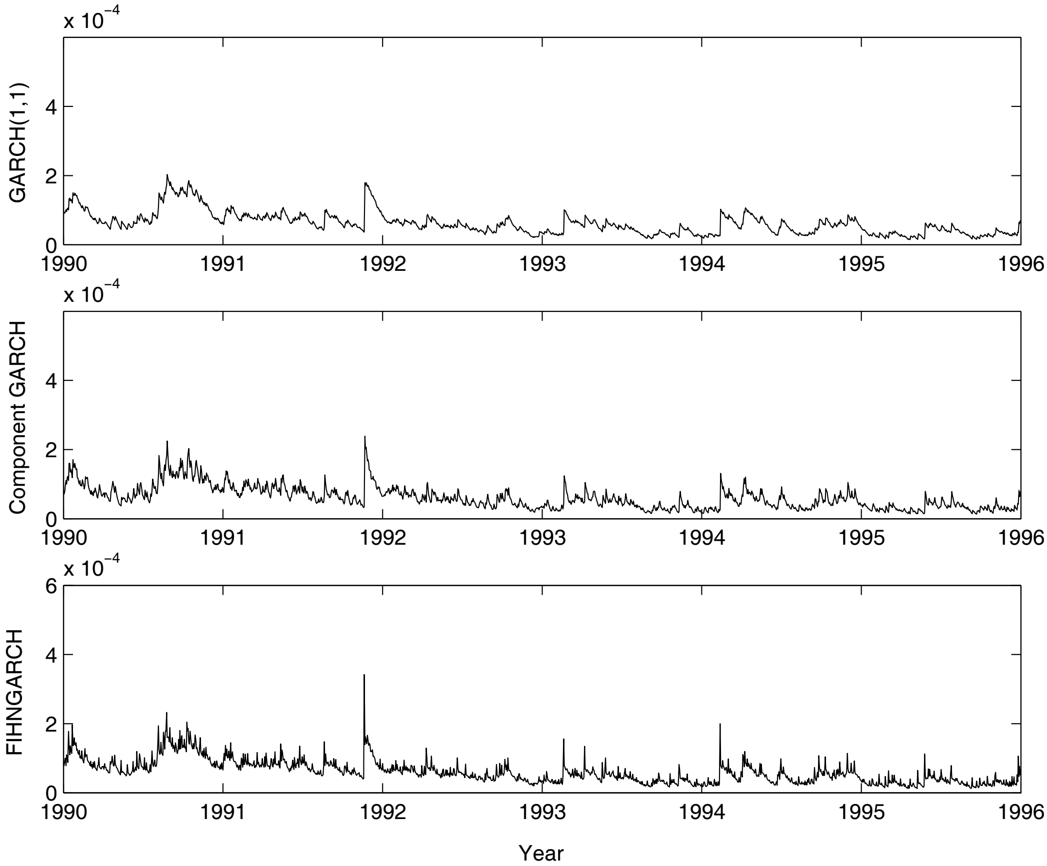
$$B_t = \phi\left(-0.5 + \gamma_1\right) + \frac{\left(B_{t+1}\overline{\lambda}_1 + \Lambda_{t+1,1}\right)\gamma_1^2 + \frac{\phi^2}{2} - \phi\gamma_1}{1 - 2\left(B_{t+1}\overline{\lambda}_1 + \Lambda_{t+1,1}\right)}$$

$$\Lambda_{t,1} = B_{t+1}\overline{\lambda}_2 + \Lambda_{t+1,2};\ \Lambda_{t,2} = B_{t+1}\overline{\lambda}_3 + \Lambda_{t+1,3}, ...$$

$$\Lambda_{t,n} = B_{t+1}\overline{\lambda}_{n+1} + \Lambda_{t+1,n+1}$$

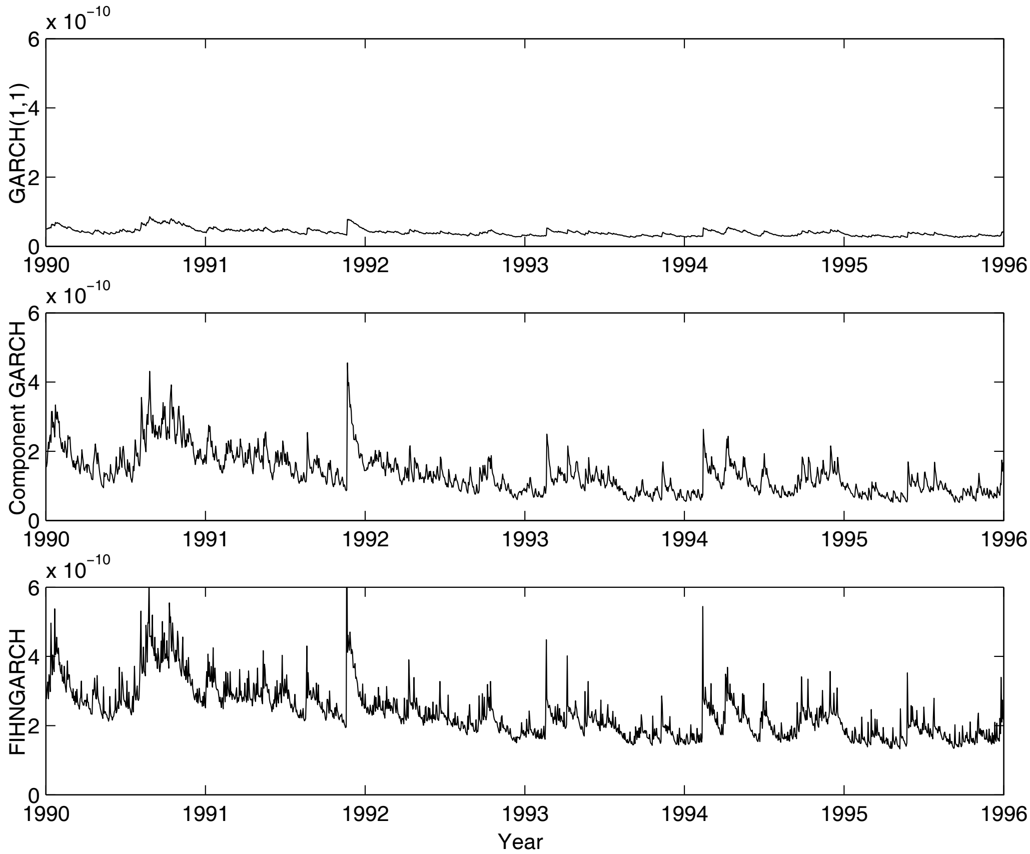
7 Figures and Tables

Figure 1. Conditional Variance



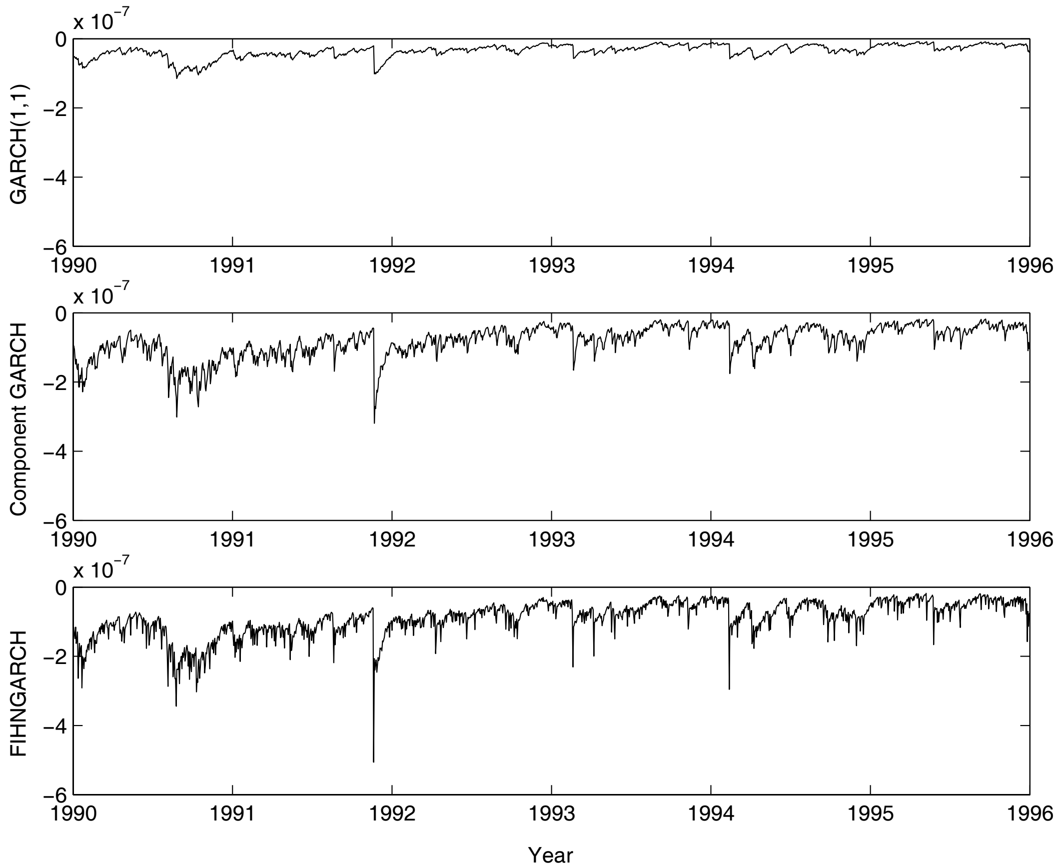
Note to Figure: In Figure 1, we plot the variance paths from the GARCH(1,1), the component GARCH, and the FIHNGARCH model. The parameters are obtained from the MLE estimates on returns in Table 2.

Figure 2. Conditional Variance of Variance



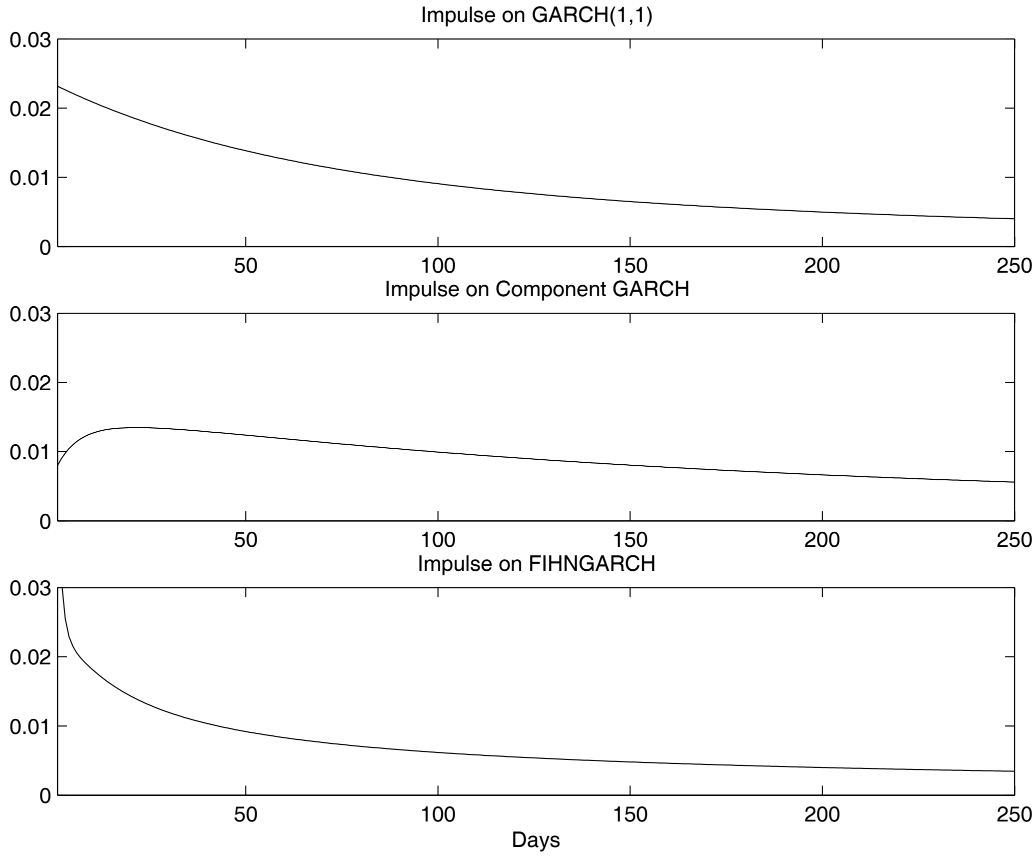
Note to Figure: In Figure 2, we plot the conditional variance of next day's variance as implied by the GARCH(1,1), the component GARCH and the FIHNGARCH models. The scales are identical across panels to facilitate comparison across models. The parameters are obtained from the MLE estimates on returns in Table 2.

Figure 3. Conditional Leverage Effect



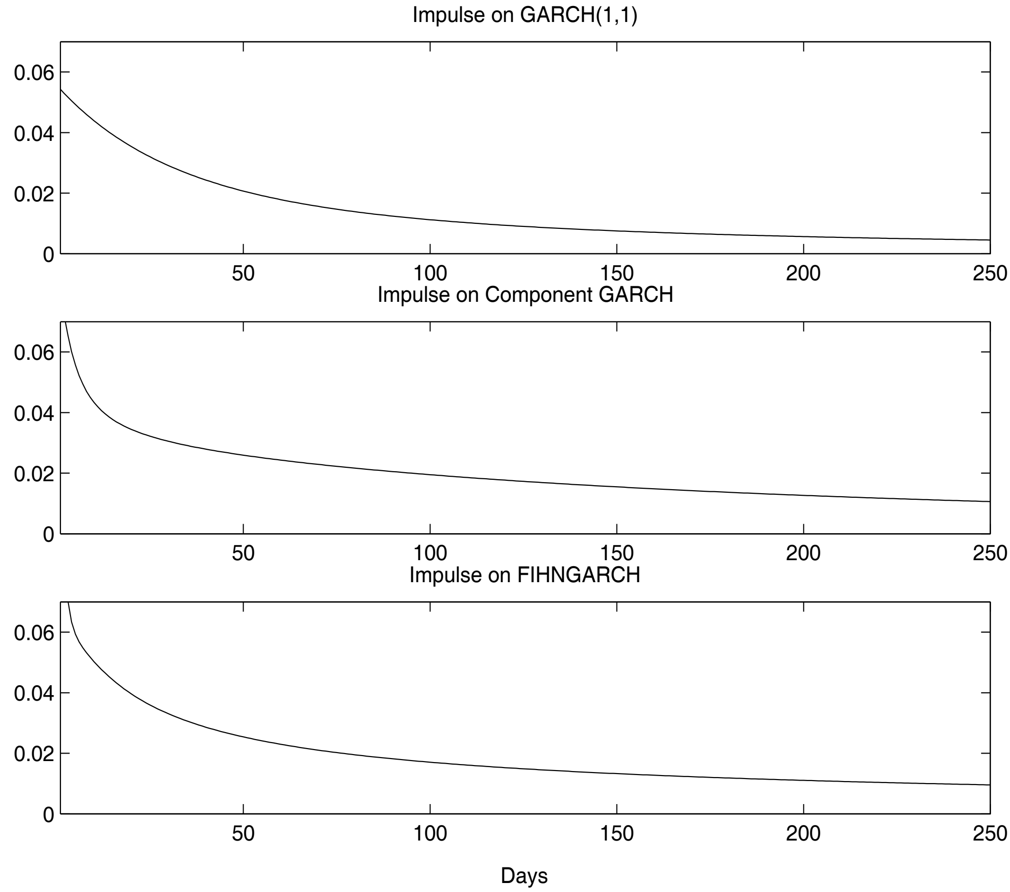
Note to Figure: In Figure 3, we plot the conditional leverage between the return and the next-day variance as implied by the component GARCH, and FIHNGARCH models and refer to it as conditional leverage. The scales are identical across panels to facilitate comparison across models. The parameters are obtained from the MLE estimates on returns in Table 2.

Figure 4a. Term Structure Impulse Response to a Positive Return Shock ($z_t = 2$)



Note to Figure: In Figure 4a, we plot the variance term structure response to a $z_t = 2$ shock in the GARCH(1,1), the component GARCH model and the FIHNGARCH model. The parameters are obtained from the MLE estimates in Table 2. The current variance is set equal to its unconditional value. All values are normalized by the unconditional variance.

Figure 4b. Term Structure Impulse Response to A Negative Return Shock ($z_t = -2$)



Note to Figure: In the Figure 4b, we plot the variance term structure response to a $z_t = -2$ shock in the component GARCH model, and in the FIHNGARCH model. The parameters are obtained from the MLE estimates in Table 2. The current variance is set equal to the unconditional value. All values are normalized by the unconditional variance.

Figure 5. Term Structure Impulse Response to a Positive Return Shock ($z_t = 2$) under Different Values of d

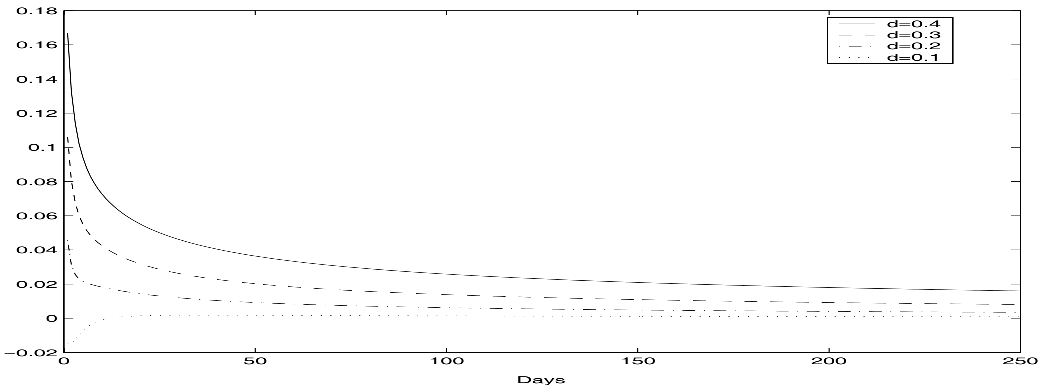
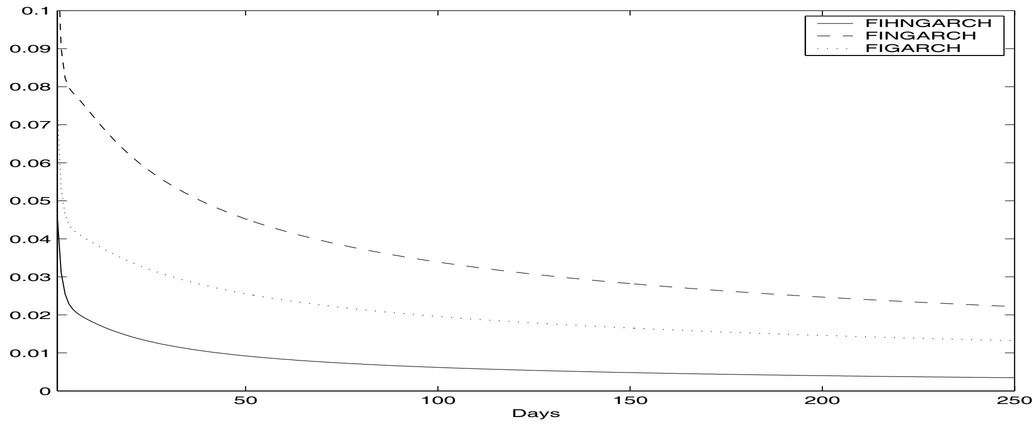
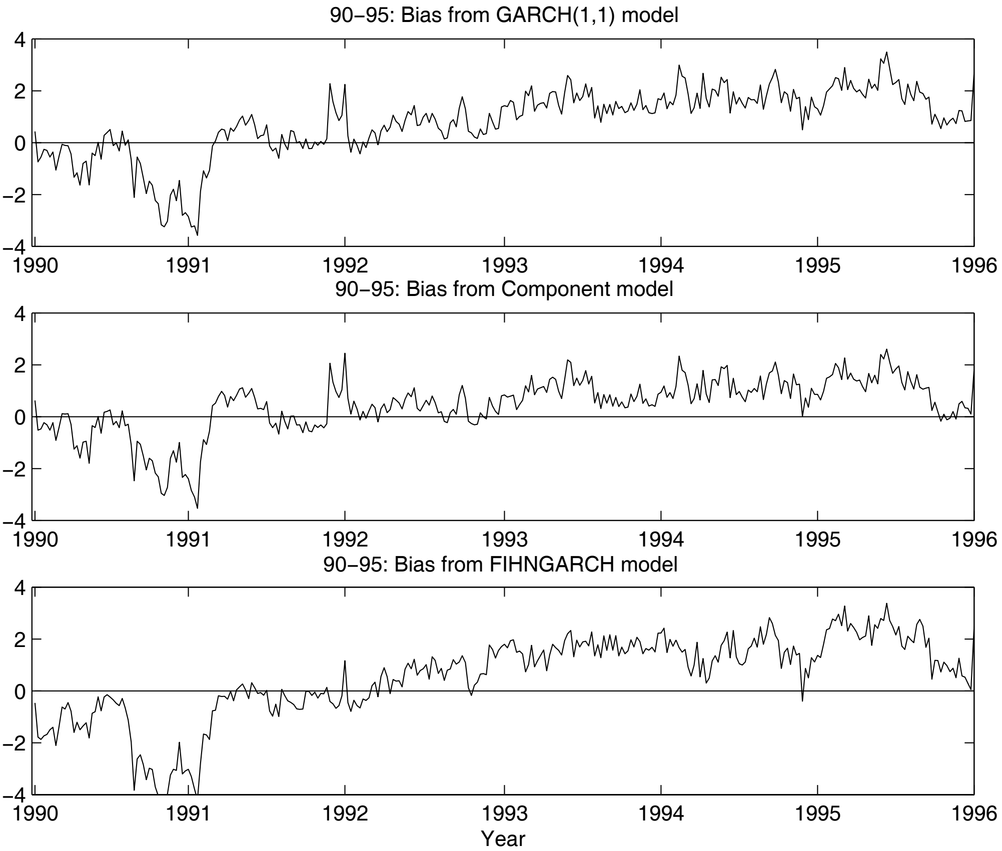


Figure 6. Term Structure Impulse Response to a Positive Return Shock ($z_t = 2$) for Different Models



Note to Figure: In Figure 5. we plot the variance term structure impulse response to a shock $z_t = 2$ in the FIHNGARCH model by varying d, while keeping all other MLE parameters unchanged as in Table 2. In Figure 6. we plot the variance term structure impulse response to a shock $z_t = 2$ for three different GARCH models. All values are normalized by the unconditional variance. The parameters are obtained from the MLE estimates in Table 2. The current variance is set equal to the unconditional value. All values are normalized by the unconditional variance.

Figure 7. Weekly Average Dollar Bias



Note to Figure: We plot the average weekly RMSE (modeled prices less market prices) for the GARCH(1,1), the component GARCH, and the FIHNGARCH during the option data sample (1990-1995). The parameters are obtained from the MLE estimates on returns in Table 2.

Figure 8a. The Cumulative 30-day Ahead Forecasted Conditional Variance

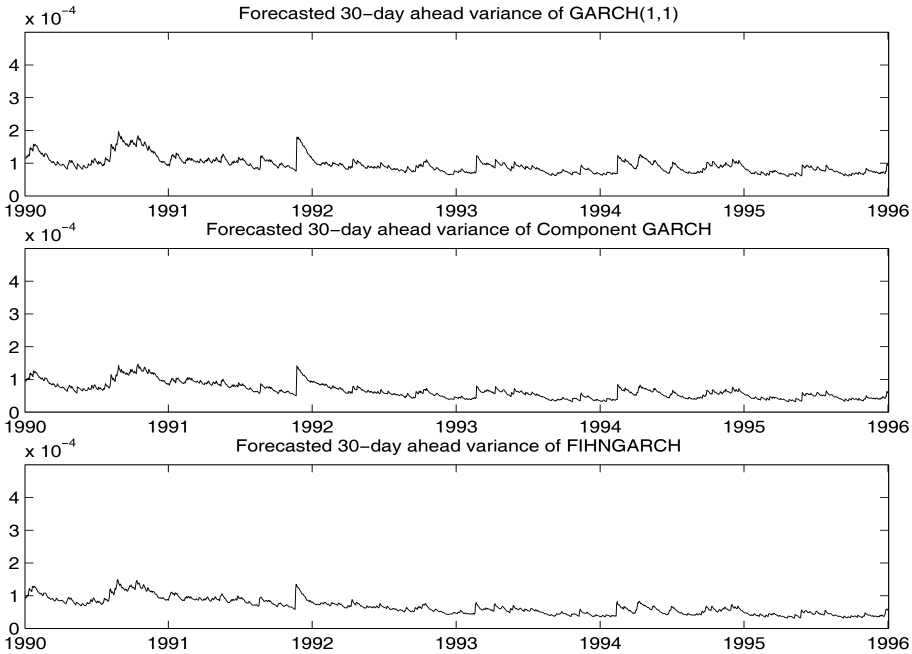
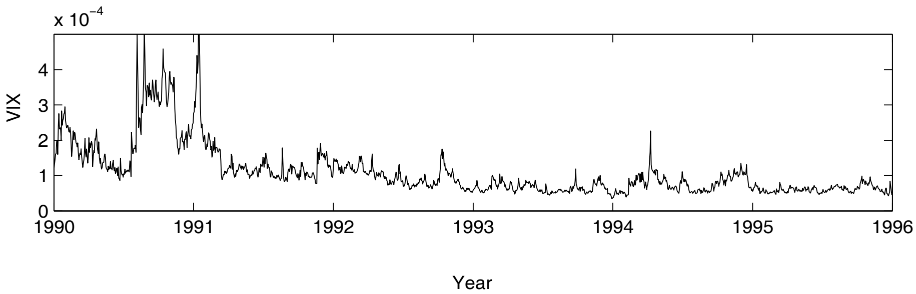
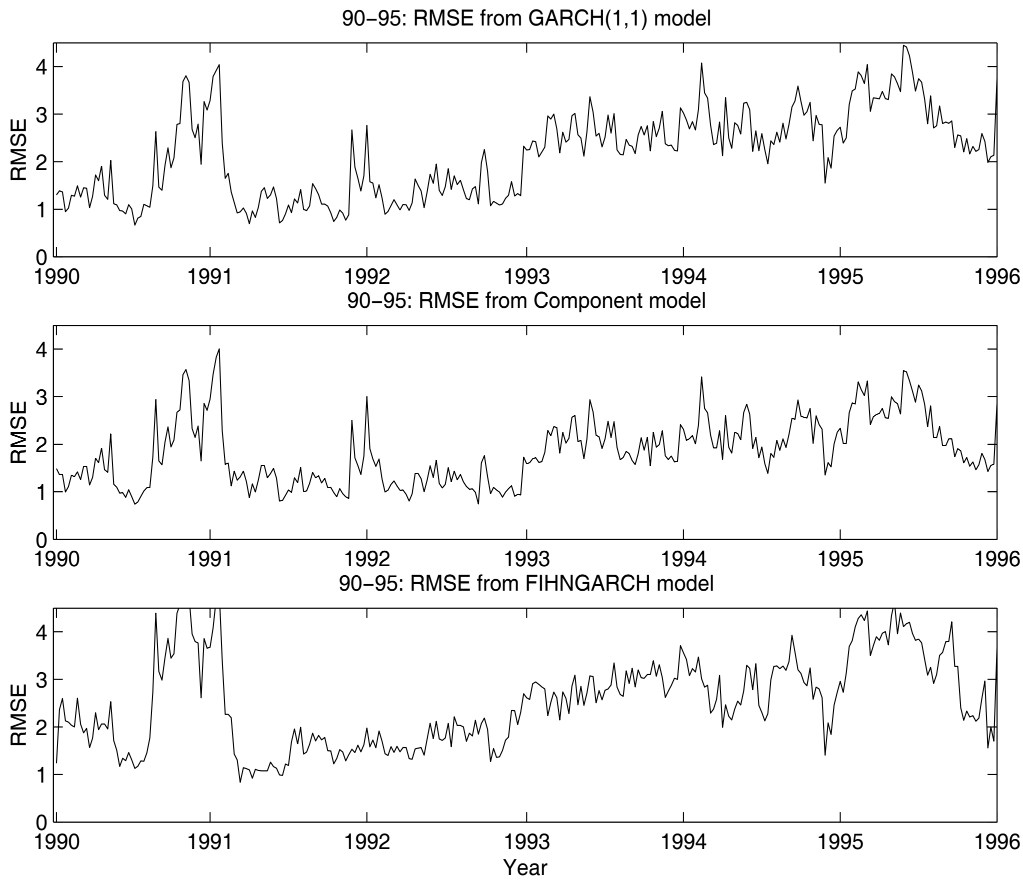


Figure 8b. VIX



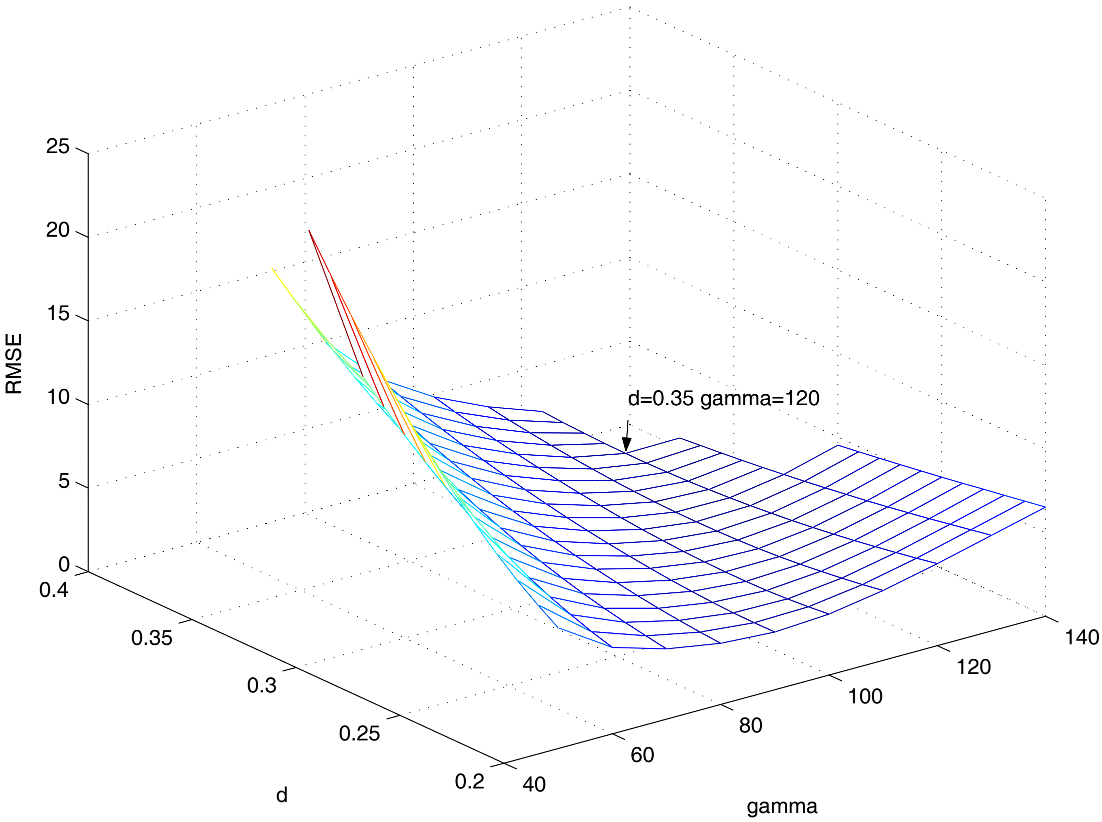
Note to Figure: In panel a, we plot the cumulative 30-day ahead forecasted variance paths from the GARCH(1,1), the component GARCH and the FIHN-GARCH model. The parameters are obtained from the MLE estimates on returns in Table 2. In Panel b, we plot the VIX index from the CBOE for comparison. The scales are identical across panels to facilitate comparison across models.

Figure 9. Weekly Average Dollar RMSE



Note to Figure: We plot the average weekly bias (modeled prices less market prices) for the component GARCH and FIHNGARCH during the option data sample (1990-1995). The parameters are obtained from the MLE estimates on returns in Table 2.

Figure 10. Surfaces of RMSEs



Note to Figure: We plot the RMSE surface for the FIHNGARCH model for varying d and γ, keeping other MLE estimates unchanged as in Table 2

Table 1: S&P 500 Index Call Option Data (1990-1995)

Panel A. Number of Call Option Contracts

	DTM<20	20<DTM<80	80<DTM<180	DTM>180	All
S/X<0.975	101	1,884	1,931	1,769	5,685
0.975<S/X<1.00	283	1,272	706	477	2,738
1.00<S/X<1.025	300	1,212	726	526	2,764
1.025<S/X<1.05	261	1,167	654	409	2,491
1.05<S/X<1.075	245	1,039	582	390	2,256
1.075<S/X	549	2,345	1,679	1,245	5,818
All	1,739	8,919	6,278	4,816	21,752

Panel B. Average Call Price

	DTM<20	20<DTM<80	80<DTM<180	DTM>180	All
S/X<0.975	0.88	2.30	6.25	11.94	6.62
0.975<S/X<1.00	2.29	6.83	15.19	27.50	12.12
1.00<S/X<1.025	8.35	13.60	22.48	34.41	19.32
1.025<S/X<1.05	17.57	22.00	30.11	42.14	26.97
1.05<S/X<1.075	27.11	30.84	38.14	48.83	35.43
1.075<S/X	50.67	52.79	58.99	68.34	57.70
All	24.32	23.66	28.68	36.07	27.91

Panel C. Average Implied Volatility from Call Options

	DTM<20	20<DTM<80	80<DTM<180	DTM>180	All
S/X<0.975	0.1625	0.1269	0.1350	0.1394	0.1342
0.975<S/X<1.00	0.1308	0.1296	0.1449	0.1562	0.1383
1.00<S/X<1.025	0.1527	0.1459	0.1558	0.1606	0.1520
1.025<S/X<1.05	0.1915	0.1647	0.1665	0.1656	0.1681
1.05<S/X<1.075	0.2433	0.1828	0.1775	0.1739	0.1865
1.075<S/X	0.3897	0.2356	0.1961	0.1868	0.2283
All	0.2434	0.1703	0.1622	0.1607	0.1717

Notes to Table: We use European call options on the S&P 500 index. The prices are taken from quotes within 30 minutes from closing on each Wednesday during the January 1, 1990 to December 31, 1995 period. The moneyness and maturity filters used by Bakshi, Cao and Chen (1997) are applied here as well. The implied volatilities are calculated using the Black-Scholes formula.

Table 2: Panel a. MLE Estimates and Properties
Sample: Daily Returns, 1966-2001

GARCH(1,1)			**Component GARCH**			**FIHNGARCH**		
Parameter	Estimate	Std. Error	Parameter	Estimate	Std. Error	Parameter	Estimate	Std. Error
b	0.977	0.012	$\widetilde{\beta}$	7.659E-01	0.163	β_1	6.636E-01	3.414E-07
a	3.210E-06	2.810E-06	α	1.770E-06	1.110E-06	d	2.032E-01	5.778E-04
c	88.192	15.623	γ_1	3.129E+02	108.430	γ_1	1.016E+02	1.430E-02
λ	1.815	0.224	γ_2	5.904E+01	30.196	φ_1	5.355E-01	7.147E-04
		0.054	φ	1.960E-06	3.720E-07	η	1.945E+00	3.358E-03
			ρ	9.888E-01	0.002			
			λ	1.809E+00	5.26E-01			
Annual Vol	0.147		Annual Vol	0.145		Annual Vol	0.145	
Vol of Vol	4.574E-06		Vol of Vol	1.329E-05		Vol of Vol	1.700E-05	
Leverage	-5.662E-04		Leverage	-1.339E-03		Leverage	-1.481E-03	
Ln Likelihood	30059.800		Ln Likelihood	30112.480		Ln Likelihood	30104.500	

Notes to Table: We use daily total returns from July 1, 1966 to December 31, 2001 on the S&P 500 index to estimate the three GARCH models using Maximum Likelihood. Robust standard errors are calculated from the outer product of the gradient at the optimum parameter values. Annual Vol refers to the annualized unconditional standard deviation as implied by the parameters in each model. Var of Var refers to the unconditional variance of the conditional variance in each model. For FIGARCH models where the unconditional varaince does not exist, we use the average of the conditional variance. Leverage refers to the unconditional covariance between the return and the conditional variance. Ln Likelihood refers to the logarithm of the likelihood at the optimal parameter values.

Table 2: Panel b. RMSE of MLE Estimates
Sample: Option Data, 1990-1995

GARCH(1,1)		**Component GARCH**		**FIHNGARCH**	
RMSE(90-95)	2.461	RMSE(90-95)	2.040	RMSE(90-95)	2.787
Normalized	1	Normalized	0.829	Normalized	1.133
RMSE(90	1.920	RMSE(90	1.859	RMSE(90)	2.804
Normalized	1	Normalized	0.968	Normalized	1.461
RMSE (91)	1.608	RMSE (91)	1.630	RMSE (91)	1.871
Normalized	1	Normalized	1.014	Normalized	1.164
RMSE (92)	1.433	RMSE (92)	1.263	RMSE (92)	1.801
Normalized	1	Normalized	0.881	Normalized	1.256
RMSE (93)	2.584	RMSE (93)	2.045	RMSE (93)	2.891
Normalized	1	Normalized	0.791	Normalized	1.119
RMSE (94)	2.786	RMSE (94)	2.245	RMSE (94)	2.852
Normalized	1	Normalized	0.806	Normalized	1.024
RMSE (95)	3.239	RMSE (95)	2.559	RMSE (95)	3.583
Normalized	1	Normalized	0.790	Normalized	1.106

Notes to Table: Option RMSE refers to the fit of the models on the 21,752 contracts quoted from 1990 to 1995 in Table 1. The RMSEs are computated at the MLE in Panel a, Table 1.

Table 3: MLE Estimates
Sample: Daily Returns, 1966-2001

FIGARCH			FINGARCH		
Parameter	Estimate	Std. Error	Parameter	Estimate	Std. Error
β	6.731E-01	9.989E-02	β	7.20E-01	1.81E-02
d	4.424E-01	5.585E-02	δ	4.81E-01	1.83E-02
φ	3.487E-01	9.025E-02	γ	5.85E-01	2.84E-02
l	4.961E+00	9.785E-01	φ	3.80E-01	2.12E-02
ϖ	6.491E-06	8.192E-07	λ	4.35160324	1.63E-01
			ϖ	1.23E-12	1.16E-12
Ln Likelihood	30093.000		Ln Likelihood	30143.9	

Notes to Table: We use daily total returns from July 1, 1966 to December 31, 2001 on the S&P 500 index to estimate the two GARCH models using Maximum Likelihood. Robust standard errors are calculated from the outer product of the gradient at the optimum parameter values. Ln Likelihood refers to the logarithm of the likelihood at the optimal parameter values.

Table 4: RMSE and Ratio RMSE by Moneyness and Maturity

Panel A. GARCH(1,1)

	DTM<20	20<DTM<80	80<DTM<180	DTM>180	All
S/X<0.975	0.438	1.825	3.305	5.060	3.310
0.975<S/X<1.00	0.661	2.059	3.289	4.363	2.633
1.00<S/X<1.025	0.597	1.549	2.648	3.676	2.139
1.025<S/X<1.05	0.580	1.102	2.043	3.071	1.618
1.05<S/X<1.075	0.744	0.931	1.663	2.354	1.346
1.075<S/X	0.758	0.988	1.211	1.697	1.182
All	0.674	1.467	2.544	3.842	2.240

Panel B. Ratio of Component to GARCH(1,1) RMSE

	DTM<20	20<DTM<80	80<DTM<180	DTM>180	All
S/X<0.975	0.766	0.808	0.824	0.849	0.833
0.975<S/X<1.00	0.789	0.784	0.771	0.786	0.781
1.00<S/X<1.025	0.898	0.774	0.747	0.758	0.764
1.025<S/X<1.05	0.977	0.861	0.774	0.744	0.800
1.05<S/X<1.075	0.995	1.012	0.881	0.772	0.896
1.075<S/X	0.999	1.046	1.048	0.930	1.008
All	0.947	0.843	0.820	0.829	0.833

Panel C. Ratio of FIHNGARCH to GARCH(1,1) RMSE

	DTM<20	20<DTM<80	80<DTM<180	DTM>180	All
S/X<0.975	1.370	1.449	1.180	1.003	1.145
0.975<S/X<1.00	1.750	1.458	1.157	0.958	1.224
1.00<S/X<1.025	1.483	1.431	1.169	0.914	1.178
1.025<S/X<1.05	1.076	1.342	1.198	0.962	1.166
1.05<S/X<1.075	0.921	1.269	1.283	1.069	1.192
1.075<S/X	1.011	1.163	1.391	1.329	1.272
All	1.228	1.402	1.194	1.008	1.180

Notes to Table: We use the MLE estimates from Table 2 to compute the root mean squared option valuation error (RMSE) for various moneyness and maturity bins during 1990-1995. Panel A shows the RMSEs for the GARCH(1,1) model. Panel B shows the ratio of the Component GARCH MSEs to the GARCH(1,1) RMSEs from Panel A. Panel C shows the ratio of the FIHNGARCH RMSEs to the GARCH(1,1) RMSEs.