Three Essays on Volatility Long Memory and European Option Valuation

Yintian Wang

Desautels Faculty of Management
McGill University
Montreal, Canada

March 2007

A thesis submitted to
Faculty of Graduate Studies and Research
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy

Your file Votre référence
ISBN: 978-0-494-32334-2
Our file Notre référence
ISBN: 978-0-494-32334-2

Abstract

This dissertation is in the form of three essays on the topic of component and long memory GARCH models. The unifying feature of the thesis is the focus on investigating European index option evaluation using these models.

The first essay presents a new model for the valuation of European options. In this model, the volatility of returns consists of two components. One of these components is a long-run component that can be modeled as fully persistent. The other component is short-run and has zero mean. The model can be viewed as an affine version of Engle and Lee (1999), allowing for easy valuation of European options. The model substantially outperforms a benchmark single-component volatility model that is well established in the literature. It also fits options better than a model that combines conditional heteroskedasticity and Poisson normal jumps. While the improvement in the component model's performance is partly due to its improved ability to capture the structure of the smirk and the path of spot volatility, its most distinctive feature is its ability to model the term structure. This feature enables the component model to jointly model long-maturity and short-maturity options.

The second essay derives two new GARCH variance component models with non-normal innovations. One of these models has an affine structure and leads to a closed-form option valuation formula. The other model has a non-affine structure and hence, option valuation is carried out using Monte Carlo simulation. We provide an empirical comparison of these two new component models and the respective special cases with normal innovations. We also compare the four

component models against GARCH(1,1) models which they nest. All eight models are estimated using MLE on S&P500 returns. The likelihood criterion strongly favors the component models as well as non-normal innovations. The properties of the non-affine models differ significantly from those of the affine models. Evaluating the performance of component variance specifications for option valuation using parameter estimates from returns data also provides strong support for component models. However, support for non-normal innovations and non-affine structure is less convincing for option valuation.

The third essay aims to investigate the impact of long memory in volatility on European option valuation. We mainly compare two groups of GARCH models that allow for long memory in volatility. They are the component Heston-Nandi GARCH model developed in the first essay, in which the volatility of returns consists of a long-run and a short-run component, and a fractionally integrated Heston-Nandi GARCH (FIHNGARCH) model based on Bollerslev and Mikkelsen (1999). We investigate the performance of the models using S&P500 index returns and cross-sections of European options data. The component GARCH model slightly outperforms the FIGARCH in fitting return data but significantly dominates the FIHNGARCH in capturing option prices. The findings are mainly due to the shorter memory of the FIHNGARCH model, which may be attributed to an artificially prolonged leverage effect that results from fractional integration and the limitations of the affine structure.

Abstract

La dissertation ci-dessous comporte trois essais consacrés aux modèles GARCH et à l'évaluation européenne du prix d'option. Ces trois parties ont en commun d'étudier les techniques européennes d'évaluation du prix d'option en utilisant les modèles GARCH.

La première partie présente un nouveau modèle pour l'évaluation des options européennes. Dans notre modèle, la volatilité du rendement se compose de deux attributs. Le premier est une composante de longue durée qui peut être modelée de manière persistante. L'autre composante est à court terme et a une moyenne de zéro. Notre modèle a l'ambition de préciser la version mise au point par Engle et Lee (1999) en facilitant l'évaluation des options européennes. Ce modèle surpasse considérablement le modèle établi de volatilité à composante simple. Les options s'adaptent mieux que dans un modèle qui combine l'hétéroskedasticité conditionnelle et la composante saut de poisson. L'amélioration du modèle est partiellement due à son identification de la structure ascendante et de la volatilité du marché ponctuel. Son dispositif le plus distinctif réside en sa capacité de modeler la structure de limite. Ce dispositif permet de modeler conjointement des options avec un délai de remboursement à longue terme et à court terme.

Dans le deuxième essai, nous dérivons deux nouveaux modèles GARCH à composante aléatoire et aux innovations inédites ou anormales. Un de ces modèles a une structure d'affinage et aboutit à une formule d'évaluation d'option close. L'autre modèle a un non-affinage dynamique. Son évaluation du prix d'option doit être faite par l'intermédiaire de la simulation de Monte Carlo. Nous procédons une comparaison empirique de ces deux nouveaux modèles et de leurs cas spéciaux respectifs avec les

innovations normales. Nous comparons également les quatre modèles à ceux de GARCH (1.1), qu'ils emboîtent. Chacun des huit modèles est évalué en utilisant l'estimateur maximum de vraisemblance (MLE) sur les retours de S&P500. Le critère de probabilité favorise fortement les modèles composants et les innovations anormales. Les propriétés des modèles sans affinage diffèrent significativement de celles des modèles d'affinage. En projetant les paramètres pour évaluer les prix d'option, nous aboutissons encore à des résultats en faveur des composantes alatoires, mais les résultats des innovations anormales et des structures de non-affinage sont moins convaincants.

Le troisième essai étudie l'impact de la volatilité mémoire longue sur l'évaluation européenne du prix d'option en utilisant différents modèles. Nous comparons principalement deux groupes de modèles GARCH permettant la volatilité mémoire longue. Il s'agit premièrement du composant Heston-Nandi du modèle GARCH développé par Christoffersen, Jacobs, et Wang (2005), dont la volatilité du rendement est à court et long terme ; il s'agit ensuite du modèle GARCH partiellement intégré de Heston-Nandi basé sur Baillie, Bollerslev et Mikkelsen (1996). Nous étudions les modèles au travers des retours d'indice S&P 500 et des données sur les options européennes basées sur une coupe statistique. Les données de rendement du modèle GARCH surpassent légèrement le FIGARCH, mais c'est dans l'évaluation des prix d'option que GARCH domine de manière significative le FIGARCH. La supériorité du modèle GARCH est due à la mémoire plus courte du modèle FIGARCH, qui pourrait être attribuée à l'effet de levier artificellement prolongé par l'intégration et la limitation partielles de la structure d'affinage.

Acknowledgments

I would like to express my gratitude to all those who gave me the opportunity to complete this thesis.

I convey my sincere thanks to my co-supervisors Prof. Peter Christoffersen and Prof. Kris Jacobs for their unwavering encouragement and support during my Ph.D. study. I would have never finished without their enlightenment, guidance and generous financial support. It is difficult to overstate my gratitude to Prof. Peter Christoffersen, who tracked the progress of my work and was always available when I needed his advice. His enthusiasm, and his great efforts to explain things clearly and simply, helped rendering academic problems intuitive and fun for me. I acknowledge deep appreciation to him for introducing me to this strand of research. I have learnt a lot from him.

I also wish to thank Prof. Lawrence Kryzanowski for serving on my thesis committee and providing me with valuable comments. Indeed, I am indebted to all the faculty members of the Finance Ph.D. program of McGill. Without their talent, dedication, and patience, the Finance Ph.D. program would not exist.

I would like to thank my fellow finance Ph.D. students, Lei Lu and Hao Wang, for helpful. instructive discussions.

I am also grateful to the program manager Stella Scalia, who is always assisting Ph.D. students with great patience and dedication.

A journey is easier when you travel together. Finally, I would give my special thanks to my husband, Xiaochuan Qiu. Although we lived apart for more than two years, he nevertheless accompanied me with great love and emotional support at every moment. I cannot imagine how I could have finished my thesis without him. I dedicate this thesis to him with deepest love and gratitude.

Contribution of the Authors

I am solely responsible for Chapter 3 "The Impact of Volatility Long Memory on Option Valuation: Component GARCH versus FIGARCH." Nevertheless, I would like to thank Prof. Christoffersen and Prof. Jacobs for inspiration, support, and guidance.

Chapter 1 and Chapter 2, "Option Valuation with Long-run and Short-run Volatility Component" and "Volatility Components: Affine Restrictions and Non-normal Innovations" are joint work with the co-supervisors my thesis, Peter Christoffersen and Kris Jacobs. Prof. Christoffersen, Prof. Jacobs and I have made equally substantial contributions to the two essays. The responsibility of any remaining error is shared accordingly.

Contents

General Introduction

The seminal work of Black and Scholes (1973) and Merton (1973) on option pricing theory, commonly known as the Black-Scholes model, has not only spawned a huge literature on derivative contracts but also transformed the financial industry. However, this influential option pricing model has several shortcomings. Many empirical studies, including the empirical work in Black and Scholes (1972), have shown that the Black-Scholes model exhibits systematic pricing biases. It tends to overprice call options with high strike prices, and underprice call options with low strike prices. Recent empirical studies typically focus on the pricing biases in terms of implied volatilities, and the bias phenomenon is referred to as the "volatility smile" or the "volatility smirk". The "volatility smirk" refers to the phenomenon that the Black-Scholes implied volatilities for stock call options often exhibit a downward-sloping, convex pattern when plotted against their exercise prices. This persistent feature of option data contradicts the prediction of the Black-Scholes model, which implies constant implied volatility.

It has been documented in the existing literature that the volatility smirk is partly due to the unrealistic assumption of normally distributed returns in the Black-Scholes model. Empirical evidence suggests that the return distribution has fatter tails, and that the distribution implicit in option prices is substantially negatively skewed after the 1987 crash. It is therefore necessary to build in skewness and excess kurtosis in the return process. This can be done in several ways.

Heston (1993) proposes a continuous-time stochastic volatility model that allows for correlation between volatility and spot asset returns. In the discrete-time literature, the NGARCH(1,1) option valuation model proposed by Duan (1995) allows for time variation in the conditional variance as well as a leverage effect that generates skewness in returns. However, the discrete time model of Duan does not provide a closed form solution for option valuation. Heston and Nandi (2000) proposed a closely related GARCH option pricing model that provides a closed form solution (up to a numerical integration) for European option valuation.

Another approach is to assume distributions other than Gaussian for the return innovations. Candidate distributions should contain more shape parameters than the normal distribution in order to accommodate fatter tails, for example the GED distribution, and/or skewness, for example the Inverse Gaussian distribution. Christoffersen, Heston and Jacobs develop a new discrete-time dynamic model of stock returns with Inverse Gaussian innovation. The model allows for conditional skewness as well as heteroskedasticity and a leverage effect, and gives a closed-form solution. Their empirical results suggest that the model improves the pricing of out-of-money put options.

A large number of papers have added jump components to the dynamics of returns or to both returns and volatility. In stochastic volatility models, the volatility smile decreases with maturity. This contradicts the stylized fact that shorter maturity options have a more pronounced smile. Furthermore, diffusive stochastic volatility can only increase gradually by a sequence of small, normally distributed increments. However, while jumps in returns can generate large movements and more skewness during short time intervals, the impact of a jump is transient. The general consensus is that both jumps and stochastic volatility are needed. Jumps generate return non-normality over the short term while a persistent stochastic volatility process slows down the convergence of the return to normality as the maturity increases.

In summary, while stochastic volatility models, jump processes and non-normal innovations improve on the Black-Scholes model in a qualitative sense, they are still biased in a quantitative sense, because the strength of the effects is insufficient. To further improve on these existing models, we need models that possess the same qualitative feature but contain stronger quantitative effects.

This dissertation attempts to provide such models by focusing on the strong evidence of long memory in return volatility. The variance is highly persistent over long horizons (see Ding, Granger, and Engle (1993) and Andersen, Bollerslev, Diebold and Labys (2003)). Various long-memory models have been developed to capture this stylized fact. Engle and Lee (1999) introduced a component GARCH model to capture the long memory in volatility. Baillie, Bollerslev and Mikkelsen (1996) and Bollerslev and Mikkelsen (1999) incorporate the idea of long-memory fractional differencing into the GARCH model. Comte, Coutin and Renault (2001) propose an extension of Heston's (1993) model to capture the long-run dependencies in volatility. This model can disentangle short and long-memory properties in the resulting option prices. Despite the appeal of these models, empirical work that applies long-memory models to option pricing is quite limited. No empirical research has ever been carried out to compare the performance of long-memory models with that of other popular benchmarks. This dissertation aims to develop novel long-memory volatility models that allows for easier and improved European option evaluation.

To model the variance, we can either use a continuous-time stochastic volatility model or a discrete-time GARCH model. The advantages of the continuous-time models lie in their mathematical elegance, and that they sometimes lead to closed-form option pricing formulas. However, GARCH models may offer distinct advantages over stochastic volatility models from an estimation perspective. Continuous-time stochastic volatility models are difficult to implement because, with discrete observations on the underlying asset price process, the volatility is not readily identifiable. We therefore use a GARCH framework. The dissertation takes the form of three essays on the topic of component GARCH models. The unifying feature of the entire thesis is the focus on investigating European index option valuation with component GARCH models.

There are two cornerstones in the first dissertation essay. One is the component GARCH model of Engle and Lee (1993) and the other is an affine GARCH(1,1) model proposed by Heston and Nandi (2000). Building on these two papers, the first essay presents a new component GARCH model that allows for easy valuation of European options. In the model, the volatility of returns consists of two components. One of these components is a long-run component that can be modeled as fully persistent. The other component is short-run and has zero mean. Due to the flexibility in variance term structure and the flexibility in generating more higher moments, the model can forecast the conditional density functions of 7-day to 360-day returns well. It therefore generates more accurate European option prices. This model substantially outperforms a benchmark single-component volatility model that is well established in the literature. It also fits options better than a model that combines conditional heteroskedasticity and Poisson normal jumps.

In the first essay, some very fundamental assumptions are imposed, namely normally distributed return innovations and an affine structure. The second essay relaxes these assumptions and derives two new component GARCH models with non-normal innovations. One of these models has an affine structure with Inverse Gaussian return innovations and leads to a closed-form option valuation formula. The other model has a non-affine structure with GED return innovations. Since non-affine models do not lead to closed form solutions, we use Monte Carlo simulations for option valuation. An empirical comparison of these two new component models and the respective special cases with normal innovations is provided. All four component models are also compared with the GARCH(1,1) models which they nest. All eight models are estimated using MLE on S&P500 returns. The likelihood criterion strongly favors the component models as well as non-normal innovations. The properties of the non-affine models differ significantly from those of the affine models. Evaluating the performance of component variance specifications for option valuation using parameter estimates from returns data also provides strong support for component models. However, support for non-normal innovations and non-affine structure is less convincing.

The component GARCH model is not the only GARCH model to capture long memory in volatility. Bollerslev and Mikkelsen (1999) developed a fractionally integrated GARCH model or a FIGARCH, in which a shock to variance decays at a hyperbolic rate. The third essay compares the performance of the component GARCH from the first essay and a fractionally integrated Heston-Nandi GARCH model (FIHNGARCH) in terms of fitting option data. We investigate the performance of the models using S&P500 index returns and cross-sectional European options data. The component Heston-Nandi GARCH model slightly outperforms the FIHNGARCH in fitting returns data, but significantly dominates the FIHNGARCH in capturing option prices. These results are mainly due to the shorter memory of the FIHNGARCH model, which can be attributed to the artificially prolonged leverage effect and the limitation of the affine structure.

Chapter 1
Option Valuation with Long-run and Short-run Volatility Components

Peter Christoffersen Kris Jacobs Yintian Wang

Abstract

This paper presents a new model for the valuation of European options. In our model, the volatility of returns consists of two components. One of these components is a long-run component and it can be modeled as fully persistent. The other component is short-run and has a zero mean. Our model can be viewed as an affine version of Engle and Lee (1999), allowing for easy valuation of European options. The model substantially outperforms a benchmarks single-component volatility model that is well established in the literature, and it fits options better than a model that combines conditional heteroskedasticity and Poisson normal jumps .The improvement in the component model's performance is partly due to its improved ability to capture the structure of the smirk and the path of spot volatility, it its most distinctive feature is its ability to model the term structure. This feature enables the component model to jointly mode long-maturity and short-maturity options.

JEL Classification: G12
Keywords: Volatility term structure; GARCH; Out-of-sample

1.1 Introduction

There is a consensus in the equity options literature that combining time-variation in the conditional variance of asset returns (Engle (1982), Bollerslev (1986)) with a leverage effect (Black (1976)) constitutes a potential solution to well-known biases associated with the Black-Scholes (1973) model, such as the implied volatility smirk. These asymmetric dynamic volatility models generate negative skewness in the distribution of asset returns which in turn generates higher prices for out-of-the-money put options as compared to the Black-Scholes formula. In the continuous-time option valuation literature, the Heston (1993) model addresses some of these biases. This model contains a leverage effect as well as stochastic volatility.[1] In the discrete-time literature, the NGARCH$(1,1)$ option valuation model proposed by Duan (1995) contains time-variation in the conditional variance as well as a leverage effect. The model by Heston and Nandi (2000) is closely related to Duan's model.

Many existing empirical studies have confirmed the importance of time-varying volatility, the leverage effect and negative skewness in continuous-time and discrete-time setups, using parametric as well as non-parametric techniques.[2] However, it has become clear that while these models help explain the biases of the Black-Scholes model in a qualitative sense, they come up short in a quantitative sense. Using parameters estimated from returns or options data, these models reduce the biases of the Black-Scholes model, but the mag-

[1] The importance of stochastic volatility is also studied in Hull and White (1987), Melino and Turnbull (1990), Scott (1987) and Wiggins (1987).

[2] See for example Ait-Sahalia and Lo (1998), Amin and Ng (1993), Bakshi, Cao and Chen (1997), Bates (2000), Benzoni (1998), Bollerslev and Mikkelsen (1999), Chernov and Ghysels (2000), Duan, Ritchken and Sun (2005, 2006), Engle and Mustafa (1992), Eraker (2004), Heston and Nandi (2000), Jones (2003), Nandi (1998) and Pan (2002).

nitude of the effects is insufficient to completely resolve the biases. The resulting pricing errors have the same sign as the Black-Scholes pricing errors, but are smaller in magnitude. We therefore need models that possess the same qualitative features as the models in Heston (1993) and Duan (1995), but that contain stronger quantitative effects. These models need to generate more flexible skewness and volatility of volatility dynamics in order to fit observed option prices. Existing studies have attempted to address this by combing stochastic volatility specifications with jump processes, or by using non-normal innovations in heteroskedastic models.[3]

The shortcomings of existing models in modeling the moneyness dimension are compounded by their shortcomings in modeling the term structure of volatility, as well as the path of spot volatility. It has been observed using a variety of diagnostics that it is difficult to fit the dynamics of return volatility using a benchmark model such as a GARCH(1, 1). A similar observation applies to stochastic volatility models such as Heston (1993). The main problem is that volatility autocorrelations are too high at longer lags to be explained by a GARCH(1, 1), unless the process is extremely persistent. This extreme persistence may impact negatively on other aspects of option valuation, such as the valuation of short-maturity options.

In fact, it has been observed in the literature that volatility may be better modeled using a fractionally integrated process, rather than a stationary GARCH process.[4] Andersen, Bollerslev, Diebold and Labys (2003) confirm this finding using realized volatility.

[3] See for example Bakshi, Cao and Chen (1997), Bates (2000), Broadie, Chernov and Johannes (2004), Christoffersen, Heston and Jacobs (2006), Eraker, Johannes and Polson (2003), Eraker (2004), Huang and Wu (2004) and Pan (2002).

[4] See Baillie, Bollerslev and Mikkelsen (1996).

Bollerslev and Mikkelsen (1996, 1999) and Comte, Coutin and Renault (2001) investigate and discuss some of the implications of long memory for option valuation. Using fractional integration models for option valuation is somewhat cumbersome. Optimization is time-intensive and certain ad-hoc choices have to be made regarding implementation.

This paper attempts to remedy remaining option biases by modeling richer volatility dynamics. We use a model that is relatively easy to implement and that captures the stylized facts addressed by long-memory models at horizons relevant for option valuation. The model builds on Heston and Nandi (2000) and Engle and Lee (1999). In our model, the volatility of returns consists of two components. One of these components is a long-run component, and it can be modeled as (fully) persistent. The other component is short-run and mean zero. We study two models: one where the long-run component is constrained to be fully persistent and one where it is not. We refer to these models as the persistent component model and the component model respectively. These models are able to generate autocorrelations that are richer than those of a GARCH(1. 1) model while using just a few additional parameters.

Unobserved component or factor models are very popular in the finance literature. See Fama and French (1988), Poterba and Summers (1988) and Summers (1986) for applications to stock prices. In the option pricing literature, Bates (2000) and Taylor and Xu (1994) investigate two-factor stochastic volatility models. Duffie, Pan and Singleton (2000) provide a general continuous-time framework for the valuation of contingent claims using multifactor affine models. Eraker (2004) suggests the usefulness of a multifactor approach based on his empirical results. Alizadeh, Brandt and Diebold (2002) uncover two factors in stochastic volatility models of exchange rates using range-based estimation. Bollerslev and Zhou (2002), Brandt and Jones (2006), Chacko and Viceira (2003), Chernov, Gallant, Ghysels and Tauchen (2003), and Maheu (2002) also find that two-factor stochastic volatility models outperform single factor models when modeling daily asset return volatility. Adrian and Rosenberg (2005) investigate the relevance of a two-component volatility model for pricing the cross-section of stock returns. Unobserved component models are also very popular in the term structure literature, although in this literature the models are more commonly referred to as multifactor models.[5] There are very interesting parallels between our approach and results and stylized facts in the term structure literature. In the term structure literature, it is customary to model short-run fluctuations around a time-varying long-run mean of the short rate. In our framework we model short-run fluctuations around a time-varying long-run volatility.

Dynamic factor and component models can be implemented in continuous or discrete time.[6] We choose a discrete-time approach because of the ease of implementation. In particular, our model is related to the GARCH class of processes, and therefore volatility filtering and forecasting are relatively straightforward, which is critically important for option valuation.[7] An additional advantage of our model is parsimony: the most general

[5] See for example Dai and Singleton (2000), Duffee (1999), Duffie and Singleton (1999) and Pearson and Sun (1994).

[6] Duffie, Pan and Singleton (2000) suggest a multifactor continuous-time model that captures the spirit of our approach, but do not investigate the model empirically.

[7] Because the filtering problem is extremely simple in the GARCH framework, we are able to analyze an extensive option sample. See also Heston and Nandi (2000). See among others Bates (2000, 2006), Chernov and Ghysels (2000), Eraker (2004) and Pan (2002) for other empirical studies that estimate model parameters using options data.

model we investigate has seven parameters. We speculate that parsimony may help our model's out-of-sample performance.

Because our component model is a generalization of the GARCH$(1,1)$ model, and because its implementation uses similar techniques, the GARCH$(1,1)$ is a natural benchmark. Moreover, Christoffersen, Jacobs and Mimouni (2005) find that the performance of the GARCH$(1,1)$ model is similar to that of the Heston (1993) model, which is the most commonly used benchmark in the literature. Heston and Nandi (2000) find that the GARCH$(1,1)$ slightly outperforms the ad-hoc implied volatility benchmark model in Dumas, Fleming and Whaley (1998). Finally, because there is substantial evidence that Poisson-normal jump processes can alleviate some of the biases associated with the Heston (1993) model and the GARCH$(1,1)$ model, we also include a GARCH$(1,1)$ model augmented with Poisson-normal jumps in our analysis.

We provide two different analyses of the component model. We first estimate the physical model parameters by maximum likelihood estimation (MLE) on historical S&P returns for 1962-2001. We compare the component model and the persistent component model to the GARCH$(1,1)$ benchmark as well as to the more general GARCH$(1,1)$-Jump model. Based on the log-likelihood criterion, the GARCH$(1,1)$-Jump model performs the best, followed by the component model, the persistent component model and the GARCH$(1,1)$ model. However, when we compare the models based on option fit using MLE parameters, the best fit is obtained using the component and persistent component model, followed by the GARCH$(1,1)$-Jump model. The GARCH$(1,1)$ model is again the worst performer. We also use the MLE parameters to emphasize differences in important model features, such as the conditional volatility of variance, the correlation between returns and conditional variance, the term structure of conditional skewness and kurtosis, the volatility smirk and the volatility term structure. The improvement in the model's performance is due to its richer dynamics, which result in different modeling of the term structure, and which enable the component model to capture patterns in long-maturity as well as short-maturity options.

In a second empirical investigation, we estimate the models using options data, while filtering the latent volatility from the underlying returns data. When the persistence of the long-run component is freely estimated, it is very close to one. The performance of the component model is impressive when compared with a benchmark GARCH(1, 1) model. When using all available option data, the dollar RMSE of the component model is 11.3-22.7% lower than that of the benchmark GARCH model in-sample and 21.8-23.3% out-of-sample. Our out-of-sample results strongly suggest that these results are not simply due to spurious in-sample overfitting. The persistent component model performs better than the benchmark GARCH(1, 1) model, but in contrast to the results obtained using MLE parameters, it is clearly inferior to the component model both in- and out-of-sample.

The paper proceeds as follows. Section 2 introduces the model. Section 3 discusses the volatility term structure and Section 4 discusses option valuation. Sections 5 and 6 present the two empirical investigations, and Section 7 concludes.

1.2 Return Dynamics with Volatility Components

In this section we first present the Heston-Nandi GARCH$(1,1)$ model which will serve as the benchmark model throughout the paper. We then construct the component model as a natural extension of a rearranged version of the GARCH$(1,1)$ model. We finally present the persistent component model as a special case of the component model.

1.2.1 The Heston and Nandi GARCH(1,1) Model

Heston and Nandi (2000) propose a class of GARCH models that allow for a closed-form solution for the price of a European call option. They present an empirical analysis of the GARCH$(1,1)$ version of this model, which is given by

$$\begin{aligned} R_{t+1} &\equiv \ln(S_{t+1}/S_t) = r + \lambda h_{t+1} + \sqrt{h_{t+1}} z_{t+1} \qquad (1.1) \\ h_{t+1} &= w + b h_t + a(z_t - c\sqrt{h_t})^2 \end{aligned}$$

where S_{t+1} denotes the underlying asset price, r the risk free rate, λ the price of risk and h_{t+1} the daily variance on day $t+1$ which is known at the end of day t. The z_{t+1} shock is assumed to be i.i.d. $N(0,1)$. The Heston-Nandi model captures time variation in the conditional variance as in Engle (1982) and Bollerslev (1986),[8] and the parameter c captures the leverage effect. The leverage effect captures the negative relationship between shocks to returns and volatility (Black (1976)), which results in a negatively skewed distribution of returns.[9] Note that the GARCH$(1,1)$ dynamic in (1.1) is slightly different from the more

[8] For an early application of GARCH to stock returns, see French, Schwert and Stambaugh (1987).

[9] Its importance for option valuation has been emphasized among others by Benzoni (1998), Chernov and Ghysels (2000), Christoffersen and Jacobs (2004), Eraker (2004), Eraker, Johannes and Polson (2003), Heston (1993), Heston and Nandi (2000) and Nandi (1998).

conventional NGARCH model used by Engle and Ng (1993) and Hentschel (1995), which is used for option valuation in Duan (1995). The reason is that the dynamic in (1.1) is engineered to yield a closed-form solution for option valuation, whereas a closed-form solution does not obtain for the more conventional GARCH dynamic. Hsieh and Ritchken (2000) provide evidence that the more traditional GARCH model may actually slightly dominate the fit of (1.1). Our main point can be demonstrated using either dynamic. Because of the convenience of the closed-form solution provided by dynamics such as (1.1), we use this as a benchmark in our empirical analysis and we model the richer component structure within the Heston-Nandi framework.[10]

To better appreciate the workings of the component models presented below, note that by using the expression for the unconditional variance

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

to substitute out w, the variance process can be rewritten as

$$h_{t+1} = \sigma^2 + b\left(h_t - \sigma^2\right) + a\left((z_t - c\sqrt{h_t})^2 - (1 + c^2\sigma^2)\right). \tag{1.2}$$

1.2.2 Building a Component Volatility Model

The expression for the GARCH(1, 1) variance process in (1.2) highlights the role of the parameter σ^2 as the constant unconditional mean of the conditional variance process. A natural generalization is then to specify σ^2 as time-varying. Denoting this time-varying

[10] See Bollerslev and Mikkelsen (1996), Engle and Mustafa (1992), Christoffersen and Jacobs (2004), and Hsieh and Ritchken (2000) for other empirical studies of European option valuation using GARCH dynamics. Ritchken and Trevor (1999) discusses the pricing of American options with GARCH processes.

component by q_{t+1}, the expression for the variance in (1.2) can be generalized to

$$h_{t+1} = q_{t+1} + \beta\left(h_t - q_t\right) + \alpha\left(\left(z_t - \gamma_1\sqrt{h_t}\right)^2 - \left(1 + \gamma_1^2 q_t\right)\right). \tag{1.3}$$

This model is similar in spirit to the component model of Engle and Lee (1999). The difference between our model and Engle and Lee (1999) is that the functional form of the GARCH dynamic (1.3) allows for a closed-form solution for European option prices. This is similar to the difference between the Heston-Nandi (2000) GARCH$(1, 1)$ dynamic and the more traditional NGARCH$(1, 1)$ dynamic discussed in the previous subsection. In specification (1.3), the conditional volatility h_{t+1} can most usefully be thought of as having two components. Following Engle and Lee (1999), we refer to the component q_{t+1} as the long-run component, and to $h_{t+1} - q_{t+1}$ as the short-run component. We will discuss this terminology in some more detail below. Note that by construction the unconditional mean of the short-run component $h_{t+1} - q_{t+1}$ is zero.

The model can also be written as

$$\begin{aligned} h_{t+1} &= q_{t+1} + \left(\alpha\gamma_1^2 + \beta\right)\left(h_t - q_t\right) + \alpha\left(\left(z_t - \gamma_1\sqrt{h_t}\right)^2 - \left(1 + \gamma_1^2 h_t\right)\right) \\ &= q_{t+1} + \tilde{\beta}\left(h_t - q_t\right) + \alpha\left(\left(z_t - \gamma_1\sqrt{h_t}\right)^2 - \left(1 + \gamma_1^2 h_t\right)\right) \end{aligned} \tag{1.4}$$

where $\tilde{\beta} = \alpha\gamma_1^2 + \beta$. This representation is useful because we can think of

$$\begin{aligned} v_{1,t} &\equiv \left(z_t - \gamma_1\sqrt{h_t}\right)^2 - \left(1 + \gamma_1^2 h_t\right) \\ &= \left(z_t^2 - 1\right) - 2\gamma_1\sqrt{h_t}z_t \end{aligned} \tag{1.5}$$

as a mean-zero innovation.

The model is completed by specifying the functional form of the long-run volatility component. In a first step, we assume that q_{t+1} follows the process

$$q_{t+1} = \omega + \rho q_t + \varphi \left(\left(z_t^2 - 1 \right) - 2\gamma_2 \sqrt{h_t} z_t \right). \tag{1.6}$$

Note that $E\left[q_{t+1}\right] = E\left[h_{t+1}\right] = \sigma^2 = \dfrac{\omega}{1-\rho}$ as long as $\rho < 1$. We can therefore write the component volatility model as

$$\begin{aligned} h_{t+1} &= q_{t+1} + \tilde{\beta}\left(h_t - q_t\right) + \alpha v_{1,t} \\ q_{t+1} &= \omega + \rho q_t + \varphi v_{2,t} \\ &= \sigma^2 + \rho(q_t - \sigma^2) + \varphi v_{2,t} \end{aligned} \tag{1.7}$$

with

$$v_{i,t} = \left(z_t^2 - 1 \right) - 2\gamma_i \sqrt{h_t} z_t, \text{ for } i = 1, 2. \tag{1.8}$$

and $E_{t-1}\left[v_{i,t}\right] = 0$, $i = 1, 2$. Also note that in addition to the price of risk, λ, the model contains seven parameters: $\alpha, \tilde{\beta}, \gamma_1, \gamma_2, \omega, \rho$ and φ.

1.2.3 A Fully Persistent Special Case

In our empirical work, we also investigate a special case of the model in (1.7). Notice that in (1.7) the long-run component of volatility will be a mean reverting process for $\rho < 1$. We also estimate a version of the model which imposes $\rho = 1$. The resulting process is

$$\begin{aligned} h_{t+1} &= q_{t+1} + \tilde{\beta}\left(h_t - q_t\right) + \alpha v_{1,t} \\ q_{t+1} &= \omega + q_t + \varphi v_{2,t} \end{aligned} \tag{1.9}$$

and $v_{i,t}$, $i = 1, 2$ are as in (1.8). In addition to the price of risk, λ, the model now contains six parameters: $\alpha, \tilde{\beta}, \gamma_1, \gamma_2, \omega$ and φ.

In this case the process for long-run volatility contains a unit root and shocks to the long-run volatility never die out: they have a "permanent" effect. Recall that following Engle and Lee (1999) in (1.7) we refer to q_{t+1} as the long-run component and to $h_{t+1} - q_{t+1}$ as the short-run component. In the special case (1.9) we can also refer to q_{t+1} as the "permanent" component, because innovations to q_{t+1} are truly "permanent" and do not die out. It is then customary to refer to $h_{t+1} - q_{t+1}$ as the "transitory" component, which reverts to zero. It is in fact this permanent-effects version of the model that is most closely related to models which have been studied more extensively in the finance and economics literature, rather than the more general model in (1.7).[11] We will refer to this model as the persistent component model.

It is clear that (1.9) is nested by (1.7). It is therefore to be expected that the in-sample fit of (1.7) is superior. However, out-of-sample this may not necessarily be the case. It is often the case that more parsimonious models perform better out-of-sample if the restriction imposed by the model is a sufficiently adequate representation of reality. The persistent component model may also be better able to capture structural breaks in volatility out-of-sample, because a unit root in the process allows it to adjust to a structural break, which not possible for a mean-reverting process. It will therefore be of interest to verify how close ρ is to one when estimating the more general model (1.7).

---

[11] See Fama and French (1988), Poterba and Summers (1988) and Summers (1986) for applications to stock prices. See Beveridge and Nelson (1981) for an application to macroeconomics.

1.3 Variance Term Structures

To intuitively understand the shortcomings of existing models such as the GARCH$(1,1)$ model in (1.1) and the improvements provided by our model (1.7), it is instructive to graphically illustrate some of the models' statistical properties that are key for option valuation. In this section we therefore illustrate the models' variance term structures and impulse response functions.

1.3.1 The Variance Term Structure for the GARCH(1,1) Model

Following the logic used for the component model in (1.7), we can rewrite the GARCH$(1,1)$ variance dynamic in (1.2). We have

$$h_{t+1} = \sigma^2 + \tilde{b}\left(h_t - \sigma^2\right) + a\left(\left(z_t^2 - 1\right) - 2c\sqrt{h_t}z_t\right) \tag{1.10}$$

where $\tilde{b} = b + ac^2$ and where the innovation term has a zero conditional mean. From (1.10) the multi-step forecast of the conditional variance is

$$E_t\left[h_{t+k}\right] = \sigma^2 + \tilde{b}^{k-1}(h_{t+1} - \sigma^2)$$

where the conditional expectation is taken at the end of day t. Notice that $\tilde{b}$ is directly interpretable as the variance persistence in this representation of the model.

We can now define a convenient measure of the variance term structure for maturity K as

$$h_{t+1:t+K} \equiv \frac{1}{K}\sum_{k=1}^{K} E_t\left[h_{t+k}\right] = \frac{1}{K}\sum_{k=1}^{K} \sigma^2 + \tilde{b}^{k-1}(h_{t+1} - \sigma^2) = \sigma^2 + \frac{1-\tilde{b}^K}{1-\tilde{b}}\frac{\left(h_{t+1} - \sigma^2\right)}{K}.$$

This variance term structure measure succinctly captures important information about the model's potential for explaining the variation of option values across maturities.[12] To compare different models, it is convenient to set the current variance, h_{t+1}, to a simple m multiple of the long run variance. In this case the variance term structure relative to the unconditional variance is given by

$$h_{t+1:t+K}/\sigma^2 \equiv 1 + \frac{1-\tilde{b}^K}{1-\tilde{b}} \frac{(m-1)}{K}.$$

The dash-dot lines in the top panels of Figures 1 and 2 show the term structure of variance for the GARCH$(1,1)$ model for a low and high initial conditional variance respectively. We use parameter values estimated via MLE on daily S&P500 returns (the estimation details are in Table 1 and will be discussed further below). We set $m = \frac{1}{2}$ in Figure 1 and $m = 2$ in Figure 2. 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. It can be clearly seen from Figures 1 and 2 that for the GARCH$(1,1)$ model, the conditional variance converges to the long-run variance rather fast.

We can also learn about the dynamics of the variance term structure though impulse response functions. For the GARCH$(1,1)$ model, the effect of a shock at time t, z_t, on the expected k-day ahead variance is

$$\partial(E_t[h_{t+k}])/\partial z_t^2 = \tilde{b}^{k-1} a \left(1 - c\sqrt{h_t}/z_t\right)$$

[12] Notice that due to the price of risk term in the conditional mean of returns, the term structure of variance as defined here is not exactly equal to the conditional variance of cumulative returns over K days.

and thus the effect on the variance term structure is

$$\partial E_t\left[h_{t:t+K}\right]/\partial z_t^2 = \frac{1-\tilde{b}^K}{1-\tilde{b}}\frac{a}{K}\left(1 - c\sqrt{h_t}/z_t\right).$$

The bottom-left panels of Figures 3 and 4 plot the impulse responses to the term structure of variance for $h_t = \sigma^2$ and $z_t = 2$ and $z_t = -2$ respectively, again using the parameter estimates from Table 1. The impulse responses are normalized by the unconditional variance. Notice that the effect of a shock dies out rather quickly for the GARCH$(1,1)$ model. Comparing across Figures 3 and 4 we see the asymmetric response of the variance term structure from a positive versus negative shock to returns. This can be thought of as the term structure of the leverage effect. Due to the presence of a positive c, a positive shock has less impact than a negative shock along the entire term structure of variance.

1.3.2 The Variance Term Structure for the Component Model

In the component model we have

$$\begin{aligned} h_{t+1} &= q_{t+1} + \tilde{\beta}\left(h_t - q_t\right) + \alpha v_{1,t} \\ q_{t+1} &= \sigma^2 + \rho(q_t - \sigma^2) + \varphi v_{2,t}. \end{aligned}$$

The multi-day forecast of the two components are

$$\begin{aligned} E_t\left[h_{t+k} - q_{t+k}\right] &= \tilde{\beta}^{k-1}\left(h_{t+1} - q_{t+1}\right) \\ E_t\left[q_{t+k}\right] &= \sigma^2 + \rho^{k-1}\left(q_{t+1} - \sigma^2\right). \end{aligned}$$

The simplicity of these multi-day forecasts is a key advantage of the component model. The multi-day variance forecast is a simple sum of two exponential components. Notice that

$\tilde{\beta}$ and ρ correspond directly to the persistence of the short-run and long-run components respectively.

We can now calculate the variance term structure in the component model for maturity K as

$$\begin{aligned} h_{t+1:t+K} &\equiv \frac{1}{K}\sum_{k=1}^{K} E_t\left[q_{t+k}\right] + E_t\left[h_{t+k} - q_{t+k}\right] \\ &= \frac{1}{K}\sum_{k=1}^{K} \sigma^2 + \rho^{k-1}\left(q_{t+1} - \sigma^2\right) + \tilde{\beta}^{k-1}\left(h_{t+1} - q_{t+1}\right) \\ &= \sigma^2 + \frac{1-\rho^K}{1-\rho}\frac{q_{t+1}-\sigma^2}{K} + \frac{1-\tilde{\beta}^K}{1-\tilde{\beta}}\frac{h_{t+1}-q_{t+1}}{K}. \end{aligned}$$

If we set q_{t+1} and h_{t+1} equal to m_1 and m_2 multiples of the long run variance respectively, then we get the variance term structure relative to the unconditional variance simply as

$$h_{t+1:t+K}/\sigma^2 = 1 + \frac{1-\rho^K}{1-\rho}\frac{m_1-1}{K} + \frac{1-\tilde{\beta}^K}{1-\tilde{\beta}}\frac{m_2-m_1}{K}. \tag{1.11}$$

The solid lines in the top panels in Figures 1 and 2 show the term structure of variance for the component model using parameters estimated via MLE on daily S&P500 returns from Table 1. We set $m_1 = \frac{3}{4}, m_2 = \frac{1}{2}$ in Figure 1 and $m_1 = \frac{7}{4}, m_2 = 2$ in Figure 2. By picking m_2 equal to the m used for the GARCH$(1,1)$ model, we ensure comparability across models within each figure because the spot variances relative to their long-run variances are identical.[13] The main conclusion from Figures 1 and 2 is that compared to the dash-dot GARCH$(1,1)$, the conditional variance converges more slowly to the unconditional variance in the component model. This is particularly so on days with a high spot

[13] Note that we need $m_1 \neq m_2$ in this numerical experiment to generate a "short-term" effect in (1.11). Changing m_1 will change the picture but the main conclusions stay the same.

variance. The middle and bottom panels show the contribution to the total variance from each component. Notice the strong persistence in the long-run component.

We can also calculate impulse response functions in the component model. The effects of a shock at time t, z_t on the expected k-day ahead variance components are

$$\begin{aligned}\partial E_t\left[q_{t+k}\right]/\partial z_t^2 &= \rho^{k-1}\varphi\left(1-\gamma_2\sqrt{h_t}/z_t\right)\\ \partial E_t\left[h_{t+k}-q_{t+k}\right]/\partial z_t^2 &= \tilde{\beta}^{k-1}\alpha\left(1-\gamma_1\sqrt{h_t}/z_t\right)\\ \partial E_t\left[h_{t+k}\right]/\partial z_t^2 &= \tilde{\beta}^{k-1}\alpha\left(1-\gamma_1\sqrt{h_t}/z_t\right)+\rho^{k-1}\varphi\left(1-\gamma_2\sqrt{h_t}/z_t\right).\end{aligned}$$

Notice again the simplicity due to the component structure. The impulse response on the term structure of variance is then

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

The top-left panels of Figures 3 and 4 plot the impulse responses to the term structure of variance for $h_t = \sigma^2$ and $z_t = 2$ and $z_t = -2$ respectively. The figures reinforce the message from Figures 1 and 2 that using parametrization estimated from the data, the component model is quite different from the GARCH(1, 1) model. The effects of shocks are much longer lasting in the component model using estimated parameter values because of the parameterization of the long-run component. Comparing across Figures 3 and 4 it is also clear that the term structure of the leverage effect is more flexible. As a result current shocks and the current state of the economy potentially have a much more profound impact on the pricing of options across maturities in the component model.

It has been argued in the literature that the hyperbolic rate of decay displayed by long memory processes may be a more adequate representation for the conditional variance of returns.[14] We do not disagree with these findings. Instead, we argue that Figures 1 through 4 demonstrate that in the component model the combination of two variance components with exponential decay gives rise to a slower decay pattern that sufficiently adequately captures the hyperbolic decay pattern of long memory processes for the horizons relevant for option valuation. This is of interest because although the long-memory model may be a more adequate representation of the data, it is harder to implement.

1.4 Option Valuation

We now turn to the ultimate purpose of this paper, namely the valuation of derivatives on an underlying asset with dynamic variance components. For the purpose of option valuation we first derive the conditional moment generating function for the return process and then present the risk-neutral return dynamics.

1.4.1 The Moment Generating Function

For the return dynamics in this paper we can characterize the moment generating function (MGF) of the log stock price with a set of difference equations using the techniques in Heston and Nandi (2000). Appendix A demonstrates that for the component GARCH model we have that the MGF defined by

$$f(t, T; \phi) \equiv E_t \left[\exp\left(\phi \ln\left(S_T\right)\right)\right]$$

[14] See Bollerslev and Mikkelsen (1996,1999), Baillie, Bollerslev and Mikkelsen (1996) and Ding, Granger and Engle (1993).

can be written

$$f(t,T;\phi) = S_t^{\phi}\exp[A_t + B_{1,t}(h_{t+1} - q_{t+1}) + B_{2,t}q_{t+1}] \tag{1.12}$$

with coefficients

$$A_t = A_{t+1} + r\phi - (\alpha B_{1,t+1} + \varphi B_{2,t+1}) - 1/2\ln(1 - 2\alpha B_{1,t+1} - 2\varphi B_{2,t+1}) + B_{2,t+1}\omega$$

$$B_{1,t} = B_{1,t+1}\tilde{\beta} + \lambda\phi + 2\frac{(\alpha\gamma_1 B_{1,t+1} + \varphi\gamma_2 B_{2,t+1} - 0.5\phi)^2}{1 - 2\alpha B_{1,t+1} - 2\varphi B_{2,t+1}}$$

$$B_{2,t} = B_{2,t+1}\rho + \lambda\phi + 2\frac{(\alpha\gamma_1 B_{1,t+1} + \varphi\gamma_2 B_{2,t+1} - 0.5\phi)^2}{1 - 2\alpha B_{1,t+1} - 2\varphi B_{2,t+1}}$$

and terminal conditions

$$A_T = B_{1,T} = B_{2,T} = 0.$$

For the moment generating function in the GARCH(1, 1) case we refer to Heston and Nandi (2000).

In Figures 1 and 2 we illustrated differences across models in terms of variance term structures that are key for option valuation. Following Das and Sundaram (1999), we can use the moment generating function in (1.12) to further investigate the conditional term structure of higher moments. Specifically, we can derive conditional skewness and excess kurtosis for maturity T using the logarithm of the conditional moment generating function as follows

$$Skewness(t,T) = \frac{\partial^3 \ln f(t,T;\phi)/\partial\phi^3\big|_{\phi=0}}{Var(t,T)^{3/2}} \qquad Kurtosis(t,T) = \frac{\partial^4 \ln f(t,T;\phi)/\partial\phi^4\big|_{\phi=0}}{Var(t,T)^2}$$

where

$$Var(t,T) = \partial^2 \ln f(t,T;\phi)/\partial\phi^2\big|_{\phi=0}.$$

We compute these moments by taking numerical derivatives of the log of the moment generating function in (1.12).

In Figure 5 we plot the term structure of skewness and kurtosis in the three GARCH models. The initial volatility is set to its long run value in the GARCH$(1, 1)$ and component GARCH models. In the persistent component model the initial volatility is set to the unconditional volatility from the component model. The parameter estimates are again taken from Table 1

Figure 5 reveals important differences between the term structures of these moments for the GARCH$(1, 1)$ model, the component model and the persistent component model. While the term structures of skewness and kurtosis are hump-shaped for the GARCH$(1, 1)$ model over the maturities relevant for pricing the options in our sample, they are downward sloping and upward sloping respectively for the skewness and kurtosis of the persistent component model. For the component model, the minimum and the maximum respectively for the conditional skewness and kurtosis occur for options with approximately a six month maturity, but the skewness and kurtosis for longer maturities are very close to these extrema. These fundamental differences in higher moment term structures may have important implications for the option valuation properties across models.

1.4.2 The Risk-Neutral GARCH(1,1) Dynamic

The risk-neutral dynamics for the GARCH$(1,1)$ model are given in Heston and Nandi (2000)[15] as

$$\begin{aligned} R_{t+1} &= r - \tfrac{1}{2}h_{t+1} + \sqrt{h_{t+1}}z^*_{t+1} \\ h_{t+1} &= w + bh_t + a(z^*_t - c^*\sqrt{h_t})^2 \end{aligned} \tag{1.13}$$

with $c^* = c + \lambda + 0.5$ and $z^*_t \sim N(0,1)$.

1.4.3 The Risk-Neutral Component GARCH Dynamic

Appendix B demonstrates that the risk-neutral component GARCH dynamic is given by

$$\begin{aligned} h_{t+1} &= q_{t+1} + \tilde{\beta}^*(h_t - q_t) + \alpha\left(\left(z^*_t - \gamma^*_1\sqrt{h_t}\right)^2 - \left(1 + \gamma^{*2}_1 h_t\right)\right) \\ q_{t+1} &= \omega + \rho^* q_t + \varphi\left(\left(z^*_t - \gamma^*_2\sqrt{h_t}\right)^2 - \left(1 + \gamma^{*2}_2 h_t\right)\right) \end{aligned} \tag{1.14}$$

where the risk neutral parameters are defined as follows

$$\begin{aligned} \tilde{\beta}^* &= \tilde{\beta} + \alpha\left(\gamma^{*2}_1 - \gamma^2_1\right) + \varphi\left(\gamma^{*2}_2 - \gamma^2_2\right) \\ \rho^* &= \rho + \alpha\left(\gamma^{*2}_1 - \gamma^2_1\right) + \varphi\left(\gamma^{*2}_2 - \gamma^2_2\right) \\ \gamma^*_i &= \gamma_i + \lambda + 0.5,\ i = 1, 2. \end{aligned}$$

The moment generating function for the risk-neutral component GARCH process is therefore equal to the one for the physical component GARCH process, setting $\lambda = -0.5$ and using the risk neutral parameters $\gamma^*_1, \gamma^*_2, \rho^*, \tilde{\beta}^*$ as well as ω, α and φ.

[15] For the underlying theory on risk neutral distributions in discrete time option valuation see Rubinstein (1976), Brennan (1979), Amin and Ng (1993), Duan (1995), Camara (2003), and Schroder (2004).

1.4.4 The Option Valuation Formula

Given the moment generating function and the risk-neutral dynamics, and option valuation is relatively straightforward. We use the result of Heston and Nandi (2000) that at time t, a European call option with strike price K that expires at time T is worth

$$\begin{aligned}\text{Call Price} &= e^{-r(T-t)}E_t^*[Max(S_T-K,0)] \qquad (15)\\ &= S_t\left(\frac{1}{2}+\frac{1}{\pi}\int_0^\infty \mathrm{Re}\left[\frac{K^{-i\phi}f^*(t,T;i\phi+1)}{i\phi S_t e^{r(T-t)}}\right]d\phi\right)\\ &\quad -Ke^{-r(T-t)}\left(\frac{1}{2}+\frac{1}{\pi}\int_0^\infty \mathrm{Re}\left[\frac{K^{-i\phi}f^*(t,T;i\phi)}{i\phi}\right]d\phi\right)\end{aligned}$$

where $f^*(t,T;i\phi)$ is the conditional characteristic function of the logarithm of the spot price under the risk neutral measure.

1.5 Empirical Results

This section presents the core of our empirical results. We first study the models estimated using a long time series of S&P500 index returns. We then report the differences between the models when used for option valuation and compare the models to a GARCH(1,1) model allowing for jumps in returns. Finally, we analyze the option valuation differences along various dimensions.

1.5.1 Properties of the Physical Return Process

Table 1 presents maximum likelihood estimates (MLE) of the physical model parameters obtained using returns data for 1962-2001. We use a long sample of returns on the S&P 500 because it is well-known that it is difficult to estimate GARCH parameters precisely using relatively short samples of return data. We compare the models using goodness-of-fit statistics, and we discuss differences in model properties. We present results for three models: the GARCH$(1,1)$ model (1.1), the component model (1.7) and the persistent component model (1.9). Almost all parameters are estimated significantly different from zero at conventional significance levels.[16] The price of risk, λ, is marginally significant in the case of the GARCH$(1,1)$ and not significant in the persistent component models. The log likelihood values indicate that the fit of the component model is much better than that of the persistent component model, which in turn fits much better than the GARCH$(1,1)$ model.

The dynamic variance models can be compared by assessing their persistence properties. The variance persistence in the GARCH$(1,1)$ model is defined by $\tilde{b} = b + ac^2$ from (1.10). In the component model, the total variance persistence is a confluence of the persistence in the two factors. If we substitute out q_{t+1} and q_t from the h_{t+1} equation in (1.7), then persistence can be computed as the as the sum of the coefficients on h_t and h_{t-1}. This way, the component persistence formula can be derived to be $\rho + \tilde{\beta}(1-\rho)$.

The improvement in fit for the component GARCH model over the persistent component GARCH model is perhaps somewhat surprising when inspecting the persistence of the component GARCH model. The persistence is equal to 0.9963. It therefore would appear that equating this persistence to 1, as is done in the persistent component model, is an interesting hypothesis, but apparently modeling these small differences from one is important. It must of course be noted that while the persistence of the long-run component (ρ) is

[16] The standard errors are computed using the outer product of the gradient at the optimal parameter values.

0.9896 for the component model as opposed to 1 for the persistent component model, the persistence of the short-run component ($\tilde{\beta}$) is 0.6437 versus 0.8822 and this may account for the differences in likelihood. Note that the persistence of the GARCH(1, 1) model is estimated at 0.9553, which is consistent with earlier literature. It is slightly lower than the estimate in Christoffersen, Heston and Jacobs (2006) and a bit higher than the average of the estimates in Heston and Nandi (2000).

Figures 6 and 7 further analyze the component models' improvement in performance over the benchmark GARCH(1, 1) model. These figures present the 1990-1995 sample path for the spot variance in the GARCH(1, 1) model, the component model and the persistent component model, as well as the sample path for volatility components for the component and persistent component models.[17] In each figure, the sample path is obtained by using the parameter estimates in Table 1 to iterate on the variance dynamic starting from the unconditional volatility 500 days before the first volatility included in the figure, as is done in estimation. Initial conditions are therefore unlikely to affect model comparisons. Figure 6 contains the results for the component model. The overall conclusion seems to be that the mean zero short run component in the top-right panel adds short-horizon noise around the long-run component in the bottom-right panel. This results in a volatility dynamic for the component model in the top-left panel that is more noisy than the volatility dynamic for the GARCH(1, 1) model in the bottom-left panel. This more noisy sample path suggests a higher value for the variance of variance in the component model. The results for the persistent component model in Figure 7 suggest similar conclusions, even though the sample

[17] We plot results for the 1990-1995 subsample here because it will be used for option valuation subsequently.

paths for the components in Figure 7 are somewhat different from those in Figure 6. Table 1 gives the average annual volatility (standard deviation) for 1990-1995 in the three models as 12.06% (GARCH$(1,1)$), 11.74% (component) and 12.39% (persistent component).

We now investigate in more detail differences between the models in the modeling of the standard deviation of the conditional variance, as well as differences in the modeling of the covariance and correlation between returns and variance. For option valuation, the conditional versions of these quantities and their variation through time are just as important as the unconditional versions. The conditional versions of return-variance covariance and variance of variance are computed as follows. For the GARCH$(1,1)$ model the conditional variance of variance is

$$\begin{aligned} Var_t(h_{t+2}) &= E_t\left[h_{t+2} - E_t\left[h_{t+2}\right]\right]^2 \\ &= 2a^2 + 4a^2c^2h_{t+1} \end{aligned} \tag{1.16}$$

and for the component and persistent component models, the conditional variance of variance is

$$Var_t(h_{t+2}) = 2\left(\alpha + \varphi\right)^2 + 4\left(\gamma_1\alpha + \gamma_2\varphi\right)^2 h_{t+1}. \tag{1.17}$$

In Figure 8 we use the parameters from Table 1 to plot the standard deviation of variance in the three GARCH models. Notice that the standard deviation of variance in the component model is in general much higher than in the GARCH$(1,1)$ model and it is also more volatile. The average level of the conditional standard deviation of variance in the persistent component is in between that of the other two models. Table 1 gives the average volatility of variance during 1990-1995 in the row labeled "Average Vol of Var".

If we think of the option price as being a function of the spot variance, then we can view variation in the option prices as being driven by the volatility of variance. The volatility of variance is also related to kurtosis. Figure 8 shows that the component model is able to generate richer time-varying kurtosis dynamics than the GARCH$(1,1)$ model and thus potentially richer option price dynamics.

The conditional covariance between return and variance in the GARCH$(1,1)$ model is given by

$$\begin{aligned} Cov_t(R_{t+1}, h_{t+2}) &= E_t\left[(R_{t+1} - E_t\left[R_{t+1}\right])\left(h_{t+2} - E_t\left[h_{t+2}\right]\right)\right] \\ &= E_t\left[\sqrt{h_{t+1}} z_{t+1}\left(a z_{t+1}^2 - 2ac z_{t+1}\sqrt{h_{t+1}} - a\right)\right] \\ &= -2ac h_{t+1}. \end{aligned} \tag{1.18}$$

Conditional correlation is easier to interpret than conditional covariance. The conditional correlation in the GARCH$(1,1)$ model is

$$Corr_t(R_{t+1}, h_{t+2}) = \frac{-2c\sqrt{h_{t+1}}}{\sqrt{2 + 4c^2 h_{t+1}}} \tag{1.19}$$

where we have used the conditional variance of variance from (1.16).

The conditional covariance in the component model is

$$Cov_t(R_{t+1}, h_{t+2}) = -2\left(\gamma_1 \alpha + \gamma_2 \varphi\right) h_{t+1} \tag{1.20}$$

and the conditional correlation in the component model is thus given by

$$Corr_t(R_{t+1}, h_{t+2}) = \frac{-2\left(\gamma_1 \alpha + \gamma_2 \varphi\right)\sqrt{h_{t+1}}}{\sqrt{2\left(\alpha + \varphi\right)^2 + 4\left(\gamma_1 \alpha + \gamma_2 \varphi\right)^2 h_{t+1}}}. \tag{1.21}$$

Figure 9 plots the conditional covariance (left panels) and correlation (right panels) for the three models. The conditional covariance and correlation is clearly more negative in the component models than in the GARCH$(1,1)$ model, and furthermore the component covariance paths are much more volatile. Table 1 gives the average correlations during 1990-1995 as -79.40% (GARCH$(1,1)$), -88.49% (component), and -90.14% (persistent component).

For the component and persistent component models, we can also compute the conditional correlations between the return and each volatility component separately

$$\begin{aligned} Corr_t(R_{t+1}, h_{t+2} - q_{t+2}) &= \frac{-2\gamma_1\sqrt{h_{t+1}}}{\sqrt{2+4\gamma_1^2 h_{t+1}}} \\ Corr_t(R_{t+1}, q_{t+2}) &= \frac{-2\gamma_2\sqrt{h_{t+1}}}{\sqrt{2+4\gamma_2^2 h_{t+1}}}. \end{aligned} \tag{1.22}$$

Figure 10 indicates that for both component models, the conditional correlation of the return with the short-run variance component is on average more negative than the conditional correlation between the return and the long-run variance component. This difference can be traced to Table 1 where $\gamma_1 > \gamma_2$ in both models. The correlations with the long-run factor are relatively more negative in the persistent component model whereas the correlations with the short-run factor are relatively more negative in the component model. This can also be traced back to Table 1 where γ_1 is larger in the component model than in the persistent component model; whereas γ_2 is largest in the persistent component model.

Figure 11 shows the correlation between returns and variances from a different perspective. We plot the correlations from (1.19) and (1.21) against levels of the conditional variance expressed in annual standard deviations. Notice that for all three models the relationship between the level of volatility and the correlation is negative. This is shown by Jones (2003) to be a desirable feature for option valuation and it is a feature missing in the standard Heston (1993) SV model where the correlation is constant. Interestingly, the Heston and Nandi (2000) GARCH$(1, 1)$ model does have this negative relationship as Figure 11 shows. Figure 11 also shows that when fitted on the more general component model, the return data wants a correlation which is more negative than the simple GARCH$(1, 1)$ model for all levels of volatility. The differences in correlation are quite large for the most common levels of volatility.

We conclude from Figures 8-11 that the more flexible component model is capable of generating not only more flexible term structures of variance, but also more flexible dynamics for the conditional correlation between returns and variance, and the conditional variance of variance. These dynamics are critically important for skewness and kurtosis dynamics which in turn are key for explaining the variation in index options prices. This is the topic to which we now turn.

1.5.2 Option Valuation Performance

We use a sample of six years of data on S&P 500 call options, for the period 1990-1995. Following Bakshi, Cao and Chen (1997), we apply standard filters to the data. We only use Wednesday options data. 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. For those weeks where Wednesday is a holiday, we use the next trading day. The decision to pick one day every week is to some extent motivated by computational constraints. Using only Wednesday data allows us to study a fairly long time-series, which is useful considering the highly persistent volatility processes. An additional motivation for using Wednesday data is that following the work of Dumas, Fleming and Whaley (1998), several studies have used this setup.[18]

Table 2 presents descriptive statistics for the options data for 1990-1995 by moneyness and maturity. Panels A and B indicate that the data are standard. We can observe the volatility smirk from Panel C and it is clear that the slope of the smirk differs across maturities. Descriptive statistics for different sub-periods (not reported here) demonstrate that the slope changes over time, but that the smirk is present throughout the sample. The top panel of Figure 12 gives some indication of the pattern of implied volatility over time. For the 312 days of options data used in the empirical analysis, we present the average implied volatility of the options on that day. It is evident from Figure 12 that there is substantial clustering in implied volatilities. It can also be seen that volatility is higher in the early part of the sample. The bottom panel of Figure 12 presents a time series for the 30-day at-the-money volatility (VIX) index from the CBOE for our sample period. A comparison with the top panel clearly indicates that the options data in our sample are representative of market conditions, although the time series based on our sample is of course a bit more noisy due to the presence of options with different moneyness and maturities.

The last row of Table 1 compares the performance of the four models for option valuation. We use the MLE parameter estimates in Table 1 to compute root mean squared errors (RMSEs) for the 1990-1995 option sample described above and various subsamples.[19] The most important conclusion is that the models' ranking is similar to the ranking

---

18 See for instance Heston and Nandi (2000).

19 As the price of risk parameter, λ, is poorly estimated in Table 1, and in order to keep the persistence at unity under both measures in the persistent model, we simply set $\lambda = -0.5$ across models. This way the other parameters are identical under the two measures.

based on the log likelihood. The GARCH$(1, 1)$ model is the worst performer based on the RMSE, as is the case using the log likelihood criterion, but the persistent component model achieves the lowest RMSE, followed by the component model–although the differences between the component and persistent models is much smaller in RMSE terms, than it were in log likelihood terms.

Table 3 provides additional evidence on the option fit of the three models. We report option RMSE by moneyness and maturity. The top panel reports the RMSE for the GARCH$(1, 1)$ model, while the two other panels report the ratio of the RMSE for the two other models to that of the GARCH$(1, 1)$ The improvements of the component models over the GARCH$(1, 1)$ model are fairly robust across maturity and moneyness. Importantly, the component models are never much worse than the GARCH$(1, 1)$ model and they fail to improve on the GARCH$(1, 1)$ model only for short term deep-in-the-money call options. This finding leads us to consider jumps in returns which by way of adding non-normality to the conditional density may lead to improvements in the valuation of short-term options.

1.5.3 Comparing with a GARCH(1,1)-Jump Model

The Heston-Nandi GARCH$(1, 1)$ model is a useful first benchmark, but it has well-known empirical biases. These biases are similar to those displayed by the Heston (1993) model. The continuous-time literature has attempted to improve the performance of the Heston (1993) model by adding to it (potentially correlated) jumps in returns and volatility, and this strategy has been partly successful. Poisson-normal jumps in returns and volatility improve option valuation when parameters are estimated using historical time series of returns. When model parameters are estimated using the cross-section of option prices, Poisson jumps usually do not lead to improved model fit, but Broadie, Chernov and Johannes (2004) find evidence of the importance of jumps for option pricing when imposing consistency between the physical and risk-neutral parameters.[20] Carr and Wu (2004) and Huang and Wu (2004) analyze Levy processes and find that they are better able to fit options.

In the discrete-time literature, some studies have attempted to address these model biases by combining conditional heteroskedasticity with non-normal innovations. This strategy may seem very different from including jumps in the return process, but both approaches essentially introduce conditional non-normalities in the return distribution. However, Christoffersen, Heston and Jacobs (2006) find that inverse Gaussian innovations do not improve out-of-sample model fit. We therefore use the approach proposed by Duan, Ritchken and Sun (2005, 2006), which combines the GARCH$(1,1)$ dynamic with a Poisson jump process similar to the one used in the continuous-time option valuation literature.[21]

We refer to the resulting model as the GARCH$(1,1)$-Jump model. The return dynamics can be written

$$\begin{aligned} R_{t+1} &= r + \lambda h_{t+1} + \chi\left(1 - \exp\left(\sqrt{h_{t+1}}\mu + \frac{h_{t+1}\tau^2}{2}\right)\right) + \sqrt{h_{t+1}}J_{t+1} \quad (1.23) \\ h_{t+1} &= w + bh_t + a\left(J_t - c\sqrt{h_t}\right)^2 \end{aligned}$$

where the variance dynamic h_{t+1} has the affine structure from Heston and Nandi (2000), and J_t is a standard normal random variable plus a Poisson random sum of normal random

[20] For evidence on the importance of Poisson-normal jumps, see for example Andersen, Benzoni and Lund (2002), Bakshi, Cao and Chen (1997), Bates (1996, 2000), Chernov, Gallant, Ghysels and Tauchen (2003), Eraker, Johannes and Polson (2003), Eraker (2004) and Pan (2002).

[21] See also Maheu and McCurdy (2004).

variables. In particular,

$$J_t = X_t^{(0)} + \sum_{j=1}^{N_t} X_t^{(j)}$$

with

$$X_t^{(0)} \sim N(0,1) \text{ and } X_t^{(j)} \sim N\left(\mu, \tau^2\right) \text{ for } j = 1, 2, ..., N_t.$$

and N_t is the Poisson random variable with constant intensity χ. When $\chi = 0$, J_t is a standard normal variable.

Duan, Ritchken and Sun (2005, 2006) formulate sufficient conditions to derive a risk-neutral process for the GARCH(1, 1)-Jump model that takes the same form as (1.23), and that has the following parameterization $\lambda^* = -0.5$, $J_t^* = J_t + \frac{1}{2}\sqrt{h_t} + \lambda\sqrt{h_t}$, $c^* = c + \frac{1}{2} + \lambda$. Unfortunately, no closed-form solution exists for the GARCH(1, 1)-Jump model so that option prices must be computed by Monte Carlo simulation.

Table 4 reports the empirical results for the GARCH(1, 1)-Jump model. Panel A reports the parameter estimates from maximum likelihood estimation on the sample of daily S&P500 returns used in Table 1. Again, all the parameters except for λ are significant. Notice that the log likelihood value is considerably larger than for the three models in Table 1. The GARCH(1, 1)-Jump model thus gives a good description of the conditional density for daily S&P500 returns. Notice however, that the Option RMSE for the GARCH(1, 1)-Jump model is \$2.138 which is only marginally better than the \$2.236 for GARCH(1, 1) in Table 1, and much worse than the \$1.706 and \$1.705 for the component model and persistent model respectively.

Panel B in Table 4 shows the ratio of the RMSE of the GARCH(1, 1)-Jump to the GARCH(1, 1) model. The Jump model in general performs close to the GARCH(1, 1) across moneyness and maturity. The best relative performances is for short term in-the-money calls (0.897) and the worst is for long term deep in-the-money calls (1.043). Somewhat surprisingly, the GARCH(1, 1)-Jump model outperforms the GARCH(1, 1) model by a smaller margin for shorter maturities in general than for longer maturities.

The lack of improvement offered by the jump model is surprising. We suspect that any of the following reasons could be the culprit. First, Poisson jumps may be quantitatively more important for short maturity options when combined with a continuous-time stochastic volatility model than a discrete-time GARCH model, because of the continuous sample path. Second, Eraker (2004) finds that adding jumps do not improve the out-of-sample option valuation performance of a standard SV model. The jump parameters may simply be difficult to estimate reliably–perhaps because they are changing over time. Third, jumps may improve the likelihood function for daily returns without improving much the conditional density function for 7-365 day returns that is relevant for option valuation. Fourth, other specifications of conditional normality may work better than the jump specification chosen here, but a full investigation of non-normal innovations in GARCH models is beyond the scope of this paper. Panel B of Table 4 shows strong similarities between the pricing error patterns of the GARCH(1, 1) and GARCH(1, 1)-Jump models, we will therefore restrict attention to the conditional normal models in the analysis below.

1.5.4 Analyzing the Option Valuation Performance

It must be emphasized that the component models' performance is remarkable and to some extent surprising. First, the GARCH(1, 1) model is a good benchmark which itself has a very solid empirical performance (see Heston and Nandi (2000)). The model captures important stylized facts about option prices such as volatility clustering and the leverage effect (or equivalently negative skewness). When estimating models from option prices, Christoffersen and Jacobs (2004) find that GARCH models with richer news impact parametrization do not improve the model fit out-of-sample. Christoffersen, Heston and Jacobs (2006) find that a GARCH model with non-normal innovations improves the model's fit in-sample and for short out-of-sample horizons, but not for long out-of-sample horizons. Although we do not report the results in the paper, we have also compared the performance of the GARCH$(1, 1)$ model with the implied Black-Scholes model in Dumas, Fleming and Whaley (1998). We confirm the finding of Heston and Nandi (2000) that the GARCH$(1, 1)$ model outperforms the implied Black-Scholes model out-of-sample. Furthermore, the analysis in Tables 3 and 4 demonstrates that the component model also provides a better option fit than the GARCH$(1, 1)$ model augmented with Poisson-normal jumps.

We now provide some more insight behind the improved performance of the component models by analyzing the differences across models along three critical dimensions: the (spot) volatility level, the volatility term structure and the modeling of the smirk. First, component models may better match the volatility patterns over time. We investigate this by comparing the differences in the time paths between implied volatilities from the data and the models. Second, it may be the case that the component models more adequately capture the term structure of volatility than the GARCH$(1, 1)$ model. We investigate this by comparing the models' term structures of implied volatility for at-the-money options.

Third, it may be the case that the component models better capture the implied volatility smirk at various maturities. We study the differences between the models in this dimension for different initial levels of volatility.

Figure 13 presents the average weekly implied volatility bias (average observed market implied volatility less average model implied volatility) over the 1990-1995 option sample, using the MLE estimates from Table 1. Clearly the component models outperform the GARCH$(1, 1)$ model in this dimension: The GARCH$(1, 1)$ model shows significant underpricing (positive bias) during the high volatility episode in 1990-1991 and extended periods of overpricing (negative bias) during the low volatility period in 1993-1995. In comparison, the component model has smaller (positive) bias in 1990-1991 and also smaller (negative) bias in 1993-1995, suggesting that it is much better able to capture the dynamics of market volatility. The persistent model has the smallest (positive) bias in late 1990 but instead has significant (negative) bias in early 1990 and in late 1991.

Figure 14 studies the implied volatility term structure for at-the-money options in the three models. For each model, we use three different levels of initial spot volatility: we set spot volatility to 1/2, 1 or 2 times the unconditional volatility respectively. We again use the MLE parameters from Table 1 to compute option prices. The differences between the models are very pronounced. The critical difference between the models in the term structure dimension is that in the component model, the initial volatility is much more important for the valuation of longer maturity options than in the GARCH$(1, 1)$ model, and even more so in the persistent component model. Put differently, in the GARCH$(1, 1)$ model, today's level of volatility has virtually no impact on the implied volatility for 1-year to maturity options. For the component model, the initial volatility has an effect on the implied volatility for 1-year maturity options, and in the fully persistent model the effect of initial volatility is as large at the 1-year maturity as it is at short maturities. The three models are thus fundamentally different along this dimension.

Figure 15 analyzes a third source of differences in fit between the models. For each of the models, we plot moneyness smirks at three different maturities: 30, 90 and 365 days to maturity (DTM). Following the exposition in Figure 14, we repeat the analysis for three different levels of initial volatility. Figure 15 indicates that there are some differences between the models with respect to their ability to generate steep slopes in the smirk. The smirk for a one-year option is steeper for the persistent component model than it is for the GARCH$(1,1)$ model. However, the evidence suggests that the model differences in the moneyness dimension may be less important than the model differences in the term structure dimension. In the case of the GARCH$(1,1)$ model, the smirk is nearly identical for a one-year option, regardless of the level of initial volatility. For the persistent component model, the initial volatility level impacts on the level of the smirk, but does not greatly impact on the slope. The component model falls between these two cases.

We conclude that there are important differences between the GARCH$(1,1)$, component and persistent component model in terms of the path of spot volatility and the term structure of volatility, but all three models seem to be able to generate volatility smirks at different maturities. The differences between the models in this dimension seem quantitatively less important than differences in the models' volatility term structures.

1.6 Estimation Using Option Price Information

So far we have used the option price information only to evaluate the different models. However, it stands to reason that the observed option prices should be helpful in estimating the models as well. In this section we therefore implement the GARCH$(1,1)$, component, and persistent component model by minimizing the mean squared option valuation error rather than maximizing the daily return likelihood as we did in Table 1.

To be specific, we obtain parameters by minimizing the dollar mean squared error

$$\$MSE = \frac{1}{N^T} \sum_{t,i} \left(C_{i,t}^D - C_{i,t}^M \right)^2 \tag{1.24}$$

where $C_{i,t}^D$ is the market price of option i at time t, $C_{i,t}^M$ 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. The variance dynamic is used to update the variance from one Wednesday to the next using daily returns and the option valuation formula in (1.15) is used to compute the model prices on each Wednesday. The volatility updating rule is applied to the 500 days predating the Wednesday used in the estimation exercise, and it is initialized at the model's unconditional variance.

Needless to say, this nonlinear least squares (NLS) estimation techniques is much more computationally intensive than the simple MLE on returns in Table 1. For each function evaluation performed by the numerical optimizer, thousands of option prices must be calculated. The optimizer performs many function evaluations for each parameter update and consequently it is crucial to be able to compute option prices quickly and reliably. The pricing formula in (1.15) makes this estimation technique feasible. As we unfortunately do not have a closed-form pricing formula for the GARCH(1, 1)-Jump model we do not consider that model in this section.

Table 5 presents parameter estimates obtained using the 1990-1992 options data and in-sample RMSEs for the 1990-1992 data, as well as out-of-sample RMSEs using the 1993 data. Note that the shortest maturity is seven days because options with very short maturities were filtered out. Table 6 presents parameter estimates obtained using options data for 1992-1994, as well as 1992-1994 in-sample and 1995 out-of-sample option RMSEs. Tables 7-10 present RMSE results by moneyness and maturity for the two in-sample and two out-of-sample periods.[22]

In Table 5 we present results for the 1990-1992 period (in-sample) and the 1993 period (out-of-sample). The standard errors indicate that almost all parameters are estimated significantly different from zero.[23] There are some interesting differences with the parameters estimated from returns in Table 1, but the parameters are mostly of the same order of magnitude. This is also true for critical determinants of the models' performance, such as average annual volatility, average volatility of variance and average return correlation. Note also that the persistence of the short-run components and the long-run components is not dramatically different from Table 1. The persistence of the GARCH(1, 1) process is higher than in Table 1 though. In fact, it is interesting to note that the persistence of the GARCH(1, 1) model and the component GARCH model is close to one. This of course motivates the use of the persistent component model, where the persistence is restricted to

[22] Notice from the risk-neutral dynamics (1.13) and (1.14) that the parameter λ is not separately identified using option prices. We therefore simply set $\lambda = -0.5$ and we do not report λ in Tables 5-6.

[23] The standard errors are again computed using the outer product of the gradient at the optimum.

be one. Note also that the average correlation between return and volatility is now close to minus one in all three models.

Table 5 contains two sets of RMSEs. The RMSEs in the leftmost columns (NLS) are obtained using the parameter values in the Table. In the rightmost column, we report RMSEs based on parameter values obtained from MLE in Table 1. First consider the RMSEs obtained using NLS. In the in-sample 1990-1992 period, the RMSE of the component model is 89.7% of that of the benchmark GARCH$(1,1)$ model. For the out-of-sample 1993 period, the ratio of the RMSEs is 76.5%. For the persistent component model, the ratios are 94.8% and 93.3% respectively. Using the MLE estimates, the relative RMSEs are similar for the component model: 84.9% in 1990-1992 and 71.0% in 93. Using the MLE estimates the persistent models performs relatively worse in 1990-1992 with 113.0% but better in 1993 with 56.5% of the RMSE for the GARCH$(1,1)$. Naturally, when comparing across MLE and NLS estimates the RMSEs from NLS are typically much smaller than those from MLE. The information in option prices is clearly very valuable for estimating the models. Interestingly, the only example where the RMSE from MLE is close to that of the NLS counterpart is for the persistent model in the 1993 out-of-sample period.

Table 6 presents the results for the 1992-1994 period (in-sample) and the 1995 period (out-of-sample). The results largely confirm those obtained in Table 5. The most important difference is that the in-sample and out-of-sample performance of the component model is even better relative to the benchmark, as compared with the results in Table 5. When using NLS estimates component model's RMSE is 77.3% of that of the GARCH$(1,1)$ model for the 1992-1994 in-sample period, and for the 1993 out-of-sample period the ratio is 79.2%.

For the persistent model the ratios are 95.6% and 95.7%. When using MLE estimates the non-normal model is 76.2% of the GARCH(1, 1) in 1992-1994 and 69.2% in 1995. The persistent model performs very well relative to the GARCH(1, 1) MLE generating a 70.1% relative RMSE for 1992-1994 and 45.0% in 1995.

Comparing RMSEs across NLS and MLE parameters, we again find that the option prices add important information and drive the NLS RMSEs down from their MLE levels. Interestingly, the only case where the RMSE from MLE comes close to that from NLS is for the out-of-sample persistent model. Other interesting differences with Table 5 are that the persistence of the short-run non-normal is much higher, and that the persistence of the GARCH(1, 1) process in Table 6 is lower than in Table 5 but in line with the MLE estimate in Table 1.

Tables 7-10 provide a more detailed analysis of moneyness and maturity effects by presenting RMSE results by moneyness and maturity, using the parameter estimates from Tables 5 and 6. In each table, Panel A contains the RMSE for the GARCH(1, 1) model. To facilitate the interpretation of the table, panels B and C contain RMSEs that are normalized by the corresponding RMSE for the GARCH(1, 1) model. It is clear that an overall RMSE which is not too different across the three models in Tables 5 and 6 can mask large differences in the models' performance for a given moneyness/maturity cell. Inspection of the out-of-sample results in Tables 8 and 10 is especially instructive. We conclude that the improved out-of-sample performance of the non-normal models is due to the improved valuation of long-maturity options. This is an interesting affirmation of the intuition obtained previously in Figures 1-5 and 14. The richer volatility dynamics in the non-normal model enables richer explanations of variations in long-term option prices.

Overall, we conclude that based on the parameter values obtained using NLS, the performance of the component GARCH model is very impressive. Its RMSE is between 76.5% and 89.7% of the RMSE of the benchmark GARCH(1, 1) model. The performance of the persistent component model is less impressive, both in-sample and out-of-sample. However the persistent component model performs relatively well in the out-of-sample experiments when MLE parameters are used. This suggests that the persistent component model may be valuable for option valuation in cases where no option price information is available.

1.7 Conclusion and Directions for Future Work

This paper presents a new option valuation model based on the work by Engle and Lee (1999) and Heston and Nandi (2000). The empirical performance of the new variance component model is significantly better than that of the benchmark GARCH(1, 1) model, in-sample as well as out-of-sample, and regardless of the information used in estimation. This is an important finding because the literature has demonstrated that it is difficult to find empirical models that improve on the GARCH(1, 1) model or the Heston (1993) model. We also compare the component model to a GARCH(1, 1)-Jump model, which combines conditional heteroskedasticity with Poisson-normal jumps. The GARCH(1, 1)-Jump model achieves a better statistical fit than the component model in-sample, but the component model performs far better when using the parameter estimates to fit options.

An important aspect of the model's improved performance is that its richer parameterization allows for improved joint modeling of long-maturity and short-maturity options. The model captures the stylized fact that shocks to current conditional volatility impact on the conditional variance forecast up to a year in the future, which results in a very different implied volatility term structure for at-the-money options. The component model also results in a different path for spot volatility compared to the GARCH$(1, 1)$ model, but in the moneyness dimension the differences with the GARCH$(1, 1)$ model seem relatively less important. The component model is also characterized by term structures of skewness and kurtosis that are very different from those of the GARCH$(1, 1)$ model.

Because the estimated persistence of the model is close to one, we also investigate a special case of our model in which shocks to the variance never die out. When estimating model parameters by maximum likelihood using a historical time series of returns, the persistent component model is somewhat inferior to the component model when judged by the likelihood criterion. When the MLE parameters are used to price options, the persistent component model performs similarly to the component model in terms of overall fit. When model parameters are estimated from option prices, the component model significantly outperforms the other models both in and out-of-sample. We also find that for a given model the parameters obtained from historical return data always lead to higher RMSEs than the parameters directly estimated from option data.

Given the success of the proposed component models, a number of further extensions to this work are warranted. First, the empirical performance of the model should of course be validated using other datasets. In particular, it would be interesting to test the model using LEAPS data, because the model may excel at modeling long-maturity LEAPS options. In this regard a direct comparison between component and fractionally integrated volatility models may be interesting. It could also be useful to combine the stylized features of the model with other modeling components that improve option valuation. One interesting experiment could be to replace the jump innovations considered in this paper by a another non-Gaussian distribution. Combining the model in this paper with the inverse Gaussian shock model in Christoffersen, Heston and Jacobs (2006) may be a viable approach. Finally, in this paper we have proposed a component model that gives a closed form solution using results from Heston and Nandi (2000) who rely on an affine GARCH model. We believe that this is a logical first step, but the affine structure of the model may be restrictive in ways that are not immediately apparent. It may therefore prove worthwhile to investigate non-affine variance component models.

1.8 Appendix

1.8.1 MGF of the Component GARCH model

This Appendix derives the moment generating function for the component GARCH process. The component GARCH process is given by

$$\begin{aligned} h_{t+1} &= q_{t+1} + \tilde{\beta}\left(h_t - q_t\right) + \alpha\left((z_t - \gamma_1\sqrt{h_t})^2 - (1+\gamma_1^2 h_t)\right) \\ q_{t+1} &= \omega + \rho q_t + \varphi\left((z_t - \gamma_2\sqrt{h_t})^2 - (1+\gamma_2^2 h_t)\right). \end{aligned}$$

Let $x_t = \ln(S_t)$. For convenience we will denote the time t conditional generating function of S_T (or equivalently the conditional moment generating function (MGF) of x_T) by f_t instead of the more cumbersome $f(t;T,\phi)$. By definition

$$f_t = E_t[\exp(\phi x_T)].$$

We shall guess that the moment generating function has the log-linear form. We again use the more parsimonious notation A_t to indicate $A(t;T,\phi)$.

$$f_t = \exp\left(\phi x_t + A_t + B_{1,t}(h_{t+1} - q_{t+1}) + B_{2,t}q_{t+1}\right). \tag{1.26}$$

We have the terminal condition $A_T = B_{i,T} = 0$. Applying the law of iterated expectations to f_t we get

$$f_t = E_t\left[f_{t+1}\right] = E_t \exp\left(\phi x_{t+1} + A_{t+1} + B_{1,t+1}(h_{t+2} - q_{t+2}) + B_{2,t+1}q_{t+2}\right).$$

Substituting the dynamics of x_t gives

$$
\begin{aligned}
f_t &= E_t \exp\left(\begin{array}{c} \phi(x_t + r) + \phi\lambda h_{t+1} + \phi\sqrt{h_{t+1}} z_{t+1} + A_{t+1} + B_{1,t+1}(h_{t+2} - q_{t+2}) + \\ B_{2,t+1} q_{t+2} \end{array}\right) \\
&= E_t \exp\left(\begin{array}{c} \phi(x_t + r) + \phi\lambda h_{t+1} + \phi\sqrt{h_{t+1}} z_{t+1} + A_{t+1} + \\ B_{1,t+1}\left(\tilde{\beta}\left(h_{t+1} - q_{t+1}\right) + \alpha\left((z_{t+1} - \gamma_1\sqrt{h_{t+1}})^2 - (1 + \gamma_1^2 h_{t+1})\right)\right) + \\ B_{2,t+1}\left(\omega + \rho q_{t+1} + \varphi\left((z_{t+1} - \gamma_2\sqrt{h_{t+1}})^2 - (1 + \gamma_2^2 h_{t+1})\right)\right) \end{array}\right) \\
&= E_t \exp\left(\begin{array}{c} \phi(x_t + r) + \phi\lambda h_{t+1} + A_{t+1} + B_{1,t+1}\tilde{\beta}\left(h_{t+1} - q_{t+1}\right) + \\ B_{2,t+1}(\omega + \rho q_{t+1}) - (\alpha B_{1,t+1} + \varphi B_{2,t+1}) + \\ (\alpha B_{1,t+1} + \varphi B_{2,t+1})\left(z_{t+1} - \frac{\alpha\gamma_1 B_{1,t+1} + \varphi\gamma_2 B_{2,t+1} - 0.5\phi}{(\alpha B_{1,t+1} + \varphi B_{2,t+1})}\sqrt{h_{t+1}}\right)^2 - \\ \frac{(\alpha\gamma_1 B_{1,t+1} + \varphi\gamma_2 B_{2,t+1} - 0.5\phi)^2}{(\alpha B_{1,t+1} + \varphi B_{2,t+1})} h_{t+1} \end{array}\right).
\end{aligned}
$$

Using the result

$$
E\left[\exp(x(z+y)^2)\right] = \exp\left(-\frac{1}{2}\ln(1 - 2x) + xy^2/(1 - 2x)\right)
$$

we get

$$
f_t = E_t \exp\left(\begin{array}{c} \phi(x_t + r) + A_{t+1} - (\alpha B_{1,t+1} + \varphi B_{2,t+1}) \\ -1/2\ln\left(1 - 2\alpha B_{1,t+1} - 2\varphi B_{2,t+1}\right) + B_{2,t+1}\omega + \\ B_{1,t+1}\tilde{\beta}\left(h_{t+1} - q_{t+1}\right) + B_{2,t+1}\rho q_{t+1} + \\ \left(\lambda\varphi + 2\frac{(\alpha\gamma_1 B_{1,t-1} + \varphi\gamma_2 B_{2,t+1} - 0.5\phi)^2}{1 - 2\alpha B_{1,t+1} - 2\varphi B_{2,t+1}}\right) h_{t+1} \end{array}\right). \tag{1.27}
$$

Matching terms in (1.27) and (1.26) gives

$$
\begin{aligned}
A_t &= A_{t+1} + r\phi - (\alpha B_{1,t+1} + \varphi B_{2,t+1}) - 1/2\ln\left(1 - 2\alpha B_{1,t+1} - 2\varphi B_{2,t+1}\right) + B_{2,t+1}\omega \\
B_{1,t} &= B_{1,t+1}\tilde{\beta} + \lambda\phi + 2\frac{(\alpha\gamma_1 B_{1,t+1} + \varphi\gamma_2 B_{2,t+1} - 0.5\phi)^2}{1 - 2\alpha B_{1,t+1} - 2\varphi B_{2,t+1}} \\
B_{2,t} &= B_{2,t+1}\rho + \lambda\phi + 2\frac{(\alpha\gamma_1 B_{1,t+1} + \varphi\gamma_2 B_{2,t+1} - 0.5\phi)^2}{1 - 2\alpha B_{1,t+1} - 2\varphi B_{2,t+1}}.
\end{aligned}
$$

1.8.2 Risk Neutralization of the Component GARHC model

The physical Component GARCH dynamic is given by

$$\begin{aligned}
\ln(S_{t+1}) &= \ln(S_t) + r + \lambda h_{t+1} + \sqrt{h_{t+1}} z_{t+1} \\
h_{t+1} &= q_{t+1} + \tilde{\beta}\left(h_t - q_t\right) + \alpha\left(\left(z_t - \gamma_1\sqrt{h_t}\right)^2 - \left(1 + \gamma_1^2 h_t\right)\right) \quad (1.28) \\
q_{t+1} &= \omega + \rho q_t + \varphi\left(\left(z_t - \gamma_2\sqrt{h_t}\right)^2 - \left(1 + \gamma_2^2 h_t\right)\right) \quad (1.29)
\end{aligned}$$

Under the risk neutral measure, we need $E^*\left[S_{t+1}/S_t\right] = \exp(r)$, which requires that

$$\ln(S_{t+1}) = \ln(S_t) + r - 0.5h_{t+1} + \sqrt{h_{t+1}} z_{t+1}^*.$$

This implies in turn that

$$z_{t+1}^* = z_{t+1} + (\lambda + 0.5)\sqrt{h_{t+1}}. \quad (1.30)$$

We also need to ensure that

$$Var_t\left(R_{t+1}\right) = Var_t^*\left(R_{t+1}\right).$$

In order to have the same conditional variances under the two measures, we need to have the same variance innovations under the two measures. Thus we need

$$\left(z_t - \gamma_i\sqrt{h_t}\right)^2 = \left(z_t^* - \gamma_i^*\sqrt{h_t}\right)^2 \qquad i = 1, 2$$

which can be achieved by defining a new risk neutral parameter

$$\gamma_i^* = \gamma_i + \lambda + 0.5, i = 1, 2.$$

Consider the following candidate for the risk-neutral Component GARCH dynamic

$$\begin{aligned}
h_{t+1} &= q_{t+1} + \tilde{\beta}^*\left(h_t - q_t\right) + \alpha\left(\left(z_t^* - \gamma_1^*\sqrt{h_t}\right)^2 - \left(1 + \gamma_1^{*2} h_t\right)\right) \quad (1.31) \\
q_{t+1} &= \omega + \rho^* q_t + \varphi\left(\left(z_t^* - \gamma_2^*\sqrt{h_t}\right)^2 - \left(1 + \gamma_2^{*2} h_t\right)\right) \quad (1.32)
\end{aligned}$$

where $z_t^* \sim N(0,1)$ and the risk neutral parameters are defined as follows

$$\begin{aligned} \tilde{\beta}^* &= \tilde{\beta} + \alpha\left(\gamma_1^{*2} - \gamma_1^2\right) + \varphi\left(\gamma_2^{*2} - \gamma_2^2\right) \\ \rho^* &= \rho + \alpha\left(\gamma_1^{*2} - \gamma_1^2\right) + \varphi\left(\gamma_2^{*2} - \gamma_2^2\right). \end{aligned} \tag{1.33}$$

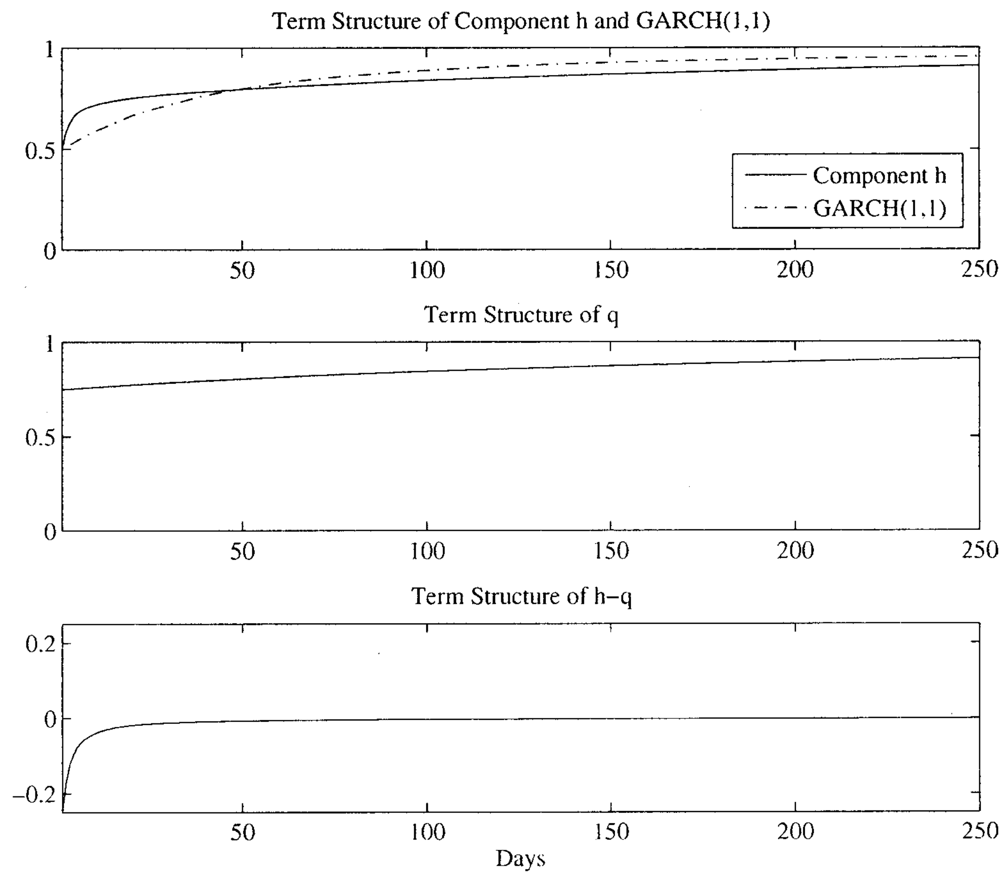
For this candidate risk-neutral dynamic to be valid, we have to verify that it is consistent with (1.28) and (1.29). Using (1.30), (1.33) and (1.32) in (1.31) we get

$$\begin{aligned} h_{t+1} &= \omega + \rho q_t + \varphi\left(\left(z_t - \gamma_2\sqrt{h_t}\right)^2 - \left(1 + \gamma_2^2 h_t\right)\right) + \tilde{\beta}\left(h_t - q_t\right) + \ldots \\ &\quad \alpha\left(\left(z_t - \gamma_1\sqrt{h_t}\right)^2 - \left(1 + \gamma_1^2 h_t\right)\right) \end{aligned}$$

which is identical to what we get using the physical component GARCH dynamic (1.28) and (1.29).

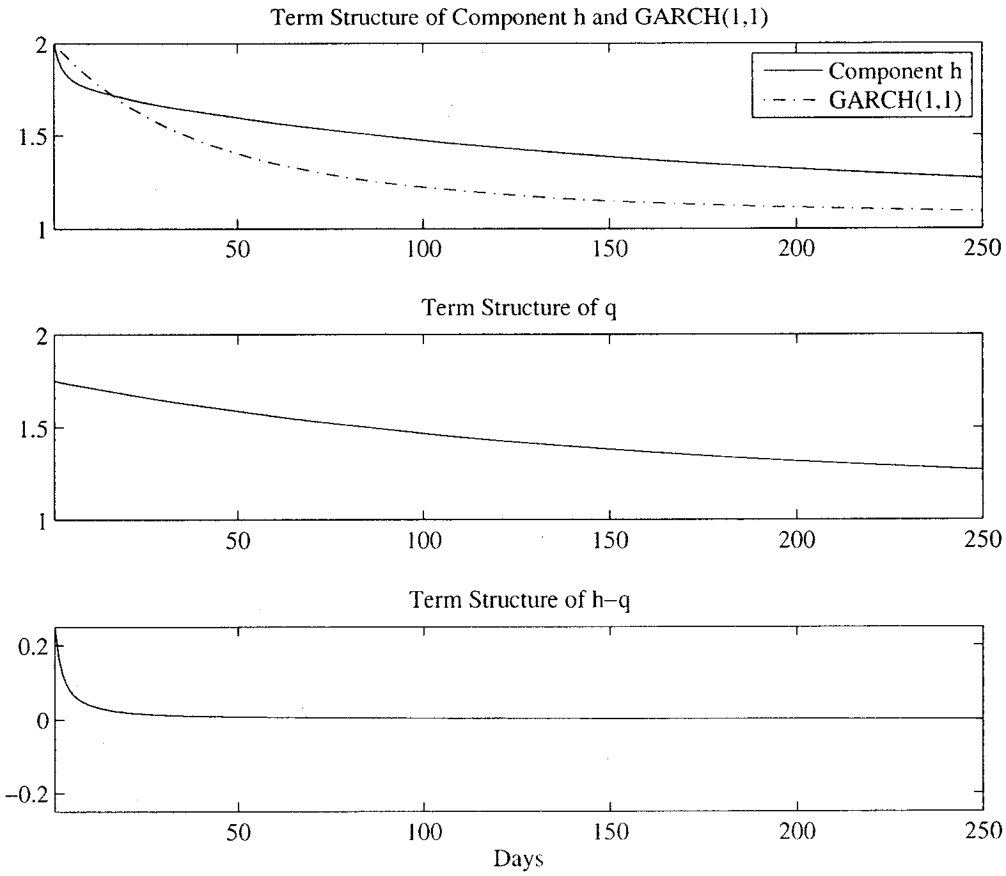
1.9 Figures and Tables

Figure 1. Term Structure of Variance with Low Initial Variance, Component Model Versus GARCH(1,1). Normalized by Unconditional Variance



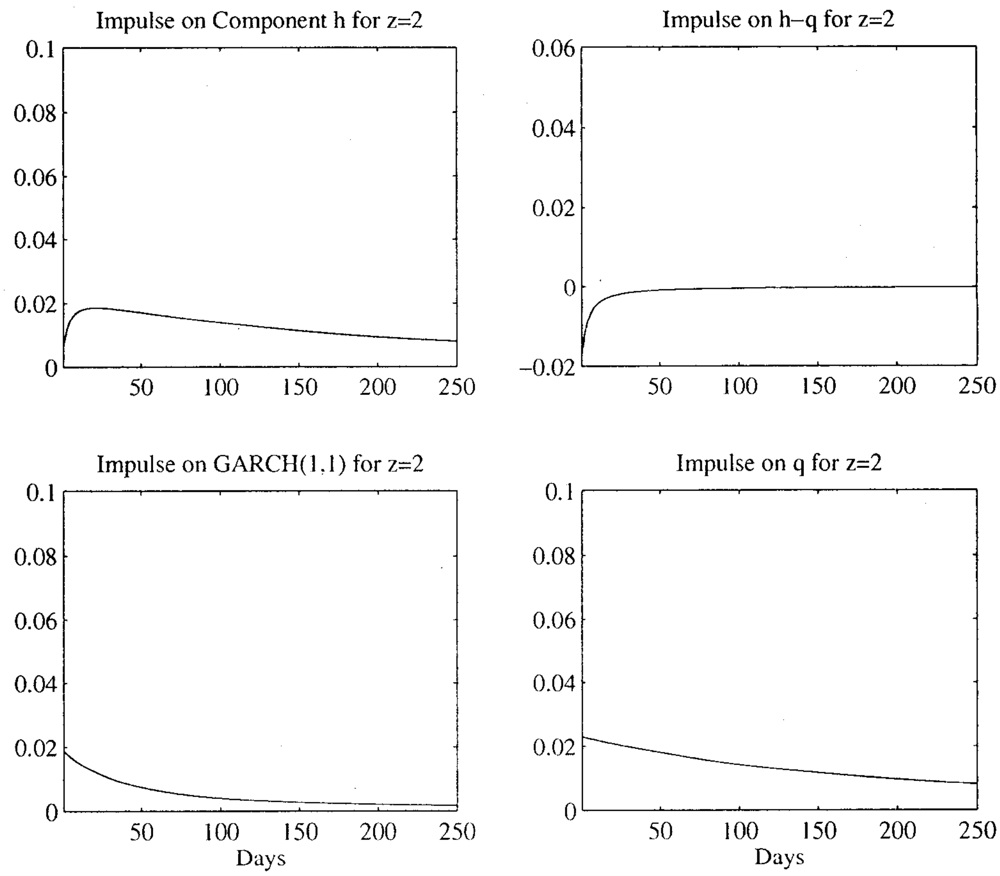
Notes to Figure: In the top panel we plot the variance term structure implied by the component GARCH and GARCH(1, 1) models for 1 through 250 days. In the second and third panel we plot the term structure of the individual components for the component model. The parameter values are obtained from MLE estimation on returns in Table 1. The initial value of q_{t+1} is set to $0.75\sigma^2$ and the initial value of h_{t+1} is set to $0.5\sigma^2$. The initial value for h_{t+1} in the GARCH(1, 1) is set to $0.5\sigma^2$ as well. All values are normalized by the unconditional variance σ^2.

Figure 2. Term Structure of Variance with High Initial Variance, Component Model Versus GARCH(1,1). Normalized by Unconditional Variance



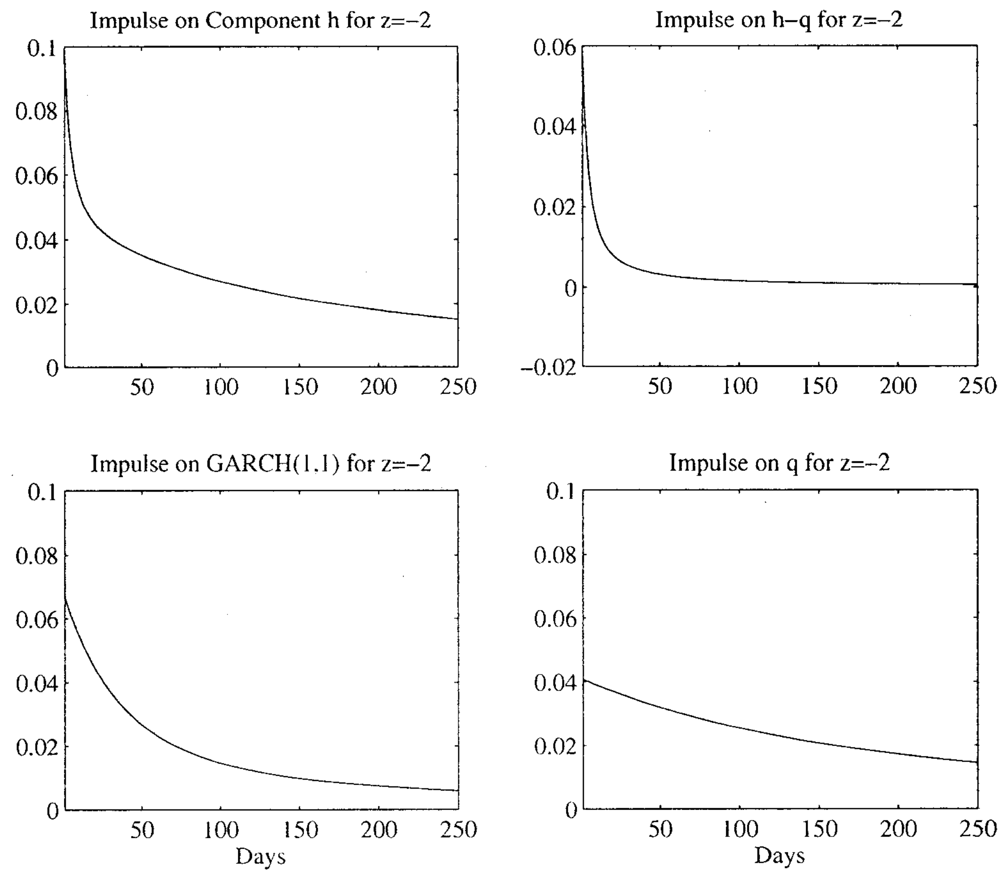
Notes to Figure: In the top panel we plot the variance term structure implied by the component GARCH and GARCH(1, 1) models for 1 through 250 days. In the second and third panel we plot the term structure of the individual components for the component model. The parameter values are obtained from MLE estimation on returns in Table 1. The initial value of q_{t+1} is set to $1.75\sigma^2$ and the initial value of h_{t+1} is set to $2\sigma^2$. The initial value for h_{t+1} in the GARCH(1, 1) is set to $2\sigma^2$ as well. All values are normalized by the unconditional variance σ^2.

Figure 3. Term Structure Impulse Response to Positive Return Shock ($z_t = 2$), Component Model Versus GARCH(1,1). Normalized by Unconditional Variance



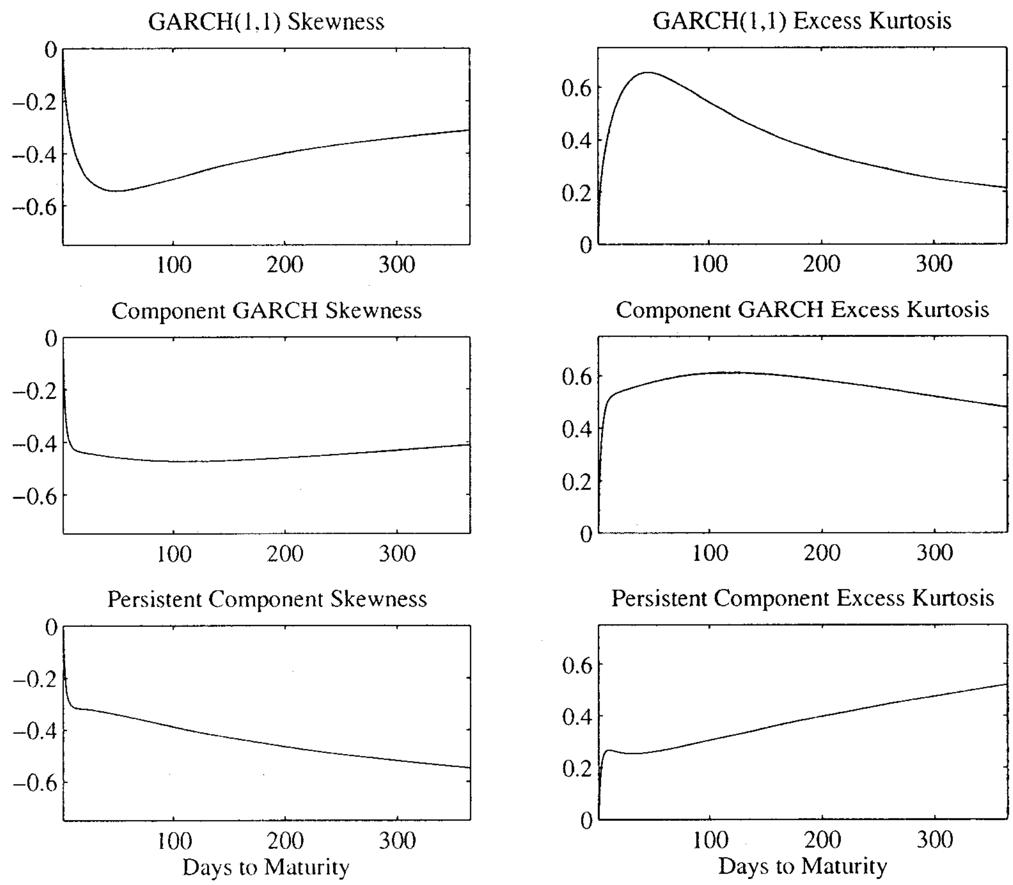
Notes to Figure: In the left-hand panels we plot the variance term structure response to a $z_t = 2$ shock to the return in the component and GARCH(1, 1) models. For the component model, the right-hand panels show the response of the individual components. The parameter values are obtained from the MLE estimation on returns in Table 1. The current variance is set equal to the unconditional value. All values are normalized by the unconditional variance.

Figure 4. Term Structure Impulse Response to Negative Return Shock ($z_t = -2$), Component Model Versus GARCH(1,1). Normalized by Unconditional Variance



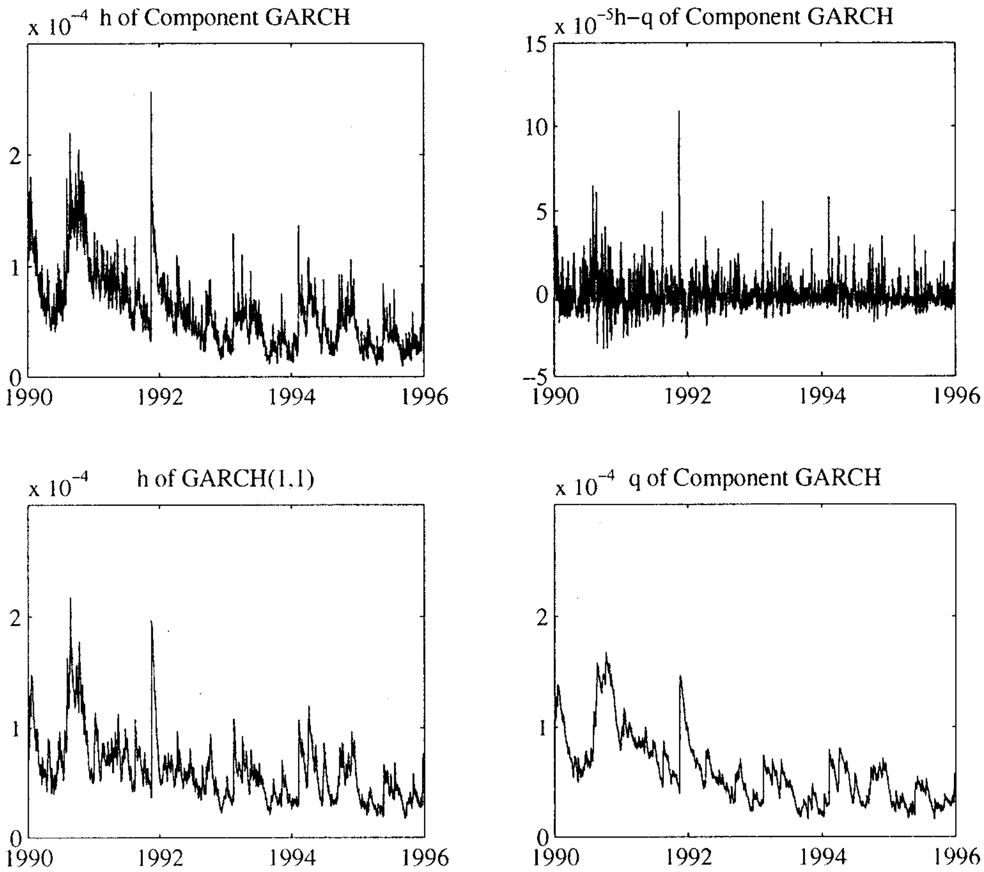
Notes to Figure: In the left-hand panels we plot the variance term structure response to a $z_t = -2$ shock to the return in the component and GARCH$(1, 1)$ models. For the component model, the right-hand panels show the response of the individual components. The parameter values are obtained from the MLE estimation on returns in Table 1. The current variance is set equal to the unconditional value. All values are normalized by the unconditional variance.

Figure 5: Term Structure of Skewness and Kurtosis



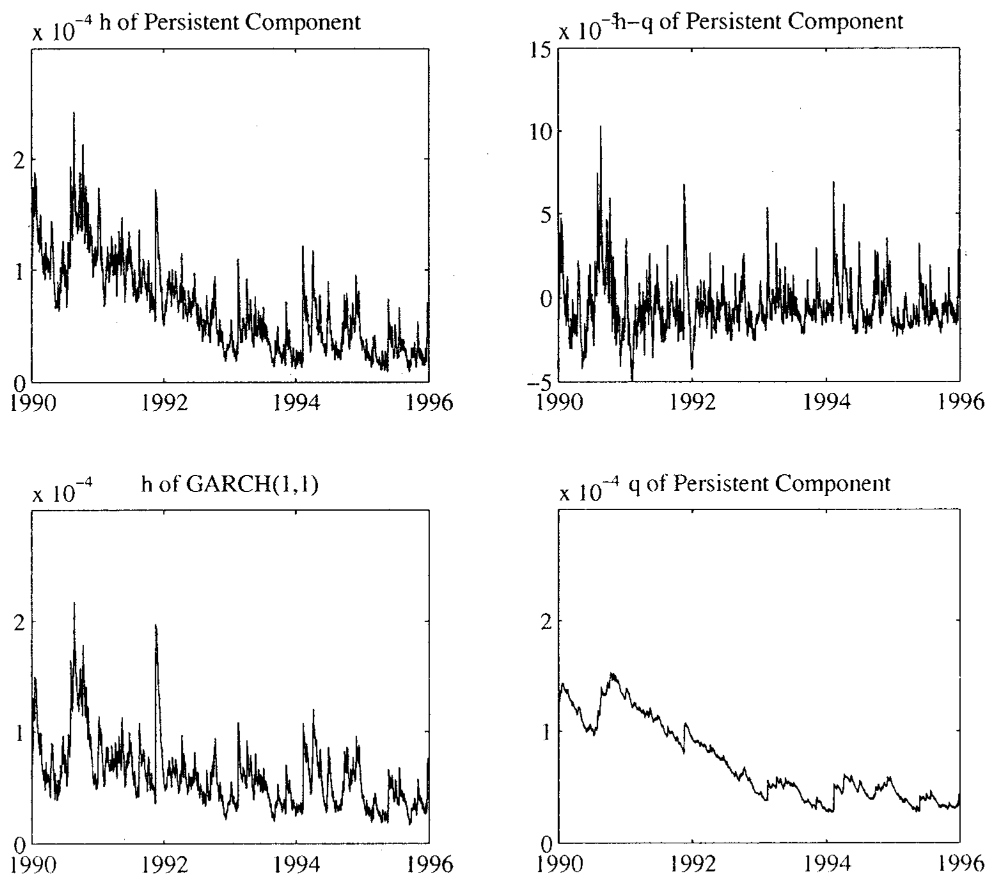
Notes to Figure: We use the numerical derivatives of the log conditional moment generating function to compute the term structure of skewness and kurtosis in the three GARCH models. The initial volatility is set to its long run value in the GARCH(1, 1) and component GARCH models. In the persistent component model the initial volatility is set to the unconditional volatility from the component model. The parameter values are obtained from the MLE estimates on returns in Table 1.

Figure 6. Spot Variance of Component GARCH versus GARCH(1,1)



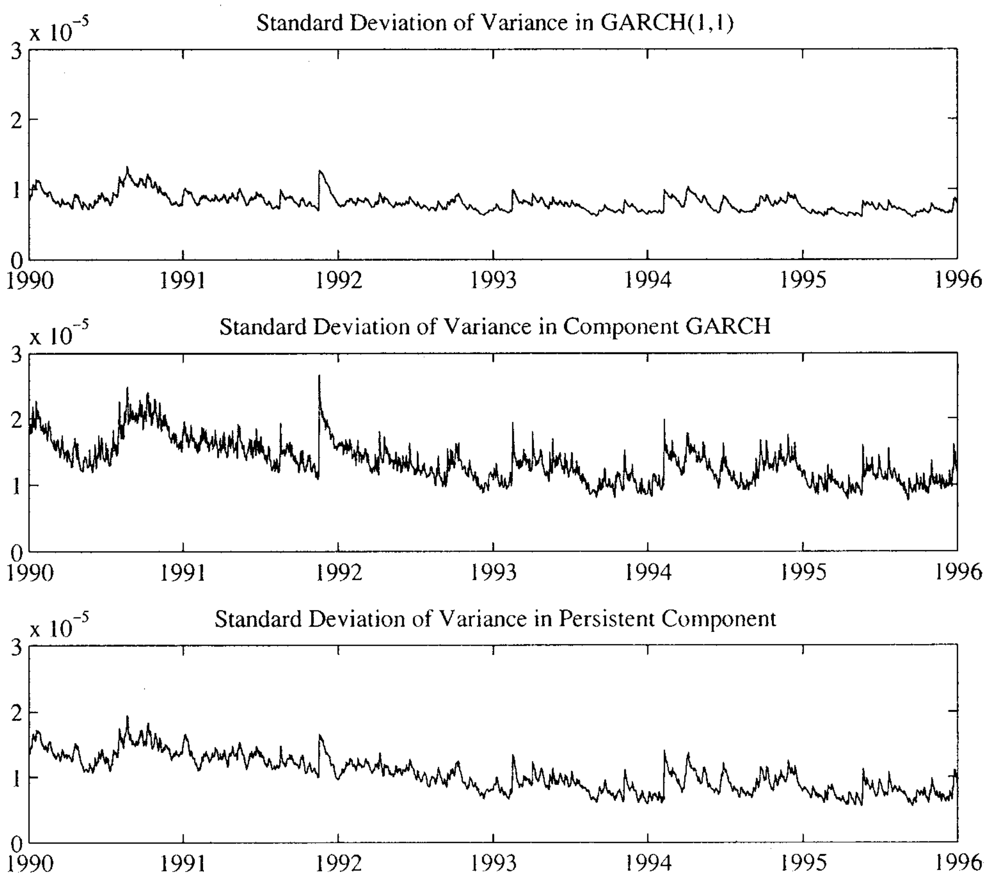
Notes to Figure: The left-hand panels plot the variance paths from the component and GARCH(1, 1) models. The right-hand panels plot the individual components. The parameter values are obtained from MLE estimation on returns in Table 1.

Figure 7. Spot Variance of Persistent Component Model versus GARCH(1,1)



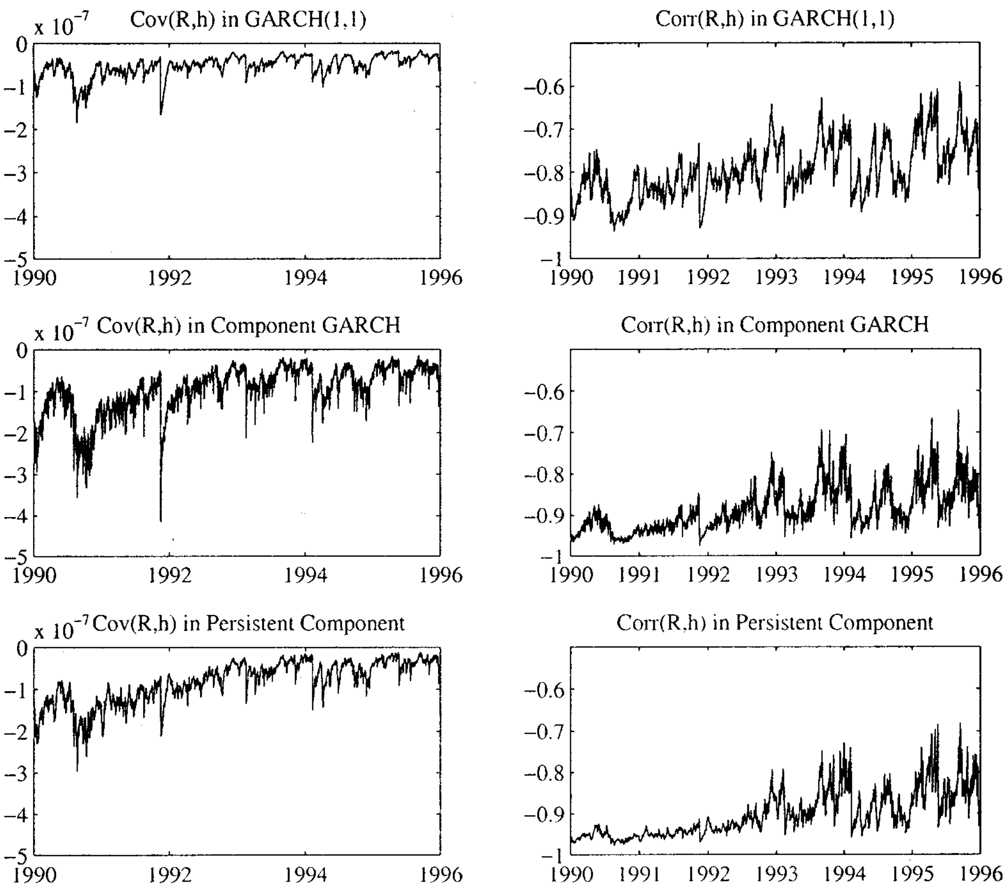
Notes to Figure: The left-hand panels plot the variance paths from the persistent component ($\rho = 1$) and GARCH$(1, 1)$ models. The right-hand panels plot the individual components. The parameter values are obtained from MLE estimation on returns in Table 1.

Figure 8: Conditional Variance of Variance Paths



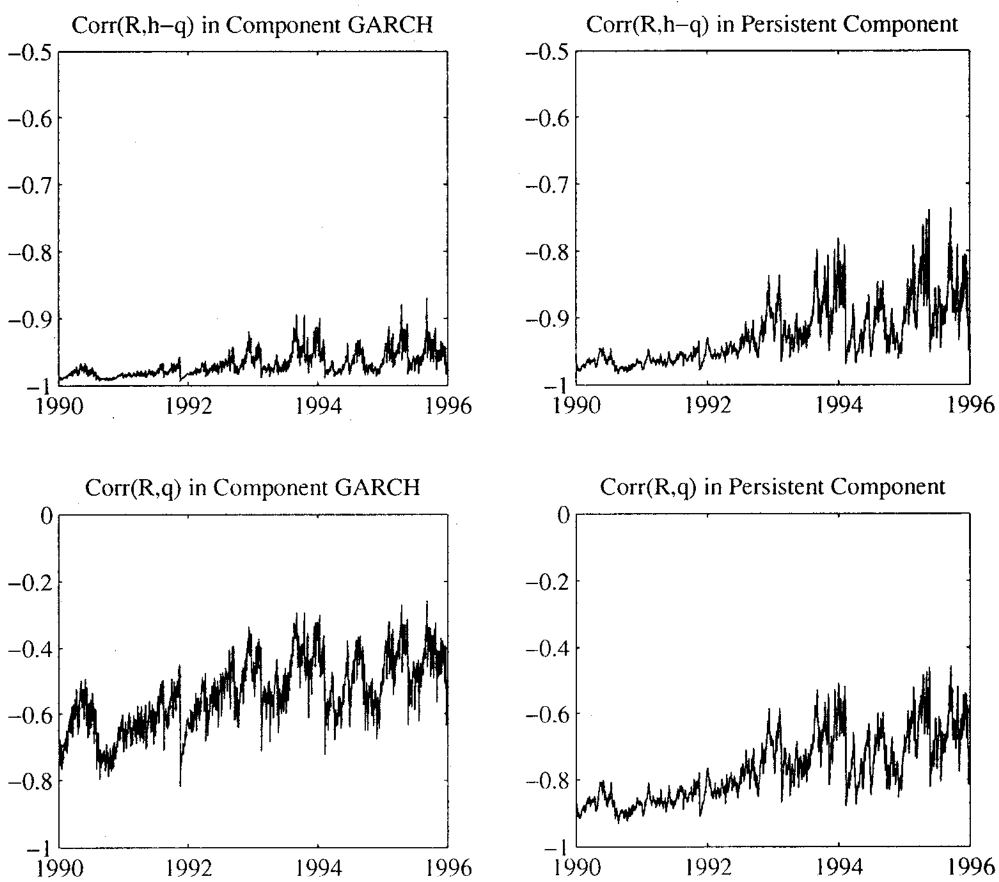
Notes to Figure: We plot the conditional variance of next day's variance as implied by the GARCH models. The top panel shows the GARCH(1. 1) model, the middle panel shows the component model and the bottom panel shows the persistent component model. The scales are identical across panels to facilitate comparison across models. The parameter values are obtained from the MLE estimates on returns in Table 1.

Figure 9. Conditional Covariance and Correlation



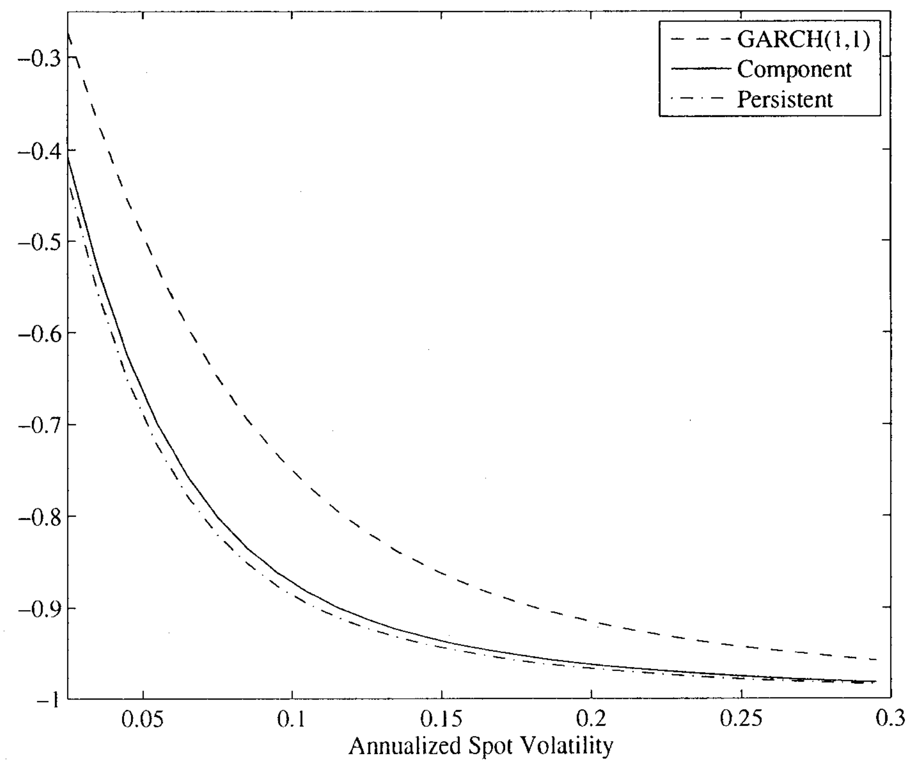
Notes to Figure: In the left panels we plot the conditional covariance between return and next-day variance as implied by the GARCH models and in the right panels we plot the corresponding conditional correlations. The scales are identical across top and bottom panels in order to facilitate comparison across models. The parameter values are obtained from the MLE estimates on returns in Table 1.

Figure 10. Conditional Correlations between Returns and Volatility Components



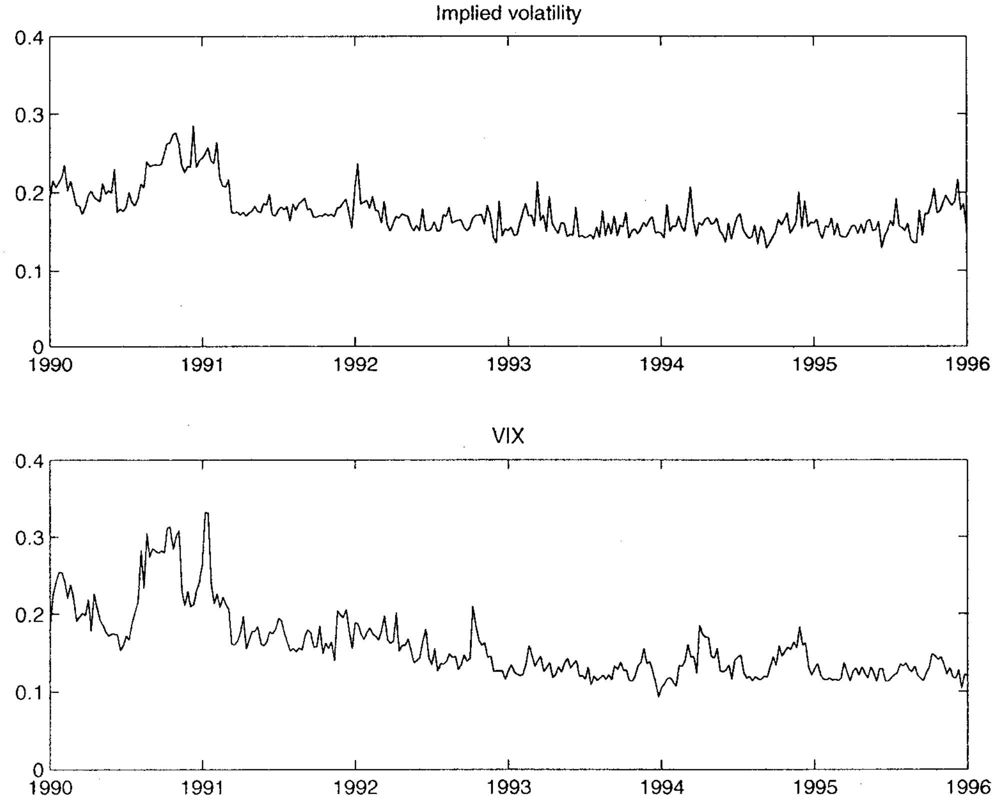
Notes to Figure: In the top row we plot the conditional correlation between return and the short-run volatility component. In the middle row we plot the conditional correlation between return and the long-run volatility component. In the bottom row we plot conditional correlation between the short-run and the long-run volatility components. The left column shows the component GARCH model and the right column shows the persistent component model. The parameter estimates are from Table 1.

Figure 11: Correlation Between Return and Variance as a Function of Volatility Level



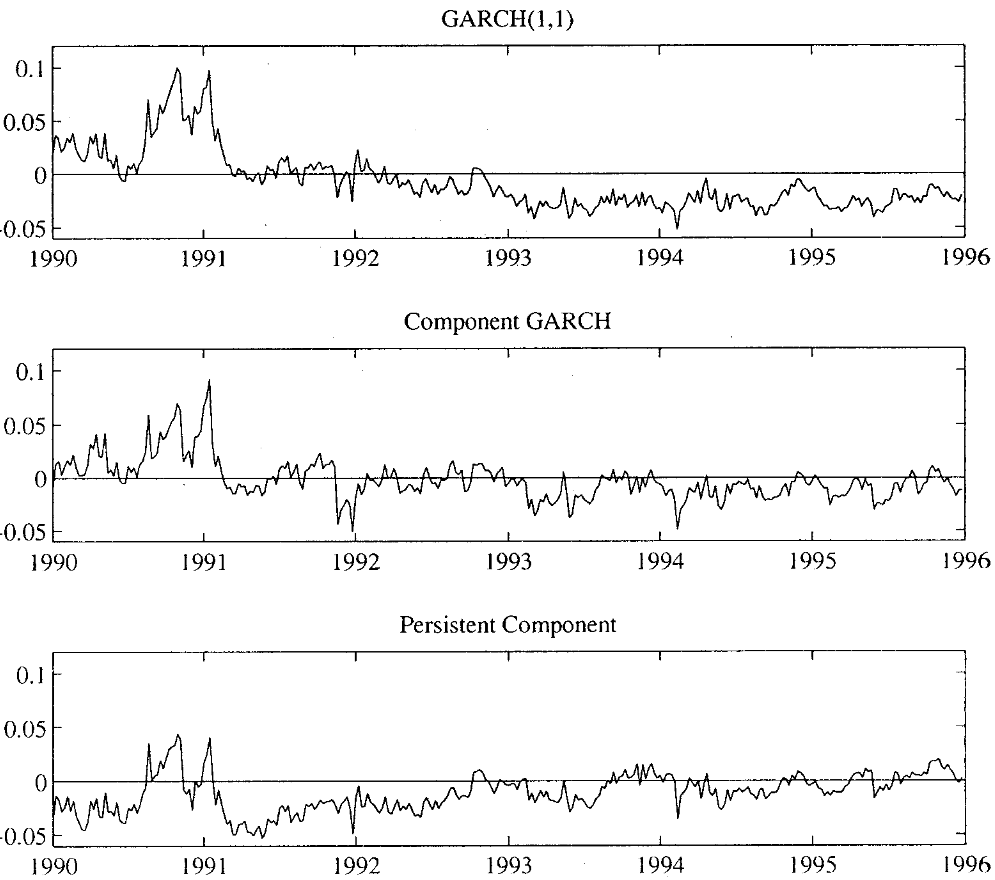
Notes to Figure: The figure shows the conditional correlation between the return on the underlying index and the daily variance. This conditional correlation is plotted against the level of volatility annualized. The dashed line corresponds to the GARCH$(1, 1)$, the solid line to the component model and the dash-dots to the persistent component model. The parameter estimates are from Table 1

Figure 12: Sample Average Weekly Implied Volatility and VIX



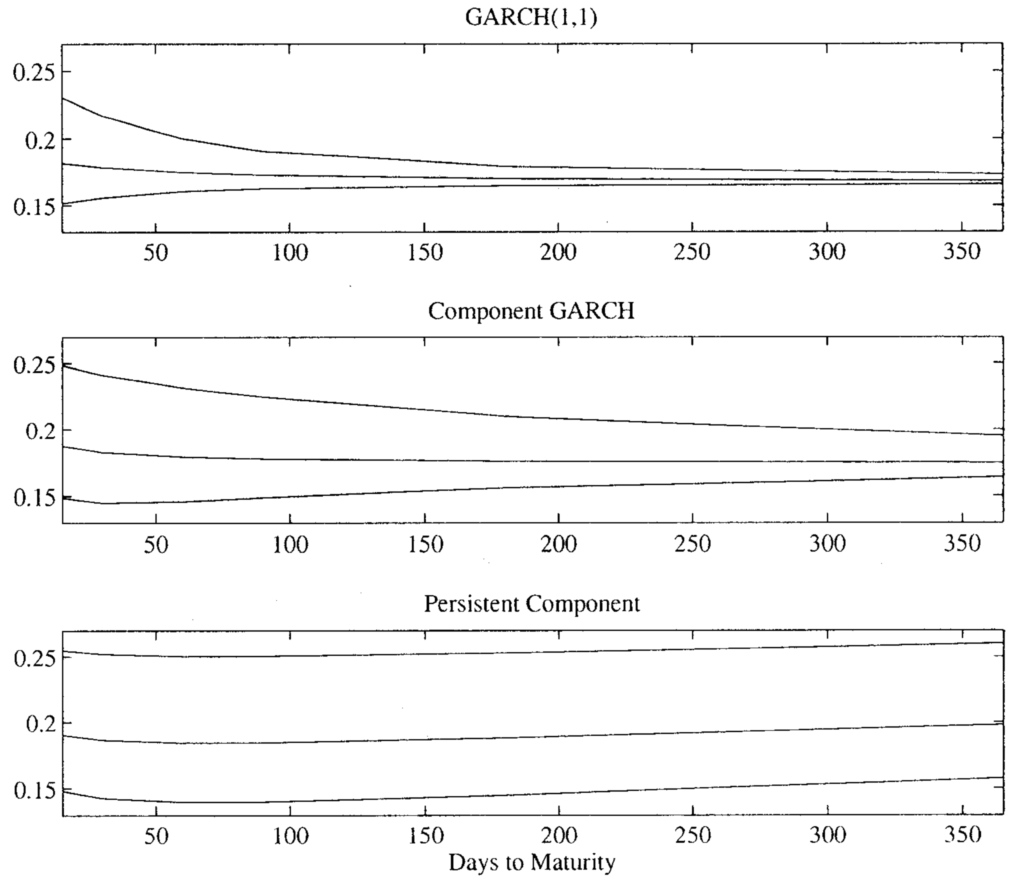
Notes to Figure: The top panel plots the average weekly implied Black-Scholes volatility for the S&P500 call options in our sample. The bottom panel plots the VIX index from the CBOE for comparison.

Figure 13: Weekly Implied Volatility Bias for At-the-Money Options



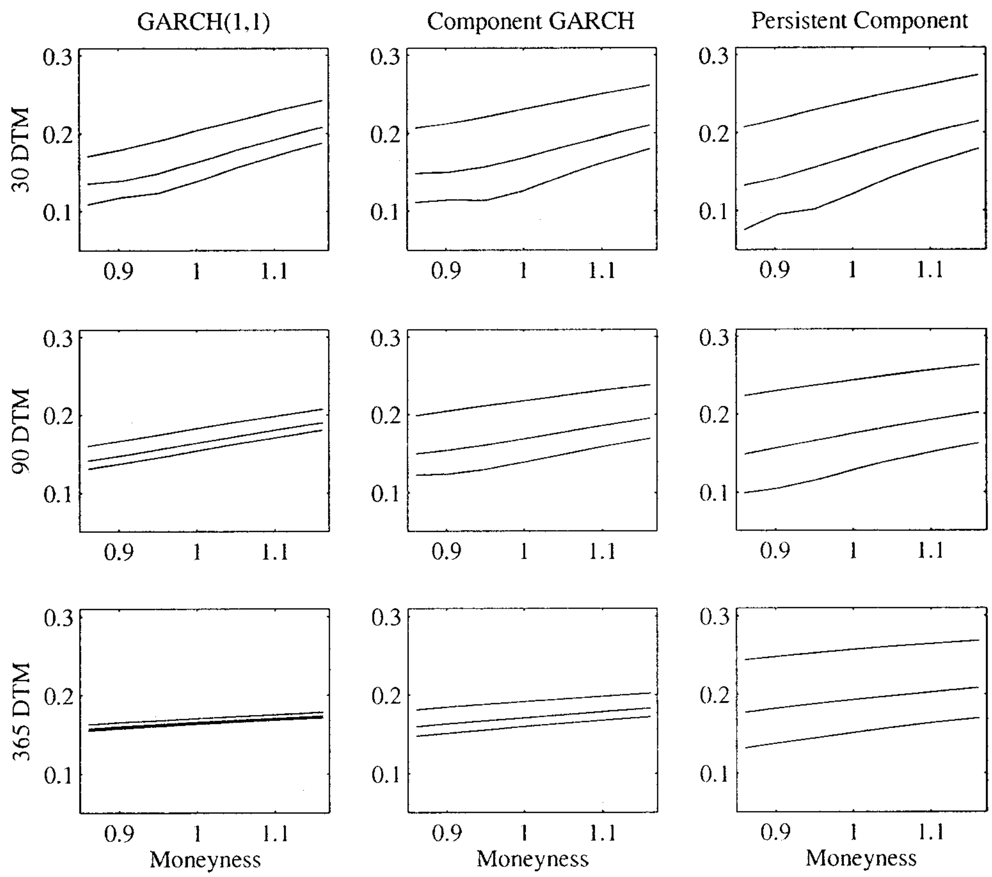
Notes to Figure: Each Wednesday we compute the Black-Scholes implied volatility for each at-the-money option contract. Options with moneyness (index value over strike price) between 0.975 and 1.025 are considered at-the-money. The implied volatility is computed both for the market price and for each model price. We plot the weekly average difference between the market and model implied volatility. The top panel shows the GARCH(1, 1) model, the middle panel shows the component model and the bottom panel shows the persistent component model. The MLE estimates from Table 1 are used.

Figure 14: Implied Volatility Term Structures for At-the-Money Options



Notes to Figure: We compute option prices and then implied annualized Black-Scholes volatilities from the three GARCH models for at-the-money options. The time to maturity is on the horizontal axis, and the three lines in each panel corresponds to an initial volatility half the unconditional (bottom line), equal the unconditional (middle line), and twice the unconditional (top line) volatility. The MLE estimates from Table 1 are used.

Figure 15: Implied Volatility Smirks for Various Maturities



Notes to Figure: We compute option prices and then implied annualized Black-Scholes volatilities from the three GARCH models for various moneyness, maturity and initial volatility. The moneyness is on the horizontal axis, each row of panels corresponds to a different maturity, and the three lines in each panel correspond to an initial volatility half the unconditional (bottom line), equal the unconditional (middle line), and twice the unconditional (top line) volatility. The MLE estimates from Table 1 are used.

Table 1: MLE Estimates and Properties

Estimation Sample: Daily Returns, 1962-2001

	GARCH(1,1)-Normal			**Component GARCH**			**Persistent Component**	
Parameter	Estimate	Std. Error	Parameter	Estimate	Std. Error	Parameter	Estimate	Std. Error
λ	2.231E+00	1.123E+00	λ	2.092E+00	7.729E-01	λ	2.017E-07	4.316E-01
w	2.101E-17	1.120E-07	$\widetilde{\beta}$	6.437E-01	2.759E-02	$\widetilde{\beta}$	8.822E-01	9.931E-03
b	9.012E-01	4.678E-03	α	1.580E-06	2.430E-07	α	2.057E-06	1.539E-07
a	3.317E-06	1.380E-07	γ_1	4.151E+02	6.341E+01	γ_1	2.516E+02	2.237E+01
c	1.276E+02	8.347E+00	γ_2	6.324E+01	5.300E+00	γ_2	1.187E+02	1.126E+01
			ω	8.208E-07	7.620E-08	ω	1.187E-07	1.393E-08
			φ	2.480E-06	1.160E-07	φ	7.966E-07	4.599E-08
			ρ	9.896E-01	9.630E-04	ρ	1.000E+00	
Ln Likelihood	33,955		Ln Likelihood	34,102		Ln Likelihood	34,005	
Persistence	0.9553		Persistence	0.9963		Persistence	1.0000	
Average Annual Vol	0.1206		Average Annual Vol	0.1174		Average Annual Vol	0.1239	
Average Vol of Var	7.997E-06		Average Vol of Var	1.341E-05		Average Vol of Var	1.044E-05	
Average Correlation	-0.7940		Average Correlation	-0.8849		Average Correlation	-0.9014	
Option RMSE	2.236		Option RMSE	1.706		Option RMSE	1.705	
Normalized	1.000		Normalized	0.763		Normalized	0.763	

Notes to Table: We use daily total returns from July 1, 1962 to December 31, 2001 on the S&P500 index to estimate the four models using Maximum Likelihood. Robust standard errors are calculated from the outer product of the gradient at the optimum parameter values. Persistence refers to the persistence of the conditional variance as defined in the text. Average Annual Vol refers to the average annualized standard deviation during 1990-95. Average Vol of Var refers to the average standard deviation of the conditional variance during 1990-95. Average Correlation refers to the average correlation between the return and the conditional variance during 1990-95. Ln Likelihood refers to the logarithm of the likelihood at the optimal parameter values. Option RMSE refers to the fit of the models on option prices observed during 1990-95.

Table 2: 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&P500 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 3: 1990-1995 RMSE and Ratio RMSE by Moneyness and Maturity Parameters Estimated from Daily Returns 1962-2001

Panel A. GARCH(1,1) RMSE

	DTM<20	20<DTM<80	80<DTM<180	DTM>180	All
S/X<0.975	0.454	1.778	3.032	4.155	3.090
0.975<S/X<1.00	0.671	2.116	3.087	3.548	2.603
1.00<S/X<1.025	0.638	1.650	2.574	2.955	2.154
1.025<S/X<1.05	0.595	1.204	2.099	2.487	1.700
1.05<S/X<1.075	0.735	1.013	1.879	2.227	1.516
1.075<S/X	0.759	1.024	1.424	1.917	1.360
All	0.683	1.503	2.448	3.228	2.236

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

	DTM<20	20<DTM<80	80<DTM<180	DTM>180	All
S/X<0.975	0.782	0.684	0.712	0.782	0.749
0.975<S/X<1.00	0.788	0.631	0.657	0.739	0.678
1.00<S/X<1.025	0.870	0.655	0.669	0.733	0.691
1.025<S/X<1.05	0.968	0.832	0.744	0.755	0.773
1.05<S/X<1.075	1.043	1.000	0.849	0.800	0.870
1.075<S/X	1.000	1.037	0.974	0.907	0.962
All	0.949	0.750	0.735	0.784	0.763

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

	DTM<20	20<DTM<80	80<DTM<180	DTM>180	All
S/X<0.975	0.985	0.726	0.774	0.842	0.808
0.975<S/X<1.00	0.978	0.640	0.635	0.722	0.669
1.00<S/X<1.025	0.982	0.672	0.600	0.681	0.653
1.025<S/X<1.05	0.947	0.773	0.613	0.660	0.675
1.05<S/X<1.075	1.010	0.909	0.663	0.718	0.750
1.075<S/X	1.002	0.981	0.836	0.784	0.856
All	0.990	0.746	0.719	0.796	0.763

Notes to Table: We use the MLE estimates from Table 1 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. Panels B and C show the ratio of the RMSEs of the Component and Persistent Component models over the RMSE of the GARCH(1,1) model.

Table 4: GARCH(1,1)-Jump Model

Estimation Sample: Daily Returns, 1962-2001

Panel A. GARCH(1,1)-Jump Parameter ML Estimates

Parameter	Estimate	Std. Error
λ	4.431E-01	1.425E+00
w	9.727E-11	2.179E-10
b	9.082E-01	7.158E-03
a	2.345E-08	5.090E-09
c	1.472E+03	2.175E+02
χ	5.790E+00	2.922E-01
μ	2.208E-03	3.449E-04
τ	1.330E+00	1.202E-01

Ln Likelihood	34,153
Option RMSE	2.138

Panel B. Ratio of the GARCH(1,1)-Jump to GARCH(1,1) RMSE

	DTM<20	20<DTM<80	80<DTM<180	DTM>180	All
S/X<0.975	1.001	0.941	0.946	0.940	0.942
0.975<S/X<1.00	0.978	0.931	0.940	0.937	0.936
1.00<S/X<1.025	0.993	0.928	0.950	0.945	0.943
1.025<S/X<1.05	1.055	0.965	0.974	0.962	0.968
1.05<S/X<1.075	0.897	1.015	1.013	1.011	1.010
1.075<S/X	1.003	1.023	1.036	1.043	1.035
All	0.987	0.953	0.960	0.954	0.956

Notes to Table: In Panel A We use daily total returns from July 1, 1962 to December 31, 2001 on the S&P500 index to estimate the GARCH(1,1)-Jump models using Maximum Likelihood. Robust standard errors are calculated from the outer product of the gradient at the optimum parameter values. In Pabel B we compute the ratio of the option root mean squared error (RMSE) from the GARCH(1,1)-Jump model to the RMSE of the GARCH(1,1) in Table 3 Panel A.

Table 5: NLS Estimates and Properties

Sample: 1990-1992 (in-sample) 1993 (out-of-sample).

GARCH(1,1)			**Component GARCH**			**Persistent Component**		
Parameter	Estimate	Std. Error	Parameter	Estimate	Std. Error	Parameter	Estimate	Std. Error
w	3.891E-14	3.560E-12	$\tilde{\beta}$	7.050E-01	2.565E-01	$\tilde{\beta}$	7.201E-01	1.021E-01
b	6.801E-01	3.211E-03	α	1.770E-06	3.444E-07	α	1.597E-06	2.279E-06
a	2.666E-07	6.110E-09	γ_1	5.617E+02	1.494E+02	γ_1	7.481E+02	8.974E+01
c	1.090E+03	5.432E+01	γ_2	5.638E+02	1.555E+02	γ_2	4.767E+02	1.246E+02
			ω	2.424E-07	1.212E-07	ω	5.343E-08	1.345E-08
			φ	5.249E-07	3.525E-07	φ	5.123E-07	1.285E-07
			ρ	9.981E-01	3.519E-03	ρ	1.000E+00	

Persistence	0.9970	Persistence	0.9994	Persistence	1.0000
Average Annual Vol	0.1347	Average Annual Vol	0.1405	Average Annual Vol	0.1431
Average Vol of Var	4.283E-06	Average Vol of Var	1.962E-05	Average Vol of Var	2.197E-05
Average Correlation	-0.9967	Average Correlation	-0.9876	Average Correlation	-0.9914

	NLS	MLE		NLS	MLE		NLS	MLE
RMSE (90-92)	1.038	1.896	RMSE (90-92)	0.931	1.609	RMSE (90-92)	0.984	2.143
Normalized	1.000	1.000	Normalized	0.897	0.849	Normalized	0.948	1.130
RMSE (93)	1.284	2.229	RMSE (93)	0.983	1.584	RMSE (93)	1.198	1.260
Normalized	1.000	1.000	Normalized	0.765	0.710	Normalized	0.933	0.565

Notes to Table: We use Wednesday option prices from from January 1, 1990 to December 31, 1992 on the S&P500 index to estimate the three GARCH models using Nonlinear Least Squares on the valuation errors. Robust standard errors are calculated from the outer product of the gradient at the optimum parameter values. RMSE refers to the square root of the mean-squared valuation errors. RMSE(in) refers to 1990-1992 and RMSE(out) to 1993. NLS refers to the model estimated using option data and MLE refers to the model estimated using returns only. Normalized values are divided by the respective RMSE from GARCH(1,1).

Table 6: NLS Estimates and Properties

Sample: 1992-1994 (in-sample) 1995 (out-of-sample)

GARCH(1,1)			**Component GARCH**			**Persistent Component**		
Parameter	Estimate	Std. Error	Parameter	Estimate	Std. Error	Parameter	Estimate	Std. Error
w	7.521E-16	3.498E-09	$\widetilde{\beta}$	9.297E-01	3.346E-02	$\widetilde{\beta}$	9.587E-01	3.821E-05
b	4.694E-01	1.251E-01	α	1.808E-06	1.320E-07	α	1.943E-06	1.614E-06
a	1.936E-06	3.986E-07	γ_1	5.854E+02	2.362E+02	γ_1	2.589E+02	8.383E+01
c	5.061E+02	1.041E+02	γ_2	5.749E+02	4.025E+02	γ_2	2.254E+02	5.063E+02
			ω	2.204E-07	3.470E-08	ω	6.927E-08	1.262E-08
			φ	2.835E-07	1.586E-07	φ	6.971E-07	1.253E-08
			ρ	9.966E-01	1.277E-03	ρ	1.000E+00	
Persistence	0.9654		Persistence	0.9998		Persistence	1.0000	
Average Annual Vol	0.1074		Average Annual Vol	0.1129		Average Annual Vol	0.1082	
Average Vol of Var	1.423E-05		Average Vol of Var	1.838E-05		Average Vol of Var	1.085E-05	
Average Correlation	-0.9701		Average Correlation	-0.9781		Average Correlation	-0.9095	
	NLS	MLE		NLS	MLE		NLS	MLE
RMSE (92-94)	1.107	2.000	RMSE (92-94)	0.855	1.524	RMSE (92-94)	1.058	1.402
Normalized	1.000	1.000	Normalized	0.773	0.762	Normalized	0.956	0.701
RMSE (95)	1.227	2.775	RMSE (95)	0.972	1.920	RMSE (95)	1.174	1.249
Normalized	1.000	1.000	Normalized	0.792	0.692	Normalized	0.957	0.450

Notes to Table: See notes to Table 5. RMSE(in) now refers to 1992-1994 and RMSE(out) to 1995.

Table 7: 1990-1992 (in-sample) RMSE and Ratio RMSE by Moneyness and Maturity

Panel A. GARCH(1,1) RMSE

	DTM<20	20<DTM<80	80<DTM<180	DTM>180	All
S/X<0.975	0.437	0.889	1.098	1.276	1.078
0.975<S/X<1.00	0.664	1.054	1.123	1.139	1.054
1.00<S/X<1.025	0.575	0.956	1.049	0.993	0.956
1.025<S/X<1.05	0.556	0.907	1.030	0.949	0.919
1.05<S/X<1.075	0.687	0.989	1.166	1.112	1.032
1.075<S/X	0.642	1.075	1.229	1.022	1.079
All	0.610	0.976	1.124	1.151	1.038

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

	DTM<20	20<DTM<80	80<DTM<180	DTM>180	All
S/X<0.975	0.939	0.816	0.847	0.925	0.873
0.975<S/X<1.00	0.895	0.872	0.923	1.049	0.927
1.00<S/X<1.025	0.923	0.916	0.955	1.000	0.947
1.025<S/X<1.05	0.872	0.881	0.956	1.063	0.936
1.05<S/X<1.075	0.902	0.849	0.906	1.003	0.902
1.075<S/X	0.971	0.877	0.848	0.949	0.883
All	0.923	0.865	0.883	0.960	0.897

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

	DTM<20	20<DTM<80	80<DTM<180	DTM>180	All
S/X<0.975	0.845	0.802	0.898	1.145	0.987
0.975<S/X<1.00	0.833	0.887	0.959	1.194	0.977
1.00<S/X<1.025	0.952	0.927	0.965	1.211	1.003
1.025<S/X<1.05	0.863	0.876	0.932	1.163	0.941
1.05<S/X<1.075	0.878	0.837	0.859	1.061	0.891
1.075<S/X	0.964	0.855	0.810	0.951	0.861
All	0.903	0.859	0.891	1.121	0.948

Notes to Table: We use the NLS estimates from Table 5 to compute the root mean squared option valuation error (RMSE) for various moneyness and maturity bins during 1990-1992. Panel A shows the RMSEs for the GARCH(1,1) model. Panel B shows the ratio of the component GARCH RMSEs to the GARCH(1,1) RMSEs from Panel A. Panel C shows the ratio of the persistence component GARCH RMSEs to the GARCH(1,1) RMSEs.

Table 8: 1993 (out-of-sample) RMSE and Ratio RMSE by Moneyness and Maturity

Panel A. GARCH(1,1) RMSE

	DTM<20	20<DTM<80	80<DTM<180	DTM>180	All
S/X<0.975	0.289	1.157	1.328	1.944	1.461
0.975<S/X<1.00	0.579	1.511	1.800	2.434	1.631
1.00<S/X<1.025	0.498	1.147	1.460	2.290	1.356
1.025<S/X<1.05	0.593	0.724	1.144	2.014	1.008
1.05<S/X<1.075	0.650	0.654	0.834	1.580	0.860
1.075<S/X	1.147	1.160	0.991	1.402	1.166
All	0.813	1.124	1.258	1.822	1.284

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

	DTM<20	20<DTM<80	80<DTM<180	DTM>180	All
S/X<0.975	0.717	0.609	0.583	0.509	0.556
0.975<S/X<1.00	0.799	0.673	0.658	0.533	0.644
1.00<S/X<1.025	1.087	0.844	0.754	0.589	0.749
1.025<S/X<1.05	0.966	1.076	0.871	0.689	0.883
1.05<S/X<1.075	1.121	1.037	1.039	0.621	0.902
1.075<S/X	0.991	0.971	1.044	0.928	0.977
All	0.994	0.828	0.778	0.652	0.765

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

	DTM<20	20<DTM<80	80<DTM<180	DTM>180	All
S/X<0.975	0.626	0.547	0.822	1.293	1.028
0.975<S/X<1.00	0.786	0.579	0.678	0.832	0.671
1.00<S/X<1.025	1.060	0.709	0.815	0.754	0.763
1.025<S/X<1.05	0.918	0.995	0.933	0.681	0.876
1.05<S/X<1.075	1.163	1.012	1.252	0.977	1.071
1.075<S/X	0.994	0.994	1.228	1.070	1.068
All	0.994	0.782	0.914	1.079	0.933

Notes to Table: See Table 7. We use the NLS estimates from Table 5 to compute the out-of-sample root mean squared option valuation error (RMSE) for various moneyness and maturity bins during 1993.

Table 9: 1992-1994 (in-sample) RMSE and Ratio RMSE by Moneyness and Maturity

Panel A. GARCH(1,1) RMSE

	DTM<20	20<DTM<80	80<DTM<180	DTM>180	All
S/X<0.975	0.482	0.929	1.095	1.364	1.122
0.975<S/X<1.00	0.988	1.283	1.293	1.398	1.275
1.00<S/X<1.025	0.904	1.212	1.184	1.499	1.228
1.025<S/X<1.05	0.589	0.953	0.999	1.348	1.002
1.05<S/X<1.075	0.786	0.804	0.886	1.549	0.991
1.075<S/X	0.922	0.866	0.857	1.447	1.032
All	0.857	1.009	1.045	1.422	1.107

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

	DTM<20	20<DTM<80	80<DTM<180	DTM>180	All
S/X<0.975	0.813	0.641	0.609	0.577	0.604
0.975<S/X<1.00	0.773	0.770	0.728	0.692	0.747
1.00<S/X<1.025	0.866	0.824	0.727	0.628	0.760
1.025<S/X<1.05	0.960	0.818	0.715	0.787	0.790
1.05<S/X<1.075	0.919	0.840	0.716	0.751	0.785
1.075<S/X	0.996	0.952	0.859	0.911	0.919
All	0.909	0.808	0.717	0.744	0.773

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

	DTM<20	20<DTM<80	80<DTM<180	DTM>180	All
S/X<0.975	0.835	0.920	1.036	1.041	1.010
0.975<S/X<1.00	0.757	0.924	1.054	1.005	0.964
1.00<S/X<1.025	0.791	0.867	0.973	0.981	0.918
1.025<S/X<1.05	0.943	0.855	0.949	0.920	0.900
1.05<S/X<1.075	0.994	0.898	0.903	0.774	0.856
1.075<S/X	0.993	1.046	0.964	0.917	0.969
All	0.899	0.927	1.000	0.959	0.956

Notes to Table: See Table 7. We use the NLS estimates from Table 6 to compute the root mean squared option valuation error (RMSE) for various moneyness and maturity bins during 1992-1994.

Table 10: 1995 (out-of-sample) RMSE and Ratio RMSE by Moneyness and Maturity

Panel A. GARCH(1,1) RMSE

	DTM<20	20<DTM<80	80<DTM<180	DTM>180	All
S/X<0.975	0.387	0.863	1.456	2.456	1.771
0.975<S/X<1.00	0.995	1.175	1.719	2.093	1.546
1.00<S/X<1.025	0.752	1.065	1.514	1.872	1.389
1.025<S/X<1.05	0.538	0.909	1.265	1.450	1.110
1.05<S/X<1.075	0.903	0.617	0.867	1.401	0.896
1.075<S/X	0.644	0.617	0.571	0.964	0.681
All	0.744	0.846	1.187	1.848	1.227

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

	DTM<20	20<DTM<80	80<DTM<180	DTM>180	All
S/X<0.975	1.304	1.185	0.973	0.615	0.757
0.975<S/X<1.00	0.994	0.903	0.755	0.693	0.783
1.00<S/X<1.025	0.839	0.750	0.737	0.651	0.708
1.025<S/X<1.05	1.036	0.748	0.657	0.682	0.705
1.05<S/X<1.075	0.979	0.831	0.767	0.741	0.796
1.075<S/X	1.008	1.077	1.091	0.933	1.026
All	0.978	0.929	0.847	0.670	0.792

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

	DTM<20	20<DTM<80	80<DTM<180	DTM>180	All
S/X<0.975	0.674	1.093	0.947	0.683	0.781
0.975<S/X<1.00	0.895	1.072	0.903	0.992	0.985
1.00<S/X<1.025	0.938	0.967	0.961	0.933	0.950
1.025<S/X<1.05	1.105	1.080	0.928	1.040	1.021
1.05<S/X<1.075	1.042	1.325	1.222	0.882	1.092
1.075<S/X	0.903	1.342	1.548	1.140	1.286
All	0.952	1.129	1.016	0.848	0.957

Notes to Table: See Table 7. We use the NLS estimates from Table 6 to compute the out-of-sample root mean squared option valuation error (RMSE) for various moneyness and maturity bins during 1995.

Chapter 2
Volatility Components: Affine Restrictions and Non-normal Innovations

Peter Christoffersen Kris Jacobs Yintian Wang

Abstract

We derive two new GARCH variance component models with non-normal innovations. One of these models has an affine structure and leads to a closed form option valuation formula. The other model has a non-affine dynamic and option valuation must be done via Monte Carlo simulation. We provide an empirical comparison of these two new component models and the respective special cases with normal innovations. We also compare the four component models with the GARCH(1,1) models which they nest. All eight models are estimated using MLE on S&P500 returns. The likelihood criterion strongly favors the component models, and favors the non-normal innovations. The properties of the non-affine models differ significantly from those of the affine models. When using the estimated parameters for option valuation, we again find strong support for the component variance specifications, but the support for the non-normal innovations and the non-affine structures is less convincing.

JEL Classification: G22 G13
Keywords: Volatility, GARCH, Component Model, Affine, Long memory, Option valuation, Normal

2.1 Introduction

Following the path-breaking work of Engle (1982) and Bollerslev (1986), GARCH models have become an ubiquitous toolkit in empirical finance. In this paper we derive and empirically implement two new conditional non-normal GARCH variance component models. The first model builds on Engle and Lee (1999), who use a non-affine variance component dynamic. We modify the model of Engle and Lee (1999) by modeling the return innovation using a Generalized Error Distribution (GED). This innovation is more general than the more traditional normal innovation, and therefore this new model has the ability to better fit the return distribution. Option valuation in this model must be done via Monte Carlo simulation. The second new component model follows an affine variance dynamic, and assumes a conditional inverse Gaussian shock distribution as in Christoffersen, Heston and Jacobs (2004). We derive a closed form option valuation formula for this model. We compare the empirical performance of these two new non-normal component models with the corresponding special cases characterized by normal innovations. These models are discussed in Engle and Lee (1999) and Christoffersen, Jacobs and Wang (2005) respectively. These four component models are also compared with the four GARCH(1,1) models that they nest. For each of the component models, we provide two-way parameter mappings between the component models and their respective GARCH(2,2) counterparts.

We estimate these eight models using MLE on S&P500 returns. This empirical comparison allows us to compare the importance of three types of modeling assumptions: first, the importance of the component structure versus the simpler and more parsimonious GARCH(1,1) structure; second, the importance of non-normal return innovations; third, the importance of the affine structure. The likelihood criterion strongly favors the component models in all cases, as well as the non-normal return innovations. Using the MLE estimates, we characterize key properties of each model, including autocorrelation functions for the squared innovations and conditional leverage and variance of variance paths. We find important differences between affine and non-affine models, as well as between GARCH(1,1) and component models and between models with normal and non-normal innovations. These results suggest that non-normal innovations and the non-affine structure provide more flexibility in a parsimonious fashion.

When we use the estimated model parameters for option valuation, we again find strong support for the component variance specifications, but less support for the non-normal return innovations and non-affine specifications. These findings are of interest because they provide a perspective that differs from the available GARCH literature. Many papers in the literature compare volatility models via mean-squared-error type comparisons computed from volatility point forecasts and some measure of realized volatility. For a recent example see Hansen and Lunde (2005) and the references therein. Such papers often find that it is difficult to outperform the simple GARCH(1,1) specification from Bollerslev (1986). While those studies are clearly important and useful, we proceed instead by comparing the suggested volatility models based on their ability to fit observed option prices. Such a comparison is arguably richer in that it makes use of each model's entire (risk neutral) conditional density forecast at many horizons corresponding to the maturity of each option. We believe that the richer model evaluation criterion at least in part explains our novel empirical findings.

The literature on GARCH variance component models is rapidly expanding. Component GARCH models can be viewed as a convenient way of incorporating long-memory-like features into a short-memory model, at least for the horizons relevant for option valuation. Bollerslev and Mikkelsen (1999) find support for a long-memory GARCH option valuation model applied to long maturity LEAPs options. We consider options with up to one-year maturity where the component models are likely to provide good approximations to true long-memory processes. Maheu (2002) presents Monte-Carlo evidence that a component model similar to the ones in this paper can capture long-range volatility dynamics. Adrian and Rosenberg (2005) demonstrate the relevance of the component volatility structure for cross-sectional asset pricing. GARCH component variance models are also related to the stochastic volatility component models which have received empirical support (see for example Alizadeh, Brandt and Diebold (2002), Chernov, Gallant, Ghysels and Tauchen (2003), and Taylor and Xu (1994)).

The remainder of the paper is structured as follows. In Sections 2 and 3 we introduce two new GARCH component models. Section 2 introduces a non-affine conditional non-normal GARCH component model, derives a number of its properties, and discusses option valuation for this component dynamic. Section 3 introduces a new affine conditional inverse Gaussian component model and derives the corresponding option valuation formula. The special cases of conditional normality for these two new models are discussed at the end of each section. Section 4 presents empirical model comparisons based on maximum likelihood estimation of returns and root mean squared errors from valuing options on the S&P500 index. Section 5 concludes.

2.2 A Non-affine, Non-normal GARCH Component Model

In this section, we build on the work of Engle and Lee (1999) and Duan (1999) to construct a conditionally non-normal, non-affine component GARCH model that can be used for option valuation. The non-affine models considered in this section are somewhat more cumbersome to use in option valuation than the affine models considered in Section 3 below, because they do not allow for a closed-form solution for option prices. However, non-affine GARCH models may provide a better fit to the option data.

2.2.1 Return Dynamics

We first introduce the benchmark model NGARCH(1,1) option valuation model of Engle and Ng (1993) used for option valuation by Duan (1995).

$$\begin{aligned} R_{t+1} &\equiv \ln \frac{S_{t+1}}{S_t} = r + \lambda h_{t+1} + \sqrt{h_{t+1}} z_{t+1} \\ h_{t+1} &= w + \bar{b}_1 h_t + a_1 h_t \left(z_t - c_1\right)^2 \end{aligned}$$

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. Note that we use the risk premium specification of Heston and Nandi (2000) rather than that of Duan (1995) in order to facilitate comparison with the affine models in Section 3. Using the unconditional variance equation

$$\sigma^2 \equiv E\left(h_{t+1}\right) = \frac{w}{1 - \bar{b}_1 - a_1\left(1 + c_1^2\right)}$$

we can rewrite the conditional variance as

$$h_{t+1} = \sigma^2 + b_1\left(h_t - \sigma^2\right) + a_1 h_t \left(\left(z_t^2 - 1\right) - 2c_1 z_t\right)$$

where $b_1 = \bar{b}_1 + a_1 \left(1 + c_1^2\right)$.

The component NGARCH 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 itself is autoregressive of the first order. Using Greek letters for component model parameters, we write

$$\begin{aligned} h_{t+1} &= q_{t+1} + \beta(h_t - q_t) + \alpha h_t v_{1,t} \\ q_{t+1} &= \sigma^2 + \rho\left(q_t - \sigma^2\right) + \varphi h_t v_{2,t} \end{aligned}$$

where $v_{i,t} = (z_t^2 - 1) - 2\gamma_i z_t$ for $i = 1, 2$ can be viewed as zero-mean innovations to the volatility components. Henceforth, we denote the component NGARCH model as NGARCH(C).

Following Duan (1999), we will assume that the $i.i.d.$ return innovation z_t follows the Generalized Error Distribution (GED) which, after normalizing to get a zero mean and unit variance, is given by[24]

$$g(z; v) = \frac{v}{2^{1+\frac{1}{v}}\theta\Gamma\left(\frac{1}{v}\right)} \exp\left(-\frac{1}{2}\left|\frac{z}{\theta}\right|^{v}\right) \qquad \text{for } 0 < v \leq \infty$$

where $\Gamma(.)$ is the gamma function and $\theta = \left(\frac{2^{-\frac{2}{v}}\Gamma\left(\frac{1}{v}\right)}{\Gamma\left(\frac{3}{v}\right)}\right)^{\frac{1}{2}}$. The parameter v determines the thickness of the density tails. For $v < 2$, the density function has a tail fatter than the normal distribution and vice versa. The innovation z_t has a skewness of zero and a kurtosis of $\frac{\Gamma\left(\frac{5}{v}\right)\Gamma\left(\frac{1}{v}\right)}{\Gamma\left(\frac{3}{v}\right)^2}$.

---

24 See Hamilton (1994) and Nelson (1991) on the properties of the GED.

We now derive a number of properties of the NGARCH(C)-GED and NGARCH(1,1)-GED models that are key for understanding their performance in capturing the salient features of speculative returns and in fitting option prices.

2.2.2 Conditional Leverage and Variance of Variance

In order 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 NGARCH(1,1)-GED model the conditional leverage effect is given by

$$Cov_t\left(R_{t+1}, h_{t+2}\right) = -2a_1c_1h_{t+1}^{3/2}$$

For the NGARCH(C)-GED model we get

$$Cov_t\left(R_{t+1}, h_{t+2}\right) = -2\left(\alpha\gamma_1 + \varphi\gamma_2\right)h_{t+1}^{3/2}$$

Notice that in neither case does the conditional leverage effect depend on the GED distribution's tail parameter v.

We define the conditional variance of variance as $Var_t\left(h_{t+2}\right)$, which in the NGARCH(1,1)-GED is given by

$$Var_t\left(h_{t+2}\right) = \left(\kappa(v) - 1 + 4c_1^2\right)a_1^2h_{t+1}^2$$

where $\kappa(v) = \frac{\Gamma\left(\frac{5}{v}\right)\Gamma\left(\frac{1}{v}\right)}{\Gamma\left(\frac{3}{v}\right)^2}$. In the NGARCH(C)-GED model we get

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

2.2.3 The Autocorrelation Function for the Squared Innovations

The conditional autocorrelation function (ACF) of the squared GARCH innovation defined as $\varepsilon_{t+1}^2 = z_{t+1}^2 h_{t+1}$ parsimoniously captures the models' volatility memory properties. The ACF is defined as

$$Corr_t(\varepsilon_{t+1}^2, \varepsilon_{t+k}^2) \equiv \frac{Cov_t(\varepsilon_{t+1}^2, \varepsilon_{t+k}^2)}{\sqrt{Var_t\left(\varepsilon_{t+1}^2\right)}\sqrt{Var_t\left(\varepsilon_{t+k}^2\right)}} = \frac{Cov_t(\varepsilon_{t+1}^2, h_{t+k})}{\sqrt{Var_t\left(\varepsilon_{t+1}^2\right)}\sqrt{Var_t\left(\varepsilon_{t+k}^2\right)}} \tag{2.34}$$

For the NGARCH(1,1)-GED we have

$$Cov_t(\varepsilon_{t+1}^2, h_{t+k}) = (\kappa(v) - 1)\, b_1^{k-2} a_1 h_{t+1}^2, \text{ for } k \geq 2$$

$$Var_t\left(\varepsilon_{t+1}^2\right) = (\kappa(v) - 1)\, h_{t+1}^2$$

$$\begin{aligned} Var_t\left(\varepsilon_{t+k}^2\right) &= 3\sum_{i=2}^{k} \left(b^{k-i}\right)^2 a_1^2 \left(\kappa(v) - 1 + 4c_1^2\right) E_t\left(h_{t+i-1}^2\right) + \ldots \\ & \quad (\kappa(v) - 1)\left(\sigma^2\left(1 - b_1^{k-1}\right) + b_1^{k-1} h_{t+1}\right)^2 \end{aligned}$$

with the expected future variance given by

$$E_t\left(h_{t+k-1}\right) = \sigma^2 + b_1^{k-2}(h_{t+1} - \sigma^2)$$

The ACF for the NGARCH(C)-GED model can be obtained using

$$Cov_t(\varepsilon_{t+1}^2, h_{t+k}) = (\kappa(v) - 1)\,(\beta^{k-2}\alpha + \rho^{k-2}\varphi) h_{t+1}^2$$

$$Var_t\left(\varepsilon_{t+1}^2\right) = (\kappa(v) - 1)\, h_{t+1}^2$$

$$Var_t\left(\varepsilon_{t+k}^2\right) = 3\sum_{i=2}^{k}\left(\left(\kappa(v)-1\right)\left(\beta^{k-i}\alpha+\rho^{k-i}\varphi\right)^2+4\left(\beta^{k-i}\alpha\gamma_1+\rho^{k-i}\varphi\gamma_2\right)^2\right)E_t h_{t+i-1}^2$$
$$+\left(\kappa(v)-1\right)\left(\sigma^2\left(1-\rho^{k-1}\right)+\rho^{k-1}q_{t+1}+\beta^{k-1}(h_{t+1}-q_{t+1})\right)^2$$

where the expected future variance is now given by

$$E_t\left(h_{t+k-1}\right) = \sigma^2+\beta^{k-2}(h_{t+1}-q_{t+1})+\rho^{k-2}\left(q_{t+1}-\sigma^2\right)$$

2.2.4 GARCH(2,2) Mappings

Engle and Lee (1999) demonstrate that component GARCH models can be rewritten as GARCH(2,2) models. In order to better understand the component GARCH model, we investigate this relationship in more detail. The component model can be mapped into a GARCH(2,2) as follows

$$h_{t+1} = w+\bar{b}_1 h_t+\bar{b}_2 h_{t-1}+a_1 h_t\left(z_t-c_1\right)^2+a_2 h_{t-1}\left(z_{t-1}-c_2\right)^2$$

where

$$\bar{b}_1 = \rho+\beta-\alpha-\varphi-\frac{\left(\alpha\gamma_1+\varphi\gamma_2\right)^2}{\alpha+\varphi} \qquad a_1 = \alpha+\varphi \qquad c_1 = \frac{\alpha\gamma_1+\varphi\gamma_2}{\alpha+\varphi}$$
$$\bar{b}_2 = \frac{\left(\rho\alpha\gamma_1+\beta\varphi\gamma_2\right)^2}{\rho\alpha+\beta\varphi}-\rho\beta+\rho\alpha+\beta\varphi \quad a_2 = -\left(\rho\alpha+\beta\varphi\right) \quad c_2 = -\frac{\rho\alpha\gamma_1+\beta\varphi\gamma_2}{\rho\alpha+\beta\varphi}$$

and $w = (1-\rho)\,\sigma^2(1-\beta)$.

This mapping was provided in Engle and Lee (1999). We also provide the reverse mapping where the component parameters are solved as a function of the GARCH(2,2) parameters, as follows

$$\beta = \tfrac{1}{2}\left(b_1-\sqrt{A}\right) \quad \alpha = \frac{a_2+a_1\beta}{\beta+\rho} \quad \gamma_1 = \frac{-\beta\varphi c_2-\alpha\rho c_2-\beta\varphi c_1-\alpha\beta c_1}{\alpha(\rho-\beta)}$$
$$\rho = \tfrac{1}{2}\left(b_1+\sqrt{A}\right) \quad \varphi = \frac{a_2+a_1\rho}{\beta+\rho} \quad \gamma_2 = \frac{\alpha\rho c_1+\rho\varphi c_1+\beta\varphi c_2+\alpha\rho c_2}{\varphi(\rho-\beta)}$$

where $A = -b_1^2 - 4b_2$. Notice that β and ρ are the inverse of the roots of

$$1 - b_1 L - b_2 L^2$$

which implies that by imposing $\beta < 1$ and $\rho < 1$, the necessary conditions for stationarity and non-negativity are imposed.

The relationship between the component GARCH model and the GARCH(2,2) model deserves further discussion. The mappings above imply that the component model can be viewed as a GARCH(2,2) model with nonlinear parameter restrictions. These restrictions yield the component structure, which enables interpretation of the model as having a potentially persistent long-run component and a rapidly mean-reverting short-run component.

In our empirical work we will demonstrate that the component model significantly outperforms the GARCH(1,1) model. Given the relationship between the component model and the GARCH(2,2) model, one may wonder about this result, because it is well-known that it is difficult to outperform a GARCH(1,1) model in standard volatility forecasting comparisons (see for instance the results in Hansen and Lunde (2005) and the references therein). However, our main metric of comparison is option valuation, which makes use of the entire (risk neutral) conditional distribution (at many horizons) implied by the GARCH model and not just the conditional variance. Component models may thus produce conditional risk neutral density forecasts that are superior to the GARCH(1,1) model even if the conditional variance forecasts are rather similar.

There are also some rather subtle but potentially important explanations for why the empirical performance of the component model might differ from that of a GARCH(2,2) model. In the GARCH(2,2) model, the stationarity requirements are quite complicated, but in the component model we simply need $\beta < 1$ and $\rho < 1$. The component structure also restricts the roots in the implied GARCH(2,2) model to be real, which turns out to be one of the necessary conditions for non-negativity, as illustrated in Nelson and Cao (1992). The component model structure is therefore much easier to implement from the point of view of finding reasonable starting values and enforcing stationarity and non-negativity of variance in estimation. This may result in better performance of the component model. Another (related) explanation can be seen by thinking of the component model as a GARCH(2,2) model with nonlinear parameter restrictions. It is well-known that such restrictions may improve the performance out-of-sample if they describe salient features of the data.

2.2.5 Risk Neutralization and Option Valuation

Given the mappings between the component model and GARCH(2,2) model, the most straightforward approach to risk neutralization and option valuation is to use the risk neutralization for the GARCH(2,2) model. Duan (1999) derives a Generalized Local Risk Neutral Valuation Relationship, under which the risk-neutral NGARCH(2,2) process can be written as

$$\begin{aligned}\ln S_{t+1} &= \ln S_t + r - \ln\left[E_t^*\left(\exp\left(\sqrt{h_{t+1}}G^{-1}\left[\Phi\left(z_{t+1}^*\right);v\right]\right)\right)\right] + \sqrt{h_{t+1}}G^{-1}\left[\Phi\left(z_{t+1}^*\right);v\right]\\ h_{t+1} &= w + \bar{b}_1 h_t + \bar{b}_2 h_{t-1} + a_1 h_t\left(G^{-1}\left[\Phi\left(z_t^*\right);v\right] - c_1\right)^2 + a_2 h_{t-1}\left(G^{-1}\left[\Phi\left(z_{t-1}^*\right);v\right] - c_2\right)^2\end{aligned}$$

where z_{t+1}^* is a standard normal random variable under the risk neutral measure, $G^{-1}[.;v]$ is the inverse CDF of the GED distribution, and Φ is the standard normal CDF. This risk neutralization involves a slight approximation which is suggested by Duan (1999) to speed up the computation. See Duan (1999) for the details.

The European call option price is calculated by Monte Carlo, simulating the above risk-neutral process and computing the sample analogue of the discounted risk neutral expectation

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

2.2.6 The Conditional Normal Special Case

In the empirical section below, we also consider the special case of the standard normal distribution, which corresponds to $v = 2$, which gives $\frac{\Gamma(\frac{5}{v})\Gamma(\frac{1}{v})}{\Gamma(\frac{3}{v})^2} = 3$. In this case the conditional variance of variance in the GARCH(1,1) case simplifies to

$$Var_t(h_{t+2}) = (2 + 4c_1^2)\, a_1^2 h_{t+1}^2$$

We refer to this model as the NGARCH(1,1)-N. We refer to the component model with a normal innovation as NGARCH(C)-N, and we get

$$\begin{aligned} Var_t(h_{t+2}) &= \left(2(\alpha + \varphi)^2 + 4(\alpha\gamma_1 + \varphi\gamma_2)^2\right) h_{t+1}^2 \\ Cov_t(R_{t+1}, h_{t+2}) &= -2(\alpha\gamma_1 + \varphi\gamma_2)\, h_{t+1}^{3/2} \end{aligned}$$

In the normal case the risk neutral dynamics are

$$\ln S_{t+1} = \ln S_t + r - \frac{1}{2} h_{t+1} + \sqrt{h_{t+1}} z_{t+1}^*$$

$$\begin{aligned} h_{t+1} = {} & w + \bar{b}_1 h_t + \bar{b}_2 h_{t-1} + a_1 h_t \left(z_t^* - c_1 - \left(\lambda + \frac{1}{2}\right)\sqrt{h_t} \right)^2 + \ldots \\ & a_2 h_{t-1} \left(z_{t-1}^* - c_2 - \left(\lambda + \frac{1}{2}\right)\sqrt{h_{t-1}} \right)^2 \end{aligned}$$

Notice that in the normal case there is a simple mapping from physical to risk neutral innovations: $z_t^* = z_t + (\lambda + 0.5)\sqrt{h_t}$. Such a simple relationship is not available in the

GED case above. See Duan (1995) for the details of risk neutralization in the conditional normal GARCH model.[25]

2.3 An Affine, Non-Normal GARCH Component Model

The affine Inverse Gaussian GARCH(1,1) model (henceforth AGARCH(1,1)-IG) was first proposed by Christoffersen, Heston and Jacobs (2005) (henceforth CHJ). The model allows for conditional skewness as well as conditional heteroskedasticity and a leverage effect, which provides flexibility to capture moneyness effects for short-term as well as long-term options. We now develop an affine component AGARCH(C)-IG model and derive a number of useful properties of the model including a closed-form option valuation formula.

2.3.1 Return Dynamics

The AGARCH(1,1)-IG model can be written as

$$
\begin{aligned}
\ln\left(\frac{S_{t+1}}{S_t}\right) &= r + \lambda h_{t+1} + z_{t+1}\sqrt{h_{t+1}}, \qquad \text{where} \\
z_t &= \frac{\left(y_t - \frac{h_t}{\eta^2}\right)}{\sqrt{h_t}/\eta}, \qquad y_t \sim IG(\delta_t) \qquad \text{and} \\
h_{t+1} &= w + \bar{b}_1 h_t + c_1 y_t + a_1 \frac{h_t^2}{y_t} \\
&= \sigma^2 + b_1\left(h_t - \sigma^2\right) + c_1\left(y_t - \frac{h_t}{\eta^2}\right) + a_1\left(\frac{h_t^2}{y_t} - h_t\eta^2 - \eta^4\right)
\end{aligned}
$$

with $b_1 = \bar{b}_1 + \frac{c_1}{\eta^2} + a_1\eta^2$ and $\sigma^2 = \frac{w + a_1\eta^4}{1 - b_1}$. The innovation y_t has an Inverse Gaussian distribution with degrees of freedom parameter $\delta_t = \frac{h_t}{\eta^2}$ and the variance innovation term

[25] See also Amin and Ng (1993). See Brennan (1979), Camara (2003), Rubinstein (1976) and Schroder (2004) for option valuation in discrete time for the constant volatility case.

$c_1\left(y_t-\frac{h_t}{\eta^2}\right)+a_1\left(\frac{h_t^2}{y_t}-h_t\eta^2-\eta^4\right)$ has a conditional mean equal to zero. If η is negative then stock returns display negative conditional skewness. We refer to CHJ for further discussion of the Inverse Gaussian process.

Employing a similar reparameterization as in the NGARCH(C)-GED model and generalizing σ^2 to q_{t+1}, the component AGARCH(C)-IG is defined as

$$\begin{aligned} h_{t+1} &= q_{t+1}+\beta\left(h_t-q_t\right)+v_{1,t} \\ q_{t+1} &= \sigma^2+\rho\left(q_t-\sigma^2\right)+v_{2,t} \end{aligned}$$

where $v_{1,t}=\gamma_1\left(y_t-\frac{h_t}{\eta^2}\right)+\alpha\left(\frac{h_t^2}{y_t}-h_t\eta^2-\eta^4\right)$ and $v_{2,t}=\gamma_2\left(y_t-\frac{h_t}{\eta^2}\right)+\varphi\left(\frac{h_t^2}{y_t}-h_t\eta^2-\eta^4\right)$.

2.3.2 Conditional Leverage and Variance of Variance

The conditional variance of variance and the conditional leverage effect for the AGARCH(1,1)-IG model can be derived as

$$\begin{aligned} Var_t\left(h_{t+2}\right) &= 2a_1^2\eta^8+\left(\frac{c_1^2}{\eta^2}-2\eta^2a_1c_1+a_1^2\eta^6\right)h_{t+1} \\ Cov_t\left(R_{t+1},h_{t+2}\right) &= \left(\frac{c_1}{\eta}-\eta^3a_1\right)h_{t+1} \end{aligned}$$

The conditional variance of variance and the conditional leverage effect for the AGARCH(C)-IG are

$$\begin{aligned} Var_t\left(h_{t+2}\right) &= \left(\alpha+\varphi\right)^2 2\eta^8+\ldots \\ & \left(\left(\alpha+\varphi\right)^2\eta^6+\left(\gamma_1+\gamma_2\right)^2\frac{1}{\eta^2}-2\left(\alpha+\varphi\right)\left(\gamma_1+\gamma_2\right)\eta^2\right)h_{t+1} \\ Cov_t\left(R_{t+1},h_{t+2}\right) &= \left(\left(\gamma_1+\gamma_2\right)\frac{1}{\eta}-\left(\alpha+\varphi\right)\eta^3\right)h_{t+1} \end{aligned}$$

2.3.3 The Autocorrelation Function for the Squared Innovations

We now provide results related to the autocorrelation function (2.34) for the AGARCH(1,1)-IG and AGARCH(C)-IG models. For the AGARCH(1,1)-IG model we can derive

$$\begin{aligned} Cov_t(\varepsilon_{t+1}^2, h_{t+k}) &= b_1^{k-2}\left(3c_1 - a_1\eta^4\right)h_{t+1} \\ Var_t\left(\varepsilon_{t+1}^2\right) &= \left(\left(15\frac{\eta^2}{h_{t+1}} + 3\right) - 1\right)h_{t+1}^2 \end{aligned}$$

and the expected future variance is

$$E_t\left[h_{t+k-1}\right] = \sigma^2 + b_1^{k-2}(h_{t+1} - \sigma^2)$$

For the AGARCH(C)-IG we have

$$\begin{aligned} Cov_t(\varepsilon_{t+1}^2, h_{t+k}) &= \beta^{k-2}\left(3\gamma_1 - \alpha\eta^4\right)h_{t+1} + \rho^{k-2}\left(3\gamma_2 - \varphi\eta^4\right))h_{t+1} \\ Var_t\left(\varepsilon_{t+1}^2\right) &= \left(\left(15\frac{\eta^2}{h_{t+1}} + 3\right) - 1\right)h_{t+1}^2 \end{aligned}$$

and the expected future variance is

$$E_t\left(h_{t+k-1}\right) = \sigma^2 + \beta^{k-2}(h_{t+1} - q_{t+1}) + \rho^{k-2}\left(q_{t+1} - \sigma^2\right)$$

However, due to the fact that y_t appears in the denominator in the variance dynamics, a closed form solution for the conditional variance of the k-period-ahead squared innovation is not available. We therefore compute $Var_t\left(\varepsilon_{t+k}^2\right)$ for $k > 1$ by simulation in both models.

2.3.4 GARCH(2,2) Mappings

The AGARCH(C)-IG model can be mapped into the AGARCH(2,2)-IG

$$h_{t+1} = w + \bar{b}_1 h_t + \bar{b}_2 h_{t-1} + c_1 y_t + a_1\frac{h_t^2}{y_t} + c_2 y_{t-1} + a_2\frac{h_{t-1}^2}{y_{t-1}}$$

where

$$
\begin{array}{lll}
\bar{b}_1 = \rho + \beta - \dfrac{\gamma_1 + \gamma_2}{\eta^2} - (\alpha + \varphi)^2 \eta^2 & a_1 = \alpha + \varphi & c_1 = \gamma_1 + \gamma_2 \\
\bar{b}_2 = -\rho\beta + \dfrac{\rho\gamma_1 + \beta\gamma_2}{\eta^2} + (\alpha\rho + \varphi\beta)\,\eta^2 & a_2 = -\rho\alpha - \beta\varphi & c_2 = -\rho\gamma_1 - \beta\gamma_2
\end{array}
$$

and $w = (1-\rho)\,\sigma^2(1-\beta) + (\rho\alpha + \beta\varphi - \alpha - \varphi)\,\eta^4$. The reverse mapping is

$$
\begin{array}{lll}
\beta = \frac{1}{2}\left(b_1 - \sqrt{A}\right) & \alpha = \dfrac{a_2 + a_1\beta}{\beta - \rho} & \gamma_1 = \dfrac{c_2 + c_1\beta}{\beta - \rho} \\
\rho = \frac{1}{2}\left(b_1 + \sqrt{A}\right) & \varphi = -\dfrac{a_2 + a_1\rho}{\beta - \rho} & \gamma_2 = -\dfrac{c_2 + c_1\rho}{\beta - \rho}
\end{array}
$$

where $A = b_1 \pm \sqrt{b_1^2 + 4b_2}$. Therefore, β and ρ are the inverse of the roots of

$$1 - b_1 L - b_2 L^2$$

which again implies that by restricting $\beta < 1$ and $\rho < 1$, the necessary conditions for stationarity and non-negativity are imposed.

2.3.5 Risk Neutralization and Option Valuation

Relying on the mappings above, we once again limit ourselves to discussing risk neutralization and option valuation for the AGARCH(2,2)-IG model, from which everything else follows as a special case. Under the risk neutral measure, the AGARCH(2,2)-IG dynamic is given by

$$
\begin{aligned}
\ln(S_{t+1}) &= \ln(S_t) + r + \lambda^* h_{t+1}^* + z_{t+1}^* \sqrt{h_{t+1}^*} \\
z_t^* &= z_t \Pi^{-\frac{3}{4}} \\
h_{t+1}^* &= w^* + \bar{b}_1 h_t^* + \bar{b}_2 h_{t-1}^* + c_1^* y_t^* + c_2^* y_{t-1}^* + a_1^* \frac{h_t^{*2}}{y_t^*} + a_2^* \frac{h_{t-1}^{*2}}{y_{t-1}^*}
\end{aligned}
$$

where $\sigma^{*2} = \sigma^2\Pi^{3/2}, w^* = \sigma^{*2}(1-b_1-b_2) - (a_1^*+a_2^*)\Pi^4\eta^4$, $h_t^* = h_t\Pi^{3/2}$, $y_t^* = y_t\Pi^{-1}$, $c_1^* = c_1\Pi^{5/2}$, $a_1^* = a_1\Pi^{-5/2}$, $c_2^* = c_2\Pi^{5/2}$, $a_2^* = a_2\Pi^{-5/2}$, $\lambda^* = \lambda\Pi^{-3/2}$, $\eta^* = \lambda^2\eta^3(1+0.5\lambda^2\eta^3)^2$, and $\Pi = \left(\frac{\eta^*}{\eta}\right)$.

Given the risk-neutral dynamics, option valuation is relatively straightforward. We use the result of Heston and Nandi (2000) that at time t, a European call option with strike price K that expires at time T is worth

$$
\begin{aligned}
C &= e^{-r(T-t)}E_t^*[Max(S_T-K,0)] \\
&= S_t\left(\frac{1}{2}+\frac{e^{-r(T-t)}}{\pi}\int_0^\infty \mathrm{Re}\left[\frac{K^{-i\phi}f^*(t,T;i\phi+1)}{i\phi}\right]d\phi\right) - \ldots \\
& \quad Ke^{-r(T-t)}\left(\frac{1}{2}+\frac{1}{\pi}\int_0^\infty \mathrm{Re}\left[\frac{K^{-i\phi}f^*(t,T;i\phi)}{i\phi}\right]d\phi\right)
\end{aligned}
$$

where $f^*(t,T;i\phi)$ is the conditional characteristic function of the logarithm of the spot price under the risk neutral measure. Christoffersen, Heston, and Jacobs (2004) provide the moment generating function for the AGARCH(1,1)-IG model; here we provide the result for the higher order model.

First, let $f_t(\phi)$ denote the conditional generating function of the asset price $f_t(\phi) = E_t\left[S_T^\phi\right]$, which is also the moment generating function of the logarithm of S_T. In Appendix A we show that the generating function for the AGARCH(2,2)-IG model takes the form

$$
f_t(\phi) = \exp\left(\phi x_t + A_t + B_t h_t + C_t h_{t+1} + D_{1,t}y_t + D_{2,t}\frac{h_t^2}{y_t}\right)
$$

where $x_t = \ln(S_t)$. The coefficients $\{A_t, B_t, C_t, D_{1,t}, D_{2,t}\}$ depend on the parameters of the model. Appendix A shows that the coefficients in the moment generating functions are

$$A_t = A_{t+1} + \phi r + C_{t+1} w + \ln\left(\frac{\eta^{-2}}{\sqrt{\eta^{-4} - 2\left(C_{t+1} a_1 + D_{2,t+1}\right)}}\right)$$

$$\begin{aligned} C_t &= \left(\phi\widetilde{\lambda} + B_{t+1} + C_{t+1}\bar{b}_1\right) + \ldots \\ & \left(\eta^{-2} + 1\right)\sqrt{\eta^{-4} - 2\left(C_{t+1} a_1 + D_{2,t+1}\right)}\sqrt{1 - 2\left(\phi\eta + C_{t+1} c_1 + D_{1,t+1}\right)} \end{aligned}$$

$$B_t = C_{t+1}\bar{b}_2 \qquad D_{1,t} = C_{t+1} c_2 \qquad D_{2,t} = C_{t+1} a_2$$

where $\widetilde{\lambda} = \lambda - \eta^{-1}$ and where we have the terminal conditions $A_T = B_T = C_T = D_{1,T} = D_{2,T} = 0$.

2.3.6 The Conditional Normal Limiting Case

Christoffersen, Heston and Jacobs (2004) show that as η approaches zero, the AGARCH(1,1)-IG model converges to the conditionally normal affine GARCH(1,1) model in Heston and Nandi (2000). We refer to the latter model as AGARCH(1,1)-N in this paper. We now derive a corresponding result for the component IG model derived above, showing that it converges to the conditionally normal affine component GARCH model in Chapter 1 as η approaches zero. We refer to the latter model as AGARCH(C)-N.

We first write $\frac{1}{y_t}$ as $\frac{1}{\delta_t\left(\frac{y_t}{\delta_t}-1+1\right)}$ and then do a Taylor expansion of the term $\frac{y_t}{\delta_t}-1$ around zero. This changes the short-run and long-run components into

$$
\begin{aligned}
h_{t+1} &= q_t+\beta h_t+\gamma_1\left(y_t-\frac{h_t}{\eta^2}\right) \\
&\quad +\alpha\left(\frac{h_t^2}{\delta_t}\left(1-\left(\frac{y_t}{\delta_t}-1\right)+\left(\frac{y_t}{\delta_t}-1\right)^2+O\left(\frac{y_t}{\delta_t}-1\right)^3\right)-h_t\eta^2-\eta^4\right) \\
q_{t+1} &= \sigma^2+\rho\left(q_t-\sigma^2\right)+\gamma_2\left(y_t-\frac{h_t}{\eta^2}\right)+\varphi\left(\frac{h_t^2}{\delta_t}\left(\begin{array}{c}1-\left(\frac{y_t}{\delta_t}-1\right)+\left(\frac{y_t}{\delta_t}-1\right)^2+ \\ O\left(\frac{y_t}{\delta_t}-1\right)^3\end{array}\right)\cdots \atop -h_t\eta^2-\eta^4\right)
\end{aligned}
$$

Then we substitute $\frac{y_t}{\delta_t}-1=\frac{z_t}{\sqrt{\delta_t}}$, $\delta_t=\frac{h_t}{\eta^2}$, and we reparameterize the model using

$$
\begin{array}{ll}
\alpha=\dfrac{\alpha^N}{\eta^4} & \gamma_1=\alpha^N-2\alpha^N\gamma_1^N\eta \\
\varphi=\dfrac{\varphi^N}{\eta^4} & \gamma_2=\varphi^N-2\varphi^N\gamma_2^N\eta
\end{array}
$$

where superscript N denotes the parameter from the conditional normal component model. This gives a quadratic function of z_t that exactly matches the AGARCH(C)-N component model in Chapter 1

$$
\begin{aligned}
h_{t+1} &= q_{t+1}+\beta\left(h_t-q_t\right)+\alpha\left(\left(z_t^2-1\right)-2\gamma_1\sqrt{h_t}z_t\right) \\
q_{t+1} &= \sigma^2+\rho\left(q_t-\sigma^2\right)+\varphi\left(\left(z_t^2-1\right)-2\gamma_2\sqrt{h_t}z_t\right)
\end{aligned}
$$

plus two cubic remainder terms

$$
\frac{\alpha h_t}{\eta^2}O\left(\frac{z_t^3}{\delta_t^{3/2}}\right)=\frac{\alpha\eta}{\sqrt{h_t}}O\left(z_t^3\right) \quad \text{and} \quad \frac{\varphi h_t}{\eta^2}O\left(\frac{z_t^3}{\delta_t^{3/2}}\right)=\frac{\varphi\eta}{\sqrt{h_t}}O\left(z_t^3\right)
$$

For a fixed h_t, these remainders vanish as η approaches zero. Thus the AGARCH(C)-IG converges to the affine normal component model. By letting η go to zero, the skewness disappears and the affine Inverse Gaussian component model converges to the affine Normal component model.

2.3.7 Properties of the Affine Normal Component Model

Many useful properties of the AGARCH(C)-N and AGARCH(1,1)-N models are derived in Chapter 1 and Heston and Nandi (2000). Here we briefly report the autocorrelation function and GARCH(2,2) mappings. The ACF (2.34) can be computed using

$$
\begin{aligned}
Cov_t(\varepsilon_{t+1}^2; h_{t+k}) &= 2b_1^{k-2}a_1 h_{t+1} \\
Var_t\left(\varepsilon_{t+1}^2\right) &= 2h_{t+1}^2 \\
Var_t\left(\varepsilon_{t+k}^2\right) &= 3\sum_{i=2}^{k}\left(b_1^{k-i}\right)^2 a_1^2(2+4c_1^2 E_t\left(h_{t+i-1}\right)) + 2\left(\begin{array}{c} (\sigma^2\left(1-b_1\right)+a_1)\frac{1-b_1^{k-1}}{1-b_1}+\ldots \\ b_1^{k-1}h_{t+1} \end{array}\right)^2
\end{aligned}
$$

and where the expected future variance is

$$
E_t\left(h_{t+k-1}\right) = \sigma^2 + b_1^{k-2}(h_{t+1}-\sigma^2)
$$

The AGARCH(C)-N has the following ACF

$$
\begin{aligned}
Cov_t(\varepsilon_{t+1}^2, h_{t+k}) &= 2(\beta^{k-2}\alpha + \rho^{k-2}\varphi)h_{t+1} \\
Var_t\left(\varepsilon_{t+1}^2\right) &= 2h_{t+1}^2 \\
Var_t\left(\varepsilon_{t+k}^2\right) &= 3\sum_{i=2}^{k}\left(2\left(\beta^{k-i}\alpha+\rho^{k-i}\varphi\right)^2 + 4\left(\beta^{k-i}\alpha\gamma_1+\rho^{k-i}\varphi\gamma_2\right)^2 E_t\left(h_{t+i-1}\right)\right) \\
&\quad +2\left(\sigma^2\left(1-\rho^{k-1}\right)+\rho^{k-1}q_{t+1}+\beta^{k-1}(h_{t+1}-q_{t+1})\right)^2
\end{aligned}
$$

where the expected future variance is

$$
E_t\left(h_{t+k-1}\right) = \sigma^2 + \beta^{k-2}(h_{t+1}-q_{t+1}) + \rho^{k-2}\left(q_{t+1}-\sigma^2\right)
$$

The AGARCH(C)-N can be mapped into an AGARCH(2,2)-N as follows

$$\bar{b}_1 = (\rho+\beta) - \frac{(\alpha\gamma_1+\varphi\gamma_2)^2}{a_1} \quad a_1 = \alpha+\varphi \quad c_1 = \frac{\gamma_1\alpha+\gamma_2\varphi}{a_1}$$
$$\bar{b}_2 = -\frac{(\rho\alpha\gamma_1+\beta\varphi\gamma_2)^2}{a_2} - \rho\beta \quad a_2 = -(\rho\alpha+\beta\varphi) \quad c_2 = -\frac{\rho\gamma_1\alpha+\varphi\gamma_2\beta}{a_2}$$

and $w = (\omega-\varphi)(1-\beta) - \alpha(1-\rho)$. The reverse mapping is

$$\beta = \tfrac{1}{2}\left(b_1 - \sqrt{A}\right) \quad \alpha = a_1 - \frac{a_1 b_1^2 + 4a_1 b_2}{2A} - \frac{2a_2 + a_1 b_1}{2\sqrt{A}} \quad \gamma_1 = \frac{\beta\varphi c_2 + \alpha\rho c_2 - \alpha\beta c_1 - \beta\varphi c_1}{\alpha(\rho-\beta)}$$
$$\rho = \tfrac{1}{2}\left(b_1 + \sqrt{A}\right) \quad \varphi = \frac{a_1 b_1^2 + 4a_1 b_2}{2A} + \frac{2a_2 + a_1 b_1}{2\sqrt{A}} \quad \gamma_2 = \frac{\alpha\rho c_1 + \rho\varphi c_1 - \beta\varphi c_2 - \alpha\rho c_2}{\varphi(\rho-\beta)}$$

where $A = b_1^2 + 4b_2$. Notice again that the solutions for β and ρ above are the roots of the polynomial $Y^2 - b_1 Y - b_2$. Therefore, $\beta < 1$ and $\rho < 1$ are required for the variance to be stationary in the GARCH(2,2).

Option valuation can be done as in the AGARCH-IG model. The MGF for the AGARCH(2,2)-N model is shown in Appendix B, which corrects some typos in Heston and Nandi (2000).

2.4 Empirical Results

This section presents the empirical results. We use MLE on a long time series of S&P500 return data to estimate the eight models discussed above: NGARCH(1,1)-GED, NGARCH(C)-GED, NGARCH(1,1)-N, NGARCH(C)-N, AGARCH(1,1)-IG, AGARCH(C)-IG, AGARCH(1,1)-N and AGARCH(C)-N. We discuss the parameter estimates and their implications for the salient properties of the models. The eight models allow us to make three types of comparisons: component models versus GARCH(1,1) models; affine models versus non-affine models; and non-normal innovations versus normal innovations. Subsequently we introduce the options data. We use each of our eight models to price the option contracts and we compare model and market prices for various maturities, strike prices, and sample years.

2.4.1 Parameter Estimates on Daily Return Data

Table 1 presents the Maximum Likelihood estimation results obtained using daily returns data from July 1, 1962 through December 31, 2001. The returns data were obtained from CRSP. Standard errors are calculated from the outer product of the gradient and are given in parentheses. Table 1 reports the physical conditional variance parameters as well as the price of risk, λ. We use variance targeting for all models, we use variance targeting, forcing the annualized return standard deviation to be 14.7%. This technique fixes the parameter w in each model, and we therefore do not report on w in Table 1.

We first note from Table 1 that the price of risk, λ. is positive and significant in all models–although only marginally so in the case of AGARCH(1,1)-N. Next, notice that b_1, the variance persistence in the GARCH(1,1) models, is close to one in all four models. The fourth row from the bottom reports the overall variance persistence in the component models, $\rho + \beta(1-\rho)$, as well as b_1 for the GARCH(1,1) models. Notice that while the GARCH(1,1) models have high persistence, for each corresponding component model the persistence is even higher. The very large component variance persistence is driven by a large long-run component persistence ρ, plus the contribution from ($(1-\rho)$ times) the less persistent short-run component β.

In the GARCH(1,1) model the conditional leverage is driven by c_1. which as expected is significantly positive in all cases. In the component models, the conditional leverage effect is driven by a combination of γ_1 and γ_2, which are both significantly positive in all four component models. Thus, both the long-run and short-run components contribute to the overall leverage effect with the expected sign. The unconditional leverage effects are reported in the second row from the bottom. They are all negative, as expected. The results show that for each set of models, the component model displays a more pronounced leverage effect than the corresponding GARCH(1,1).

The variance of variance is driven mainly by the a_1 parameter in the GARCH(1,1) models and by the α and φ parameters in the component models. The overall unconditional variance of variance is reported in the third row from the bottom. Notice again that in each case the component model displays a larger variance of variance than its GARCH(1,1) counterpart. Thus three important empirical regularities emerge when comparing component models to their GARCH(1,1) counterparts: The component models allow us to (simultaneously) capture a larger variance persistence, a larger leverage effect, and a larger variance of variance than their GARCH(1,1) counterparts. Finally, Table 1 also presents standard likelihood ratio tests of the component model versus the corresponding nested GARCH(1,1) model. As the reported p-values show, each GARCH(1,1) model is strongly rejected in favor of the corresponding component model in all cases.

2.4.2 Dynamic Model Properties

In order to explore the models further, Figure 1 plots the conditional variances for the period 1989-2001. This period includes the dates for the option valuation exercise we present in Section 4.4. Recall that each model is estimated forcing the annual standard deviation to be 14.7%, corresponding to an unconditional variance per day of $8.5750E-005$. Notice that the conditional variance patterns across the four GARCH(1,1) models in the left column and the corresponding four component models in the right column display some similarities. The models all capture the low variance during the equity market run-up in 1993-1998, preceded by higher volatility during the first Gulf war and the 1990-1991 recession. The LTCM and Russia debacles in the fall of 1998 are evident, as is the higher volatility during the dot-com bust in the later part of the sample.

However, Figure 1 also reveals some important differences between models. The non-affine models (in the two top rows) appear to display much more variation in the conditional variance during the more recent period than do the two affine models (in the bottom two rows). This difference is also evident in Figure 2 which plots the long-run variance component, q_{t+1} (left column) and short-run variance component, $h_{t+1} - q_{t+1}$ (right-column) for the four component models. The non-affine components appear to be more variable than the affine components, both in the case of the long-run and short-run components. This is again particularly evident during the 1998-2001 period.

We plot the conditional variance of variance path, $Var_t\left(h_{t+2}\right)$ for each of the eight models in Figure 3. Figure 3 confirms the findings in Figures 1 and 2. The non-affine models in the two top rows of Figure 3 display a much larger variance of variance than the two affine models in the bottom two panels. This is true for both the GARCH(1,1) models in the left column and the component models in the right column.

Figure 4 plots the conditional leverage path, $Cov_t\left(R_{t+1}, h_{t+2}\right)$ for each of the eight models we consider. The left-hand column contains the single component GARCH(1,1) models and the right-hand column contains the component models. Notice that in each case the component model has a larger (more negative) and more variable leverage effect than the corresponding GARCH(1,1) model. This is particularly true for the two affine models and the NGARCH-GED, and less so for the NGARCH model. The large discrepancy in the leverage effect between the NGARCH(1,1)-GED and the NGARCH(1,1)-N may seem puzzling. It is however confirmed by the much larger c_1 parameter in the latter model in Table 1.

While Figures 1-4 depict various aspects of the dynamics of the one-day ahead conditional distribution, Figure 5 captures the properties of the variance dynamics across longer horizons. We plot the conditional autocorrelation function of the squared innovations, $Corr_t(\varepsilon_{t+1}^2, \varepsilon_{t+k}^2)$ across $k = 1, \ldots 250$ days for each of the eight models we consider. The top-left panel contains the non-affine GARCH model with GED shocks, the top-right panel depicts the non-affine GARCH model with normal innovations, the bottom-left panel represents the affine GARCH with inverse Gaussian shocks and the bottom-right panel contains the affine GARCH model with normal shocks. Each panel contains the component GARCH (solid line) and the GARCH(1,1) (dashed line) model. The conditional variance is set equal to the unconditional sample variance for all models, and the parameters are the MLE estimates reported in Table 1.

For each of the four pairwise comparisons, the autocorrelation function for the component model is below that of the GARCH(1,1) model for short horizons but above it for longer horizons. In this sense, the component model displays long-memory like features. While both the GARCH(1,1) and the component models are truly short-memory exponentially decaying models, the dynamic properties of the component models are similar to those of long-memory models for the horizons we care about for option valuation, namely 1-250 trading days. Another interesting observation from Figure 5 is that the non-affine models in the top two panels have larger autocorrelations than the two affine models in the bottom row. This difference may have important implications for the valuation of long-maturity options. We now turn to an option valuation exercise to investigate this further.

2.4.3 Option Data and Valuation Methodology

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. We also only use Wednesday options data. 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 considering the highly persistent volatility processes.

Table 2 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 with 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 2 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 is 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 result in systematic pricing errors, which motivates the use of stochastic volatility and GARCH models for option valuation.

When calculating option prices according to the eight GARCH models, we use the MLE parameters in Table 1 transformed to the risk neutral measure. These risk-neutral parameters as well as the conditional variance paths from Figure 1 are used as inputs into the option pricing formula. In the case of the non-affine models, the formula requires Monte Carlo simulation to calculate the price, whereas in the case of the affine models numerical integration is used.

2.4.4 Option Valuation Results

The overall $RMSE$s for the eight GARCH model are reported in the last row of Table 1. The $RMSE$ is computed as

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

where N is equal to 21,752, the total number of option contracts in the sample.

The results in Table 1 allow us to make three types of comparisons. We first focus on the performance of the component models versus the GARCH(1,1) models. Most importantly, note that the best overall model (i.e. the one with the lowest $RMSE$) is the NGARCH(C)-GED. Moreover, for each of the four pairwise comparisons, the $RMSE$ of the component model is much lower than the $RMSE$ of the corresponding GARCH(1,1) model. The differences are large, ranging from a 21% improvement in the AGARCH-IG case (1.705 versus 2.162) to a 38% improvement in the NGARCH-N case (1.466 versus 2.356).

The second comparison is between models with normal and non-normal innovations. In this case, the differences are smaller but systematic. The NGARCH-GED improves on the NGARCH-N model by 13% in the GARCH(1,1) case and by 1% in the component case. The AGARCH-IG improves on the AGARCH-N by 7% in the GARCH(1,1) case and 6% in the component case.

The third comparison is between affine and non-affine models.[26] The $RMSE$ of the best non-affine model (the component NGARCH-GED) is 14% lower than that of the best affine model (the component AGARCH-IG). When conducting pairwise comparisons, the non-affine models generally have lower $RMSE$s than their affine counterparts. The only exception is the NGARCH(1,1)-N which has a slightly larger $RMSE$ than the corresponding AGARCH(1,1)-N.

Table 3 provides more detail on the option valuation results. In Panel A we report the $RMSE$ for each of six moneyness bins, where the $RMSE$ has been divided by the average market option price for that bin (from Table 2, Panel B). Looking across the rows of Panel A, we see that in each row but one, the best model is a component model. The only

[26] Hsieh and Ritchken (2000) compare the fit of affine and non-affine single component conditional Gaussian models. Our main focus of course is on two-component, non-Gaussian models.

exception is for deep-in-the-money options where the AGARCH(1,1)-IG is best. We also see that the overall best model, namely the component NGARCH-GED is best or near-best in every row. Interestingly, the non-affine models tend to do well for the out-of-the-money options in the top rows, whereas the affine models do well for the in-the-money options in the bottom rows.

In Panel B of Table 3, we report the $RMSE$ for each of four moneyness bins, where the $RMSE$ again has been divided by the average market option price for that bin. In each of the four rows, a component model is the best performer. The component NGARCH-GED is once again best or near-best in every row. Finally, Panel C reports the normalized $RMSE$ for each of the years in the option sample. A component model performs the best in all but one year, namely 1991, when the AGARCH(1,1)-IG is the top performer.

2.4.5 Discussion

We have considered eight GARCH models that differ along three important dimensions. Four of the models have non-affine dynamics while four have affine dynamics, four models are of the GARCH(1,1) type while the other four are component volatility models, and we have four models each with normal and non-normal innovations.

The most important empirical regularity we observe is that component models are strongly favored by the data over GARCH(1,1) models. This is the case when we use likelihood values based on returns data, but also when we use $RMSE$s based on options data (judging from $RMSE$s). When using returns data, non-affine models display very different properties than affine models, and non-normal innovations outperform normal innovations. However, differences in option fit are much less significant for these types of comparisons.

The $RMSE$ criterion is clearly different from the likelihood-based criterion, and this in itself can explain the results. However, it is important to note that the $RMSE$-based comparison also differs from the likelihood-based comparison in a methodological sense. The comparisons based on option prices are of an out-of-sample nature, while this is not the case for the likelihood-based comparison. While the option sample time period is part of the sample period used for ML estimation, the GARCH model parameters are estimated on returns only. Our finding regarding the performance of the component models is therefore much more robust than the findings regarding non-normal innovations and affine restrictions, because these results are not as strongly supported out-of-sample. Moreover, because the option valuation results are out-of-sample, the finding that the more richly parameterized component GARCH models are outperforming more parsimonious models is completely non-trivial.

It is also important to note that other studies have documented that the benchmark NGARCH(1,1)-N and AGARCH(1,1)-N work very well. Christoffersen and Jacobs (2004) find that the NGARCH(1,1)-N model is almost impossible to improve upon by changing the news impact specification of the GARCH(1,1) model. Heston and Nandi (2000) find that the AGARCH(1,1)-N model performs well relative to the standard model-free benchmark in Dumas, Fleming and Whaley (1998).

2.5 Conclusion and Directions for Future Work

This paper presents two new conditional non-normal GARCH variance component models. The first model allows for GED innovations to the variance dynamic. Because the model is characterized by a more traditional non-affine GARCH variance dynamic, option valuation must be done by Monte Carlo simulation. The second model is characterized by a conditional inverse Gaussian innovation and by affine variance dynamics. A closed-form option valuation formula is derived for this model. The two new non-normal component models are compared with the corresponding special cases with normal innovations, and the resulting four component models are compared with the GARCH(1,1) models which they nest. All eight models are estimated using MLE on a long time series of S&P500 returns. The likelihood criterion strongly favors the component models in all cases, and it also favors non-normal innovations. Non-affine models and affine models differ along several critical dimensions, such as conditional leverage and variance of variance. When we use the models' parameter estimates for option valuation, we find very strong support for the component variance specifications. The support for non-normal innovations and for the non-affine structure is less strong.

The empirical results leave a few questions unanswered. First, it remains to be seen if the differences in performance between models are confirmed when using model parameters estimated from option prices, or when using an integrated analysis that uses option prices as well as underlying returns (see Bates (2000), Chernov and Ghysels (2000), Eraker (2004) and Pan (2002)). Second, it would be useful to reconcile the relationship between the superior option valuation performance of the component models we find here and the less than superior performance of GARCH(2,2) models in traditional volatility forecasting studies. Comparing the density forecasts implied by the different models could be an avenue to explore. Finally, looking forward, it would be interesting to compare the range of discrete-time GARCH models considered here with the continuous-time stochastic volatility models that are popular in the finance literature. Bakshi, Cao and Chen (1997), Bates (1996), and Eraker (2004) study stochastic volatility models with Poisson jumps, and Bates (2000) analyses models with Poisson jumps and multiple volatility factors. Recently, Carr and Wu (2004) and Huang and Wu (2004) have considered Levy processes with infinitely many jumps. The relationships between the continuous-time and discrete-time models are very interesting, and comparing the models for the purpose of option valuation may provide more insight into the strengths and weaknesses of the component models.

2.6 Appendix

2.6.1 The AGARCH(2,2)-IG MGF

Let $x_t = \ln(S_t)$ and let f_t be the conditional generating function of S_T, or equivalently the conditional moment generating function (MGF) of x_T, i.e.

$$f_t = E_t[\exp(\phi x_T)]$$

We shall guess that the moment generating function takes the log-linear form

$$f_t = \exp\left(\phi x_t + A_t + B_t h_t + C_t h_{t+1} + D_{1,t} y_t + D_{2,t}\frac{h_t^2}{y_t}\right)$$

Since x_T is known at time T, we have the terminal condition

$$A_T = B_T = C_T = D_{1,T} = D_{2,T} = 0$$

Applying the law of iterated expectations to f_t we get,

$$f_t = E_t\left[f_{t+1}\right] = E_t \exp\left(\phi x_{t+1} + A_{t+1} + B_{t+1} h_{t+1} + C_{t+1} h_{t+2} + D_{1,t+1} y_{t+1} + D_{2,t+1}\frac{h_{t+1}^2}{y_{t+1}}\right)$$

We first rewrite the return dynamic as

$$\ln\left(\frac{S_{t+1}}{S_t}\right) = r + \widetilde{\lambda} h_{t+1} + \eta y_{t+1}$$

where $\widetilde{\lambda} = \lambda - \frac{1}{\eta}$. Substituting in the dynamics of x_{t+1} and h_{t+2} yields

$$\begin{aligned}
f_t &= E_t \exp \left(\begin{array}{c} \phi(x_t + r) + \phi\widetilde{\lambda} h_{t+1} + \phi\eta y_{t+1} + A_{t+1} + B_{t+1} h_{t+1} + \ldots \\ C_{t+1} h_{t+2} + D_{1,t+1} y_{t+1} + D_{2,t+1} \frac{h_{t+1}^2}{y_{t+1}} \end{array} \right) \\
&= E_t \exp \left(\begin{array}{c} \phi(x_t + r) + \phi\widetilde{\lambda} h_{t+1} + \phi\eta y_{t+1} + A_{t+1} + B_{t+1} h_{t+1} + \\ C_{t+1} \left(w + \bar{b}_1 h_{t+1} + \bar{b}_2 h_t + c_1 y_{t+1} + \frac{a_1 h_{t+1}^2}{y_{t+1}} + c_2 y_t + \frac{a_2 h_t^2}{y_t} \right) + \ldots \\ D_{1,t+1} y_{t+1} + D_{2,t+1} \frac{h_{t+1}^2}{y_{t+1}} \end{array} \right) \\
&= E_t \exp \left(\begin{array}{c} \phi(x_t + r) + A_{t+1} + C_{t+1} w + \left(\phi\widetilde{\lambda} + B_{t+1} + C_{t+1} \bar{b}_1 \right) h_{t+1} + C_{t+1} \bar{b}_2 h_t + \\ C_{t+1} c_2 y_t + C_{t+1} a_2 \frac{h_t^2}{y_t} + \left(\phi\eta + C_{t+1} c_1 + D_{1,t+1} \right) y_{t+1} + \left(C_{t+1} a_1 + D_{2,t+1} \right) \frac{h_{t+1}^2}{y_{t+1}} \end{array} \right)
\end{aligned}$$

where we have applied the general result for an $IG(\delta)$ variable, y, and constants, a and b,

$$E\left[\exp\left(ay + b/y\right)\right] = \frac{\delta}{\sqrt{\delta^2 - 2b}} \exp\left(\delta - \sqrt{\left(\delta^2 - 2b\right)\left(1 - 2a\right)}\right)$$

Solving this expectation and equating coefficients demonstrates

$$\begin{aligned}
f_t &= E_t \exp \left(\begin{array}{c} \phi(x_t + r) + A_{t+1} + C_{t+1} w + \left(\phi\widetilde{\lambda} + B_{t+1} + C_{t+1} \bar{b}_1 \right) h_{t+1} + \\ C_{t+1} \bar{b}_2 h_t + C_{t+1} c_2 y_t + C_{t+1} a_2 \frac{h_t^2}{y_t} + \ln\left(\frac{h_{t+1}\eta^{-2}}{\sqrt{h_{t+1}^2 \eta^{-4} - 2(C_{t+1}a_1 + D_{2,t+1}) h_{t+1}^2}} \right) + h_{t+1}\eta^{-2} + \\ \sqrt{h_{t+1}^2 \eta^{-4} - 2\left(C_{t+1}a_1 + D_{2,t+1}\right) h_{t+1}^2} \sqrt{1 - 2\left(\phi\eta + C_{t+1}c_1 + D_{1,t+1}\right)} \end{array} \right) \\
&= E_t \exp \left(\begin{array}{c} \phi(x_t + r) + A_{t+1} + C_{t+1} w + \left(\phi\widetilde{\lambda} + B_{t+1} + C_{t+1} \bar{b}_1 \right) h_{t+1} + \\ C_{t+1} \bar{b}_2 h_t + C_{t+1} c_2 y_t + C_{t+1} a_2 \frac{h_t^2}{y_t} + \ln\left(\frac{\eta^{-2}}{\sqrt{\eta^{-4} - 2(C_{t+1}a_1 + D_{2,t+1})}} \right) + \\ \left(\eta^{-2} + 1\right) h_{t+1} \sqrt{\eta^{-4} - 2\left(C_{t+1}a_1 + D_{2,t+1}\right)} \sqrt{1 - 2\left(\phi\eta + C_{t+1}c_1 + D_{1,t+1}\right)} \end{array} \right)
\end{aligned}$$

Therefore

$$\begin{aligned}
A_t &= A_{t+1} + \phi r + C_{t+1} w + \ln\left(\frac{\eta^{-2}}{\sqrt{\eta^{-4} - 2\left(C_{t+1}a_1 + D_{2,t+1}\right)}} \right) \\
C_t &= \left(\phi\widetilde{\lambda} + B_{t+1} + C_{t+1} \bar{b}_1 \right) + \ldots \\
&\quad \left(\eta^{-2} + 1\right) \sqrt{\eta^{-4} - 2\left(C_{t+1}a_1 + D_{2,t+1}\right)} \sqrt{1 - 2\left(\phi\eta + C_{t+1}c_1 + D_{1,t+1}\right)} \\
B_t &= C_{t+1} \bar{b}_2,\ D_{1,t} = C_{t+1} c_2,\ D_{2,t} = C_{t+1} a_2
\end{aligned}$$

2.6.2 The AGARCH(2,2)-N MGF

We shall guess that the MGF takes the log-linear form

$$f_t = \exp\left(\phi x_t + A_t + B_{1,t}h_{t+1} + B_{2,t}h_t + C_t(z_t - c_2\sqrt{h_t})^2\right)$$

where $x_t = \ln(S_t)$. We have

$$f_t = E_t\left[f_{t+1}\right] = E_t\left[\exp\left(\phi x_{t+1} + A_{t+1} + B_{1,t+1}h_{t+2} + B_{2,t+1}h_{t+1} + C_{t+1}(z_{t+1} - c_2\sqrt{h_{t+1}})^2\right)\right] \tag{2.40}$$

Since x_T is known at time T, we require the terminal condition

$$A_T = B_{i,T} = C_T = 0$$

Substituting the dynamics of x_{t+1} into (2.40) and rewriting we get

$$f_t = E_t \exp\left(\begin{array}{c} \phi(x_t + r) + (B_{1,t+1}a_1 + C_{t+1})\left(z_{t+1} - (\bar{c}_{t+1} - \frac{\phi}{2(B_{1,t+1}a_1 + C_{t+1})})\sqrt{h_{t+1}}\right)^2 + \\ A_{t+1} + B_{1,t+1}w + B_{1,t+1}\bar{b}_2h_t + B_{1,t+1}a_2(z_t - c_2\sqrt{h_t})^2 + \\ \left(\begin{array}{c} \phi\lambda + B_{1,t+1}\bar{b}_1 + B_{2,t+1} + (\phi\bar{c}_{t+1} - \frac{\phi^2}{4(B_{1,t+1}a_1 + C_{t+1})}) + \\ (B_{1,t+1}a_1c_1^2 + C_{t+1}c_2^2) - \bar{c}_{t+1}\left(B_{1,t+1}a_1c_1 + C_{t+1}c_2\right) \end{array}\right) h_{t+1} \end{array}\right) \tag{2.41}$$

where

$$\bar{c}_{t+1} = \frac{B_{1,t+1}a_1c_1 + C_{t+1}c_2}{B_{1,t+1}a_1 + C_{t+1}}$$

and we have used

$$\begin{aligned} &(B_{1,t+1}a_1 + C_{t+1})\left(z_{t+1} - (\bar{c}_{t+1} - \frac{\phi}{2(B_{1,t+1}a_1 + C_{t+1})})\sqrt{h_{t+1}}\right)^2 \\ &= B_{1,t+1}a_1(z_{t+1} - c_1\sqrt{h_{t+1}})^2 + C_{t+1}(z_{t+1} - c_2\sqrt{h_{t+1}})^2 \\ &+ z_{t+1}\phi\sqrt{h_{t+1}} + \left(-\phi\bar{c}_{t+1} + \frac{\phi^2}{4(B_{1,t+1}a_1 + C_{t+1})}\right) h_{t+1} \\ &- \left(B_{1,t+1}a_1c_1^2 + C_{t+1}c_2^2 - \bar{c}_{t+1}\left(B_{1,t+1}a_1c_1 + C_{t+1}c_2\right)\right) h_{t+1} \end{aligned}$$

Using the general result for a standard normal variable, z, and constants, a and b,

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

in (2.41) we get

$$f_t = \exp\left(\begin{array}{c} \phi(x_t + r) + A_{t+1} + B_{1,t+1}w - \frac{1}{2}\ln(1 - 2B_{1,t+1}a_1 - 2C_{t+1}) + \\ B_{1,t+1}\bar{b}_2 h_t + B_{1,t+1}a_2(z_t - c_2\sqrt{h_t})^2 + \frac{(B_{1,t+1}a_1 + C_{t+1})(\bar{c}_{t+1} - \frac{\phi}{2(B_{1,t+1}a_1 + C_{t+1})})^2}{1 - 2B_{1,t+1}a_1 - 2C_{t+1})} h_{t+1} + \\ \left(\begin{array}{c} \phi\lambda + B_{1,t+1}\bar{b}_1 + B_{2,t+1} + (\phi\bar{c}_{t+1} - \frac{\phi^2}{4(B_{1,t+1}a_1 + C_{t+1})}) + \ldots \\ (B_{1,t+1}a_1c_1^2 + C_{t+1}c_2^2) - \bar{c}_{t+1}\left(B_{1,t+1}a_1c_1 + C_{t+1}c_2\right) \end{array}\right) h_{t+1} \end{array}\right)$$

Matching terms gives

$$A_t = A_{t+1} + \phi r + B_{1,t+1}w - \frac{1}{2}\ln(1 - 2B_{1,t+1}a_1 - 2C_{t+1})$$

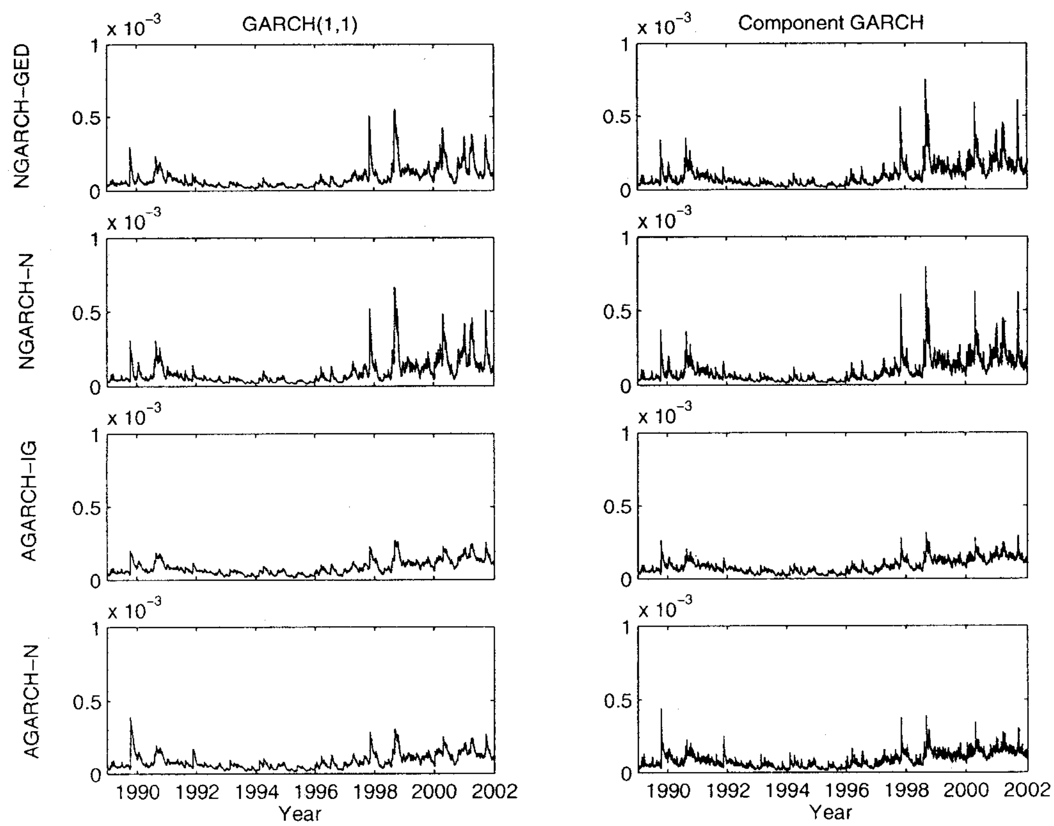
$$\begin{aligned} B_{1,t} &= \phi\lambda + B_{1,t+1}\bar{b}_1 + B_{2,t+1} + \left(B_{1,t+1}a_1c_1^2 + C_{t+1}c_2^2\right) \\ &\quad + \frac{1/2\phi^2 + 2\left(B_{1,t+1}a_1c_1 + C_{t+1}c_2\right)\left(B_{1,t+1}a_1c_1 + C_{t+1}c_2 - \phi\right)}{1 - 2B_{1,t+1}a_1 - 2C_{t+1}} \\ B_{2,t} &= B_{1,t+1}\bar{b}_2, \qquad C_t = B_{1,t+1}a_2 \end{aligned}$$

where we have used the fact that

$$\begin{aligned} &\phi\bar{c}_{t+1} - \frac{\phi^2}{4(B_{1,t+1}a_1 + C_{t+1})} - \bar{c}_{t+1}\left(B_{1,t+1}a_1c_1 + C_{t+1}c_2\right) \\ &+ \frac{(B_{1,t+1}a_1 + C_{t+1})(\bar{c}_{t+1} - \frac{\phi}{2(B_{1,t+1}a_1 + C_{t+1})})^2}{1 - 2B_{1,t+1}a_1 - 2C_{t+1}} \\ = &\ \frac{1/2\phi^2 + 2\left(B_{1,t+1}a_1c_1 + C_{t+1}c_2\right)\left(B_{1,t+1}a_1c_1 + C_{t+1}c_2 - \phi\right)}{1 - 2B_{1,t+1}a_1 - 2C_{t+1}} \end{aligned}$$

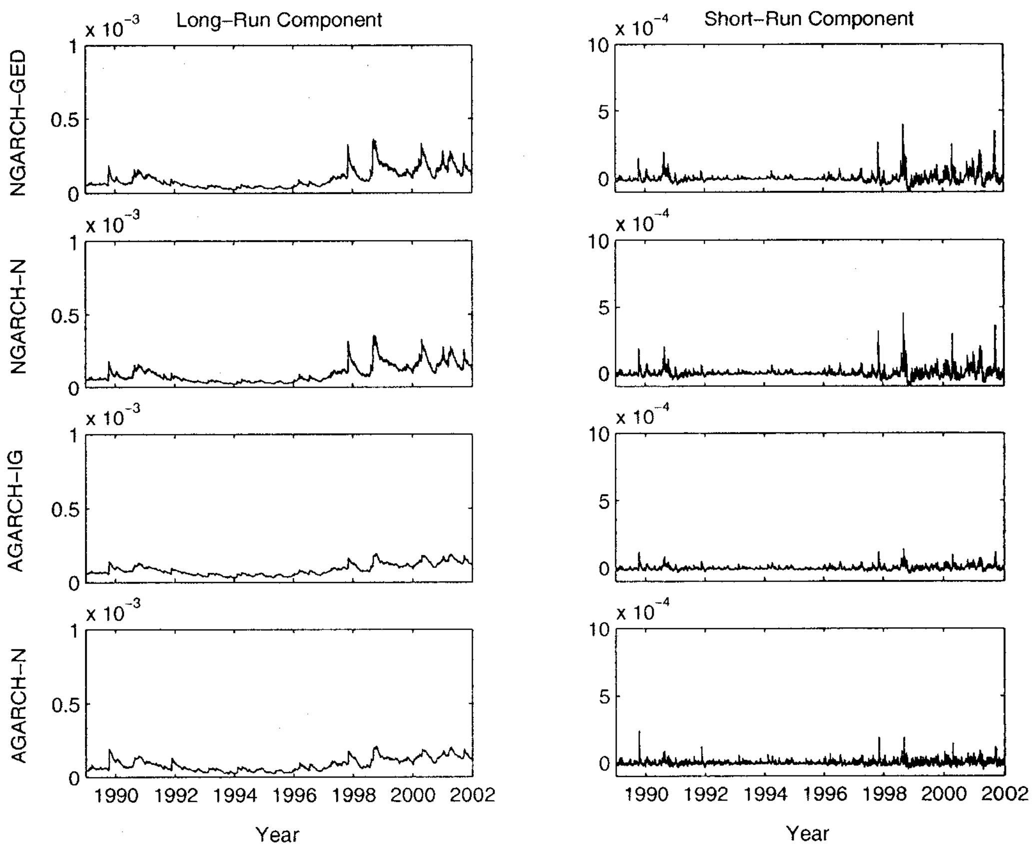
2.7 Figures and Tables

Figure 1. Conditional Variance Paths



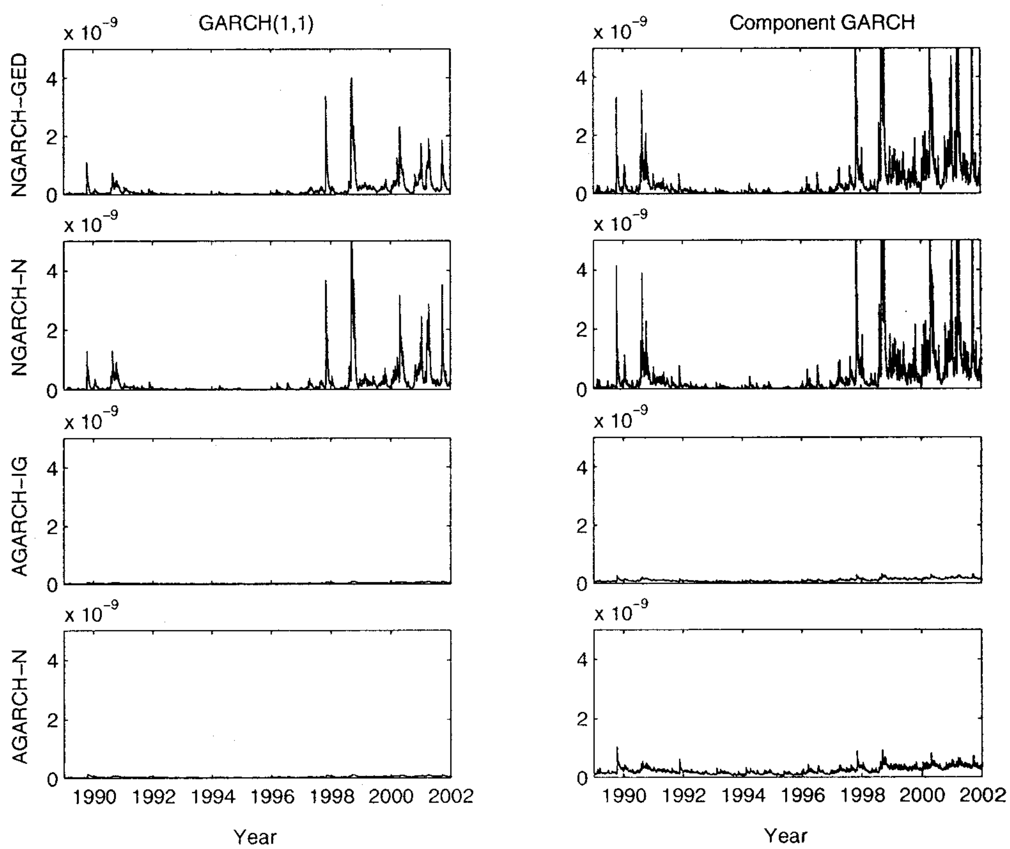
Notes to Figure: We plot the conditional variance path, h_{t+1}, for each of the eight models we consider. The left-hand column contains the single component GARCH(1,1) models and the right-hand column contains the two-component models. Rows 1 and 2 contain the non-affine GARCH models with GED shocks and Normal shocks, followed by the affine GARCH models with IG and Normal shocks in rows 3 and 4. The parameter values from the underlying GARCH models are obtained from MLE estimation on S&P500 returns as reported in Table 1.

Figure 2. Variance Component Paths



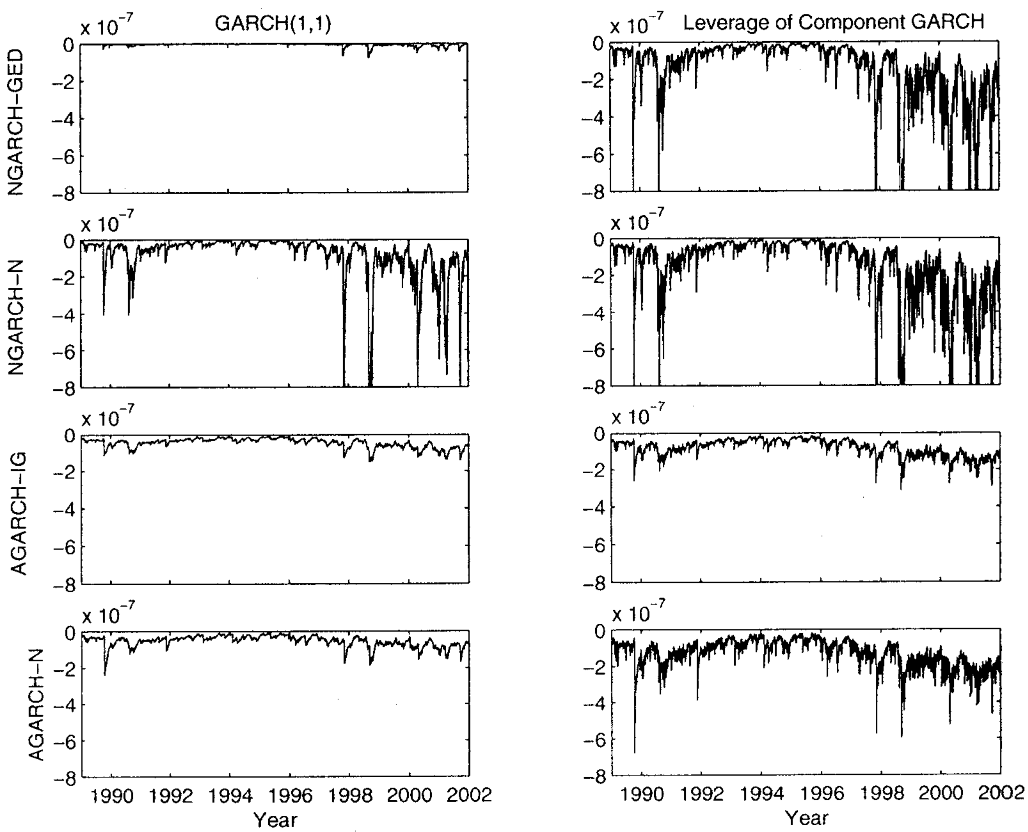
Notes to Figure: We plot the two variance components for the four component models we consider. For each model, the left-hand column contains the long-run component, q_{t+1} and the right-hand panel contains the short-run component, $h_{t+1} - q_{t+1}$. Rows 1 and 2 contain the non-affine GARCH models with GED shocks and Normal shocks, followed by the affine GARCH models with IG and Normal shocks in rows 3 and 4. The parameter values from the underlying GARCH models are obtained from MLE estimation on S&P500 returns as reported in Table 1.

Figure 3. Conditional Variance of Variance Paths



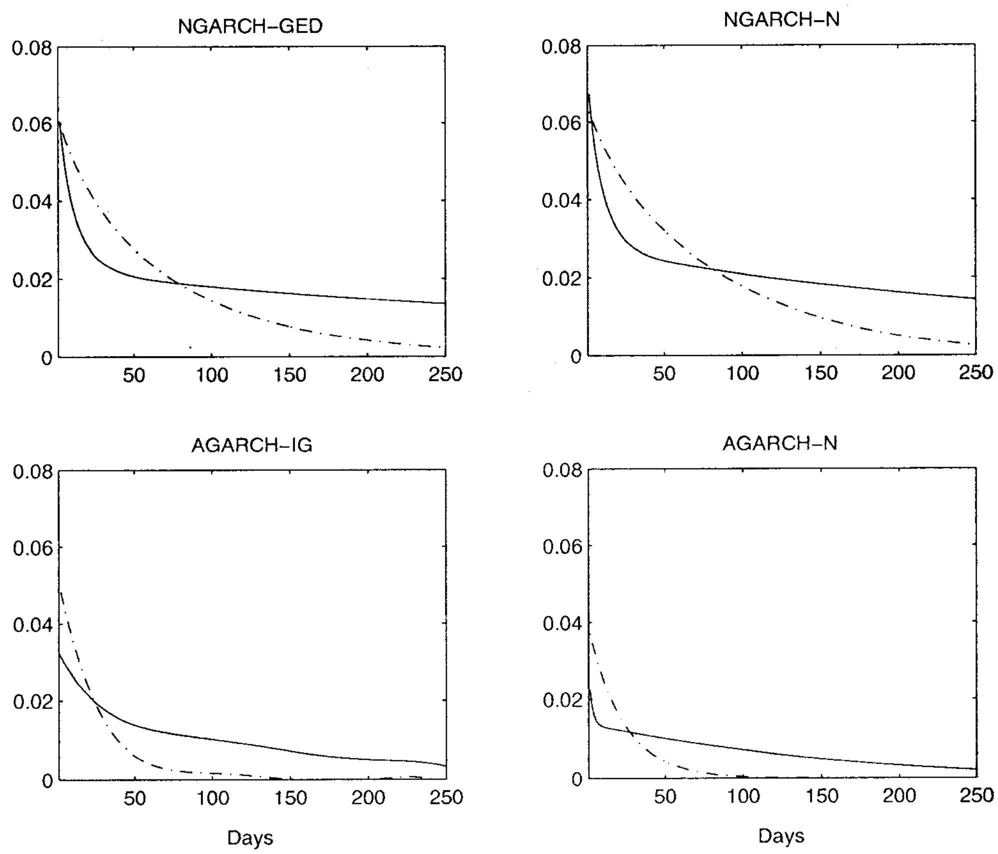
Notes to Figure: We plot the conditional variance of variance path, $Var_t\left(h_{t+2}\right)$, for each of the eight models we consider. The left-hand column contains the single component GARCH(1,1) models and the right-hand column contains the two-component models. Rows 1 and 2 contain the non-affine GARCH models with GED shocks and Normal shocks, followed by the affine GARCH models with IG and Normal shocks in rows 3 and 4. The parameter values from the underlying GARCH models are obtained from MLE estimation on S&P500 returns as reported in Table 1.

Figure 4. Conditional Leverage Paths



Notes to Figure: We plot the conditional leverage path, $Cov_t(R_{t+1}, h_{t+2})$ for each of the eight models we consider. The left-hand column contains the single component GARCH(1,1) models and the right-hand column contains the two-component models. Rows 1 and 2 contain the non-affine GARCH models with GED shocks and Normal shocks, followed by the affine GARCH models with IG and Normal shocks in rows 3 and 4. The parameter values from the underlying GARCH models are obtained from MLE estimation on S&P500 returns as reported in Table 1.

Figure 5. Autocorrelation Function of Component GARCH and GARCH(1,1)



Notes to Figure: We plot the conditional autocorrelation function of the squared innovations, $Corr_t(\varepsilon_{t+1}^2, \varepsilon_{t+k}^2)$, for each of the eight models we consider. The top-left panel contains the non-affine GARCH model with GED shocks, the top-right panel contains has normal shocks, the bottom-left panel contains the affine GARCH with inverse Gaussian shocks and the bottom-right panel has normal shocks. Each panel contains a component GARCH (solid line) and a GARCH(1,1) (dashed line) model. The conditional variance is set to the unconditional sample variance in each model. The parameter values from the underlying GARCH models are obtained from MLE estimation on S&P500 returns as reported in Table 1.The Impact of Volatility Long Memory on Option Valuation: component GARCH versus FIGARCH

Table 1: Parameter Estimates and Model Properties

Parameter	NGARCH-GED		NGARCH-N		AGARCH-IG		AGARCH-N	
	GARCH(1,1)	Component	GARCH(1,1)	Component	GARCH(1,1)	Component	GARCH(1,1)	Component
λ	5.00E+00	2.74E+00	2.37E+00	2.48E+00	2.50E-01	9.41E-01	7.26E-01	1.06E+00
	(1.03E+00)	(1.31E+00)	(7.03E-01)	(6.10E-01)	(6.75E-03)	(3.44E-03)	(4.57E-01)	(1.67E-01)
b_1 , β	9.88E-01	9.24E-01	9.92E-01	9.08E-01	9.88E-01	8.93E-01	9.80E-01	7.47E-01
	(5.04E-03)	(5.06E-02)	(2.30E-03)	(8.87E-03)	(1.33E+00)	(4.71E-02)	(1.59E-02)	(2.22E-02)
a_1 , α	6.04E-02	3.01E-02	6.26E-02	3.70E-02	3.52E+07	5.56E+07	3.01E-06	2.13E-06
	(6.09E-03)	(8.41E-03)	(2.09E-03)	(4.05E-03)	(4.91E+06)	(8.76E+06)	(1.67E-06)	(2.78E-07)
c_1 , γ_1	5.08E-02	1.85E+00	5.92E-01	1.66E+00	2.58E-06	1.66E-06	1.01E+02	2.98E+02
	(1.12E-02)	(5.00E-01)	(4.51E-02)	(1.98E-01)	(1.16E-09)	(2.36E-07)	(6.55E+01)	(4.15E+01)
ρ		9.98E-01		9.98E-01		9.93E-01		9.92E-01
		(7.84E-04)		(4.13E-04)		(1.43E-03)		(8.41E-04)
φ		3.28E-02		3.23E-02		5.83E+07		1.77E-06
		(9.17E-03)		(2.23E-03)		(1.13E+07)		(1.28E-07)
γ_2		3.04E-01		3.10E-01		1.49E-06		7.07E+01
		(1.10E-01)		(7.44E-02)		(3.05E-07)		(8.06E+00)
ν , η	1.23E+00	1.45E+00			-5.05E-04	-3.94E-04		
	(1.57E-02)	(9.12E-03)			(1.72E-05)	(5.37E-06)		
Properties								
LogLikelihood	34,215	34,384	34,124	34,196	34,105	34,159	34,029	34,126
LR Test	0.000		0.000		0.000		0.000	
SR Persistence		0.9244		0.9080		0.8928		0.7470
LR Persistence		0.9982		0.9980		0.9933		0.9915
Variance Persistence	0.9877	0.9999	0.9920	0.9998	0.9880	0.9993	0.9800	0.9979
Variance of Variance	9.609E-11	2.087E-10	9.696E-11	2.176E-10	3.771E-11	9.594E-09	4.979E-11	1.678E-10
Leverage	-4.834E-09	-1.035E-07	-5.838E-08	-1.123E-07	-4.812E-08	-8.549E-08	-5.201E-08	-1.296E-07
Option RMSE	2.060	1.458	2.356	1.466	2.162	1.705	2.316	1.813

Notes to Table: We use daily total returns from July 1, 1962 to December 31, 1995 on the S&P500 index to estimate the GARCH models using Maximum Likelihood. Robust standard errors are calculated from the outer product of the gradient at the optimum parameter values. Variance Persistence refers to the persistence of the conditional variance in each model. For the component models, SR Persistence refers to the persistence of the short-run component and LR Persistence refers to the persistence of the long-run component. Variance of Variance refers to the unconditional variance of the conditional variance in each model. Leverage refers to the unconditional covariance between the return and the conditional variance. LogLikelihood refers to the logarithm of the likelihood at the optimal parameter values, and LR test refers to the likelihood ratio test of the component model versus the corresponding nested GARCH(1,1) model. Option RMSE refers to the dollar root mean squared option valuation error (RMSE) calculated using the risk-neutralized MLE parameters.

Table 2: 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	52.78	58.98	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	0.1448	0.1562	0.1383
1.00<S/X<1.025	0.1527	0.1459	0.1558	0.1605	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&P500 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. We use the moneyness and maturity filters used by Bakshi, Cao and Chen (1997). The implied volatilities are calculated using the Black-Scholes formula.

Table 3: Root Mean Squared Error (RMSE) over Average Call Price

Panel A: RMSE over Average Call Price for Options with Various Moneyness

	NGARCH-GED		NGARCH-N		AGARCH-IG		AGARCH-N	
	GARCH(1,1)	Component	GARCH(1,1)	Component	GARCH(1,1)	Component	GARCH(1,1)	Component
S/X<0.975	0.3236	0.2690	0.4789	0.1976	0.4690	0.3649	0.5039	0.3996
0.975<S/X<1.00	0.1391	0.1080	0.2014	0.1096	0.2025	0.1462	0.2205	0.1604
1.00<S/X<1.025	0.0965	0.0695	0.1170	0.0792	0.1070	0.0746	0.1146	0.0809
1.025<S/X<1.05	0.0779	0.0504	0.0746	0.0588	0.0608	0.0452	0.0638	0.0476
1.05<S/X<1.075	0.0643	0.0386	0.0511	0.0459	0.0380	0.0339	0.0400	0.0338
1.075<S/X	0.0368	0.0224	0.0286	0.0264	0.0202	0.0213	0.0209	0.0211
All	0.0738	0.0523	0.0844	0.0525	0.0775	0.0611	0.0830	0.0654

Panel B: RMSE over Average Call Price for Options with Various Maturities

	NGARCH-GED		NGARCH-N		AGARCH-IG		AGARCH-N	
	GARCH(1,1)	Component	GARCH(1,1)	Component	GARCH(1,1)	Component	GARCH(1,1)	Component
DTM<20	0.0301	0.0262	0.0278	0.0264	0.0268	0.0260	0.0277	0.0263
20<DTM<80	0.0567	0.0428	0.0467	0.0434	0.0510	0.0461	0.0584	0.0473
80<DTM<180	0.0776	0.0534	0.0657	0.0557	0.0726	0.0612	0.0828	0.0653
DTM>180	0.0841	0.0588	0.1177	0.0574	0.0985	0.0721	0.1006	0.0785
All	0.0738	0.0523	0.0844	0.0525	0.0775	0.0611	0.0830	0.0654

Panel C: RMSE over Average Call Price for Various Sample Years

	NGARCH-GED		NGARCH-N		AGARCH-IG		AGARCH-N	
	GARCH(1,1)	Component	GARCH(1,1)	Component	GARCH(1,1)	Component	GARCH(1,1)	Component
1990	0.1256	0.0658	0.0874	0.0725	0.0672	0.0812	0.0850	0.0803
1991	0.1115	0.0910	0.0925	0.0798	0.0660	0.0705	0.0714	0.0735
1992	0.0809	0.0510	0.0612	0.0570	0.0534	0.0469	0.0561	0.0501
1993	0.0610	0.0486	0.0835	0.0483	0.0838	0.0613	0.0899	0.0662
1994	0.0672	0.0499	0.1020	0.0519	0.0896	0.0612	0.0911	0.0685
1995	0.0465	0.0352	0.0717	0.0355	0.0731	0.0533	0.0786	0.0572
All	0.0738	0.0523	0.0844	0.0525	0.0775	0.0611	0.0830	0.0654

Notes to Table: We use the MLE estimates from Table 1 to compute the dollar root mean squared option valuation error (RMSE) divided by the average call price. In Panel A, we show the RMSEs according to moneyness bins. In Panel B, we show the RMSEs according to maturity bins. In Panel C, we show the RMSEs on a year- by-year basis.

Chapter 3
The Impact of Volatility Long Memory on Option Valuation: Component GARCH versus FIGARCH

Yintian Wang

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.

3.1 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. The first Chapter 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 Chapter 1. Please note that we refer to the fractionally integrated Heston-Nandi GARCH model as FIHN-GARCH, 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 Chapter 1 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.

3.2 The Component Heston-Nandi GARCH Model

3.2.1 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

$$
\begin{aligned}
R_{t+1} &\equiv \ln \frac{S_{t+1}}{S_t} = r + \lambda h_{t+1} + \sqrt{h_{t+1}} z_{t+1} \\
h_{t+1} &= w + b h_t + a \left(z_t - c\sqrt{h_t} \right)^2
\end{aligned}
\tag{3.42}
$$

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

$$
h_{t+1} = \sigma^2 + b\left(h_t - \sigma^2\right) + a\left(\left(z_t - c\sqrt{h_t}\right)^2 - \left(1 + c^2\sigma^2\right)\right) \tag{3.43}
$$

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

$$
\begin{aligned}
h_{t+1} &= q_{t+1} + \widetilde{\beta}(h_t - q_t) + \alpha h_t v_{1,t} \\
q_{t+1} &= \omega + \rho q_t + \varphi h_t v_{2,t}
\end{aligned}
\tag{3.44}
$$

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 Chapter 1 for more details.

3.2.2 Variance Term Structures

Following Chapter 1, 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.

$$\frac{h_{t+1:t+K}}{\sigma^2} \equiv \frac{\frac{1}{K}\sum_{k=1}^{K} E_t\left[h_{t+k}\right]}{\sigma^2} = 1 + \frac{\frac{q_{t+1}}{\sigma^2}-1}{K}\frac{1-\rho^K}{1-\rho} + \frac{\frac{h_{t+1}-q_{t+1}}{\sigma^2}}{K}\frac{1-\tilde{\beta}^K}{1-\tilde{\beta}} \tag{3.45}$$

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

$$\partial E_t\left[h_{t:t+K}\right]/\partial z_t^2 = \frac{\alpha\left(1-\gamma_1\sqrt{h_t}/z_t\right)}{K}\frac{1-\tilde{\beta}^K}{1-\tilde{\beta}} + \frac{\varphi\left(1-\gamma_2\sqrt{h_t}/z_t\right)}{K}\frac{1-\rho^K}{1-\rho} \tag{3.46}$$

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.

3.2.3 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

$$Cov_t\left(R_{t+1}, h_{t+2}\right) = -2\left(\alpha\gamma_1 + \varphi\gamma_2\right) h_{t+1} \tag{3.47}$$

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

$$Var_t\left(h_{t+2}\right) = 2\left(\alpha + \varphi\right)^2 + 4\left(\gamma_1\alpha + \gamma_2\varphi\right)^2 h_{t+1} \tag{3.48}$$

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.

3.3 An Affine FIGARCH Model

3.3.1 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

$$(z_t - \gamma_1 \sqrt{h_t})^2 (1 - \varphi_1 L) = 1 + \gamma_1^2 \omega_1 - \beta_1 + (1 - \beta_1 L) v_t \tag{3.49}$$

where $v_t = \left(z_t - \gamma_1 \sqrt{h_t} \right)^2 - (1 + \gamma_1^2 h_t)$ 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 (3.49) 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

$$(z_t - \gamma_1 \sqrt{h_t})^2 (1 - \varphi_1 L) (1 - L)^d = 1 + \gamma_1^2 \omega_1 - \beta_1 + (1 - \beta_1 L) v_t \tag{3.50}$$

An alternative representation is

$$\begin{aligned} (1 - \beta_1 L) \left(1 + \gamma_1^2 h_t \right) &= \left(1 + \gamma_1^2 \omega_1 - \beta_1 \right) + \dots \\ &+ \left(1 - \beta_1 L - (1 - \varphi_1 L) (1 - L)^d \right) (z_t - \gamma_1 \sqrt{h_t})^2 \end{aligned} \tag{3.51}$$

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,

$$
\begin{aligned}
h_t &= \frac{\omega_1}{1-\beta_1} + \left(\frac{1 - \frac{(1-\varphi_1 L)(1-L)^d}{(1-\beta_1 L)}}{\gamma_1^2} \right) (z_t - \gamma_1 \sqrt{h_t})^2 \\
&= \varpi + \frac{\lambda(L)}{\gamma_1^2} (z_t - \gamma_1 \sqrt{h_t})^2 \qquad (3.52)
\end{aligned}
$$

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

$$
\begin{aligned}
\lambda_1 &= (\varphi_1 - \beta_1 + d) = \alpha_1 \gamma_1^2 + d \\
\lambda_k &= \beta_1 \lambda_{k-1} + \left((k-1-d) k^{-1} - \varphi_1 \right) \delta_{d,k-1} \text{ for } k \geq 2 \qquad (3.53)
\end{aligned}
$$

where

$$
\begin{aligned}
\delta_{d.1} &= d \qquad (3.54) \\
\delta_{d.k} &= \delta_{d.k-1} k^{-1} (k-1-d) \quad \text{for } k \geq 2
\end{aligned}
$$

$\left(1 - \frac{(1-\varphi_1 L)(1-L)^d}{(1-\beta_1 L)}\right)$ 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 (3.52) 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.

3.3.2 Variance Term Structures

We again define the variance term structure

$$\frac{h_{t+1:t+k}}{\sigma^2} = \frac{1}{K}\frac{1}{\sigma^2}\sum_{k=1}^{K} E_t\left(h_{t+k}\right) \tag{3.55}$$

where

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

and the impulse response functions are

$$\begin{aligned} \partial E_t\left[h_{t:t+K}\right]/\partial z_t^2 &= \sum_{i=1}^{K} \frac{\partial E_t h_{t+i}}{\partial z_t^2} \\ \frac{\partial E_t h_{t+k}}{\partial z_t^2} &= \sum_{i=1}^{k-1} \frac{\lambda_i}{\gamma_1^2} \frac{\partial E_t h_{t+k-i}}{\partial z_t^2} + \frac{\lambda_k}{\gamma_1^2}\left(1-\gamma_1 \frac{\sqrt{h_t}}{z_t}\right) \\ \frac{\partial E_t h_{t+1}}{\partial z_t^2} &= \frac{\lambda_1}{\gamma_1^2}\left(1-\gamma_1 \frac{\sqrt{h_t}}{z_t}\right) \end{aligned} \tag{3.56}$$

3.3.3 Conditional Leverage and Variance of Variance

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

$$\begin{aligned} Var_t(h_{t+2}) &= E_t\left[h_{t+2} - E_t\left[h_{t+2}\right]\right]^2 \\ &= \left(2 + 4\gamma_1^2 h_{t+1}\right)\frac{\lambda_1^2}{\gamma_1^4} \end{aligned} \tag{3.57}$$

$$\begin{aligned} Cov_t(\ln\left(S_{t+1}\right), h_{t+2}) &= E_t\left[\left(\ln\left(S_{t+1}\right) - E_t\left[\ln\left(S_{t+1}\right)\right]\right)\left(h_{t+2} - E_t\left[h_{t+2}\right]\right)\right] \\ &= E_t\left[\sqrt{h_{t+1}} z_{t+1} \frac{\lambda_1}{\gamma_1^2}\left(z_{t+1}^2 - 2\gamma_1 z_{t+1}\sqrt{h_{t+1}} - 1\right)\right] \\ &= -2\frac{\lambda_1}{\gamma_1} h_{t+1} \end{aligned} \tag{3.58}$$

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.

3.3.4 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.

$$Corr_t(\varepsilon_{t+1}^2, \varepsilon_{t+k}^2) = \frac{Cov_t(\varepsilon_{t+1}^2, h_{t+k})}{\sqrt{Var_t\left[\varepsilon_{t+1}^2\right]}\sqrt{Var_t\left[\varepsilon_{t+k}^2\right]}} \tag{3.59}$$

where

$$\begin{aligned} Cov_t(\varepsilon_{t+1}^2, h_{t+k}) &= \frac{2\lambda_{k-1}}{\gamma_1^2} h_{t+1} \\ Var_t\left[\varepsilon_{t+1}^2\right] &= 2h_{t+1}^2 \\ Var_t\left[\varepsilon_{t+k}\right] &= 2\begin{pmatrix} \Delta^2 + 2\Delta \sum_{i=1}^{k-1} \frac{\lambda_i}{\gamma_1^2}\left(1 + \gamma_1^2 E_t h_{t+k-i}\right) + \ldots \\ \sum_{i=1}^{k-1} \frac{\lambda_i^2}{\gamma_1^4}\left(4 + 8\gamma_1^2 E_t h_{t+k-i} + \gamma_1^4 E_t h_{t+k-1}^2\right) \end{pmatrix} \end{aligned} \tag{3.60}$$

and

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

3.3.5 Risk Neutralization and Option Valuation

As in Chapter 1, we assume Duan (1995)'s Locally Risk-Neutral Valuation Relationship assumption. In the risk-neutral world, the asset price S_t follows

$$\begin{aligned} \ln(S_{t+1}) &= \ln(S_t) + r - 0.5h_{t+1} + \sqrt{h_{t+1}} z_{t+1}^* \\ h_t &= \varpi + \frac{\lambda(L)}{\gamma_1^2}(z_t^* - \gamma_1^* \sqrt{h_t})^2 \end{aligned} \tag{3.61}$$

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

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

$$
\begin{aligned}
&= \frac{1}{2}S_t + \frac{e^{-r(T-t)}}{\pi}\int_0^{\infty}\mathrm{Re}\left[\frac{K^{-i\phi}f^*(t,T;i\phi+1)}{i\phi}\right]d\phi... \\
&\quad -Ke^{-r(T-t)}\left(\frac{1}{2}+\frac{1}{\pi}\int_0^{\infty}\mathrm{Re}\left[\frac{K^{-i\phi}f^*(t,T;i\phi)}{i\phi}\right]d\phi\right)
\end{aligned} \tag{3.62}
$$

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

$$
\begin{aligned}
f_t &= E_t \exp\left(\phi \ln\left(S_T\right)\right) \\
&= \exp\left(\phi x_t + A_t + B_t h_{t+1} + \Lambda_t(L)\left(z_{t+1} - \gamma_1\sqrt{h_{t+1}}\right)^2\right)
\end{aligned} \tag{3.63}
$$

where $x_t = \ln\left(S_t\right)$. 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

$$
\begin{aligned}
A_t &= \phi r + A_{t+1} + B_{t+1}\overline{\omega} - \frac{1}{2}\ln\left(1 - 2\left(B_{t+1}\overline{\lambda}_1 + \Lambda_{t+1,1}\right)\right) \\
B_t &= -0.5 + \gamma_1 + \frac{\left(B_{t+1}\overline{\lambda}_1 + \Lambda_{t+1,1}\right)\gamma_1^2 + \frac{\phi^2}{2} - \phi\gamma_1}{1 - 2\left(B_{t+1}\overline{\lambda}_1 + \Lambda_{t+1,1}\right)} \\
\Lambda_{t,1} &= B_{t+1}\overline{\lambda}_2 + \Lambda_{t+1,2} \\
\Lambda_{t,2} &= B_{t+1}\overline{\lambda}_3 + \Lambda_{t+1,3} \\
\Lambda_{t,n} &= B_{t+1}\overline{\lambda}_{n+1} + \Lambda_{t+1,n+1}
\end{aligned} \tag{3.64}
$$

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.

3.4 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.[27] 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.[28] 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.

[27] 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.

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

3.4.1 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 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 (3.48). 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 (3.57). To allow a comparison of the unconditional leverage for models, we report the moments in (3.47) and (3.58) 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.

3.4.2 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, $Var_t\left(h_{t+2}\right)$ 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, $Cov_t\left(R_{t+1}, h_{t+2}\right)$ 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 (3.46). 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.[29] From an economic point of view, the debt-equity ratio may be hard to adjust in the short run, but

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

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, fractionally 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. Chapter 1 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 (3.58) and (3.57) 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.

3.4.3 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 (3.61). 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 (3.62).

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

$$RMSE = \sqrt{\frac{1}{N^T} \sum_{i,t} \left(C_{i,t}^{MKT} - C_{i,t}^{GARCH} \right)^2} \tag{3.65}$$

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 3.45. 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.[30] 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 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

[30] 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.

$d = 0.35$, 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.

3.5 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 Chapter 1, 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 FIHN-GARCH 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.

3.6 Appendix

3.6.1 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[31]

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

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

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

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

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

[31] 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.

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

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

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

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

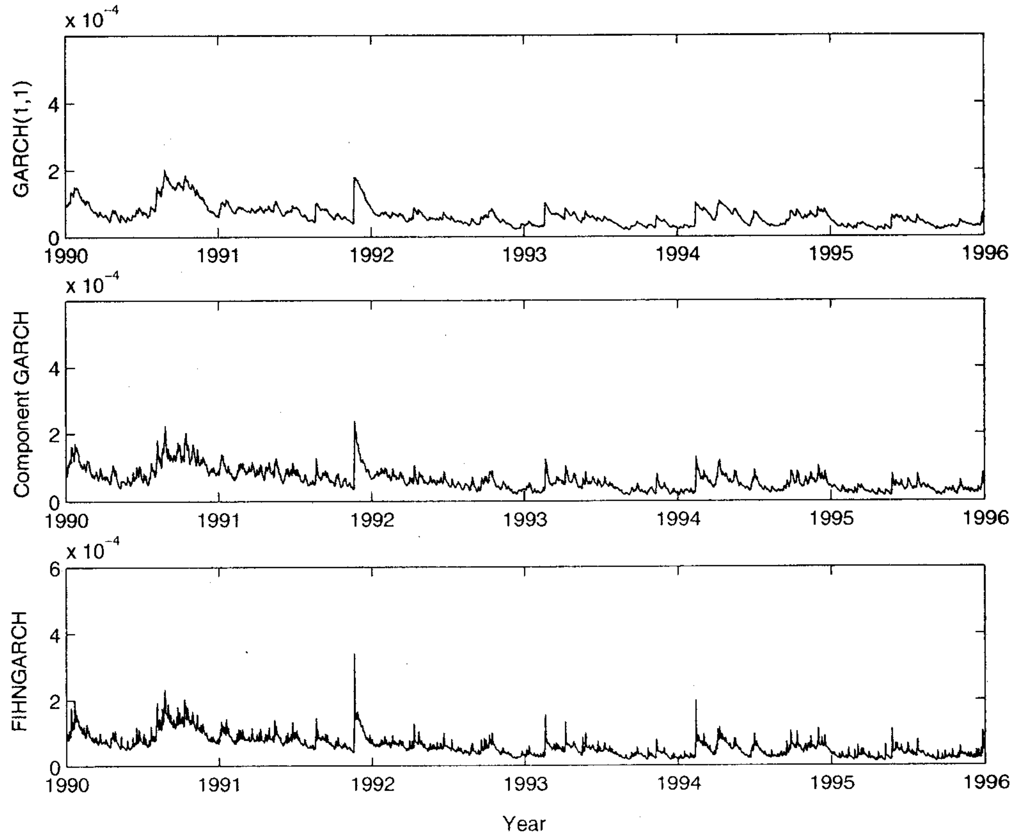
$$
B_t = \phi\left(-0.5 + \gamma_1\right) + \frac{\left(B_{t+1}\overline{\lambda}_1 + \Lambda_{t+1,1}\right)\gamma_1^2 + \frac{\phi^2}{2} - \phi\gamma_1}{1 - 2\left(B_{t+1}\overline{\lambda}_1 + \Lambda_{t+1,1}\right)}
$$

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

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

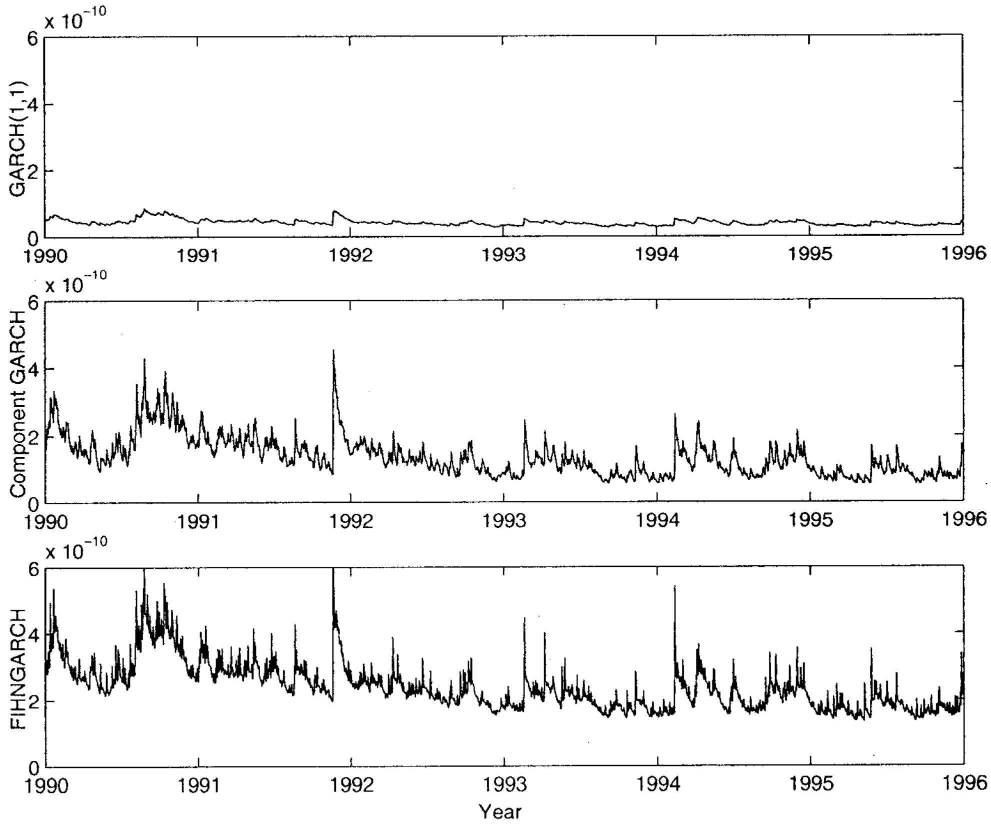
3.7 Figures and Tables

Figure 1. Conditional Variance



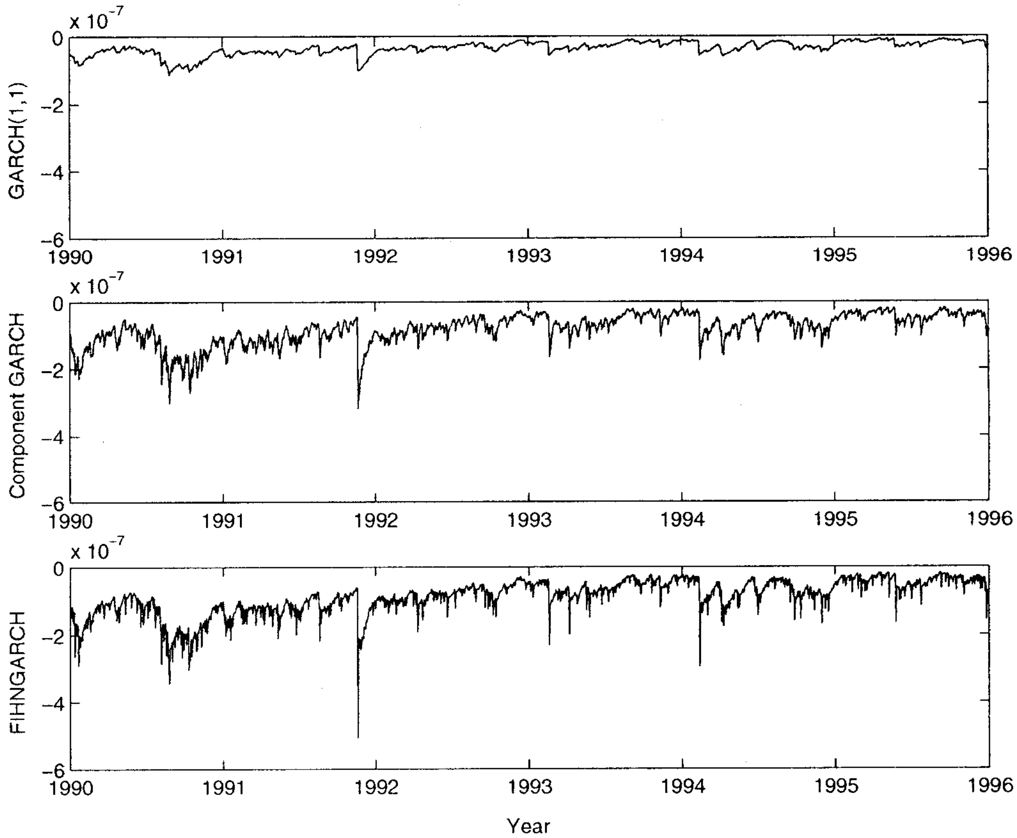
Note to Figure: In Figure 1, we plot the variance paths from the GARCH(1,1), the component GARCH, and the FIHNGARCH model. The parameters are obtained from the MLE estimates on returns in Table 2.

Figure 2. Conditional Variance of Variance



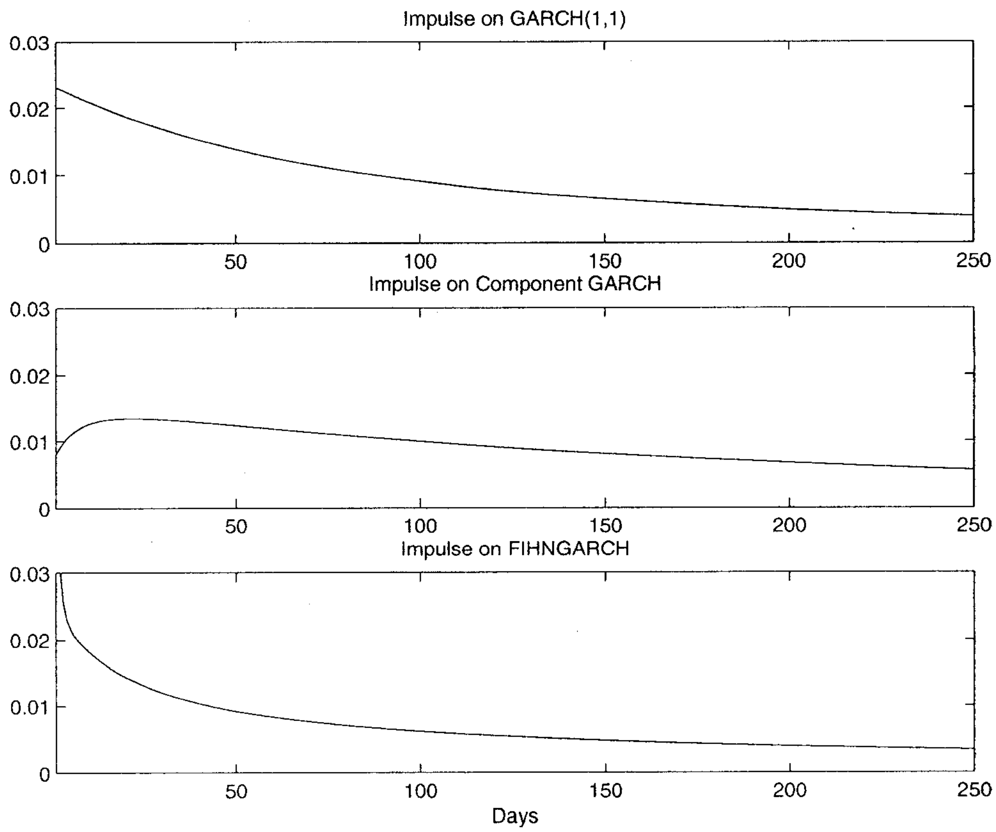
Note to Figure: In Figure 2, we plot the conditional variance of next day's variance as implied by the GARCH(1,1), the component GARCH and the FIHNGARCH models. The scales are identical across panels to facilitate comparison across models. The parameters are obtained from the MLE estimates on returns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

Figure 3. Conditional Leverage Effect



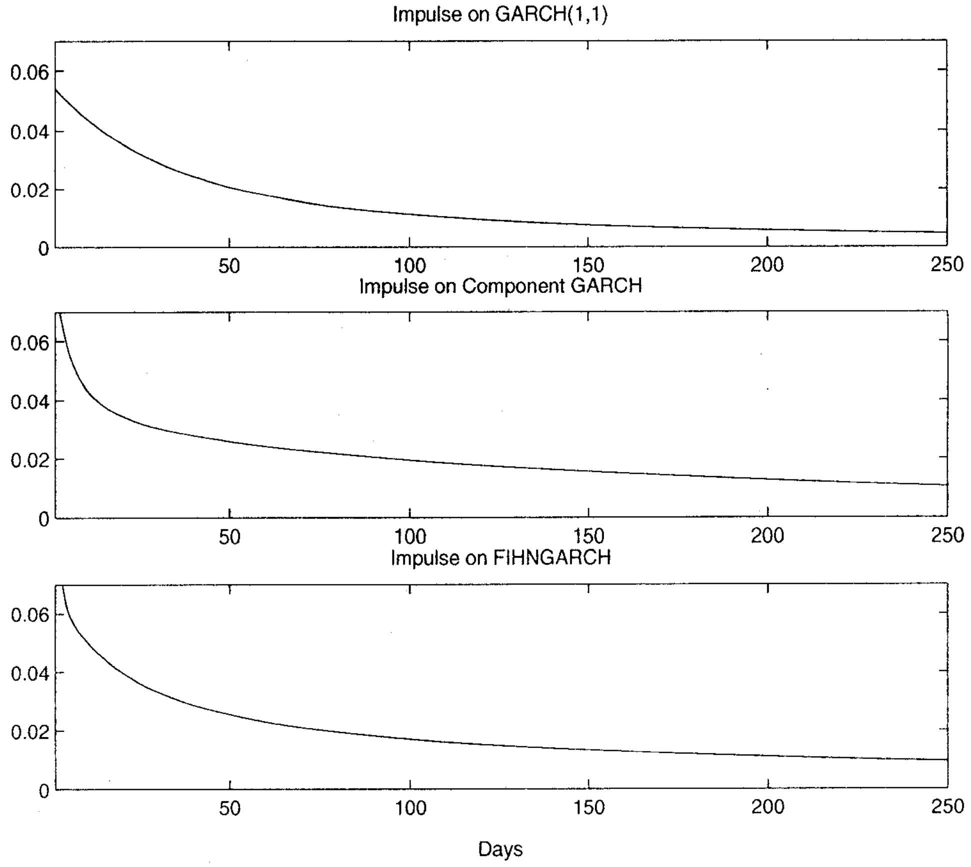
Note to Figure: In Figure 3, we plot the conditional leverage between the return and the next-day variance as implied by the component GARCH, and FIHNGARCH models and refer to it as conditional leverage. The scales are identical across panels to facilitate comparison across models. The parameters are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

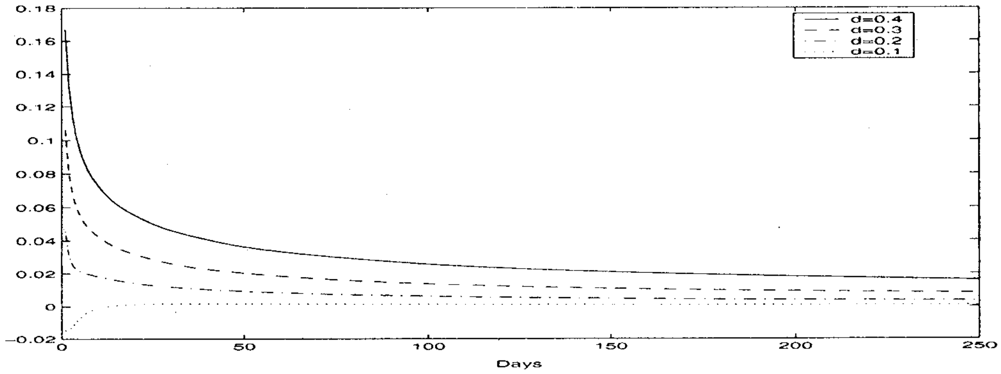
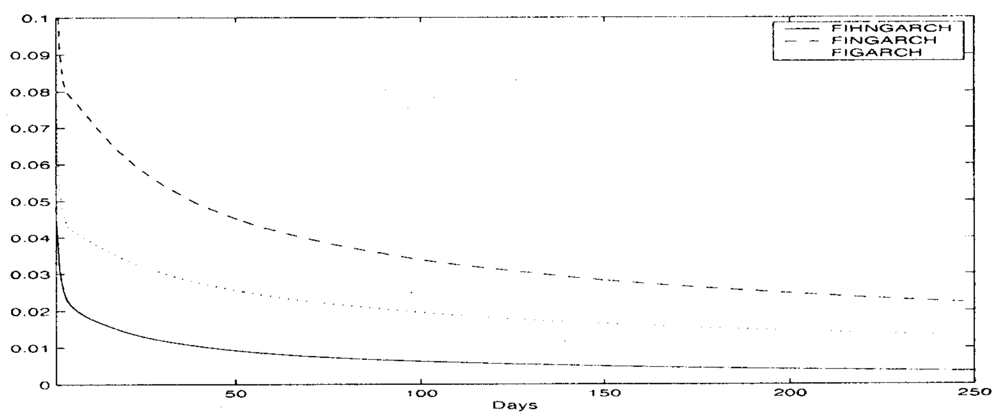
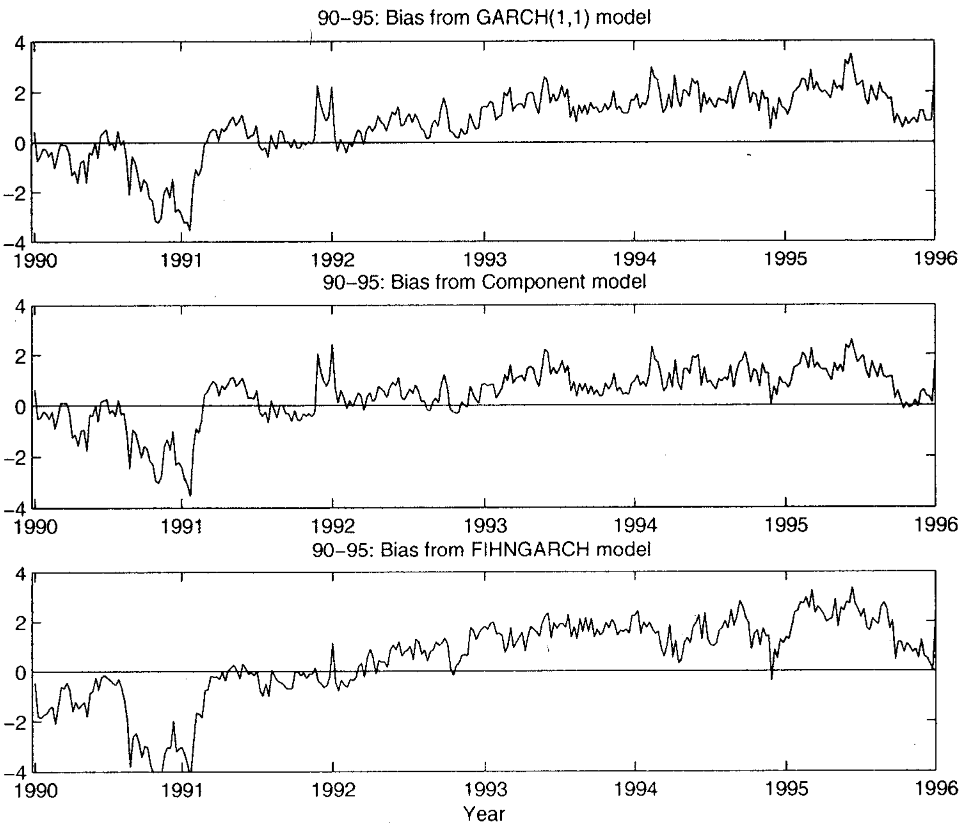


Figure 6. Term Structure Impulse Response to a Positive Return Shock ($z_t = 2$) for Different Models



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

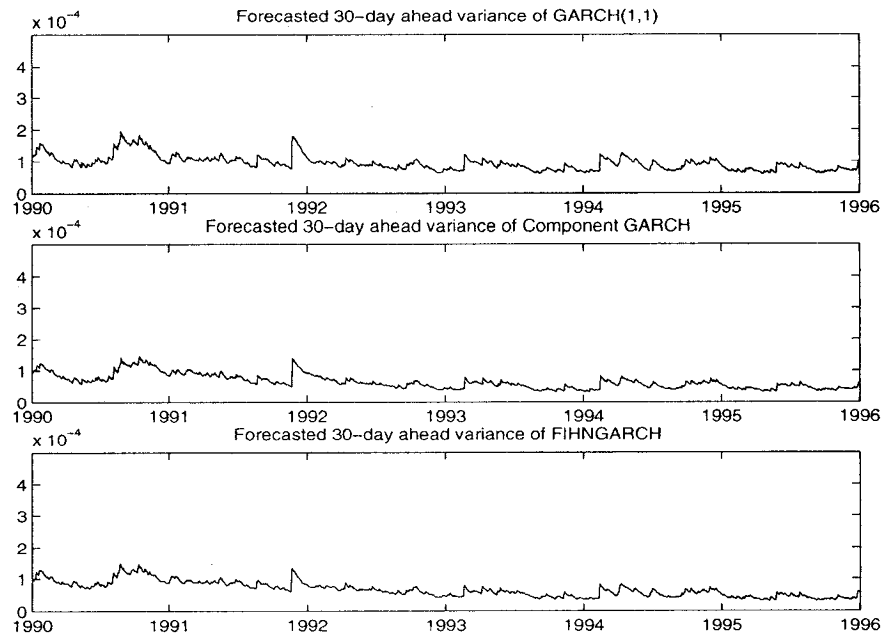
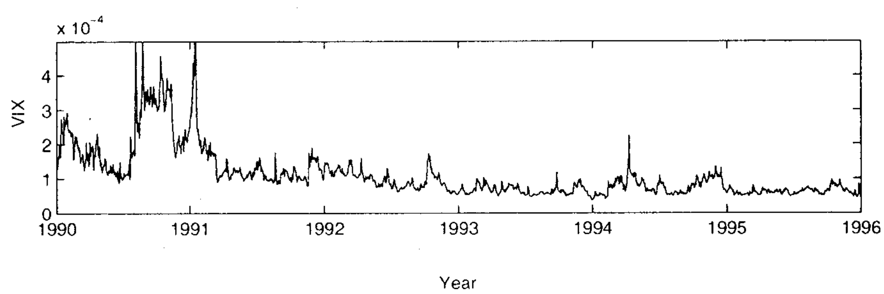
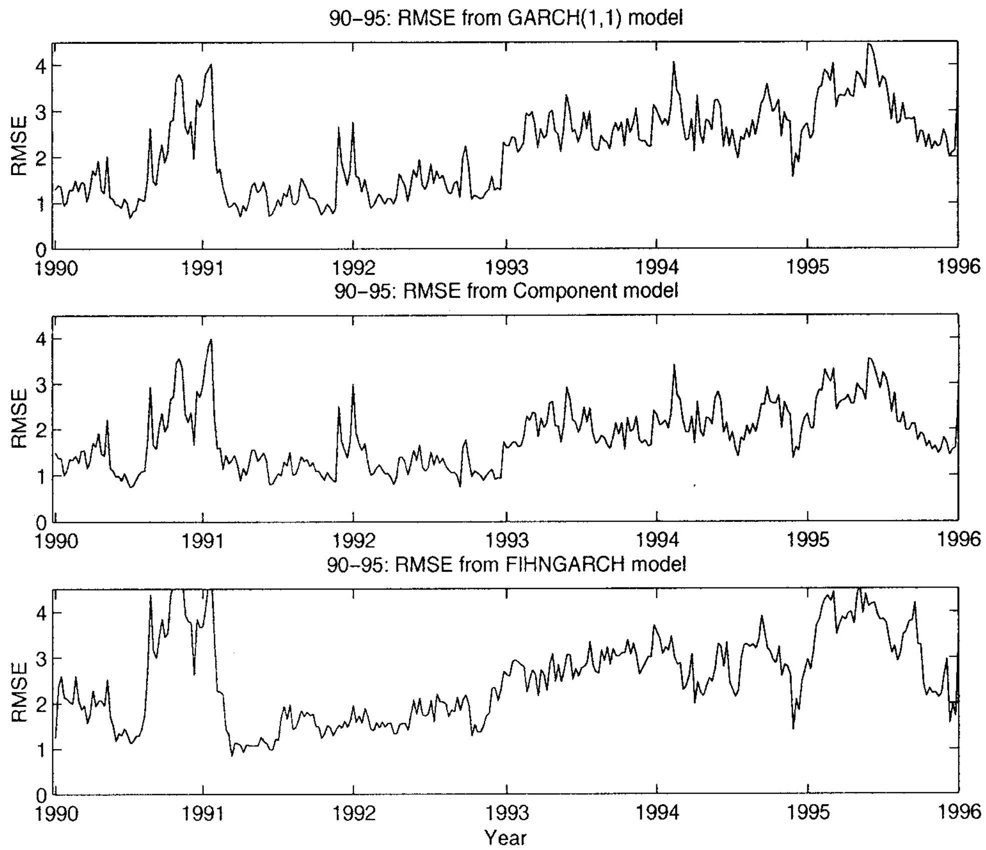


Figure 8b. VIX



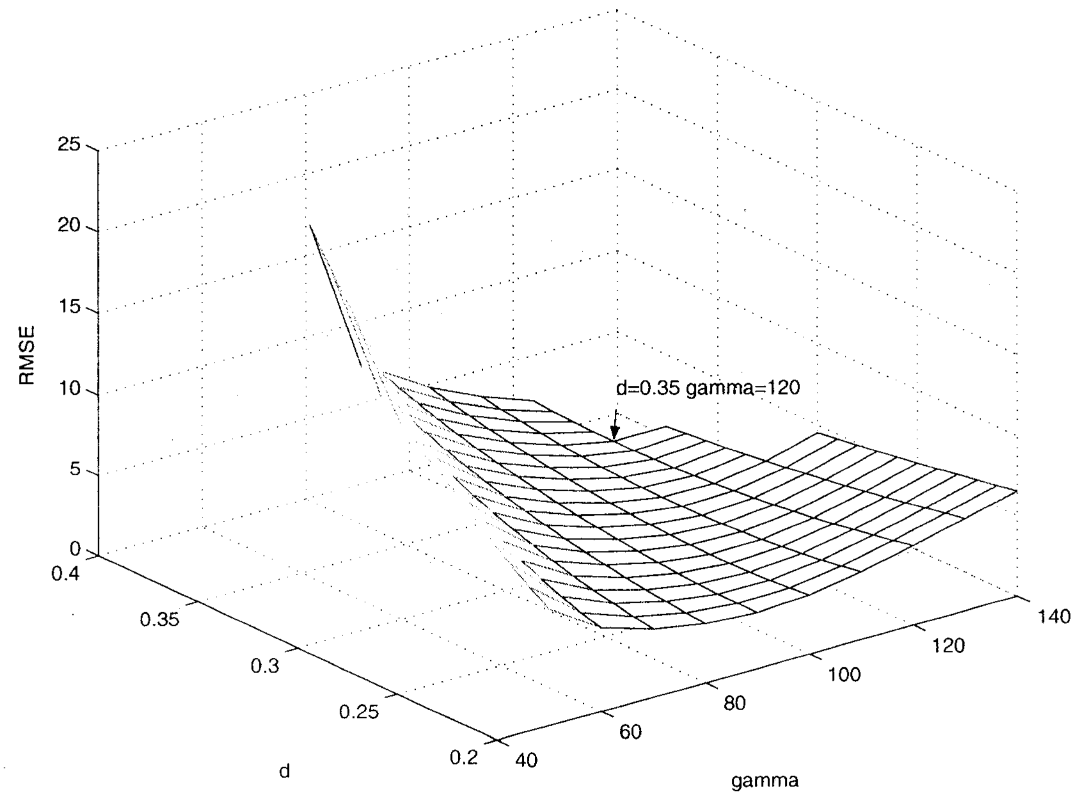
Note to Figure: In panel a, we plot the cumulative 30-day ahead forecasted variance paths from the GARCH(1,1), the component GARCH and the FIHNGARCH model. The parameters are obtained from the MLE estimates on returns in Table 2. In Panel b, we plot the VIX index from the CBOE for comparison. The scales 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	$\tilde{\beta}$	0.766	0.163	β_1	0.664	3.414E-07
a	3.210E-06	2.810E-06	α	1.770E-06	1.110E-06	d	0.203	5.778E-04
c	88.192	15.623	γ_1	312.880	108.430	γ_1	101.594	1.430E-02
λ	1.815	0.224	γ_2	59.043	30.196	φ_1	0.536	7.147E-04
		0.054	φ	0.000	0.000	η	1.945	3.358E-03
			ρ	0.989	0.002			
			λ	1.809	0.526			
Annual Vol	0.147		Annual Vol	0.145		Annual Vol	0.145	
Vol of Var	4.574E-06		Vol of Var	1.329E-05		Vol of Var	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. Vol of Var refers to the unconditional standard deviation 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 computed at the MLE estimates in Panel A of Table 2.

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

FIGARCH			FINGARCH		
Parameter	Estimate	Std. Error	Parameter	Estimate	Std. Error
β	0.673	0.100	β	0.720	0.018
d	0.442	0.056	d	0.481	0.018
φ	0.349	0.090	γ	0.585	0.028
l	4.961	0.979	φ	0.380	0.021
ϖ	6.491E-06	8.192E-07	λ	4.352	0.163
			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.

Chapter 4
Conclusion and Future Work

This dissertation is in the form of three essays on the topic of component GARCH models. The unifying feature that permeates the entire thesis is the focus on investigating European option evaluation with component GARCH models.

The dissertation presents a new option valuation model based on the work by Engle and Lee (1999) and Heston and Nandi (2000). The empirical performance of the new variance component model is significantly better than that of the benchmark GARCH$(1, 1)$ model, in-sample as well as out-of-sample, and regardless of the information used in estimation. This is an important finding because the literature has demonstrated that it is difficult to find empirical models that improve on the GARCH$(1, 1)$ model or the Heston (1993) model. The component GARCH model is also compared to a GARCH$(1, 1)$-Jump model, which combines conditional heteroskedasticity with Poisson-normal jumps. The GARCH$(1, 1)$-Jump model achieves a better statistical fit than the component model in-sample, but the component model performs far better when using the parameter estimates to fit options.

Two extensions have been made to this novel component GARCH model to allow non-normal innovations as well as non-affine structures. One extended model allows for GED innovations to the variance dynamic. The second model is characterized by a conditional inverse Gaussian innovation and by affine variance dynamics. A closed-form option valuation formula is derived for this model. The two new non-normal component models are compared with the corresponding special cases with normal innovations, and the resulting four component models are compared with the GARCH(1,1) models which they nest. All eight models are estimated using MLE on a long time series of S&P500 returns. The likelihood criterion strongly favors the component models in all cases, and it also favors non-normal innovations. When the models' parameters are used for option valuation, there is very strong support for the component variance specifications. The support for non-normal innovations and for the non-affine structure is less strong.

Overall, an important aspect of the component GARCH model's improved performance is that its richer parameterization allows for improved joint modeling of long-maturity and short-maturity options. The model captures the stylized fact that shocks to current conditional volatility impact on the conditional variance forecast up to a year in the future, which results in a very different implied volatility term structure for at-the-money options. The component model also results in a different path for spot volatility compared to the GARCH(1, 1) model, but in the moneyness dimension the differences with the GARCH(1, 1) model seem relatively less important. The component model is also characterized by term structures of skewness and kurtosis that are very different from those of the GARCH(1, 1) model.

In the dissertation, the affine component GARCH model is also compared with a fractionally integrated affine GARCH model that allows for volatility long memory. The dissertation investigates the models through S&P 500 index returns and cross-sectional European options data. The component GARCH model is slightly better than the FIGARCH in fitting S&P 500 returns, and significantly outperforms FIGARCH in fitting option prices. In return, the FIGARCH model dominates the GARCH(1,1) in terms of log-likelihood function while yielding higher RMSE of option pricing than does the GARCH(1,1) model. This superiority is mainly due to the shorter memory of the FIGARCH 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 have been investigated in previous literature, this finding is novel.

Given the success of the proposed volatility component models, a number of further extensions to this work are warranted. First, the empirical performance of the model should of course be validated using other datasets. In particular, it would be interesting to test the model using LEAPS data, because the model may excel at modeling long-maturity LEAPS options. Second, it remains to be seen if the differences in performance between the models are confirmed when using model parameters estimated from option prices, or when using an integrated analysis that uses option prices as well as underlying returns. Third, it would be useful to reconcile the relationship between the superior option valuation performance of the component models and the less than superior performance of GARCH(2,2) models in traditional volatility forecasting studies. Comparing the density forecasts implied by the different models could be an avenue to explore. Finally, looking forward, it would be interesting to compare the range of discrete-time GARCH models considered here with the continuous-time stochastic volatility models that are popular in the finance literature.

References

Adrian, T., Rosenberg, J., 2005. "Stock returns and volatility: pricing the long-run and short-run components of market risk," Unpublished working paper, the Federal Reserve Bank of New York.

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

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

Andersen, T., Bollerslev, T., Diebold, F., Labys, P., 2003. "Modeling and forecasting realized volatility," Econometrica 71, 529-626.

Amin, K., Ng, V., 1993. "ARCH processes and option valuation," Unpublished working paper, University of Michigan.

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

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

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

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

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

Bates, D., 2006. "Maximum likelihood estimation of latent affine processes," Review of Financial Studies 19, 909-965.

Benzoni, L., 1998. "Pricing options under stochastic volatility: an econometric analysis," Unpublished working paper, University of Minnesota.

Beveridge, S., Nelson, C., 1981. "A new approach to decomposition of economic time series into permanent and transitory components with particular attention to measurement of the business cycle," Journal of Monetary Economics 7, 151-174.

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.

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

Bollerslev, T., 1986. "Generalized autoregressive conditional heteroskedasticity," Journal of Econometrics 31, 307-327.

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

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

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

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

Breidt, F. J., N. Crato, and P. de Lima, 1998, "The detection and estimation of long memory in stochasticvolatility," Journal of Econometrics, 83, 325–348.

Brennan, M., 1979. "The pricing of contingent claims in discrete-time Models," Journal of Finance 34, 53-68.

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

Camara, A., 2003. "A generalization of the Brennan-Rubinstein approach for the pricing of derivatives," Journal of Finance 58, 805-819.

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

Chacko, G., Viceira, L., 2003. "Spectral GMM estimation of continuous-time processes," Journal of Econometrics 116, 259-292.

Chernov, M., Ghysels, E., 2000. "A study towards a unified approach to the joint estimation of objective and risk neutral measures for the purpose of option valuation," Journal of Financial Economics 56, 407-458.

Chernov, M., Gallant, R., Ghysels, E., Tauchen, G., 2003. "Alternative models for stock price dynamics," Journal of Econometrics 116, 225-257.

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

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

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.

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

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

Das, S., Sundaram, R., 1999. "Of smiles and smirks: A term structure perspective," Journal of Financial and Quantitative Analysis 34, 211-239.

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.

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

Duan, J.-C. (1999), "Conditionally Fat-Tailed Distributions and the Volatility Smile in Options," Manuscript, Hong Kong University of Science and Technology.

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.

Duan, J.-C., Ritchken, P., Sun, Z., 2006. "Approximating GARCH-jump models, jump-diffusion processes, and option pricing," Mathematical Finance 16, 21-52.

Duffee, G., 1999. "Estimating the price of default risk," Review of Financial Studies 12, 197-226.

Duffie, D., Singleton, K., 1999. "Modeling term structures of defaultable bonds," Review of Financial Studies 12, 687-720.

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

Dumas, B., Fleming, J., Whaley, R., 1998. "Implied volatility functions: empirical tests," Journal of Finance 53, 2059-2106.

Engle, R., 1982. "Autoregressive conditional heteroskedasticity with estimates of the variance of UK inflation," Econometrica 50, 987-1008.

Engle, R., Mustafa, C., 1992. "Implied ARCH models from options prices," Journal of Econometrics 52, 289-311.

Engle, R., Lee, G., 1999. "A permanent and transitory component model of stock return volatility," in Cointegration, Causality, and Forecasting, edited by R. Engle and H. White, Oxford University Press, New York, pp. 475-497.

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

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

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

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

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

Hamilton, J. (1994), "Time Series Analysis," Princeton University Press.

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.

Hentschel, L., 1995. "All in the family: nesting symmetric and asymmetric GARCH models," Journal of Financial Economics 39, 71-104.

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.

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

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

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

Hull, J., White, A., 1987. "The pricing of options with stochastic volatilities," Journal of Finance 42, 281-300.

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

Maheu, J., 2002. "Can GARCH models capture the long-range dependence in financial market volatility?" Unpublished working paper, University of Toronto.

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

Melino, A., Turnbull, S., 1990. "Pricing foreign currency options with stochastic volatility," Journal of Econometrics 45, 239-265.

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.

Nelson, D. B. (1991). "Conditional Heteroskedasticity in Asset Returns: A New Approach," Econometrica, 59, 347-370.

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

Pearson, N., Sun, T., 1994. "Exploiting the conditional density in estimating the term structure: an application to the Cox, Ingersoll, and Ross model," Journal of Finance 49, 1279-1304.

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

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

Rubinstein, M., 1976. "The valuation of uncertain income streams and the pricing of options," Bell Journal of Economics 7, 407-425.

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

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

Summers, L., 1986. "Does the stock market rationally reflect fundamental values?" Journal of Finance 41, 591-600.

Taylor, S., Xu, X., 1994. "The term structure of volatility implied by foreign exchange options," Journal of Financial and Quantitative Analysis 29, 57-74.

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

Wiggins, J., 1987. "Option values under stochastic volatility: theory and empirical evidence," Journal of Financial Economics 19, 351-372.