

# OPTION PRICING WITH ASYMMETRIC HETEROSKEDASTIC NORMAL MIXTURE MODELS

Jeroen V.K. Rombouts<sup>1</sup> and Lars Stentoft<sup>2</sup>

February 18, 2013

## Abstract

This paper uses asymmetric heteroskedastic normal mixture models to fit return data and to price options. The models can be estimated straightforwardly by maximum likelihood, have high statistical fit when used on S&P 500 index return data, and allow for substantial negative skewness and time varying higher order moments of the risk neutral distribution. When forecasting out-of-sample a large set of index options between 1996 and 2009, substantial improvements are found compared to several benchmark models in terms of dollar losses and the ability to explain the smirk in implied volatilities. Overall, the dollar root mean squared error of the best performing benchmark component model is 39% larger than for the mixture model. When considering the recent financial crisis this difference increases to 69%.

**Keywords:** Asymmetric heteroskedastic models, finite mixture models, option pricing, out-of-sample prediction, statistical fit

**JEL Classification:** C11; C15; C22; G13

---

<sup>1</sup>Institute of Applied Economics at HEC Montréal, CIRANO, CIRPEE, Université catholique de Louvain (CORE), 3000 Cote Sainte Catherine, H3T 2A7 Montréal (QC), Canada. Email: jeroen.rombouts@hec.ca. Phone: (+1) 514 340-6466

<sup>2</sup>Department of Finance at HEC Montréal, CIRANO, CIRPEE, CREATES, 3000 Cote Sainte Catherine, H3T 2A7 Montréal (QC), Canada. Email: lars.stentoft@hec.ca. Phone: (+1) 514 340-6671.

The authors thank Eric Renault, Kris Jacobs, as well as seminar participants at Center (Tilburg University), CORE (Catholic University of Louvain), CREATES (University of Aarhus), and the Economics department at Boston University for comments. They also thank participants and various discussants at the 2009 FMA, NFA, and CIRPEE annual meetings as well as the Conference on Periodicity, Non-stationarity, and Forecasting of Economic and Financial Time Series and the Financial Econometrics Conference. The second author thanks CREATES (Center for Research in Econometric Analysis of Time Series, funded by the Danish National Research Foundation) for financial support.

# 1 Introduction

While the Black-Scholes-Merton (BSM) model of Black and Scholes (1973) and Merton (1973) remains a standard tool for practitioners, when it comes to option pricing many attempts have been made on extending the model to obtain a better fit to actually observed prices. Several studies show that the BSM model, which assumes constant volatility and Gaussian returns, severely underprices out of the money options, particularly those with short maturity. In terms of implied volatilities, this leads to the well-known smile or smirk across moneyness, which is found to be particularly pronounced for index options. Intuitively, such findings can be the result of either non constant volatility or non Gaussian returns, or a combination of the two. Thus, extensions to the BSM model have been developed focusing on relaxing these assumptions. In this paper, we provide a new approach that accommodates flexible conditional return distributions with time varying volatility.

Several studies have examined models which allow for more flexible specifications of the volatility process compared to the BSM model. In particular, the stochastic volatility (SV) models have successfully been applied, see Hull and White (1987), Johnson and Shanno (1987), Scott (1987), Stein and Stein (1991), Wiggins (1987), Amin and Ng (1993), and Heston (1993). When comparing these models to the BSM model, empirically support is found for the stochastic volatility specification. This is documented in Bakshi, Cao, and Chen (1997), Bates (2000), and Nandi (1996) when considering options on the S&P 500 index or the index futures. In addition to the SV models, the generalized autoregressive conditional heteroskedasticity (GARCH) framework has been used for option pricing using the model of Duan (1995). This model has also been used with success for empirical option pricing by, among others, Christoffersen and Jacobs (2004), Heston and Nandi (2000), and Hsieh and Ritchken (2005). A recent contribution is Christoffersen, Jacobs, Ornthanalai, and Wang (2008) where the volatility is allowed to have both short and long run components. These studies have all analyzed options on the S&P 500 index and found that

the GARCH models diminish the mispricings found when using the constant volatility BSM model. Stentoft (2005) documents the same findings using a sample of individual stock options and options on the S&P 100 index.

However, while the improvements in option pricing performance of SV and GARCH models are important, mispricings still exist when comparing these models to actual option data as documented by e.g. Nandi (1996). This has led to the development of option pricing models which rely on alternative conditional distributions in addition to having non constant volatility. In the SV literature, jumps have been introduced in the return and volatility processes. Classical references are, among others, Bakshi, Cao, and Chen (1997), Bakshi, Cao, and Chen (2000), Bates (1991), Bates (2000), Pan (2002), and Eraker (2004). When examining the empirical performance of these models, most of the above papers find support for the existence of jumps again when looking at S&P 500 index options. One exception is Bakshi, Cao, and Chen (2000) which, however, analyzes long term options and finds only small gains of allowing for jumps. In addition to the jump-diffusion processes, models based on infinite activity Lévy processes have been proposed in Barndorff-Nielsen and Shephard (2001) and Carr, Geman, Madan, and Yor (2003), and on time changed Lévy processes in Carr and Wu (2004) with applications to currency options in Carr and Wu (2007) and Bakshi, Carr, and Wu (2008).

Within the GARCH literature extensions include the use of distributions which are either skewed or leptokurtic, or both, as is done in Duan (1999) and Christoffersen, Elkamhi, Feunou, and Jacobs (2010). For empirical applications of this framework, see Christoffersen, Heston, and Jacobs (2006), Christoffersen, Jacobs, Dorion, and Wang (2010) and Stentoft (2008). Although Christoffersen, Jacobs, Dorion, and Wang (2010) find little improvement for the non-normal models Christoffersen, Heston, and Jacobs (2006) observe that allowing for non-normal innovations is important when pricing out of the money put options on the S&P 500 index. Moreover, Stentoft (2008) documents improvements for both call and put options in terms of fitting the smile across moneyness for a sample of individual stock options as well as for options on the S&P 100 index. In addition to

models with non Gaussian innovations, GARCH models with jumps have been developed by Duan, Ritchken, and Sun (2006) and Christoffersen, Jacobs, and Ornthanalai (2008). The latter of these papers also examines the empirical performance and shows that jumps are important empirically when pricing S&P 500 index options.

The discussion above documents that important advances have been made in the empirical option pricing literature when it comes to extending the BSM model. However, it is also clear that there is still much room for improvements as discussed in the reviews by Bates (2003) and Garcia, Gysels, and Renault (2010). In particular, the existing research has shown that there are large differences between the conditional distribution of the underlying asset and the distribution implied from option pricing. One difference is that the volatility implied by at the money options is significantly different from that observed over the life of the option. However, much more importantly is the finding that the implied volatility curve, that is the implied volatility plotted against moneyness, is not only asymmetric but also changes through time. The first of these findings implies substantial negative skewness, more than is often found in the underlying process, whereas the latter indicates that moments of higher order are time varying.

In the present paper, we price options using asymmetric heteroskedastic normal mixture models, and we argue that our proposed model can address the above points. In particular, the type of finite mixture model we use is flexible enough to approximate arbitrarily well any kind of conditional distribution, for example highly skewed and leptokurtic, and to allow for stochastic volatility of the returns on the underlying asset of the option contract. By proposing a NGARCH in mean mixture model we suggest a feasible way for option pricing within this general framework and we derive the appropriate risk neutral dynamics. In our application, we show that our model allows for significant negative skewness and time varying higher order moments.

We use our model to forecast out-of-sample prices of a large sample of options on the S&P 500 index from 1996 through 2009 and compute dollar losses and implied standard deviation losses. Our results show that the added flexibility of the finite mixture model in terms of both skewness and

excess kurtosis provides important improvements for predicted option prices. Overall, the dollar root mean squared error of the best performing benchmark component model is 39% larger than for the mixture model. In particular, improvements are found for all categories of maturity and moneyness. Next, using implied volatility losses, we show that the heteroskedastic normal mixture model performs significantly better. Moreover, when the losses are considered across moneyness it is observed that our model does a much better job than the benchmark models in explaining the smirk found in our sample of options. Finally, when considering separately the dot-com bubble and burst, the subsequent calm period, and the financial crises starting in 2008 the performance of the mixture model is particularly stable compared to the benchmark models. In particular, when considering the financial crisis beginning in 2008 the dollar root mean squared error of the best performing benchmark component model relative to the mixture model increases to 69%.

Finite mixture models, which are convex combinations of densities, are becoming a standard tool in financial econometrics. They are attractive because of the parsimonious flexibility they provide in the specification of the distribution of the underlying random variable, which gives them a semiparametric flavor. In this framework, each distribution in the mixture can have its own mean and conditional variance process. Moreover, if required by the data, some conditional variance processes may even be weakly nonstationary, for example to capture turbulent periods, while the overall conditional variance remains weakly stationary. Finite mixture textbooks are for example McLachlan and Peel (2000) and Frühwirth-Schnatter (2006). Early applications are Kon (1982) and Kim and Kon (1994) who investigate the statistical properties of stock returns using mixture models. Boothe and Glassman (1987), Tucker and Pond (1988), and Pan, Chan, and Fok (1995) use mixtures of normals to model exchange rates. Recent examples are Wong and Li (2000) and Wong and Li (2001) who model the conditional return distribution, extended by Haas, Mittnik, and Paoletta (2004) with an application of value at risk prediction, and Bauwens and Rombouts (2007) for the clustering of financial time series. Durham (2007) investigates the power of finite mixtures of normal densities with stochastic volatility to model the return of the S&P 500 index.

Using statistical criteria such as QQ-plots, goodness-of-fit tests and information criteria he finds that the finite mixture can capturing most of the salient features of the data. We extend his work by analyzing the performance of finite mixtures in out-of-sample option pricing. Hence, we focus both on the statistical and financial properties of the finite mixture model. In doing so, we examine both the physical and risk neutral measures.

Option pricing with our approach is straightforward since we only use historical data on the underlying asset, and therefore the model can be estimated directly by maximum likelihood. However, historical option prices themselves contain important information on the model parameters. Therefore, an alternative approach is to infer these parameters either from historical option data alone or by using both returns and options data. For example, this can be done by calibrating the option pricing model to existing option data, as detailed in Chernov and Ghysels (2000). For this to be feasible option pricing models within the affine class are considered, a choice which is motivated by the existence of closed or semiclosed form pricing formulas. However, this class of models could be considered as restrictive since many of the well known specifications within for example the GARCH framework are non-affine. Moreover, research has shown that in terms of option pricing, non-affine models often perform better than affine models, see for example Hsieh and Ritchken (2005) and Christoffersen, Jacobs, Dorion, and Wang (2010).

The rest of the paper is organized as follows: Section 2 presents the asymmetric heteroskedastic normal mixture model and details how options can be priced using this framework. Section 3 reports estimation results and explores the implications for the risk neutral dynamics. Section 4 contains the results of the empirical application to options on the S&P 500 index, and Section 5 investigates the performance through time and in the recent financial crisis. Finally, Section 6 contains the conclusion.

## 2 Model and risk neutral dynamics

Letting  $\mathcal{F}_t$  denote the information set up to time  $t$ , we assume that the underlying return process  $R_t \equiv \ln(S_t/S_{t-1})$  can be characterized by

$$R_t = m_t(\cdot) - \Psi_t(-1) + \varepsilon_t, \quad (1)$$

where  $S_t$  is the index level on day  $t$ , and where  $\varepsilon_t | \mathcal{F}_{t-1} \sim P(0, \sigma_t^2)$  with  $\sigma_t^2$  being the conditional variance and  $P(\cdot)$  is a continuous distribution function. In (1), the term  $\Psi_t(\cdot)$  denotes the conditional cumulant generating function. This, on the other hand, corresponds to the logarithm of the conditional moment generating function and is given by

$$\Psi_t(u) = \ln(E_{t-1}[\exp(-u\varepsilon_t)]). \quad (2)$$

Evaluating this at  $u = -1$  it is seen that

$$\begin{aligned} E_{t-1}[S_t/S_{t-1}] &= E_{t-1}[\exp(m_t(\cdot) - \Psi_t(-1) + \varepsilon_t)] \\ &= \exp(m_t(\cdot)). \end{aligned} \quad (3)$$

Thus, it makes indeed sense to define the return process as above since with this particular specification  $m_t(\cdot)$  can be interpreted as the expected gross rate of return. In Section 2.3 we specify functional form for  $m_t(\cdot)$  that allows us to perform option pricing.

### 2.1 The asymmetric heteroskedastic normal mixture model

We propose to use a mixture of normals coupled with an asymmetric heteroskedastic specification for the conditional distribution  $P(0, \sigma_t^2)$ . In particular, we specify the conditional distribution of  $\varepsilon_t$  is a combination of  $K$  densities

$$P(\varepsilon_t | \mathcal{F}_{t-1}) = \sum_{k=1}^K \pi_k \Phi\left(\frac{\varepsilon_t - \mu_k}{\sigma_{k,t}}\right), \quad (4)$$

where

$$\sigma_{k,t}^2 = \omega_k + \alpha_k(\varepsilon_{t-1} + \gamma\sigma_{k,t-1})^2 + \beta_k\sigma_{k,t-1}^2, \quad (5)$$

and  $\Phi(\cdot)$  is the standard Gaussian distribution. We denote this the asymmetric heteroskedastic normal mixture (MN-NGARCH) model, since the conditional variance specification in (5) corresponds to the well known NGARCH model of Engle and Ng (1993). The MN-GARCH model developed by Haas, Mittnik, and Paoletta (2004) corresponds to the case when  $\gamma = 0$  in (5). We note that from an econometric point of view the mixture model is attractive since a common innovation term  $\varepsilon_t$  feeds in the  $K$  conditional variance equations. This makes the evaluation of the likelihood function possible since there is no path dependence problem as would be the case in a Markov switching GARCH model. This allows us to estimate the mixture model by standard maximum likelihood techniques.

In the MN-NGARCH model at each  $t$  the innovation  $\varepsilon_t$  is drawn from one of the  $K$  conditional distributions with probabilities  $\pi_1, \dots, \pi_K$ . Consequently, the parameter  $\pi_k$  is restricted to be positive for all  $k$  and  $\sum_{k=1}^K \pi_k = 1$ , which is imposed by setting  $\pi_K = 1 - \sum_{k=1}^{K-1} \pi_k$ . Moreover, the zero mean assumption on  $\varepsilon_t$  is ensured by the restriction

$$\mu_K = - \sum_{k=1}^{K-1} \frac{\pi_k \mu_k}{\pi_K}. \quad (6)$$

Note that this zero mean restriction does not imply a symmetric distribution. A symmetric distribution would obtain only if all  $\mu_k$ 's are zero, something that can be tested.

The MN-NGARCH provides a very flexible model which can capture, for example, conditional skewness and kurtosis. In particular, in the MN-NGARCH model conditional moments of  $\varepsilon_t$  are combinations of the  $K$  distribution moments in (4). The conditional variance of  $\varepsilon_t$  is given by

$$\sigma_t^2 = \sum_{k=1}^K (\pi_k \mu_k^2 + \pi_k \sigma_{k,t}^2). \quad (7)$$

The conditional third moment is given by

$$E_{t-1}(\varepsilon_t^3) = \sum_{k=1}^K (\pi_k \mu_k^3 + 3\pi_k \sigma_{k,t}^2 \mu_k), \quad (8)$$

and the conditional fourth moment is given by

$$E_{t-1}(\varepsilon_t^4) = \sum_{k=1}^K (\pi_k \mu_k^4 + 6\pi_k \mu_k^2 \sigma_{k,t}^2 + 3\pi_k \sigma_{k,t}^4). \quad (9)$$



These formulas illustrate that we can have flexible dynamics compared to the classical GARCH model which arises when  $K = 1$ . For example, the skewness of the conditional distribution would be forced to zero in the latter case.

Moreover, note that to have an overall variance process that is weakly stationary, only one of the conditional variance processes is required to be weakly stationary. The other  $K - 1$  conditional variance processes are allowed to be explosive ( $\alpha_k(1 + \gamma_k^2) + \beta_k > 1$ ) as long as their combined probability is not too high. More formally, the weak stationarity condition for the model is

$$\left[ \sum_{k=1}^K \frac{\pi_k}{(1 - \beta_k)} \left( 1 - \alpha_k(1 + \gamma_k^2) - \beta_k \right) \right] \prod_{k=1}^K (1 - \beta_k) > 0. \quad (10)$$

More details on the unconditional moments of the symmetric heteroskedastic normal mixture model can be found in Haas, Mittnik, and Paoletta (2004).

It should be noted that the parameters of the mixture model are not identified as such because of the label switching problem which leaves the model likelihood unchanged when we change the order of the distributions in the finite mixture. This is not a problem if the objects of interest are label invariant, an example would be the predictive density of future returns. However, if we want to give a financial interpretation of the parameters, like in this paper, we add an identification restriction like  $\pi_1 \geq \pi_2 \geq \dots > \pi_K$ . Other restrictions, for example on the mean of the distributions, are possible as explained in Hamilton, Zha, and Waggoner (2007).

Finally, we note that the mixture model can incorporate extreme events by having a distribution with very low probability and with a large mean and a small constant variance, for example. Moreover, the model can be modified if needed by considering other conditional variance models. Examples include, but are not restricted to, the GJR-GARCH model of Glosten, Jagannathan, and Runkle (1993) or the EGARCH model of Nelson (1991). Finally, other distributions than the normal can be considered in the finite mixture like the exponential power distribution proposed by Bouaddi and Rombouts (2009).

Alternative to the mixture approach of this paper, Student type distributions with time varying skewness and kurtosis have been introduced in the literature to model the conditional distribution

of stock returns. For example, using skewed Student distributions Hansen (1994), Harvey and Siddique (1999), and Jondeau and Rockinger (2003) find also insample evidence for time varying higher moments. However, these conditional distributions are fully parametric and miss the semiparametric flavor of mixtures, and therefore can be misleading when severely misspecified. Furthermore, in our context of option pricing, Student type distributions (or even mixtures of them) are not used because the conditional Laplace transform is unknown and their moments do not necessarily always exist.

## 2.2 Risk neutral dynamics

In the MN-NGARCH model markets are incomplete and hence there is no unique way to derive the equivalent martingale measure (EMM) needed for option pricing. In this paper, we follow the approach of Christoffersen, Elkamhi, Feunou, and Jacobs (2010) in which a candidate EMM is specified from the following Radon-Nikodym derivative

$$\frac{dQ}{dP} \Big|_{\mathcal{F}_t} = \exp \left( - \sum_{i=1}^t (\nu_i \varepsilon_i + \Psi_i(\nu_i)) \right), \quad (11)$$

where  $\Psi_t(u)$  is the conditional cumulant generating function from above. Note that though our model has  $K$  components, there is only one innovation term and hence in (11) there is also only one  $\nu_t$  sequence. However, if the model had multiple innovations a natural generalization of this would involve several such sequences. For example, this is the case in the general multivariate discrete time framework developed in Rombouts and Stentoft (2011). Alternatively, one could consider a stochastic volatility model in which it would also be natural to include a second sequence for the innovation to the volatility process (see Christoffersen, Elkamhi, Feunou, and Jacobs (2010)).

It is immediately observed that by using a Radon-Nikodym of this type we are guaranteed that the two measures are equivalent as they have the same null sets. Moreover, to ensure that discounted asset prices are martingales under the risk neutral measure it can be shown that the

sequence  $\nu_t$  has to satisfy the following equation

$$0 = \Psi_t(\nu_t - 1) - \Psi_t(\nu_t) - \Psi_t(-1) + m_t(\cdot) - r_t. \quad (12)$$

Finally, it can be shown that under the risk neutral measure the conditional cumulant generating function of  $\varepsilon_t$  is given by

$$\Psi_t^Q(u) = \Psi_t(\nu_t + u) - \Psi_t(\nu_t). \quad (13)$$

For more details and proofs see Christoffersen, Elkamhi, Feunou, and Jacobs (2010). Using the Inversion Theorem (see for example Billingsley (1995, Theorem 26.2) or Davidson (1997, Theorem 11.12)) the expression in (13) can be used to obtain the distribution under  $Q$ .

Note that an alternative method to the one used here which would provide a similar set of conditions is to specify a candidate stochastic discount factor directly as is done in Gouriéroux and Monfort (2007). For a discussion of the relationship between the two probability measures  $P$  and  $Q$  and the corresponding stochastic discount factor see Bertholon, Monfort, and Pegoraro (2008). One can also work within a general equilibrium setup as is done in Duan (1999). While this method also yields the dynamics to be used for option pricing, the specification is generally less explicit. In particular, an actual application of the method is computationally complex (see e.g. Stentoft (2008)), and this approach therefore appears to be more restrictive.

### 2.2.1 Riskneutralization of asymmetric heteroskedastic normal mixture models

Conditional on (11), equations (12) and (13) completely characterize the risk neutral process and hence this is, in fact, all that is needed for option pricing purposes. In particular, equation (12) provides a link between the original measure  $P$  and a risk neutral measure  $Q$  through the sequence  $\nu_t$  and (13) characterizes the risk neutral distribution in terms of this sequence. Thus, in order to apply the method for pricing all that is left is to derive these dynamics explicitly.

In the case of the MN-NGARCH model the conditional cumulant generating function is given

by

$$\Psi_t(u) = \ln \left( \sum_{k=1}^K \pi_i \exp \left( -u\mu_k + \frac{u^2\sigma_{k,t}^2}{2} \right) \right). \quad (14)$$

This is just the logarithm of a convex combination of Gaussian moment generating functions and thus very simple to calculate. For option pricing purposes, the fact that it may easily be calculated is very convenient. Substituting (14) into (13) the conditional cumulant generation function of  $\varepsilon_t$  under the risk neutral measure  $Q$  is easily obtained as

$$\begin{aligned} \Psi_t^Q(u) &= \Psi_t(\nu_t + u) - \Psi_t(\nu_t) \\ &= \ln \left( \frac{\sum_{k=1}^K \pi_i \exp \left( -(\nu_t + u)\mu_k + \frac{(\nu_t + u)^2\sigma_{k,t}^2}{2} \right)}{\sum_{k=1}^K \pi_i \exp \left( -\nu_t\mu_k + \frac{\nu_t^2\sigma_{k,t}^2}{2} \right)} \right) \\ &= \ln \left( \sum_{k=1}^K \pi_i^* \exp \left( -u\mu_k^* + \frac{u^2\sigma_{k,t}^2}{2} \right) \right), \end{aligned} \quad (15)$$

where

$$\mu_{k,t}^* = \mu_k - \nu_t\sigma_{k,t}^2, \quad (16)$$

and

$$\pi_{k,t}^* = \frac{\pi_k \exp \left( -\nu_t\mu_k + \frac{\nu_t^2\sigma_{k,t}^2}{2} \right)}{\sum_{k=1}^K \pi_k \exp \left( -\nu_t\mu_k + \frac{\nu_t^2\sigma_{k,t}^2}{2} \right)}, \quad (17)$$

for  $k = 1, \dots, K$ . Thus, the risk neutral distribution of  $\varepsilon_t$  remains within the family of normal mixtures. Using the stochastic discount factor principle of Gouriéroux and Monfort (2007) similar results are derived by Bertholon, Monfort, and Pegoraro (2006) in the particular restricted case with  $K = 2$  and for the GARCH specification only.

We remark that for the risk neutral distribution, the weak stationarity condition in (10) is not appropriate anymore, since the finite mixture under  $Q$  has time varying distribution probabilities  $\pi_t^*$ . Therefore, it can in principle occur, that the physical distribution is weakly stationary, but

the risk neutral distribution is not. This has been noted also by Duan (1995) in the Gaussian framework.

### 2.2.2 Interpreting the impact of riskneutralization

Equations (16) and (17) show that in the special case with no unit risk premium the risk neutral distribution correspond to the original distribution. However, in general the distribution of  $\varepsilon_t$  will have changed means and probabilities under  $Q$ . In particular, with respect to the risk neutral means from (16), it is immediately seen that the correction is very similar to what is obtained with a Gaussian model, where the mean of  $\varepsilon_t$  under  $Q$  is equal to  $-\nu_t\sigma_t^2$ . The intuition behind this is the following: If volatility risk carries a positive premium, then in the risk neutral world the mean of the innovations is shifted downwards to compensate for this.

With respect to the risk neutral probabilities, the relationship is somewhat less straightforward. Thus, for simplicity we now consider the special case where  $K = 2$  in which (17) simplifies to

$$\pi_t^* = \frac{\pi \exp\left(-\nu_t\mu_1 + \frac{\nu_t^2\sigma_{1,t}^2}{2}\right)}{\pi \exp\left(-\nu_t\mu_1 + \frac{\nu_t^2\sigma_{1,t}^2}{2}\right) + (1 - \pi) \exp\left(-\nu_t\mu_2 + \frac{\nu_t^2\sigma_{2,t}^2}{2}\right)}. \quad (18)$$

Moreover, by dividing through with  $\exp\left(-\nu_t\mu_1 + \frac{\nu_t^2\sigma_{1,t}^2}{2}\right)$ , it is seen that

$$\pi_t^* \begin{matrix} \geq \\ \leq \end{matrix} \pi \text{ if } \exp\left(-\nu_t\mu_1 + \frac{\nu_t^2\sigma_{1,t}^2}{2}\right) \begin{matrix} \geq \\ \leq \end{matrix} \exp\left(-\nu_t\mu_2 + \frac{\nu_t^2\sigma_{2,t}^2}{2}\right). \quad (19)$$

Thus, if we further assume that the variance term  $\frac{\nu_t^2\sigma_{k,t}^2}{2}$  is negligible compared to the mean term  $-\nu_t\mu_k$ , for  $k = 1, 2$ , this restriction simplifies to

$$\pi_t^* \begin{matrix} \geq \\ \leq \end{matrix} \pi \text{ if } \nu_t\mu_2 \begin{matrix} \geq \\ \leq \end{matrix} \nu_t\mu_1. \quad (20)$$

However, by construction  $\mu_2 < \mu_1$  and hence the effect on the probabilities depends entirely on the sign of the unit risk premium. Since we expect this premium to be positive, it follows that we will in general have that  $\pi_t^* < \pi$ . The intuition behind this result is the following: If volatility risk

carries a positive premium then the probability attributed to the explosive part, that is  $1 - \pi_t^*$ , is increased in the risk neutral world to compensate appropriately for this.

### 2.3 Feasible option pricing

As mentioned above, we effectively choose the appropriate EMM by solving (12) for  $\nu_t$  given  $m_t(\cdot)$  and the assumed distribution of  $\varepsilon_t$ . However, this is potentially complicated due to the nonlinearity of this relationship and an analytical expression for  $\nu_t$  may therefore not be available in general. In particular, this is the case with the asymmetric heteroskedastic normal mixture model used here, and it may thus seem to be impossible to derive the EMM given  $m_t(\cdot)$  for this model.

However, (12) may equally well be solved for  $m_t(\cdot)$  given  $\nu_t$  and the assumed distribution of  $\varepsilon_t$  as

$$m_t = r_t - \Psi_t(\nu_t - 1) + \Psi_t(\nu_t) + \Psi_t(-1). \quad (21)$$

From this, we note that for any choice of  $\nu_t$  a closed form expression exist for  $m_t$  given that the cumulant generating function exists. Substituting this into the return equation in (1) we obtain

$$R_t = r_t - \Psi_t(\nu_t - 1) + \Psi_t(\nu_t) + \varepsilon_t, \quad (22)$$

which can be used for estimation directly. In fact, the model we use for option pricing is a NGARCH-in-mean mixture model.

Note that this way of implying the gross rate of return  $m_t$ , given a particular specification of  $\nu_t$ , is used in Stentoft (2008) in the Normal Inverse Gaussian framework using the option pricing model of Duan (1999). However, it is equally applicable here and the method allows us to incorporate coherently a risk premium in the MN-NGARCH. This was, for instance, neglected in Badescu, Kulperger, and Lazar (2008).

### 2.3.1 Interpreting $\nu_t$ in the Gaussian special case

It should be noted that, depending on the specification of  $\nu_t$  and the assumed distribution of  $\varepsilon_t$ , we may interpret the relationship in (22) differently. We now illustrate this for the Gaussian case in which (22) corresponds to

$$\begin{aligned} R_t &= r_t - \frac{1}{2}\sigma_t^2(\nu_t - 1)^2 + \frac{1}{2}\sigma_t^2(\nu_t)^2 + \varepsilon_t \\ &= r_t - \frac{1}{2}\sigma_t^2 + \nu_t\sigma_t^2 + \varepsilon_t. \end{aligned} \tag{23}$$

Hence, we see that  $\nu_t$  is directly related to the unit risk premium. In particular, if we were to specify  $\nu_t = \nu$ , that is as a constant, the implied mean specification corresponds to assuming a unit risk premium proportional to the level of the variance. Alternatively, if  $\nu_t = \nu/\sigma_t$ , the unit risk premium becomes proportional to the level of the standard deviation, and finally with  $\nu_t = \nu/\sigma_t^2$  a constant unit risk premium independent of the level of the variance is obtained.

Thus, while it may appear that we, by implying the gross rate of return through (22), are constraining the potential mean specification in an unreasonably way from an econometric point of view, this is in fact not the case. In particular, all of the specifications used in the existing literature can be implemented. Also note that  $\nu_t = 0$  is permitted, and which particular specification is the most appropriate one can be tested by simple likelihood ratio type tests.

### 2.3.2 Interpreting $\nu_t$ in the general case

In the general case, i.e. with the asymmetric heteroskedastic normal mixture model, the relation is less obvious. However, we may analyze the effect of  $\nu_t$  using a Taylor series expansion in (22). To do this, we first note that the two terms involving the cumulant generating functions may be approximated by

$$\begin{aligned} \Psi_t(\nu_t - 1) &\approx \Psi_t(0) + \Psi_t'(0)(\nu_t - 1) + \frac{1}{2}\Psi_t''(0)(\nu_t - 1)^2 \\ &\quad + \frac{1}{6}\Psi_t'''(0)(\nu_t - 1)^3 + \frac{1}{24}\Psi_t''''(0)(\nu_t - 1)^4 \end{aligned} \tag{24}$$

and

$$\Psi_t(\nu_t) \approx \Psi_t(0) + \Psi_t'(0)\nu_t + \frac{1}{2}\Psi_t''(0)\nu_t^2 + \frac{1}{6}\Psi_t'''(0)\nu_t^3 + \frac{1}{24}\Psi_t''''(0)\nu_t^4, \quad (25)$$

respectively. Furthermore, by the definition of  $\varepsilon_t$  as a zero mean random variable with conditional variance  $\sigma_t^2$  we have that  $\Psi_t'(0) = 0$  and  $\Psi_t''(0) = \sigma_t^2$ . Moreover, by the definition of  $\Psi_t(u)$ , we have that  $\Psi_t'''(0) = -skew_t\sigma_t^3$  and  $\Psi_t''''(0) = exkurt_t\sigma_t^4$ , where  $skew_t$  and  $exkurt_t$  denotes the conditional skewness and excess kurtosis at time  $t$  respectively.

With the above expressions, (22) may be rewritten as

$$R_t \approx r_t - \Psi_t(-1) + \nu_t\sigma_t^2 + \frac{(3\nu_t^2 - 3\nu_t)}{6}(-skew_t)\sigma_t^3 + \frac{(4\nu_t^3 - 6\nu_t^2 + 4\nu_t)}{24}exkurt_t\sigma_t^4 + \varepsilon_t, \quad (26)$$

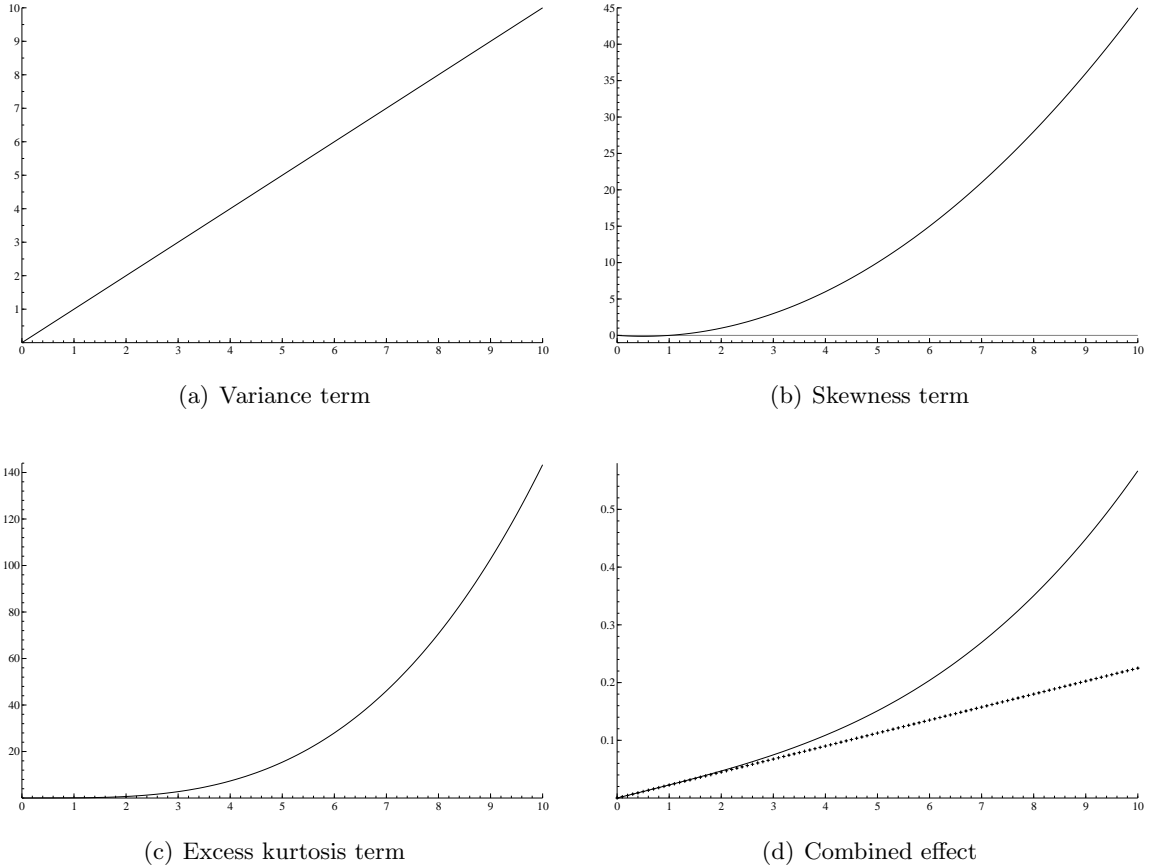
where we have used the fact that

$$\Psi_t(-1) \approx \Psi_t(0) - \Psi_t'(0) + \frac{1}{2}\Psi_t''(0) - \frac{1}{6}\Psi_t'''(0)(-1) + \frac{1}{24}\Psi_t''''(0), \quad (27)$$

which again follows from a Taylor series expansion, to collect all the terms that do not involve  $\nu_t$  in  $\Psi_t(-1)$ . Equation (26) shows that even in the more general setting it makes sense to interpret  $\nu_t$  as the unit risk premium. In particular, it may be observed that for reasonable values of  $\nu_t$ , say larger than 1, the coefficients above are all positive. In Figure 1, we plot in panels (a) through (c) the coefficients of the variance term, the skewness term, and the excess kurtosis term as a function of  $\nu_t$ . Thus, in a general setting, the premium is increasing in the variance, decreasing in the skewness, and increasing in the excess kurtosis.

In the case of our model for the S&P 500 index, for which we obtain negative skewness and excess kurtosis, equation (26) shows that the higher the value of  $\nu_t$  the larger the premium required by investors for holding this particular risky asset. Figure 1(d) plots the combined effect of the three terms with a solid line. Note also, that if we were to neglect these higher order moment properties of the return process and assume them equal to zero, then to compensate for this a higher value of  $\nu_t$  would be required to generate the same overall level of compensation for risk. This is clear from the dotted line in Figure 1(d) which shows the effect without the higher order terms.





**Figure 1: Unit risk premium effects**

This figure plots the coefficients of the variance term, the skewness term, and the excess kurtosis term in (26) as a function of  $\nu_t$  in panels (a) through (c). Panel (d) plots the combined effect. The dotted line in this panel shows the effect when the higher order terms are neglected.

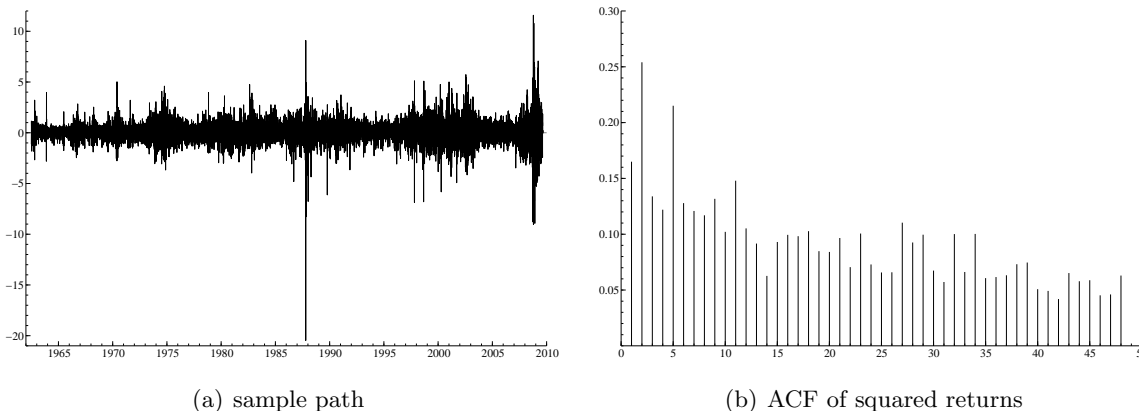
### 3 Application to the S&P 500 index

We provide an application of our MN-NGARCH framework to a long sample of data on the S&P 500 index. We compare several versions of our mixture model to the Gaussian GARCH and NGARCH models, which are special cases of our MN-GARCH model. We also analyze the implied dynamics under the risk neutral measure and compare these to what is obtained under the original measure. In the first section we describe the data. The second section contains the estimation results, and the last section discusses the implied dynamics to be used for option pricing.

**Table I: Descriptive statistics for S&P 500 index percentage returns**

Sample period: July 2, 1962, to August 26, 2009, for a total of 11,869 observations.

Minimum	-20.467	Standard deviation	1.016
Mean	0.030	Skewness	-0.663
Maximum	11.580	Kurtosis	25.944



**Figure 2: Properties of the S&P 500 returns**

This figure plots the sample path and the ACF of squared returns using the sample period from July 2, 1962 to August 26, 2009, for a total of 11,869 observations.

### 3.1 Data

As return data, we use the S&P 500 index returns excluding dividends from July 2, 1962, which is the first day data from CRSP is available, to August 26, 2009, which corresponds to the last week for which we have option data, for a total of 11,869 observations. Table I provides the standard descriptive statistics for the S&P 500 index return series. The numbers in this table show that the return data is negatively skewed and very leptokurtic. Figure 2(a) displays the sample path and shows the well known pattern of time varying volatility with periods of high volatility levels followed by periods of low levels of volatility. Figure 2(b) provides evidence of the strong persistence in squared returns, a proxy of the second moment of the series. Considering the evidence of Table I and Figure 2 it is clear that a model which allows for conditional heteroskedasticity and non-Gaussian features is needed for this type of data.

### 3.2 Estimation results

In Tables II and III, we provide maximum likelihood estimation results for the mixture models with GARCH and NGARCH volatility specifications. The parametrization used is given by  $R_t = r - \Psi_t(\nu - 1) + \Psi_t(\nu) + \varepsilon_t$ , with the distribution of  $\varepsilon_t$  defined in (4).<sup>1</sup> Note that in our empirical application we choose to use a constant value for  $\nu_t$ , which according to (26) then can be interpreted as the unit risk premium. Though an interesting alternative would be to specify a separate risk premium parameter for each component, it is difficult to estimate precisely more than one such parameter. In general, all parameters are statistically significantly different from zero with the exception of the  $\alpha$ 's of the most explosive component in the three component model, the  $\gamma$ 's of the most explosive NGARCH variance process, and the  $\nu$ 's for the MN2- and MN3-NGARCH models.

In terms of statistical fit, Tables II and III clearly show the value of using the mixture model. In particular, for both the MN2GARCH and MN2NGARCH specifications the likelihood value, which is shown in the second to last row, increases dramatically. Adding a third distribution in the mixture also increases the likelihood value although less importantly. When considering the Bayesian Information Criteria, or BIC, shown in the last row, the best model is the MN3 model for both volatility specifications. Tables II and III also clearly show the value of allowing for asymmetries in the volatility specification. In particular, for each specification when the NGARCH model is used the likelihood value increases dramatically. When considering the BIC, the NGARCH models are again preferred to the GARCH specifications. Thus, overall the results provide clear support for mixture models with NGARCH volatility specifications.

When considering the actual volatility processes in the mixture distributions, the estimation results in the Tables II and III show that these are very different, and the second variance process in the MN2 model is explosive. For the GARCH volatility specification the persistence, defined as  $\alpha_i + \beta_i$ , of 1.387. For the NGARCH the persistence, defined as  $\alpha_i(1 + \gamma_i^2) + \beta_i$ , is 1.365. The first distributions in the mixture, which have a probability of 94% and 96%, respectively, are however

---

<sup>1</sup>As the interest rate we take the LIBOR rate, which on August 26, 2009, was 0.45% annualized.

**Table II: Estimation results for the GARCH mixture models**

This table reports maximum likelihood estimates for the various GARCH specifications using percentage returns. A - symbol means that the parameter is set to 0. The date used is from July 2, 1962, to August 26, 2009, for a total of 11,869 observations. Robust standard errors are reported in brackets. The row headed Loglik reports the log likelihood value, and the row headed BIC reports the value of the Bayesian Information Criteria defined as  $2(\log(p) - \text{Loglik})/obs$ , where  $p$  is the number of parameters and  $obs$  is the number of observations. The last four rows report the overall pricing errors for our sample of options (see also Section 4).

Parameter	MN1GARCH		MN2GARCH		MN3GARCH	
	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.
$\nu$	6.088	(0.790)	6.107	(1.096)	5.354	(0.841)
$\mu_1$	-	-	0.001	(0.009)	0.073	(0.018)
$\mu_2$	-	-	-	-	-0.675	(0.310)
$\pi_1$	-	-	0.941	(0.019)	0.203	(0.035)
$\pi_2$	-	-	-	-	0.010	(0.004)
$\omega_1$	0.005	(0.001)	0.003	(0.001)	0.027	(0.014)
$\alpha_1$	0.074	(0.004)	0.052	(0.005)	0.014	(0.005)
$\beta_1$	0.923	(0.004)	0.935	(0.005)	0.932	(0.021)
$\omega_2$	-	-	0.170	(0.100)	0.751	(0.329)
$\alpha_2$	-	-	0.608	(0.197)	1.756	(0.941)
$\beta_2$	-	-	0.779	(0.063)	0.656	(0.163)
$\omega_3$	-	-	-	-	0.065	(0.007)
$\alpha_3$	-	-	-	-	0.073	(0.006)
$\beta_3$	-	-	-	-	0.932	(0.005)
Loglik	-14542.66		-14340.78		-14273.69	
BIC	2.4508		2.4169		2.4056	
\$ BIAS	4.193		2.782		2.727	
\$ RMSE	13.616		13.105		10.925	
ISD BIAS	2.729		2.101		2.084	
ISD RMSE	7.522		7.660		6.626	

**Table III: Estimation results for the NGARCH mixture models**

This table reports maximum likelihood estimates for the various NGARCH specifications using percentage returns. A - symbol means that the parameter is set to 0. The date used is from July 2, 1962, to August 26, 2009, for a total of 11,869 observations. Robust standard errors are reported in brackets. The row headed Loglik reports the log likelihood value, and the row headed BIC reports the value of the Bayesian Information Criteria defined as  $2(\log(p) - \text{Loglik})/obs$ , where  $p$  is the number of parameters and  $obs$  is the number of observations. The last four rows report the overall pricing errors for our sample of options (see also Section 4).

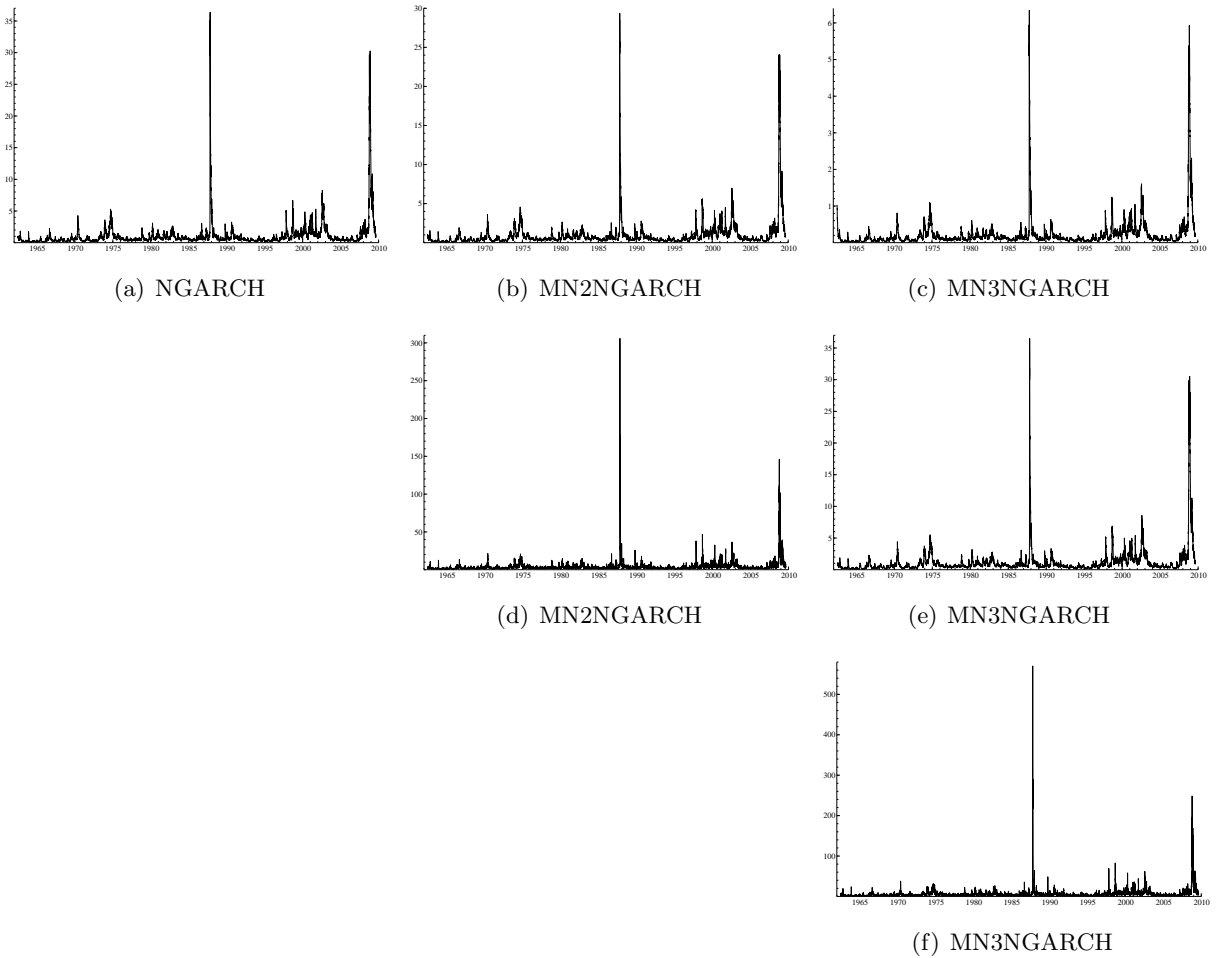
Parameter	MN1NGARCH		MN2NGARCH		MN3NGARCH	
	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.
$\nu$	1.791	(0.769)	-0.250	(1.082)	1.039	(0.808)
$\mu_1$	-	-	0.037	(0.009)	0.066	(0.021)
$\mu_2$	-	-	-	-	-0.751	(0.366)
$\pi_1$	-	-	0.962	(0.012)	0.165	(0.038)
$\pi_2$	-	-	-	-	0.010	(0.004)
$\omega_1$	0.007	(0.001)	0.004	(0.001)	0.028	(0.014)
$\gamma_1$	-0.683	(0.047)	-0.815	(0.060)	-1.985	(1.127)
$\alpha_1$	0.062	(0.004)	0.050	(0.004)	0.010	(0.006)
$\beta_1$	0.906	(0.005)	0.910	(0.006)	0.909	(0.044)
$\omega_2$	-	-	0.242	(0.137)	0.679	(0.270)
$\gamma_2$	-	-	-0.303	(0.160)	-0.172	(0.295)
$\alpha_2$	-	-	0.566	(0.197)	1.104	(0.792)
$\beta_2$	-	-	0.747	(0.084)	0.714	(0.127)
$\omega_3$	-	-	-	-	0.072	(0.007)
$\gamma_3$	-	-	-	-	-0.669	(0.058)
$\alpha_3$	-	-	-	-	0.062	(0.005)
$\beta_3$	-	-	-	-	0.913	(0.006)
Loglik	-14404.08		-14208.43		-14159.94	
BIC	2.4274		2.3946		2.3865	
\$ BIAS	2.771		1.145		1.879	
\$ RMSE	10.418		7.033		9.476	
ISD BIAS	1.866		1.019		1.522	
ISD RMSE	6.326		4.817		5.542	

very similar to what is found in the MN1 Gaussian models. Given that the probability to be in this state is very large the second order stationarity condition in (10) is still met. However, it is thanks to the second distribution that we can accommodate for example the high kurtosis in the index returns. When adding the third distribution in the mixture, Tables II and III first of all show that the persistence in the second variance process becomes even larger, that is 2.412 for the GARCH model and 1.851 for the NGARCH model. However, in this specification the persistence of the first variance process on the other hand decreases to 0.946 for the GARCH model and 0.958 for the NGARCH model.

In Figure 3, we plot the conditional variances of each of the processes. While all plots have the same overall shape the differences in persistence is clear. For example, the second volatility path in the MN2NGARCH model reacts much stronger to the crash in 1987 and because of this the first volatility path is now somewhat more stable. In the MN3NGARCH model the 3 conditional volatility paths are also very different. In particular, while the second path is similar to the conditional volatility path from the NGARCH model the volatility path of the first variance process is much more stable and the third volatility path much less so.

### 3.3 Implications for the risk neutral dynamics

When it comes to option pricing an important parameter is the risk premium parameter,  $\nu$ . In particular, if this parameter is significant there will be differences between the dynamics under  $P$  and  $Q$ . Table II shows that this parameter is in fact statistically significant and large for all models when considering a GARCH volatility specification. Moreover, the table shows that the size of the premium tends to decrease when the number of distributions in the mixture is increased. We note that this is to be expected since the MN1 Gaussian model neglects the impact of higher order conditional moments. In particular, as it was mentioned above the higher order terms in (26) varies negatively with conditional skewness and positively with conditional excess kurtosis for a given level of  $\nu$ . Hence, when these terms are neglected, a higher  $\nu$  is required to compensate



**Figure 3: Conditional volatility paths**

This figure plots the conditional volatility paths for the NGARCH mixture models. For each mixture model, the paths are plotted in terms of increasing persistence from top to bottom.

investors appropriately for the risk of the asset. Table III, on the other hand, shows that with a NGARCH volatility specification  $\nu$  is in fact no longer statistically significant for the MN2 and MN3 models. For the MN1NGARCH model the point estimate is much lower than for the GARCH specification. Thus, our results show that when asymmetries are allowed for and once sufficiently flexible distributions are used the risk premium becomes insignificant. This again has important implications for option pricing as the distribution under the physical and risk neutral measures are then identical.

Figure 4 illustrates the term structure of the skewness and kurtosis of predicted returns under  $P$  and  $Q$  until the highest maturity for the last week of the option dataset, i.e. 213 trading days. The plots are shown for the MN2GARCH and MN2NGARCH models only, using the estimated parameters from Tables II and III. The figures show that for the mixture models the term structures are very different under  $P$  and  $Q$ . In particular, this figure shows that the model has more negative skewness and a higher kurtosis under  $Q$  than under  $P$ . These differences are a result of the risk premium parameter  $\nu$  which drives a wedge between the physical and risk neutral dynamics, especially so for the MN2GARCH model for which  $\nu$  is largest as Table II shows.

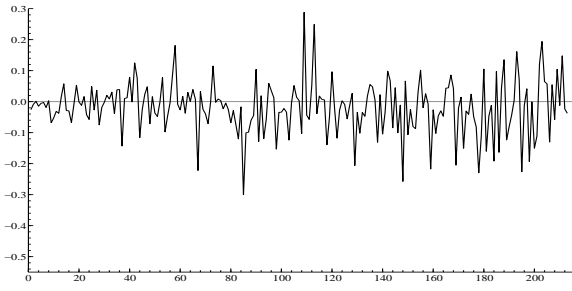
## 4 Pricing S&P 500 options

We now evaluate our model’s out-of-sample forecast performance by pricing a sample of 27,137 call option contracts on the S&P 500 index from 1996 through 2009. In the next section, we describe the option data in detail. We then report detailed results on the pricing performance using both the dollar losses and the implied standard deviation losses.

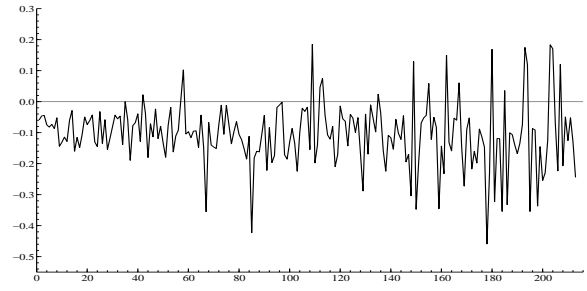
### 4.1 Data and option pricing methodology

In this paper, we follow the literature and use data on call options on the S&P 500 index. Our data covers the period from 1996 through 2009 which is the entire period for which data is provided by OptionMetrics. We impose the following standard restrictions on our sample: Firstly, we consider

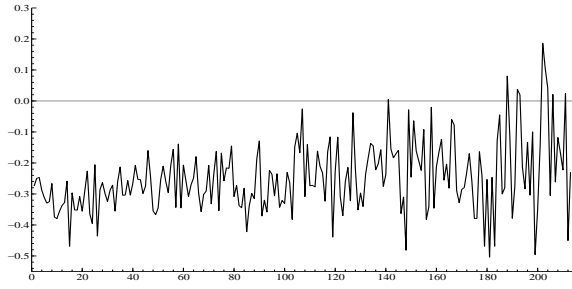




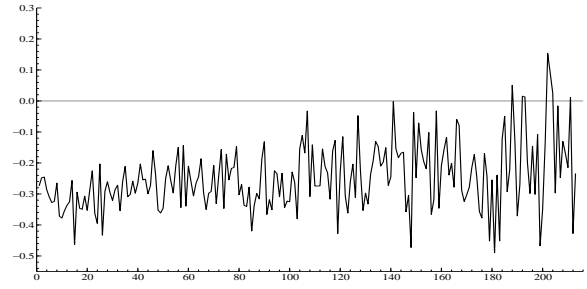
(a) MN2GARCH skewness under  $P$



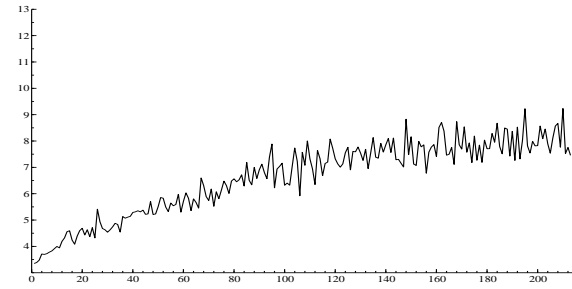
(b) MN2GARCH skewness under  $Q$



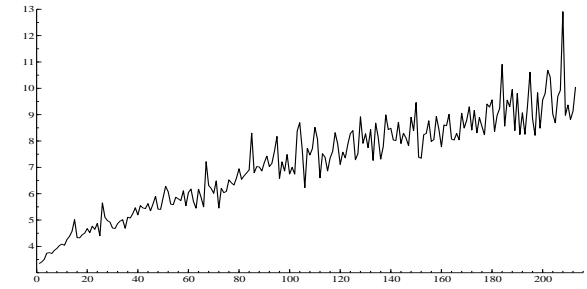
(c) MN2NGARCH skewness under  $P$



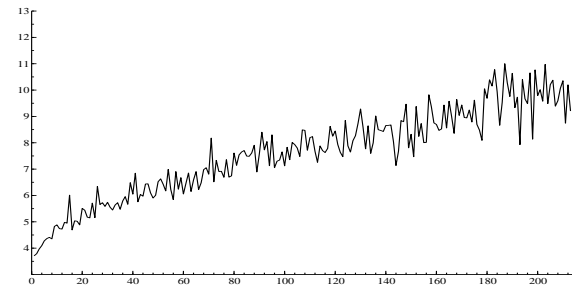
(d) MN2NGARCH skewness under  $Q$



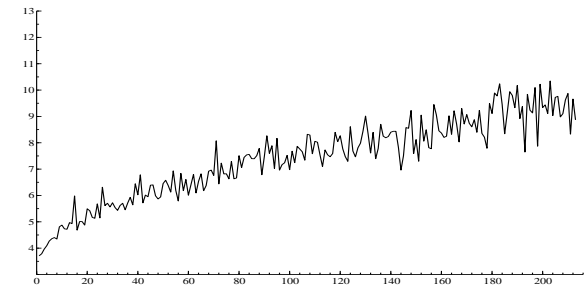
(e) MN2GARCH kurtosis under  $P$



(f) MN2GARCH kurtosis under  $Q$



(g) MN2NGARCH kurtosis under  $P$



(h) MN2NGARCH kurtosis under  $Q$

**Figure 4: Skewness and Kurtosis term structure under  $P$  and  $Q$**

This figure plots the term structure of skewness and kurtosis of the predictive returns under  $P$  and  $Q$  for the MN2GARCH and MN2NGARCH models. The moments are calculated using 20,000 draws in a Monte Carlo simulation using the parameter estimates from Tables II and III.

**Table IV: Properties of the S&P 500 index options data set**

Mean price in USD, mean ISD, and number of contracts in the cells of this table. The maturity categories are divided into very short term (VST), with  $T < 22$ , short term (ST), with  $22 \leq T < 43$ , medium term (MT), with  $43 \leq T < 85$ , long term (LT), with  $85 \leq T < 169$ , and very long term (VLT), with  $T \geq 169$ . The moneyness categories are divided into deep out of the money (DOTM), with  $M < 0.95$ , out of the money (OTM), with  $0.95 \leq M < 0.98$ , at the money (ATM), with  $0.98 \leq M < 1.02$ , in the money (ITM), with  $1.02 \leq M < 1.05$ , and deep in the money (DITM), with  $M \geq 1.05$ .

Maturity					
VST	ST	MT	LT	VLT	ALL
\$ 20.05	\$ 23.80	\$ 34.44	\$ 44.12	\$ 61.55	\$ 30.90
% 17.43	% 17.68	% 18.04	% 17.43	% 16.76	% 17.60
7077	7974	6455	3561	2070	27137
Moneyness					
DOTM	OTM	ATM	ITM	DITM	ALL
\$ 10.02	\$ 14.07	\$ 29.98	\$ 58.07	\$ 140.53	\$ 30.90
% 19.89	% 15.73	% 16.51	% 18.40	% 18.61	% 17.60
7615	5938	9341	2397	1846	27137

weekly data only and choose the options traded on Wednesdays. If Wednesday is not a trading day we pick the date closest to it. This choice is made to balance the tradeoff between having a long time period against the computational complexity. We choose Wednesdays as these options are the least affected by weekend effects. Secondly, we choose to work only with those contracts which had a daily traded volume of at least 100 contracts. Thirdly, we exclude options which have an ask price below 50 cents. Fourthly, we exclude options with less than 7 or more than 252 calendar days to maturity. Finally, we eliminate options in the LEAPS series as the contract specifications for these options do not correspond to that of the standard options. In total, we end up with a sample of 27,137 call options.

In Table IV, we provide descriptive statistics for the options in terms of the average prices, the average implied standard deviations (ISD) from the BSM model, and the number of options. We tabulate data for various categories of maturity measured in trading days,  $T$ , and moneyness measured as  $M = S / (K \exp(-rT))$ , where  $S$  is the value of the underlying,  $K$  is the strike price,

and  $r$  is the risk free interest rate. The table shows that our data contains a diverse sample of traded options. First of all, in terms of the number of contracts even the VLT category contains a large number of options and so does the DITM category. Naturally, most of the options in our sample are ATM options and most have ST maturity. Next, in terms of average option prices, the table shows that these vary from \$10.02 to \$140.53 thus spanning a large interval. Finally and most importantly, we observe the well known volatility smirk, i.e. Table IV shows that the mean ISD is higher for DOTM options than for DITM options.

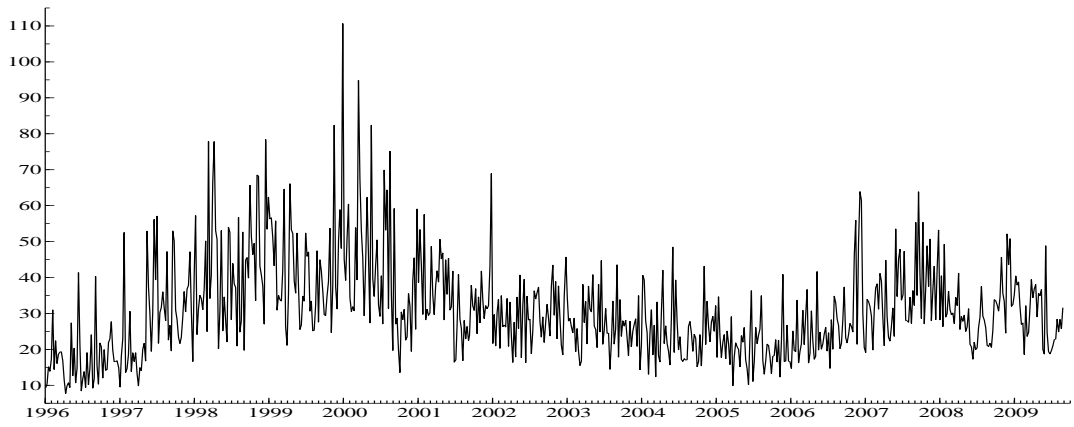
In total, we price options for 713 weeks, from January 4, 1996, to August 26, 2009. However, the average prices, volatilities, and the number of contracts are not constant over time. In Figure 5 we plot these through our sample on a weekly basis. In terms of the prices, Figure 5(a) shows that the average prices are highest around the peak of the dot-com bubble in 2000 and 2001. In terms of the implied volatilities, Figure 5(b) shows that they peak around the time of the financial crisis of 2008. This comes after a prolonged period with low volatility. Implied volatilities are also relatively high at the time of the dot-com bubble. Finally, Figure 5(c) shows a significant increase in the number of contracts throughout the sample. The number of contracts peaks around the time of the financial crisis of 2008, at the same time as the implied standard deviations.

For each week, parameter estimation and pricing is performed using the return data available at that time, and as we move forward through time the sample used for estimation increases. Given the estimated parameters the risk neutral dynamics can be derived and options may be priced as the expected value of their future cash flows discounted using the risk free interest rate. For example, the theoretical value of a European call option at time  $t$  with a strike price equal to  $K$  and maturity in  $T$  days is

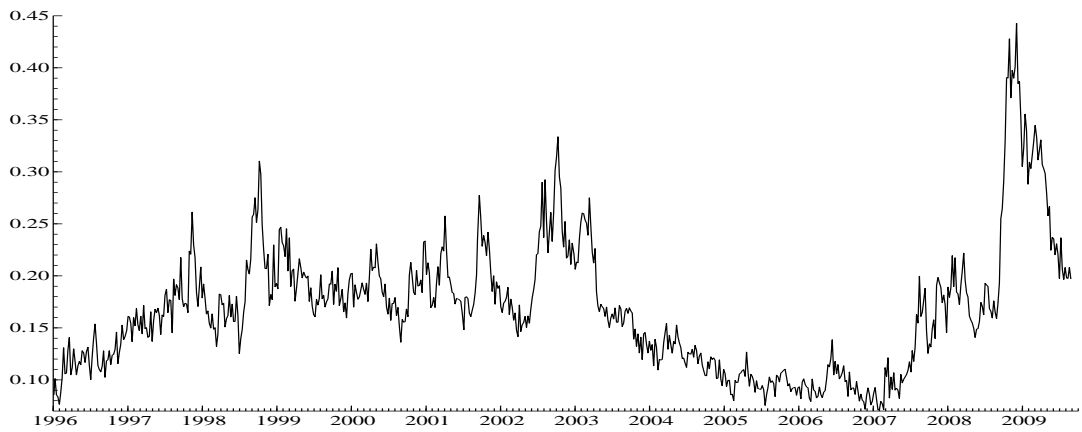
$$C_t(T, K) = e^{-r(T-t)} \mathbb{E}_t^Q [\max(K - S_T, 0)], \quad (28)$$

where  $\mathbb{E}_t^Q$  means that the expectation is taken under the risk neutral measure  $Q$  and where  $S_T$  is the value at time  $T$  of the underlying index.

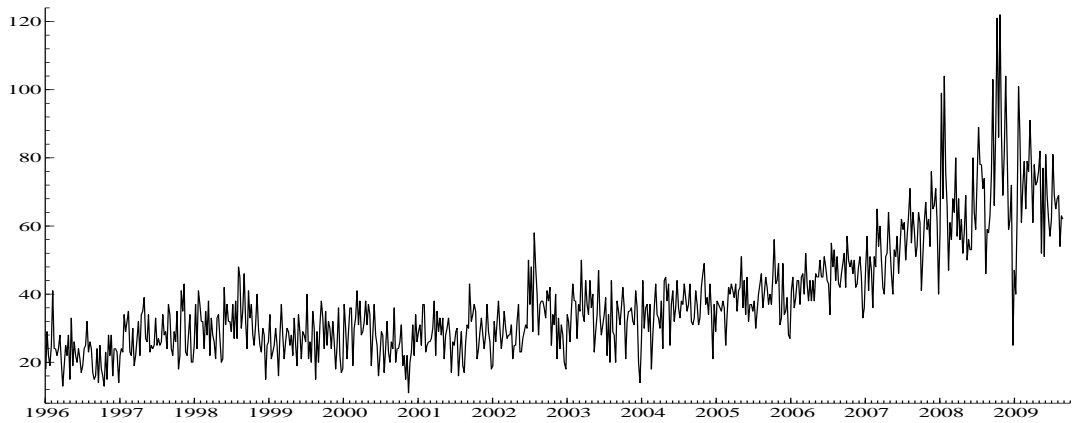
While closed form solutions to (28) exist in a few cases, i.e. the constant volatility Gaussian



(a) Option prices



(b) Implied volatilities



(c) Number of contracts

**Figure 5: Option properties over time**

This figure plots the average weekly prices, implied volatilities, and number of contracts over time for the sample of option data.

case, this is not so for more general specifications of the underlying dynamics. However, even in this case it remains possible to price the options using numerical methods. We choose to use a Monte Carlo approach which is easy to implement in the current setting as the models are simple to simulate from under the risk neutral distribution. For the option example in (28), an estimate of the price is given by

$$\hat{C}_t(T, K) = e^{-r(T-t)} \frac{1}{M} \sum_{j=1}^M \max \left( K - S_T^{(j)}, 0 \right), \quad (29)$$

where  $M$  is the number of paths chosen to be sufficiently large, and where  $S_T^{(j)}$  is the terminal index value simulated under the risk neutral dynamics for the  $j$ th path.<sup>2</sup>

## 4.2 A benchmark - the component NGARCH model

The MN1 Gaussian GARCH and NGARCH models have been used extensively in the literature, see e.g. Christoffersen and Jacobs (2004), Heston and Nandi (2000), and Hsieh and Ritchken (2005). In particular, they have been shown to improve significantly on the performance of the constant volatility BSM model. However, more recently Christoffersen, Jacobs, Ornathanalai, and Wang (2008), inspired by Engle and Lee (1999), propose a component model for option pricing and illustrate its excellent pricing performance compared to the NGARCH model. Further evidence in favor of the component model is provided by Christoffersen, Jacobs, Dorion, and Wang (2010). In light of these findings, we choose this model as an additional benchmark model.

The component volatility model incorporates in a convenient way long range dependence in volatility. For example, in Engle and Lee (1999), one component captures the long run movements in volatility while a second component accounts for the short run volatility movements. In the notation of Christoffersen, Jacobs, Dorion, and Wang (2010) the component NGARCH model,

---

<sup>2</sup>In our application we use  $M = 20,000$ , and as the interest rate we take the prevailing LIBOR rate at the day of pricing.

henceforth CNGARCH, is specified as

$$\begin{aligned} q_t &= \omega + \rho(q_{t-1} - \omega) + \varrho(\varepsilon_{t-1}^2 - \sigma_{t-1}^2 - 2\gamma_1\varepsilon_{t-1}\sigma_{t-1}) \\ \sigma_t^2 &= q_t + \beta(\sigma_{t-1}^2 - q_{t-1}) + \alpha(\varepsilon_{t-1}^2 - \sigma_{t-1}^2 - 2\gamma_1\varepsilon_{t-1}\sigma_{t-1}) \end{aligned} \quad (30)$$

where  $q_t$  can be interpreted as the long term variance component. It can easily be shown that this model has a NGARCH(2,2) type structure and therefore allows for richer variance dynamics than the typical NGARCH(1,1) dynamics. The MN-NGARCH model we propose in this paper is, however, substantially different from the component volatility model. In fact, while each component in the CNGARCH model allows the variance innovations to decay at a different rate there is only one conditional variance process. Therefore, skewness in the conditional return distribution of the component model is only possible when the innovation distribution is skewed. Christoffersen, Jacobs, Dorion, and Wang (2010) estimate the model under a Gaussian and a more general GED distributional assumption and find that the latter does not outperform. Therefore, we fit the model with a Gaussian distribution.

Parameter estimates and standard errors are given in the second and third columns of Table V for the CNGARCH model. As is often found in the literature the persistence of the long term variance component given by  $\rho$  is very close to unity. We therefore also consider a fully persistent or integrated component specification, the ICNGARCH model, which sets  $\rho = 1$ . The table shows that this restriction is not binding given the respective likelihood values. Finally, in terms of statistical in sample fit Tables II, III, and V show that the component NGARCH models are preferred to the MN1NGARCH model as can be seen from the BIC values. However, the MN3NGARCH model still has the smallest BIC.

### 4.3 Dollar losses

We now proceed to compare out-of-sample our MN-NGARCH model to the benchmark models in terms of actual prices. In this section, we report results using the two most used metrics in the

**Table V: Estimation results for the component NGARCH models**

This table reports maximum likelihood estimates for the CNGARCH and ICNGARCH using percentage returns. The date used is from July 2, 1962, to August 26, 2009, for a total of 11,869 observations. Robust standard errors are reported in brackets. The row headed Loglik reports the log likelihood value, and the row headed BIC reports the value of the Bayesian Information Criteria defined as  $2(\log(p) - \text{Loglik})/obs$ , where  $p$  is the number of parameters and  $obs$  is the number of observations. The last four rows report the overall pricing errors for our sample of options.

CNGARCH			ICNGARCH		
Parameter	Estimate	S.E.	Parameter	Estimate	S.E.
$\nu$	3.346	(0.854)	$\nu$	3.081	(0.900)
$\beta$	0.934	(0.012)	$\beta$	0.937	(0.012)
$\alpha$	0.027	(0.006)	$\alpha$	0.027	(0.004)
$\gamma_1$	2.225	(0.471)	$\gamma_1$	2.226	(0.478)
$\omega$	1.490	(0.478)	$\omega$	1.538	(0.431)
$\rho$	0.999	(0.001)	$\rho$	1	-
$\varrho$	0.035	(0.004)	$\varrho$	0.036	(0.004)
$\gamma_2$	0.292	(0.088)	$\gamma_2$	0.306	(0.090)
Loglik	-14328.75			-14329.88	
BIC	2.4148			2.4150	
\$ BIAS	2.900			3.536	
\$ RMSE	9.777			10.644	
ISD BIAS	1.956			2.280	
ISD RMSE	5.464			5.786	

literature, the mean error, or BIAS, and the root mean squared error, or RMSE. The BIAS is calculated as the price predicted by a particular model minus the observed price.

We first examine the overall performance of the models shown in the bottom part of Tables II, III, and V. First of all, the results show that the MN2NGARCH model has the smallest errors using both metrics and outperforms the next best model, the MN3NGARCH model, by 64% in terms of BIAS and 35% in terms of RMSE. The tables also show that, in line with Christoffersen, Jacobs, Ornathanalai, and Wang (2008) and Christoffersen, Jacobs, Dorion, and Wang (2010), the CNGARCH model has superior performance compared to the MN1NGARCH model in terms of RMSE for this sample of option data.<sup>3</sup> However, its performance is still 39% worse than our best mixture model. In terms of the BIAS the mixture model, with an error of 1.145, also outperforms the CNGARCH model, with an error of 2.900.

In Table VI, we further examine the performance of the models across the different maturity categories as defined in Section 4.1. The table shows that the MN2NGARCH model is uniformly the best model in terms of BIAS and its relative performance increases with maturity. The MN2NGARCH model also has excellent performance in terms of RMSE compared to the benchmark models. For example, for the VLT options the RMSE is 12.696 and 22.974 for the MN2NGARCH and CNGARCH models, respectively. For the ST options the RMSE's are similar, that is 5.074 and 4.813 respectively.

We next consider the performance across the different categories of moneyness. Table VII shows that the MN2NGARCH model again performs well across all the categories, though it is dominated by another mixture model, the MN2GARCH model, for the ITM and DITM options. Compared to the mixture models, the benchmarks have errors of at least twice the size. In terms of RMSE, the MN2NGARCH is uniformly the best model and the relative performance is highest for the DOTM

---

<sup>3</sup>The reported errors for the CNGARCH model may appear large, however this is due to our sample period. When using an earlier sample from 1990 to 1995 we obtain results similar to what is found in Christoffersen, Jacobs, Dorion, and Wang (2010).



**Table VI: Dollar losses for different maturities**

This table reports the dollar bias and root mean squared error for each model across different maturities. The maturity categories are defined in Table IV.

<b>Panel A: BIAS</b>						
	VST	ST	MT	LT	VLT	ALL
BSM	-1.317	-2.325	-4.691	-5.745	-6.104	-3.362
MN1GARCH	1.738	2.751	3.827	8.289	12.243	4.193
MN1NGARCH	1.393	1.947	2.473	4.709	8.252	2.771
MN2GARCH	1.391	2.331	2.911	4.869	5.280	2.782
MN2NGARCH	0.993	1.134	0.913	1.163	2.402	1.145
MN3GARCH	1.436	2.162	2.599	4.906	5.975	2.727
MN3NGARCH	1.011	1.271	1.492	3.364	5.841	1.879
CNGARCH	1.235	1.673	2.449	5.719	9.878	2.900
ICNGARCH	1.372	1.944	3.016	7.083	12.590	3.536
<b>Panel B: RMSE</b>						
	VST	ST	MT	LT	VLT	ALL
BSM	5.722	9.051	12.997	16.319	21.356	11.939
MN1GARCH	4.545	7.502	11.768	23.271	27.923	13.616
MN1NGARCH	4.033	6.573	9.942	16.588	20.472	10.418
MN2GARCH	4.265	8.184	13.460	21.831	23.380	13.105
MN2NGARCH	3.254	5.074	7.423	10.231	12.696	7.033
MN3GARCH	4.178	6.823	10.479	18.349	20.093	10.924
MN3NGARCH	3.187	5.041	7.817	15.181	21.391	9.476
CNGARCH	3.066	4.813	7.707	15.599	22.974	9.777
ICNGARCH	3.186	5.034	8.173	16.782	25.695	10.644

**Table VII: Dollar losses for different moneyness**

This table reports the dollar bias and root mean squared error, for each model across different moneyness categories. The moneyness categories are defined in Table IV.

<b>Panel A: BIAS</b>						
	DOTM	OTM	ATM	ITM	DITM	ALL
BSM	-3.953	-1.488	-3.168	-5.698	-4.898	-3.362
MN1GARCH	8.600	3.821	1.907	1.719	1.998	4.193
MN1NGARCH	4.933	2.362	1.464	1.982	2.807	2.771
MN2GARCH	6.631	2.853	0.716	-0.074	0.831	2.782
MN2NGARCH	2.275	1.158	0.349	0.506	1.302	1.145
MN3GARCH	6.344	2.562	0.764	0.414	1.279	2.727
MN3NGARCH	4.324	1.394	0.484	0.707	1.936	1.879
CNGARCH	5.415	2.204	1.578	1.944	2.697	2.900
ICNGARCH	6.251	2.682	2.127	2.704	3.300	3.536
<b>Panel B: RMSE</b>						
	DOTM	OTM	ATM	ITM	DITM	ALL
BSM	10.135	11.172	12.814	14.590	12.762	11.939
MN1GARCH	19.096	11.283	9.942	10.588	12.886	13.616
MN1NGARCH	13.389	8.997	8.388	9.443	11.126	10.418
MN2GARCH	17.916	11.137	9.965	10.234	12.777	13.105
MN2NGARCH	8.223	6.136	6.280	7.399	7.515	7.033
MN3GARCH	14.934	9.077	8.349	8.964	10.593	10.924
MN3NGARCH	12.276	7.766	7.781	8.274	10.443	9.476
CNGARCH	13.049	7.957	7.765	8.405	10.111	9.777
ICNGARCH	14.248	8.607	8.440	9.249	10.881	10.644

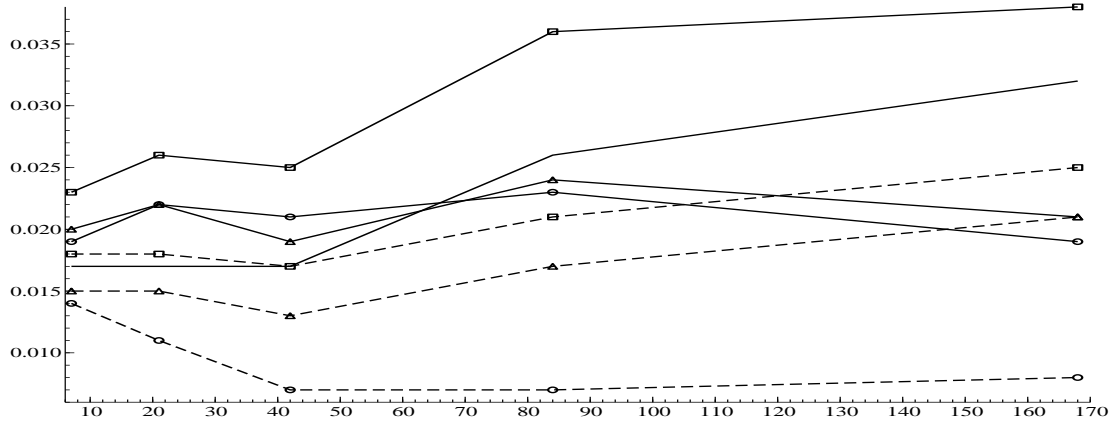
options. For this category of options the RMSE is 8.223 and 13.049 for the MN2NGARCH and CNGARCH models, respectively.

#### 4.4 Implied standard deviation losses

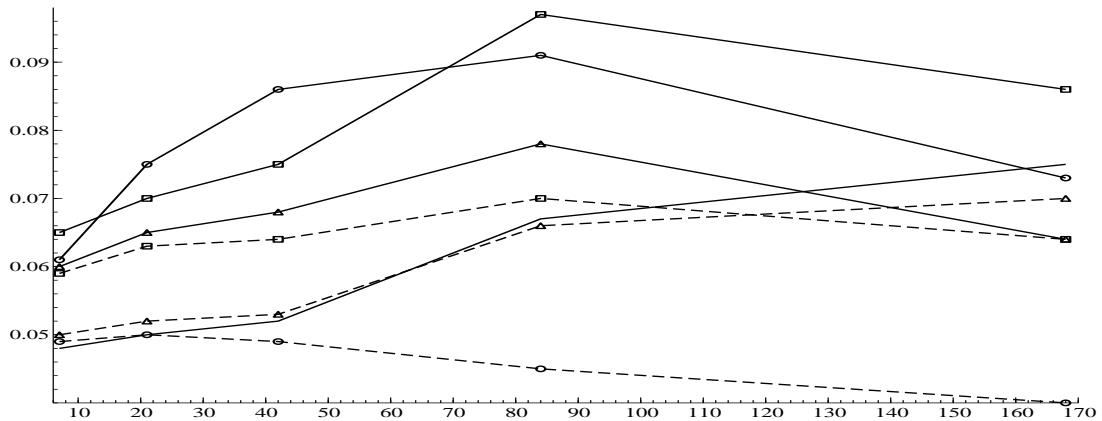
The previous section considers the model's option pricing performance in terms of actual dollar losses. However, an interesting alternative is to compare the performance in terms of ISD, or implied standard deviation, losses. In particular, these losses allow firstly to gauge whether or not the models generate a sufficiently high level of volatility under the risk neutral measure in general. The bottom part of Tables II, III, and V also provide results for the overall performance of the models using ISD losses. Comparing across the tables it is again seen that the MN2NGARCH model has the smallest errors using both metrics and it outperforms the next best models, the MN3NGARCH model for the BIAS by 49% and the CNGARCH model for the RMSE by 13%, respectively. Note that the size of the BIAS for the MN2NGARCH model only amounts to 1.02%, an indication of the solid empirical performance of the mixture model.

Secondly, when considering the ISD losses across maturity we assess which model best fits the volatility term structure. Figure 6 plots the BIAS and the RMSE of the various models considered against maturity and provides evidence on their ability to accommodate the volatility terms structure. The first thing to notice from Figure 6(a) is that the MN2NGARCH model clearly outperforms the other models across maturity in terms of BIAS, and its relative performance is even more striking for higher maturities. In fact, for MT and higher maturities the BIAS for the MN2NGARCH is merely about 0.75%. On the contrary, for the CNGARCH model the bias starts at 1.75% for the ST options but then deteriorates until 3.23% for the VLT category. Figure 6(b) shows that the performance of the MN2NGARCH is equally spectacular in terms of RMSE.

Finally, we analyze which model best captures the implied volatility smirk. Figure 7 plots the BIAS and the RMSE of the various models considered against moneyness. Figure 7(a) shows that mixture models have the lowest BIAS for all categories. In particular, for DOTM, OTM, and



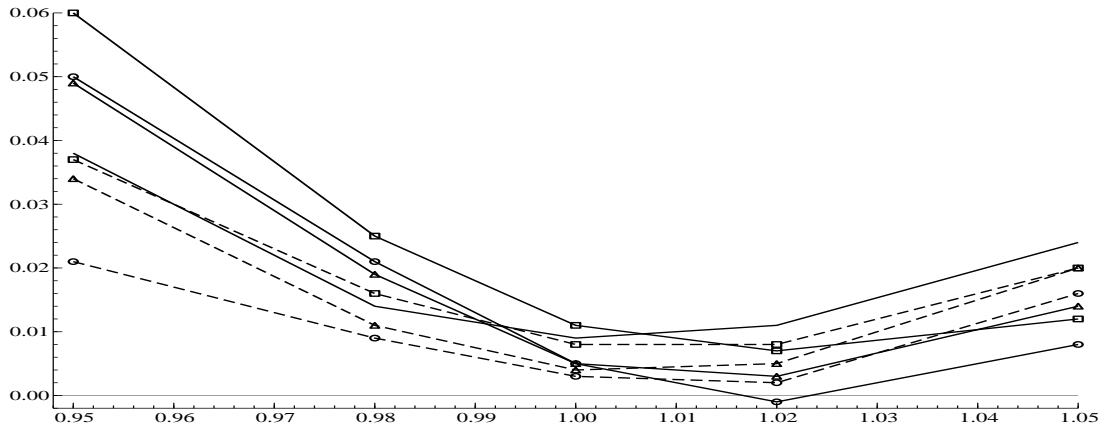
(a) BIAS



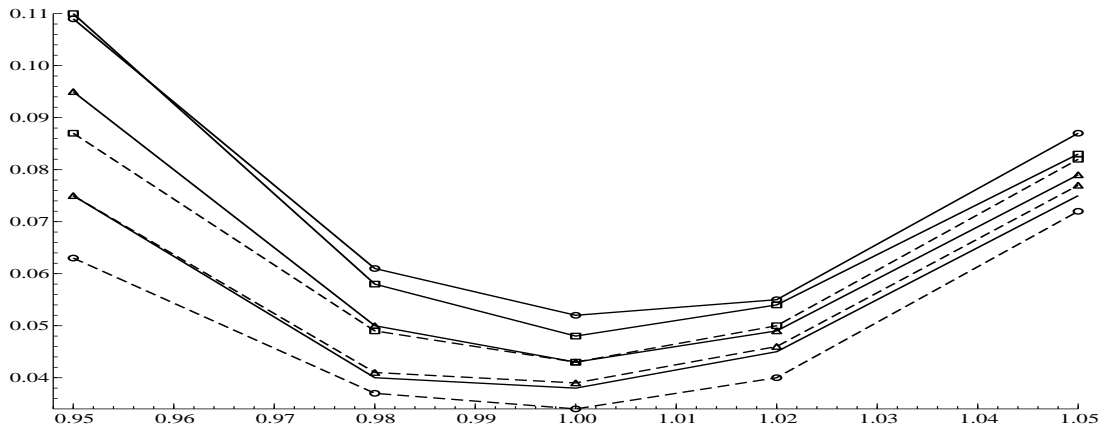
(b) RMSE

**Figure 6: Implied Standard Deviation BIAS and RMSE across maturity**

This figure plots the BIAS and RMSE of the Implied Standard Deviation, or ISD, as a function of maturity. A "□" denotes the MN1 or Gaussian distribution, a "o" the MN2 distribution, and a "Δ" the MN3 distribution. For each distribution a solid line denotes the GARCH specification and a dotted line the NGARCH specification. The solid line without symbols is for the benchmark component NGARCH model.



(a) BIAS



(b) RMSE

**Figure 7: Implied Standard Deviation BIAS and RMSE across moneyness**

This figure plots the BIAS and RMSE of the Implied Standard Deviation, or ISD, as a function of moneyness. A "□" denotes the MN1 or Gaussian distribution, a "o" the MN2 distribution, and a "Δ" the MN3 distribution. For each distribution a solid line denotes the GARCH specification and a dotted line the NGARCH specification. The solid line without symbols is for the benchmark component NGARCH model.

ATM options the MN2NGARCH has the lowest BIAS whereas for the ITM and DITM options the MN2GARCH model has the best performance. Figure 7(b), however, illustrates nicely how the MN2NGARCH model consistently has the lowest RMSE. To conclude, Figure 7 provides solid evidence on the MN2NGARCH models ability to explain the volatility smile in our sample of index options.

## 5 Performance through time and in financial crises

The option data we consider spans a period of approximately 14 years which covers several interesting subperiods: the dot-com bubble and its burst in 2002 (10,182 options), the calm period which followed (10,857 options), and the period of financial turmoil starting in 2008 (6,098 options). We note that each of these periods affected the options market in different ways as can be seen from the time series of average option prices and implied volatilities in Figure 5. Thus, we now examine the performance of the option pricing models through these different market conditions.

Table VIII reports the dollar losses for these different subperiods and clearly shows that the performance changes through time. First of all, when considering the period of the dot-com bubble and burst we see that the mixture models and the component models have similar performance in terms of both BIAS and RMSE. However, for the calm period from 2003-2007 this is not the case and the MN3NGARCH mixture model outperforms all the other models. In fact, for the latter model the BIAS and the RMSE is as small as 0.209 and 2.827, respectively. Note that, for this calm period the errors are relatively small compared to the full sample for all models, and this so especially for the RMSEs.

Finally, we consider the last period of financial turmoil which started in 2008. For this period Table VIII shows that the MN2NGARCH performs extraordinarily well both in terms of the BIAS and the RMSE. In particular, the BIAS for this model is 3.934 compared to 8.414 for the next best performing specification which is the MN1NGARCH model. The CNGARCH has, with an error of 9.876, a slightly worse performance for this subperiod in terms of BIAS although when

**Table VIII: Dollar losses for different subperiods**

This table reports the overall dollar bias and root mean squared error, for each model across different subperiods. The DotCom period is from 1996 to 2002, for a total of 10,182 options, the calm period is from 2003 to 2007, for a total of 10,857 options, and the period of the financial crisis is from 2008 to 2009, for a total of 6,098 options.

<b>Panel A: BIAS</b>				
	DotCom	Calm period	Financial crisis	Full period
BSM	-6.971	4.593	-14.592	-3.362
MN1GARCH	1.601	2.691	11.196	4.193
MN1NGARCH	0.372	1.850	8.414	2.771
MN2GARCH	-0.549	2.129	9.505	2.782
MN2NGARCH	-1.034	1.622	3.934	1.145
MN3GARCH	0.252	1.694	8.700	2.727
MN3NGARCH	-0.539	0.209	8.891	1.879
CNGARCH	0.835	0.919	9.876	2.900
ICNGARCH	1.241	1.314	11.325	3.536
<b>Panel B: RMSE</b>				
	DotCom	Calm period	Financial crisis	Full period
BSM	11.056	8.815	21.163	11.939
MN1GARCH	9.135	5.821	25.009	13.616
MN1NGARCH	7.907	4.779	18.383	10.418
MN2GARCH	7.425	5.152	25.000	13.105
MN2NGARCH	6.441	4.319	10.847	7.033
MN3GARCH	7.756	4.590	19.826	10.924
MN3NGARCH	6.521	2.827	17.730	9.476
CNGARCH	6.387	3.319	18.374	9.777
ICNGARCH	6.648	3.742	20.137	10.644

considering the RMSE its performance is similar to that of the MN1NGARCH model. However, for this metric the MN2NGARCH model again outperforms the benchmark models with an error as low as 10.847 compared to 18.374 for the CNGARCH benchmark. In conclusion, the results for this period provide strong evidence in favor of the mixture model and show that they are capable of adapting to the rapid changes in market conditions in this period which pose a challenge to the models considered here.

## 6 Conclusion

In this paper, we perform option pricing using asymmetric heteroskedastic normal mixture models. We provide details on how to obtain the appropriate risk neutral dynamics and we suggest a feasible way for option pricing within this general framework. When estimated on a long sample of index returns we find support for three distributions in the mixture model. We compare our option pricing model to several benchmarks including component models and find substantial improvements in terms of both dollar losses and implied standard deviation losses for a large sample of 27,137 options on the S&P 500 index over a period of almost 14 years.

In terms of the risk neutralization, we show that the risk neutral dynamics stay within the class of asymmetric heteroskedastic normal mixture models, although the parameters of the distribution are changed. These risk neutral parameters are easily interpreted as providing investors with compensation for specific features of the model like the distribution with an explosive variance process found in our data. Moreover, when comparing the properties of the risk neutral distribution these may differ in a pronounced way from the properties under the original measure used for inference.

We document that our model is capable of generating negative skewness and significant amounts of excess kurtosis. In terms of pricing performance, our results confirm the importance of both these features for our sample of index options. When forecasting out-of-sample substantial improvements are found compared to the benchmark models in terms of dollar losses and the ability to explain the smirk in implied volatilities. Overall, the dollar root mean squared error of the best performing benchmark component model is 39% larger than for the mixture model. When considering the recent financial crisis this difference increases to 69%.

There are several interesting extensions for further research. The most important extension would be to incorporate option prices for inference on the model parameters. Another extension would be to consider Markov switching volatility models in which returns can have a conditional



variance driven by parameters that switch according to a Markov process. Examples are Gray (1996) and Bauwens, Preminger, and Rombouts (2010). Within this framework, it is possible to allow for state dependent unit risk premiums that further drive a wedge between the physical and risk neutral dynamics. However, this framework would require more complex estimation procedures like e.g. Bayesian inference, and the riskneutralization would become less appealing since the conditional Laplace transforms are not yet available.

## References

- AMIN, K. I., AND V. K. NG (1993): “Option Valuation with Systematic Stochastic Volatility,” *Journal of Finance*, 48(3), 881–910.
- BADESCU, A., R. KULPERGER, AND E. LAZAR (2008): “Option Valuation with Normal Mixture GARCH Models,” *Studies in Nonlinear Dynamics and Control*, 12(2), Article 5.
- BAKSHI, G., C. CAO, AND Z. CHEN (1997): “Empirical Performance of Alternative Option Pricing Models,” *Journal of Finance*, 52, 2003–2049.
- (2000): “Pricing and Hedging Long-Term Options,” *Journal of Econometrics*, 94, 277–318.
- BAKSHI, G., P. CARR, AND L. WU (2008): “Stochastic Risk Premiums, Stochastic Skewness in Currency Options, and Stochastic Discount Factors in International Economies,” *Journal of Financial Economics*, 87, 132–156.
- BARNDORFF-NIELSEN, O., AND N. SHEPHARD (2001): “Non-Gaussian Ornstein-Uhlenbeck based models and some of their uses in financial economics,” *Journal of the Royal Statistical Society-Series B*, 63, 167–241.
- BATES, D. (1991): “The Crash of ’87: Was it expected? The Evidence from Option Markets,” *Journal of Finance*, 46, 1009–1044.

- BATES, D. (2000): “Post-’87 Crash Fears in the S&P 500 Futures Option Market,” *Journal of Econometrics*, 94, 181–238.
- BATES, D. (2003): “Empirical Option Pricing: a Retrospection,” *Journal of Econometrics*, 116, 387–404.
- BAUWENS, L., A. PREMINGER, AND J. ROMBOUTS (2010): “Theory and inference for a Markov switching GARCH model,” *Econometrics Journal*, 13, 218–244.
- BAUWENS, L., AND J. ROMBOUTS (2007): “Bayesian Clustering of Many GARCH Models,” *Econometric Reviews*, 26, 365–386.
- BERTHOLON, H., A. MONFORT, AND F. PEGORARO (2006): “Pricing and Inference with Mixtures of Conditional Normal Processes,” *CREST Working Paper 28*.
- (2008): “Econometric Asset Pricing Modelling,” *Journal of Financial Econometrics*, 6(4), 407–458.
- BILLINGSLEY, P. (1995): *Probability and Measure*. Wiley-Interscience, New York.
- BLACK, F., AND M. SCHOLES (1973): “The Pricing of Options and Corporate Liabilities,” *Journal of Political Economy*, 81, 637–654.
- BOOTHE, P., AND D. GLASSMAN (1987): “The Statistical Distribution of Exchange Rates,” *Journal of International Economics*, 22, 297–319.
- BOUADDI, M., AND J. ROMBOUTS (2009): “Mixed Exponential Power Asymmetric Conditional Heteroskedasticity,” *Studies on Nonlinear Dynamics and Econometrics*, 13(3), Article 3.
- CARR, P., H. GEMAN, D. MADAN, AND M. YOR (2003): “Stochastic volatility for Lévy processes,” *Mathematical Finance*, 13, 345–382.
- CARR, P., AND L. WU (2004): “Time-changed Lévy processes and option pricing,” *Journal of Financial Economics*, 71, 113–141.

- CARR, P., AND L. WU (2007): “Stochastic Skew in Currency Options,” *Journal of Financial Economics*, 86, 213–247.
- CHERNOV, M., AND E. GHYSELS (2000): “Towards a Unified Approach to the Joint Estimation of Objective and Risk Neutral Measures for the Purpose of Options Pricing Evaluation,” *Journal of Financial Economics*, 56, 407–458.
- CHRISTOFFERSEN, P., R. ELKAMHI, B. FEUNOU, AND K. JACOBS (2010): “Option Valuation with Conditional Heteroskedasticity and Non-Normality,” *Review of Financial Studies*, 23, 2139–2183.
- CHRISTOFFERSEN, P., S. HESTON, AND K. JACOBS (2006): “Option valuation with Conditional skewness,” *Journal of Econometrics*, 131, 253–284.
- CHRISTOFFERSEN, P., AND K. JACOBS (2004): “Which GARCH Model for Option Valuation?,” *Management Science*, 50(9), 1204–1221.
- CHRISTOFFERSEN, P., K. JACOBS, C. DORION, AND Y. WANG (2010): “Volatility Components, Affine Restrictions, and Non-Normal Innovations,” *Journal of Business and Economic Statistics*, 28(4), 483–502.
- CHRISTOFFERSEN, P., K. JACOBS, AND C. ORNTHANALAI (2008): “Exploring Time-Varying Jump Intensities: Evidence from S&P500 Returns and Options,” *Manuscript*.
- CHRISTOFFERSEN, P., K. JACOBS, C. ORNTHANALAI, AND Y. WANG (2008): “Option Valuation with Long-run and Short-run Volatility Components,” *Journal of Financial Economics*, 90, 272–297.
- DAVIDSON, J. (1997): *Stochastic Limit Theory*. Oxford University Press, Oxford.
- DUAN, J. (1995): “The GARCH Option Pricing Model,” *Mathematical Finance*, 5, 13–32.
- (1999): “Conditionally Fat-Tailed Distributions and the Volatility Smile in Options,” *Mimeo, Hong Kong University of Science and Technology*.

- DUAN, J.-C., P. RITCHKEN, AND Z. SUN (2006): “Approximating GARCH-Jump Models, Jump-Diffusion Process, and Option Pricing,” *Mathematical Finance*, 16(1), 21–52.
- DURHAM, G. (2007): “SV mixture models with application to S&P 500 index returns,” *Journal of Financial Economics*, 85, 822–856.
- ENGLE, R., AND G. LEE (1999): *A Permanent and Transitory Component Model of Stock Return Volatility*, pp. 475–497, *Cointegration, Causality and Forecasting: A Festschrift in Honor of Clive W.J. Granger*. R.F. Engle and H. White (eds), Oxford University Press.
- ENGLE, R., AND G. NG (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.
- FRÜHWIRTH-SCHNATTER, S. (2006): *Finite Mixture and Markov Switching Models*. Springer, New York.
- GARCIA, R., E. GYSELS, AND E. RENAULT (2010): *The Econometrics of Option Pricing* pp. 479–552, *Handbook of Financial Econometrics*, Vol 1. Ait-Sahalia and Hansen (eds.), Elsevier.
- GLOSTEN, L., R. JAGANNATHAN, AND D. RUNKLE (1993): “On the Relation Between the Expected Value and the Volatility of the Nominal Excess Return on Stocks,” *Journal of Finance*, 48(5), 1779–1801.
- GOURIEROUX, C., AND A. MONFORT (2007): “Econometric Specification of Stochastic Discount Factor Models,” *Journal of Econometrics*, 136, 509–530.
- GRAY, S. (1996): “Modeling the conditional distribution of interest rates as a regime-switching process,” *Journal of Financial Economics*, 42, 27–62.

- HAAS, M., S. MITTNIK, AND M. PAOLELLA (2004): “Mixed Normal Conditional Heteroskedasticity,” *Journal of Financial Econometrics*, 2, 211–250.
- HAMILTON, J., T. ZHA, AND D. WAGGONER (2007): “Normalization in Econometrics,” *Econometric Reviews*, 26, 221–252.
- HANSEN, B. (1994): “Autoregressive Conditional Density Estimation,” *International Economic Review*, 35(3), 705–730.
- HARVEY, C., AND A. SIDDIQUE (1999): “Autoregressive Conditional Skewness,” *Journal of Financial and Quantitative Analysis*, 34, 465–487.
- HESTON, S. L. (1993): “A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options,” *Review of Financial Studies*, 6(2), 327–343.
- HESTON, S. L., AND S. NANDI (2000): “A Closed-Form GARCH Option Valuation Model,” *Review of Financial Studies*, 13(3), 585–625.
- HSIEH, K. C., AND P. RITCHKEN (2005): “An Empirical Comparison of GARCH Option Pricing Models,” *Review of Derivatives Research*, 8, 129–150.
- HULL, J., AND A. WHITE (1987): “The Pricing of Options on Assets with Stochastic Volatilities,” *Journal of Finance*, 42(2), 281–300.
- JOHNSON, H., AND D. SHANNO (1987): “Option Pricing When the Variance is Changing,” *Journal of Financial and Quantitative Analysis*, 22(2), 143–151.
- JONDEAU, E., AND M. ROCKINGER (2003): “Conditional Volatility, Skewness, and Kurtosis: Existence, Persistence, and Comovements,” *Journal of Economic Dynamics and Control*, 27, 1699–1737.
- KIM, D., AND S. KON (1994): “Alternative Models for the Conditional Heteroscedasticity of Stock Returns,” *Journal of Business*, 67, 563–598.

- KON, S. (1982): “Models of Stock Returns, a Comparison,” *Journal of Finance*, 39, 147–165.
- MCLACHLAN, G., AND D. PEEL (2000): *Finite Mixture Models*. Wiley Interscience, New York.
- MERTON, R. C. (1973): “Theory of Rational Option Pricing,” *Bell Journal of Economics and Management Science*, 4, 141–183.
- NANDI, S. (1996): “Pricing and Hedging Index Options under Stochastic Volatility: An Empirical Examination,” *Federal Reserve Bank of Atlanta Working Paper 96-9*.
- NELSON, D. (1991): “Conditional Heteroskedasticity In Asset Returns: A New Approach,” *Econometrica*, 59(2), 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.
- PAN, M., K. CHAN, AND C. FOK (1995): “Currency Futures Price Changes: A Two-Piece Mixture of Normals Approach,” *International Review of Economics and Finance*, 4, 69–78.
- ROMBOUTS, J., AND L. STENTOFT (2011): “Multivariate Option Pricing with Time Varying Volatility and Correlations,” *forthcoming in Journal of Banking and Finance*.
- SCOTT, L. O. (1987): “Option Pricing When the Variance Changes Randomly: Theory, Estimation, and an Application,” *Journal of Financial and Quantitative Analysis*, 22(4), 419–438.
- STEIN, E. M., AND J. C. STEIN (1991): “Stock Price Distributions with Stochastic Volatility: An Analytical Approach,” *Review of Financial Studies*, 4(4), 727–752.
- STENTOFT, L. (2005): “Pricing American Options When the Underlying Asset follows GARCH Processes,” *Journal of Empirical Finance*, 12(4), 576–611.
- (2008): “American Option Pricing using GARCH models and the Normal Inverse Gaussian distribution,” *Journal of Financial Econometrics*, 6(4), 540–582.

- TUCKER, A., AND L. POND (1988): “The Probability Distribution of Foreign Exchange Price Changes: Tests of Candidate Processes,” *The Review of Economics and Statistics*, 70, 638–647.
- WIGGINS, J. B. (1987): “Option Values Under Stochastic Volatility: Theory and Empirical Estimates,” *Journal of Financial Economics*, 19, 351–372.
- WONG, C., AND W. LI (2000): “On a Mixture Autoregressive Model,” *Journal of the Royal Statistical Society, Series B*, 62, 95–115.
- WONG, C., AND W. LI (2001): “On a Mixture Autoregressive Conditional Heteroscedastic Model,” *Journal of the American Statistical Association*, 96, 982–995.