



Centre Interuniversitaire sur le Risque,
les Politiques Économiques et l'Emploi

Cahier de recherche/Working Paper **10-20**

Multivariate Option Pricing with Time Varying Volatility and Correlations

Jeroen V.K. Rombouts
Lars Stentoft

Mai/May 2010

Rombouts : Institute of Applied Economics at HEC Montréal, CIRANO, CIRPÉE, Université Catholique de Louvain (CORE), 3000
Chemin de la Côte-Ste-Catherine, Montréal, QC, Canada H3T 2A7

jeroen.rombouts@hec.ca

Stentoft : Department of Finance at HEC Montréal, CIRANO, CIRPÉE, CREATES, 3000 Chemin de la Côte-Ste-Catherine, Montréal,
QC, Canada H3T 2A7

lars.stentoft@hec.ca

The authors thank Jean-Guy Simonato and seminar participants at the CIRPÉE workshop on Applied Financial Time Series for useful comments. The second author thanks CREATES (Center for Research in Econometric Analysis of Time Series, funded by the Danish National Research Foundation) for financial support.

Abstract:

In recent years multivariate models for asset returns have received much attention, in particular this is the case for models with time varying volatility. In this paper we consider models of this class and examine their potential when it comes to option pricing. Specifically, we derive the risk neutral dynamics for a general class of multivariate heteroskedastic models, and we provide a feasible way to price options in this framework. Our framework can be used irrespective of the assumed underlying distribution and dynamics, and it nests several important special cases. We provide an application to options on the minimum of two indices. Our results show that not only is correlation important for these options but so is allowing this correlation to be dynamic. Moreover, we show that for the general model exposure to correlation risk carries an important premium, and when this is neglected option prices are estimated with errors. Finally, we show that when neglecting the non-Gaussian features of the data, option prices are also estimated with large errors.

Keywords: Multivariate risk premia, Option pricing, GARCH models

JEL Classification: C11, C15, C22, G13

1 Introduction

Multivariate models have been used extensively in recent years to model the behavior of financial data. The resulting dynamics are generally more realistic as they take into account the interaction and correlation between such assets. Classical applications of such models have been to model the interaction between stock returns, interest rates and exchange rates, and the framework has been used for prediction of e.g. asset volatilities. However, multivariate models are also potentially important for option pricing. In particular, the most traded options, the index options, are in principle options on an average of multiple assets. Neglecting the correlation between the constituent assets could provide severe mispricings of these claims.

Index options however are not the only example of what is also referred to as basket options. In fact, many other types of options exist where the basket is not an actual index but rather an arbitrarily chosen portfolio of assets. Options written directly on this portfolio provide an efficient method of hedging the risk involved in such exposure. Moreover, in addition to basket options other types of multivariate options exist where the option payoff is on something different than the the average of the underlying assets. For example, a spread option derives its value from the difference between the prices of two or more assets. Spread options can be written on all types of financial products including equities, bonds, currencies, and commodities, and they are especially important in the market for energy futures. Lastly, options exist which pay the best or worst of N assets. This type of option, which is also referred to as a rainbow option, often occur as an element of structured products.

Although the literature on multiple asset options is somewhat limited as the survey by Broadie and Detemple (2004) shows, we are not the first to consider pricing these claims. In fact pricing formulas in constant volatility Gaussian models were proposed by among others Margrabe (1978), Stulz (1982), Johnson (1987), and Boyle, Evninie, and Gibbs (1989) for European claims, and results on American options have been provided by Broadie and Detemple (1997), Detemple, Feng, and Tian (2003), and Villeneuve (1999). However, the assumption of constant volatility and correlations stands in stark contrast to empirical findings, and more recently this assumption has been relaxed and models with time varying volatility

have been used. In particular, models with conditional heteroskedasticity have been used in e.g. Werker, Genest, and van den Goorbergh (2005) and Zhang and Guégan (2008) together with dynamic copulas and in Duan and Pliska (2004) together with co-integrated asset prices.

However, common to all of the previous papers is that they treat the risk neutralization on an individual stock basis. Considering the empirical evidence on the presence of correlation between financial assets the assumption that exposure to this risk carries no price may seem counterintuitive. In fact, this assumption is at odds with the fundamental theories in finance such as the modern portfolio theory and the capital asset pricing model. Moreover, the existing papers all assume a Gaussian model. However, it is often found that financial asset returns are skewed and leptokurtic and therefore this restriction may be an unreasonable assumption.

In this paper we provide a theoretical framework for pricing options in a general multivariate framework. To be specific, building on work by Christoffersen, Elkamhi, Feunou, and Jacobs (2010) we derive the risk neutral dynamics for a general class of multivariate heteroskedastic models, and we provide a feasible way to price options in this framework. Our approach can be used to price any of the multivariate options mentioned above as well as any other type of multivariate claim, and it can be used in any dimension. The framework allows for very general specifications of the risk premia, and it is easily applicable with non-Gaussian models. To our knowledge this is the first paper to consider multivariate option pricing at this level of generality.

A class of volatility models of particular interest are multivariate GARCH models. These models are essentially multivariate extensions of the generalized autoregressive conditional heteroskedastic (GARCH) framework of Engle (1982) and Bollerslev (1986), and they are used extensively in practice because they are relatively easy to estimate. In particular, they have been extensively used to model volatility spillovers and in applications such as conditional CAPM and futures hedging. Examples are respectively Karolyi (1995), Bali (2008) and Moschini and Meyers (2002). Moreover, with this type of models forecasting multivariate volatility is straightforward since the conditional variance matrix is modeled as a function of past innovations. In this paper, we use our theoretical framework together with multivariate GARCH models for option pricing.

In our application, we consider options written on the minimum of the two indices. We first analyze the importance of dynamic correlations on the estimated option prices. Our results show that neglecting correlation has an important effect on the price obtained, and we conclude that models with conditional heteroskedasticity and time varying correlation are important for multivariate option pricing. Secondly, we incorporate a risk premium on correlations and show that this is also important to consider in pricing this type of options. Our results show that neglecting the risk premium on correlations leads to poorly estimated prices of out of the money options. Thus, we conclude that appropriate riskneutralization is important for option pricing in a multivariate setting. Finally, we consider a generalization of the Gaussian model where the conditional distribution is a multivariate mixture of normals. Our results show that neglecting the non-Gaussian features of the data leads to very poor estimates of the option prices. Thus, the final conclusion of the paper is that taking non-Gaussian features into consideration is extremely important when pricing options in a multivariate setting.

It should be noted that an alternative to the discrete time framework used here is models formulated in continuous time. In particular, such models often benefit from allowing elegant option pricing formulas. However, while there is an abundance of univariate continuous time option pricing models, the number of multivariate extensions is limited. This is most likely due to the complexity of such models, and their implementation is quite challenging compared to the discrete time framework. In particular, existing applications often rely on having option data readily available also for model calibration. Contrary to this, the framework used here can be implemented using straightforward estimation techniques and requires only observations on the underlying assets. As many multivariate options are traded over the counter, and therefore the available data is limited, the discrete time framework may in fact provide the only consistent way to price such claims.

The rest of the paper proceeds as follows: In Section 2 we introduce the multivariate heteroskedastic model and we derive the risk neutral dynamics for this class of models. In Section 3 we provide detailed results on some important special cases of our framework. In Section 4 we introduce the data and the particular volatility models to be used in the empirical application in Section 5. Finally, Section 6 contains concluding remarks.

2 Multivariate heteroskedastic asset return dynamics

Letting \mathcal{F}_t denote the information set up to time t , we assume that the underlying return process $R_{j,t} = \ln(S_{j,t}/S_{j,t-1})$ for $j = 1, \dots, N$ can be characterized by

$$R_{j,t} = \mu_{j,t} - \Psi_t(-e_j) + \varepsilon_{j,t}, \quad (1)$$

where $S_{j,t}$ is the price level of asset j on day t . The term $\Psi_t(\cdot)$ denotes the conditional cumulant generating function, and e_j is an N dimensional vector of zeros except for position j where it is 1. Thus, we have that $\ln(E_{t-1}[\exp(\varepsilon_{j,t})]) = \Psi_t(-e_j)$, and it follows that

$$\begin{aligned} E_{t-1}[S_{j,t}/S_{j,t-1}] &= E_{t-1}[\exp(\mu_{j,t} - \Psi_t(-e_j) + \varepsilon_{j,t})] \\ &= \exp(\mu_{j,t}). \end{aligned} \quad (2)$$

The specification in (1) hence implies that $\mu_{j,t}$ can be interpreted as the expected gross rate of return for asset j .

While the above holds irrespective of the assumed dynamics for ε_t , for now we will restrict attention to the case of multivariate heteroskedastic models which can be written as

$$\varepsilon_t = H_t z_t, \quad (3)$$

where H_t is a \mathcal{F}_t -measurable $N \times N$ matrix of full rank, and where z_t is i.i.d. and admits an absolutely continuous N -variate distribution function $P(\cdot)$ with $E[z_t] = 0$ and $E[z_t z_t'] = I_N$. The model implies that the covariance matrix of ε_t conditional on \mathcal{F}_{t-1} is given by $\Sigma_t = H_t H_t'$, which is symmetric. It is of full rank because H_t is of full rank. This also implies that Σ_t is positive definite.

2.1 Specification of an equivalent martingale measure

In the multivariate heteroskedastic model markets are incomplete and hence there is no unique way to derive the equivalent martingale measure, or EMM, needed for option pricing. However, by making additional assumptions it becomes possible to derive such a unique characterization. Our additional assumption is that the relevant measure can be derived from an exponential affine Radon-Nikodym derivative of the following form

$$\frac{dQ}{dP} \Big|_{\mathcal{F}_t} = \exp \left(- \sum_{i=1}^t (\nu_i' \varepsilon_i + \Psi_i(\nu_i)) \right), \quad (4)$$

where ν_i is a N -dimensional vector sequence. The specification in (4) may be seen as a multivariate generalization of the suggested Radon-Nikodym derivative from Christoffersen, Elkamhi, Feunou, and Jacobs (2010).

We now proceed to show that it is possible to derive an EMM from the above. To do this, we proceed in two steps. The first step is to show that (4) is in fact a Radon-Nikodym derivative. The following lemma provides the details.

Lemma 1 $\frac{dQ}{dP} \Big| \mathcal{F}_t = \exp \left(- \sum_{i=1}^t (\nu'_i \varepsilon_i + \Psi_i(\nu_i)) \right)$ is a Radon-Nikodym derivative.

Proof We first of all need to show that $\frac{dQ}{dP} \Big| \mathcal{F}_t > 0$. However, due to the exponential affine structure this is immediate. Next we need to show that $E_0^P \left[\frac{dQ}{dP} \Big| \mathcal{F}_t \right] = 1$. However, using the law of iterated expectations we obtain the following

$$\begin{aligned}
E_0^P \left[\frac{dQ}{dP} \Big| \mathcal{F}_t \right] &= E_0^P \left[\exp \left(- \sum_{i=1}^t (\nu'_i \varepsilon_i + \Psi_i(\nu_i)) \right) \right] \\
&= E_0^P \left[E_1^P \dots E_{t-1}^P \exp \left(- \sum_{i=1}^t \nu'_i \varepsilon_i - \sum_{i=1}^t \Psi_i(\nu_i) \right) \right] \\
&= E_0^P \left[E_1^P \dots E_{t-2}^P \exp \left(- \sum_{i=1}^{t-1} \nu'_i \varepsilon_i - \sum_{i=1}^t \Psi_i(\nu_i) \right) E_{t-1}^P \exp(-\nu'_t \varepsilon_t) \right] \\
&= E_0^P \left[E_1^P \dots E_{t-2}^P \exp \left(- \sum_{i=1}^{t-1} \nu'_i \varepsilon_i - \sum_{i=1}^t \Psi_i(\nu_i) \right) \Psi_t(\nu_t) \right] \\
&= E_0^P \left[E_1^P \dots E_{t-2}^P \exp \left(- \sum_{i=1}^{t-1} \nu'_i \varepsilon_i - \sum_{i=1}^{t-1} \Psi_i(\nu_i) \right) \right].
\end{aligned}$$

Iterating on this yields the required result since

$$\begin{aligned}
E_0^P \left[\frac{dQ}{dP} \Big| \mathcal{F}_t \right] &= E_0^P [\exp(\nu'_1 \varepsilon_1 + \Psi_1(\nu_1))] \\
&= \exp(-\Psi_1(\nu_1)) \exp(\Psi_1(\nu_1)) = 1.
\end{aligned}$$

This completes the proof. ■

Having completed the first step, the second step is to show that the resulting measure is in fact an EMM, i.e. that under the transformed measure discounted returns are martingales. The following proposition shows that this is the case provided the vector sequence ν_t satisfy a specific set of conditions.

Proposition 1 The probability measure Q defined by the Radon-Nikodym derivative in (4)

is an EMM if and only if

$$0 = \Psi_t(\nu_t - e_j) - \Psi_t(\nu_t) - \Psi_t(-e_j) + \mu_{j,t} - r_t, \quad (5)$$

for all $j = 1, \dots, N$.

Proof We need to show that under Q the expected asset returns equals the risk free interest rate, i.e. that discounted asset prices are martingales. To do this we proceed by considering asset j for which we have to show that $E_{t-1}^Q \left[\frac{S_{j,t}}{S_{j,t-1}} \right] = \exp(r_t)$, or equivalently that $E_{t-1}^Q \left[\frac{S_{j,t}}{S_{j,t-1}} \exp(-r_t) \right] = 1$. Using (4) we obtain

$$\begin{aligned} E_{t-1}^Q \left[\frac{S_{j,t}}{S_{j,t-1}} \exp(-r_t) \right] &= E_{t-1}^P \left[\left(\frac{\frac{dQ}{dP} | \mathcal{F}_t}{\frac{dQ}{dP} | \mathcal{F}_{t-1}} \right) \frac{S_{j,t}}{S_{j,t-1}} \exp(-r_t) \right] \\ &= E_{t-1}^P \left[\exp(-\nu_t' \varepsilon_t - \Psi_t(\nu_t)) \exp(\mu_{j,t} - \Psi_t(-e_j) + \varepsilon_{j,t}) \exp(-r_t) \right] \\ &= \exp(-\Psi_t(\nu_t) + \mu_{j,t} - \Psi_t(-e_j) - r_t) E_{t-1}^P \left[\exp((e_j - \nu_t)' \varepsilon_t) \right] \\ &= \exp(-\Psi_t(\nu_t) + \mu_{j,t} - \Psi_t(-e_j) - r_t + \Psi_t(\nu_t - e_j)). \end{aligned}$$

Thus, it follows that if we construct the probability measure Q from the proposed Radon-Nikodym derivative and ensure that

$$0 = \Psi_t(\nu_t - e_j) - \Psi_t(\nu_t) - \Psi_t(-e_j) + \mu_{j,t} - r_t,$$

for all $j = 1, \dots, N$, then the probability measure is in fact an EMM since it makes discounted asset prices martingales. This completes the proof. ■

2.2 The risk neutral dynamics

Having derived the appropriate restrictions on the sequence ν_t , conditional on the specification of the Radon-Nikodym derivative, we can now derive the risk neutral dynamics. In order to do this, the first step is to derive the conditional moment generating function under Q . The following lemma provides the details.

Lemma 2 The conditional moment generating function under Q is given by

$$E_{t-1}^Q [\exp(-u' \varepsilon_t)] = \exp(\Psi_t(\nu_t + u) - \Psi_t(\nu_t)).$$

Proof The proof follows easily by substituting the values from the specification of the Radon-Nikodym derivative and rearranging as follows:

$$\begin{aligned}
E_{t-1}^Q [\exp(-u'\varepsilon_t)] &= E_{t-1}^P \left[\left(\frac{\frac{dQ}{dP} | \mathcal{F}_t}{\frac{dQ}{dP} | \mathcal{F}_{t-1}} \right) \exp(-u'\varepsilon_t) \right] \\
&= E_{t-1}^P [\exp(-\nu_t'\varepsilon_t - \Psi_t(\nu_t)) \exp(-u'\varepsilon_t)] \\
&= E_{t-1}^P [\exp(-(\nu_t + u)'\varepsilon_t) - \Psi_t(\nu_t)] \\
&= \exp(\Psi_t(\nu_t + u) - \Psi_t(\nu_t)).
\end{aligned}$$

This completes the proof. ■

Next, using the above lemma it follows easily that under the risk neutral measure the conditional cumulant generating function of ε_t is given by

$$\Psi_t^Q(u) = \Psi_t(\nu_t + u) - \Psi_t(\nu_t). \quad (6)$$

Using the Inversion Theorem (see for example Billingsley (1995, Theorem 26.2) or Davidson (1997, Theorem 11.12)) the conditional cumulant generating function can be used to obtain the distribution under Q , provided that this is recognized as belonging to a known family of distributions.

As a verification we note that for any choice of ν_t we can substitute (5) and (6) into the mean equation in (1). Doing so, we obtain the following specification of the risk neutral dynamics

$$R_{j,t}^Q = r_t - \Psi_t^Q(-e_j) + \varepsilon_{j,t}^Q, \quad (7)$$

where the superscript Q indicates that the variables are considered under the risk neutral distribution. From (7), we can calculate the gross rate of return on asset j under Q as

$$\begin{aligned}
E_{t-1}^Q [S_{j,t}/S_{j,t-1}] &= E_{t-1}^Q \left[\exp \left(r_t - \Psi_t^Q(-e_j) + \varepsilon_{j,t}^Q \right) \right] \\
&= \exp(r_t).
\end{aligned} \quad (8)$$

Thus, it equals the risk free interest rate as required.

With the derived EMM claims may now be priced as the expected value, under the EMM, of their future cash flows discounted using the risk free interest rate. For example, the price of

a European option at time t with maturity T and terminal payoff function $\varrho_\alpha(S_{1,T}, \dots, S_{N,T})$, α being the set of parameters like e.g. the strike price, can be computed as the following discounted expectation

$$C_t(T, \varrho_\alpha) = e^{-r(T-t)} \int_0^\infty \dots \int_0^\infty \varrho_\alpha(S_{1,T}, \dots, S_{N,T}) f^Q(S_{1,T}, \dots, S_{N,T}) dS_{1,T} \dots dS_{N,T}, \quad (9)$$

where $f^Q(S_{1,T}, \dots, S_{N,T})$ is the density of the underlying asset prices at expiration under the EMM.

2.3 Feasible option pricing

Equations (5) and (6) above completely characterize the risk neutral process and hence this is, in fact, all that is needed for option pricing purposes. In particular, equation (5) provides the link between the distributional properties under the original measure P and the riskneutral distribution under Q from the sequence ν_t , and equation (6) then 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 the sequence and the dynamics explicitly given the choice of dynamics and underlying distribution. However, unless the cumulant generating function is of a simple form this is potentially problematic since it involves solving (5) in order to obtain the actual sequence of ν_t . The multivariate Gaussian distribution we consider below is a special case where it is in fact possible to solve (5) directly for any given specification of $\mu_{j,t}$. However, it is obvious that this may not always be the case for a general choice of distribution.

An alternative way to proceed is to realize that we may equally well use equation (5) to provide a link between a particular choice of sequence ν_t and the distributional properties under the original measure P instead. Say, we could impose constant values for ν_t through time or we could let ν_t depend on the level of the conditional variances and covariances. For any choice we can then derive the restriction on the gross rate of return by rearranging (5) as

$$\mu_{j,t} = r_t - \Psi_t(\nu_t - e_j) + \Psi_t(\nu_t) + \Psi_t(-e_j). \quad (10)$$

From (10), we note that for any choice of ν_t a closed form expression exist for $\mu_{j,t}$ given that the cumulant generating function exists. Substituting this into the return equation in (1)

we obtain

$$R_{j,t} = r_t - \Psi_t(\nu_t - e_j) + \Psi_t(\nu_t) + \varepsilon_{j,t}. \quad (11)$$

Since this only depends on known parameters it can be used for estimation directly and it thus provides a method for option pricing.

It should be noted that for most choices of distribution and hence cumulant generating functions evaluation of the expression in (11) is easy and hence estimation based on this specification is not only theoretically possible but also feasible in practice. Moreover using the estimated parameters, option pricing can be easily performed since this again depends only on parameters identified from estimating the model on historical data on the underlying assets. Hence no procedure of calibrating the model to option data is required. In particular, if the options under consideration are European style, price estimates can be obtained very quickly using a simulation approach as averages of predicted payoffs at maturity using the risk neutral dynamics.

3 Important special cases

In this section we explore several special cases of our general framework which are of particular interest and which will be used in our application in Section 5. The first example is the case of a multivariate Gaussian distribution, of which the bivariate case is of particular interest. The second example is a generalization of this benchmark case where the innovations are from a multivariate mixture of normals.

3.1 The multivariate Gaussian distribution

When ε_t is multivariate Gaussian distributed, the conditional cumulant generating function is given by

$$\Psi_t(u) = \frac{1}{2}u'\Sigma_t u. \quad (12)$$

This very simple form means that option pricing is straightforward even in the case with conditional heteroskedasticity.

First of all, under the assumption of a multivariate Gaussian distribution we may solve (5) directly for any specification of $\mu_{j,t}$. This can be done by substituting (12) into (5) upon which we obtain the following

$$\begin{aligned}
0 &= \Psi_t(\nu_t - e_j) - \Psi_t(\nu_t) - \Psi_t(-e_j) + \mu_{j,t} - r_t \\
&= \frac{1}{2} [(\nu_t - e_j)' \Sigma_t (\nu_t - e_j) - \nu_t' \Sigma_t \nu_t - e_j' \Sigma_t e_j] + \mu_{j,t} - r_t \\
&= -e_j' \Sigma_t \nu_t + \mu_{j,t} - r_t,
\end{aligned}$$

for all $j = 1, \dots, N$. Solving for ν_t and writing this in vector form we obtain

$$\nu_t = \Sigma_t^{-1}(\mu_t - r_t), \quad (13)$$

where μ_t and r_t are the corresponding vectors of the gross rate of return and the interest rate, respectively. Thus, for a given choice of μ_t it is possible to solve explicitly for ν_t .

Secondly, under the multivariate Gaussian assumption the risk neutral dynamics needed for option pricing to proceed are easily obtained. In particular, as already mentioned these may be derived from the cumulant generating function under P using the specification in (6). Substituting (12) into (6) we obtain

$$\begin{aligned}
\Psi_t^Q(u) &= \Psi_t(\nu_t + u) - \Psi_t(\nu_t) \\
&= \frac{1}{2} (\nu_t + u)' \Sigma_t (\nu_t + u) - \frac{1}{2} \nu_t' \Sigma_t \nu_t \\
&= u' \Sigma_t \nu_t + \frac{1}{2} u' \Sigma_t u.
\end{aligned} \quad (14)$$

Moreover, using the expression in (13) the following characterization of the distribution under Q is obtained

$$\Psi_t^Q(u) = u'(\mu_t - r_t) + \frac{1}{2} u' \Sigma_t u. \quad (15)$$

Thus, it follows that the risk neutral dynamics remain Gaussian although with a shifted mean. The shift in the mean is exactly what is required to compensate investors for the risk associated with investing in the underlying risky assets.

The general formulations above show that it is possible to choose the specification of μ_t freely in the multivariate Gaussian framework since the ν_t sequence can always be obtained

from (13). Alternatively, one could use the technique of implying the conditional mean μ_t for a given choice of ν_t using (10). However, unless one chooses the sequence of ν_t in a clever way this may lead to a complicated specification of μ_t , which could be difficult to estimate.

3.1.1 A simple mean specification

We start by considering a simple specification in which the conditional mean is given by

$$\mu_t = r_t + \text{diag}\Sigma_t\lambda, \quad (16)$$

where $\text{diag}\Sigma_t$ is the matrix containing only the diagonal elements of Σ_t . In this specification the compensation for asset j only depends on the asset's own variance, and λ is readily interpreted as the unit risk premium. Compared to a more general specification this assumption has the effect of making estimation easier. In particular, conditional on the choice of dynamics and underlying distribution, estimation can sometimes be done for each asset individually.

The particular choice of conditional mean also implies that the risk neutral dynamics are simple. To see this, substitute (16) into (15) to obtain

$$\Psi_t^Q(u) = u'\text{diag}\Sigma_t\lambda + \frac{1}{2}u'\Sigma_t u. \quad (17)$$

Thus, for the particular choice the risk neutral mean equals $-\text{diag}\Sigma_t\lambda$, which is readily interpreted as the compensation an investor requires for holding the risky assets with λ interpreted as the unit price of risk.

The particular specification used in (16) is a straightforward generalization of the univariate specification used in Heston and Nandi (2000). Note however, that we could equally well let the unit risk premium be proportional to the volatility or to be constant also without any loss of generality. The first of these specifications, which corresponds to setting $\mu_{i,t} = r_t + \sigma_{i,t}\lambda_i$, is a straightforward generalization of the univariate specification used in Duan (1995). This also corresponds to the approach used in e.g. Werker, Genest, and van den Goorbergh (2005) and Zhang and Guégan (2008) in a multivariate context.

3.1.2 A general mean specification in the bivariate case

The above specification corresponds in terms of the risk neutralization to what is used in the existing bivariate option pricing literature, and is a special case of our more general framework. In particular, the above specification eliminates any consideration of the correlation between assets and the risk associated with exposure to the risk stemming from this. This may seem counterintuitive considering that the models used often allow for time varying correlations or covariances.

Although various specifications are possible, we will consider the case where both volatility and correlation risk carries a premium. For notational convenience we limit the exposition to the bivariate case, but all the results easily generalize to the N -dimensional situation. To be specific, we consider the following specification of the conditional mean for the i 't asset

$$\mu_{i,t} = r_t + \sigma_i^2 \lambda_{i,i} + \rho_{i,j} \lambda_{i,j}. \quad (18)$$

Substituting (18) into (15) it is easily seen that with this more general mean specification the risk neutral distribution is again Gaussian but with a mean which is changed in order to compensate investors for holding the asset which is exposed to the two sources of risk. Note that it is also possible to include σ_j^2 in the above equation for asset i , although this makes less sense from a theoretical point of view.

It should be noted that while the case with correlation risk is easily accommodated using our framework this is not the case in the existing literature. In particular, the approach based on riskneutralization on an equation by equation basis cannot be used with a more general mean specification. Whether these extra terms are important is, of course, essentially an empirical question.

3.2 The multivariate mixed normal distribution

As convenient as it is, the Gaussian distribution on the innovations is often not suitable because it is symmetric and has a kurtosis equal to three. In particular, this is the case for financial asset returns which are often found to be skewed and leptokurtic and it is therefore desirable to use a more flexible distribution. One alternative is to use finite mixtures of

Gaussian distributions which are extremely flexible and which are becoming a standard tool in financial econometrics. This choice is attractive because of the parsimonious flexibility it provides in the specification of the distribution of the underlying random variable, which has a semiparametric flavor. Recently, finite mixtures have been used for option pricing, although in a univariate framework, by Badescu, Kulperger, and Lazar (2008), Bertholon, Monfort, and Pegoraro (2006), Durham (2007), and Rombouts and Stentoft (2009).

In the multivariate mixed normal framework, the conditional distribution of ε_t is a combination of K densities

$$F(\varepsilon_t | \mathcal{F}_{t-1}) = \sum_{k=1}^K \pi_k \Phi(\mu_k, \Sigma_{k,t}), \quad (19)$$

where $\Phi(\cdot)$ is the standard multivariate Gaussian distribution. For each t in this finite mixture framework, ε_t is drawn from one of the K conditional distributions with probabilities π_1, \dots, π_K . Consequently, the parameter π_k is restricted to be positive for all k and $\sum_{k=1}^K \pi_k = 1$, which is imposed by setting $\pi_K = 1 - \sum_{k=1}^{K-1} \pi_k$. The zero mean assumption on ε_t is ensured by the restriction

$$\mu_K = - \sum_{k=1}^{K-1} \frac{\pi_k \mu_k}{\pi_K}. \quad (20)$$

For more on the theoretical properties of the multivariate mixed normal distribution see Bauwens, Hafner, and Rombouts (2007).

When ε_t has a multivariate mixed normal distribution, the conditional cumulant generating function is given by

$$\Psi_t(u) = \ln \left(\sum_{k=1}^K \pi_i \exp \left(-u' \mu_k + \frac{u' \Sigma_{k,t} u}{2} \right) \right). \quad (21)$$

That is, the conditional moment generating function is a convex combination of multivariate Gaussian moment generating functions. The fact that the expression in (21) remains simple is convenient when it comes to option pricing since this is needed to obtain the appropriate risk neutral dynamics. To be specific, the conditional cumulant generation function of ε_t under the risk neutral measure Q is easily obtained by substituting (21) into (6) which

yields

$$\begin{aligned}
\Psi_t^Q(u) &= \Psi_t(\nu_t + u) - \Psi_t(\nu_t) \\
&= \ln \left(\frac{\sum_{k=1}^K \pi_i \exp \left(-(\nu_t + u)' \mu_k + \frac{(\nu_t + u)' \Sigma_{k,t} (\nu_t + u)}{2} \right)}{\sum_{k=1}^K \pi_i \exp \left(-\nu_t' \mu_k + \frac{\nu_t' \Sigma_{k,t} \nu_t}{2} \right)} \right) \\
&= \ln \left(\sum_{k=1}^K \pi_{i,t}^* \exp \left(-u' \mu_{k,t}^* + \frac{u' \Sigma_{k,t} u}{2} \right) \right), \tag{22}
\end{aligned}$$

where

$$\mu_{k,t}^* = \mu_k - \Sigma_{k,t} \nu_t, \tag{23}$$

and

$$\pi_{k,t}^* = \frac{\pi_k \exp \left(-\nu_t' \mu_k + \frac{\nu_t' \Sigma_{k,t} \nu_t}{2} \right)}{\sum_{k=1}^K \pi_k \exp \left(-\nu_t' \mu_k + \frac{\nu_t' \Sigma_{k,t} \nu_t}{2} \right)}, \tag{24}$$

for $k = 1, \dots, K$. Thus, the risk neutral distribution of ε_t remains within the family of multivariate normal mixtures.

With respect to the risk neutral means from (23), it is immediately seen that the correction is very similar to what is obtained with the Gaussian model, where the mean of ε_t under Q is equal to $r_t - \mu_t = -\Sigma_t \nu_t$. The intuition behind this is the following: If variance and covariance risk carries a positive premium, i.e. ν_t is positive, then in the risk neutral world the mean of the innovations is shifted downwards to compensate for this.

4 Data and volatility models

In this section we introduce the return data, the particular choice of variance dynamics which will be used, and we explain how options can be priced in the empirical application of the multivariate heteroskedastic model. We note though that the framework introduced in this paper is not restricted to the particular choice of data, dynamics, or of options to be priced.

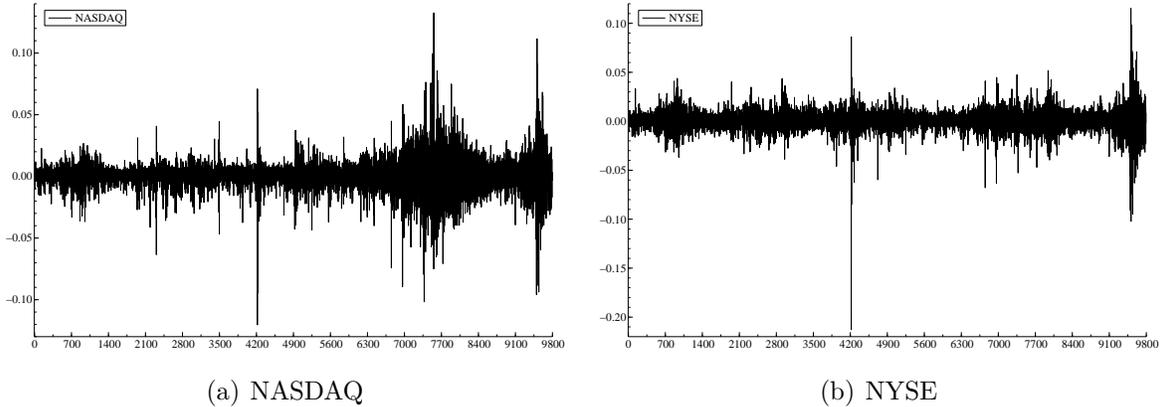


Figure 1: Return time series

This figure plots the time series of returns for the NASDAQ and NYSE indices for the period from February 5, 1971, through November 27, 2009.

4.1 Data

In this paper we use data on the NASDAQ and NYSE composite indices, which are among the most important financial indices in the world. Both of these indices are market-value weighted, which means that they are easy to replicate and therefore potentially investable. Options written directly on these indices are therefore potentially important for hedging as well as for speculative purposes. However, the two indices reflect the performance of two quite distinct markets. In particular, NASDAQ is heavy on technology stocks, such as Microsoft and Intel, whereas NYSE contains a large fraction of mostly well established large industrial companies, such as General Electric and Ford. Thus, when modeling the dynamics of these two time series a multivariate model is clearly needed.

Figure 1 plots the time series of returns for the two indices and clearly shows that they share similar properties. In particular, for both time series the classical pattern of time varying volatility and volatility clustering are found. However, the figure also shows that there are differences between the NASDAQ and the NYSE when it comes to the effect of two of the most important events in the last 25 years, the crash in 1987 and the dot-com bubble. To be specific, looking at the figure we see that the NYSE was much more affected by the crash in 1987, whereas the dot-com bubble effected the NASDAQ much more than the broader based NYSE index. The most recent financial crises, on the other hand, appears

Table I: Descriptive statistics for the NASDAQ and NYSE indices

This table provides descriptive statistics for the two indices using data from February 5, 1971, through November 27, 2009.

Index	min	mean	max	stddev	skewn	kurt
NASDAQ	-0.12043	0.000313	0.13255	0.012608	-0.27920	13.268
NYSE	-0.21286	0.000258	0.11526	0.010391	-1.15310	31.186

to have had a similar effect on both indices.

Table I provides descriptive statistics for first four moments of the two indices. The table shows that the two series are negatively skewed and very leptokurtic. Again this is in line with previous research which has documented these properties for other financial asset returns. However, the table also shows that while the means and variances are of similar magnitude this is not so for the third and fourth moments. In particular, when comparing the two time series the table shows that NYSE is much more skewed and leptokurtic than NASDAQ. Again this is due to the larger effect of the crash in 1987 on the NYSE index which caused a negative return of 21%.

4.2 Multivariate heteroskedastic models

Considering Figure 1 and Table I it is obvious that multivariate heteroskedastic models provide an interesting framework for this type of series. In this paper we chose to specify the conditional variance matrix Σ_t by multivariate GARCH models for which estimation is easily done by maximum likelihood. Although many types of models exist, see Bauwens, Laurent, and Rombouts (2006) and Silvennoinen and Terasvirta (2008) for surveys, they all deliver very similar out-of-sample forecasts. Therefore, we will limit our attention to two well known models that we define next.

First, a general formulation of Σ_t has been proposed by Bollerslev, Engle, and Wooldridge (1988). In the general VEC model, each element of Σ_t is a linear function of the lagged squared errors and cross products of errors and lagged values of the elements of Σ_t . To be specific, the dynamics are given by

$$\sigma_t = c + A \eta_{t-1} + B \sigma_{t-1}, \quad (25)$$

where $\sigma_t = \text{vech}(\Sigma_t)$, $\eta_t = \text{vech}(\epsilon_t \epsilon_t')$ with ϵ_t defined in (3), and $\text{vech}(\cdot)$ denotes the operator that stacks the lower triangular portion of a $N \times N$ matrix as a $N(N+1)/2 \times 1$ vector. A and B are square parameter matrices of order $(N+1)N/2$ and c is a $(N+1)N/2 \times 1$ parameter vector. Given the large number of parameters, this general model is typically estimated only when $N = 2$.

Second, the dynamic conditional correlation model (DCC) in the formulation of Engle (2002) is defined as

$$\Sigma_t = D_t^{1/2} R_t D_t^{1/2} \quad (26)$$

$$D_t = \text{diag}(\sigma_{11,t}^2, \dots, \sigma_{NN,t}^2) \quad (27)$$

$$R_t = \text{diag}(q_{11,t}^{-1/2} \dots q_{NN,t}^{-1/2}) Q_t \text{diag}(q_{11,t}^{-1/2} \dots q_{NN,t}^{-1/2}) \quad (28)$$

$$Q_t = (1 - \theta_1 - \theta_2) \bar{Q} + \theta_1 u_{t-1} u_{t-1}' + \theta_2 Q_{t-1}, \quad (29)$$

where $u_{i,t} = \epsilon_{i,t} / \sigma_{i,t}$ define the devolatilized innovations and \bar{Q} is the $N \times N$ unconditional variance matrix of u_t . The DCC model is a generalization of the constant conditional correlation (CCC) model of Bollerslev (1990) which fixes $R_t = R$. In our application we use the GARCH model of Bollerslev (1986) for $\sigma_{ii,t}^2$.

Estimation of both multivariate GARCH models is done by maximum likelihood using the full sample. Under the Gaussian innovation assumption the sample log-likelihood is given, up to a constant, by

$$-\frac{1}{2} \sum_{t=1}^T \log |\Sigma_t| - \frac{1}{2} \sum_{t=1}^T \epsilon_t' \Sigma_t^{-1} \epsilon_t, \quad (30)$$

where T denotes the sample size. We maximize numerically for the parameters in ϵ_t and Σ_t . The particular structure of the DCC model in (26) allows to split the loglikelihood function in two parts making estimation easier from a numerical viewpoint. The first part depends only on D_t implying that estimation of each of the conditional variances can be done separately. Given the conditional variances, the second part involves the conditional correlation process for which only θ_1 and θ_2 have to be estimated (the matrix \bar{Q} is fitted in advance using the devolatilized innovations). This two step approach is particularly convenient when N is large. For the mixture model, which allows incorporating departures from normality, the estimation

of the parameters is necessarily done in one step given the more involved likelihood which is directly implied by (19).

4.3 Option pricing methodology

In this paper we consider European put options on the minimum of two indices. The theoretical value at time t of this option with a strike price equal to K and maturity in T days is

$$C_t(T, K) = e^{-r(T-t)} E_t^Q[\max[K - \min(S_{1,T}, S_{2,T}), 0]], \quad (31)$$

where E_t^Q means that the expectation is taken under the risk neutral measure Q as specified by (9) and where $S_{1,T}$ and $S_{2,T}$ are the values at time T of the two indices, respectively. While closed form solutions to (31) do exist in a few cases, i.e. the constant volatility Gaussian case, this is not so for more general specifications of the underlying dynamics. Even in this case, however, it remains possible to price the options using numerical methods together with the dynamics derived in this paper.

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 European option example in (31), an estimate of the price is given by

$$\hat{C}_t(T, K) = e^{-r(T-t)} \frac{1}{M} \sum_{j=1}^M \max\left(K - \min(S_{1,T}^{(j)}, S_{2,T}^{(j)}), 0\right), \quad (32)$$

where $S_{1,T}^{(j)}$ and $S_{2,T}^{(j)}$ are the terminal index values simulated under the risk neutral dynamics for M paths. In our empirical application we use $M = 20,000$ paths, and for the time being we do not use any variance reduction techniques such as the ones suggested in e.g. Barraquand (1995) or Duan and Simonato (1998). More generally, the advantages of this approach is that it can be used in any dimension. In addition, with a Monte Carlo approach American options could also be priced using e.g. the Least-Squares Monte Carlo method of Longstaff and Schwartz (2001), for which the mathematical foundation was provided in Stentoft (2004).

The bivariate options we consider are, however, not readily traded at exchanges, and therefore we consider a set of artificial options. To be specific, we consider options with 5

different maturities: 21, 42, 63, 126 and 252 days to maturity. This corresponds to options with 1, 2, 3, 6, and 12 months to maturity. For each of these maturities we consider 9 different strike prices from 80 to 120 in increments of 5. Moreover, in order to put the two indices on equal footing it will be assumed that the starting values are 100 for both indices. Doing this has the effect of essentially considering options on the worst performing of the indices.

5 Empirical application

In this section we report results on three empirical applications which not only illustrate the flexibility of our framework but also examines several important issues related to the pricing of multivariate options. To be specific, in the first section we examine the importance of allowing for dynamic correlations. In the second section we consider the effect of having priced correlation risk. In the final section we analyze the importance of allowing for non-Gaussian distributed innovations.

5.1 The importance of dynamic correlation

In this section we consider the case where risk neutralization is done on an equation by equation basis. That is, we specify the conditional mean as

$$\mu_t = r_t + \text{diag}\Sigma_t\lambda. \quad (33)$$

For the dynamics we choose the DCC model as this allows us to estimate the model separately for each of the underlying assets. This also means that the same dynamics are used for the individual variances irrespective of the dynamics used for the correlations. We compare the results of the DCC model to the CCC model and to a model without correlation, the NoC model, where R_t in (28) is fixed to the identity matrix.

5.1.1 Estimation results

Table II shows the estimation results for the univariate variance processes for each of the two indices as well as the results for the DCC model defined in (26)-(29). Comparing the first

Table II: Estimation results for DCC model using the NASDAQ and NYSE indices

This table reports the univariate GARCH estimation results and results for the DCC model defined in (26)-(29). Estimation is done with Maximum Likelihood using both variance and correlation targeting.

Dynamics	$\hat{\alpha}_1$	SE($\hat{\alpha}_1$)	$\hat{\beta}_1$	SE($\hat{\beta}_1$)	$\hat{\nu}_1$	SE($\hat{\nu}_1$)
NASDAQ	0.10799	(0.00617)	0.88840	(0.00687)	3.3806	(0.73415)
NYSE	0.07639	(0.00470)	0.91360	(0.00570)	3.0226	(0.90739)
Correlation	0.02557	(0.00244)	0.96790	(0.00343)		

two rows we see that the estimated parameters for the conditional variances are generally of similar size. In particular, although the estimated α is somewhat higher for NASDAQ than for NYSE the opposite is true for the estimated β , and overall this leads to a persistency, given by the sum of α and β , of similar magnitude for the two series. The last row, which reports the estimates of the DCC specification, shows that these parameters are also of similar size as those from the individual variance equations. In particular, this holds when considering the overall persistency of the time series. Finally, the table shows that the risk premia are estimated to be significantly positive for both indices, and taking these into account is clearly important. Moreover, because of this the dynamics under the risk neutral measure differ from those under the original measure and appropriately riskneutralization is important.

The estimation results thus indicate that there is a high degree of persistency in the conditional volatilities and in the conditional correlation. In Figure 2 we plot these three time series and the plots confirm this. The two top plots, Figure 2(a) and 2(b), are for the conditional variances, and show that these vary a lot through time and do not always move together. This is reflected in Figure 2(c), which plots the conditional correlation time series, and shows that the conditional correlation changes through time. In fact, while it is positive through the entire sample it varies between roughly 40% and 95%. The minimum value of 38.9% is found on April 4, 2000, which is right after the climax of the dot-com bubble which happened on March 10 that year with the NASDAQ peaking at 5132.52. This again reflects the fact that this event was primarily something that effected the technology heavy

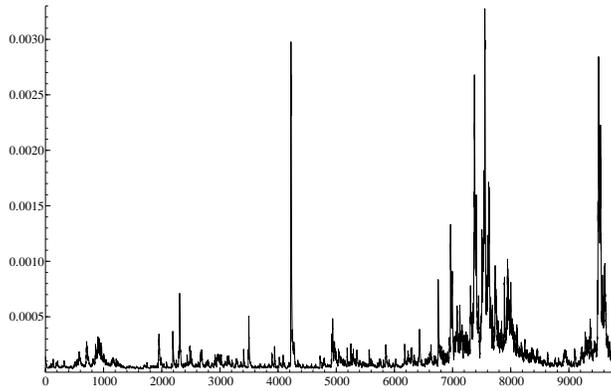
NASDAQ index.

Based on the estimation results and the plotted conditional variances and correlations it should be clear that using a model which treats conditional correlations, or for that sake conditional variances, as constant is clearly not in line with the data. While this holds true from a statistical point of view the results tells us nothing about the effects on option prices. This is explored in the the next section, where we examine the importance of dynamic correlations for option pricing. We choose to price the options on three dates which reflect different levels of the conditional correlation: April 4, 2000, as this is the minimum correlation observed, the last observation in the sample on November 27, 2009, for which the correlation is 91.6% and close to the maximum correlation, and March 10, 2005, which is right between these two dates since at this day the correlation is very close to the unconditional correlation of 78%.

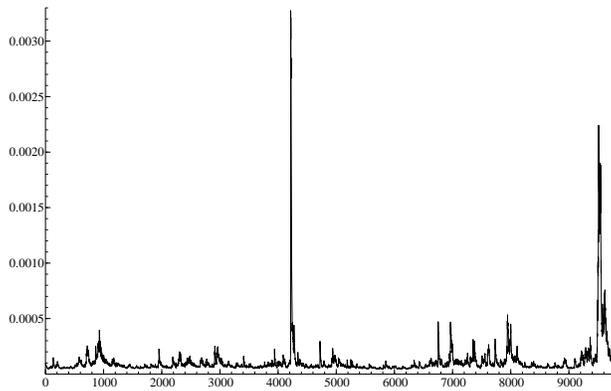
5.1.2 Pricing results

In order to perform option pricing with the above models the appropriate risk neutral dynamics have to be derived. However, this is easily done for the model considered here and the result is given in (17), which shows that the innovations remain Gaussian, although with a mean shifted such that investors are compensated sufficiently for being exposed to variance risk. Instead of reporting the estimated prices, in Figure 3 we plot the relative error surfaces as a function of maturity and moneyness. The errors are calculated as the difference between the true price from the DCC model and the CCC model price, respectively the NoC price, divided by the true DCC price. Thus, the values may be interpreted as the errors that would be committed when using a wrong model, i.e. the CCC or the NoC model, when the true model is in fact the DCC model. The left hand panels are for the CCC model and the right hand panels are for the NoC model. From top to bottom the plots are for the three dates: April 4, 2000, March 10, 2005, November 27, 2009.

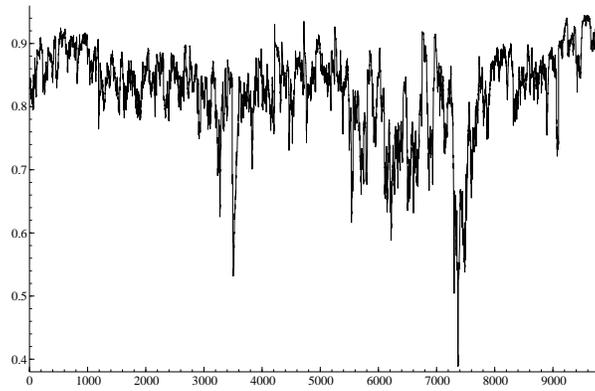
The first thing to notice from Figure 3 is that errors are clearly committed when assuming constant correlations or when neglecting correlations altogether. In particular, this holds irrespective of the assumed model and the time period considered, and errors are committed for all combinations of moneyness and maturity. The errors are clearly important and may



(a) Conditional variance for NASDAQ



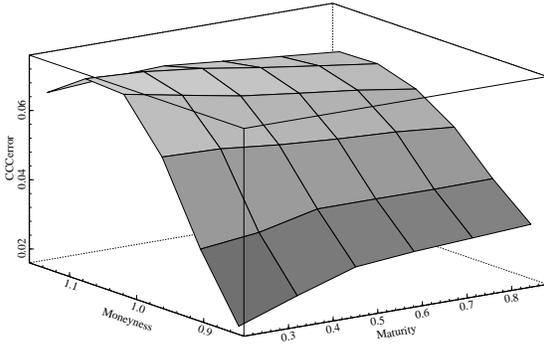
(b) Conditional variance for NYSE



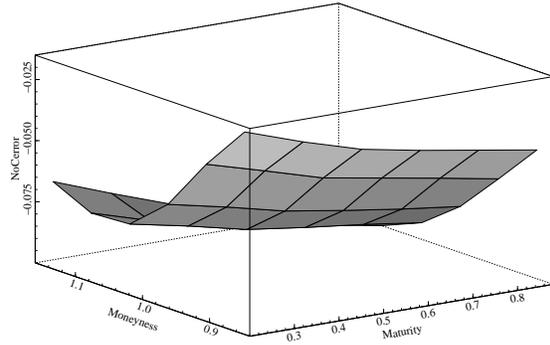
(c) Conditional correlation

Figure 2: Conditional variances and correlations

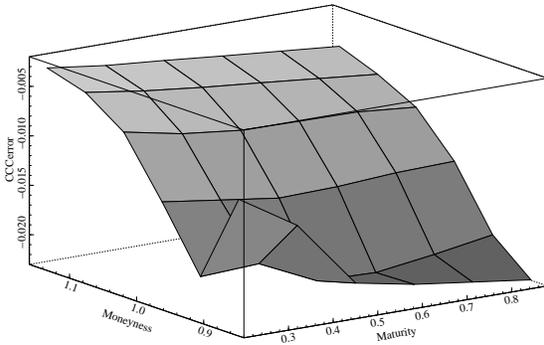
This figure plots the time series of filtered conditional variances and correlations from the DCC model for the period from February 5, 1971, through November 27, 2009.



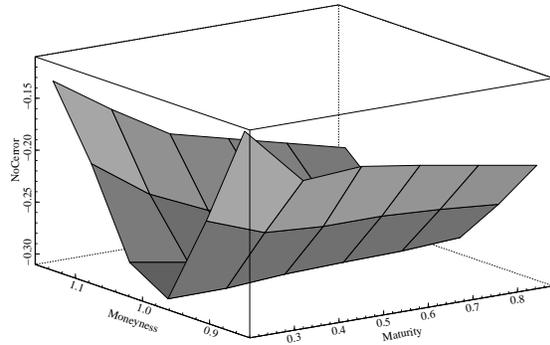
(a) Errors for CCC model in 2000



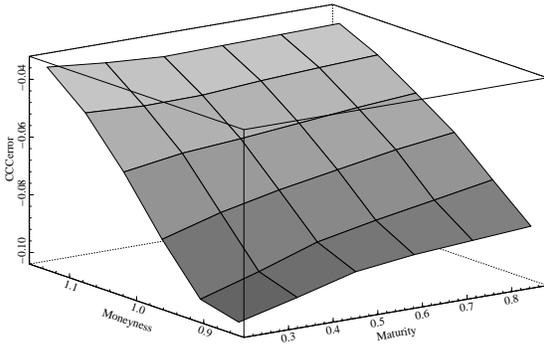
(b) Errors for NoC model in 2000



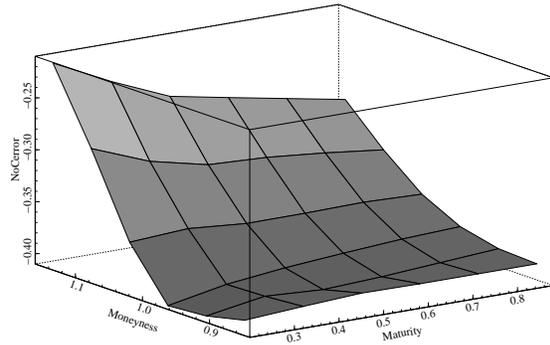
(c) Errors for CCC model in 2005



(d) Errors for NoC model in 2005



(e) Errors for CCC model in 2009



(f) Errors for NoC model in 2009

Figure 3: Pricing errors with CCC and NoC models

This figure plots the pricing errors for the constant correlation model, the CCC model, and the model without correlation, the NoC model, compared to the true dynamic conditional correlation model, the DCC model. The left hand panels are for the CCC model and the right hand panels are for the NoC model. From top to bottom the plots are for April 4, 2000, March 10, 2005, and November 27, 2009.

be as large as 10% when neglecting correlation dynamics and exceed 40% when neglecting correlation altogether. Moreover, when considering the pricing errors from e.g. the NoC model in the right hand panels the figures show that the errors increase with the actual value of the correlation as expected. Finally, while the errors are slightly less important for the short term options they increase with maturity and when the option becomes out of the money.

When considering the pricing errors from the CCC model in the left hand panels, Figure 3 show that for this model the errors are in all cases smaller than for the NoC model. The smallest errors are found on March 10, 2005 as Panel 3(c) shows, although this is to be expected since at this date the conditional correlation is close to the unconditional level used in the CCC model. However, while the sign of the errors and their magnitude change through time the shape of the error surface remains remarkably stable and shows a smirk across moneyness akin to the one often found empirically in implied volatilities. This smirk though is purely correlation induced as the variance dynamics are the same for all models. It is important to note that that the correlation smirk does not vanish with maturity as is often the case with the implied volatility smirk.

In conclusion, Figure 3 shows that neglecting correlation altogether leads to large pricing errors when pricing multivariate options. However, even when allowing for correlation important errors occur when the dynamics are neglected. Thus, the results in this section show that correctly modeling the dynamics of the conditional correlation is important for all options irrespective of the moneyness or the maturity.

5.2 The importance of correlation risk

In this section we examine the results when the correlation risk, in addition to the variance risk, is priced. That is, for each stock i we consider the following specification of the conditional mean

$$\mu_{i,t} = r_t + \sigma_i^2 \lambda_{i,i} + \rho_{i,j} \lambda_{i,j}. \quad (34)$$

In addition to this mean specification we consider two restricted versions: the first of which sets $\lambda_{i,j} = 0$ for $i \neq j$, and the second where $\lambda_{i,j} = 0$ for all i and j . That is, the first

Table III: Estimation results for the bivariate VEC model for NASDAQ and NYSE indices with correlation and variance risk premia

This table reports the estimation results for the VEC model with estimates of both variance and covariance premiums in Panel A, the full model, for the restricted model with only variance risk premia in Panel B, and for the model without risk premia in Panel C. The estimates of the correlations risk premia has been multiplied by 100. Estimation is done with Maximum Likelihood using both variance and correlation targeting.

Panel A: Full model								
Dynamics	$\hat{\alpha}_1$	SE($\hat{\alpha}_1$)	$\hat{\beta}_1$	SE($\hat{\beta}_1$)	$\hat{\lambda}_{11}$	SE($\hat{\lambda}_{11}$)	$\hat{\lambda}_{12}$	SE($\hat{\lambda}_{12}$)
NASDAQ	0.08157	(0.00205)	0.91164	(0.00226)	1.521	(0.7411)	3.943	(0.9719)
NYSE	0.06898	(0.00193)	0.92170	(0.00224)	2.281	(0.8860)	2.280	(1.0130)
Covariance	0.06956	(0.00152)	0.92155	(0.00167)				
Panel B: Restricted model with only variance risk								
Dynamics	$\hat{\alpha}_1$	SE($\hat{\alpha}_1$)	$\hat{\beta}_1$	SE($\hat{\beta}_1$)	$\hat{\lambda}_{11}$	SE($\hat{\lambda}_{11}$)		
NASDAQ	0.08037	(0.00180)	0.91294	(0.00197)	2.874	(0.6540)		
NYSE	0.06855	(0.00173)	0.92222	(0.00198)	2.897	(0.7548)		
Covariance	0.06886	(0.00127)	0.92235	(0.00133)				
Panel C: Restricted model without risk premia								
Dynamics	$\hat{\alpha}_1$	SE($\hat{\alpha}_1$)	$\hat{\beta}_1$	SE($\hat{\beta}_1$)				
NASDAQ	0.07966	(0.00202)	0.91377	(0.00222)				
NYSE	0.06789	(0.00190)	0.92354	(0.00220)				
Covariance	0.06743	(0.00151)	0.92359	(0.00165)				

alternative specification has no correlation risk but only variance risk premia whereas the latter specification has no risk premia at all. We use the VEC specification above instead of the DCC specification because in this model all the dynamics are estimated at once. This is convenient since the correlations are needed in order to estimate the risk premia simultaneously.

5.2.1 Estimation results

Table III provides the estimation results for the three versions of the VEC model. In Panel A the results with the mean specification from (34) are provided. In Panel B results for the restricted model where only the variance risk carries a premium with it are reported, and finally in Panel C results for the model with no risk premia is reported. When comparing

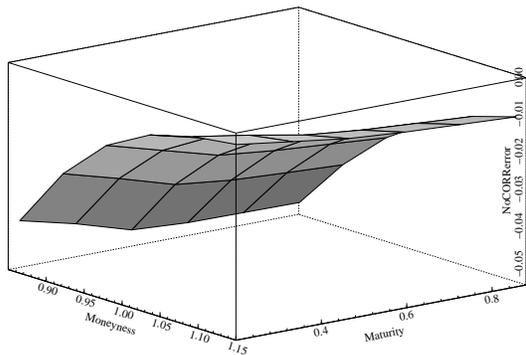
the three panels it is immediately seen that the parameter estimates for α and β are very similar. Moreover, the estimates are also very close to those found for the DCC model in Table II. Thus, it is seen that finding time variation and high persistency is not particular to a specific model or to the assumptions on the risk premia.

However, when comparing the actual estimates of the risk premia between the two panels differences are found. To be specific, it is immediately seen that when a premium on correlation risk is included the variance risk premia drops. The decrease in the estimated variance premia is most pronounced for NASDAQ where the size is roughly half that of the estimates when this is the only source of risk which is priced. For NYSE the variance risk premia on the other hand only drops marginally, but for this index the size of the estimated correlation risk premia is also much smaller than for NASDAQ also.

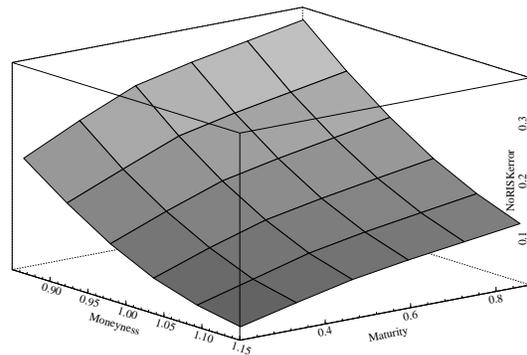
Thus, Table III clearly shows that it is important to consider a premium on correlations in addition to that on variances. This finding is in line with the capital asset pricing model, or CAPM. In particular, interpreting the NYSE as the market portfolio the CAPM would tell us that the excess returns of any other portfolio should only depend on its correlation with the market. This is largely what is found above. Moreover, our finding of priced correlation risk is in line with e.g. Driessen, Maenhout, and Vilkov (2009) who find this using model free implied variances from option prices. We note that the ability to incorporate such premiums and to properly risk neutralize with respect to these in a coherent way is possible only in a framework as general as the one put forth here.

5.2.2 Pricing results

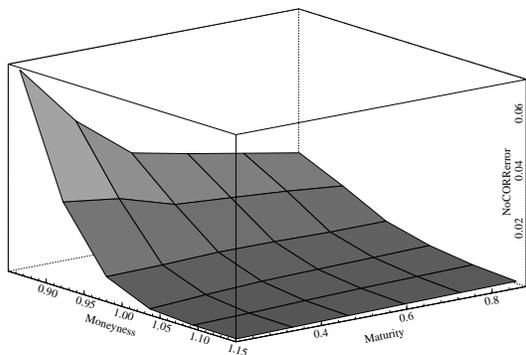
We use the models estimated above to price the same sample of artificial options as in the previous section after appropriately risk neutralization. The dynamics to be used are easily derived by substituting the values for $\lambda_{i,j}$ into equation (15). Doing so it is easy to realize that the innovations, as before, remain Gaussian. However, the distribution now has a mean which is shifted such that investors are compensated for both the variance risk and the correlation risk. In Figure 4 we plot the relative errors in the estimated option prices from using the two restricted models, one without correlation risk premia, the NoCORR model, and one with neither variance nor correlation premia, the NoRISK model.



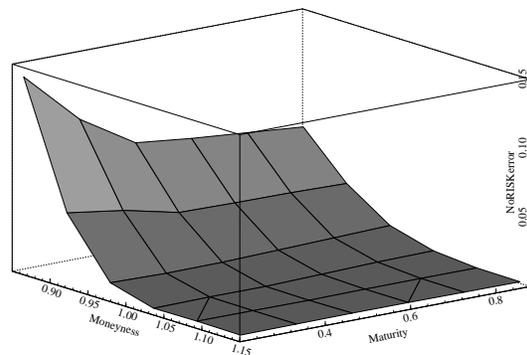
(a) Errors for NoCORR model in 2000



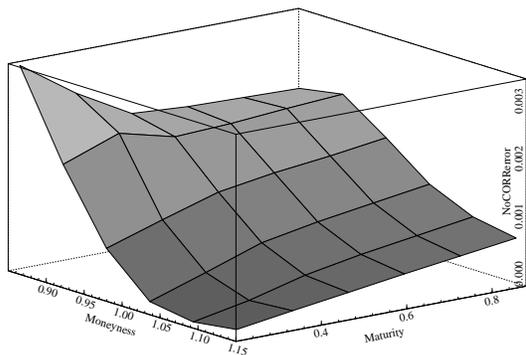
(b) Errors for NoRISK model in 2000



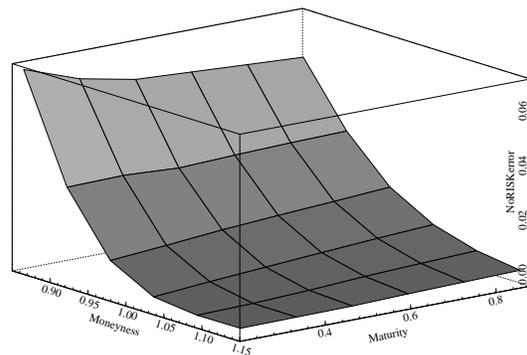
(c) Errors for NoCORR model in 2005



(d) Errors for NoRISK model in 2005



(e) Errors for NoCORR model in 2009



(f) Errors for NoRISK model in 2009

Figure 4: Pricing errors with NoCORR and NoRISK models

This figure plots the pricing errors for the model with only variance risk, the NoCORR model, and the model without any risk premia, the NoRISK model, compared to the true model with both correlation and variance risk premia. The left hand panels are for the NoCORR model and the right hand panels are for the NoRISK model. From top to bottom the plots are for April 4, 2000, March 10, 2005, and November 27, 2009.

The first thing to notice from Figure 4 is that the errors are generally smaller than the ones in Figure 3, which may indicate that it is more important to incorporate time varying conditional variances and correlations than the actual risk premia. This is somewhat to be expected and it corresponds to what is often found when pricing options in a univariate framework. In particular, in the univariate case the time variation modeled by e.g. a GARCH process often has a much larger effect on option prices than including the estimated risk premia (see e.g. Stentoft (2008)).

However, although the errors are smaller they are still present and for some options they may be quite large in relative terms. This is particularly so for the out of the money options. Here errors of up to 30% can be observed when neglecting the risk premia altogether as Panel 4(b) shows. When only the correlation risk premia is neglected the estimated error is generally smaller, but they may still be as large as 6% of the true estimate of the price given by the unrestricted VEC model as Panel 4(c) shows. Comparing across maturity all the panels furthermore show that the mispricings persist even for the options with long maturities.

In conclusion, Figure 4 shows that neglecting variance and correlation risk premia may lead to large pricing errors when pricing multivariate options. Moreover, while the variance risk premia is important so is the premia on correlation risk, and finally when neglecting the correlation risk premia our results show that the sign of the committed errors depend on the level of the correlations. Thus, the results in the section show that correctly taking account of the variance and correlation risk premia is important for all options irrespective of the moneyness or the maturity.

5.3 The importance of non-Gaussian features

In this section we examine the results when the underlying distribution is allowed to be non-Gaussian. That is, we consider the pricing results when the multivariate mixed normal distribution is used. For the present we restrict attention to the case when $K = 2$ in (19) and we compare this to the Gaussian case which corresponds to a mixture model with $K = 1$. As an additional alternative we consider a restricted version of the two component mixture model with only one variance component, that is with $\Sigma_{1,t} = \Sigma_{2,t}$ in (19). While this model

has the ability to generate skewness and excess kurtosis, it is less flexible when it comes to modeling the dynamics of the conditional variances and correlation.

When considering this more general model it is difficult, if not impossible, to derive the sequence of ν_t given a generally specified equation for the mean μ_t . Therefore we use the method of implying μ_t given ν_t instead. To be specific, we assume a constant value for ν_t and specify the mean as

$$R_{j,t} = r_t - \Psi_t(\nu - e_j) + \Psi_t(\nu) + \varepsilon_{j,t}. \quad (35)$$

Since (35) only depends on known parameters it can be used for estimation directly. We use a VEC specification for the dynamics and for simplicity we only consider that the conditional variance risk is priced. Thus, the Gaussian special case corresponds to the restricted VEC model used previously.

5.3.1 Estimation results

In Table IV we report the estimation results for the multivariate mixed normal model. The first thing to note from the table is that two distinctly different components are found. The first of these is very similar to the Gaussian special case, and in this component the estimated α and β and the implied persistence are of roughly the same size as was found in Panel B of Table III. The probability of being in this component is very high, 93%. The second component, however, is very different. In particular, in this component both variance and covariance processes are explosive. The estimated persistence of this component is 1.38 for NASDAQ and 1.23 for NYSE and the correlation dynamics. Although the probability of being in this component is rather low, only around 7%, it remains statistically significant.

In Table V we report the estimation results for the restricted version of the multivariate mixed normal model with only one specification of the variance and covariance dynamics. Again it is seen that the dynamics are very similar to those found for the Gaussian special case. Furthermore, when looking at the mixture parameters it is seen that now both mean parameters are very large numerically and significant. Together with the higher probability, which is close to the boundary, this could indicate that this type of restricted mixture, although capable of generating skewness and excess kurtosis, is not very suitable for the

Table IV: Estimation results for the multivariate mixed normal model for NASDAQ and NYSE indices

This table reports the estimation results for the multivariate mixed normal model. The top panel reports the results for the first component, the second panel the results for the second components, the third panel the results for the mixture distribution, and the bottom panel the risk premia.

Component 1						
Dynamics	$\hat{\omega}_1$	SE($\hat{\omega}_1$)	$\hat{\alpha}_1$	SE($\hat{\alpha}_1$)	$\hat{\beta}_1$	SE($\hat{\beta}_1$)
NASDAQ	0.00279	(0.000498)	0.05252	(0.002684)	0.93173	(0.003067)
NYSE	0.00274	(0.000457)	0.04406	(0.002364)	0.94004	(0.002830)
Covariance	0.00377	(0.000587)	0.04477	(0.002599)	0.93955	(0.003086)
Component 2						
Dynamics	$\hat{\omega}_1$	SE($\hat{\omega}_1$)	$\hat{\alpha}_1$	SE($\hat{\alpha}_1$)	$\hat{\beta}_1$	SE($\hat{\beta}_1$)
NASDAQ	0.42730	(0.068318)	0.74526	(0.138920)	0.64049	(0.031282)
NYSE	0.31368	(0.071851)	0.51655	(0.100510)	0.70981	(0.045008)
Covariance	0.31796	(0.077974)	0.49360	(0.067608)	0.73205	(0.044121)
Mixture parameters						
	$\hat{\mu}_1$	SE($\hat{\mu}_1$)	$\hat{\mu}_2$	SE($\hat{\mu}_2$)	$\hat{\pi}$	SE($\hat{\pi}$)
Mixture	-0.01082	(0.010493)	-0.01811	(0.013648)	0.93043	(0.007312)
Risk premia						
	$\hat{\nu}_1$	SE($\hat{\nu}_1$)	$\hat{\nu}_2$	SE($\hat{\nu}_2$)		
Risk premia	1.0254	(0.35105)	1.6574	(0.71452)		

data considered here.

Lastly, the estimation results show that the risk parameters, that is the vector of ν 's, are positive and significantly different from zero for both indices and for both types of mixtures. The estimates are, however, smaller in size than for the Gaussian special case. However, their smaller size is to be expected as the higher order moments are considered in the multivariate mixed normal model. Thus, in this model skewness and excess kurtosis affect the risk premium in addition to the variance through the cumulant generating function in (35). In the Gaussian model, on the other hand, the variance is the only moment considered. Hence, if asset returns are in fact non-Gaussian, which is the case here, the estimates will be biased upward when using a Gaussian model.

Table V: Estimation results for the restricted multivariate mixed normal model for NASDAQ and NYSE indices

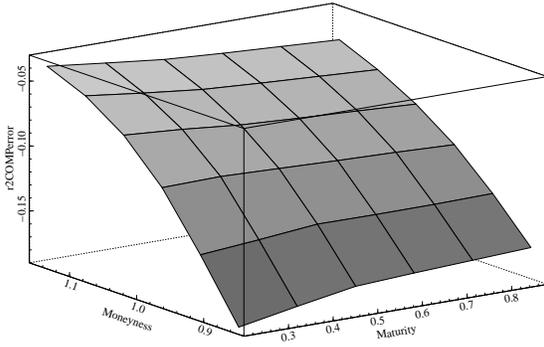
This table reports the estimation results for the restricted multivariate mixed normal model. The top panel reports the results for the variances and the covariance dynamics, the middle panel the results for the mixture distribution, and the bottom panel the risk premia.

Component 1						
Dynamics	$\hat{\omega}_1$	SE($\hat{\omega}_1$)	$\hat{\alpha}_1$	SE($\hat{\alpha}_1$)	$\hat{\beta}_1$	SE($\hat{\beta}_1$)
NASDAQ	0.00915	(0.000720)	0.08157	(0.002615)	0.90904	(0.002662)
NYSE	0.00797	(0.000548)	0.06911	(0.001912)	0.91978	(0.002158)
Covariance	0.00932	(0.000697)	0.07226	(0.001545)	0.91835	(0.002030)
Mixture parameters						
	$\hat{\mu}_1$	SE($\hat{\mu}_1$)	$\hat{\mu}_2$	SE($\hat{\mu}_2$)	$\hat{\pi}$	SE($\hat{\pi}$)
Mixture	-0.08707	(0.037580)	-0.03081	(0.018140)	0.96489	(0.29608)
Risk premia						
	$\hat{\nu}_1$	SE($\hat{\nu}_1$)	$\hat{\nu}_2$	SE($\hat{\nu}_2$)		
Risk premia	0.27512	(0.076209)	1.2524	(0.50986)		

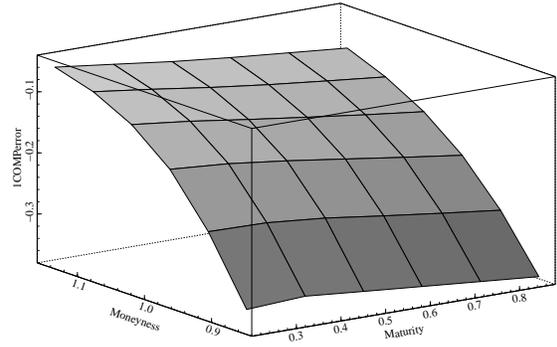
5.3.2 Pricing results

Once again, we use the estimated models to price the sample of artificial options after appropriately risk neutralization. The risk neutral dynamics in the multivariate mixed normal model follow from (22) which shows that the innovations remains within the multivariate mixed normal model. In Figure 5 we plot the relative errors in the estimated option prices when using either the restricted two component model, the r2COMP model, or the one component model, the 1COMP model, which as already mentioned corresponds to the Gaussian model used before.

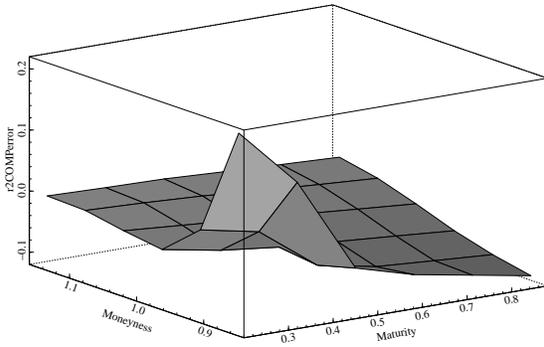
The first thing to notice from Figure 5 is that the errors committed when using a Gaussian model instead of the correct mixture model are generally large, and as the figure shows these errors may be as large as 40% . In fact, they are much larger than when neglecting correlation risk and often they are of similar size as the errors committed when neglecting time varying correlation altogether. Thus, neglecting the non-Gaussian features of the data can lead to very poor estimates of the option prices. Moreover, while the errors are smaller when considering the restricted two component model they remain large in size.



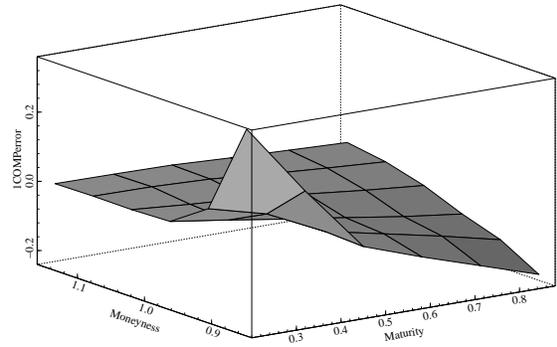
(a) Errors for r2COMP model in 2000



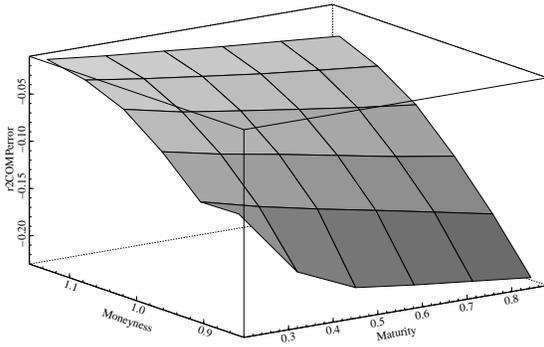
(b) Errors for 1COMP model in 2000



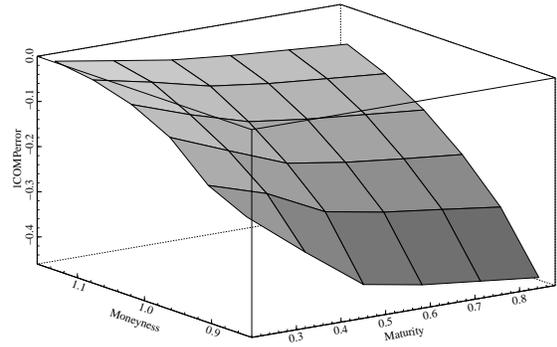
(c) Errors for r2COMP model in 2005



(d) Errors for 1COMP model in 2005



(e) Errors for r2COMP model in 2009



(f) Errors for 1COMP model in 2009

Figure 5: Pricing errors with r2COMP and 1COMP models

This figure plots the pricing errors for the restricted two component model, the r2COMP model, and the one component model, the 1COMP model, compared to the true two component model, the 2COMP model. The left hand panels are for the r2COMP model and the right hand panels are for the 1COMP model. From top to bottom the plots are for April 4, 2000, March 10, 2005, and November 27, 2009.

Figure 5 also shows that the pricing errors committed are again relatively largest for the out of the money options. In fact, in the majority of the panels a smirk in terms of the pricing errors can be seen. This time though, the smirk is driven by the non-Gaussian features in the model and occurs in addition to the correlation induced smirk. Lastly, it should be noted that this smirk does not vanish when considering long term options.

In conclusion, Figure 5 shows that neglecting non-Gaussian features of the underlying distribution may lead to large pricing errors when pricing multivariate options. Moreover, the mispricings for the restricted two component model also shows that not only is non-Gaussian features important but so is allowing for a flexible specification of the dynamics. Thus, the results of this section show that correctly taking account of non-Gaussian features in the data is important for all options irrespective of the moneyness and maturity.

6 Conclusion

In this paper we consider option pricing when asset dynamics can be described by multivariate time varying volatility models. We derive the risk neutral dynamics in the general model and provide detailed results for some special cases which are of particular interest. These results show that our method can be used for various different specifications of risk premia and when the model is used with non-Gaussian distributions.

We apply our method to price bivariate options on the minimum of the NASDAQ and the NYSE indices. We compare the general multivariate model with conditional heteroskedasticity and Gaussian innovations to the various special cases, and find large differences in the estimated prices. We then use the framework to examine the importance of premiums for exposure to correlation risk, and find that this has important implications for out of the money options in particular. Finally, we compare the estimates obtained in the Gaussian case to those obtained when the conditional distribution is a multivariate mixture of normals. Again we find very large differences in the estimated prices. Based on these findings our overall conclusion is that all of the features considered here have potentially important implications when it comes to option pricing in a multivariate setting and should not be neglected. A framework, such as the one developed in this paper, which is flexible enough

to accommodate each of these issues is therefore clearly required.

With the framework developed in the present paper several issues of interest can be examined. First of all, we report here results only for minimum options although other multivariate options could easily be considered. A second interesting question is to apply the framework to options on a large number of underlying assets. Finally, one could use the multivariate framework to price options on the individual assets. Considering the importance of correlations and the premia on this source of risk, this could potentially lead to better estimates of the price of such options.

References

- BADESCU, A., R. KULPERGER, AND E. LAZAR (2008): “Option Valuation with Normal Mixture GARCH Models,” *Studie in Nonlinear Dynamics and Control*, 12(2), Article 5.
- BALI, T. (2008): “The Intertemporal Relation between Expected Returns and Risk,” *Journal of Financial Economics*, 87, 101–131.
- BARRAQUAND, J. (1995): “Numerical Valuation of High Dimensional Multivariate European Securities,” *Management Science*, 41(12), 1882–1891.
- BAUWENS, L., C. HAFNER, AND J. ROMBOUTS (2007): “Multivariate Mixed Normal Conditional Heteroskedasticity,” *Computational Statistics and Data Analysis*, 51, 3551–3566.
- BAUWENS, L., S. LAURENT, AND J. ROMBOUTS (2006): “Multivariate GARCH Models: A Survey,” *Journal of Applied Econometrics*, 21, 79–109.
- BERTHOLON, H., A. MONFORT, AND F. PEGORARO (2006): “Pricing and Inference with Mixtures of Conditional Normal Processes,” *CREST Working Paper 28*.
- BILLINGSLEY, P. (1995): *Probability and Measure*. Wiley-Interscience, New York.
- BOLLERSLEV, T. (1986): “Generalized Autoregressive Conditional Heteroskedasticity,” *Journal of Econometrics*, 31, 307–327.

- (1990): “Modeling the Coherence in Short-run Nominal Exchange Rates: A Multivariate Generalized ARCH model,” *Review of Economics and Statistics*, 72, 498–505.
- BOLLERSLEV, T., R. ENGLE, AND J. WOOLDRIDGE (1988): “A Capital Asset Pricing Model with Time Varying Covariances,” *Journal of Political Economy*, 96, 116–131.
- BOYLE, P. P., J. EVNINIE, AND S. GIBBS (1989): “Numerical Evaluation of Multivariate Contingent Claims,” *The Review of Financial Studies*, 2, 241–250.
- BROADIE, M., AND J. DETEMPLE (1997): “The Valuation of American Options on Multiple Assets,” *Mathematical Finance*, 7(3), 241–286.
- BROADIE, M., AND J. B. DETEMPLE (2004): “Option Pricing: Valuation Models and Applications,” *Management Science*, 50(9), 1145–1177.
- 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.
- DAVIDSON, J. (1997): *Stochastic Limit Theory*. Oxford University Press, Oxford.
- DETEMPLE, J., S. FENG, AND W. TIAN (2003): “The Valuation of American Call Options on the Minimum of Two Dividend-paying Assets,” *The Annals of Applied Probability*, 13(3), 953–983.
- DRIESSEN, J., P. J. MAENHOUT, AND G. VILKOV (2009): “The Price of Correlation Risk: Evidence from Equity Options,” *Journal of Finance*, LXIV(3), 1377–1406.
- DUAN, J. (1995): “The GARCH Option Pricing Model,” *Mathematical Finance*, 5, 13–32.
- DUAN, J., AND S. R. PLISKA (2004): “Option Valuation with Co-integrated Asset Prices,” *Journal of Economic Dynamics and Control*, 28, 727–754.
- DUAN, J.-C., AND J.-G. SIMONATO (1998): “Empirical Martingale Simulation for Asset Prices,” *Management Science*, 44(9), 1218–1233.

- DURHAM, G. (2007): “SV mixture models with application to S&P 500 index returns,” *Journal of Financial Economics*, 85, 822–856.
- ENGLE, R. (1982): “Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation,” *Econometrica*, 50, 987–1007.
- (2002): “Dynamic Conditional Correlation - a Simple Class of Multivariate GARCH Models,” *Journal of Business and Economic Statistics*, 20, 339–350.
- HESTON, S. L., AND S. NANDI (2000): “A Closed-Form GARCH Option Valuation Model,” *Review of Financial Studies*, 13(3), 585–625.
- JOHNSON, H. (1987): “Options on the Maximum or the Minimum of Several Assets,” *Journal of Financial and Quantitative Analysis*, 22, 277–283.
- KAROLYI, G. (1995): “A Multivariate GARCH Model of International Transmission of Stock Returns and Volatility: The Case of United States and Canada,” *Journal of Business and Economic Statistics*, 13, 11–25.
- LONGSTAFF, F. A., AND E. S. SCHWARTZ (2001): “Valuing American Options by Simulation: A Simple Least-Squares Approach,” *Review of Financial Studies*, 14, 113–147.
- MARGRABE, W. (1978): “The Value of an Option to Exchange One Asset for Another,” *Journal of Finance*, 33, 177–186.
- MOSCHINI, G., AND R. MEYERS (2002): “Testing for Constant Hedge Ratios in Commodity Markets: a Multivariate GARCH Approach,” *Journal of Empirical Finance*, 9, 589–603.
- ROMBOUTS, J., AND L. STENTOFT (2009): “Bayesian Option Pricing using Mixed Normal Heteroskedastity Models,” *SSRN Working Paper*.
- SILVENNOINEN, A., AND T. TERASVIRTA (2008): “Multivariate GARCH models,” in *Handbook of Financial Time Series*, ed. by T. Andersen, R. Davis, J. Kreiss, and T. Mikosch. Springer.
- STENTOFT, L. (2004): “Convergence of the Least Squares Monte Carlo Approach to American Option Valuation,” *Management Science*, 50(9), 1193–1203.

- (2008): “American Option Pricing using GARCH models and the Normal Inverse Gaussian distribution,” *Journal of Financial Econometrics*, 6(4), 540–582.
- STULZ, R. (1982): “Options on the Minimum or Maximum of Two Risky Assets: Analysis and Applications,” *Journal of Financial Economics*, 10, 161–185.
- VILLANEUVE, S. (1999): “Exercise Regions of American Options on Several Assets,” *Finance and Stochastics*, 3, 295–322.
- WERKER, B., C. GENEST, AND R. VAN DEN GOORBERGH (2005): “Bivariate Option Pricing Using Dynamic Copula Models,” *Insurance: Mathematics and Economics*, 37, 101–114.
- ZHANG, J., AND D. GUÉGAN (2008): “Pricing Bivariate Option under GARCH Processes with Time-varying Copula,” *Insurance: Mathematics and Economics*, 42, 1095–1103.