S Knox and P Ouwehand*

Pricing rainbow options: Nonparametric methods using copulas

1. INTRODUCTION

The aim of this paper is to investigate the pricing of bivariate options on the Johannesburg Stock Exchange All Share Index (ALSI) and Bond Exchange of South Africa All Bond Index (ALBI).

Here follow some remarks on mathematical asset pricing theory (cf. Bingham and Kiesel (2004):

A probability measure $\bar{P}$ is said to be risk-neutral if, under $\bar{P}$, every asset has the same expected return as the riskless bank account, i.e. if for all assets S, we have $S_0 = e^{-rT}\bar{E}[S_T]$, where S_t is the time t-value of S, and r is the risk-free rate over the period (ignoring dividends and other intermediate payoffs.) The *Fundamental Theorem of Mathematical Finance* states that a market model is arbitrage-free if and only if there is a risk-neutral measure for that model. The *Principle of Risk-Neutral Valuation* states that if a financial derivative X can be hedged, then its price is the discounted expected value of its future payoff, i.e. $X_0 = e^{-rT}\bar{E}[X_T]$. This relationship allows one to compute derivatives prices by Monte Carlo simulation. It must be stressed that the expectation *must* be taken with respect to a risk-neutral measure; the use of "real world" distributions (obtained from historical prices) will generally introduce arbitrage opportunities.

When pricing multivariate financial derivatives, one needs to estimate the *joint* distribution of the underlying variables, in the "risk-neutral world", which cannot be directly observed. Such estimation can be done parametrically or nonparametrically. The standard Black-Scholes option pricing framework, for example, is parametric: it assumes that (i) each "real world" asset price process is completely characterized by its drift and volatility, and (ii) asset returns are jointly normally distributed. When one moves to the risk-neutral world, only the drifts (i.e. rates of return) of the underlying assets change. The volatilities and correlations between assets are not affected by the change of measure, and these parameters can be estimated directly from historical price data.

The existence of the *volatility smile* or *skew* bears testimony to this fact that the Black-Scholes framework does not price options accurately. To price options on more than one asset it is necessary to model asset returns by non-multinormal distributions.

Instead of positing another parametric form for the (risk-neutral) distribution of asset returns, one may estimate the distribution *nonparametrically*, with no prior restrictions on the form of the distribution. This is relatively straightforward for univariate asset returns: Breeden and Litzenberger (1978) note that the risk-neutral distribution of future asset prices can be extracted from market-observed option prices. If there is liquid market for vanilla options with a wide range of strikes, this result can be used in practice (Ait-Sahalia and Lo (1998).

In pricing multi-asset options, we face two problems: Firstly, the estimation of the risk-neutral multivariate distribution is *not* straightforward. In principle, the distribution can be estimated from the values of traded multi-asset options (Rosenberg, 2003). In practice, these are mostly OTC contracts, and thus prices are illiquid and difficult to obtain. Simply ignoring the problem by using a sophisticated model to price single-asset options, and a multi-asset Black-Scholes model to price multi-asset options may well introduce arbitrage (Coutant, Durrleman, Rapuch and Roncalli, 2001) In addition, because the single-asset options are typically used to hedge the multi-asset option, the use of two different models may result in bad hedging strategies.

Secondly, we need to estimate the dependence structure between the underlying assets. Outside the Gaussian framework, *correlation* is an inadequate measure of dependence: (Embrechts, MacNeil and Straumann, 2002) discuss several dependence concepts in a financial setting, and highlight many shortcomings of correlation. To build an adequate multi-asset pricing model, a thorough understanding of the possible dependencies between assets is therefore necessary. Blind reliance on correlation will lead to mispricing.

The use of *copulas* to model multivariate distributions brings important benefits. As we shall see, a copula captures the notion of *dependence*. The estimation of multi-asset risk-neutral distributions is divided into two steps:

- Estimate the univariate risk-neutral distributions.
- Model the dependence structure with a copula.

* Department of Mathematics and Applied Mathematics, University of Cape Town, Private Bag, Rondebosch 7707, Republic of South Africa.
Email: peter@maths.uct.ac.za

The first step is fairly well understood and has generated a large literature (see Jackwerth (1999) for a comprehensive literature review). The second step is more contentious. Various authors have addressed the problem of estimating copulas from data. However, the estimated copula will model the "real world" dependence structure. To price options, one needs to know the risk-neutral copula. Authors such as Van den Goorbergh, Genest and Werker (2005) and Rosenberg (2003) argue that the "real world" and risk-neutral dependence structure will be the same, on grounds that the move to the risk-neutral world should only adjust the expected returns of the underlying assets, but leave everything else the same.

The aim of this paper, then, is to investigate the pricing of bivariate options, taking into account the fact that the marginal returns non-normal, and that the dependence structure of joint returns is non-Gaussian. We use nonparametric methods to estimate the marginal risk-neutral asset returns distributions, and then couple these marginals into a joint returns distribution using a copula. Option prices are obtained via Monte Carlo simulation, and compared to the prices obtained within the Black-Scholes framework.

2. PROBLEM DESCRIPTION AND DATA

We will use risk-neutral valuation methodology to price bivariate options on the ALSI and ALBI indices, denoted by S_t (shares) and B_t (bonds), respectively. We have in mind options of the following types (where K is the strike, and T the expiry date):

- Rainbow call on the maximum:
  Payoff = $\max(\max(S_T, B_T) - K, 0)$
- Rainbow call on the minimum:
  Payoff = $\max(\min(S_T, B_T) - K, 0)$
- Exchange (Margrabe) option:
  Payoff = $\max(S_T - B_T, 0)$

Generally, the options are on returns, i.e. the initial prices are normalized, so that $S_0 = 1 = B_0$.

Closed form solutions for option prices exist for all but the spread option in the Black-Scholes framework, and can be found in many texts. See Ouwehand and West (2006) for a modern derivation.

To price options by simulation, we estimate the risk-neutral joint distribution of returns, in two steps:

- First, we obtain the implied risk-neutral price/returns distributions for S_T and B_T.
- Then, we obtain a realistic dependence structure (copula) to join the marginal distributions into a plausible risk-neutral joint distribution.

All options have a one-month expiry, i.e. $T = \frac{1}{12}$. The SAFEX implied volatility skew of one-month Top 40 futures options (expiring 16 September 2004) was used to recover the implied risk-neutral distributions for equity using the method of Breeden and Litzenberger (1978) (see Appendix A.1 for a description).

Bond options are not exchange-traded, and in the absence of a market-visible bond option volatility smile, we use a second method, due to (Duan, 2002), to obtain a risk-neutral distribution for bond returns that is "close" to the "real world" distributions (see Appendix A.2 for a description).

Like Van den Goorbergh *et al.* (2005) and Rosenberg (2003), we will assume that the "real world" and risk-neutral dependence structure are the same. This is a flaw in the methodology, but it seems impossible to recover an implied risk-neutral joint distribution without access to bivariate option prices, and no other method readily suggests itself. We therefore obtain the "real world" dependence structure from monthly (total) returns data of the ALSI and ALBI. To get joint distributions that are likely to be quite different from a Black-Scholes distribution, we've chosen a data set that exhibits strong dependence (because of a market crash). Specifically, the sample used in this exercise is the five-year period from 30 September 1995 to 30 September 2000. For rolling 60 month periods taken over the last 20 years, this period exhibits the strongest dependence, as measured by correlation, Spearman's rho and Kendall's tau.

3. ESTIMATING THE RISK-NEUTRAL JOINT DENSITY

3.1. Estimating the risk-neutral marginal distribution for the ALSI

We use Breeden and Litzenberger (1978) to extract the implied risk-neutral equity price distribution $\bar{F}_S$ from futures option prices (given by the SAFEX skew). To obtain a smooth density, we used kernel smoothing, using simulation (discussed in Appendices A.5 and A.4: Recall that if U is a standard [0, 1]-uniformly distributed random variable, then $\bar{F}_S^{-1}(U)$ has distribution $\bar{F}_S$. One can use simulation to generate a large vector of $\bar{F}_S$-distributed sample points, and then use kernel smoothing to recover a smoothed version of the density. The results are shown in Figure 1.

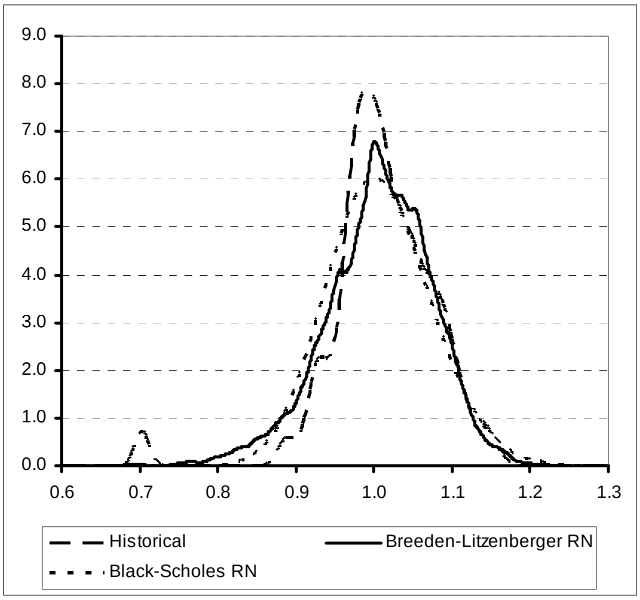

Figure 1: Historical, risk-neutral, and Black-Scholes equity price densities

3.2 Estimating a risk-neutral marginal distribution for the ALBI

Because there are no exchange-traded options on the ALBI, a volatility skew is not directly observable. Instead, we use the method of Duan (2002) to obtain a candidate for the marginal risk-neutral distribution. This method works by finding the risk-neutral distribution that is "closest" (as measured by relative entropy) to the historical distribution.

Recalling that the main feature of a risk-neutral distribution is that the expected return of an asset is the riskless rate, Duan (2002) notes that

> *"Other than the expected return condition, there is no* a priori *reason to expect the risk-neutral distribution to deviate from the physical distribution. In other words, it is natural to have a risk-neutral distribution that is as close to the physical distribution as possible, while satisfying the condition on the expected value".*

Plausible, but there is a problem: The risk-neutral density obtained by Duan's method is close to the historical measure, and "remembers" the past. In particular, in August 1998, the bond market fell by about 15%, and this is clearly visible in the risk-neutral bond density (see Figure 2). Applying Duan's method to equity, a similar bump is observed at -35%. This is not seen in the density obtained by Breeden and Litzenberger's method; the density obtained from the volatility skew is forward looking, whereas the one obtained from historical data is backwards looking. These densities are shown in Figure 3.

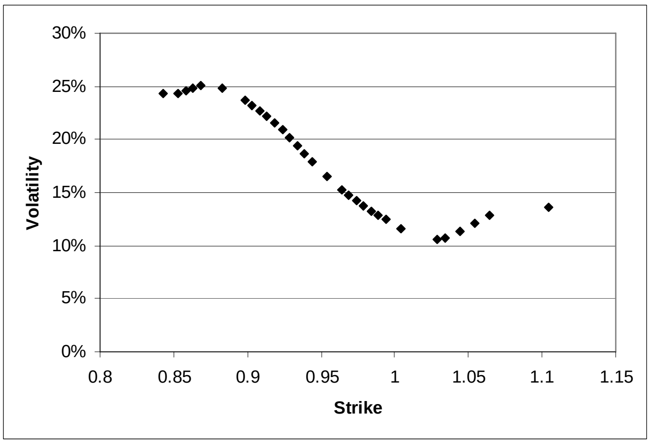

Figure 2: Volatility skew for bond index options obtained from Duan's method

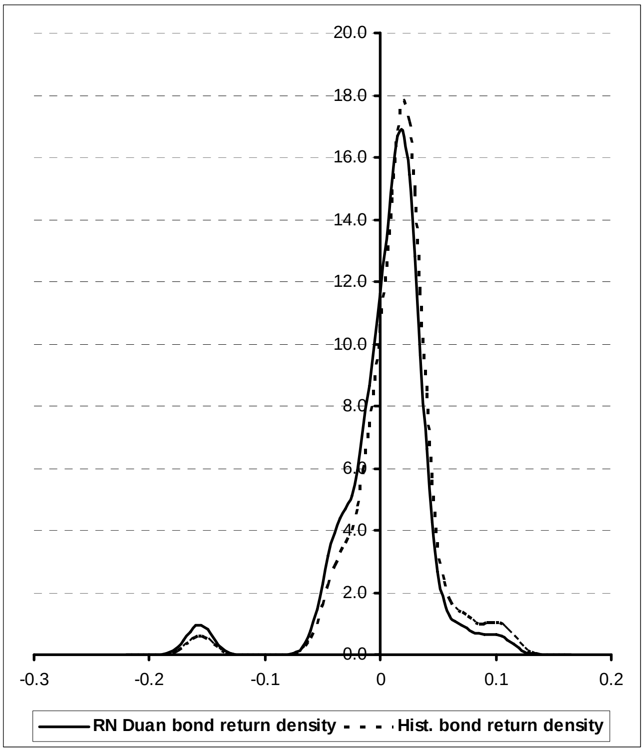

Figure 3: Risk-neutral and historical bond return densities

3.3 Estimating the dependence structure

As stated earlier, we model dependence using copulas. It is a straightforward matter to obtain the empirical copula directly from monthly joint returns data (see Appendix A.3). Figure 4 shows a contour diagram of the market copula. This is jagged, because there are only 60 sample points (monthly data for five years).

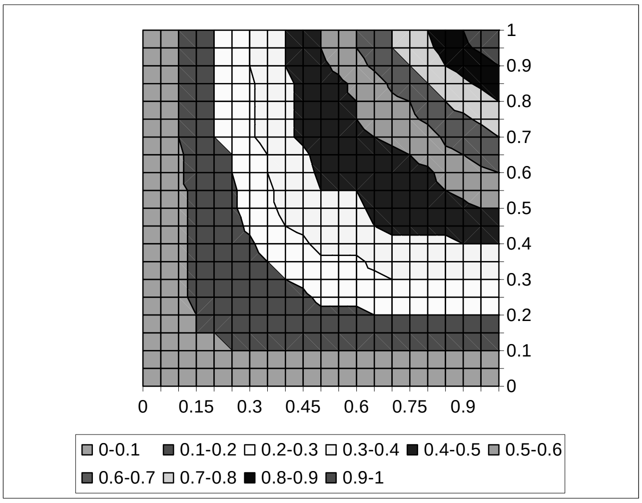

Figure 4: Contour diagram of copula obtained from historical price data

To obtain a smoother version of the risk-neutral joint distribution, proceed as follows: given assets S, B, let $F_{S,B}, F_S$ and F_B be the "real world" joint distribution function and its univariate margins, and let $\bar{F}_{S,B}, \bar{F}_S, \bar{F}_B$ denote the corresponding risk-neutral entities. Then

$$F_{S,B}(x,y) = C(F_S(x), F_B(y)) \quad \bar{F}_{S,B}(x,y) = C(\bar{F}_S(x), \bar{F}_B(y))$$

so that

$$\bar{F}_{S,B}(x,y) = F_{S,B}(F_S^{-1}(\bar{F}_S(x)), F_B^{-1}(\bar{F}_B(y)))$$

i.e. the risk-neutral joint distribution can be obtained directly from the historical joint-and marginal distributions, and the estimated risk-neutral marginal distributions. Now apply kernel smoothing techniques to obtain reasonable versions of $F_{S,B}, F_S$ and F_B, and thus also of $\bar{F}_{S,B}$. Figure 5 shows the risk-neutral joint density obtained by this method. For comparison, Figure 6 shows the difference between the estimated (non-parametric) joint price density and the corresponding density within the Black-Scholes framework.

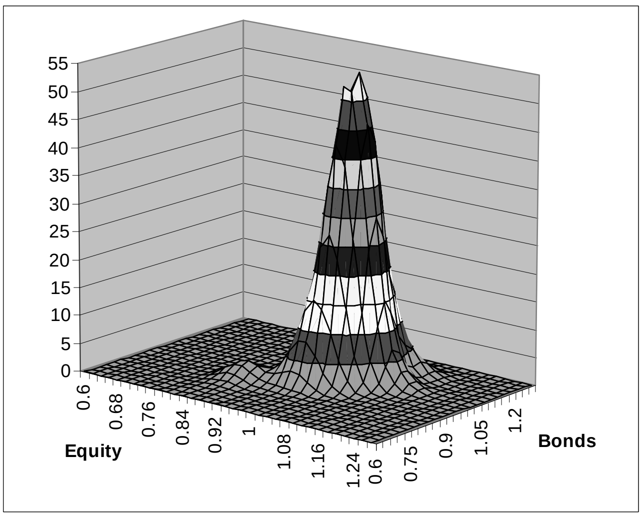

Figure 5: Risk-neutral joint price density for equity and bonds

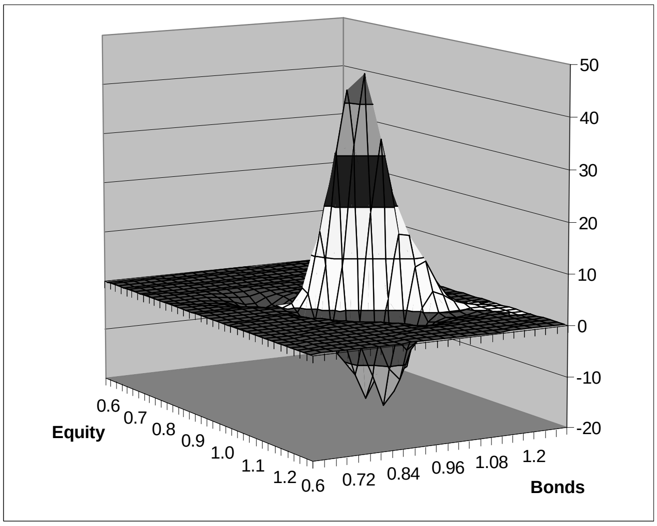

Figure 6: Difference between the Black-Scholes density and the nonparametric density

4. OPTION PRICING

Recall the principle of risk-neutral valuation: If X_t is the time-t value of a European-style derivative, with payoff $f(S_T, B_T)$ at expiry T, then

$$X_0 = e^{-rT}\,\bar{E}[f(S_T, B_T)] \ .$$

(where $\bar{E}$ denotes the expectation under the risk-neutral measure). To price options by simulation, construct a large number N of draws (U_n, V_n) from a bivariate distribution with uniform marginals and market copula C (see Appendix A.4). Then

$$S_n = \bar{F}_S^{-1}(U_n)$$

$$B_n = B_0 \exp[\sigma G^{-1}(\Phi(\Phi^{-1}(V_n) + \lambda)) + \mu]$$

are N draws from the risk-neutral joint distribution of (S_T, B_T) with copula C. (Refer to Appendix A.2 for an explanation of the symbols $\sigma, \mu, G, \lambda, \Phi$.) The option price is estimated by

$$X_0 = \frac{e^{-rT}}{N}\sum_{n=1}^{N} f(S_n, B_n)$$

To increase the accuracy of the simulations, the methods of antithetic variates and control variates were implemented (Glasserman, 2004).

Various bivariate options were priced, and the prices compared with their Black-Scholes counterparts. Here, prices have been normalized (i.e. $S_0 = 1 = B_0$), and expiry is one month (i.e. $T = \frac{1}{12}$).

- Exchange (Margrabe) option with payoff $\max(S_T - B_T, 0)$:

  The Black-Scholes and nonparametric prices were, respectively, BS: 0.0199 and NP: 0.0219 i.e. the NP price is approximately 10% greater than the BS price.

- Rainbow call on the maximum with payoff $\max(\max(S_T, B_T) - K, 0)$: See Figures 7 and 8.

- Rainbow call on the minimum with payoff = $\max(\min(S_T, B_T) - K, 0)$: See Figure 9.

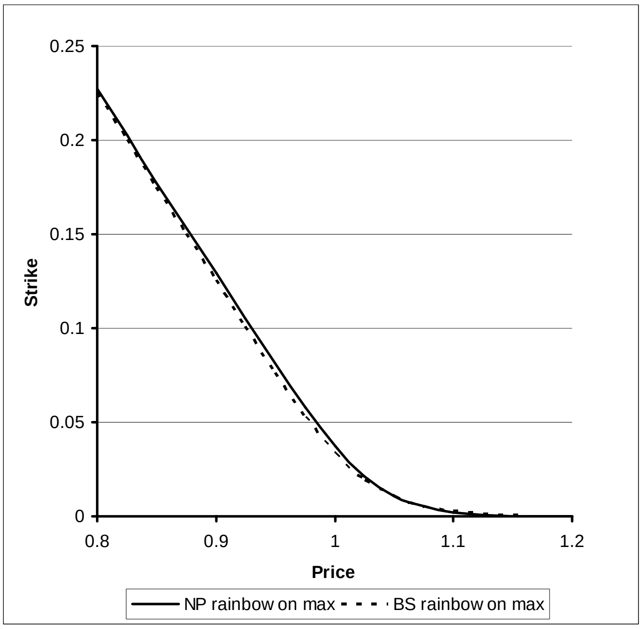

Figure 7: Black-Scholes and nonparametric prices for rainbow call options on the maximum.

5. CONCLUSIONS

We have obtained bivariate option prices nonparametrically by Monte Carlo simulation. For near-the-money options, the Black-Scholes and nonparametric prices differences were in the range of 0-25%, and often larger far away from the money. (But note that simulated distributions are likely to be inaccurate in the tails).

The methods presented in this paper suffer from several maladies:

- The assumption that Duan's risk-neutral marginal distribution corresponds to the market risk-neutral distribution is not entirely justifiable. Nor is the assumption that the market and risk-neutral copulas are the same.

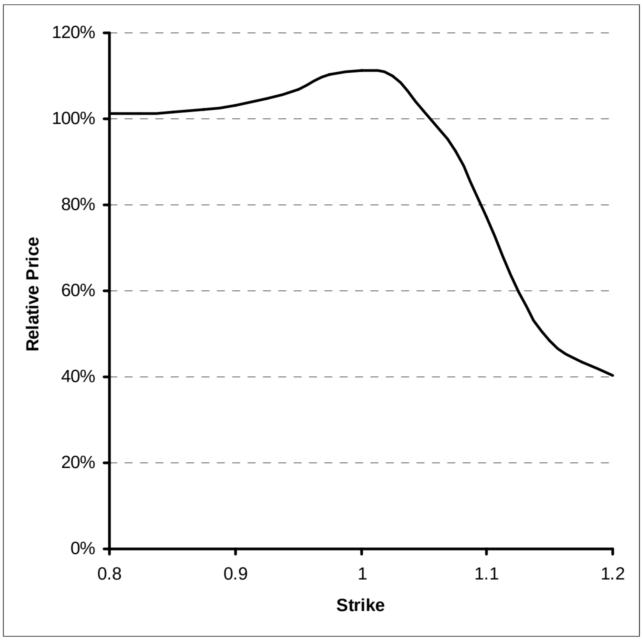

Figure 8: Ratio of nonparametric and Black-Scholes prices for calls on the maximum

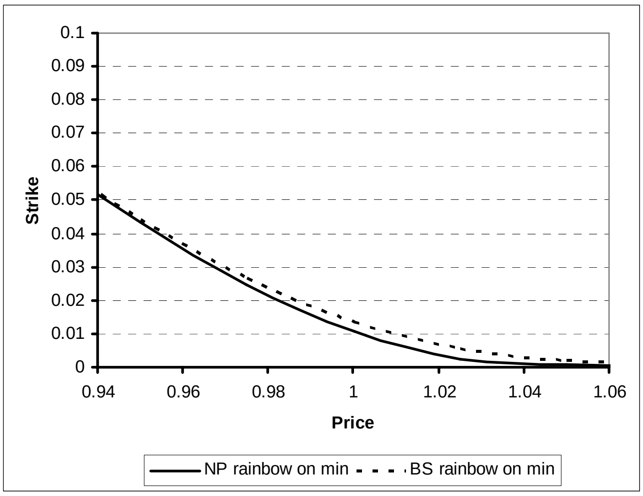

Figure 9: Black-Scholes and nonparametric prices for rainbow call options on the minimum

- To price 1-month options, we must estimate a copula obtained from monthly returns data. We used a 5-year period, resulting in 60 sample points. To price 3-month options, we'd have to use a copula obtained from 3-monthly returns data. One would have to use data going back 15 years to get 60 sample points. This is an unrealistically large time period, because copulas change over time (Van den Goorbergh *et al.,* 2005).

- Risk-neutral valuation of contingent claims works only when the claims can be hedged. Without some assumption on the joint dynamics of asset prices, it is difficult to construct hedging portfolios. Thus we have obtained prices for the bivariate claims, but haven't indicated how to hedge them, and thus given no indication of how to extract arbitrage profits should the options be mispriced in the market.

- The methodology presented above is computationally very expensive. It is not clear that it is worthwhile.

REFERENCES

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

Bingham N and Kiesel R. 2004. *Risk-neutral Valuation*. Springer-Verlag.

Breeden D and Litzenberger RH. 1978 Prices of state-contingent claims implicit in option prices. *Journal of Business,* 51:621-651.

Coutant S, Durrleman V Rapuch T and Roncalli G. 2001. Copulas, multi-variate risk-neutral distributions and implied dependence functions. Unpubl. manuscript, http://gro. creditlyonais.fr/content/rtd/home_copulas.htm.

Duan JC. 2002. Option pricing by transformation. Unpubl. manuscript.
http://www. rotman.utoronto.ca/~jcduan.

Embrechts P, Lindskog F and MacNeil A. 2001. Modelling dependence with copulas. Unpubl. manuscript. http://www.ethz.ch/finance.

Embrechts P, MacNeil A and Straumann D. 2002. Correlation and dependence in risk management: pitfalls and properties, chapter in *Risk Management: Value at Risk and Beyond,* 176-223. Cambridge University Press.

Glasserman P. 2004. *Monte Carlo Methods in Financial Engineering.* Springer-Verlag.

Jackwerth J. 1999. Option implied risk-neutral distributions and implied binomial trees: a literature review. *Journal of Derivatives*, 7:66-82.

Nelsen RB. 1999. *An Introduction to Copulas.* Springer-Verlag.

Ouwehand P and West G. 2006. Pricing rainbow options. *Wilmott Magazine,* May 2006

Rosenberg JV. 2003. Nonparametric pricing of multivariate contingent claims. *Journal of Derivatives*, 10:9-26.

Silverman BW. 1986. *Density Estimation for Statistics and Data Analysis*. Chapman and Hall.

Van den Goorbergh R, Genest C and Werker B. 2005. Bivariate option pricing using dynamic copula models. *Insurance Mathematics and Economics*, 37:101-114.

West G. 2005. *South African Financial Markets*. Unpubl. notes. http://www.finmod.co.za

APPENDIX: THEORETICAL COMPLEMENTS

A.1 The method of Breeden and Litzenberger (1978)

This is a method for obtaining a market-implied risk-neutral distribution from option prices: Let $F_{t,T}$ denote the time t-futures price of an asset S, for delivery at time T, so that $S_T = F_{t,T}$. Let $C(x,T)$ be the time $t = 0$-price of a futures call option with strike x and maturity T.

Then $C(x,T) = \bar{E}[\max(F_{T,T} - x, 0)] = \bar{E}[\max(S_T - x, 0)]$ (cf. West (2005) for a description of SAFEX futures options)

Using the fact that if X is a non-negative random variable, then $E[X] = \int_0^\infty P(X > t)\ dt$, we see that

$$\frac{\partial C}{\partial x} = \frac{\partial}{\partial x}\int_x^\infty \bar{P}(S_T > t)dt = F(x) - 1$$

where
$F(x)$ is the risk-neutral distribution function of S_T. Hence

$$\frac{\partial^2 C}{\partial x^2} = f(x)$$

is the risk-neutral density of S_T.

A.2 The method of Duan (2002)

Duan (2002) shows how to obtain *a* univariate market risk-neutral asset price distribution directly from (historical) market prices. Briefly, the method works as follows:

- Assume the historical sequence $Y_t = \frac{R_t - \mu_t}{\sigma_t}$ is i.i.d., where R_t is the 1-period log-return of stock S, i.e. $S_t = S_{t-1}\exp(R_t)$, and μ_t, σ_t^2 are the conditional mean and variance of R_t.

- Let G be distribution function of Y_t, and let Φ be the standard normal distribution function. Then $Z_t = \Phi^{-1}(G(Y_t))$ are independent standard normal

random variables, and we have "real world" asset price dynamics
$S_t = S_{t-1}\exp(\sigma_t G^{-1}(\Phi(Z_t)) + \mu_t)$

- Now find a *new* density f for Z_t which corresponds to a risk-neutral return, and *minimizes the relative entropy* with respect to the standard normal density ϕ: Solve
$\min_f \int_{-\infty}^{\infty} f(x)\ln\frac{f(x)}{\phi(x)}dx$
subject to
$\int_{-\infty}^{\infty} xf(x)dx = c_t$ and
$\int_{-\infty}^{\infty} f(x)dx = 1$

The value of c_t depends on the mean that we want the new density to have.

- It is known (Duan, 2002) that the solution to the above optimization problem is of the form $f(x) = \phi(x - \lambda_t)$ where λ depends on c.

- To ensure that corresponding asset distribution asset is *risk-neutral*, we must ensure that the expected rate of return of S under the new density f is the riskless rate, i.e. we must calculate λ so that
$$\int_{-\infty}^{\infty} \exp(\sigma_t G^{-1}(\Phi(x)) + \mu_t)\phi(x - \lambda_t)dx = e^{r_t}$$
where r_t is the 1-period riskless rate. Finding λ will therefore require numerical integration techniques; we used Gauss-Legendre quadrature, and also Romberg's method.

- Consequently, the risk-neutral asset price dynamics are
$S_t = S_{t-1}\exp(\sigma_t G^{-1}(\Phi(Z_t)) + \mu_t)$
where Z_t is a normal random variable with mean λ_t and variance 1.
- This allows one to simulate risk-neutral asset prices, and hence extract the risk-neutral price distribution.

A.3 Copulas

Correlation is not a good measure of dependence for non-Gaussian distributions – see (Embrechts et al., 2002). To model dependence, we use copulas. Nelsen (1999) is a good introduction to copulas, and the source for the ideas expounded below. Briefly, a (bivariate) copula is a function that *couples* two marginal distribution functions into a joint distribution function: Specifically, if F_S, F_B and $F_{S,B}$ are, the marginal and joint distributions of S and B, then

$$F_{S,B}(x,y) = C(F_S(x), F_B(y))$$

i.e. feed the marginal distributions into the copula, and you get the joint distribution. If one thinks joint distribution = marginal distributions + dependence structure it is clear that C captures the dependence structure between S and B. Sklar's Theorem (cf. Nelsen (1999)) states that every joint distribution function has a copula. This is easy to see if F_S, F_B are invertible: If C is the joint distribution function of $F_S(S), F_B(B)$, then

$$F_{S,B}(x,y) = P(S \le x, B \le y) = P(F_S(S) \le F_S(x), F_B(B) \le F_B(y)) = C(F_S(x), F_B(y))$$

(because $F_S(S), F_B(B)$ are uniform-[0, 1] random variables).

It is easy to see that if f, g are increasing functions, then the pairs (X, Y) and $(f(X), g(Y))$ have the same copula, and thus the same dependence structure. In particular, equity-bond prices and equity-bond returns have the same copula. Correlation is not stable under increasing transformations however, which is another indication that it is unsuitable as a measure of dependence.

If $(x_1, y_1), \ldots, (x_n, y_n)$ is a sample from a bivariate distribution, then the *Deheuvels* or *empirical* copula is given by

$$C(\tfrac{i}{n}, \tfrac{j}{n}) = \frac{\text{no. of pairs } (x,y) \text{ in sample such that } x \le x_{(i)}, y \le y_{(i)}}{n}$$

where $x_{(i)}, y_{(i)}$ are the order statistics of the sample.

A.4 Random number generation

There is a simple algorithm to generate random pairs (U, V) with uniform margins and copula C (Nelsen, 1999):

- Let $c_u(v) = P(V \le v \mid U = u) = \frac{\partial C(u,v)}{\partial u}$;

- Generate two independent uniform-[0,1] variates U, W;

- Set $c_u^{-1}(W)$, where c_u^{-1} is a quasi-inverse of c_u. (U, V) is the required pair.

To obtain random pairs (S, B) with marginal distributions F_S, F_B and copula C is straightforward: Since F_S, F_B are increasing functions, the pairs (U, V) and $(F_S^{-1}(U), F_B^{-1}(V))$ have the same copula. Since $F_S^{-1}(U)$ has distribution F_S and $F_B^{-1}(V)$ has distribution F_B we see that $(F_S^{-1}(U), F_B^{-1}(V))$ is the required pair.

A.5 Kernel smoothing

In order to get reasonably smooth densities, we used kernel smoothing (Silverman, 1986). To estimate the density f of a random variable X from a sample $X_1, \ldots, X_n$, a naive estimate is obtained by putting

$$\tilde{f}(x) = \frac{\text{no. of } X_1, \ldots, X_n \in (x-h, x+h)}{2nh} = \frac{1}{n}\sum_{k=1}^{n} w\left(\frac{x - X_k}{h}\right)$$

where w is the uniform density on [-1,1]. Instead of using the uniform density w, which has sharp cut-offs, one may also use smoother densities, which lead to smoother estimates, e.g. the standard normal density $w(x) = \frac{e^{-x^2/2}}{\sqrt{2\pi}}$. The smoothing parameter (or bandwidth) h must be selected to so that $\tilde{f}$ looks sufficiently smooth, but still retains the main features of the unsmoothed estimate.