Dynamic Density Forecasts for Multivariate Asset Returns

Arnold Polanski
Evarist Stoja

September 2009

Department of Economics
University of Bristol
8 Woodland Road
Bristol BS8 1TN

Dynamic Density Forecasts for Multivariate Asset Returns

Arnold Polanski, *Queen's University Belfast*

Evarist Stoja, *University of Bristol*

September 2009

Abstract

We propose a simple and flexible framework for forecasting the joint density of asset returns. The multinormal distribution is augmented with a polynomial in (time-varying) non-central co-moments of assets. We estimate the coefficients of the polynomial via the Method of Moments for a carefully selected set of co-moments. In an extensive empirical study, we compare the proposed model with a range of other models widely used in the literature. Employing a recently proposed technique to evaluate multivariate forecasts, we conclude that the augmented joint density provides highly accurate forecasts of the negative tail of the joint distribution.

Keywords: Time-varying higher co-moments; Joint Density Forecasting, Method of Moments, Multivariate Value-at-Risk.

JEL classification: C22, C51, C52, G11

Address for correspondence: *Arnold Polanski*, Queen's University Management School, 25 University Square, Queen's University Belfast, Belfast, BT7 1NN, UK. Tel: +44(0)2890975024, Email: a.polanski@qub.ac.uk. *Evarist Stoja*, School of Economics, Finance and Management, University of Bristol, 8 Woodland Road, Bristol, BS8 1TN, UK. Tel: +44(0)1173310603, Email: e.stoja@bristol.ac.uk.

1. Introduction

The finance literature is replete with models and applications of point forecasts. For example, the celebrated ARCH model of Engle (1982) projects the next period's level of volatility as a function of current and lagged squared returns. However, the amount of information contained and exploited in point forecasts is considerably lower than in the case of interval or density forecasting (see, for example, Christoffersen, 1998). A density forecast is an estimate of the probability distribution of the possible realizations of a variable, thereby providing a full description of the uncertainty associated with the forecast. This contrasts sharply with the point forecast which, by definition, provides no such information. This simple observation in itself provides a strong case for density forecasting. Moreover, the recent prominence of the risk management industry which heavily relies on density forecasting has strengthened this approach. Indeed, financial companies such as Reuters, Bloomberg, J.P. Morgan regularly provide density forecasts for their clients. The aim of this practice is to provide the user with a procedure which generates density forecasts of tailored portfolio returns over a specified horizon. One example is the Standard Portfolio Analysis of Risk (SPAN) framework which has become an industry standard for calculating margin requirements for customers and clearing house members. SPAN is essentially a mixture of stress tests performed on each underlying asset in the portfolio (see, for example, Artzner et al., 1999). Moreover, density forecasts are also of particular interest to regulators where the obvious example is Value-at-Risk (VaR), which since its adoption by the Basle Committee on Banking

2

Supervision (1996) has become the most widely used risk measurement tool in the banking sector.

As already pointed out, the ARCH model is mainly concerned with point forecasting. Under certain circumstances, it is possible to construct density forecasts with the volatility estimate obtained from an ARCH model and, hence, it could be argued that the former stems from the latter. Indeed, under normal or, more generally, elliptical distributions, one can obtain a forecast for the entire distribution from a volatility forecast. The simplifying assumption of elliptically distributed returns (in some cases augmented with a time-varying variance), which ensures tractability, is paramount in finance with wide ranging applications in risk management, asset and option pricing and portfolio decisions. There is however, increasing evidence that asset returns are not elliptically distributed. For example, Singleton and Wingender (1986) find evidence of skewness in the distribution of stock returns. Simkowitz and Beedles (1978) show that the degree of skewness preference of investors will impact on the extent of their diversification since the higher the degree of skewness, the fewer assets investors will hold given that diversification reduces the skewness of a portfolio, while Cotner (1991) documents the impact of asymmetries on option prices. More recent evidence on non-normality of return distribution is provided *i.a.* by Bae et al. (2003), Bali and Weinbaum (2007), and Polanski and Stoja (2009a). When the normality assumption is violated then a specification and forecast of the entire conditional joint distribution is necessary. More generally, decision making under uncertainty with asymmetric loss function and non-

Gaussian variables involves density forecasts. Applications of density forecasting in finance and economics are surveyed and discussed in Tay and Wallis (2000).

The literature on modeling the entire density of a random variable is more limited. Gallant et al. (1991) employ a semi-parametric framework to forecast the density which relies on a series expansion of the normal. Hansen (1994) proposes the auto regressive conditional density model which employs a skewed student-t conditional distribution. Other approaches to modelling the conditional density of returns include the exponential generalized beta of McDonald and Xu (1995) and skewed generalized-t distribution of Theodossiou (1998). Another strand of literature on density forecasting originates in the seminal contribution of Breeden and Litzenberger (1978) and relies on extracting the implied density from the option prices via the Black-Scholes pricing model (see, for example, Fackler and King, 1990; Jackwerth and Rubinstein, 1996; Bahra, 1997).

The cited literature is exclusively focused on modeling the univariate density of returns. However, financial decision making usually involves more than one risky asset. The standard practice here, as in the early univariate density literature, is to assume that the density is multinormal with a possibly time-varying covariance matrix (see Diebold et al., 1999). In this case, the correlation coefficient adequately captures the dependence between assets. However, joint normality is not supported by empirical evidence (see, for example, Guidolin and Timmermann, 2006). Moreover, correlation is only a measure of linear dependence and suffers from a number of limitations (see Embrechts et al., 2002). For example, Patton (2004) argues that the dependence between assets is stronger during

market downturns than during market upturns while Polanski and Stoja (2009a) show that the probability of extreme events, as measured by the tail thickness, varies over time. These findings imply that results obtained via the standard elliptical distributions would generally be invalid.

To address this shortcoming, Patton (2007) advocates the copula approach in which the multivariate density is modelled as the product of the marginal densities of variables and the copula function which captures the dependence between the variables. The copula technique is flexible as the marginal densities can all be different from each other (and from the copula). However, in the finance literature it is mainly employed in the bivariate case. Recent attempts to generalise it to the multivariate case turned out to be technically and computationally demanding which detracts somewhat from their usefulness (see, for example, Aas et al., 2009).

In this paper, we propose a novel technique to modelling the multivariate density of asset returns. We approximate the joint density as the product of a multivariate normal and a polynomial which adjusts it for the estimated time-varying (co-)moments of the variables of interest. We estimate the coefficients of the polynomial via the Method of Moments (MM). This approach provides a flexible tool for modelling the empirical joint distribution of financial data, which in addition to volatility, exhibits time-varying higher (co-)moments. While our method maintains the simplicity of the multivariate normal distribution, it can be easily adapted to more complex distribution functions and generalized moments. Furthermore, we extend the extant literature by employing the

MM-estimated augmented joint density (AJD) to forecast multivariate VaR (MVaR). Similar to its univariate counterpart, MVaR determines the probability of extreme losses for two or more assets.

Our analysis employs daily returns of highly traded stock indices and exchange rates. The results suggest that the time-varying conditional (co-)moments are very important in characterising the joint return distributions and yield significant improvements in the forecasting of MVaR. While a specifically developed test shows that our density forecasts cannot adequately approximate the distribution over the entire domain, they are successful in approximating the negative tail of the joint distribution of returns which is the focus of risk management.

The outline of the remainder of this paper is as follows. In Section 2 we discuss the theoretical framework for the MM multivariate density forecasting. Section 3 describes the statistical evaluation method. Section 4 presents the data and the empirical results while Section 5 concludes.

2.   Theoretical Framework

We approximate the multivariate distribution of a vector $X = (X_1,...,X_n)$ of returns by a polynomially-adjusted multinormal probability density function (pdf) $\hat{f}(x)$. If $f(x;\mu,S)$ is the n-dimensional normal pdf with mean μ and the variance-covariance matrix S, then the AJD $\hat{f}(x)$ is the product of $f(x;\mu,S)$ and a polynomial in $x = (x_1,...,x_n)$,

6

$$\widehat{f}(\mathrm{x}) = \boldsymbol{j}(\mathrm{x}; \boldsymbol{m}, \Sigma)(\sum_{s} \boldsymbol{l}_s x_1^{s_1} ... x_n^{s_n})^2 \qquad (1)$$

where $s = (s_1,..., s_n) \in N^n$ is a vector of exponents and $\boldsymbol{l}_s$ is the coefficient of the term $x_1^{s_1} ... x_n^{s_n}$. Note that the polynomial, and hence the pdf $\widehat{f}(x)$, assumes only non-negative values. We estimate the coefficients $\boldsymbol{l}_s$ via the MM procedure. From historical returns $x = (x_1,..., x_n)$, where x_i is the return vector of asset $i = (1,..., n)$, we compute the co-moment estimates $\widehat{E}[x_1^{v_1} ... x_n^{v_n}]$ for a set V of exponent vectors $v = (v_1,..., v_n) \in N^n$ and then solve the multivariate quadratic system,

$$E_{\widehat{f}}[X_1^{v_1} ... X_n^{v_n}] = \widehat{E}[x_1^{v_1} ... x_n^{v_n}] \quad \text{for all} \quad v \in V, \qquad (2)$$

for the coefficients $\boldsymbol{l}_s$, where $E_{\widehat{f}}[.]$ is the expectation operator with respect to $\widehat{f}$. Since solving quadratic equations is in general an NP-hard problem (with several applications in cryptography), we can approximate the solution to (2) for a large number of assets by minimizing the weighted sum of squared deviations,

$$\sum_{v} w_v (E_{\widehat{f}}[X_1^{v_1} ... X_n^{v_n}] - \widehat{E}[x_1^{v_1} ... x_n^{v_n}])^2, \qquad (2a)$$

with respect to the coefficients $\mathbf{l}_s$, where $\mathbf{w}_v$ is the weight of the co-moment corresponding to v. Note that for a sufficiently high weight $\mathbf{w}_v$ for $v = (0,...,0)$, $\widehat{f}(x)$ integrates to one,

$$\mathbf{w}_v (E_{\widehat{f}}[X_1^0...X_n^0] - \widehat{E}[x_1^0...x_n^0])^2 = \mathbf{w}_v (\int_{R^n} \widehat{f}(r)dr - 1)^2 \downarrow 0 \; as \; \mathbf{w}_v \uparrow$$

The estimation of $\widehat{f}(x)$ requires, therefore, a set of non-central co-moments, each defined by a vector of exponents v. In the empirical applications of Section 4, each co-moment vector v in (2) is also used as a co-moment vector s in the polynomial part of (1) and vice versa. The co-moments employed in the MM estimation (2) mirror, therefore, the polynomial terms in the forecast (1), although this is not a necessary condition for the applicability of our model.

The AJD (1) provides a flexible modelling framework. Interestingly, by including in it higher co-moments, the multinormal distribution can be tailored to the specific features of financial data such as fat tails, joint asymmetry and – in a slightly modified version – asymmetric dependence among assets.[1]

---

[1] For example, the AJD (1) could be modified to capture the higher dependence among assets during market downturns than during market upturns. To this end, we would include a term which is activated whenever all returns are negative as illustrated in the bivariate density forecast $\widehat{f}(x) = \mathbf{j}(x_1, x_2;.)(... + \mathbf{l}_{11}x_1x_2 \cdot I(x_1 < 0 \; \& \; x_2 < 0) + ...)^2$.

The model allows also for the incorporation of time-varying co-moments into dynamic forecasts. For example, the exponentially weighted moving average (EWMA) specifies the next period's k^{th} moment $u_{k,t+1}$ of the portfolio return to be a weighted average of the current k^{th} moment $u_{k,t}$ and the current actual return x_t, raised to the power of k,

$$u_{k,t+1} = g_k u_{k,t} + (1-g_k)x_t^k = (1-g_k)\sum_{i=0}^{\infty} g_k^i x_{t-i}^k \qquad (3)$$

Therefore, an EWMA moment can be interpreted as a weighted average of past returns, raised to the appropriate power, encapsulating information from all past shocks with exponentially declining importance attached. This can be extended to n assets and non-central co-moments defined by a vector of exponents $v = (v_1,...,v_n) \in N^n$,

$$u_{1,t+1}^{v_1}....u_{n,t+1}^{v_n} = g_v u_{1,t}^{v_1}....u_{n,t}^{v_n} + (1-g_v)x_{1,t}^{v_1}....x_{n,t}^{v_n} = (1-g_v)\sum_{i=0}^{\infty} g_v^i x_{1,t-i}^{v_1}....x_{n,t-i}^{v_n} \qquad (4)$$

The forecast $u_{1,t+1}^{v_1}....u_{n,t+1}^{v_n}$ can then be employed as the co-moment estimate $\widehat{E}[x_1^{v_1}...x_n^{v_n}]$ in the MM optimization (2). The polynomial coefficients that minimize (2), together with the estimates of the mean and variance of the multinormal part in (1), fully define the density forecast $\widehat{f}$. Note that the forecasts are dynamic in the sense that they evolve over time. In the next section, we discuss the evaluation of multidimensional density forecasts.

3. Statistical Evaluation Method

A correct density forecast should be unconditionally and conditionally accurate. Unconditional accuracy of a continuous univariate (time-varying) density forecast $\widehat{f}_t$ implies that the frequency of observations x_t, for which the probability integral transformation (PIT) $z_t := \widehat{F}_t(x_t) = \Pr(X_t < x_t)$ is less than $a \in [0,1]$, approaches a as the sample size T increases (Rosenblatt, 1952). In other words, the PIT sequence $\{z_t\}_{t=1}^{T}$, computed from observations $\{x_t\}_{t=1}^{T}$, must be uniformly distributed on the unit interval when the forecasts are correct. Polanski and Stoja (2009b) showed that this criterion does not generalize directly to multivariate density forecasts. The PIT sequence $z_t = \widehat{F}_t(x_t = (x_{1,t},...,x_{n,t}))$ from multidimensional density forecasts $\widehat{f}_t$ is not necessarily uniform even when the forecasts are correct. However, a simple modification in the PIT computation restores the uniformity of the scores. First, each observation $x_t = (x_{1,t},...,x_{n,t})$ is transformed into a vector $x_t^M := Max(x_{1,t},...,x_{n,t}) * (1,....,1)$ and then the score $z_t^M = \widehat{F}_t(x_t^M)$ is computed. Note that for unidimensional forecasts, z_t^M and the standard PIT z_t are identical. Polanski and Stoja (2009b) proved that the scores z_t^M are *i.i.d.* according to the uniform distribution U[0,1] if the AJDs $\widehat{f}_t$ are correct.

Testing the conditional accuracy of forecasts entails the proof that the current score z_t^M does not convey any information on the score in the next period or, alternatively, that the scores z_t^M are distributed independently across the time. The unconditional and conditional

accuracy can be, therefore, tested with the usual goodness-of-fit tests (see Noceti et al., 2003 for a comparison of the existing tests) and serial independence tests (e.g., portmanteau test) on z_t^M-scores.

From the latter scores, we can also compute the exceedance rates for the MVaR. For a unidimensional cdf $\widehat{F}_t$, the VaR at the 1-a coverage level is defined as (the negative of) the quantile q_t^a for which $\widehat{F}_t(q_t^a) = a$. By analogy, for an n-dimensional cdf $\widehat{F}_t$, we require that the MVaR $(q_t^a,...,q_t^a)$ satisfies the condition $\widehat{F}_t((q_t^a,...,q_t^a)) = a$. Then, from the definition $z_t^M = \widehat{F}_t(x_t^M)$ it can be deduced that z_t^M is less than a whenever all components of the observation $x_t = (x_{1,t},...,x_{n,t})$ exceed the critical value q_t^a,

$$z_t^M < a \iff x_t < (q_t^a,...,q_t^a).$$

As in the univariate case, the assessment of MVaR forecasts is inherently difficult. Since the actual MVaR is not observable, the accuracy of the forecasted MVaR cannot be directly evaluated. One can, however, compute the proportion of observations, which exceed the MVaR forecast and compare this number with a required significance level. We refer to this procedure as unconditional accuracy. On the other hand, the conditional accuracy requires that the number of observations that exceed the MVaR forecast should be unpredictable when conditioned on the available information. To assess both types of accuracy, we employ the unconditional accuracy test of Kupiec (1995) and the conditional accuracy test of Christoffersen (1998). Although both tests are designed for

11

univariate densities, they still apply for joint distributions, because the score computation has effectively converted a multivariate problem into a univariate one. Indeed, as the exceedence of the MVaR at level a is equivalent to $z_t^M < a$, the test statistics for both tests can be computed directly from the z_t^M-scores.

4. Empirical Study

Data

To evaluate the performance of our model, we tested out-of-sample the forecasts for the joint distribution of the daily returns of S&P500, Dow Jones and Nasdaq equity indices as well as the forecasts for the joint distribution of the exchange rates of GBP, CHF and JPY measured against the USD. We investigated also the performance of our model in the inter-temporal dependence context by modelling the joint distribution of returns of an asset over five consecutive business days.

Table 1 (Panel A) presents summary statistics for the continuously compounded daily return series of equity indices computed from the raw prices. The mean returns are almost identical for all series, and close to zero. The ARCH(4) portmanteau test for up to fourth order serial correlation in squared returns shows that all returns display significant volatility clustering and are highly leptokurtic (which is consistent with the existence of time-varying volatility). The returns for Dow Jones are slightly negatively skewed while for the remaining two series the returns are positively skewed. The three return series are relatively strongly correlated, with the highest correlation between S&P500 and Dow Jones, and the lowest between Dow Jones and Nasdaq. In line with previous evidence for

daily returns, the null hypothesis of normality is strongly rejected by the Bera-Jarque statistic. Panel B in Table 1 reports the same summary statistics for the exchange rates, where, with the exception of the correlation coefficients which are significantly lower, similar observations apply. In particular, the null hypothesis of normality is again strongly rejected for the exchange rates.

[Table 1]

The joint distribution of the log returns of n assets,

$$x_t = (x_{1,t},...,x_{n,t}) = 100\,(\ln(\,p_{1,t}\,/\,p_{1,t-1}),...,\ln(\,p_{n,t}\,/\,p_{n,t-1}))\,,$$

where $p_{i,t}$ is the closing price at date t, were forecasted according to three models.

In the basic model N, the forecast $\widehat{f}_{t+1}$ at date $t+1$ was simply the multinormal pdf with zero mean and the variance-covariance matrix estimated from the historical returns in the moving window $(r_{t-T},...,r_t)$.

Model N2 used the EWMA specification (4) to forecast the second (co-)moments,

$$\mathrm{u}_{1,t+1}^{v_1}....\mathrm{u}_{n,t+1}^{v_n} = \boldsymbol{g}_{v,t}\,\mathrm{u}_{1,t}^{v_1}....\mathrm{u}_{n,t}^{v_n} + (1-\boldsymbol{g}_{v,t})\mathrm{x}_{1,t}^{v_1}....\mathrm{x}_{n,t}^{v_n} \qquad (5)$$

for all vectors $v = (v_1, \ldots, v_n)$ such that $v_1 + \ldots + v_n = 2$ and with the parameters $\boldsymbol{g}_{v,t}$ in (5) set to minimize the squared sum of the historical forecast errors in the moving window $(x_{t-T}, \ldots, x_t)$,[2]

$$\sum_{s=t-T}^{t} (u_{1,s}^{v_1} \ldots u_{n,s}^{v_n} - x_{1,s}^{v_1} \ldots x_{n,s}^{v_n})^2 \tag{6}$$

The forecasted latter moments were used to construct the variance-covariance matrix Σ_{t+1} for the AJD $\hat{f}_{t+1} = N(0, \Sigma_{t+1})$. N2 corresponds, therefore, to the EWMA-multinormal model (see Diebold et al., 1999).

Finally, in the model N24, the forecasted multinormal pdf from N2 was multiplied by the polynomial term,

$$(\sum_v \boldsymbol{l}_v x_1^{v_1} \ldots x_n^{v_n})^2 \quad \text{for all } v = (v_1, \ldots, v_n) \text{ such that } v_i \in \{0,2,4\}, v_1 + \ldots + v_n \in \{0,4\} \tag{7}$$

that included the fourth (co-)moments.[3] The coefficients $\boldsymbol{l}_n$ in the polynomial (7) were estimated by MM with the moment forecasts (5) for the vectors of exponents v from (7).

---

[2] Polanski and Stoja (2009a) show that this procedure is equivalent to the ML estimation of $\boldsymbol{g}_{v,t}$.

As in N2, the parameters $\boldsymbol{g}_{v,t}$ in (5) were set to minimize the squared sum of historical errors (6).

Results

The out-of-sample evaluation of density forecasts was based on the scores $z_t^M = \hat{F}_t(x_t^M)$, where $x_t^M = Max(x_{1,t},...,x_{n,t}) * (1,....,1)$. For all three models and for all data sets under study, the Pearson's $\boldsymbol{c}^2$ test strictly rejected the null of uniform distribution of z_t^M-scores with the resulting p-values virtually equal to zero in each case. None of the models generates, therefore, an acceptable forecast for the multivariate distribution of returns over the whole domain. However, risk managers and regulators are interested generally in the likelihood of large losses. If this is the case, then a model which accurately describes the extreme events while failing in the interior of the distribution will not be rejected. In other words, a model superior in forecasting the central part of the distribution will be eschewed in favour of another model which accurately forecasts the negative tail. This objective motivates the censored likelihood test of Berkowitz (2001), in which the observations not falling into the negative tail of the distribution (with cut-off point being decided by the user's requirements) are truncated.

---

[3] We also experimented with a polynomial including the odd co-moments, but in our datasets they were found to have little impact on the negative tail of the joint density which was our main focus.

Henceforth, we focus only on the negative tail of the distribution which is the relevant part of the distribution for risk management purposes. We observed that the MVaR forecasts from N24 were accurate for relatively low values of the nominal level a in all data sets under study. Note that the latter MVaR are of particular importance for the assessment of the joint risk of financial assets of interest.

In Table 2, we report the actual MVaR exceedance rates for different nominal levels a, together with Kupiec's (1995) unconditional (t_u) and Christoffersen's (1998) conditional (LR_c) test statistics for the joint density of S&P500, Dow Jones and Nasdaq indices. The simple constant-higher-moments N and N2 models perform poorly, conditionally and unconditionally, with exception rates often higher than twice the nominal exception level. In sharp contrast to N and N2, the inclusion of the fourth co-moments in the specification of N24 dramatically improves the quality of the forecasts both conditionally and unconditionally. Similar conclusions can be drawn from Table 3, where we report the same statistics for the joint density of GBP, CHF and JPY exchange rates. These findings confirm the importance of (even) higher (co-)moments in the density forecasting of financial data as observed previously for unidimensional variables (Polanski and Stoja, 2009a).

[Table 2 and 3]

In addition to the joint density of returns on different assets, financial institutions may be interested in the joint density of extreme losses on the same asset over consecutive

16

business days. Table 4 presents the actual MVaR exceedance rates for different nominal levels a, together with Kupiec's (1995) unconditional (t_u) and Christoffersen's (1998) conditional (LR_c) test statistics for the joint density of returns on five consecutive business days for Dow Jones. The out-of-sample period is 09/12/1976 to 07/09/2005 (7500 daily observations were used to construct a sequence of 1500 quintuple observations). Although for high a the improvements on MVaR forecasts are not significant, for lower levels of a the dominance of the AJD over its simpler counterparts is overwhelming both conditionally and unconditionally. This illustrates that incorporating higher co-moments in the joint density of returns yields significant improvements in MVaR forecasting, not only cross-sectionally but also in an intertemporal context.

[Table 4]

5. Summary and Conclusion

We propose a simple and flexible framework for forecasting the joint density of asset returns. The multinormal distribution is augmented with a polynomial in time-varying higher co-moments, where the coefficients of the polynomial are MM estimated for a carefully selected set of co-moments. In an empirical study, we compare the proposed model with a range of other models widely used in the literature. Although a recently proposed goodness-of-fit test (Polanski & Stoja, 2009b) shows that none of the models examined provides an accurate description of the entire joint distribution of returns, the

17

AJD performs well in the negative tail of the distribution. By focusing on the negative tail, which relates to the probability of joint extreme losses, we keep with the standard risk management practice. In spite of its conceptual and computational simplicity, our framework appears to deliver a highly accurate forecast of the joint risk.

A consequence of the conceptual simplicity and flexibility is its intuitive appeal. The co-moments, used in the MM estimation, relate directly to the shape of the joint distribution. Furthermore, the structure of density forecasts is the same for an arbitrary number of assets. Regarding the computational cost, we note that the estimation of multi-dimensional functions can be limited by specifying only a few co-moments in the objective function (2) whereas each forecast is evaluated by computing a single multidimensional integral. Possible extensions of the basic model could include the asymmetric dependence of returns on joint positive and negative shocks and/or augmenting other functions than the standard normal.

References

Aas, K., Czado, C., Frigessi, A., and Bakken, H., (2009), "Pair-copula constructions of multiple dependence", *Insurance: Mathematics and Economics* 44, 182–198.

Artzner, P., Delbaen, F., Eber, J.-M., and Heath, D. (1999), "Coherent Measures of Risk", *Mathematical Finance*, 9, 203-228.

Bae, K.-H., Karolyi, G.A., and Stulz, R.M., (2003), "A New Approach to Measuring Financial Contagion". *Review of Financial Studies*, 16, 717-764.

Bali, T.G., and Weinbaum, D., (2007), "A Conditional Extreme Value Volatility Estimator Based on High Frequency Returns", *Journal of Economic Dynamics and Control,* 31, 361–397.

Bahra, B., (1997), "Implied risk!neutral probability density functions from option prices: theory and application". Working Paper No. 66, Bank of England, London.

Basle Committee on Banking Supervision, (1996), "Overview of the Amendment to the Capital Accord to Incorporate Market Risks", January.

Berkowitz, J., (2001) "Testing Density Forecasts with Applications to Risk Management". *Journal of Business and Economic Statistics,* 19, 465-474.

Breeden, D.T., and Litzenberger, R.H., (1978), "Prices of state-contingent claims implicit in options prices", *Journal of Business*, 51, 621-651.

Christoffersen, P. F., (1998), "Evaluating Interval Forecasts", *International Economic Review*, 39, 841–862.

Cotner, J.S., (1991), "Index option pricing: do investors pay for skewness?", *Journal of Futures Markets,* 11, 1-8.

Diebold, F.X., Hahn, J., and Tay, A.S., (1999), "Multivariate Density Forecast Evaluation and Calibration in Financial Risk Management: High Frequency Returns on Foreign Exchange". *Review of Economics and Statistics*, 81, 661-673.

Embrechts, P., McNeil, A., and Straumann, D., (2002), "Correlation and dependence properties in risk management: properties and pitfalls", in M. Dempster, (ed.), Risk Management: Value at Risk and Beyond, Cambridge University Press.

Engle, R.F., (1982), "Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of U.K. Inflation," *Econometrica*, 50, 987-1008.

Fackler, P.L., and King, R.P., (1990), "Calibration of options-based probability assessments in agricultural commodity markets", *American Journal of Agricultural Economics,* 72, 73-83.

Gallant, R., Hsieh, D.A., and Tauchen, G., (1991), "On fitting a recalcitrant series: the pound/dollar exchange rate 1974-83". In *Nonparametric and Semiparametric Methods in Econometrics and Statistics: Proceedings of the Fifth International Symposium in Economic Theory and Econometrics*, Barnett, W.A., Powell, J., Tauchen, G. (eds). Cambridge University Press: Cambridge.

Guidolin, M., and Timmermann, A., (2006), "An econometric model of nonlinear dynamics in the joint distribution of stock and bond returns", *Journal of Applied Econometrics,* 21: 1–22.

Hansen, B.E., (1994), "Autoregressive conditional density estimation", *International Economic Review,* 35, 422-438.

Kupiec, P. H., (1995), "Techniques for verifying the accuracy of risk measurement models", *Journal of Derivatives* 3, 73–84.

Jackwerth, J.C., Rubinstein, M., (1996), "Recovering probability distributions from option prices", *Journal of Finance,* 7, 1611-1631.

McDonald, J., Xu, Y., (1995), "A Generalisation of the Beta Distribution with Applications", *Journal of Econometrics*, 66, 133-152.

Noceti, P., J. Smith and Hodges, S., (2003), "An Evaluation of Tests of Distributional Forecasts", *Journal of Forecasting,* 22, 447–455.

Patton, A.J., (2004), "On the out-of-sample importance of skewness and asymmetric dependence for asset allocation", *Journal of Financial Econometrics*, 2, 130-168.

Patton, A.J., (2007), "Copula-based models for financial time series", in *Handbook of Financial Time Series*, T. G. Andersen, R. A. Davis, J.-P. Kreiss and T. Mikosch (eds.), Springer Verlag.

Polanski, A., and Stoja, E., (2009a), "Incorporating Higher Moments into Value at Risk Forecasting", *Journal of Forecasting,* forthcoming.

Polanski, A., and Stoja, E., (2009b), "An Efficient Evaluation Method of Multidimensional Time-Varying Density Forecasts", under review at *Journal of Business and Economic Statistics.*

Rosenblatt, M., (1952), "Remarks on a multivariate transformation". *Annals of Mathematical Statistics*, 23, 470–472.

Simkowitz, M.A., and Beedles, W.L., (1978), "Diversification in a three-moment world", *Journal of Finance,* 22, 177- 181.

Singleton, J.F., and Wingender, J., (1986) "Skewness persistence in common stock returns", *Journal of Financial and Quantitative Analysis,* 10, 224-230.

Tay, A.S., and Wallis, K.F., (2000), "Density Forecasting: A Survey", *Journal of Forecasting*, 19, 235-254.

Theodossiou, P., (1998), "Financial Data and the Skewed Generalized t Distribution", *Management Science* 44, 1650–1661.

Table 1 Summary Statistics

Panel A: Equity Indices

Returns

	S&P500	**Dow Jones**	**Nasdaq**
Mean (%)	0.008	0.014	0.012
Stand Dev (%)	0.011	0.010	0.018
Skewness	0.051	-0.064	0.116
Kurtosis	4.984	6.004	6.614
B-J	413.509	946.749	1372.523
ARCH(4)	658.680	635.390	721.279

Correlations

	S&P500	**Dow Jones**	**Nasdaq**
S&P500	1.000	0.945	0.855
Dow Jones		1.000	0.718
Nasdaq			1.000

Notes: The table reports the mean, standard deviation, skewness, kurtosis, Bera-Jarque statistic, ARCH(4) statistic and correlation for log returns for S&P500, Dow Jones and Nasdaq for the sample period 15/09/2000 to 29/08/08 (2000 daily observations). The five percent critical values of the B-J and ARCH(4) statistics are 5.99 and 9.49, respectively.

Table 1 Summary Statistics (continued)

Panel B: Foreign Exchange

Returns

	GBP	CHF	JPY
Mean (%)	0.001	-0.003	-0.007
Stand Dev (%)	0.589	0.715	0.689
Skewness	-0.180	-0.619	-0.135
Kurtosis	6.111	8.443	4.616
B-J	662.268	763.295	839.624
ARCH(4)	727.058	759.304	854.435

Correlations

	GBP	CHF	JPY
GBP	1.000	0.010	0.006
CHF		1.000	-0.010
JPY			1.000

Notes: The table reports the mean, standard deviation, skewness, kurtosis, Bera-Jarque statistic, ARCH(4) statistic and correlation for log returns for GBP, CHF and JPY against USD for the sample period 6/12/1988 to 6/08/1996 (2000 daily observations). The five percent critical values of the B-J and ARCH(4) statistics are 5.99 and 9.49, respectively.

Table 2: MVaR Forecast Accuracy Results for Equity Indices

Models	$a = 0.5\%$			$a = 1\%$			$a = 1.5\%$		
	%x	t_u	LR_c	%x	t_u	LR_c	%x	t_u	LR_c
N	0.014	3.554	0.851	0.023	3.878	0.702	0.026	3.091	0.286
N2	0.010	2.247	0.402	0.018	2.821	1.754	0.027	3.310	1.336
N24	0.006	0.579	0.143	0.013	1.184	0.683	0.018	1.009	1.318

Models	$a = 2\%$			$a = 2.5\%$			$a = 5\%$		
N	0.031	2.838	0.003	0.036	2.537	0.092	0.059	1.708	2.289
N2	0.032	2.944	0.482	0.034	2.220	1.096	0.064	2.558	0.427
N24	0.020	-0.161	0.071	0.025	-0.144	0.469	0.05	0.000	0.795

Notes: The table details the results for 2000 daily observations (15/09/2000 to 29/08/08) on triplets S&P500, Dow Jones and Nasdaq. It reports the percentage of z_t^M-scores that are less than the nominal level a (i.e., observations that exceed the a-MVaR), the Kupiec's t-statistic to test the unconditional accuracy and the Christoffersen's likelihood ratio statistic to test conditional accuracy for different nominal levels.
Model N is the normal distribution with zero mean and the variance-covariance matrix estimated for each t in the window [t-T, t] for T=498 and t=1,…,2000. Model N2 is normal distribution with zero mean and the variance-covariance matrix estimated according to EMWA (5). Model N24 is normal distribution with zero mean and the variance-covariance matrix and all 4th moments estimated according to EMWA (5).

Table 3: M VaR Forecast Accuracy Results for Exchange Rates

Models	$a = 0.5\%$			$a = 1\%$			$a = 1.5\%$		
	%x	t_u	LR_c	%x	t_u	LR_c	%x	t_u	LR_c
N	0.004	-0.874	0.175	0.006	-2.969	0.436	0.010	-3.377	1.049
N2	0.004	-0.002	0.249	0.008	-1.414	0.676	0.012	-1.525	0.050
N24	0.005	0.744	0.336	0.010	0.000	0.396	0.014	-0.854	0.006
	$a = 2\%$			$a = 2.5\%$			$a = 5\%$		
N	0.013	-4.058	0.011	0.016	-4.511	1.408	0.035	-5.404	0.056
N2	0.016	-2.132	0.083	0.020	-1.977	0.270	0.023	-3.092	0.537
N24	0.018	-1.406	0.204	0.022	-1.759	0.210	0.041	-3.288	0.241

Notes: The table details the results for 2000 daily observations (6/12/1988 to 6/08/1996) on triplets GBP, CHF and JPY against USD. It reports the percentage of z_t^M-scores that are less than the nominal level a (i.e., observations that exceed the a-MVaR), the Kupiec's t-statistic to test the unconditional accuracy and the Christoffersen's likelihood ratio statistic to test the conditional accuracy for different nominal levels.

Model N is the normal distribution with zero mean and the variance-covariance matrix estimated for each t in the window [t-T, t] for T=498 and t=1,…,2000. Model N2 is normal distribution with zero mean and the variance-covariance matrix estimated according to EMWA (5). Model N24 is normal distribution with zero mean and the variance-covariance matrix and all 4th moments estimated according to EMWA (5).

Table 4: Intertemporal M VaR Forecast Accuracy Results for Dow Jones

Models	$a = 0.5\%$ %x	t_u	LR_c	$a = 1\%$ %x	t_u	LR_c	$a = 1.5\%$ %x	t_u	LR_c
N	0.002	-2.600	0.011	0.007	-1.586	0.132	0.008	-3.043	0.192
N2	0.003	-1.752	0.020	0.008	-0.869	0.192	0.011	-1.341	0.388
N24	0.003	-1.119	0.032	0.009	-0.557	0.225	0.012	-1.341	0.388

Models	$a = 2\%$ %x	t_u	LR_c	$a = 2.5\%$ %x	t_u	LR_c	$a = 5\%$ %x	t_u	LR_c
N	0.011	-3.171	0.388	0.018	-1.812	1.064	0.051	0.118	0.006
N2	0.013	-2.251	0.539	0.019	-1.811	1.064	0.043	-1.268	0.012
N24	0.015	-1.718	0.653	0.019	-1.811	1.064	0.048	-0.362	0.089

Notes: The table details the results for 1500 quintuple observations on Dow Jones returns (i.e., each quintuple contains Dow Jones returns on five consecutive working days). It covers the period of 7500 days (09/12/1976 to 07/09/2005). The table reports the percentage of z_t^M-scores that are less than the nominal level a (i.e., observations that exceed the a-MVaR), the Kupiec's t-statistic to test the null hypothesis of the unconditional accuracy and the Christoffersen's likelihood ratio statistic to test the null hypothesis of conditional accuracy for different nominal levels.
Model N is the normal distribution with zero mean and the variance-covariance matrix estimated for each t in the window [*t-T, t*] for T=300 and t=1,…,1500. Model N2 is normal distribution with zero mean and the variance-covariance matrix estimated according to EMWA (5). Model N24 is normal distribution with zero mean and the variance-covariance matrix and all 4th moments estimated according to EMWA (5).