Review of Applied Economics, Vol. 1, No. 2, (2005) : 143-169

Nonlinear Properties of Multifactor Financial Models

***Kim Heng Chen*[a,*], *Venkata K. Jandhyala*[b], *& Stergios B. Fotopoulos*[c]**

This paper provides a comprehensive analysis of the nonlinear properties of multifactor pricing models. Beginning with the generalized geometric Brownian motion, we develop a method whereby the log returns of a set of d*-assets or portfolios admit a scale mixture model. This is followed by an analytical study on the conditional behavior of a subset of assets given another subset. Expressions for the first two conditional moments are provided under the scale mixture family. The regression equation when the scaling variable is constant (unity) corresponds with the renowned APT. Computable conditional moment expressions for the scaling variable are derived under both inverse gamma and gamma distributions. These moment equations are nonlinear in parameters, apart from containing the usual linear terms under the APT. We then apply the above nonlinear methodology to the log asset returns of four major companies in the U.S. stock market.*

JEL classification: *C31, C44, G12*

Keywords: *Generalized Geometric Brownian Motion; Heteroskedasticity; Scale Mixture of Normal Distributions.*

INTRODUCTION

One of the major developments in modern financial economics is the Capital Asset Pricing Model (CAPM), which enabled economists to quantify the tradeoff between risk and expected return associated with holding a particular asset. Based on the foundation laid by Markowitz (1959), Sharpe (1964) and Lintner (1965) develop economy-wide implications of the CAPM. They postulate that investors of homogeneous expectations will hold mean-variance efficient portfolios and in the absence of market frictions, the market portfolio of its own accord is a mean-variance efficient portfolio. However, empirical evidence has indicated that the CAPM beta does not completely explain the cross section of expected asset returns: an interesting review of such literature can be found in Fama (1991).

---

[a] Assistant Professor, School of Business and Management, American University of Sharjah, P.O. Box 26666, Sharjah, United Arab Emirates.

[*] Corresponding Author. Tel: (971) - 6-5152365; fax: (971) - 6-5585065; e-mail: *kachen@aus.edu*

[b] Professor, Department of Mathematics and Department of Statistics, Washington State University, Pullman, WA99164-3113.

[c] Professor, Department of Management & Operations and Department of Statistics, Washington State University, Pullman, WA99164-4736.

As such, the Arbitrage Pricing Theory (APT) was introduced by Ross (1976) as an alternative to the CAPM and the former has since developed into a powerful tool for modeling the return-generating process of a population of assets or portfolios. Numerous studies on APT have emerged: Roll and Ross (1980), Huberman (1982), Jobson (1982), Chamberlain (1983), Chamberlain and Rothschild (1983), Dhrymes et al. (1984), Ingersoll (1984), Connor and Korajczyk (1988), Milne (1988), Bansal and Viswanathan (1993), Ferson and Korajczyk (1995) *inter alia*. The APT begins with the assumption that the return-generating process can be governed by a linear relationship with a number of systematic factors. Under this and other assumptions, it establishes that the excess return for any arbitrary asset or portfolio may be captured by a linear relationship with the excess returns of a few portfolios with systematic risks. For empirical verification (estimation and testing) purposes, the APT is generally treated as a multifactor model, relating conditional expected returns of assets with diversified portfolios.

Jobson (1982) presents this conditional relationship between expected asset returns and expected portfolio returns in a multivariate regression setup. Relying on the homoskedasticity and normality assumptions, he develops the likelihood ratio test for APT. Financial analysts rather routinely rely upon the normality assumption in this framework for parameter estimation and the associated finite sample inferential properties. Otherwise, inferences are usually restricted to an asymptotic setting, where again, the normal distribution of the underlying parameter estimates is utilized via the central limit theory. However, such dependence upon normality-based theory may not be practical when high-frequency financial data is used in the data analyses. For example, the normality assumption appears to be adequate for characterizing aggregated returns such as monthly or yearly asset returns. But it does not work well with high-frequency financial data such as intraday asset returns. Empirical evidence has shown that these high-frequency financial data exhibits a leptokurtic distribution, that is, the distribution of the high-frequency financial data has “heavier” tails and “sharper” peak relative to the normal distribution, see e.g., Bollerslev (1986), Bollerslev et al. (1992), Feinstone (1987), Gerhard and Hautsch (2002). The violation of the normality assumption will result in biased or even erroneous parameter estimation, which in turn, produces inferences that are misleading at best.

In this paper, we take a rather different approach in the analysis of asset returns from that of Roll and Ross (1980) and Jobson (1982). Beginning with the generalized geometric Brownian motion, we develop a model for the cross-sectional log asset returns (natural logarithm of asset returns) where the volatility matrix is assumed to be stochastic. The stochastic model admits a general form of scale mixtures of the multivariate Gaussian family where the latter is a broad family of distributions that encompasses an array of heavy-tailed distributions and may be closely related to the stable family. This enables us to capture sharp departures from normality in log asset returns. With the derived scale mixture model for log asset returns as the basis, we investigate the conditional behavior of a set of log asset returns given another set. Thus, the basic goal overlaps with that of the APT.

The remainder of this paper is organized as follows. The next section reviews the econometric approach to APT by Roll and Ross (1980) and Jobson (1982). The section after that details the development and properties of the generalized multifactor financial model. This is followed by the section on data analysis, where we perform a detailed empirical analysis on the log returns data from four major U.S. companies, namely, Cisco Systems, Inc. (CISCO), Coca-Cola

Company (COKE), Dell Computer Corporation (DELL), and the Microsoft Corporation (MFST), conditional upon the log returns of the S&P500 index. The final section summarizes, concludes, and sketches directions for future research.

REVIEW OF APT

Roll and Ross (1980) derive the APT by postulating that the returns of a set of *d* assets are governed by a *K*-factor model of the form:

$$R_i = \mu_i + \sum_{j=1}^{K} b_{ij}\delta_j + \varepsilon_i, \quad i = 1,\ldots,d, \qquad ...(1)$$

where R_i and μ_i are the random return and the expected return associated with the i^{th} asset, δ_j is the j^{th} systematic factor risk assumed to be random with mean zero and is common to all assets, b_{ij} quantifies the sensitivity of the i^{th} asset to the j^{th} systematic factor, and finally, ε_i is the unsystematic risk component or the noise term. They argue that under equilibrium, there exist weights $\lambda_1,\ldots,\lambda_K$ $(K << d)$ such that the expected asset returns will satisfy

$$\mu_i - \mu_0 = \sum_{j=1}^{K} \lambda_j b_{ij}, \quad i = 1,\ldots,d, \qquad ...(2)$$

where μ_0 is the rate of return of a riskfree asset. Upon identifying the factor risk premia λ_j, they arrive at the APT of the following form:

$$\mu_i - \mu_0 = \sum_{j=1}^{K} \left(\mu^j - \mu_0\right) b_{ij}, \quad i = 1,\ldots,d, \qquad ...(3)$$

where $\mu^j - \mu_0$ denotes the excess returns on portfolios with only j^{th} systematic factor risk, $j = 1,\ldots,K$. Thus, the central conclusion of the APT is that the mean premium returns lie in a *K*-dimensional subspace spanned by the factor loadings.

Inspired by the work of Roll and Ross (1980), Jobson (1982) formulated the APT into a multivariate regression framework. Specifically, he begins with the time-dependent version of Roll and Ross (1980) *K*-factor model by rewriting eq. (1) as

$$\mathbf{R}_t = \boldsymbol{\mu} + \mathbf{B}\boldsymbol{\delta}_t + \boldsymbol{\varepsilon}_t, \quad t = 1,\ldots,n, \qquad ...(4)$$

where $\boldsymbol{\mu}$ is the $d \times 1$ mean premium return vector, $\mathbf{B}$ is a $d \times K$ matrix of factor loadings, $\boldsymbol{\delta}_t$ is a $K \times 1$ vector of systematic factors, and $\boldsymbol{\varepsilon}_t$ is an $d \times 1$ vector of error terms with $\boldsymbol{\delta}_t$ being independent of $\boldsymbol{\varepsilon}_t$. Assuming that there exists a subset of *K* independent assets or portfolios with return vector $\mathbf{R}_{2t}$, he partitions the model in eq. (4) into the following two equations:

$$\mathbf{R}_{1t} = \boldsymbol{\mu}_1 + \mathbf{B}_1\boldsymbol{\delta}_t + \boldsymbol{\varepsilon}_{1t}, \quad ...(5)$$

$$\mathbf{R}_{2t} = \boldsymbol{\mu}_2 + \mathbf{B}_2\boldsymbol{\delta}_t + \boldsymbol{\varepsilon}_{2t}, \quad ...(6)$$

where the partitions of $\boldsymbol{\mu}$, $\mathbf{B}$, and $\boldsymbol{\varepsilon}_t$ conform to the partitioning of $\mathbf{R}_t$ into $\mathbf{R}_{1t}$ and $\mathbf{R}_{2t}$. He then derives the APT, namely, the conditional model for $\mathbf{R}_{1t}$ given $\mathbf{R}_{2t}$ as

$$\mathbf{R}_{1t} = \boldsymbol{\mu}^* + \mathbf{B}^*\mathbf{R}_{2t} + \boldsymbol{\eta}_t, \quad ...(7)$$

where $\boldsymbol{\mu}^*$ is $(d-K)\times 1$, $\mathbf{B}^*$ is $(d-K)\times K$, and $\boldsymbol{\eta}_t$ is $(d-K)\times 1$. Under the assumption that the premium return vector $\mathbf{R}_t$ follows a multivariate normal distribution with mean $\boldsymbol{\mu}$ and covariance $\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}$, he then continues to investigate the inferential properties for the model in eq. (7). In short, Jobson (1982) states that, typically, in regression analysis, one assumes normality to hold or asymptotic theory is employed to justify this unequivocal commitment to normality as well as the linearity associated with the APT in eq. (7).

ASSET PRICES AND RETURNS MODELING

As mentioned in Section I, there is growing evidence in the finance literature that asset returns exhibit strong stochasticity in volatility, especially for high-frequency data. To better model the behavior of the asset price dynamics, researchers have developed various methodologies: Hull and White (1987) and Melino and Turnbull (1990) on processes with diffusion stochastic volatility, Rachev and SenGupta (1993) and Cambanis et al. (2000) on mixtures of normal and other distributions, McCulloch (1996) and Willinger et al. (1999) on Stable Paretian models, Madan and Milne (1991) and Barndoff-Nielson (1998) on subordinated Levy processes in finance.

This leads us to question whether the linearity conditions, commonly imposed on the conditional asset returns in financial modeling continue to hold under strong departures from the deterministic nature of volatility. As shown below, when stochastic volatility is incorporated into financial returns, it will have important implications to various salient features such as linearity, homoskedasticity, as well as the normality of the usual APT-type financial models.

In developing the methodology, we adapt the conditional and scale mixture of normal distribution approaches of Jobson (1982) and Cambanis et al. (2000), respectively. In particular, we derive a cross-sectional model for the log returns of a set of assets or portfolios from the classical generalized Brownian formulation for asset prices. We incorporate stochasticity into the volatility matrix by formulating it as a scale mixture of normal distribution. This establishes a general structural form for the log asset returns, which subsequently plays a fundamental role in understanding the behavior of conditional asset returns.

Generalized Geometric Brownian Motion and Model for Log Asset Returns

As presented in Karatzas and Shreve (1998), we consider a financial market $\mathbb{M}$ comprising several (d) assets, $d \in \mathbb{N}$. Let $\{S_i(t), i = 1,\ldots,d, t \in [0,T)\}$ denote the prices of these financial instruments at time t with $\{S_i(0), i = 1,\ldots,d\}$ being positive constants. Then, the prices evolve according to the generalized geometric Brownian motion:

$$dS_i(t) = S_i(t)\left\{\mu_i(t)dt + \left\langle \boldsymbol{\sigma}_i^{1/2}(t), d\mathbf{W}(t)\right\rangle\right\}, \quad i = 1,\ldots,d, \qquad \ldots(8)$$

where $\langle .,. \rangle$ denote the inner product, $\mathbf{W} = \left(W^{(1)},\ldots,W^{(d)}\right)'$ denotes a standard Brownian motion in $\mathbb{R}^d$ defined on a complete probability space $(\Omega, \mathcal{F}, \mathbb{P}), \{\mathcal{F}_t, t \in [0,T]\}$ denotes the $\mathbb{P}$-augmentation of the filtration $\mathcal{F}_t^W = \sigma(\mathbf{W}(s) : 0 \le s \le t)$ (s-algebra generated by $\mathbf{W}$), $\mu_i(t)$ and $\boldsymbol{\sigma}_i^{1/2}(t)$ denotes the appreciation rate and volatility of i^{th} asset at time t, respectively. Note that both $\mu_i(t)$ and $\boldsymbol{\sigma}_i^{1/2}(t)$, $i = 1,\ldots,d$, are assumed to be progressively measurable with respect to $\{\mathcal{F}_t\}$ (i.e., independent of the standard Brownian motion $\mathbf{W}$) and are bounded uniformly in $(t,w) \in [0,T] \times \Omega$.

The process of interest in this study is the log asset price process, i.e., $Y_i = \log S_i(t)$, $i = 1,\ldots,d$. Since $S_i(t)$ is an Ito process, Y_i is also an Ito process and is given by

$$dY_i(t) = \left\{\mu_i(t) - \frac{1}{2}\boldsymbol{\sigma}_i^{1/2}(t)\mathbf{1}\right\}dt + \left\langle \boldsymbol{\sigma}_i^{1/2}(t), d\mathbf{W}(t)\right\rangle, \quad i = 1,\ldots,d, \qquad \ldots(9)$$

or equivalently in vector form

$$d\mathbf{Y}(t) = \left\{\boldsymbol{\mu}(t) - \frac{1}{2}\sigma(t)\mathbf{1}\right\}dt + \left\langle \sigma^{1/2}(t), d\mathbf{W}(t)\right\rangle, \qquad \ldots(10)$$

where $\mathbf{1} = (1,\ldots,1)'$. By imposing the strong non-degeneracy condition $\mathbf{y}'\sigma(t)\mathbf{y} \ge \varepsilon\|\mathbf{y}\|^2, \forall (t,\mathbf{y}) \in [0,T] \times \mathbb{R}^d$ *a.s.* on the volatility matrix σ where $\varepsilon > 0$, and assuming $\int_0^t \left\{ \left\| \boldsymbol{\mu}(u) \right\|^2 + \left\| \sigma(u) \right\| \right\} du < \infty$ *a.s.*, together with the Lipschitz condition

$$\left\| \boldsymbol{\mu}(u) - \boldsymbol{\mu}(s) \right\|^2 + \left\| \sigma(u) - \sigma(s) \right\| \le c|u-s|, \quad u,s > 0, \qquad \text{...(11)}$$

the following may be shown to hold as a solution of the differential equation in eq. (10):

$$\mathbf{Y}(t) = \mathbf{Y}_0 + \int_0^t \left\{ \boldsymbol{\mu}(u) - \frac{1}{2}\sigma(u)\mathbf{1} \right\} du + \int_0^t \left\langle \sigma^{1/2}(u), d\mathbf{W}(t) \right\rangle, \quad t \in [0,T]. \qquad \text{...(12)}$$

Next, we partition the time horizon $[0,T]$ into $0 = t_{0k} < t_{1k} < \ldots < t_{kk} \le T$ such that $\lim_{k\to\infty} \sup_j \left(t_{jk} - t_{j-1,k} \right) = 0$. In practice, the length of time between two successive partitioned periods may represent years, months, or even weeks. The latter can be further sub-partitioned into *n* intervals with length *d*, i.e., $\left[t_{jk} + (i-1)\delta, t_{jk} + i\delta \right), i = 1,\ldots,n;\ j = 1,\ldots,k,\ k \ge 1$. The log return of the asset for the i^{th} interval at j^{th} period can then be expressed as

$$\begin{aligned}\mathbf{R}\left(t_{jk},\delta,i\right) &= \mathbf{Y}\left(t_{jk} + i\delta\right) - \mathbf{Y}\left(t_{jk} + (i-1)\delta\right) \\ &\doteq \left(\boldsymbol{\mu}\left(t_{jk} + (i-1)\delta\right) - \frac{1}{2}\sigma\left(t_{jk} + (i-1)\delta\right)\mathbf{1} \right)\delta \\ &\quad + \left\langle \sigma^{1/2}\left(t_{jk} + (i-1)\delta\right), \mathbf{W}\left(t_{jk} + i\delta\right) - \mathbf{W}\left(t_{jk} + (i-1)\delta\right) \right\rangle. \end{aligned} \qquad \text{...(13)}$$

Applying the Lipschitz condition in eq. (11) upon the appreciation rate vector and the volatility matrix, for sufficiently large *k*, the above return vector may then be approximated as

$$\mathbf{R}\left(t_{jk},\delta,i\right) \cong \left(\boldsymbol{\mu}\left(t_{jk}\right) - \frac{1}{2}\sigma\left(t_{jk}\right)\mathbf{1} \right)\delta + \left\langle \sigma^{1/2}\left(t_{jk}\right), \delta^{1/2}\mathbf{Z} \right\rangle = \mathbf{R}\left(t_{jk},\delta\right) \ \textit{a.s.} \qquad \text{...(14)}$$

Eq. (14) suggests that the returns (or log returns) within any *d* interval are independent of the position at which the interval is considered. In the modeling literature, researchers usually let *d* represents the length of a day, hour, minute, or even second. Rydberg (1999) and Rydberg and Shephard (2001) consider an intraday analysis of the U.S. stock prices by letting *d* be in seconds. Note that the stochastic difference equation in eq. (14) can be interpreted as the Euler approximation of the log asset returns. The discretized log asset price process then converges almost surely to its continuous counterpart in eq. (10) as $\delta \to 0$ and $k \to \infty$. Refer to Kloeden and Platten (1995) for details.

When the appreciation rate vector $\boldsymbol{\mu}$ and the volatility matrix σ are deterministic, the process $\left\{ S_i(t), i = 1,\ldots,d, t \in [0,T) \right\}$ follows the classical lognormal model. Though elegant from an analytic point of view, empirical evidence does not support this situation: high-frequency data indicates that the distributions of asset returns are by and large leptokurtic, i.e., the distributions of asset returns have relatively heavier tails than that of the Gaussian process. This suggests that either the appreciation vector or the volatility matrix within the framework of the geometric Brownian motion is non-deterministic. The *modus operandi* is to propose stochastic versions of these parameter processes within the generalized geometric Brownian motion framework. Retaining the deterministic nature of the appreciation vector $\boldsymbol{\mu}(.)$, we incorporate stochasticity into the volatility matrix by letting $\sigma(.) = A\Sigma(.)$ *a.s.*, where A is a strictly positive random variable and $\Sigma(.)$ is a deterministic variance-covariance matrix.

The focus of this study is on the returns from a fixed partitioned interval. Thus, we can drop the time component t_{jk} from eq. (14) and without loss of generality, we let the sub-partitioned size parameter d be 1. As a result, eq. (14) can be written as

$$\mathbf{R} = \boldsymbol{\mu} - \frac{A}{2}\Sigma\mathbf{1} + A^{1/2}\mathbf{G}, \qquad \text{...(15)}$$

where $\mathbf{G}$ follows d-multivariate normal distribution with variance-covariance matrix Σ. Note that the scale random variable A is independent of $\mathbf{G}$. The above model can be expressed as a special case of the following more general form:

$$\mathbf{R} = \boldsymbol{\mu} + A\mathbf{m} + A^{1/2}\mathbf{G}, \qquad \text{...(16)}$$

where $\mathbf{m}$ is a d-dimensional parameter vector. A special case of the model in eq. (16) is considered in the monograph by Kotz et al. (2001), where they study some distributional behavior of **R** when the scaling variable A follows the exponential distribution with mean one.

In the studies of stock market returns, scale mixture type volatility processes as mentioned above have been considered by several authors to capture the excess kurtosis. For instance, Praetz (1972), Blattberg and Gonedes (1974), Madan and Seneta (1990), Kuchler et al. (1995), Hurst and Platen (1997), and Barndorff-Nielsen (1997, 1998) *inter alia* have proposed gamma, inverse gamma, lognormal, and inverse Gaussian distributions for the mixing variable while modeling the volatility process of a single asset by scale mixtures.

Conditional Properties of Log Asset Returns

Conditional distributions of log asset returns and their conditional moments play important roles in financial modeling and forecasting of returns. The CAPM and APT amply illustrate the importance of such conditional modeling. While these are well studied for the case where the volatility process is in equilibrium, such models and their properties are not fully explored and well established for the case of stochastic volatility, such as those represented by scale mixtures as shown in eq. (16). As such, recent works by Cambanis and Fotopoulos (1995), Fotopoulos

(1998), Fotopoulos and He (1999, 2001), and Cambanis et al, (2000) are fundamental. Several of their results are applied directly in the study of conditional behavior of log asset returns.

We begin by partitioning $\mathbf{R}$, $\boldsymbol{\mu}$, $\mathbf{m}$, $\mathbf{G}$, and Σ into $\mathbf{R}' = (\mathbf{R}_1', \mathbf{R}_2')$, $\boldsymbol{\mu}' = (\boldsymbol{\mu}_1', \boldsymbol{\mu}_2')$, $\mathbf{m}' = (\mathbf{m}_1', \mathbf{m}_2')$, $\mathbf{G}' = (\mathbf{G}_1', \mathbf{G}_2')$, and $\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}$ where $\mathbf{R}_2'$, $\boldsymbol{\mu}_2'$, $\mathbf{m}_2'$, and $\mathbf{G}_2'$ are K-dimensional and Σ_{22} is a $K \times K$ variance-covariance matrix, $K < d$, etc. Utilizing the methodology in Fang and Zhang (1990), i.e., by defining $\Sigma = B'B$ and $\Sigma_{22} = B_2'B_2$, where B_2 is a positive definite matrix, it follows that

$$\Sigma = B'B = \begin{bmatrix} I & \Sigma_{12}\Sigma_{22}^{-1} \\ 0 & I \end{bmatrix} \begin{bmatrix} \Sigma_{11.2} & 0 \\ 0 & \Sigma_{22} \end{bmatrix} \begin{bmatrix} I & 0 \\ \Sigma_{22}^{-1}\Sigma_{21} & I \end{bmatrix} = F'F, \qquad \text{...(17)}$$

where $F = \begin{bmatrix} B_{11.2} & 0 \\ 0 & B_2 \end{bmatrix} \begin{bmatrix} I & 0 \\ \Sigma_{22}^{-1}\Sigma_{21} & I \end{bmatrix} = \begin{bmatrix} B_{11.2} & 0 \\ B_2\Sigma_{22}^{-1}\Sigma_{21} & B_2 \end{bmatrix}$. Note that $\Sigma_{11.2} = B_{11.2}'B_{11.2}$ and with these results, eq. (16) may be equivalently expressed as

$$\mathbf{R} = \boldsymbol{\mu} + A\mathbf{m} + A^{1/2}B'\mathbf{Z} \overset{d}{=} \boldsymbol{\mu} + A\mathbf{m} + A^{1/2}F'\mathbf{Z},$$

which can further be expressed as

$$\mathbf{R} = \begin{pmatrix} \mathbf{R}_1 \\ \mathbf{R}_2 \end{pmatrix} = \begin{pmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{pmatrix} + A\begin{pmatrix} \mathbf{m}_1 \\ \mathbf{m}_2 \end{pmatrix} + A^{1/2}\begin{pmatrix} B_{11.2}'\mathbf{Z}_1 + \Sigma_{12}\Sigma_{22}^{-1}B_2'\mathbf{Z}_2 \\ B_2'\mathbf{Z}_2 \end{pmatrix}, \qquad \text{...(18)}$$

where $\mathbf{Z}$ is a d-dimensional standard normal vector.

The above representation of the log asset returns turns out to be well suited for in-depth study of the conditional properties of $(\mathbf{R}_1 \mid \mathbf{R}_2)$. This class of scale mixture type random vectors, which encompasses the Gaussian systems as distinguished members, exhibits conditional behavior that is distinct from the normal theory. For example, regression involving jointly Gaussian random variables is always linear. However, this is not the case for the current model. Moreover, the conditional covariance of jointly normal random variables leads to degenerate, non-random quantities. Again, this is not the case for the model in eq. (16). Cambanis et al. (1981), Hardin (1982), Cambanis and Wu (1992), Samorodnitsky and Taqqu (1994), *inter alia*, consider situations in which the conditional regression is nonlinear and the conditional variance-covariance matrix is non-degenerate.

The theorem below provides expressions for the regression and skedastic (conditional variance of $\mathbf{R}_1|\mathbf{R}_2$ *a.s.*) functions of $\mathbf{R}_2$ along with necessary and sufficient conditions for their existence. The proof can be obtained by simple conditioning arguments on eq. (16) and is thus omitted.

THEOREM 1. I. *The regression equation of* $\mathbf{R}_2$ *(conditional mean of* $\mathbf{R}_1$ *given* $\mathbf{R}_2$*) is given as*

$$E\left[\mathbf{R}_1 \middle| \mathbf{R}_2\right] = \boldsymbol{\mu}_1 + \mathbf{m}_{11.2} E\left[A \middle| \mathbf{R}_2\right] + \Sigma_{12}\Sigma_{22}^{-1}\left(\mathbf{R}_2 - \boldsymbol{\mu}_2\right), \ a.s.$$

if and only if $E\left[A \middle| \mathbf{R}_2\right] < \infty$ *a.s., where* $\mathbf{m}_{11.2} = \mathbf{m}_1 - \Sigma_{12}\Sigma_{22}^{-1}\mathbf{m}_2$.

II. *The skedastic equation of* $\mathbf{R}_2$ *(conditional variance of* $\mathbf{R}_1$ *given* $\mathbf{R}_2$*) is given as*

$$V\left[\mathbf{R}_1 \middle| \mathbf{R}_2\right] = \Sigma_{11.2} E\left[A \middle| \mathbf{R}_2\right] + \mathbf{m}_{11.2}\mathbf{m}_{11.2}' V\left[A \middle| \mathbf{R}_2\right], \ a.s.$$

if and only if $E\left[A^2 \middle| \mathbf{R}_2\right] < \infty$ *a.s., where* $\Sigma_{11.2} = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}$.

By setting $\mathbf{m}_1 = -\frac{1}{2}\left(\Sigma_{11}\mathbf{1}_1 + \Sigma_{12}\mathbf{1}_2\right)$ and $\mathbf{m}_2 = -\frac{1}{2}\left(\Sigma_{21}\mathbf{1}_1 + \Sigma_{22}\mathbf{1}_2\right)$ leads one to the model in eq. (15). Consequently, the corresponding regression and skedastic functions for this special case can be expressed as

$$E\left[\mathbf{R}_1 \middle| \mathbf{R}_2\right] = \boldsymbol{\mu}_1 - \frac{1}{2}\Sigma_{11.2}\mathbf{1}_1 E\left[A \middle| \mathbf{R}_2\right] + \Sigma_{12}\Sigma_{22}^{-1}\left(\mathbf{R}_2 - \boldsymbol{\mu}_2\right), \ a.s. \qquad ...(19)$$

$$V\left[\mathbf{R}_1 \middle| \mathbf{R}_2\right] = \Sigma_{11.2} E\left[A \middle| \mathbf{R}_2\right] + \frac{1}{4}\Sigma_{11.2}\mathbf{1}_1\mathbf{1}_1'\Sigma_{11.2} V\left[A \middle| \mathbf{R}_2\right] \ a.s. \qquad ...(20)$$

Note that the regression equation in eq. (19) contains the conditional expectation of the mixing variable, $E\left[A \middle| \mathbf{R}_2\right]$, in addition to the usual linear term in $\mathbf{R}_2$. We will show in the next subsection that for some specific distributions of A, $E\left[A \middle| \mathbf{R}_2\right]$ is nonlinear in both $\mathbf{R}_2$ and the parameters, thus making the regression intrinsically nonlinear. Similarly, the expression for the skedastic function in eq. (20) also contains $E\left[A \middle| \mathbf{R}_2\right]$ besides the conditional variance term $V\left[A \middle| \mathbf{R}_2\right]$.

When the scaling variable A is constant, i.e., $A = 1$, the regression equation in eq. (19) coincides with the expectation part of the APT as in eq. (7). Furthermore, the conditional variance-covariance matrix is degenerate.

When the scaling variable A is random and $\mathbf{m} = \mathbf{0}$, it follows from Theorem 1 that the regression equation continues to be linear. However, the corresponding variance-covariance matrix is non-degenerate, thus, leading to heteroskedastic situations. Cambanis et al. (2000) characterize various nonlinear properties of $E\left[A\middle|\mathbf{R}_2\right]$ when $\mathbf{m}=\mathbf{0}$ and $\boldsymbol{\mu}=\mathbf{0}$. These properties enabled them to explain heteroskedasticity in the presence of the mixing random variable *A*. In particular, they consider the case where $A \sim S_{\alpha/2}\left(\cos\left(\frac{\pi\alpha}{4}\right),1,0\right)$, $0<\alpha<2$, i.e., *A* may have a stable distribution totally skewed to the right. The concentration of mass in the tail areas of the stable family of distributions has made it a suitable candidate for modeling financial data. They have noted that the conditional expectation of the mixing variable, $E\left[A\middle|\mathbf{R}_2\right]$, is finite even though its unconditional counterpart, $E\left[A\right]$, may be infinite.

When the scaling variable *A* is random and $\mathbf{m}\neq\mathbf{0}$, in addition to the linear component, the regression equation necessarily contains the nonlinear factor associated with the scaling random variable. Similarly, the variance-covariance matrix is non-degenerate, involving both $E\left[A\middle|\mathbf{R}_2\right]$ and $V\left[A\middle|\mathbf{R}_2\right]$.

Undoubtedly, the model in eq. (16) with *A* being random and $\mathbf{m}\neq\mathbf{0}$ captures the behavior of return premiums for the situation where the volatility is stochastic. Under this setup, the unsystematic risks are modeled by scale mixtures and the heavy-tailed nature of errors is accommodated. The model generalizes the APT to accommodate stochastic volatility and should be of significant interest to financial analysts. In what follows, we study the conditional properties of the model in eq. (16) for some special distributions of *A*.

Log Asset Returns Under Inverse Gamma and Gamma Mixture of Normal Distributions

The expressions for both the regression and skedastic functions in eqs. (19) and (20), respectively, involve the two conditioning moments, $E\left[A\middle|\mathbf{R}_2\right]$ and $V\left[A\middle|\mathbf{R}_2\right]$. Computation of these moments requires closed-form expressions for certain distributions of *A*. In this subsection, we shall first provide explicit analytical expressions for these conditional moments for cases where the mixing variable *A* follows an inverse gamma distribution with *p* degrees of freedom, i.e., $A \sim p\big/\chi_p^2$, and where the mixing variable *A* follows a gamma distribution with shape parameters *p* and 1, respectively, i.e., $A \sim \Gamma\left(p,1\right)$.

Letting $A \sim p\big/\chi_p^2$ implies that the vector of log returns **R** in eq. (16) follows a non-central multivariate *t*-type distribution. This is a natural non-Gaussian family that allows the modeling of both heavy-tailed (for small *p*) as well as lighted-tailed distributions (for moderate *p*). For instance, Hurst and Platten (1997) remark that the inverse gamma distribution for the mixing variable fits well for modeling log asset returns. On the other hand, if one lets $A \sim \Gamma\left(p,1\right)$, then the vector of log returns **R** in eq. (16) follows a hyperbolic-type multivariate distribution. Madan and Seneta (1990) obtain some distributional properties of the log asset returns over time under gamma scale mixtures.

The following theorem provides closed-form expressions for the conditional moments of the scaling variable A when the latter follows either the inverse gamma or gamma distributions.

THEOREM 2. I. *Suppose that* $A \sim p/\chi_p^2$ *with probability density function given as*

$$f_A(a) = \frac{p^{p/2}}{2^{p/2}\Gamma(p/2)} a^{-\frac{p}{2}-1} e^{-p/2a},\ a > 0.$$

Then,

$$E\left[A^j \middle| \mathbf{R}_2 = \mathbf{r}_2\right] = \left(\frac{x}{y}\right)^j \frac{K_{(p+k-2j)/2}(xy)}{K_{(p+k)/2}(xy)},\ \mathbf{r}_2 \in \mathbb{R}^k; j \geq 1,$$

where $x^2 = \left(p + \|\mathbf{r}_2 - \boldsymbol{\mu}_2\|^2_{\Sigma_{22}^{-1}}\right)$, $y^2 = \|\mathbf{m}_2\|^2_{\Sigma_{22}^{-1}}$, $K_\lambda(.)$ is the modified Bessel function of the third kind and the norm $\|\mathbf{r}\|^2_{\Sigma_{22}^{-1}} = \langle \Sigma^{-1}\mathbf{r}, \mathbf{r}\rangle$.

Proof: See Appendix.

II. *Suppose that* $A \sim \Gamma(p,1)$ *with probability density function given as*

$$f_A(a) = \frac{1}{\Gamma(p)} a^{p-1} e^{-a},\ a > 0.$$

Then,

$$E\left[A^j \middle| \mathbf{R}_2 = \mathbf{r}_2\right] = \left(\frac{u}{v}\right)^j \frac{K_{p+j-\frac{k}{2}}(uv)}{K_{p-\frac{k}{2}}(uv)},\ \mathbf{r}_2 \in \mathbb{R}^k; j \geq 1,$$

where $u^2 = \|\mathbf{r}_2 - \boldsymbol{\mu}_2\|^2_{\Sigma_{22}^{-1}}$, $v^2 = 2 + \|\mathbf{m}_2\|^2_{\Sigma_{22}^{-1}}$, $K_\lambda(.)$ is the modified Bessel function of the third kind and the norm $\|\mathbf{r}\|^2_{\Sigma_{22}^{-1}} = \langle \Sigma^{-1}\mathbf{r}, \mathbf{r}\rangle$.

Proof: See Appendix.

It is well known that the Bessel function $K_\lambda(z)$ is analytic on the z-plane except at branch points $z = 0$ and $z = \infty$. Moreover, if $-\pi < \arg z < \pi$, then $K_\lambda(z)$ is real and positive whenever λ is real, and z is real and positive. From the above theorem, it is clear that the index λ is real and that the arguments in both Parts I and II of Theorem 2 are positive. Thus, $K_\lambda(z)$ is always real and positive in both Parts I and II of Theorem 2.

Using Theorem 2, the two conditional moments for eqs. (19) and (20) may be expressed as

When $A \sim p/\chi_p^2$,

$$E\left[A \middle| \mathbf{R}_2 = \mathbf{r}_2\right] = \frac{x}{y}\alpha_{(p+k)/2}\left(xy\right),\ \mathbf{r}_2 \in \mathbb{R}^k, \qquad ...(21)$$

$$V\left[A \middle| \mathbf{R}_2 = \mathbf{r}_2\right] = \left(\frac{x}{y}\right)^2 \alpha_{(p+k)/2}\left(xy\right)\left\{\alpha_{\frac{(p+k)}{2}-1}\left(xy\right) - \alpha_{(p+k)/2}\left(xy\right)\right\},\ \mathbf{r}_2 \in \mathbb{R}^k. \qquad ...(22)$$

When $A \sim \Gamma\left(p,1\right)$,

$$E\left[A \middle| \mathbf{R}_2 = \mathbf{r}_2\right] = \frac{u}{v}\alpha_{p-\frac{k}{2}}\left(uv\right),\ \mathbf{r}_2 \in \mathbb{R}^k, \qquad ...(23)$$

$$V\left[A \middle| \mathbf{R}_2 = \mathbf{r}_2\right] = \left(\frac{u}{v}\right)^2 \alpha_{p-\frac{k}{2}}\left(uv\right)\left\{\alpha_{p-\frac{k}{2}-1}\left(uv\right) - \alpha_{p-\frac{k}{2}}\left(uv\right)\right\},\ \mathbf{r}_2 \in \mathbb{R}^k, \qquad ...(24)$$

where $\alpha_\lambda\left(z\right) = K_{\lambda-1}\left(z\right)/K_\lambda\left(z\right), z > 0, \lambda \in \mathbb{R}$.

DATA ANALYSIS

Data Description

The data used in this study: daily and intraday (15-minute) stock prices of the Cisco Systems, Inc. (CISCO), Coca-Cola Company (COKE), Dell Computer Corporation (DELL), Microsoft Corporation (MSFT), and S&P500 index (S&P500) for the period spanning from January 1998 to December 2000 are obtained from AnalyzerXL LLC. These stocks are chosen mainly because of their liquidity, particularly the technology stocks which were very popular in the late 90's. For example, the 15-minute trading volume of the CISCO in 1998 averaged to about 2 million and reached the year high of 18 million. As for the S&P500, it is one of the commonly-used proxies for the market portfolio.

The log asset returns are then computed by taking the natural logarithms of the price ratios. Table 1 reports the summary statistics of the daily and intraday log asset returns. The large and positive (unconditional) excess kurtosis values of the intraday log asset returns indicate that their distributions have "heavier" tails and "sharper" peaks relative to that of the normal distribution; this concurs with the previous empirical findings.

Parameter Estimation

For the purpose of conciseness, we perform the comparative study of the scale mixture

Table 1
Summary Statistics of Log Asset Returns

	CISCO					
	Daily (252 obs.)			*Intraday (6530 obs.)*		
	1998	*1999*	*2000*	*1998*	*1999*	*2000*
Mean	0.00406	0.00367	-0.00044	0.00014	0.00013	-0.00006
Std Dev	0.02886	0.02640	0.04267	0.00560	0.00542	0.00806
Skewness	-0.59128	0.05456	0.51734	-0.51730	0.65486	-0.14116
Excess Kurtosis	3.88971	0.31198	1.17345	27.6850	12.8634	13.0440
Minimum	-0.13531	-0.06269	-0.12575	-0.10114	-0.06573	-0.11219
Maximum	0.09924	0.07128	0.16667	0.05584	0.05604	0.06454

	COKE					
	Daily (252 obs.)			*Intraday (6530 obs.)*		
	1998	*1999*	*2000*	*1998*	*1999*	*2000*
Mean	0.00024	-0.00032	0.00059	0.00000	-0.00002	0.00001
Std Dev	0.01948	0.01966	0.02692	0.00392	0.00393	0.00481
Skewness	-0.56106	0.51137	0.27313	-0.42679	-1.41467	0.41457
Excess Kurtosis	3.41665	2.19799	0.89365	22.9822	45.0368	13.2502
Minimum	-0.10524	-0.06404	-0.09069	-0.05439	-0.08744	-0.04461
Maximum	0.05088	0.08344	0.09654	0.05022	0.04058	0.06027

	DELL					
	Daily (252 obs.)			*Intraday (6530 obs.)*		
	1998	*1999*	*2000*	*1998*	*1999*	*2000*
Mean	0.00559	0.00191	-0.00329	0.00019	0.00005	-0.00017
Std Dev	0.03521	0.03443	0.04367	0.00687	0.00682	0.00835
Skewness	-0.28358	-0.03126	0.01918	0.24399	-1.10832	-2.21543
Excess Kurtosis	1.53678	0.11266	1.98852	26.8646	31.0459	53.5474
Minimum	-0.15790	-0.11779	-0.18943	-0.09580	-0.10380	-0.18056
Maximum	0.10071	0.08943	0.17790	0.09422	0.07735	0.10606

	MSFT					
	Daily (252 obs.)			*Intraday (6530 obs.)*		
	1998	*1999*	*2000*	*1998*	*1999*	*2000*
Mean	0.00332	0.00235	-0.00329	0.00012	0.00008	-0.00015
Std Dev	0.02383	0.02398	0.03544	0.00437	0.00462	0.00668
Skewness	-0.08075	0.40155	-0.00784	0.35719	0.57973	-2.20855
Excess Kurtosis	0.34590	0.69517	5.72548	18.1038	18.1379	101.290
Minimum	-0.08848	-0.06494	-0.15598	-0.03870	-0.05760	-0.16476
Maximum	0.06237	0.09880	0.19565	0.06614	0.06454	0.12474

	S&P 500 Index					
	Daily (252 obs.)			*Intraday (6530 obs.)*		
	1998	*1999*	*2000*	*1998*	*1999*	*2000*
Mean	0.00109	0.00082	-0.00027	0.00003	0.00003	-0.00001
Std Dev	0.01277	0.01136	0.01416	0.00282	0.00239	0.00277
Skewness	-0.52335	0.08816	0.07537	-0.28310	0.21440	-0.33209
Excess Kurtosis	4.51938	0.10190	1.37649	22.5907	9.21598	11.5762
Minimum	-0.06799	-0.02779	-0.05947	-0.03917	-0.02449	-0.03592
Maximum	0.05058	0.03563	0.04709	0.03582	0.02051	0.02618

nonlinear model (scaling variable A is random) with the usual linear APT (scaling variable A is constant, i.e., $A = 1$) by conditioning the individual log asset returns of CISCO, COKE, DELL, and MSFT upon S&P500. In addition, we will only report the results when the scaling variable follows the inverse gamma distribution as it has been found that the inverse gamma is more appropriate for both daily and intraday log asset returns analyses.

Given a set of data on the log asset returns, computation of $E\left[R_1 \middle| R_2\right]$ and $V\left[R_1 \middle| R_2\right]$ in Theorem 1 requires that one is able to compute $E\left[A \middle| R_2\right]$ and $V\left[A \middle| R_2\right]$. Throughout the analysis, we let $m_1 = -\frac{1}{2}\left(\sigma_{11} + \sigma_{12}\right)$ and $m_2 = -\frac{1}{2}\left(\sigma_{21} + \sigma_{22}\right)$, as derived based on the geometric Brownian motion. Then, the relevant expressions for these conditional moments of the scaling variable are provided in eqs. (21) and (22) for the case where the scaling variable follows inverse gamma distribution. Computations of these expressions involve estimation of the unknown parameters σ_{11}, σ_{12}, and σ_{22} as well as the computation of the modified Bessel function, $K_\lambda(.)$. It may be noted that the mean of log asset returns $\boldsymbol{\mu}' = \left(\mu_1, \mu_2\right)'$ may be assumed to be zero under the usual fair market scenario and hence, requires no estimation.

The method of moments is well suited for estimating the unknown covariance matrix Σ in this case. The method involves the derivation of theoretical expressions for Σ through eq. (18), the first two moments of the variable A via the inverse gamma distribution and the computation of sample means and sample covariance matrix of R_1 and R_2. While maximum likelihood estimation would also be appropriate, we here choose to implement the method of moments in light of the highly nonlinear nature of the likelihood in the parameters Σ.

Utilizing eq. (18), it is not hard to obtain the following expressions:

$$\sigma_{11} = \frac{E[A]^2 V[R_1] - V[A]\ \left(\mu_1 - E[R_1]\ \right)^2}{E[A]^3}, \qquad ...(25)$$

$$\sigma_{22} = \frac{E[A]^2 V[R_2] - V[A]\ \left(\mu_2 - E[R_2]\ \right)^2}{E[A]^3}, \qquad ...(26)$$

$$\sigma_{12} = \frac{E[A]^2 C[R_1, R_2] - V[A]\ \left(\mu_1 - E[R_1]\ \right)\ \left(\mu_2 - E[R_2]\ \right)}{E[A]^3}, \qquad ...(27)$$

where $C\left[R_1, R_2\right]$ denotes the covariance of R_1 and R_2. Furthermore, when $A \sim p/\chi_p^2$, the first two moments are given by $E[A] = p/(p-2)$ and $V[A] = 2p^2/(p-2)^2(p-4)$.

The parameters σ_{11}, σ_{12}, and σ_{22} may then be estimated by substituting the sample moments for R_1 and R_2 into eqs. (25)-(27). In solving these equations, one sets $\boldsymbol{\mu}' = (0,\ 0)'$.

The Bessel function $K_\lambda(.)$ may be computed through standard software package. One thus proceeds to compute the conditional moments in eqs. (21) and (22) and then $E[R_1|R_2]$ and $V[R_1|R_2]$ in Theorem 1. The latter for the linear model under normality are computed by simply letting $A = 1$ in the Theorem.

Noting that our model and inference are based on the mixture setup leading to a general form of the APT for scale mixture family of distributions, it is convenient to let our model be MAPT, the mixture-based APT. For each of the log asset returns data considered in our analysis, we then compare the MAPT with the APT for effectiveness in performance with respect to both regression and skedastic functions. The comparison is based upon the residuals of the respective models. Thus, let $\hat{\varepsilon}_i$ and $\tilde{\varepsilon}_i$ for $i = 1,\ldots,n$, be the respective residuals for MAPT and the APT. It is intuitively appealing to capture the mixture model effects (*MME*) upon the two conditional moments by computing

$$MME_E = \frac{SSE(MAPT_E) - SSE(APT_E)}{SSE(APT_E)} \times 100\%, \quad \text{...(28)}$$

$$MME_V = \frac{SSE(MAPT_V) - SSE(APT_V)}{SSE(APT_V)} \times 100\%, \quad \text{...(29)}$$

where $SSE(MAPT_E) = \sum_{i=1}^n \hat{\varepsilon}_i^2$, $SSE(APT_E) = \sum_{i=1}^n \tilde{\varepsilon}_i^2$, $SSE(MAPT_V) = \sum_{i=1}^n \left(\left(\hat{\varepsilon}_i^2 / \hat{V}[R_1|R_2]_{MAPT}\right) - 1\right)^2$ and $SSE(APT_V) = \sum_{i=1}^n \left(\left(\tilde{\varepsilon}_i^2 / \hat{V}[R_1|R_2]_{APT}\right) - 1\right)^2$. Note that under the above definition, with respect to the expectation, the MAPT will be more effective than the APT whenever MME_E is negative. Otherwise, the APT will be more efficient. Similar criterion applies to the variance via MME_V.

Even though we let the scaling variable follow the inverse gamma distribution, in reality, we still do not know the true value of *p*, the degrees of freedom parameter. Noting that the two conditional moments exist only when $p > 4$, we compute the above scale mixture effects in (28) and (29) for various values of *p* beginning with $p = 6$. The computations for the intraday data on all the data sets against S&P500 resulted in smallest values for both MME_E and MME_V when $p = 6$. For purposes of comparison, we maintained the same $p = 6$ for the analysis of the daily returns data as well. Consequently, we only present below the computations when the scaling variable follows inverse gamma distribution with 6 degrees of freedom. We do, however, present subsequently two graphs relating to the analysis of daily returns data from COKE and DELL that illustrate the behavior of MME_V against p, the degrees of freedom. The following table provides a summary of the relevant computations including the mixture model effects upon the two conditional moments.

Table 2
Comparisons of the MAPT and APT Models through MME_E and MME_V

	CISCO					
	Daily (252 obs.)			*Intraday (6530 obs.)*		
	1998	*1999*	*2000*	*1998*	*1999*	*2000*
SSE ($MAPT_E$)	0.11153	0.08893	0.23634	0.12124	0.11378	0.22474
SSE (APT_E)	0.10990	0.08739	0.23595	0.12115	0.11371	0.22469
SSE ($MAPT_V$)	1369.86	846.70	838.78	97400	116028	46992
SSE (APT_V)	1318.42	606.91	801.84	274306	116924	85929
MME_E	1.49	1.77	0.17	0.08	0.06	0.02
MME_V	3.90	39.51	4.61	-64.49	-0.77	-45.31

	COKE					
	Daily (252 obs.)			*Intraday (6530 obs.)*		
	1998	*1999*	*2000*	*1998*	*1999*	*2000*
SSE ($MAPT_E$)	0.05792	0.08705	0.18070	0.05983	0.08325	0.14413
SSE (APT_E)	0.05786	0.08698	0.18040	0.05983	0.08326	0.14414
SSE ($MAPT_V$)	1237.81	1224.72	742.12	71798	83897	86866
SSE (APT_V)	992.09	1064.62	690.35	138057	368970	109912
MME_E	0.11	0.09	0.17	0.01	-0.01	-0.01
MME_V	24.77	15.04	7.50	-47.99	-77.26	-20.97

	DELL					
	Daily (252 obs.)			*Intraday (6530 obs.)*		
	1998	*1999*	*2000*	*1998*	*1999*	*2000*
SSE ($MAPT_E$)	0.19148	0.22048	0.34110	0.18408	0.22106	0.33951
SSE (APT_E)	0.18734	0.22044	0.33992	0.18390	0.22108	0.33945
SSE ($MAPT_V$)	1104.44	839.58	851.16	102359	136667	145873
SSE (APT_V)	713.12	699.83	884.14	205534	370057	437400
MME_E	2.21	0.02	0.35	0.09	-0.01	0.02
MME_V	54.87	19.97	-3.73	-50.20	-63.07	-66.65

	MSFT					
	Daily (252 obs.)			*Intraday (6530 obs.)*		
	1998	*1999*	*2000*	*1998*	*1999*	*2000*
SSE ($MAPT_E$)	0.07336	0.08685	0.22460	0.08233	0.09121	0.21278
SSE (APT_E)	0.07231	0.08641	0.22292	0.08228	0.09120	0.21264
SSE ($MAPT_V$)	771.66	1262.05	2448.35	96418	151145	430029
SSE (APT_V)	483.27	1025.31	2111.51	161518	210279	910136
MME_E	1.45	0.51	0.75	0.06	0.02	0.06
MME_V	59.68	23.09	15.95	-40.30	-28.12	-52.75

In what follows, we present selected graphs representing daily and intraday data analyses for CISCO in year 1998. The graphs represent the corresponding model fits, prediction curves based upon twice and thrice the conditional standard deviation and the normal probability plots for the residuals.

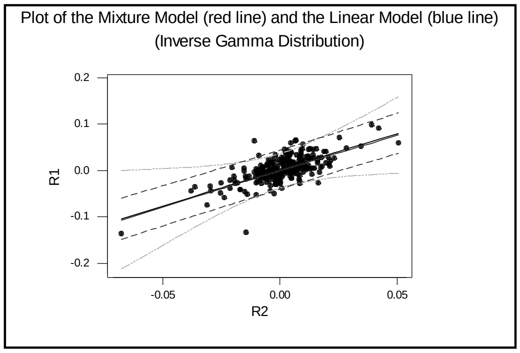

Figure 1a.

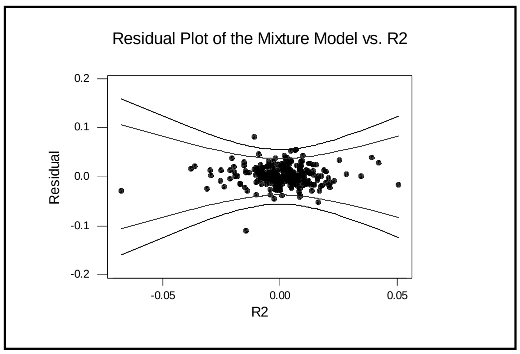

Figure 1b

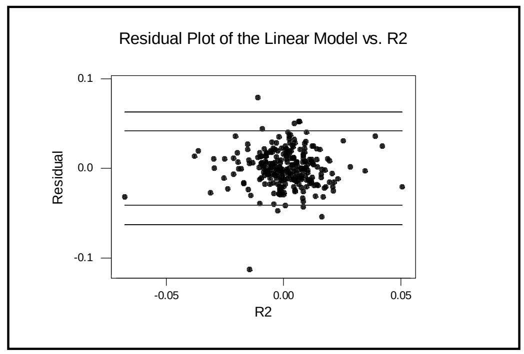

Figure 1c

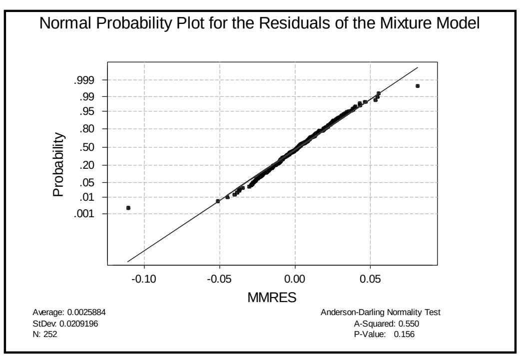

Figure 1d

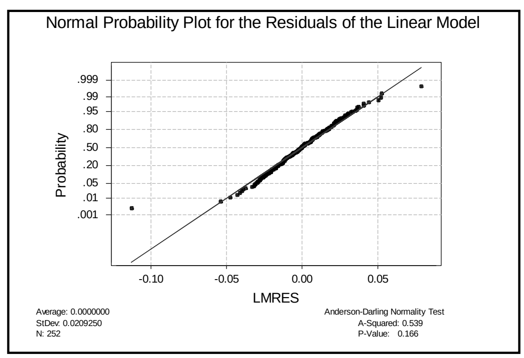

Figure 1e

Figure 1a. **Fits for both APT and MAPT models**. The APT linear fit and the MAPT scale mixture fit for the daily log returns data for CISCO, year 1998.

Figure 1b, 1c. **Residual plots and bands**. Figure 1b plots residuals for the MAPT model together with two and three standard deviation bands for CISCO, year 1998. Figure 1c represents the same for the APT model.

Figure 1d, 1e. **Normal probability plots**. Figure 1d represents the normal probability plot for residuals from the MAPT fit for CISCO, year 1998. Figure 1e represents the same for APT fit.

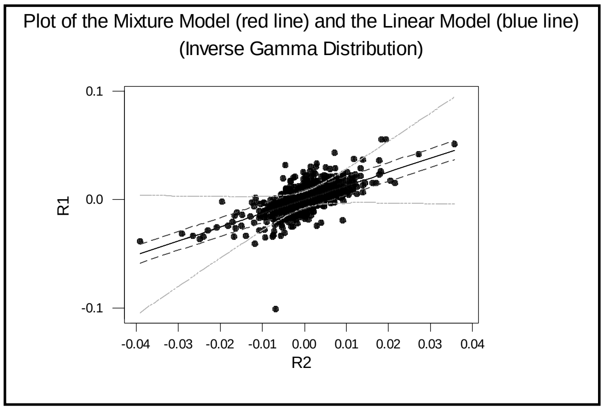

Figure 2a

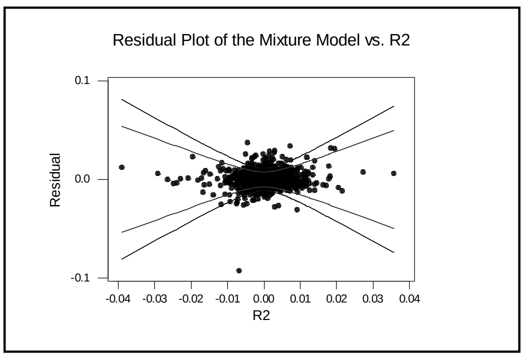

Figure 2b

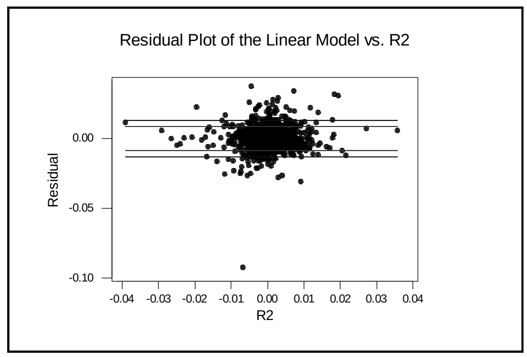

Figure 2c

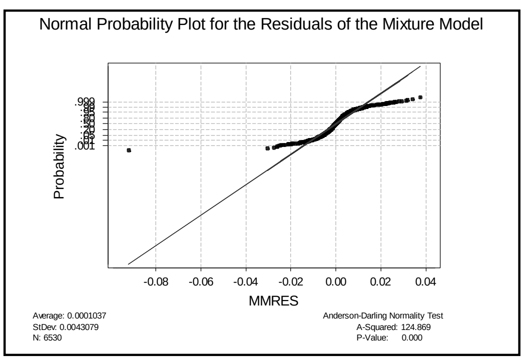

Figure 2d

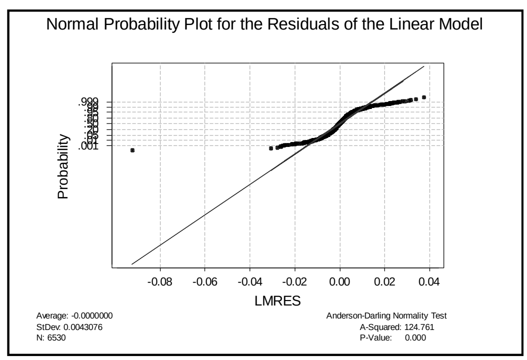

Figure 2e

Figure 2a. **Fits for both APT and MAPT models**. The APT linear fit and the MAPT scale mixture fit for the intra-day log returns data for CISCO, year 1998.

Figure 2b, 2c. **Residual plots and bands**. Figure 2b plots residuals for the MAPT model together with two and three standard deviation bands for CISCO, year 1998. Figure 2c represents the same for the APT model.

Figure 2d, 2e. **Normal probability plots**. Figure 2d represents the normal probability plot for residuals from the MAPT fit for CISCO, year 1998. Figure 2e represents the same for APT fit.

Discussion

In comparing the two models, MAPT and APT, we first note from Table 2 that the MME_E percentages range between a minimum of -0.01% (DELL, intraday, in year 1999) to a maximum of 2.21% (DELL, daily, in year 1998) with median value of 0.08%. Thus, it is quite evident that the two models are close to each other for all fits considered in the analysis. The linear APT and the nonlinear MAPT appear to behave in the same manner with regard to the expectation fits. This implies that the contribution of the nonlinear term in the MAPT is hardly significant, that is, on the whole, the regression is captured equally well by both the nonlinear MAPT and its linear counterpart in daily as well as intraday data analyses.

Now, one might ask whether the same is true for the conditional variance. Again, from Table 2, one notes that the MME_V percentages behave rather differently for the daily and the intraday analyses. In particular, all the MME_V percentages for the daily returns are positive except for DELL in year 2000, whereas the MME_V percentages for the intraday returns are all negative. The median value of the MME_V percentages for the daily returns works out to 17.96%. On the other hand, the median for the MME_V percentages for the intraday analyses is -49.10%. Thus, with regard to conditional variance, the APT seems to outperform the MAPT for the daily returns, whereas the MAPT is a far more superior fit for the analyses of the intraday returns.

We believe the difference between the MAPT and APT in capturing the variability can be traced back to the distributional properties of the log asset returns. As such, normal probability plots of the residuals from the MAPT and APT fits in the daily and intraday analyses are constructed and carefully studied. To conserve space, we only include the normal probability plots of the residuals from the MAPT and APT fits in the daily and intraday analyses of CISCO for the year 1998. It is important to note that the normal probability plots presented in Figures 1d and 1e for the daily returns are representative of what we have seen for most of the other daily return analyses. Similarly, the plots presented in Figures 2d and 2e also overwhelmingly represent the other plots for the intraday return analyses.

Figures 1d and 1e clearly demonstrate that the underlying daily log asset returns follow the normal distribution whereas Figures 2d and 2e show beyond doubt that the intraday log asset returns depart sharply from the normal distribution. Moreover, one can conclude from Figures 2d and 2e that the intraday log asset returns satisfy symmetry even as they demonstrate a heavy-tailed nature in their distribution. Note that the sharp departure from normality and strong indication of symmetry support the choice of the inverse gamma distribution for the scaling variable leading to a heavy-tailed *t*-type of distribution for the conditional intraday returns. It is this symmetry and heavy-tailed nature of the intraday return distribution as opposed to the normal distribution behavior of the daily log asset returns that together explain the differences in MAPT and APT.

Specifically, the scale mixture family with a heavy-tailed nature, particularly with small degrees of freedom $\left(p=6\right)$, is a clear misspecification of the underlying distribution for the daily returns. Hence, it is only appropriate that the linear APT fitted under the assumption of normality should outperform the scale mixture model in this case. Recall that the median MME_V

percentage for the daily returns is positive, 17.96%. The symmetric nature of the normal distribution ensures that the misspecified MAPT compares evenly with the APT in terms of the expectation fit.

On the other hand, the sharp departure from normality and strong indication of symmetry of the intraday returns support the choice of the inverse gamma distribution for the scaling variable leading to a heavy-tailed *t*-type distribution for these returns. Thus, one is able to simultaneously rationalize the observations regarding the intraday return analysis. Namely, the symmetric nature of the intraday data, on average, compares evenly with both the APT fit as well as the MAPT fit, whereas the heavy-tailed nature of the data is captured more sharply by the MAPT only (median MME_V = –49.10%).

It is easy to see the reason behind the normality of the daily returns data as well as the non-normality of the intraday data. The intraday data being high-frequency data has little scope for aggregation effect, whereas the daily data can be assumed to have substantial aggregation effect.

Choice of Degrees of Freedom, p

Finally, we wish to discuss the behavior of MME_V as a function of *p*, the degrees of freedom. Note that the expectation fit MME_E is practically unaffected by the choice of *p*. We illustrate the behavior of MME_V as a function of *p* through the analysis of daily returns for COKE in year 1999 and also the analysis of daily returns for DELL in year 1998. There is some contrasting evidence in both of these analyses. In the case of the analysis for COKE, the normal probability plot in Figure 3a shows clear departure from normality. On the other hand, the normal probability in Figure 4a shows the normal distribution to be equivocal in this case. Accordingly, one anticipates that the MAPT should fit the COKE data better perhaps with somewhat higher degrees of freedom rather than the very low value of $p = 6$. In contrast, the APT should fit the DELL data well and the value of *p* in reality should be extremely large. Figures 3b and 4b plot the behavior of MME_V against *p* for COKE and DELL, respectively. As expected, MME_V in Figure 3b drops sharply for increasing *p* and in fact becomes negative subsequent to p = 17, thus indicating that the MAPT with some *p* value below 17 would be the best fit. Due to some computational overruns associated with the Bessel functions, we are unable to go beyond p = 25 as provided in the plot. Otherwise, one can capture the best fitting *p* in this case.

In the case of the analysis of DELL, the MME_V in Figure 4b always stays positive for increasing *p* even as it monotonically decreases in *p*. Therefore, the best fitting *p* will perhaps be achieved only at a very high value. This is consistent with the fact that the underlying distribution is normal.

CONCLUDING REMARKS AND DIRECTIONS FOR FUTURE RESEARCH

The main objective of this study is to understand the nonlinear aspects of the multifactor pricing models. Empirical evidence has disproved the omnipresence of the conventional normality assumption on the financial asset returns, which prompts the need to consider situations under

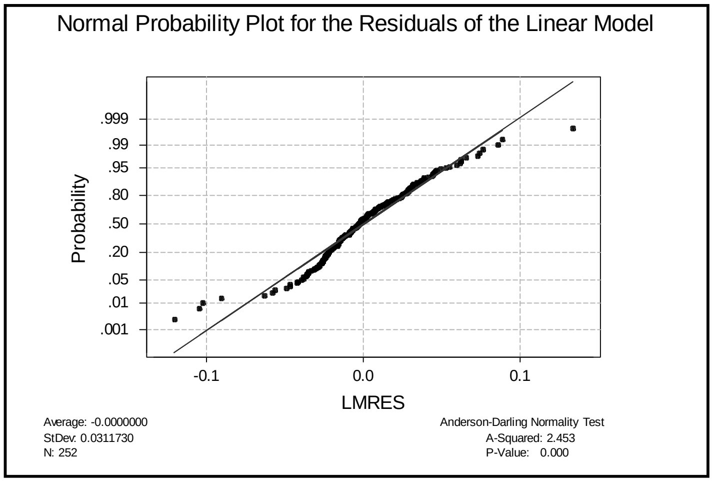

Figure 3a. **Normal probability plot**. The plot is based on residuals from the linear APT model fit to log returns data for AT&T, year 2000.

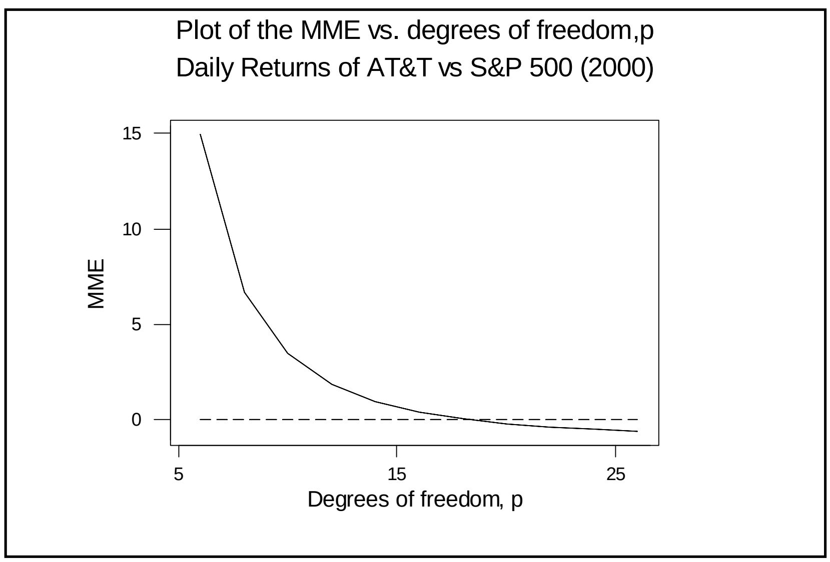

Figure 3b. MME_V **against *p*, the degrees of freedom**. The mixture model effect for the conditional variance against the APT model is computed for daily log returns data for AT&T, year 2000, at various degrees of freedom parameter p under the inverse gamma distribution for the scaling variable.

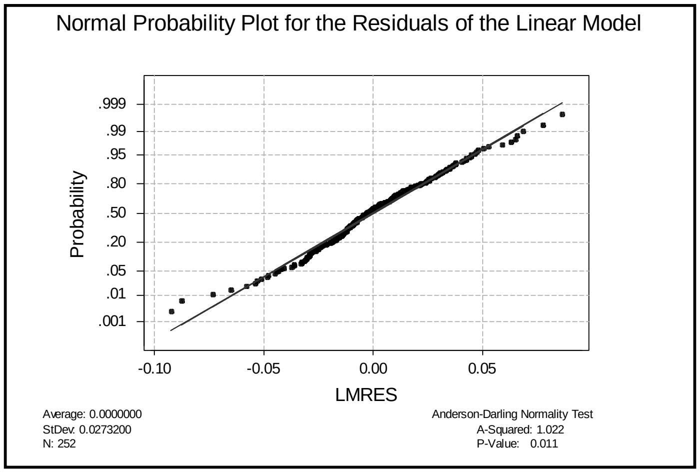

Figure 4a. **Normal probability plot**. The plot is based on residuals from the linear APT model fit to log returns data for DELL, year 1998.

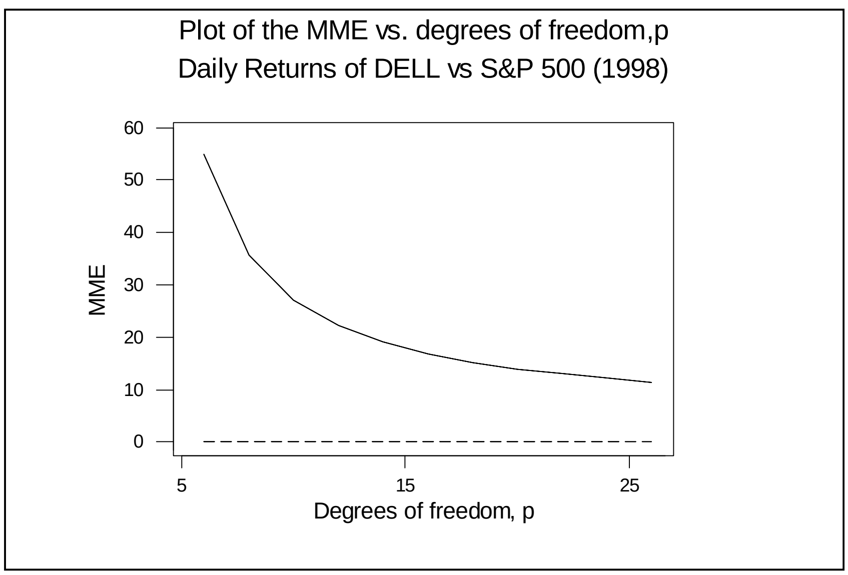

Figure 4b. MME_V **against *p*, the degrees of freedom**. The mixture model effect for the conditional variance against the APT model is computed for daily log returns data for DELL, year 1998, at various degrees of freedom parameter p under the inverse gamma distribution for the scaling variable.

strong departures from normality. We develop a general framework for understanding the properties of the multifactor pricing models when the log asset returns are modeled by members of the scale mixture family. The derived analytical expressions contain nonlinear terms such as the conditional expectation and variance-covariance of the scaling variable. The presence of these terms may explain the nonlinear and heteroskedastic features exhibited in high-frequency financial data.

Empirical analyses are performed on data sets involving both daily and intraday log returns of the stocks from the Cisco Systems, Inc. (CISCO), Coca-Cola Company (COKE), Dell Computer Corporation (DELL), and the Microsoft Corporation (MSFT) conditional upon the log returns of the S&P500 index using the methodologies discussed in Section III. The results reveal that the nonlinear model dominates over its linear counterpart in modeling high-frequency financial data with symmetrical and heavy-tailed features by a median mixture model effect (MME_ν) of -49.10%.

In conclusion, we recommend the use of the nonlinear model (MAPT) over the linear model (APT) in the analysis of (non-normal) intraday data since the former does a better job in explaining the conditional variance. However, if there is no dispute on the normality assumption, especially in the analysis of daily data, the linear model (APT) is preferred because of its simplicity.

Some open questions and directions for future research have been identified and thus proposed. First, this study only considers the formulation of nonlinear models in the cross-sectional framework. Reformulation of the techniques discussed here to incorporate time-series effects is deemed feasible, though difficult. This will definitely improve the practicability and predictability of the proposed model. Second, instead of restricting the volatility process to a scale mixture of normal distribution, it could be generalized even more by being a random matrix. Finally, as seen in Section IV, low degrees of freedom of the inverse gamma mixing variable produce better results and are thus preferred. However, the estimation of the optimal value for the degrees of freedom is outside the scope of this study. A possible way of determining this value may be attributed to the use of kernel estimation in nonparametric statistics.

APPENDIX

Proof of Theorem 2

From Cambanis et al. (2000), it can be seen that

$$E\left[A^j \middle| \mathbf{R}_2 = \mathbf{r}_2\right] = \frac{\int_0^\infty a^{j-\frac{k}{2}} \exp\left(-\frac{1}{2a}\left\|\mathbf{r}_2 - \boldsymbol{\mu}_2 - \mathbf{m}_2 a\right\|^2_{\Sigma_{22}^{-1}}\right) f_A(a)\, da}{\int_0^\infty a^{-\frac{k}{2}} \exp\left(-\frac{1}{2a}\left\|\mathbf{r}_2 - \boldsymbol{\mu}_2 - \mathbf{m}_2 a\right\|^2_{\Sigma_{22}^{-1}}\right) f_A(a)\, da}, \quad \text{(A1)}$$

where $f_A(a)$ denotes the probability density function of the scaling random variable *A*. Thus, it suffices to derive the expression for the integral in the numerator (say *NUM*) for any $j \geq 0$. We shall provide the necessary details for Part I of Theorem 2, where $A \sim p/\chi^2_p$. The details for Part II, where $A \sim \Gamma(p,1)$, can be obtained similarly and are thus omitted.

$$NUM = \frac{p^{p/2}}{2^{p/2}\Gamma(p/2)} \int_0^\infty a^{\frac{2j-(k+p)}{2}-1} \exp\left(-\frac{1}{2a}\left\|\mathbf{r}_2 - \boldsymbol{\mu}_2 - \mathbf{m}_2 a\right\|^2_{\Sigma_{22}^{-1}}\right) e^{-p/2a}\, da$$

$$= \frac{p^{p/2} e^{\left\langle (\mathbf{r}_2 - \boldsymbol{\mu}_2)\Sigma_{22}^{-1}, \mathbf{m}_2 \right\rangle}}{2^{p/2}\Gamma(p/2)} \int_0^\infty a^{\frac{2j-(k+p)}{2}-1} \exp\left[-\frac{a}{2}\left\|\mathbf{m}_2\right\|^2_{\Sigma_{22}^{-1}} - \frac{1}{2a}\left(p + \left\|\mathbf{r}_2 - \boldsymbol{\mu}_2\right\|^2_{\Sigma_{22}^{-1}}\right)\right] da$$

$$= \frac{p^{p/2} e^{\left\langle (\mathbf{r}_2 - \boldsymbol{\mu}_2)\Sigma_{22}^{-1}, \mathbf{m}_2 \right\rangle} 2^{(2j-p-k)/2}}{2^{p/2}\Gamma(p/2)\, y^{2j-p-k}} \int_0^\infty a^{\frac{2j-(k+p)}{2}-1} \exp\left[-a - \frac{1}{2a}(xy)^2\right] da,$$

where $x^2 = p + \left\|\mathbf{r}_2 - \boldsymbol{\mu}_2\right\|^2_{\Sigma_{22}^{-1}}$ and $y^2 = \left\|\mathbf{m}_2\right\|^2_{\Sigma_{22}^{-1}}$. Using eq. (8.432.6) from Gradshteyn and Rhyzik (1980), the above expression can be simplified as

$$NUM = \frac{p^{p/2} e^{\left\langle (\mathbf{r}_2 - \boldsymbol{\mu}_2)\Sigma_{22}^{-1}, \mathbf{m}_2 \right\rangle}}{2^{\frac{p}{2}-1}\Gamma(p/2)} \left(\frac{x}{y}\right)^{\frac{2j-k-p}{2}} K_{\frac{p+k-2j}{2}}(xy). \quad \text{...(A2)}$$

The expressions for the two conditional moments in Part I of Theorem 2 then follow easily by substituting eq. (A2) with $j = 1$ and 2, respectively, in the numerator of eq. (A1), and (A2) with $j = 0$ in the denominator of eq. (A1).

REFERENCES

Bansal, R. and Viswanathan, S., (1993), “No Arbitrage and Arbitrage Pricing: A New Approach,” *Journal of Finance*, 48: 1231-1262.

Barndorff-Nielsen, O. E., (1997), “Normal Inverse Gaussian Distributions and Stochastic Volatility Modelling,” *Scandinavian Journal of Statistics*, 24: 1-13.

——(1998), "Processes of Normal Inverse Gaussian Type," *Financial Stochastics*, 2: 41-68.

Blattberg, R. C., and Gonedes, N. J., (1974), "A Comparison of the Stable and Student Distributions as Statistical Models for Stock Prices," *Journal of Business*, 47: 244-280.

Bollerslev, T., (1986), "Generalized Autoregressive Conditional Heteroskedasticity," *Journal of Econometrics*, 31: 307-327.

Bollerslev, T., Chou, R. and Kroner, K., (1992), "ARCH Modeling in Finance: A Review of the Theory and Empirical Evidence," *Journal of Econometrics*, 52: 5-59.

Cambanis, S. and Fotopoulos, S. B., (1995), "Conditional Variance for Stable Random Vectors," *Probability and Mathematical Statistics*, 15, 195-214.

Cambanis, S., Fotopoulos, S. B. and He, L., (2000), "On the Conditional Variance for Scale Mixtures of Normal Distributions," *Journal of Multivariate Analysis*, 73, 163-192.

Cambanis, S., Huang, S. and Simons, G., (1981), "On the Theory of Elliptically Contoured Distributions," *Journal of Multivariate Analysis*, 11, 348-365.

Cambanis, S. and Wu, W., (1992), "Multiple Regressions on Stable Vectors," *Journal of Multivariate Analysis*, 41, 243-272.

Chamberlain, G., (1983), "Funds, Factors, and Diversification in Arbitrage Pricing Models," *Econometrica*, 51: 1305-1323.

Chamberlain, G. and Rothschild, M., (1983), "Arbitrage, Factor Structure, and Mean-Variance Analysis on Large Asset Markets," *Econometrica*, 51: 1281-1304.

Connor, G. and Korajczyk, R. A., (1988), "Risk and Return in an Equilibrium APT: Application of a New Test Methodology," *Journal of Financial Economics*, 21: 255-290.

Dhrymes, P., Friend, I., Gultekin, B. and Gultekin, M., (1984), "A Critical Reexamination of the Empirical Evidence on the Arbitrage Pricing Theory," *Journal of Finance*, 39: 323-346.

Fama, E. F., (1991), "Efficient Capital Markets: II," *Journal of Finance*, 46: 1575-1618.

Fang, K. T. and Zhang Y. T., (1990), *Generalized Multivariate Analysis*. Springer-Verlag, Berlin.

Feinstone, L. J., (1987), "Minute by Minute: Efficiency, Normality and Randomness in Intra-Daily Asset Prices," *Journal of Applied Econometrics*, 2: 193-214.

Ferson, W. E. and Korajczyk, R. A., (1995), "Do Arbitrage Pricing Models Explain the Predictability of Stock Returns?", *Journal of Business*, 68: 309-349.

Fotopoulos, S. B., (1998), "On the Conditional Variance-Covariance Matrix of Stable Random Vectors," In Karatzas, Rajput, B. S. and Taqqu, M.S. (eds), Stochastic Processes and Related Topics, Birkhauser, Boston, II: 123-140.

Fotopoulos, S. B. and He, L., (1999), "Error Bounds for Asymptotic Expansion of the Conditional Variance of the Scale Mixtures of the Multivariate Normal Distributions," *Annals of the Institute of Statistical Mathematics*, 51: 731-747.

——(2001), "The Conditional Variance for Gamma Mixtures of Normal Distributions," *Statistics*, 35: 319-330.

Gerhard, F. and Hautsch, N., (2002), "Volatility Estimation on the Basis of Price Intensities," *Journal of Empirical Finance*, 9: 57-89.

Gradshteyn, I. S. and Ryzhik, I. M., (1980), *Table of Integrals, Series, and Products,* Academic Press, New York.

Hardin, C. D. Jr., (1982), "On the Linearity of the Regression," *Zeitschrift für Wahrscheinlich - keitstheorie und Verwandte Gebiete*, 61, 293-302.

Huberman, G., (1982), “A Simple Approach to Arbitrage Pricing Theory,” *Journal of Economic Theory*, 28: 183-191.

Hull, J. and White, A., (1987), “The Pricing of Options on Assets with Stochastic Volatilities,” *Journal of Finance*, 42: 281-300.

Hurst, S. R. and Platen, E., (1997), “The Marginal Distributions of Returns and Volatility,” In Dodge, Y. (ed), L1-Statistical Procedures and Related Topics, IMS Lecture Notes-Monograph Series, 31: 301-314.

Ingersoll, J. E., (1984), “Some Results in the Theory of Arbitrage Pricing,” *Journal of Finance*, 39: 1021-1039.

Jobson, J. D., (1982), “A Multivariate Linear Regression Test for the Arbitrage Pricing Theory,” *Journal of Finance*, 37: 1037-1042.

Karatzas, I. and Shreve, S. E., (1998), *Methods of Mathematical Finance*. Springer-Verlag, New York.

Kloeden, P. E. and Platten, E., (1995), *Numerical Solution of Stochastic Differential Equations*. Springer-Verlag, Berlin.

Kotz, S., Kozubowski, T. and Podgorski, K., (2001), *The Laplace Distribution and Generalizations: A Revisit with Applications to Communications, Economics, Engineering, and Finance*. Springer, New York: Birkhauser.

Kuchler, U., Neuman, N. K., Sorensen, M. and Streller, A., (1995), *Stock Returns and Hyperbolic Distributions*. Working paper, Humboldt University of Berlin.

Lintner, J., (1965), “The Valuation of Risky Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets,” *The Review of Economics and Statistics*, 47: 13-37.

Madan, D. B. and Milne, F., (1991), “Option Pricing with VG Martingale Components,” *Mathematical Finance*, 1: 39-55.

Madan, D. B. and Seneta, E., (1990), “The Variance Gamma (V.G.) Model for Share Market Returns,” *Journal of Business*, 63: 511-524.

Markowitz, H., (1959), *Portfolio Selection: Efficient Diversification of Investments*, John Wiley, New York.

McCulloch, J., (1996), “Financial Applications of Stable Distributions,” In Maddala, G. and Rao, C. (eds), Statistical Methods in Finance, Handbook of Statistics Series, 14: 393-425.

Melino, A. and Turnbull, S., (1990), “Pricing Foreign Currency Options with Stochastic Volatility,” *Journal of Econometrics*, 45: 239-265.

Milne, F., (1988), “Arbitrage and Diversification in a General Equilibrium Asset Economy,” *Econometrica*, 56: 815-840.

Praetz, P. D., (1972), “The Distribution of Share Price Changes,” *Journal of Business*, 45: 49-55.

Rachev, S. and SenGupta, A., (1993), “Laplace-Weibull Mixtures for Modeling Price Changes,” *Management Science*, 39: 1029-1038.

Roll, R. and Ross, S. A., (1980), “An Empirical Investigation of the Arbitrage Pricing Theory,” *Journal of Finance*, 35: 1073-1103.

Ross, S. A., (1976), “The Arbitrage Pricing Theory of Capital Asset Pricing,” *Journal of Economic Theory*, 13: 341-360.

Rydberg, T. H., (1999), “Generalized Hyperbolic Diffusion Processes with Applications in Finance,” *Mathematical Finance*, 9: 183-201.

Rydberg, T. H. and Shepard, N., (2001), "A Modeling Framework for the Prices and Times of Trades made on the New York Stock Exchange," Working paper, Nuffield College, Oxford, United Kingdom.

Samorodnitsky, G. and Taqqu, M. S., (1994), *Stable Non-Gaussian Random Processes: Stochastic Models with Infinite Variance*, Chapman & Hall, New York.

Sharpe, W. F., (1964), "Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk," *Journal of Finance*, 19: 425-442.

Willinger, W., Taqqu, M. and Teverovsky, V., (1999), "Stock Market Prices and Long-Range Dependence," *Finance and Stochastics*, 3: 1-13.