Conditional Asset Allocation under Non-Normality: How Costly Is the Mean-Variance Criterion?

Eric Jondeau[a] and Michael Rockinger[b,*]

February 2005

Abstract

We evaluate how departure from normality may affect the conditional allocation of wealth. The expected utility function is approximated by a fourth-order Taylor expansion that allows for non-normal returns. Market returns are characterized by a joint model that captures the time dependency and the shape of the distribution. We show that under large departure from normality, the mean-variance criterion can lead to portfolio weights that differ significantly from those obtained using the optimal strategy accounting for non-normality. In addition, the opportunity cost for a risk-adverse investor to use the sub-optimal mean-variance criterion can be very large.

Keywords: Volatility, Skewness, Kurtosis, GARCH model, Multivariate skewed Student-t distribution, Stock returns, Asset allocation, Emerging markets.

JEL classification: C22, C51, G12.

[a] Corresponding author. HEC Lausanne and FAME, Institute of Banking and Finance, Route de Chavannes 33, CH 1007 Lausanne-Vidy, Switzerland, e-mail: eric.jondeau@unil.ch, Tel: +41 (0)21 692 33 49.

[b] Same affiliation as E. Jondeau, e-mail: michael.rockinger@unil.ch, Tel: +41 (0)21 692 33 48.

[*]Michael Rockinger is also CEPR. Both authors acknowledge help from the Swiss National Science Foundation through NCCR (Financial Valuation and Risk Management) as well as through the grant on "Multivariate modelling of asset prices using the information contained in transactions data". The usual disclaimer applies.

1 Introduction

To solve asset-allocation problems, the well-known mean-variance criterion proposed by Markowitz (1952) has provided very sensible results for a very wide range of situations. While some authors have argued that the expected utility function may be more appropriately approximated by a function of higher moments (Arditti, 1967, and Samuelson, 1970), early empirical evidence suggests that mean-variance criterion results in allocations that are very similar to the ones obtained using a direct optimization of the expected utility (Levy and Markowitz, 1979, Pulley, 1981, and Kroll, Levy, and Markowitz, 1984).

An explanation of the good performance of the mean-variance criterion reported in these papers may be that, although returns are non-normal, they are driven by an elliptical distribution, for which the mean-variance approximation of the expected utility remains exact for all utility functions (Chamberlain, 1983). In contrast, under large departure from normality, in particular when the distribution is severely asymmetric, Chunachinda et al. (1997), Athayde and Flôres (2004) and Jondeau and Rockinger (2004) show that the mean-variance criterion may fail to correctly approximate the expected utility. In such a case, a three- or four-moment optimization strategy provides a better approximation of the expected utility.[1]

A limitation of this previous evidence is that it assumes constant investment opportunities while a huge literature has argued that asset returns have time varying moments. However, extension to a conditional asset-allocation, however, is much more difficult to implement. An avenue that has recently been followed by Ang and Bekaert (2002) and Guidolin and Timmermann (2003) use the Markov-switching approach. In such framework, the mean and variance of returns are allowed to vary over time as regimes change. Although the distribution of returns is conditionally normal, it provides a measure of the opportunity cost of assuming iid normal returns. Ang and Bekaert (2002) report several results of importance. First, the cost of ignoring Markov-switching is moderate, since it ranges between 0.1 and 1.3 cents per dollar per

[1] Jurczenko and Maillet (2001) investigate how the asset-allocation problem should be modified in cases where the conditional distribution of returns is characterized by higher moments.

year, depending on the model and the risk aversion considered. Second, the cost of using a myopic strategy is essentially zero, so that the multi-period investment seems to be irrelevant. It should be noted that while this approach is appropriate to investigate the consequences of time-varying first and second moments, it is less convenient for addressing the consequences of time-varying higher moments. The reason for this is that the relationship between parameters of the typical Markov-switching model and higher moments is very non-linear, so that the effects of the different parameters on the optimal asset allocation are in some cases difficult to disentangle (see Guidolin and Timmermann, 2003).

In this paper, we propose a solution to the asset-allocation problem when the joint conditional distribution of returns is non-normal and time varying. Modelling asset returns requires rather general distributions that are able to incorporate volatility clustering, asymmetry, and fat-tailedness features found in empirical data. Choosing a distribution whose characteristics are well known generally implies that either the distribution is too restrictive to fit the data, or that it cannot be easily extended to the multivariate context. The proposed framework is both sufficiently general to capture the main statistical features of market returns and remains tractable even when several assets are included. Volatility clustering is captured using a multivariate GARCH model with dynamic conditional correlation (DCC) (Engle and Sheppard, 2002). Higher moment properties are modelled using a multivariate Student-t (Sk-t) distribution (Sahu, Dey, and Branco, 2001). In addition, it provides a convenient set-up for evaluating the opportunity cost of assuming joint normality of returns, while they actually strongly depart from normality. A clear advantage of this approach is that the properties of this distribution are rather well known, so that moments of returns can be computed analytically. Consequently, rather general asset-allocation problems can be dealt with, in real time, even with a large number of assets. The asset-allocation problem is solved using a high-order Taylor expansion of the utility function. The advantage of this approach is that, once the multivariate model has been estimated, portfolio weights can be computed in a very efficient way.

As an illustration of the proposed approach, we consider the conditional allocation of wealth between South-East Asian emerging markets. These markets are known to display very non-normal distributions that are inconsistent with the standard mean-variance criterion (Harvey, 1995). By using monthly observations covering the period from January 1975 to December 2003, we show that our model is able to capture the main characteristics of the market returns. In such a context, we find that the mean-variance criterion may result in an excessive risk taking and a significant opportunity cost, as compared to a strategy based on higher moments. For large levels of risk aversion, the opportunity cost may exceed 10 cents per dollar per year.

The outline of the paper is as follows. In section 2, we formulate our approach for modelling returns with a non-normal multivariate distribution. In section 3, we describe how the conditional asset-allocation problem can be solved in this context. We also demonstrate how portfolio moments can be computed in a very efficient way. In section 4, we present the data and discuss the result of the optimal asset allocation when investment opportunities are assumed to be constant. In section 5, we consider the case of time-varying investment opportunities and examine the consequences of using the mean-variance criterion under strong departure from normality. Section 6 concludes.

2 A multivariate model for returns

In this section, we describe a conditional set-up that incorporates most statistical features of stock market returns. First, it accounts for the well-known properties of volatility clustering (Engle, 1982) and time-varying correlations (Engle and Sheppard, 2002). Second, to capture both asymmetry and fat-tailedness often found in actual data, we adopt a multivariate skewed Student t (Sk-t) distribution.[2] This distribution has been developed, in a univariate context, by Hansen (1994) and studied by Jondeau and Rockinger (2003) and extended to the multivariate case by Sahu, Dey, and Branco (2001) and Bauwens and Laurent (2002). A nice feature of the distribution is that it is a straightforward extension of the normal and t distributions and that the associated parameters have a rather natural interpretation.

[2] The skewed Student t of Hansen (1994) differs from the density that is called a skewed Student t, in the statistics literature. Harvey and Siddique (1999) use the latter distribution.

Before describing the model, we adopt the following notations: Let $R_{i,t}$ be the rate of return of asset i from time $t-1$ to time t, in excess of the risk-free rate. Let $\mu_{i,t}$ be the expected excess return of asset i conditional on information available at time $t-1$. Then, $\varepsilon_{i,t} = R_{i,t} - \mu_{i,t}$ is the unexpected return of asset i. We define $R_t = (R_{1,t}, \cdots, R_{n,t})'$ the vector of asset returns, and $\varepsilon_t = (\varepsilon_{1,t}, \cdots, \varepsilon_{n,t})'$ the vector of unexpected returns. $\sigma_{ii,t}$ is the conditional variance of $R_{i,t}$. $\sigma_{ij,t}$ denotes the conditional covariance between $R_{i,t}$ and $R_{j,t}$. The conditional covariance matrix is denoted $\Sigma_t = \{\sigma_{ij,t}\}$.

2.1 The multivariate DCC model

We begin with a formal description of the dynamics of the first two moments of the return distribution. This model is very closely related to the DCC model proposed by Engle and Sheppard (2002). It is designed to capture both volatility clustering and persistence in correlation. The conditional mean of returns is described as an AR(1) dynamic to capture the possible first-order serial correlation of returns:

$$R_t = \mu + \varphi R_{t-1} + \Sigma_t^{1/2} z_t \tag{1}$$

where μ denotes an $(n,1)$ vector, whereas φ is a diagonal matrix. Each component of standardized residuals z_t has zero mean and unit variance. z_t is driven by a multivariate Sk-t distribution as described below. The conditional covariance matrix of returns is given by

$$\begin{aligned}
\Sigma_t &= D_t \Gamma_t D_t && (2)\\
D_t &= \begin{pmatrix} \sqrt{\sigma_{11,t}} & 0 & \cdots & 0 \\ 0 & \sqrt{\sigma_{22,t}} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \sqrt{\sigma_{nn,t}} \end{pmatrix} && (3)\\
\sigma_{ii,t} &= \omega_i + \beta_i \sigma_{ii,t-1} + \gamma_i \varepsilon_{i,t-1}^2 \qquad i = 1, \cdots, n && (4)
\end{aligned}$$

$$
\begin{aligned}
\Gamma_t &= (\mathrm{diag}\,(Q_t))^{-1} \cdot Q_t \cdot (\mathrm{diag}\,(Q_t))^{-1} && (5)\\
Q_t &= (1-\delta_1-\delta_2)\,\bar{Q} + \delta_1\left(u_{t-1}u'_{t-1}\right) + \delta_2 Q_{t-1} && (6)\\
\bar{Q} &= \begin{pmatrix} 1 & \rho_{12} & \cdots & \rho_{1n} \\ \rho_{12} & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \rho_{n-1,n} \\ \rho_{1n} & \cdots & \rho_{n-1,n} & 1 \end{pmatrix} && (7)
\end{aligned}
$$

where $\varepsilon_t = \Sigma^{1/2} z_t$ denotes the vector of unexpected returns and $u_t = D_t^{-1}\varepsilon_t$ denotes the vector of standardized unexpected returns. Equation (2) defines the covariance matrix Σ_t as a function of the univariate standard deviations contained in D_t and of the conditional correlation matrix Γ_t. Each conditional variance, $\sigma_{ii,t}$, is given by a standard GARCH model (Bollerslev, 1986).[3] The constraint $\gamma_i + \beta_i < 1$ guarantees stationarity for the variance process. The correlation matrix, Γ_t, is time-varying, following the specification of Engle and Sheppard (2002). See equations (5) and (6), where diag(Q_t) means the diagonal of Q_t. $\bar{Q}$ is the unconditional covariance matrix of u_t. We impose that $0 \leq \delta_1, \delta_2 \leq 1$ and $\delta_1 + \delta_2 \leq 1$. Once these restrictions on δ_1 and δ_2 are imposed, the conditional correlation matrix is guaranteed to be positive definite during the optimization.

2.2 The multivariate Sk-t distribution

Historically, the univariate Student t distribution was introduced as a way to model the distribution of the ratio of a normal variable and the root of a chi-square variable. Various extensions could be considered to extend the univariate t distribution to a multivariate one. One popular method is to assume that the χ^2 which appears in the definition of the t variable is the same for each component. It can be shown that this implies the same fat-tailedness for the different components, an undesirable feature for financial data. One may therefore pursue an alternative road by introducing

[3] We also investigated a model with an asymmetric effect of squared returns, as suggested by Glosten, Jagannathan, and Runkle (1993). For the data at hand, such an extension did not appear of relevance.

dependency of the various t variables via a covariance matrix. Obviously, the equality of the degree-of-freedom parameters can be tested in a second step.

We, therefore, now assume that z_t is drawn from the multivariate Sk-t distribution defined as

$$t(z_t|\nu_1,\cdots,\nu_n,\lambda_1,\cdots,\lambda_n)=\prod_{i=1}^{n} b_i c_i \left(1+\frac{\zeta_{i,t}^2}{\nu_i-2}\right)^{-\frac{\nu_i+1}{2}} \tag{8}$$

where

$$\zeta_{i,t}=\begin{cases}(b_i z_{i,t}+a_i)\,/\,(1-\lambda_i) & \text{if } z_{i,t}<-a_i/b_i,\\ (b_i z_{i,t}+a_i)\,/\,(1+\lambda_i) & \text{if } z_{i,t}\geq -a_i/b_i\end{cases}$$

and

$$a_i\equiv 4\lambda_i c_i\frac{\nu_i-2}{\nu_i-1},\qquad b_i^2\equiv 1+3\lambda_i^2-a_i^2.$$

Focusing on a given component i, we notice that λ_i introduces an asymmetry to the standard univariate t distribution. Parameters a_i and b_i are required to center and scale the asymmetric distribution so as to obtain an error term with zero mean and unit variance. Each component has a well-defined distribution if $\nu_i>2$ and $-1<\lambda_i<1$. If $\lambda_i=0$, the component reduces to the standard t distribution. If in addition $\nu_i<\infty$, the component will have positive excess kurtosis.

Defining the rth moment of the non-standardized random variable $z_{i,t}^*$ drawn from the univariate Sk-t distribution as $M_r=E\left[z_{i,t}^{*r}\right]$, it has been shown by Jondeau and Rockinger (2003) that

$$\begin{aligned}
M_{i,1} &= 4c_i\lambda_i\frac{\nu_i-2}{\nu_i-1}=a_i,\\
M_{i,2} &= 1+3\lambda_i^2=b_i^2+a_i^2,\\
M_{i,3} &= 16c_i\,\lambda_i(1+\lambda_i^2)\frac{(\nu_i-2)^2}{(\nu_i-1)(\nu_i-3)} \qquad \text{if } \nu_i>3, && (9)\\
M_{i,4} &= 3\frac{\nu_i-2}{\nu_i-4}(1+10\lambda_i^2+5\lambda_i^4) \qquad \text{if } \nu_i>4. && (10)
\end{aligned}$$

Consequently, the moments of standardized random variables, $z_{i,t}=\left(z_{i,t}^*-a_i\right)/b_i$, are obtained, using the notation $\mu_i^{(r)} = E\left[z_{i,t}^r\right]$, as

$$
\begin{aligned}
\mu_i^{(1)} &= 0, \\
\mu_i^{(2)} &= 1, \\
\mu_i^{(3)} &= \frac{M_{i,3} - 3a_i M_{i,2} + 2a_i^3}{b_i^3}, &(11) \\
\mu_i^{(4)} &= \frac{M_{i,4} - 4a_i M_{i,3} + 6a_i^2 M_{i,2} - 3a_i^4}{b_i^4}. &(12)
\end{aligned}
$$

Since $z_{i,t}$ has zero mean and unit variance, we directly deduce that the standardized skewness and kurtosis of $z_{i,t}$ are equal to $S\left[z_{i,t}\right] = \mu_i^{(3)}$ and $K\left[z_{i,t}\right] = \mu_i^{(4)}$ respectively. Therefore, skewness and kurtosis are directly related, in a non-linear way, to the asymmetry and degree-of-freedom parameters λ_i and ν_i.

2.3 The dynamic of higher moments

We investigate two alternative models for the dynamics of higher moments. In the first case, we allow for time-varying expected returns and covariances, while keeping higher moments constant. In the second case, also higher moments may vary over time. Such time variability in higher moments has also been analyzed more directly in Hansen (1994), Harvey and Siddique (1999), and Jondeau and Rockinger (2003). To model the parameters λ_i and ν_i we follow recent contributions. Considering the asymmetry parameter, Ang and Bekaert (2002) show that bearish and bullish markets tend to be persistent, suggesting that there may be some clustering in this parameter. Second, the lower the degree-of-freedom parameter, the higher the probability of extreme events in market returns. As argued by Das and Uppal (2004), such extreme events are not likely to be persistent. As a consequence, after a large shock, we expect a sharp decrease in kurtosis, so that the degree-of-freedom parameter will be negatively related to the size of shocks. Since $\nu_{i,t}$ may be close to infinity when the distribution is close to normality, for convenience we actually model the dynamics of $1/\nu_{i,t}$. After investigating several models for the degree-of-freedom and asymmetry parameters (following the approach proposed by Jondeau and Rockinger, 2003), we finally adopt the following model:

$$\frac{1}{\tilde{\nu}_{i,t}} = \frac{1}{\tilde{\nu}_i} + b_1 \sum_{i=1}^{p} \omega_i \left|\varepsilon_{t-i}\right| \tag{13}$$

$$\tilde{\lambda}_{i,t} = \tilde{\lambda}_i + b_2 \sum_{i=1}^{p} \omega_i \varepsilon_{t-i} \tag{14}$$

where unrestricted parameters $\tilde{\nu}_{i,t}$ and $\tilde{\lambda}_{i,t}$ are mapped into the authorized domain $[2;\infty[\, \times [-1;1]$ via a logistic map, and $\omega_i = 1 - i/p$ is the weight on lag i. Two main features are worth emphasizing. First, the dynamic of the degree-of-freedom parameter $\nu_{i,t}$ depends on the absolute value of residuals, because it translates in the heaviness of the distribution's tails regardless of the sign of shocks over the recent period. Since the degree-of-freedom parameter is very large in the case of very small recent shocks, we set $\tilde{\nu}_i$ to a large value to reflects normality. In contrast, the dynamic of the asymmetry parameter naturally depends on signed residuals, because $\lambda_{i,t}$ is likely to reflect the sign and size of shocks over the recent period. Second, we introduce some lags in the function of unexpected returns, so that equations (13) and (14) look like ARCH(p) models. To avoid the curse to dimensionality, we adopt the same strategy as in the DCC model of Engle and Sheppard (2002) and impose the same parameters b_1 and b_2 for all markets. We checked that estimating different parameters b_1 and b_2 would not alter the results significantly, but would increase parameter uncertainty.

We explored several values for the number of lags p (including different values for $\tilde{\nu}_{i,t}$ and $\tilde{\lambda}_{i,t}$) and retained, based on the usual information criteria, the value $p = 3$. This result suggests that for monthly data, the persistence in higher moments is in fact rather limited. This is confirmed by the inability of models with an autoregressive component to capture the dynamics of these moments.

2.4 Estimation

In model (1) to (7), the set of parameters pertaining to location and dispersion is denoted $\theta = (\mu_i,\ \varphi_i,\ \omega_i,\ \beta_i,\ \gamma_i\ (i = 1, \cdots, n),\ \delta_1,\ \delta_2,\ \rho_{jk}\ (1 \leq j < k \leq n))$. We also define ξ the set of parameters pertaining to the shape of the distribution. When innovations are drawn from a multivariate Sk-t distribution with constant shape parameters, we have $\xi = (\lambda_1, \cdots, \lambda_n, \nu_1, \cdots, \nu_n)$. When the shape parameters are time varying, we have $\xi = (\lambda_1, \cdots, \lambda_n, b_1, b_2)$.

The sample log-likelihood function of the multivariate DCC model with Sk-t distibution is then

$$\begin{aligned} \ln L(R_1, \cdots, R_T|\theta, \xi) &= \sum_{t=1}^{T} \ln\left[t\left(\Sigma_t(\theta)^{-1/2}(R_t - \mu_t(\theta))|\xi\right)\right] \\ &= \sum_{t=1}^{T} \ln\left[t(z_t|\xi)\right] - \frac{1}{2}\sum_{t=1}^{T} \ln|\Sigma_t(\theta)| \end{aligned} \tag{15}$$

where $t(z_t|\xi)$ is defined in equation (8). Maximizing expression (15) with respect to parameter vectors θ and ξ yields the maximum-likelihood (ML) estimates.

For large dimensional systems, the estimation can be significantly speed up by performing the estimation in two steps. In the first step, the quasi-ML estimation of the univariate conditional mean and variance equations is obtained assuming normality. The unconditional correlation matrix of standardized residuals is then used to estimate the matrix $\bar{Q}$. In the second step, the parameters pertaining to the dynamics of correlation (δ_1 and δ_2) and to the shape of the distribution (ξ) are estimated simultaneously. In the empirical application, we considered full ML estimation as well as the two-step procedure, and we found that both provide very similar results.[4]

3 The asset-allocation problem

We now turn to the conditional asset-allocation problem. When returns strongly depart from normality, the standard mean-variance criterion may be inappropriate to select the optimal portfolio. In such a case, as put forward by Harvey and Siddique (2000) and Dittmar (2002), incorporating the effect of higher moments on the expected utility of investors would improve the allocation of wealth. We use an approximation of expected utility up to the fourth moment to evaluate the optimal allocation of wealth. Below, we describe how the expected utility can be approximated

[4]The estimation is performed using the Constrained Maximum Likelihood procedure of GAUSS. The Gradient and the Hessian are computed numerically. The optimization is done using the BHHH algorithm and the covariance matrix is corrected for serial correlation and heteroskedasticity.

by a four-moment asset-allocation problem. We also indicate how this problem can be solved efficiently in practice.

3.1 The general investor's problem

We consider an investor who allocates her portfolio to maximize the expected utility $U(W_{t+1})$ over the end-of-period wealth W_{t+1}.[5] The initial wealth is arbitrarily set equal to one. There are n risky assets with return vector $R_{t+1} = (R_{1,t+1}, \cdots, R_{n,t+1})'$ and joint cumulative distribution function $F(R_{1,t+1}, \cdots, R_{n,t+1})$ for the period between t and $t+1$. End-of-period wealth is given by $W_{t+1} = (1 + r_{p,t+1})$, where $r_{p,t+1} = \alpha_t' R_{t+1}$ is the portfolio return and $\alpha_t = (\alpha_{1,t}, \cdots, \alpha_{n,t})'$ is the vector of fractions of wealth allocated to the various risky assets. We assume that the investor does not have access to a riskless asset, implying that the portfolio weights sum to one ($\sum_{i=1}^{n} \alpha_{i,t} = 1$).[6] In addition, portfolio weights are constrained to be positive, so that short-selling is not allowed. We also assume that the investor has forecasts for the expected mean vector μ_{t+1}, the covariance matrix Σ_{t+1}, the co-skewness matrix $M_{3,t+1}$, and the co-kurtosis matrix $M_{4,t+1}$.

Formally, the asset-allocation problem is

$$\begin{aligned} \max_{\{\alpha_t\}} \quad & E_t\left[U\left(W_{t+1}\right)\right] = E_t\left[U\left(1 + \alpha_t' R_{t+1}\right)\right] \\ \text{s.t.} \quad & \sum_{i=1}^{n} \alpha_{i,t} = 1. \end{aligned} \tag{16}$$

The optimal portfolio weights at time t are obtained by maximizing the scaled ex-

[5] We do not consider a multi-period investment problem. Brandt (1999) as well as Ang and Bekaert (2002) have shown that even if portfolio weights may be slightly affected by the investment horizon, the opportunity cost of a myopic strategy is negligible. This result suggests that hedging against unfavorable changes in the investment set does not result in any significant gain.

[6] The reason for this is that, when the investor is allowed to invest in a riskless asset, the weight affected to the riskless asset increases sharply with the degree of risk aversion. In contrast, for low degrees of risk aversion, the investor puts the emphasis on the expected return of the portfolio, implying that second and higher moments are not taken into account. In consequence, we focus on the risky part of the portfolio to evaluate the effect of the distribution assumption on the portfolio weights.

pected utility

$$\alpha_t^* = \arg\max_{\{\alpha_t\}} \quad E_t\left[U\left(1+\alpha_t' R_{t+1}\right)\right].$$

For non-normal returns, the solution to problem (16), generally does not have a closed-form solution, and numerical techniques must be used.[7] The problem for non-normal distributions (in particular for distributions with fat tails and asymmetry) is that the required number of quadrature points increases exponentially with the number of assets. Therefore, solving the optimization problem using numerical integration becomes intractable for more than two or three assets.

3.2 The four-moment investor's problem

We now investigate how higher moments and co-moments affect the optimal allocation. Early work on introducing the effect of higher moments are by Samuelson (1970) and Kraus and Litzenberger (1976). More recent work is by Ang and Bekaert (2002) or Guidolin and Timmermann (2003). Several difficulties appear in introducing higher moments in the portfolio-selection problem. While it may appear as a natural extension to the standard mean-variance criterion, a first issue is how higher moments actually affect the utility of investors. Theoretical research suggests that investors prefer high values of odd moments, and low values of even moments (see, for instance, Scott and Horvarth, 1980, or Pratt and Zeckhauser, 1987). Hence investors prefer positive skewness, because they prefer positive extreme values and dislike negative extreme value. In addition, they avoid kurtosis, because it is a measure of dispersion and, therefore, of uncertainty.

Since we are primarily interested in the effect of higher moments on the asset allocation, a convenient approach consists in approximating the utility function using a Taylor series expansion around the current value of the portfolio return. In this context, the utility function is re-written as

$$U\left(W_{t+1}\right) = \sum_{k=0}^{\infty} \frac{U^{(k)}\left(W_t\right)\left(W_{t+1}-W_t\right)^k}{k!}$$

[7] Tauchen and Hussey (1991) suggest using quadrature rules to solve this problem. This approach has been applied to normal iid returns (see Campbell and Viceira, 1999) or to Markov-switching conditionally normal returns (Ang and Bekaert, 2002).

Focusing on terms up to the fourth one, we obtain the following approximation of the expected utility

$$
\begin{aligned}
E_t\left[U\left(W_{t+1}\right)\right] \approx\ & U\left(W_t\right)+U^{(1)}\left(W_t\right) E_t\left[W_{t+1}-W_t\right]+\frac{1}{2} U^{(2)}\left(W_t\right) E_t\left[W_{t+1}-W_t\right]^2 \\
& +\frac{1}{3!} U^{(3)}\left(W_t\right) E_t\left[W_{t+1}-W_t\right]^3+\frac{1}{4!} U^{(4)}\left(W_t\right) E_t\left[W_{t+1}-W_t\right]^4
\end{aligned}
$$

up to some remainder in the Taylor expansion of U. Since $W_{t+1}-W_t=\alpha_t^{\prime} R_{t+1}= r_{p,t+1}$, the expected utility is approximated by the following preference function:

$$
\begin{aligned}
E_t\left[U\left(W_{t+1}\right)\right] \approx\ & U\left(W_t\right)+U^{(1)}\left(W_t\right) E_t\left[r_{p,t+1}\right]+\frac{1}{2} U^{(2)}\left(W_t\right) E_t\left[r_{p,t+1}^2\right] \\
& +\frac{1}{3!} U^{(3)}\left(W_t\right) E_t\left[r_{p,t+1}^3\right]+\frac{1}{4!} U^{(4)}\left(W_t\right) E_t\left[r_{p,t+1}^4\right],
\end{aligned}
$$

where the non-central moments $m_{p,t+1}^i=E_t\left[\left(r_{p,t+1}\right)^i\right]$ are functions of the expected return and the central moments

$$
\begin{aligned}
m_{p,t+1} &= \mu_{p,t+1} \\
m_{p,t+1}^2 &= \sigma_{p,t+1}^2+\mu_{p,t+1}^2 \\
m_{p,t+1}^3 &= s_{p,t+1}^3+3 m_{p,t+1}^2 \mu_{p,t+1}-2 \mu_{p,t+1}^3=s_{p,t+1}^3+3 \sigma_{p,t+1}^2 \mu_{p,t+1}+\mu_{p,t+1}^3 \\
m_{p,t+1}^4 &= \kappa_{p,t+1}^4+4 m_{p,t+1}^3 \mu_{p,t+1}-6 m_{p,t+1}^2 \mu_{p,t+1}^2+3 \mu_{p,t+1}^4 \\
&= \kappa_{p,t+1}^4+4 s_{p,t+1}^3 \mu_{p,t+1}+6 \sigma_{p,t+1}^2 \mu_{p,t+1}^2+\mu_{p,t+1}^4
\end{aligned}
$$

and where $\sigma_{p,t+1}^2$, $s_{p,t+1}^3$ and $\kappa_{p,t+1}^4$ stand for central moments $E_t\left[r_{p,t+1}-\mu_{p,t+1}\right]^j$ for $j=2$, 3 and 4 respectively.[8] At this stage, it is possible to consider various specifications of the utility function and to investigate the resulting preferences.

Example 1 (CARA) Consider an investor with constant absolute risk aversion (CARA) utility function:

$$
U\left(W_{t+1}\right)=-\exp \left(-A_0 W_{t+1}\right),
$$

[8] In finance literature, the terms $s_{p,t+1}^3$ and $\kappa_{p,t+1}^4$ are sometimes referred to as skewness and kurtosis. These definitions differ from the statistical definitions as standardized central higher moments $E\left[\left(\left(r_{p,t+1}-\mu_{p,t+1}\right) / \sigma_{p,t+1}\right)^j\right]$ for $j=3$, 4.

where A_0 measures the investor's constant absolute risk aversion. In this case, the expected utility is approximated by

$$E_t\left[U\left(W_{t+1}\right)\right] \approx -\exp\left(-A_0\right)\left[1 - A_0\mu_{p,t+1} + \frac{A_0^2}{2}m_{p,t+1}^2 - \frac{A_0^3}{3!}m_{p,t+1}^3 + \frac{A_0^4}{4!}m_{p,t+1}^4\right]. \tag{17}$$

Example 2 (CRRA) Consider now an investor with constant relative risk aversion (CRRA) utility function:

$$U\left(W_{t+1}\right) = \begin{cases} \frac{W_{t+1}^{1-A}}{1-A} & \text{if } A > 1 \\ \ln\left(W_{t+1}\right) & \text{if } A = 1, \end{cases}$$

where A measures the investor's constant relative risk aversion. For such a utility function, the approximation becomes

$$E_t\left[U\left(W_{t+1}\right)\right] \approx \frac{1}{1-A} + \mu_{p,t+1} - \frac{A}{2}m_{p,t+1}^2 + \frac{A\left(A+1\right)}{3!}m_{p,t+1}^3 - \frac{A\left(A+1\right)\left(A+2\right)}{4!}m_{p,t+1}^4. \tag{18}$$

For the two examples above, we obtain an unambiguous effect of skewness and kurtosis on the approximated expected utility function.[9] Expected utility decreases with higher negative skewness (i.e. left-skewed distributions) and higher kurtosis (i.e. fatter-tailed distributions). These effects are consistent with theoretical arguments developed by Scott and Horvath (1980).

3.3 Solving the four-moment problem

A reason for adopting a multivariate Sk-t distribution is that an analytic expression for portfolio conditional moments is easily obtained in a few steps. First, the first four moments of a univariate Sk-t distributed random variable $z_{i,t+1}$ are given by the expressions reported in section 2.2. Second, since the vector of unexpected returns is defined as $\varepsilon_{t+1} = R_{t+1} - \mu_{t+1} = \Sigma_{t+1}^{1/2} z_{t+1}$, its first four moments can be computed using matrix calculus, instead of numerical integration. Indeed, we obviously have

[9] The derivatives of equations (17) and (18) with respect to $\kappa_{p,t+1}^4$ are always positive. The derivatives with respect to $s_{p,t+1}^3$ depend on the sign of $\left(1 - A\mu_{p,t+1}\right)$ and $\left(1 - \left(A+2\right)\mu_{p,t+1}\right)$ respectively. They are therefore positive for any realistic value of the degree of risk aversion and expected returns.

$E_t\left[\varepsilon_{t+1}\right]=0$ and $V_t\left[\varepsilon_{t+1}\right]=\Sigma_{t+1}$. In addition, We denote $\Sigma_{t+1}^{1/2}=\left(\omega_{ij,t+1}\right)_{i,j=1,\cdots,n}$ the Choleski decomposition of the covariance matrix of returns. Then, using tensor notations and denoting $\otimes$ the Kronecker product, we define the (n,n^2) co-skewness matrix as

$$M_{3,t+1}=E_t\left[\left(R_{t+1}-\mu_{t+1}\right)\left(R_{t+1}-\mu_{t+1}\right)'\otimes\left(R_{t+1}-\mu_{t+1}\right)'\right]=\left\{s_{ijk,t+1}\right\},$$

and the (n,n^3) co-kurtosis matrix as

$$M_{4,t+1}=E_t\left[\left(R_{t+1}-\mu_{t+1}\right)\left(R_{t+1}-\mu_{t+1}\right)'\otimes\left(R_{t+1}-\mu_{t+1}\right)'\otimes\left(R_{t+1}-\mu_{t+1}\right)'\right]=\left\{\kappa_{ijkl,t+1}\right\}.$$

The (i,j,k) component of the third central moment is:

$$s_{ijk,t+1}=\sum_{r=1}^{n}\omega_{ir,t+1}\omega_{jr,t+1}\omega_{kr,t+1}\mu_{r,t+1}^{(3)}.$$

and the (i,j,k,l) component of the fourth central moment is:

$$\begin{aligned}\kappa_{ijkl,t+1} &= \sum_{r=1}^{n}\omega_{ir,t+1}\omega_{jr,t+1}\omega_{kr,t+1}\omega_{lr,t+1}\mu_{r,t+1}^{(4)}\\ &+\sum_{r=1}^{n}\sum_{\substack{s=1\\ s\neq r}}^{n}\left(\begin{array}{c}\omega_{ir,t+1}\omega_{jr,t+1}\omega_{ks,t+1}\omega_{ls,t+1}+\omega_{ir,t+1}\omega_{js,t+1}\omega_{kr,t+1}\omega_{ls,t+1}\\ +\omega_{ir,t+1}\omega_{js,t+1}\omega_{ks,t+1}\omega_{lr,t+1}\end{array}\right),\end{aligned}$$

where $\mu^{(3)}$ and $\mu^{(4)}$ are defined in equations (11) and (12) respectively.

These analytical expressions of higher moments are quite cumbersome to derive, yet their numerical computation is very fast.[10] As it clearly appears from the expressions above, time-variability of co-skewness and co-kurtosis between unexpected returns has two sources. On one hand, the covariance matrix between unexpected returns is time-varying, so that the $\omega_{ij,t}$ elements are time-varying. On the other hand, individual skewness and kurtosis of innovations are themselves time-varying.

The last step consists in the computation of the portfolio moments. For a given portfolio weight vector α_t, the conditional expected return, variance, skewness, and

[10] Using these notations, the matrices of co-skewness and co-kurtosis may be conveniently represented as bi-dimensional matrices.

kurtosis of the portfolio are given by:

$$\mu_{p,t+1} = \alpha_t' \mu_{t+1}, \tag{19}$$

$$\sigma_{p,t+1}^2 = \alpha_t' \Sigma_{t+1} \alpha_t, \tag{20}$$

$$s_{p,t+1}^3 = \alpha_t' M_{3,t+1} \left(\alpha_t \otimes \alpha_t \right), \tag{21}$$

$$\kappa_{p,t+1}^4 = \alpha_t' M_{4,t+1} \left(\alpha_t \otimes \alpha_t \otimes \alpha_t \right). \tag{22}$$

Derivatives of portfolio moments with respect to α_t are therefore very easy to compute:

$$\begin{aligned}
\frac{\partial \mu_{p,t+1}}{\partial \alpha_t} &= \mu_{t+1}, \\
\frac{\partial \sigma_{p,t+1}^2}{\partial \alpha_t} &= 2M_{2,t+1}\alpha_t, \\
\frac{\partial s_{p,t+1}^3}{\partial \alpha_t} &= 3M_{3,t+1} \left(\alpha_t \otimes \alpha_t \right), \\
\frac{\partial \kappa_{p,t+1}^4}{\partial \alpha_t} &= 4M_{4,t+1} \left(\alpha_t \otimes \alpha_t \otimes \alpha_t \right).
\end{aligned}$$

First-order conditions to the maximization of equation (18) are given by

$$\mu_{t+1} - \delta_1 \left(\alpha_t \right) \left[M_{2,t+1}\alpha_t \right] + \delta_2 \left(\alpha_t \right) \left[M_{3,t+1} \left(\alpha_t \otimes \alpha_t \right) \right] - \delta_3 \left(\alpha_t \right) \left[M_{4,t+1} \left(\alpha_t \otimes \alpha_t \otimes \alpha_t \right) \right] = 0 \tag{23}$$

where δ_1, δ_2, and δ_3 are non-linear scalar functions of α_t. The n equations (23) can be easily solved numerically, using a standard optimization package. The difficulty to solve this problem is of a lesser order as compared to problems involving numerical integration of the utility function. This approach also provides an alternative way of solving the asset allocation problem to the PGP approach developed by Lai (1991) and Chunhachinda et al. (1997). The main advantage of the approach adopted here is that weights attributed to the various portfolio moments in equation (23) are selected on the basis of the utility function, while they are arbitrarily chosen in the PGP approach. Solving equation (23) also provides an alternative to the rather time-consuming approach based on maximizing the expected utility numerically. Here, a very accurate solution is obtained in just a few seconds, even in the case of a large number of assets.

It is also worth noting that we have described how to solve the asset-allocation problem when the first four moments are to be incorporated. As well, we may consider as special cases the situations where only mean and variance or mean, variance, and skewness are assumed to be important.

4 Unconditional asset allocation

In the following sections, we illustrate how departure from normality can affect the asset allocation. We begin with a version of the model in which all moments are assumed to be constant over time. This approach allows to focus only on the specific effects of higher moments on the unconditional asset allocation.

4.1 The model with constant moments

We illustrate the asset-allocation problem by considering as assets a set of emerging financial markets. An abundant literature has shown that market returns in emerging markets strongly depart from non-normality (Harvey, 1995, Bekaert and Harvey, 1997, Hwang and Satchell, 1999, or Jondeau and Rockinger, 2003).[11] We thus consider monthly returns for five dollar-denominated emerging-market indices (Hong Kong, Singapore, South Korea, Taiwan, Thailand), for which data is available for a long period of time. To be consistent with the actual asset-allocation practice, we use monthly returns. The data covers the period from January 1975 to December 2003, for a total of 343 monthly observations.[12]

Let $R_{i,t}$, $t = 1, \cdots, T$, denote the monthly (log) return of market i at date t. In order to investigate the incremental effects of relaxing various assumptions, we start with a model with constant means, covariance, co-skewness and co-kurtosis matrices and Sk-t distributed innovations. Table 1 reports parameter estimates (Panel A) and summary statistics for residuals (Panel B). We also report several Wald test statistics corresponding to various restrictions on parameter values. Parameter estimates reported in Table 1 indicate that the average monthly return is positive, ranging

[11] Other examples of highly non-normal asset returns are given by hedge funds or IT stocks.

[12] The October 1987 crash has severely affected Asian stock markets. Singapore and Thailand experienced a fall by more than 50%. We dealt with this very extreme event by averaging returns over the periods just before and just after the crash. This admittedly crude correction avoids altering the statistical properties of stock-market returns too heavily.

from 0.58% to 1.37% for the five stock-market returns. While the null hypothesis that expected returns are jointly equal to zero is rejected, we do not reject the null that expected returns are equal across countries. Standard deviations are ranging from 7.02% for Singapore to 10.77% for Thailand. We reject the null hypothesis that variances are equal across countries. We observe that the degree-of-freedom parameter is rather high for South Korea, but very low for Taiwan and Thailand. The degree of freedom for Thailand is even lower than 4 ($\nu_5 = 2.98$), so that kurtosis (and even skewness) is likely to be infinite. We strongly reject the null that degree-of-freedom parameters are infinite, so that joint normality does not hold. In contrast, we do not reject the null that degree-of-freedom parameters are equal across markets ($\nu_1 = \cdots = \nu_n$). Finally, the estimate of the asymmetry parameter indicates that the distribution of returns leans leftwards for Hong Kong and Singapore, and rightwards for other countries. The asymmetry parameter turns out to be significant for South Korea only. For this market, the asymmetry parameter is strongly positive. The test statistics for the null that asymmetry parameters are jointly equal to zero ($\lambda_1 = \cdots = \lambda_n = 0$) rejects the null at the 4.5% significance level only.

[Table 1 somewhere here]

To test the ability of the model to fit the data, we perform the goodness-of-fit (GoF) test procedure suggested by Diebold, Gunther, and Tay (1998). These authors argue that, if the marginal distributions are correctly specified, the margins of the distribution should be iid Uniform$(0, 1)$. This test is performed in two steps. First, we evaluate whether the margins $u_{i,t} = \tilde{F}(z_{i,t})$ are serially correlated, where $\tilde{F}(z_{i,t})$ denotes the marginal cumulative distribution function of $z_{i,t}$. For this purpose, we examine the serial correlations of $(u_{i,t} - \bar{u}_i)^k$, for $k = 1, \cdots, 4$, thus regressing $(u_{i,t} - \bar{u}_i)^k$ on q lags of the variable.[13] The LM test statistic is defined as $LM = (T - q) R^2$, where R^2 is the coefficient of determination of the regression. Under the null, LM is asymptotically distributed as a $\chi^2(q)$. The table reveals that the LM test for no serial correlation of margins does not reject the null hypothesis of iid process

[13] Zero correlation is equivalent to independence, only under gaussianity. The correlogram is, therefore, only hinting at possible independence.

for the first and third moments for most countries. Yet, the null hypothesis is strongly rejected for the second and fourth moments in all instances. This result indicates that time dependency mainly comes from odd moments rather than from even moments. Interestingly, this result is rather well established in the empirical literature for the first and second moments. But it seems to also hold for higher moments, since the third moment does not appear to be severely autocorrelated, while the fourth moment is clearly time-varying.

Second, we test the null hypothesis that $u_{i,t}$ is distributed as an Uniform$(0,1)$. Hence, we cut the empirical and theoretical distributions into N bins and test whether the two distributions significantly differ on each bin. Table 1 reports the GoF test statistic with p-values computed with $N-1$ degrees of freedom.[14] When we consider the case $N = 20$ bins, we do not reject the null hypothesis that the theoretical distribution provides a good fit of the empirical distribution for any return series, at the 5% level.

4.2 Optimal unconditional asset allocation

We now investigate the consequences of the choice of a forecasting model on the asset-allocation performances. In particular, we evaluate the opportunity cost of assuming normality (and hence of using the mean-variance criterion) while returns are actually non-normal. For this purpose, we begin with a focus on the asset allocation based on constant conditional moments. This first step allows to describe in a simple framework the main mechanisms at work. Then, we turn to the asset allocation based on the models that incorporate time-varying conditional moments.

4.2.1 Measure of opportunity cost

As it may be expected, to evaluate the opportunity cost of assuming normality, we have to take into account the consequences of this assumption on the whole distribution of the portfolio return. In particular, it is likely to induce an excessive exposure

[14] As shown by Vlaar and Palm (1993), under the null, the correct distribution of the GoF test statistic is bounded between a $\chi^2(N-1)$ and a $\chi^2(N-K-1)$, where K is the number of estimated parameters.

to extreme events (since the asset allocation under normality does not incorporate any information on this part of the distribution). To measure the cost of assuming normality, we therefore concentrate on two particular measures: First, we acknowledge that the variance is not a relevant measure of risk under non-normality and we adopt the so-called *distributional measure of risk* proposed by Berényi (2001). When the return distribution is approximated by a fourth-order expansion, the measure is defined as

$$\begin{aligned} R\left(\mu_{p,t+1}, \sigma^2_{p,t+1}, s^3_{p,t+1}, \kappa^4_{p,t+1}\right) &= -\frac{1}{2} U^{(2)}\left(W_t\right) \left(m^2_{p,t+1} - \frac{2}{3} \frac{U^{(3)}\left(W_t\right)}{U^{(2)}\left(W_t\right)} m^3_{p,t+1} \right. \\ & \left. -\frac{1}{6} \frac{U^{(4)}\left(W_t\right)}{U^{(2)}\left(W_t\right)} m^4_{p,t+1} \right) \end{aligned} \tag{24}$$

where the term in parentheses is named "variance equivalent", because it reduces to variance under normality, but incorporates the additional effects on risk of skewness and kurtosis under more general distributions. This measure, therefore, penalizes higher variance and kurtosis, and negative skewness. In the context of the CRRA utility function, it simplifies to

$$R\left(\mu_{p,t+1}, \sigma^2_{p,t+1}, s^3_{p,t+1}, \kappa^4_{p,t+1}\right) = -\frac{A}{2} \left(m^2_{p,t+1} + \frac{A+1}{3} m^3_{p,t+1} - \frac{(A+1)(A+2)}{12} m^4_{p,t+1} \right).$$

The second tool we use to measure the cost of assuming normality is the opportunity cost (or forecast premium). If we denote $r^*_{p,t+1}$ the optimal portfolio return obtained under the true distribution, and $\hat{r}_{p,t+1}$ the optimal portfolio return obtained using the strategy based on normality, then the opportunity cost, denoted θ_{t+1}, is defined as the return that needs to be added to the portfolio return obtained under normality, so that the investor becomes indifferent to the true distribution. Formally, we have

$$E_t\left[U\left(1 + \hat{r}_{p,t+1} + \theta_{t+1}\right)\right] = E_t\left[U\left(1 + r^*_{p,t+1}\right)\right]. \tag{25}$$

The reported premium θ_{t+1} may be obtained by solving equation (25) numerically. In the context of a fourth-order Taylor approximation with CRRA utility function, expression (25) can be further simplified as

$$\begin{aligned} \theta_{t+1} &= \left(\mu^*_{p,t+1} - \hat{\mu}_{p,t+1}\right) - \frac{A}{2}\left(m^{*2}_{p,t+1} - \hat{m}^2_{p,t+1}\right) + \frac{A(A+1)}{3!}\left(m^{*3}_{p,t+1} - \hat{m}^3_{p,t+1}\right) \\ & -\frac{A(A+1)(A+2)}{4!}\left(m^{*4}_{p,t+1} - \hat{m}^4_{p,t+1}\right). \end{aligned} \tag{26}$$

This expression indicates that the opportunity cost is directly related to the moments of the portfolio return. The opportunity cost may be interpreted as the premium an investor is willing to pay to use the true data generating process, rather than an inadequate forecasting model. In the following, we evaluate the cost of using a strategy based on normality (the mean-variance criterion) when the true distribution is the Sk-t distribution. Since the investor likes odd moments and dislikes even moments, she will be willing to pay to use a strategy that decreases variance and kurtosis and increases expected return and skewness. In addition, the larger the degree of risk aversion, the higher the premium for using the correct model.

4.2.2 Optimal allocation

In the framework with constant moments, we focus on three alternative distributional assumptions: normal distribution, t distribution, and Sk-t distribution. In the first case, we only consider information on the first two moments; with the t distribution, we also incorporate information on the fatness of the distribution's tails; finally, with the Sk-t distribution, we introduce the additional effect of the asymmetry of the distribution. In this subsection, we assume that the Sk-t distribution provides a correct approximation of the true data generating process (DGP) of market returns.

Table 2 reports results for optimally selected portfolios for several values of the risk aversion parameter A. The parameter A ranges between 2 (very low risk aversion) and 20 (very large risk aversion), covering most values found in the literature. For the three alternative distributional assumptions, we report optimal portfolio weights, portfolio moments, the risk measure (24) and the opportunity cost (25) of using a given sub-optimal forecasting model. The cost of using the normal distribution and the t distribution is evaluated with respect to the Sk-t distribution.

[Table 2 somewhere here]

Several points are worth noting. First, the weights of the alternative markets in the optimal portfolio are broadly consistent with the intuition drawn from the estimated conditional moments reported in Table 1. When the investor has a very low risk aversion ($A = 2$), large weights are attributed to the two markets with the largest expected return (Hong Kong and Taiwan). In case of a large risk aversion ($A = 20$), the portfolio is much more balanced, including markets with low expected return but also with low variance (Singapore and South Korea). The moments of the portfolio return are as expected: when the degree of risk aversion increases, the optimal-portfolio expected return decreases, the variance decreases, the skewness increases and the kurtosis decreases.

Second, as expected, when the investor has some information on the fatness of the distribution's tails, a feature we model as a switch from the normal to the t distribution, significant changes in the portfolio weights occur. The investor is now reluctant to invest in markets that are characterized by a large kurtosis (Thailand and, to a lesser extent, Taiwan).[15] In contrast, markets with a low kurtosis experience a large increase in their weight (see, e.g., the weight of South Korea). Notice that the kurtosis of the portfolio return does not necessarily decrease when the investor uses a t distribution. Indeed, for low risk aversion, the investor further increases the weight of markets with large expected return, even if it implies a larger exposure to extreme events. Finally, when the asymmetry in the distribution is correctly captured (moving to the Sk-t distribution), the weight of markets with positive skewness increases and the skewness of the portfolio return increases in turn.

Typical changes in the portfolio weights are reported in Figure 1 for Hong Kong and South Korea. The figure displays how weights are affected by the level of risk aversion A and by the choice of the distribution of returns. First, Hong Kong is characterized by a large expected return and a rather large variance. Assuming normality, we notice that as the parameter of risk aversion increases, the weight associated with Hong Kong decreases. But Hong Kong also has a negative skewness, hence, if the investor focuses not only on mean and variance but on higher moments, for low levels of risk aversion, this asset will be avoided due to the negative skewness.

[15] For Thailand, the estimated degree-of-freedom parameter was found to be equal to 3.002 (for the t distribution) or 2.982 (for the Sk-t distribution), so that the fourth (and perhaps the third) moment of the distribution does not exist. For the asset allocation experiment, we assume $\nu = 4.01$, which translates into a very large kurtosis. This is the reason why the weight pertaining to Thailand decreases dramatically when the forecasting model measures tail fatness.

Hence the weight allocated to this country will be systematically decreased. We notice the opposite changes for South Korea, that is characterized by a low expected return, a low variance and a large positive skewness. Assuming normality, we notice relatively little interest to invest in this country for small values of the parameter of risk aversion. As risk aversion increases, due to the relatively low volatility, the weight allocated to this country increases. Now, if the investor also considers higher moments, given the high skewness of this country's stock market, she will increase the weight to this country.

[**Figure 1 somewhere here**]

Considering risk measures, the table reveals that the use of the standard t distribution does not necessarily result in a reduction of the distributional risk measure. Indeed, the investor alters her portfolio according to the information on the fatness of the tails, but this change may turn out to be unfortunate because of a decrease in the skewness of the portfolio. This case occurs for low risk aversion: the weight of Hong Kong increases because of a (relatively) low kurtosis, while it is also characterized by the most negative skewness. Now, for larger risk aversions, the risk measure systematically decreases when the investor uses a more sophisticated forecasting model. The reason is that the investor uses the information on higher moments to reduce the exposure of the optimal portfolio to extreme risk (and not to increase the expected return).

Finally, the opportunity cost combines the effect of the forecasting model both on expected return and on distributional risk. It, therefore, measures the global effect of the forecasting model on the portfolio performances. The table indicates that the opportunity cost is systematically positive, so that the investor is always willing to pay a premium for using the correct forecasting model. For a very low risk aversion, the opportunity cost is as expected negligible. The reason is that the investor only focuses on the expected return and therefore selects the same allocation whatever the estimates of the higher moments. In contrast, for medium to large risk aversion, the cost is relatively large: for $A = 20$, we obtain $\theta_1 = 0.1$, so that the investor is willing to pay 1.2% per year to use the strategy based on the Sk-t distribution instead of the mean-variance criterion. When using the Sk-t model, the investor obtains a slightly lower expected return (10.74% instead of 10.14% per year), but also a lower variance, a larger skewness and a much lower kurtosis. Consequently, the tails of the portfolio return are much thinner and the investor is less likely to experience an extreme event. This translates in the reduction of the risk measure (from 4.43 to 4.28).

5 Conditional asset allocation

We now consider the asset-allocation problem when the parameters characterizing investment opportunities vary over time. We, thus, focus on the model with time-varying expected returns and covariance matrix. Higher moments are supposed to be either constant or time varying.

5.1 The model with time-varying moments

As described in Section 2, the conditional mean is assumed to follow an AR(1) model, while the conditional variances are given by a GARCH(1,1) model. Conditional correlations vary over time according to a DCC model. The dynamics of degree-of-freedom parameters and asymmetry parameters are given by equations (13) and (14) respectively. Estimates of the dynamics of expected returns and variances, corresponding to the first step of the estimation, are reported in Table 3. Concerning the conditional mean, the autoregressive parameter is found to be significant for Singapore and Thailand. Conditional volatilities display patterns often found in the empirical literature. In particular, volatility is found to be strongly persistent. Also the rather large estimated values of the parameter γ pertaining to lagged squared residuals are consistent with the monthly frequency of the data. The Ljung-Box test statistics with 4 lags (QW(4)) indicate that the residuals are not serially correlated and the Engle (1982) LM test statistics indicate that the heteroskedasticity originally in the data has been filtered out.

[Table 3 somewhere here]

Table 4 reports the estimation results of the multivariate part of the model with Sk-t distribution and time-varying asymmetry and degree-of-freedom parameters. Panel A is devoted to the parameter estimates while Panel B reports summary statistics on the ability of the model to fit the data. Firstly, the conditional correlation matrix is found to vary over time. Its dynamics is stationary, since $\theta_1 + \theta_2 < 1$ and the recent cross-product of residuals has a significant effect on conditional correlation. Inspection of the dynamics of correlation across market returns shows that, after a slight decrease at the beginning of the period, correlations are steadily raising since 1985. For several pairs of markets (for instance between Hong Kong and South Korea or between South Korea and Thailand), the correlation has doubled over the sample.[16]

Table 4 also reveals that for the data sampled at monthly frequency, the asymmetry parameter is essentially constant over time for the stock markets under study. Indeed, the parameter b_1 was systematically found to be positive, yet insignificantly different from zero. Consequently, we concentrate on the case with constant asymmetry parameters and do not report the estimates with a time-varying asymmetry parameter. It should be noted that this result does not prevent market-return skewness to vary over time, since it depends on the degree-of-freedom parameter and on the covariance matrix that are actually time-varying. These findings suggest, however, that the variability of the asymmetry of the distribution is not likely to be the main source of dynamics in higher moments.

Finally, the dynamics of the degree-of-freedom parameters is correctly captured by a simple weighted average of lagged residuals. Parameter b_2 is negative and strongly significant, indicating that after an extreme event, the degree of freedom of the t distribution decreases. This result is consistent with the evidence put forward by Das and Uppal (2004) that extreme events are not likely to cluster.

As far as residual summary statistics are concerned, the test for iidness of the marginal cdfs indicates that the model is able to filter out all the dynamics found in the first four moments of the distribution. In addition, the GoF test statistic does not reject the hypothesis that the assumed Sk-t distribution fits the empirical distribution

[16] Detailed results are available from the authors.

pretty well.

Interestingly, we notice that the extent of the variability in higher moments is rather moderate. However, as it will be shown in the following subsection, it is sufficient to generate a large opportunity cost to use a mean-variance strategy when the investor is risk adverse.

[Table 4 somewhere here]

5.2 Optimal allocation

We now turn to the optimal allocation when returns have time-varying moments. We consider the following alternative distributional assumptions: normal distribution, Sk-t distribution with constant higher moments, and Sk-t distribution with time-varying higher moments (TV Sk-t). We assume here that the latter distribution provides a correct approximation of the true DGP of market returns, so that the opportunity cost of using a sub-optimal forecasting model is evaluated with respect to this model.[17]

Table 5 reports results for optimally selected portfolios for different values of the risk aversion parameter A. For each month of the sample, we use the forecasting models described above to forecast the first four moments and co-moments of market returns. Then, we maximize the approximated expected utility function (18) for the three alternative distributional assumptions. We obtain, for each month, portfolio weights and moments of the portfolio return. In the table, we report summary statistics on these time series. We also report statistics on the risk measure (24) and the opportunity cost (25) of using a given sub-optimal forecasting model.

[Table 5 somewhere here]

[17] The parameter estimates of the first two models are not reported in the paper to save space, but are available upon request from the authors. Note that the dynamics of the conditional correlation matrix is essentially unaltered when we adopt a different distributional assumption.

Considering the average weights obtained for the three distributional models, one may argue that the differences are all in all rather small. It should be emphasized, however, that the actual differences are in fact attenuated by averaging. A measure of the distance between two series of weights can be defined as $dist = \frac{1}{T}\sum_{t=1}^{T}\sum_{i=1}^{n}\left|\hat{\alpha}_{i,t} - \alpha_{i,t}^{*}\right|/2$. Similarly, we measure turnover from one month to the other using the measure $turn = \frac{1}{T}\sum_{t=1}^{T}\sum_{i=1}^{n}\left|\hat{\alpha}_{i,t} - \hat{\alpha}_{i,t-1}\right|/2$.[18] For instance, for $A = 10$, the average distance between the normal and the TV Sk-t distribution is 0.2, i.e. the difference between the two strategies amounts to 20 percentage points on average for the whole portfolio. Besides, the turnover ratio is around 0.52 whatever the distribution, indicating that each month the investor re-allocates 52% of the portfolio. We notice that as the risk aversion increases, the turnover ratio decreases, because the investor increases diversification to reduce her exposure to risk and therefore selects a portfolio that is closer and closer to an equally weighted portfolio. Consequently, the composition of the portfolio is less affected by changes in investment opportunities for large risk aversion. We also notice that the higher the risk aversion, the larger the discrepancy between the mean-variance strategy and the strategy based on higher moments. The reason for this is that the investor is more reluctant to take large risks and therefore gives an increasing importance to higher moments. For instance, for $A = 10$, the investor accepts a slight decrease in expected return (from 1.255 to 1.215) to obtain a large decrease in kurtosis (from 13.52 to 8.47). Figure 2 illustrates this phenomenon. It displays the difference in weight associated to Hong Kong, respectively South Korea, depending on using the TV Sk-t strategy or the mean-variance model. We observe that for some periods (in particular, around 1988-1992 and 2000-2002), large differences are found to hold between the weights associated with the two strategies and for both markets. Such differences correspond to periods where rises in kurtosis were expected, so that the investor who used a TV Sk-t strategy reduced her exposure to the Hong Kong market.

[**Figure 2 somewhere here**]

Returning to Table 5, for all levels of risk aversion, we obtain that the TV Sk-t

[18] The division by 2 corrects for double counting.

distribution improves Berényi's performance measure of the asset allocation as compared to the normal distribution. This result occurs because the Sk-t distribution brings information on both skewness and kurtosis that are useful for maximizing the expected utility. Moreover, the TV Sk-t distribution systematically improves the performances over the distribution with constant higher moments. We emphasize that the opportunity cost of ignoring higher moments is much larger (for medium to large risk aversion) than in the case with constant moments (as reported in Table 2). An obvious reason is that the investor is able to reallocate her portfolio at each date, in order to benefit from the information on higher moments. For $A = 10$, the investor is willing to pay almost 1.96% per year to use the optimal forecasting model. This premium has to be compared with an expected return of about 15% per year. For higher risk aversion, the relative premium is even larger. As it appears clearly in expression (26), the opportunity cost of using a mean-variance criterion increases with the degree of risk aversion. Therefore, an investor with a large risk aversion is willing to pay a large premium to increase skewness or decrease kurtosis.

6 Conclusion

In this paper, we have investigated the consequences of the non-normality of returns on the optimal asset allocation when the investment opportunities are not constant over time. Most previous work has been devoted to the case of constant investment opportunities. Recently, Ang and Bekaert (2002) and Guidolin and Timmermann (2003) have investigated the case where returns are driven by a Markov-switching process with a conditionally normal distribution and constant within-regime moments. We extend these studies in two directions:

- First, we show how to deal with the case of fully time-varying return distributions. For this purpose, we propose a model that captures most statistical features of market returns, such as volatility clustering, correlation persistence and asymmetry and fat-tailedness of the distribution. The use of a two-step estimation procedure makes the estimation of such a model tractable, even in the case of several assets.

- Second, using an approximation of the expected utility function based on a fourth-order Taylor's expansion, we compute the optimal asset allocation for the (sub-optimal) mean-variance criterion and for the four-moment-based strategy. We show that, even for moderate levels of risk aversion, the opportunity cost of using the mean-variance criterion can be very large when returns strongly depart from normality. Our evaluations of the opportunity cost are in general larger than the evaluations reported by Ang and Bekaert (2002).

References

[1] Aït-Sahalia, Y., and M. W. Brandt (2001), Variable Selection for Portfolio Choice, *Journal of Finance*, 56(4), 1297–1351.

[2] Ang, A., and G. Bekaert (2002), International Asset Allocation with Regime Shifts, *Review of Financial Studies*, 15(4), 1137–1387.

[3] Arditti, F. (1967), Risk and the Required Return on Equity, *Journal of Finance*, 22, 19–36.

[4] Athayde, G. M., and R. G. Flôres Jr (2004), Finding a Maximum Skewness Portfolio – A General Solution to Three-moments Portfolio Choice, *Journal of Economic Dynamics and Control*, 28, 1335–1352.

[5] Barone-Adesi, G. (1985), Arbitrage Equilibrium with Skewed Asset Returns, *Journal of Financial and Quantitative Analysis*, 20(3), 299–313.

[6] Bauwens, L., and S. Laurent (2002), A New Class of Multivariate Skew Densities, with Application to GARCH Models, CORE Discussion Paper 2002/20.

[7] Bekaert, G., and C. R. Harvey (1997), Emerging Equity Market Volatility, *Journal of Financial Economics*, 43, 29–77.

[8] Berényi, Z. (2001), Accounting for Illiquidity and Non-normality of Returns in the Performance Assessment, Working Paper, University of Munich.

[9] Bollerslev, T. (1986), Generalized Autoregressive Conditional Heteroskedasticity, *Journal of Econometrics*, 31(3), 307–327.

[10] Brandt, M. (1999), Estimating Portfolio and Consumption Choice: A Conditional Euler Equations Approach, *Journal of Finance*, 54, 1609–1945.

[11] Campbell, J. Y., and L. M. Viceira (1999), Consumption and Portfolio Decisions When Expected Returns Are Time Varying, *Quarterly Journal of Economics*, 114(2), 433–495.

[12] Chamberlain, G. (1983), A Characterization of the Distributions that Imply Mean-variance Utility Functions, *Journal of Economic Theory*, 29, 185–201.

[13] Chunhachinda, P., K. Dandapani, S. Hamid, A. J. Prakash (1997), Portfolio Selection and Skewness: Evidence from International Stock Markets, *Journal of Banking and Finance*, 21(2), 143–167.

[14] Das, S. R., and R. Uppal (2004), Systematic Risk and International Portfolio Choice, *Journal of Finance*, 59, 2809–2834.

[15] Diebold, F. X., T. A. Gunther, and A. S. Tay (1998), Evaluating Density Forecasts with Applications to Financial Risk Management, *International Economic Review*, 39(4), 863–883.

[16] Dittmar, R. F. (2002), Nonlinear Pricing Kernels, Kurtosis Preference, and Evidence from the Cross Section of Equity Returns, *Journal of Finance*, 57, 369–403.

[17] Engle, R. F. (1982), Auto-Regressive Conditional Heteroskedasticity with Estimates of the Variance of United Kingdom Inflation, *Econometrica*, 50(4), 987–1007.

[18] Engle, R. F., and K. Sheppard (2001), Theoretical and Empirical Properties of Dynamic Conditional Correlation Multivariate GARCH, NBER Working Paper 8554.

[19] Friend, I., and R. Westerfield (1980), Co-skewness and Capital Asset Pricing, *Journal of Finance*, 35, 897–913.

[20] Glosten, R. T., R. Jagannathan, and D. Runkle (1993), On the Relation between the Expected Value and the Volatility of the Nominal Excess Return on Stocks, *Journal of Finance*, 48(5), 1779–1801.

[21] Guidolin, M., and A. Timmermann (2003), Optimal Portfolio Choice under Regime Switching, Skew and Kurtosis Preferences, Working Paper, University of California San Diego.

[22] Hansen, B. E. (1994), Autoregressive Conditional Density Estimation, *International Economic Review*, 35(3), 705–730.

[23] Harvey, C. R. (1995), Predictable Risk and Returns in Emerging Markets, *Review of Financial Studies*, 8, 773–816.

[24] Harvey, C. R., and A. Siddique (1999), Autoregressive Conditional Skewness, *Journal of Financial and Quantitative Analysis*, 34(4), 465–487.

[25] Harvey, C. R., and A. Siddique (2000), Conditional Skewness in Asset Pricing Tests, *Journal of Finance*, 55(3), 1263–1295.

[26] Hwang, S., and S. E. Satchell (1999), Modelling Emerging Market Risk Premia Using Higher Moments, *Journal of Finance and Economics*, 4(4), 271–296.

[27] Jondeau, E., and M. Rockinger (2003), Conditional Volatility, Skewness, and Kurtosis: Existence, Persistence, and Comovements, *Journal of Economic Dynamics and Control*, 27, 1699–1737.

[28] Jondeau, E., and M. Rockinger (2004), Optimal Portfolio Allocation Under Higher Moments, forthcoming in *Journal of the European Financial Management Association.*

[29] Jurczenko, E., and B. Maillet (2001), The 3-CAPM: Theoretical Foundations and an Asset Pricing Model Comparison in a Unified Framework, in C. Dunis, A. Timmermann, and J. Moody (eds.), *Developments in Forecast Combination and Portfolio Choice*, John Wiley & Sons, 239–273.

[30] Kimball, M. S. (1993), Standard Risk Aversion, *Econometrica*, 61, 589–611.

[31] Kraus, A., and R. H. Litzenberger (1976), Skewness Preference and the Valuation of Risk Assets, *Journal of Finance*, 31(4), 1085–1100.

[32] Kroll, Y., H. Levy, and H. M. Markowitz (1984), Mean-variance versus Direct Utility Maximization, *Journal of Finance*, 36, 47–61.

[33] Lai, T. Y. (1991), Portfolio Selection with Skewness: A Multiple-objective Approach, *Review of Quantitative Finance and Accounting*, 1, 293–305.

[34] Levy, H., and H. M. Markowitz (1979), Approximating Expected Utility by a Function of Mean and Variance, *American Economic Review*, 69, 308–317.

[35] Markowitz, H. M. (1952), Portfolio Selection, *Journal of Finance*, 7, 77–91.

[36] Pratt, J. W., and R. J. Zeckhauser (1987), Proper Risk Aversion, *Econometrica*, 55(1), 143–154.

[37] Pulley, L. B. (1981), A General Mean-variance Approximation to Expected Utility for Short Holding Period, *Journal of Financial and Quantitative Analysis*, 16, 361–373.

[38] Sahu, S. K., D. K. Dey, and M. D. Branco (2001), A New Class of Multivariate Skew Distributions with Applications to Bayesian Regression Models, Working Paper, Department of Statistics, University of São Paulo.

[39] Samuelson, P. (1970), The Fundamental Approximation Theorem of Portfolio Analysis in Terms of Means, Variances and Higher Moments, *Review of Economic Studies*, 37, 537–542.

[40] Scott, R. C., and P. A. Horvath (1980), On the Direction of Preference for Moments of Higher Order than the Variance, *Journal of Finance*, 35(4), 915–919.

[41] Tauchen, G., and R. Hussey (1991), Quadrature-Based Methods for Obtaining Approximate Solutions to Nonlinear Asset Pricing Models, *Econometrica*, 59(2), 371–396.

[42] Vlaar, P., and F. Palm (1993), The Message in Weekly Exchange Rates in the European Monetary System: Mean Reversion, Conditional Heteroskedasticity and Jumps, *Journal of Business and Economic Statistics*, 11, 351–360.

Captions

Table 1: This table reports parameter estimates (Panel A), summary statistics, and specification tests (Panel B) for the models with constant moments. Returns are assumed to be drawn from a multivariate Sk-t distribution. Implied skewness and kurtosis are deduced from parameter estimates using equations (9) and (10). Conditional correlations are not reported, since they are estimated at their sample values. Following Diebold, Gunther, and Tay (1998), the table also reports the LM test statistics for the null hypothesis that the cdf of residuals is an iid process and the goodness-of-fit (GoF) test statistics for the null hypothesis that the cdf is Uniform$(0, 1)$. Under the null, the statistic is distributed as a $\chi^2(20)$. Wald tests of the type "$x_i = 0$" correspond to all moments x_i, for $i = 1, \cdots, 5$ being equal to zero. The Wald tests are distributed as a $\chi^2(5)$. Tests of the type "$x_i = x_j$" correspond to the equality of all parameters x_i, for $i = 1, \cdots, 5$. The associated Wald tests are distributed as a $\chi^2(4)$.

Table 2: This table reports statistics for optimal allocations obtained using forecasting models with constant moments, for several values of the risk-aversion parameter A ranging from 2 to 20. We report the optimal weights, the first four moments of the portfolio return (the mean, the volatility and the standardized skewness and kurtosis), the risk measure defined in equation (24) and the opportunity cost θ defined by $E[U(1 + \hat{r}_p + \theta)] = E\left[U\left(1 + r_p^*\right)\right]$, where r_p^* denotes the optimal portfolio return obtained with the Sk-t distribution, and $\hat{r}_p$ the optimal portfolio return obtained assuming the normal (or the t) distribution.

Table 3: This table reports (QML) parameter estimates (Panel A) and residual summary statistics (Panel B) for the univariate GARCH models. Parameters correspond to equations (1) and (4) in the text. $QW(4)$ is the Box-Ljung statistic for serial correlation, corrected for heteroskedasticity, computed with 4 lags. Under the null of no serial correlation, it is distributed as a $\chi^2(4)$. $LM(1)$ is the Engle (1982) statistics for heteroskedasticity. Under the null of no serial correlation of squared returns, it is distributed as a $N(0, 1)$. The superscript a (b) corresponds to estimates significant at the 5% (10%) level.

Table 4: This table reports parameter estimates (Panel A) and summary statistics (Panel B) for the models with time-varying moments. Returns are assumed to be drawn from a multivariate Sk-t distribution with time-varying higher moments. The unconditional covariance matrix $\bar{Q}$ is not reported, as it is estimated at its sample value. Parameters correspond to equations (6), (13) and (14). Following Diebold, Gunther, and Tay (1998), the table also reports the LM test statistics for the null hypothesis that the cdf of residuals is an iid process and the goodness-of-fit (GoF) test statistics for the null hypothesis that the cdf is Uniform$(0,1)$. Under the null, this statistic is distributed as a $\chi^2(20)$.

Table 5: This table reports statistics for optimal portfolios obtained using forecasting models with time-varying moments, for several values of the risk-aversion parameter A ranging from 2 to 20. We report the mean over the sample of the optimal weights, the first four moments of the portfolio return (the mean, the volatility and the standardized skewness and kurtosis), the risk measure defined in equation (24) and the opportunity cost θ_{t+1}, defined by $E_t\left[U\left(1+\hat{r}_{p,t+1}+\theta_{t+1}\right)\right]=E_t\left[U\left(1+r^*_{p,t+1}\right)\right]$, where $r^*_{p,t+1}$ denotes the optimal portfolio return obtained with the Sk-t distribution, and $\hat{r}_{p,t+1}$ the optimal portfolio return obtained assuming normal (or t) distribution. The distance between two series of weights is measured by $norm = \frac{1}{T}\sum_{t=1}^{T}\sum_{i=1}^{n}\left|\hat{\alpha}_{i,t}-\alpha^*_{i,t}\right|/2$, while the turnover ratio from one month to the other is measured by $turn = \frac{1}{T}\sum_{t=1}^{T}\sum_{i=1}^{n}\left|\hat{\alpha}_{i,t}-\hat{\alpha}_{i,t-1}\right|/2$.

Figure 1: This figure displays the optimal weights for two markets (Hong Kong and South Korea) when moments are assumed to be constant over time, for several levels of risk aversion A. We consider that returns are drawn from a normal distribution and from a Sk-t distribution in turn.

Figure 2: This figure displays the time series of the differences $(\alpha^*_{i,t}-\hat{\alpha}_{i,t})$ between the allocation weights obtained with the TV Sk-t distribution and the weights obtained with the normal distribution, for two markets (Hong Kong and South Korea).

Table 1: Estimation of the model with constant moments

	Expected return		Variance		Degree-of-freedom parameter		Asymmetry parameter		Implied skewness	Implied kurtosis
Panel A: Parameter estimate										
	Param.	Std dev.	Param.	Std dev.	Param.	Std dev.	Param.	Std dev.	Param.	Param.
Hong Kong (HK)	1.371	*(0.450)*	70.407	*(7.111)*	5.646	*(1.688)*	-0.051	*(0.076)*	-0.190	6.622
Singapor (SI)	0.583	*(0.370)*	49.415	*(4.775)*	6.441	*(2.201)*	-0.026	*(0.071)*	-0.079	4.614
S. Korea (SK)	0.733	*(0.398)*	60.579	*(5.213)*	11.948	*(7.167)*	0.206	*(0.083)*	0.318	3.226
Taiwan (TA)	1.183	*(0.547)*	116.123	*(15.464)*	4.204	*(0.841)*	0.086	*(0.070)*	0.436	27.052
Thaliand (TH)	0.863	*(0.474)*	99.090	*(18.520)*	2.982	*(0.387)*	0.078	*(0.068)*	*n.a.*	*n.a.*

Panel B: Summary statistics

Test for iid-ness of residuals

	First moment		Second moment		Third moment		Fourth moment	
	Stat.	*p-val.*	Stat.	*p-val.*	Stat.	*p-val.*	Stat.	*p-val.*
HK	13.959	*(0.833)*	34.103	*(0.025)*	22.803	*(0.299)*	35.290	*(0.019)*
SI	24.205	*(0.234)*	33.212	*(0.032)*	32.315	*(0.040)*	32.904	*(0.035)*
SK	20.042	*(0.455)*	53.481	*(0.000)*	24.754	*(0.211)*	49.877	*(0.000)*
TA	18.392	*(0.562)*	56.169	*(0.000)*	24.986	*(0.202)*	74.496	*(0.000)*
TH	41.951	*(0.003)*	53.740	*(0.000)*	34.252	*(0.025)*	55.620	*(0.000)*

GoF test

	Stat.	*p-val.*
HK	19.624	*(0.418)*
SI	5.980	*(0.998)*
SK	10.644	*(0.935)*
TA	15.542	*(0.688)*
TH	13.327	*(0.821)*

Wald tests

	Stat.	*p-val.*
$\mu_i=0$	13.048	*(0.023)*
$\mu_i=\mu_j$	4.165	*(0.384)*
$\sigma_i=\sigma_j$	26.315	*(0.000)*
$1/\nu_i=0$	86.422	*(0.000)*
$\nu_i=\nu_j$	7.903	*(0.095)*
$\lambda_i=0$	11.357	*(0.045)*
$\lambda_i=\lambda_j$	8.722	*(0.068)*

Table 2: Optimal portfolio under constant moments

	Portfolio weights					Moments of portfolio return				Risk	Opportunity cost
	HK	SI	SK	TA	TH	μ_p	σ_p	s_p	κ_p	measure	(in % per year)
A = 2											
normal	0.677	0.000	0.064	0.258	0.002	1.281	7.028	-0.117	6.115	0.511	0.005
t	0.665	0.000	0.078	0.252	0.005	1.271	6.960	-0.117	6.077	0.500	0.001
Sk-*t*	0.655	0.000	0.081	0.258	0.006	1.268	6.937	-0.112	6.083	0.497	–
A = 5											
normal	0.421	0.107	0.239	0.187	0.047	1.076	6.025	-0.074	5.108	0.946	0.032
t	0.396	0.125	0.255	0.179	0.045	1.053	5.955	-0.068	4.968	0.921	0.005
Sk-*t*	0.383	0.123	0.265	0.183	0.046	1.047	5.940	-0.059	4.915	0.915	–
A = 10											
normal	0.292	0.241	0.274	0.151	0.042	0.956	5.729	-0.051	4.502	1.805	0.179
t	0.252	0.265	0.302	0.142	0.039	0.923	5.682	-0.036	4.293	1.758	0.023
Sk-*t*	0.238	0.257	0.320	0.146	0.039	0.917	5.682	-0.024	4.216	1.750	–
A = 15											
normal	0.248	0.287	0.286	0.139	0.040	0.916	5.673	-0.043	4.309	2.932	0.541
t	0.197	0.312	0.325	0.129	0.037	0.875	5.645	-0.021	4.069	2.851	0.060
Sk-*t*	0.184	0.298	0.350	0.132	0.037	0.869	5.655	-0.006	3.985	2.841	–
A = 20											
normal	0.226	0.310	0.293	0.133	0.039	0.895	5.654	-0.039	4.220	4.429	1.195
t	0.169	0.333	0.341	0.122	0.035	0.850	5.642	-0.012	3.963	4.294	0.118
Sk-*t*	0.156	0.317	0.369	0.124	0.035	0.845	5.658	0.004	3.878	4.279	–

Table 3: Parameter estimates of the univariate GARCH models

	Hong Kong	Singapore	South Korea	Taiwan	Thailand
Panel A: Parameter estimates					
Return equation					
μ	1.320 [a]	0.399	1.032 [a]	0.610	0.496
	(0.428)	*(0.353)*	*(0.324)*	*(0.433)*	*(0.339)*
ϕ	0.068	0.131 [b]	0.057	0.100	0.188 [a]
	(0.061)	*(0.060)*	*(0.059)*	*(0.058)*	*(0.062)*
Volatility equation					
ω	5.358 [b]	4.041	1.726	3.320	1.197 [b]
	(2.614)	*(2.292)*	*(0.986)*	*(2.286)*	*(0.594)*
β	0.810 [a]	0.826 [a]	0.791 [a]	0.808 [a]	0.832 [a]
	(0.057)	*(0.070)*	*(0.056)*	*(0.060)*	*(0.036)*
γ	0.119 [a]	0.087 [b]	0.200 [a]	0.168 [a]	0.168 [a]
	(0.042)	*(0.044)*	*(0.059)*	*(0.054)*	*(0.036)*
Panel B: Summary statistics					
QW(4)	4.257	3.390	1.830	1.617	7.955
p-val.	*(0.372)*	*(0.495)*	*(0.767)*	*(0.806)*	*(0.093)*
LM(1)	0.276	0.022	0.228	0.149	0.222
p-val.	*(0.599)*	*(0.881)*	*(0.633)*	*(0.699)*	*(0.638)*

Table 4: Estimation of the models with time-varying moments

Panel A: Parameter estimate

	Param.	Std dev.
Correlation equation		
δ_1	0.933	*(0.022)*
δ_2	0.030	*(0.008)*
Degree-of-freedom equation		
b_1	-1.015	*(0.092)*
Asymmetry equation		
λ_1	-0.191	*(0.153)*
λ_2	-0.111	*(0.157)*
λ_3	0.441	*(0.163)*
λ_4	0.235	*(0.157)*
λ_5	0.262	*(0.131)*

Panel B: Summary statistics

Test for iid-ness of residuals

	First moment		**Second moment**		**Third moment**		**Fourth moment**	
	Stat.	*p-val.*	Stat.	*p-val.*	Stat.	*p-val.*	Stat.	*p-val.*
HK	5.809	*(0.669)*	11.367	*(0.498)*	11.231	*(0.940)*	21.844	*(0.349)*
SI	4.473	*(0.812)*	9.270	*(0.680)*	21.813	*(0.351)*	24.701	*(0.213)*
SK	5.003	*(0.757)*	11.824	*(0.460)*	21.266	*(0.382)*	24.657	*(0.215)*
TA	4.631	*(0.796)*	11.307	*(0.503)*	15.501	*(0.747)*	13.861	*(0.838)*
TH	17.532	*(0.025)*	20.495	*(0.058)*	32.743	*(0.036)*	21.764	*(0.353)*

GoF test	Stat.	*p-val.*
HK	20.355	*(0.436)*
SI	24.175	*(0.235)*
SK	24.964	*(0.203)*
TA	16.767	*(0.668)*
TH	24.088	*(0.239)*

Table 5: Optimal portfolio under time-varying moments

	Portfolio weights					dist	Turnover ratio	Moments of portfolio return				Risk measure	Opportunity cost (in % per year)
	HK	SI	SK	TA	TH			μ_p	σ_p	s_p	κ_p		
A=2													
Normal	0.397	0.072	0.239	0.088	0.204	0.055	1.082	1.686	6.758	0.003	14.394	0.488	0.025
Sk-*t*	0.388	0.072	0.245	0.089	0.206	0.045	1.082	1.682	6.734	0.009	14.336	0.485	0.020
TV Sk-*t*	0.375	0.073	0.251	0.091	0.210	-	1.076	1.678	6.712	0.018	13.260	0.479	–
A=5													
Normal	0.318	0.091	0.311	0.089	0.192	0.118	0.753	1.453	5.553	0.034	13.782	0.872	0.305
Sk-*t*	0.302	0.091	0.323	0.091	0.193	0.104	0.751	1.445	5.525	0.052	13.751	0.859	0.250
TV Sk-*t*	0.279	0.097	0.325	0.098	0.202	-	0.748	1.433	5.522	0.061	10.619	0.826	–
A=10													
Normal	0.255	0.130	0.325	0.097	0.194	0.201	0.523	1.255	5.011	0.058	13.517	1.664	1.957
Sk-*t*	0.231	0.130	0.344	0.101	0.195	0.186	0.519	1.241	4.990	0.090	13.532	1.619	1.585
TV Sk-*t*	0.204	0.138	0.332	0.113	0.213	-	0.533	1.215	5.025	0.076	8.472	1.461	–
A=15													
Normal	0.223	0.154	0.318	0.102	0.203	0.257	0.420	1.153	4.839	0.068	13.433	2.878	5.856
Sk-*t*	0.193	0.156	0.342	0.108	0.202	0.243	0.413	1.135	4.823	0.108	13.478	2.763	4.699
TV Sk-*t*	0.171	0.158	0.322	0.124	0.225	-	0.460	1.111	4.915	0.076	7.424	2.348	–
A=20													
Normal	0.205	0.171	0.312	0.105	0.208	0.301	0.364	1.095	4.766	0.072	13.410	4.721	13.064
Sk-*t*	0.171	0.172	0.340	0.111	0.206	0.288	0.356	1.075	4.755	0.117	13.465	4.483	10.457
TV Sk-*t*	0.156	0.167	0.315	0.132	0.230	-	0.436	1.062	4.900	0.073	6.769	3.598	–

Figure 1: Optimal weights under unconditional moments

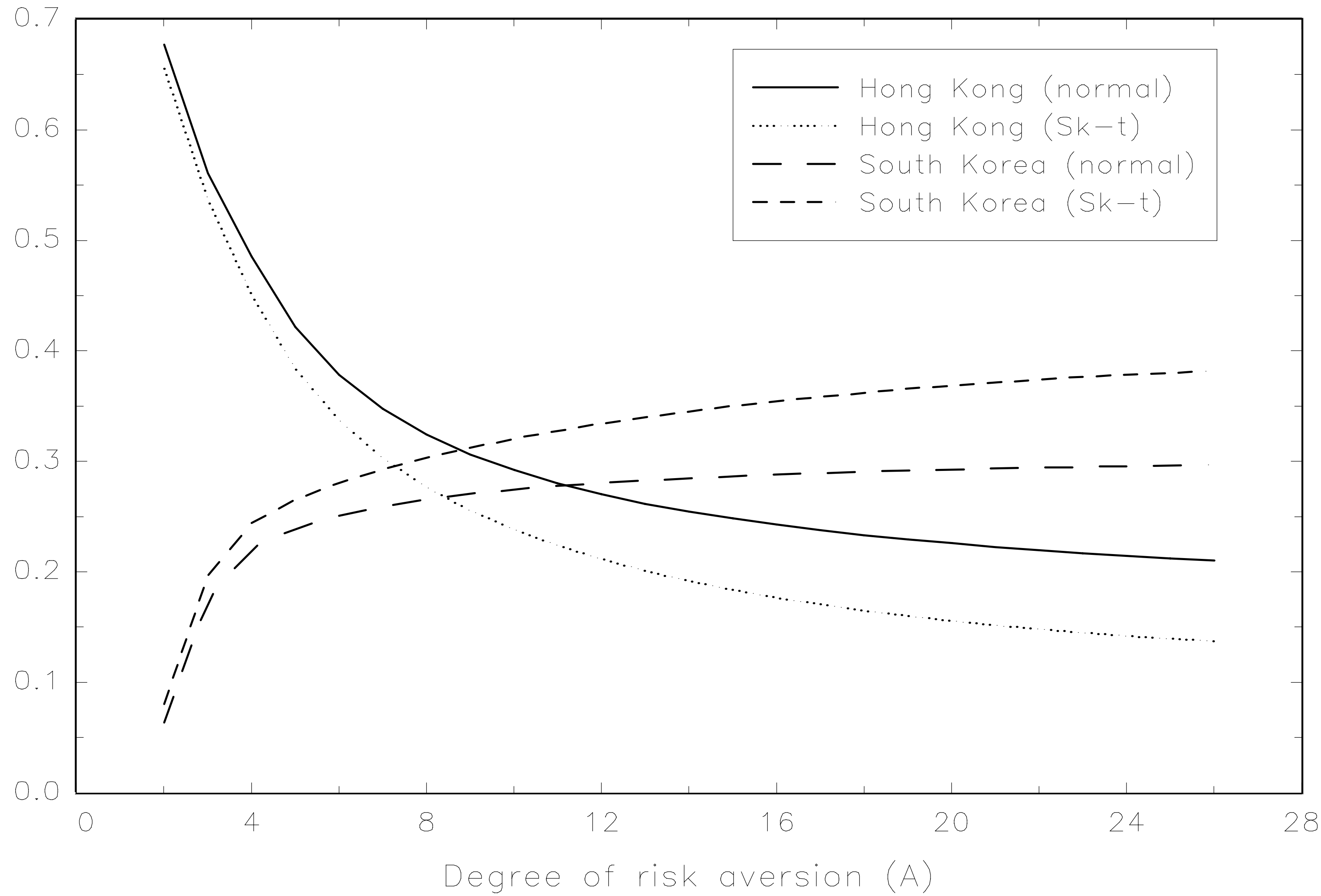

Figure 2: Model-driven allocational differences