The Coskewness Puzzle

Valerio Potì,
Dublin City University Business
School and Trinity College

1st Draft: January 2004
This Version: 22nd November 2005

COMMENTS WELCOMED

I derive the beta-pricing representation of Harvey and Siddique (2000) conditional 3M-CAPM and demonstrate the restrictions that it imposes on a quadratic market factor model. I confirm that, while the latter is surprisingly successful at capturing the cross-sectional variation of the returns on the Fama and French (1995) 30 US industry portfolios, the 3M-CAPM coskewness and gamma premia for the period 1952-2002 turn out to be much smaller when non satiation, risk aversion and non-increasing absolute risk aversion are imposed. However, I show that they are still important in explaining the expected return on certain assets and strategies, such as hedge funds and strategies that mimic the momentum factor. Alternatively, I propose to use a quadratic market factor model with an upper bound on the volatility of the implied stochastic discount factor. This approach allows the explanatory power of the quadratic specification to be salvaged yet avoiding fitting unexpected extreme outcomes rather then pricing patterns.

Key words: Asset Pricing, Coskewness

JEL Classification: G12

Contact details:

Valerio Potì, Finance Lecturer, Room Q233, Dublin City University Business School, Glasnevin, Dublin 9, Ireland; Tel: 3531-7005328, Fax: 3531- 7005446; Email: valerio.poti@dcu.ie.

The author wishes to thank the helpful comments and suggestions provided by Colm Kearney (Trinity College Dublin), Brian Lucey (Trinity College Dublin), Thierry Post (Rotterdam School of Economics), Pim Van Vliet (Rotterdam School of Economics and Robecco Asset Management), Michael Moore (Queen's University Belfast), Liam Gallagher (Dublin City University) and by the participants to the FMA 2004 European Meeting, the FMA 2004 Annual Meeting and the EFA 2005 Doctoral Colloquium. He also wishes to thank the Irish Research Council for the Humanities and the Social Sciences for the generous financial support.

The Coskewness Puzzle

I derive the beta-pricing representation of Harvey and Siddique (2000) conditional 3M-CAPM and demonstrate the restrictions that it imposes on a quadratic market factor model. I confirm that, while the latter is surprisingly successful at capturing the cross-sectional variation of the returns on the Fama and French (1995) 30 US industry portfolios, the 3M-CAPM coskewness and gamma premia for the period 1952-2002 turn out to be much smaller when non satiation, risk aversion and non-increasing absolute risk aversion are imposed. However, I show that they are still important in explaining the expected return on certain assets and strategies, such as hedge funds and strategies that mimic the momentum factor. Alternatively, I propose to use a quadratic market factor model with an upper bound on the volatility of the implied stochastic discount factor. This approach allows the explanatory power of the quadratic specification to be salvaged yet avoiding fitting unexpected extreme outcomes rather then pricing patterns.

1. Introduction

There is evidence that assets that display similar covariance but high coskewness with the market portfolio tend to display low average returns. Kraus and Litzenberger (1976), Friend and Westerfield (1980) and Harvey and Siddique (2000), among others, explain this empirical regularity on the basis of a three-moment extension of the Capital Asset Pricing Model (3M-CAPM). They find that coskewness is important and commands on average a risk premium of 3.6 percent per annum. Both Dittmar (2002) and Post, Levy and van Vliet (2003), however, find that the significance of the coskewness premium in the cross-section of industry and size-sorted portfolio average returns is greatly reduced when the shape of the representative investor's utility function is restricted to display non satiation, risk aversion and non increasing absolute risk aversion (henceforth, NS, RA and NIARA, respectively) over all values of sample wealth.

In this paper, I study the unconditional implications of stochastic discount factors that are non linear in wealth, both in a no-arbitrage framework and in a utility-based equilibrium pricing setting. I especially investigate the role of portfolio skewness in asset pricing, both under preference for skewness and in an almost preference-free setting. In the next Section, I present some background analytical results on stochastic discount factor pricing that will be useful to switch from a no-arbitrage, preference-free context to an equilibrium, utility based framework and vice versa. In Section 3, I then review the beta-pricing representation of the stochastic discount factor models. In Section 4, I apply this beta-pricing representation to stochastic discount factors that are non-linear in traded wealth. In Section 5, I show how the 3-Moment conditional CAPM (3M-CCAPM) formulated by Harvey and Siddique (2000) can be seen as a special case of this formulation. I also specify an empirical version of this beta pricing representation that explicitly uses conditioning information. In Section 6, I present my dataset. In Section 7, I estimate unconditional and conditional versions of this model. In Section 8, I present further empirical evidence on the cross-sectional explanatory power of the Quadratic Market Factor Model and I compare it with the empirical performance of Fama and French's (1995) 3-Factor Model. In Section 9, I introduce the alternative beta-gamma representation of the 3M-CAPM proposed by Kraus and Litzenberger (1976), I extend it in a conditional setting and, more importantly, I derive conditions for imposing a well behaved shape of the utility function. In Section 10, I estimate the 3M-CAPM imposing these conditions both in an unconditional and conditional setting. In Section 11, I discuss the implications of the empirical evidence for the representative investor and I show how this gives rise to a coskewness puzzle. In Section 12, 13 and 14, I discuss possible solutions to this puzzle and I show that coskewness is important even if the market price of coskewness risk is modest and that it is possible to preserve a substantial cross-sectional explanatory power while ruling out extreme Sharpe ratios. This requires bounding the volatility of the quadratic discount factor. A consequence of imposing this bound is that the coskewness puzzle becomes considerably milder. In Section 15, I summarize my findings and present my conclusions.

2. Stochastic Discount Factor Pricing

I start with as little structure as possible by invoking a theorem credited to Harrison and Kreps (1979). This theorem says that, given free portfolio formation and the law of one price, a stochastic process m_{t+1} that prices all assets exists. This process satisfies the following condition for all payoffs x_{t+1} and payoff prices p_t:

$$p_t = E_t(m_{t+1}x_{t+1}) \qquad (1)$$

Here, the expectation is taken conditional on the available information set. Under the additional assumption of no arbitrage, as shown by Hansen and Richard (1987), m_{t+1} must be positive. If the set of the assets being priced spans all possible payoffs, m_{t+1} is unique and is called the stochastic discount factor (henceforth SDF). Instead, if the priced assets are only a subset of the universe of assets, there is an infinite choice of processes m_{t+1} that satisfy (1). These processes share the same projection as the SDF on the priced payoff space. Trivially, the SDF is one of these processes. Any linear combination of these processes prices the assets.

Given observed prices and the conditional distribution of payoffs, (1) can be used to infer the process followed by m_{t+1} and thus by the relevant component of the SDF. Once we recover m_{t+1} from a set of asset prices and payoffs, we can then use (1) to price any asset that can be represented as a combination of the basis payoffs. To estimate the process followed by m_{t+1}, I model it as a linear function of a set of factors f_{t+1}:

$$m_{t+1} = a_t + b_t' f_{t+1} \tag{2}$$

Thus, from the perspective of (1) and (2), solving the asset pricing problem amounts to finding the linear combination of factors that prices traded assets. This problem in turn can be decomposed into first searching for the elementary strategies that can be used to replicate the asset payoffs and than assigning prices to the strategies. The factors in (2) should therefore be either payoffs of elementary strategies or the projection in the payoff space of processes that uniquely identify the strategies (e.g. trading rules). The linearity of the pricing equation (1) implies that the price of the assets is a linear combination of the prices of the elementary strategies used to replicate them.

3. Beta Pricing Representation

To derive the beta pricing representation of the conditional asset pricing implications of (1) and (2), it is easier to work with demeaned factors and thus it is convenient to fold the mean of the latter into the constant. Therefore, I let $\tilde{a}_t = a_t + b_t' E_t(f_{t+1})$ and $\tilde{f}_{t+1} = f_{t+1} - E_t(f_{t+1})$. The process m_{t+1} is then rewritten as follows:

$$m_{t+1} = \tilde{a}_t + b_t' \tilde{f}_{t+1} \tag{3}$$

Let $R_{i,t}$ denote the return on the asset i. The price of gross returns, i.e. $1 + R_{i,t}$, is equal to 1 by definition, thus (1) implies $1 = E_t[m_{t+1}(1 + R_{it+1})]$ for every asset i and we have

$$1 = E_t(m_{t+1}) E_t(1 + R_{i,t+1}) + Cov_t[m_{t+1}, (1 + R_{i,t+1})] \tag{4}$$

This equation can be easily rewritten in beta-pricing form, thus giving the beta-pricing representation of the asset pricing implication of (1) and (2):

$$E_t(R_{i,t+1}) = \gamma_t + \beta_{i,t}' \lambda_t \tag{5}$$

where,

$$\gamma_t = \frac{1}{\tilde{a}_t} \tag{6}$$

$$\beta_{i,t} = E_t\left(\tilde{f}_{t+1}' \tilde{f}_{t+1}\right)^{-1} E_t\left(\tilde{f}_{t+1}' R_{i,t+1}\right) = Var_t\left(f_{t+1}\right)^{-1} Cov_t\left(f_{t+1} R_{i,t+1}\right) \tag{7}$$

$$\lambda_t = -\frac{E_t\left(\tilde{f}'_{t+1}\tilde{f}_{t+1}\right)b_t}{\tilde{a}_t} = -\frac{Var_t\left(f_{t+1}\right)b_t}{\tilde{a}_t} \tag{8}$$

Recognising that the conditionally risk free rate $R_{f,t}$ is defined as $1+R_{f,t} = \frac{1}{E_t\left(m_{t+1}\right)}$, the intercept in (5) is $\gamma_t = 1+R_{f,t}$. Also, $\beta_{i,t}$ is a vector of coefficients from the regression of asset *i* on the factors and λ_t is a parameter vector. The former can be seen as the factor loadings whereas the latter are the price of the demeaned factors minus their risk neutral valuation, i.e. the factor risk premia. If the factor $f_{j,t+1}$ is a return, (5) implies that $E_t\left(f_{j,t+1}\right) = \gamma + \lambda_{j_t}$. Moreover, (5) can be rewritten using excess returns as follows:

$$E_t(r_{i,t+1}) = -E_t\left(\tilde{f}'_{t+1}r_{i,t+1}\right)b_t = -Cov_t\left(r_{i,t+1}, f_{t+1}\right)b_t \tag{9}$$

$$E_t\left(r_{i,t+1}\right) = \beta'_{i,t}\lambda_t \tag{10}$$

where,

$$\beta_{i,t} = E_t\left(\tilde{f}'_{t+1}\tilde{f}_{t+1}\right)^{-1}E_t\left(\tilde{f}'_{t+1}r_{i,t+1}\right) = Var_t\left(f_{t+1}\right)^{-1}Cov_t\left(f_{t+1}r_{i,t+1}\right) \tag{11}$$

Here, lower case letters denote excess returns. Notice that here $\beta_{i,t}$ is a vector of coefficients from the regression of asset *i* excess returns on the factors while all the other symbols are defined as before.

4. Non Linear Factor Pricing

Assume that the cheapest strategies available to investors to replicate all traded assets yield payoffs represented by polynomials of the return on stock market wealth. A buy and hold strategy of a portfolio of all assets yields the return on market wealth, whereas higher order polynomial of the latter can be replicated either buying and holding options or by resorting to dynamic trading strategies that replicate the latter. Formally, this can be modelled imposing an *n*th order polynomial structure on equation (2) with $f'_{t+1} = \left[R_{m,t+1} \quad ..... \quad R^n_{m,t+1}\right]$:

$$m_{t+1} = a_t + b_{1,t} R_{m,t+1} + ... + b_{n,t} R^n_{m,t+1} \tag{12}$$

Here, $R_{m,t+1}$ denotes the rate of return on stock market wealth. This specification, while linear in polynomials of the stock market return, implies that m_{t+1} is a non linear function of the latter. Imposing the existence of a conditionally risk free rate, (12) is equivalent to the following:

$$m_{t+1} = a_t + b_{1,t} r_{m,t+1} + ... + b_{n,t} \left(R^n_{m,t+1} - R_{f,t}\right) \tag{13}$$

Here, $f'_{t+1} = \left[r_{m,t+1} \quad \left(R^2_{m,t+1} - R_{f,t+1}\right) \quad ...... \quad \left(R^n_{m,t+1} - R_{f,t+1}\right)\right]$ can be seen as a new set of factors. The no-arbitrage conditional excess returns can be obtained by plugging these factors in (9) and (10). Using the factors in (12) amounts to pricing using the stock market return, whereas using (13) amounts to using the excess market return and the difference between its polynomials and the risk free rate. In the latter case, with n = 1, the factor loading is analogous to the CAPM beta. With n = 2, the set of factor loadings also includes the asset covariance and coskewness with the market standardized by the variance-covariance matrix of the factors. I label this specification Quadratic Market Factor Model (henceforth QMFM). Setting n = 3, yields a Cubic Market Factor Model (henceforth CMFM) that allows for cokurtosis, in addition to covariance and coskewness, to be priced in the cross section of conditional expected asset returns.

Using $q_{m,t+1} = R_{m,t+1}^2 - R_{f,t}$ as the shorthand notation for the second order market return polynomial in excess of the risk free rate, (9) can be rewritten in term of the QMFM factors in (13) as follows:

$$E_t\left(r_{i,t+1}\right) = -Cov_t(r_{i,t+1}, r_{m,t+1})b_{1,t} - Cov_t(r_{i,t+1}, q_{m,t+1})b_{2,t} \tag{14}$$

This shows that $-b_{1,t}$ can be interpreted as the price of market risk and $-b_{2,t}$ can be interpreted as the price of market volatility risk, i.e. the price of the risk arising from the quadratic market factor.

Similarly, the betas and risk premia in (10) can be rewritten in terms of the QMFM factors in (13) as follows:

$$E_t\left(r_{i,t+1}\right) = \beta_{i,t}'\lambda_t \tag{15}$$

where,

$$\beta_{i1,t} = \frac{Cov_t(r_{i,t+1}, r_{m,t+1})Var_t(q_{m,t+1}) - Cov_t(r_{m,t+1}, q_{m,t+1})Cov_t(r_{i,t+1}, q_{m,t+1})}{Var_t(r_{m,t+1})Var_t(q_{m,t+1}) - [Cov_t(r_{m,t+1}, q_{m,t+1})]^2} \quad (15a)$$

$$\beta_{i2,t} = \frac{Cov_t(r_{i,t+1}, q_{m,t+1})Var_t(r_{m,t+1}) - Cov_t(r_{m,t+1}, q_{m,t+1})Cov_t(r_{i,t+1}, r_{m,t+1})}{Var_t(r_{m,t+1})Var_t(q_{m,t+1}) - [Cov_t(r_{m,t+1}, q_{m,t+1})]^2} \quad (15b)$$

$$\lambda_t = -\begin{bmatrix} Var_t(r_{m,t+1}) & Cov_t(r_{m,t+1}, q_{m,t+1}) \\ Cov_t(r_{m,t+1}, q_{m,t+1}) & Var_t(q_{m,t+1}) \end{bmatrix} \begin{bmatrix} b_{1,t} \\ b_{2,t} \end{bmatrix} \quad (15c)$$

Here, $Cov_t(r_{m,t+1}, q_{m,t+1})$ is a proxy for the skewness of the market excess return distribution and $Cov_t(r_{i,t+1}, q_{m,t+1})$ proxies for the coskewness of asset *i* excess returns with the excess return on the market. This is perhaps easiest to see considering that, since the conditionally risk free rate is known at time *t*, these expression can be equivalently rewritten as $Cov_t(R_{m,t+1}, R^2_{m,t+1})$ and $Cov_t(R_{i,t+1}, R^2_{m,t+1})$, respectively. Thus the factor loadings $\beta_{i1,t}$ and $\beta_{i2,t}$ are functions of the market variance, its skewness and of the covariance and coskewness of the asset *i* with the market. Equation (15) is a restriction that the model in (2) imposes on the cross section of expected asset returns. Since these restrictions must hold also for the market portfolio, it follows that the no-arbitrage equilibrium market risk premium contains both a market variance and a market skewness premium

$$\lambda_{1t} = E_t\left(r_{m,t+1}\right) = -b_{1,t}Var_t(r_{m,t+1}) - b_{2,t}Cov_t(r_{m,t+1}, q_{m,t+1}) \quad (16)$$

Thus, when $\beta_{i2,t} \neq 0$, the QMFM implies the following data generating process of asset returns:

$$\begin{aligned} R_{i,t+1} &= R_{f,t} + \alpha_{i,t} + \beta_{i1,t} r_{m,t+1} + \beta_{i2,t} q_{m,t+1} + \varepsilon_{i,t+1} \\ &= R_{f,t}(1 - \beta_{i1,t} - \beta_{i2,t}) + \alpha_{i,t} + \beta_{i1,t} R_{m,t+1} + \beta_{i2,t} R_{m,t+1}^2 + \varepsilon_{i,t+1} \end{aligned} \tag{17}$$

$$\alpha_{i,t} = \beta_{i2,t} \left[\lambda_{2,t} - E_t \left(q_{m,t+1} \right) \right] \tag{18}$$

Here, $\varepsilon_{i,t+1}$ is an asset-specific regression residual. When (18) holds for all assets, this data generating process implies an APT-type no arbitrage equilibrium between asset prices and the prices of the factors, as in Ross (1976) and Chamberlain and Rothschild (1983). This data generating process is also the same as the quadratic market model used by Barone Adesi, Gagliardini and Urga (2004)[1]. However, these authors do not derive the equations for the prices of the factors nor they explicit the link between the risk premia (the elements of λ_t) and the risk prices (the negative of the elements of b_t). Describing this link is important because it clarifies that coskewness and time-variation in betas matter in asset pricing and performance evaluation even when m_{t+1} is linear in wealth. From (15c), it is clear that the quadratic market factor risk premium $\lambda_{2,t}$ does not need to be zero when $b_{2,t}$ is zero. In other words, the volatility risk premium can be non-zero even when the market price of volatility risk is zero.

[1] Barone Adesi, Gagliardini and Urga (2004) however do not derive the betas and the prices of the factors. The equations for these provided in this paper and the SDF setting are therefore novel.

5. Asset Pricing Models with Coskewness

So far I examined the role of coskewness in asset pricing and in the data generating process of asset returns purely on no arbitrage grounds. I will now discuss how specifications that include coskewness arise in the context of equilibrium models that specify investors' preferences and explicitly solve their asset allocation problem. I will start from the 3M-CAPM that solve a 3-moment mean-variance-skewness portfolio optimization problem for the representative investor. This is the classic equilibrium model traditionally used to account for the explanatory power of coskewness in the cross section of asset returns. I will then show how coskewness terms appear also in multi-period models that solve, in each period, a 2-moment mean-variance optimization problem.

The 3M-CAPM

Kraus and Litzenberger (1976) and Harvey and Siddique (2000) approximate the stochastic discount factor as a quadratic function of the market return $R_{m,t+1}$. This formulation corresponds to the case $n = 2$ in (12). Kraus and Litzenberger (1976) specify b in (12) as a vector of constants whereas Harvey and Siddique (2000) allow it to be conditionally time-varying. Both specifications are based on a third order Taylor expansion of a non polynomial representative investor utility function because, as shown by Levy (1969) and Tsiang (1972), third degree polynomial utility functions are unsuitable to model the preferences of a risk adverse investor. Under the law of one price, it is possible to impose the Euler equation equilibrium conditions for the maximization of a 2-period representative investor's utility, setting prices of the excess returns to zero $0 = E_t(m_{t+1}r_{it+1})$ for every asset *i*. This yields the 3M-CAPM asset pricing equation.

Equation (10), with beta coefficients and factor risk premia as in (15), provides a beta-pricing representation of the 3M-CAPM. In this context, the factor loadings $\beta_{i1,t}$ and $\beta_{i2,t}$ are both coefficients of the multiple regression of the asset excess-return on the market return and its square and risk measures. From this point of view, they are analogous to the CAPM beta coefficient, even though the later are simple regression coefficients. The upshot of this representation relative to the specification employed by Harvey and Siddique (2000) is that it clarifies the relation between risk premia and prices of risk, in (15c). Moreover, it can be used to test the unconditional implications of the conditional 3M-CAPM for the cross section of asset returns using a simple two pass estimation procedure, by regressing asset excess returns on the factors and the average excess returns on the estimated factor loadings. Then, on the basis of (16), the parameters b_t of the stochastic discount factor in (2) and thus of the representative investor's utility function can then be easily retrieved from estimates of the moments of the market return and its square and from the second pass estimates of their risk premia, i.e. the elements of the λ_t vector. Harvey and Siddique (2000) use instead 2-pass regressions based on mimicking portfolios from which it is difficult to infer the implied stochastic discount factor.

This setup is quite general and can price in a unified framework both non-linear payoff and market timing under a variety of circumstances. For example, it is clear from (15c) that the coskewness risk premium λ_2 is non-zero even when $b_{2,t} = 0$ as long as $b_{1,t} \neq 0$. Thus a non-zero coskewness risk premium can arise also under a linear m_{t+1} and thus also in a mean-variance framework if $\beta_{i2,t} \neq 0$ as long as the asset return distribution is non elliptic. This means that even a mean-variance investor is interested in asset with positive coskewness, as long as he is risk-averse. In this case however the coskewness risk premium is purely a function of the price of the dynamic exposure to the market risk factor generated by $\beta_{i2,t} \neq 0$ and no premium arises from preference for skewness. This implies that, even under an m_{t+1} linear in the stock market portfolio, $\alpha_{i,t}$ in (18) contains a coskewness premium for any payoff that displays a non-linear relation with the stock market return, such as assets with non zero market coskewness and strategies that pursue market timing objectives. In other words, a non-zero $\lambda_{2,t}$ implies that the squared market return risk factor is priced, i.e. it has a non-zero price, but it does not imply that it helps pricing other payoffs. For it to help price the other payoffs and, in particular, payoffs that are linear in the market return, it is necessary that $b_{2,t} \neq 0$.

Arrow (1971) argued that investors' utility functions should display non satiation (NS), risk aversion (RA) and non-increasing absolute risk aversion (NIARA). The latter is related to the notion of prudence, see Kimball (1990). With utility functions $u(W)$ defined over wealth, NS implies positive marginal utility of wealth, i.e. $u'(W) > 0$, RA implies decreasing marginal utility, i.e. $u''(W) < 0$, whereas NIARA, i.e. $\frac{d(-u''/u')}{dW} \leq 0$, implies that the rate of decrease of marginal utility does not increase in wealth. A necessary condition for NIARA, as shown in Arditti (1967), is $u''' \geq 0$. Hence NIARA implies $u''' \geq 0$ and aversion to negative skewness.

Differentiating (13) twice with respect to wealth, it becomes clear that a necessary and sufficient condition for $u''' \geq 0$ is $b_2 \geq 0$ and thus this is also a necessary condition for NIARA. When this condition holds, a necessary condition for RA is $b_1 < 0$. In my beta-pricing representation, it is not possible to fully restrict the sign of the elements of λ_t *a priori* but, if coskewness is a second order effect relative to covariance and if $b_1 < 0$, $\lambda_{1,t}$ should be positive to rule out a negative equity premium. Also, when market skewness is non positive, i.e. when $Cov_t(r_{m,t+1}, q_{m,t+1}) \leq 0$, $b_1 < 0$ and $b_2 \geq 0$ imply $\lambda_{2,t} \leq 0$. Thus, $\lambda_{1,t} > 0$ and $\lambda_{2,t} \leq 0$ are necessary conditions for 'well behaved' utility functions, i.e. utility functions that display NS, RA and NIARA. Moreover, given estimates of λ_t and of the factors variance-covariance matrix, it is possible to recover estimates of the elements of b and thus of the shape of the utility function from (15c).

Conditional Models

A premium to hold assets with negative systematic coskewness can also arise from the desire to hedge against changes to future expected uncertainty in a multi-period setting. This is one of the main results of Chen's (2002) extension of Campbell (1993) inter-temporal conditional CAPM. It is a consequence of allowing for time-varying expected returns, heteroskedasticity in conditional second moments and

NIARA. In particular, it is shown that time-varying conditional market volatility helps pricing only those assets that are conditionally co-skewed with the market. A further coskewness premium would appear if a second order Taylor expansion of the inter-temporal budget constraint was to be used in place of the first-order expansion used by Chen (2002).

The unconditional implications of Chen's (2002) model can be summarized by a 3-moment conditional CAPM specification with conditioning variables that forecast future returns and thus proxy for time-variation in the investment opportunity set. In particular, let the a_t and b_t parameters of the stochastic discount factor vary as a linear function of the conditioning information provided by the variable z_t, $a_t = a^0 + a^1 z_t$, $b_{1,t} = b_1^0 + b_1^1 z_t$ and $b_{2,t} = b_2^0 + b_2^1 z_t$. This is a simple and familiar approach to introduce time-variation in the parameters of the utility function of conditional factor models. Then, the unconditional implications of (2) can be expressed as follows:

$$\begin{aligned} m_{t+1} &= a_t + b_t' f_{t+1} \\ &= a^0 + a^1 z_t + b_1^0 r_{m,t+1} + b_1^1 z_t r_{m,t+1} + b_2^0 q_{m,t+1} + b_2^1 z_t q_{m,t+1} \end{aligned} \tag{19}$$

Here, the following can be interpreted as the new set of factors that enter the stochastic discount factor (henceforth, SDF) equation with fixed parameters:

$$f_{t+1} = \begin{bmatrix} z_t & r_{m,t+1} & z_t r_{m,t+1} & q_{m,t+1} & z_t q_{m,t+1} \end{bmatrix}' \tag{20}$$

Inter-temporal 2-moment models imply $b_2 = 0$ in (13), whereas unconditional 3-moment models imply $a^1 = 0$, $b_1^1 = 0$ and $b_2^1 = 0$. Chen's (2002) inter-temporal model allows all the coefficients to be non zero and, under greed, RA and NIARA, it implies the following restrictions:

$$a_t = E_t(m_{t+1}) - b_t' E_t(f_{t+1}) \tag{21}$$

$$E_t(m_{t+1}) \leq 1 \tag{22}$$

$$a^1 \leq 0, b_1 < 0, b_1^0 \leq 0, b_2 \geq 0 \tag{23}$$

Campbell (1993) model with no conditional time-variation in market volatility imposes the further restrictions:

$$b_2^0 = 0, b_2^1 = 0 \tag{24}$$

6. Data

I use monthly data from 1926 to 2002 on portfolios mainly formed following Fama and French[2] (1995). In particular, I use data constructed sorting stocks of the Centre for Research on Security Prices (CRSP) database into 30 US industry portfolios, 25 size and book-to-market portfolios, an overall market portfolio, size, book-to-market and momentum portfolios. I also use monthly and quarterly returns on the 1-month and 3-month US Government Treasury Bill as proxies for the risk free rate and the

[2] I thank Kenneth French for making available on his website most of the data used in this paper.

quarterly consumption-wealth ratio per capita estimates from 1952 to 2002 produced by Lettau and Ludvigson (2001), denoted by cay_t.

7. Two-Pass Regression Estimates

I estimate both unconditional and conditional specifications of the 3M-CAPM over the period 1952-2002. Since this model nests the 2M-CAPM, I then test whether the restrictions that the latter imposes on the former are statistically and economically significant. More specifically, I estimate the following unconditional beta-pricing representation of (19) based on (15):

$$r_{i,t+1} = \alpha + \beta_{i,3}\lambda_3 + \beta_{i,4}\lambda_4 + \beta_{i,5}\lambda_5 + \beta_{i,6}\lambda_{i,6} + \beta_{i,7}\lambda_{i,7} + \varepsilon_i \qquad (25)$$

Here, the elements of the λ vector are the cross-sectional parameter estimates of average asset excess returns on the corresponding elements of the β vector. The latter are the parameters estimates of the following time series regressions:

$$\begin{aligned} r_{i,t} = \alpha_i + \beta_{i,3} z_{t-1} + \beta_{i,4} r_{m,t} + \beta_{i,5} z_{m,t-1} r_{m,t} \\ + \beta_{i,6} q_{m,t} + \beta_{i,7} z_{m,t-1} q_{m,t} + \varepsilon_{i,t} \end{aligned} \qquad (26)$$

I allow for an intercept in (25) and (26). A model that is fully successful at explaining the cross section of asset excess returns should have $\alpha = 0$. I use cay_t as the conditioning variable and thus I set $z_t = cay_t$. I estimate this model by a 2-pass procedure that involves time series and cross-sectional regressions. In the first pass, I estimate in a maximum likelihood setting, by three stage least squares (3SLS), the system of time-series regressions equations in (26) for the industries in my sample. For the sake of robustness, I do not impose any constraint on the contemporaneous covariance of the residuals nor on their variance. In the second pass of the estimation procedure, I then use my estimated beta coefficients as the regressors of average industry excess returns in a cross-sectional regression based on (25). In doing so, I correct the variance and covariance matrix of the estimates for possible cross-sectional heteroskedasticity.

The empirical results are summarised in Table 1. The coefficient of determination R^2 is just under 37 percent (24 percent adjusted). The R^2 of the unconditional 3M-CAPM is slightly lower, almost 35 percent, but it is larger once we adjust for the degrees of freedom (30 percent). All the coefficient estimates are statistically significant. The CCAPM performs considerably worse than the conditional and unconditional 3M-CAPM. Its R^2 is just 7.5 percent (the adjusted one is marginally negative) and none of the coefficients estimates, with the exception of the intercept, are statistically different from zero at conventional significance levels. Because there is considerable cross-sectional dispersion, industry returns are notoriously difficult to fit. Thus, relatively low coefficient of cross-sectional determination should not surprise and are in line with the estimates reported by Harvey and Siddique (2000). These results provide evidence that systematic asset co-skewness does help explain the cross-section of average returns. Even explicitly allowing for conditional time-variation in the shape of the utility function does not drive out its cross-sectional explanatory power.

However, both the 3M-CAPM and the 3M-CCAPM estimates imply a shape of the utility function that is incompatible with the risk aversion requirement. This can be seen by computing the elements of the b_t vector implied by the risk premia λ_t estimates and by the variance-covariance matrix of the factors. Solving (16) for b_t and using the estimates of the risk premia λ reported in Table 1 and the unconditional centred sample moments of the factors, the elements of the 3M-CAPM b_t are:

$$\begin{aligned}\begin{bmatrix} b_{1,t} \\ b_{2,t} \end{bmatrix} &= -\begin{bmatrix} Var_t(r_{m,t+1}) & Cov_t(r_{m,t+1}, q_{m,t+1}) \\ Cov_t(r_{m,t+1}, q_{m,t+1}) & Var_t(q_{m,t+1}) \end{bmatrix}^{-1} \lambda_t \\ &= -\begin{bmatrix} 0.0069 & 0.000187 \\ 0.000187 & 0.000157 \end{bmatrix}^{-1} \begin{bmatrix} 0.0096 \\ -0.0055 \end{bmatrix} \\ &= \begin{bmatrix} -2.43 \\ 38.04 \end{bmatrix}\end{aligned} \qquad (26b)$$

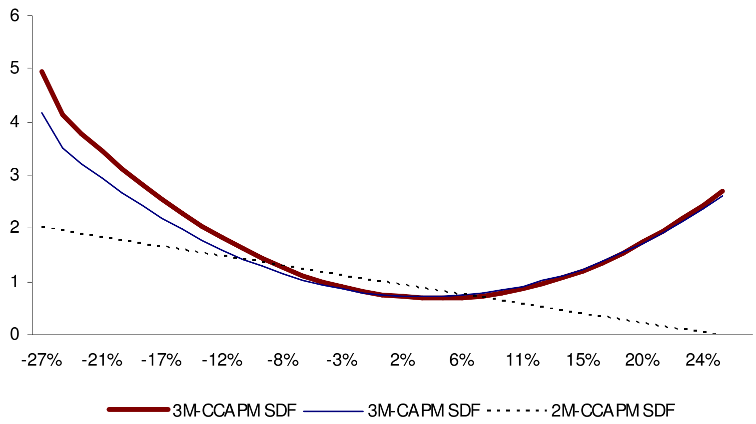
The elements of b_t for the 2M-CAPM and the 3M-CCAPM can be computed in a similar manner. I report in Figure 2 the stochastic discount factor m_{t+1} implied by the 2M-CCAPM, the 3M-CAPM and the 3M-CCAPM parameter estimates. These are consistent in all three cases with investors' non-satiation and preference for skewness. However, only the 2M-CCAPM stochastic discount factor displays risk aversion for every value taken by the market excess return over the sample. Both the 3M-CAPM and the 3M-CCAPM parameter estimates imply risk aversion only over excess returns below 1.5 percent. Above this threshold, the shape of the estimated stochastic discount factor implies risk seeking. In other words, these estimates imply an inverse S-shaped utility function.

Table 1
Factor Risk Premia Estimates
(Two-Pass Regressions)

	α	λ_3	λ_4	λ_5	λ_6	λ_7	R^2 (*Adj.* R^2)
3M-CCAPM	0.91	-0.09	0.97	-0.01	-0.59	0.002	36.9
	(1.90)	(-0.35)	(2.01)	(-1.02)	(-3.49)	(0.69)	(23.9)
3M-CAPM	0.99		0.96		-0.55		34.9
	(2.68)		(2.84)		(-3.55)		(30.0)
2M-CCAPM	1.22	0.30	0.72	0.00			7.49
	(2.54)	(1.19)	(1.61)	(0.55)			(-3.18)

Notes. This table reports percentage coefficient estimates, t-statistics (in brackets) and measures of fit (R^2 and adjusted R^2) for the 2-pass estimation of the conditional and unconditional 3M-CAPM and of the CCAPM. All the other variables are defined as in the text. The sample period is 1952-2002 and the data frequency is quarterly.

Figure 1
Estimated SDF



Notes. This Figure plots the estimated stochastic discount factor for the 3M-CCAPM, the 3M-CAPM and the 2M-CCAPM. The estimation used a 2-pass procedure with 3-Stage OLS estimates for the first step. The SDF of the conditional model is plotted for a level of the conditioning variable that corresponds to its sample average. The sample period is 1952-2002 and the data frequency is quarterly.

8. Further Results

The shape of the stochastic discount factor implied by the 3M-CAPM estimates is puzzling. To check on this result, I estimate unconditional 3M-CAPM specifications, I test the restrictions that these impose on the QMFM and compare it to estimates of the 2M-CAPM, the 4M-CAPM and the Fama and French (1995) 3-factor model (henceforth 3F-FF). I also experiment with various combinations of the factors of these models. I present estimates for both the 1952-2002 and the longer 1926-2002 sample period. In a 2-step procedure, I first regress the time series of the 30 industry portfolios excess returns on the factors allowing for an intercept in the regression equations to I estimate the factor loadings in (15a and b) and then I estimate the risk premia in (15c) using a cross sectional regression of the average portfolio returns on the factor loadings estimated in the first step.

The results are reported in Table 2. While in the 1926-2002 sample period the 3F-FF appears to fit the cross section of industry returns much better than the 2-moment and 3-moment CAPM, in the shorter sample period 1952-2002 the 3M-CAPM displays a much stronger explanatory power. However, the sign of the 3F-FF market risk premium becomes disturbingly negative thus implying a negative equity premium, while it remains positive for the 2M-CAPM and the 3M-CAPM. The coefficient of the squared market return polynomial factor in the 3M-CAPM is negative, thus satisfying a necessary condition for DIARA. Moreover, in the shorter sample period adding the squared market return factor does increase the cross sectional explanatory power, while preserving the positive sign of the market risk premium.

As shown in Figure 2, the point estimates of the factor risk premia and the sample second moments of the factors imply a well behaved shape of the investors' utility function only in the longer period, whereas this takes a theoretically unacceptable shape in the shorter sample period. In the 1926-2002 sample, $b_1 = -0.73$ and $b_2 = 0.93$, thus implying moderate risk aversion and skewness preference. In the 1952-2002 sample, $b_1 = -1.38$ implies a somewhat more pronounced risk aversion while $b_2 = 36.46$ suggests a very high rate of change of the curvature of marginal utility, rapidly decreasing risk aversion as wealth increases and thus very high preference for portfolio skewness. The speed at which risk aversion decreases in wealth implies risk seeking over a range of the market return sample realizations.

Table 2

Factor Models 2-Pass Regression Estimates

Model	r_{mt+1}	q_{mt+1}	R_{mt+1}^3-R_{ft}	**SMB**	**HML**	R^2	**σ(m)**
			Panel A 1926-2002				
2M-CAPM	0.34 (1.42)					6.7	
3M-CAPM	0.63 (1.79)	-0.07 (-0.42)				10.9	14.4
4M-CAPM	0.59 (1.16)	-0.85 (-2.21)	-0.3 (-2.23)			18.5	
3F-FF	0.59 (1.84)			0.23 (1.19)	-0.57 (-2.39)	26.8	14.6
3M-CAPM + FF	0.56 (1.49)	0.04 (0.27)		0.23 (1.13)	-0.58 (-2.28)	26.9	
			Panel B 1952-2002				
2M-CAPM	0.41 (1.06)					3.8	
3M-CAPM	0.94 (2.75)	-0.56 (-3.78)				37.9	96.7
3F-FF	-0.13 (-0.18)			0.12 (0.44)	-0.74 (-2.93)	24.8	31.6
3M-CAPM + FF	0.15 (0.23)	-0.49 (-3.28)		0.31 (1.29)	-0.84 (-3.85)	49.0	

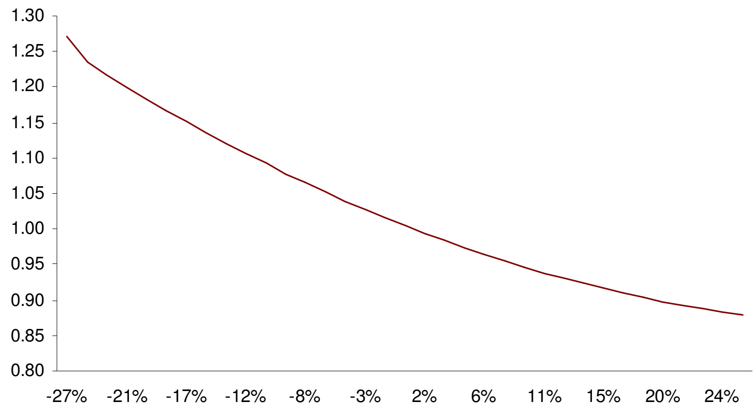
Notes. This Table reports 2-step regression estimates of the beta-pricing representation of various factor models. The top row indicates the factors included in each model. For each included factor, I report the risk premia point estimates in percentage and t-statistics in brackets. The last two columns report the coefficient of determination R^2 and the annualized volatility of the stochastic discount factor. The market Sharpe ratio is 35.2 percent in the 1926-2002 period and 40.4 percent in the 1952-2002 period. All the variables are defined as in the text. Estimates of the 4M-CAPM in the second sample period are not reported because of an extreme multi-collinearity problem between between the squared and the cubic factor (the correlation between q_{mt+1} and R_{mt+1}^3-R_{ft} is 95 percent), thus rendering the ordinary least square estimates of the factor loadings very inefficient. The data frequency is quarterly.

Figure 2

SDF

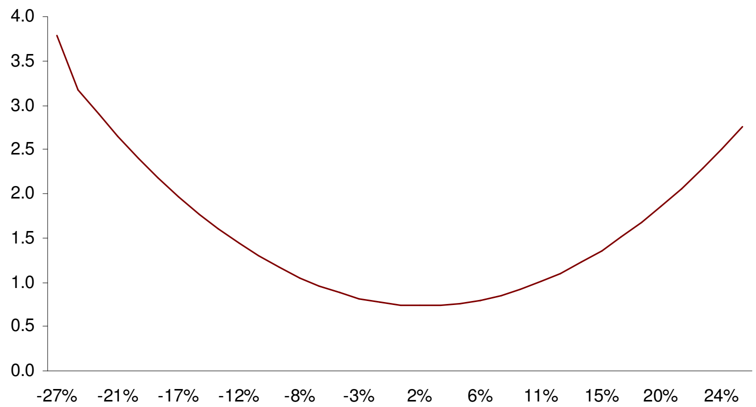
Panel A

(1927-2002)



Panel B

(1952-2002)



Notes. This Table reports the SDF implied by the 2-step regression point estimates of the 3M-CAPM obtained using quarterly data.

9. SDF Regularity Conditions

I will now derive a set of regularity conditions that ensure that the estimated parameters of the 3M-CAPM imply a shape of the investors' utility function compatible with the assumptions of greed, RA and NIARA. While similar conditions have been used by Dittmar (2002) and Post, Levy and van Vliet (2003), my formulation is more intuitive and simpler to apply than the set of conditions used by Dittmar (2002) and generalize the approach followed by Post, Levy and van Vliet (2003) to a conditional SDF setting. These conditions are easier to impose in the context of the beta-gamma representation of the 3M-CAPM proposed by Kraus and Litzenberger (1976). This is based on a standardised[3] cubic approximation of a non-polynomial utility function of the form $u(W) = W + \theta_1 W^2 + \theta_2 W^3$. Since this is an unconditional model, I first extend it to a conditional setting. I do this in Appendix I, whereas I focus here on the unconditional implications of the conditional model. These boil down to the following set of Euler equations

$$E\left[u'(R_{m,t+1} \mid \theta_t) r_{i,t+1}\right] = 0 \qquad (27)$$

with

$$u'(R_{m,t+1} \mid \theta_t) = 1 + 2\theta_{1,t} R_{m,t+1} + 3\theta_{2,t} R_{m,t+1}{}^2$$

[3] This utility function is standardized, following Post, Levy and Van Vliet (2003), such that $u(0|\theta) = 0$ and $u'(0 \mid \theta) = 1$. Since utility functions are unique up to a linear transformation, this standardization does not affect the results.

Here, the parameters of the utility function are potentially time-varying but moments are based on unconditional expectations. The orthogonality conditions corresponding to the Euler equations are equivalent to (2) with $a_1 = 1$, $b_1 = 2\theta_{1,t}$ and $b_2 = 3\theta_{2,t}$. This model imposes on the cross section of asset excess-returns a restriction similar to (10) and (15):

$$E(r_{i,t+1}) \cong \delta_1 \beta_i + \delta_2 \gamma_i \tag{28}$$

Where,

$$\delta_1 = \frac{-E\left[u''(R_{m,t+1} \mid \theta_t)\right] E\left[R_{m,t+1} - E\,(R_{m,t+1})\right]^2}{E\left[u'(R_{m,t+1} \mid \theta_t)\right]} \tag{29}$$

$$\delta_2 = \frac{-\frac{1}{2} E\left[u'''(R_{m,t+1} \mid \theta_t)\right] E\left[R_{m,t+1} - E(R_{m,t+1})\right]^3}{E\left[u'(R_{m,t+1} \mid \theta_t)\right]} \tag{30}$$

Here, β_i is the unconditional CAPM asset beta, the unconditional gamma coefficient γ_i is standardized (unconditional) asset coskewness with the market and it is defined as $E\left\{\left[r_{i,t+1} - E(r_{i,t+1})\right]\left[R_{m,t+1} - E(R_{m,t+1})\right]^2\right\} \Big/ E\left[R_{m,t+1} - E(R_{m,t+1})\right]^3$. Equation (28) thus provides a beta-gamma representation of the 3M-CAPM. This is different from my beta-pricing representation because the beta and gamma coefficients are not parameters of the regression of the asset excess returns on the factors and thus they typically must be estimated imposing and solving a set of orthogonality conditions based on (27) by GMM. Under non satiation, marginal utility must always be positive. Therefore the denominator of $\delta_{1,t}$ is always positive. Since variance is always positive, risk aversion and hence $u''(R_{m,t+1} \mid \theta_t) \leq 0$ implies $\delta_{1,t} \geq 0$. Similarly, if the skewness of the market portfolio is negative, NIARA and hence $u'''(R_{m,t+1} \mid \theta_t) \geq 0$ implies $\delta_{2,t} \geq 0$.

To model time-variation in the parameters of the utility function, I let them depend in a linear fashion on a conditioning variable z_t that represents the available information set:

$$\theta_{1,t} = \theta_3 + \theta_4 z_t \tag{31}$$

$$\theta_{2,t} = \theta_5 + \theta_6 z_t \tag{32}$$

Analytical expressions for the first, second and third derivative of utility are needed in order to be able to define the conditions for greed, NS and NIARA. I recover these by plugging (31) and (32) into (27), taking unconditional expectations and differentiating with respect to market wealth:

$$u'(R_{m,t+1} \mid \theta\ ) = 1 + 2(\theta_3 + \theta_4 z)R_m + 3(\theta_5 + \theta_6 z)R_m^2 \tag{33}$$

$$u''(R_m) = 2(\theta_3 + \theta_4 z) + 6(\theta_5 + \theta_6 z)R_m \tag{34}$$

$$u'''(R_m) = 6(\theta_5 + \theta_6 z) \tag{35}$$

In general, RA requires that $u''(R_m)$ be negative and NIARA requires $u'''(R_m)$ to be positive. Since a cubic utility function cannot be concave over its entire domain, $u''(R_m) > 0$ cannot hold for every value of the market return. Rather, it suffices to hold only over the sample values of R_m and z. Therefore, when $u'''(R_m)$ in (35) is positive and thus under NIARA, a sufficient condition for RA is the following:

$$2(\theta_3 + \theta_4 z) + 6[\theta_5 + \theta_6 Max(z)]\ Max(R_m) \leq 0 \tag{36}$$

Here, the operator $Max(\ )$ denotes the sample maximum of the argument. If θ_4 and θ_6 are not constrained to be zero, and thus in a conditional model, this condition is difficult to impose because z can take both positive and negative values and thus it is not possible to identify *a priory* the sign of both θ_3 and θ_4, on one hand, and θ_6 and θ_5 on the other hand. Under these circumstances, denoting by $Min(\ )$ the sample maximum of the argument, a set of sufficient conditions for both RA and NIARA that are relatively easy to impose is the following:

$$2[\theta_3 + \theta_4 Max(z)] + 6[\theta_5 + \theta_6 Max(z)]\ Max(R_m) \leq 0 \tag{37}$$

$$2[\theta_3 + \theta_4 Min(z)] + 6[\theta_5 + \theta_6 Max(z)]\ Max(R_m) \leq 0 \tag{38}$$

$$2[\theta_3 + \theta_4 Max(z)] + 6[\theta_5 + \theta_6 Min(z)]\ Max(R_m) \leq 0 \tag{39}$$

$$2[\theta_3 + \theta_4 Min(z)] + 6[\theta_5 + \theta_6 Min(z)]\ Max(R_m) \leq 0 \tag{40}$$

Finally, I rule out extreme solutions by imposing that the sum of beta and gamma premia equals the market risk premium, i.e

$$\delta_{1,t} + \delta_{2,t} = E(r_{m,t+1}) \tag{41}$$

This is in turn equivalent to imposing that the pricing errors are on average zero, i.e. $E\left[u'(R_{m,t+1} \mid \theta_t) r_{m,t+1}\right] = 0$, or

$$E\left[u'(R_{m,t+1} \mid \theta_t) r_{m,t+1}\right] = E\left[u'(R_{m,t+1} \mid \theta_t)\sum_{i=1}^{n} w_{i,t} r_{i,t+1}\right]$$

$$= E\left\{\sum_{i=1}^{n} w_{i,t}\left[u'(R_{m,t+1} \mid \theta_t) r_{i,t+1}\right]\right\}$$

$$= 0 \qquad (42)$$

Finally, plugging (31) and (32) into (29) and (30) and taking unconditional expectations, I recover the unconditional risk premia δ_1 and δ_2 implied by the conditional utility function parameters:

$$\delta_1 = \frac{-\left\{2\left[\theta_3 + \theta_4 E(z)\right] + 6\left[\theta_5 E(R_m) + \theta_6 E(zR_m)\right]\right\} E\left[(R_m - E(R_m)\right]^2}{1 + 2\left[\theta_3 E(R_m) + \theta_4 E(zR_m)\right] + 3\left[\theta_5 E(R_m^2) + \theta_6 E(zR_m^2)\right]} \qquad (43)$$

$$\delta_2 = \frac{-3\left[\theta_5 + \theta_6 E(z)\right] E\left[R_m - E(R_m)\right]^3}{1 + 2\left[\theta_3 E(R_m) + \theta_4 E(zR_m)\right] + 3\left[\theta_5 E(R_m^2) + \theta_6 E(zR_m^2)\right]} \qquad (44)$$

The above equations give the unconditional asset pricing implications in terms of the two risk factors, the market return and its square, of the conditional 3M-CAPM.

10. Unrestricted and Restricted GMM

Since the SDF parameters estimates for the period 1952-2002 imply a puzzling shape of the representative investor utility function, I estimate the 3M-CAPM and 3M-CCAPM imposing a well behaved utility function directly. I then test the significance of the corresponding restrictions. In empirical applications, I replace unconditional expectations by the corresponding sample moments and I estimate the 3M-CAPM and 3M-CCAPM by generalized GMM. I estimate the system of 30 orthogonality conditions in (27) with the constraint in (41) and with and without the constraints in (36) to (40). I run my estimation with instruments (the constant and a lag of the market excess return). With the instrumental variables, the set of orthogonality conditions is expanded to include the orthogonality of the pricing errors from (27) and each of the instruments. I estimate this system by iterated GMM with a continuously updating Hansen's (1982) optimal weighting matrix for the orthogonality conditions.

The empirical results are reported in Table 3. When the RA and NIARA constraints are not explicitly imposed, the unconditional 3M-CAPM estimates imply a sizeable gamma premium, even larger than the beta premium. When I impose the RA and NIARA constraints, however, the annualized gamma premium estimate collapses to 0.65 percent. This is small relative to the beta premium, almost 6.12 percent. This confirms the result reported by Dittmar (2002) and by Post, Levy and van Vliet (2003) even though my gamma premium estimate is somewhat larger than reported by the latter authors[4]. The relative magnitude of the beta and gamma premia and the impact of the RA and NIARA constraints are similar for the conditional specification.

To compare the magnitude of the covariance and coskewness premia implied by these estimates and by the estimates of my beta-pricing representation, I compute the values of the b_t parameters of the stochastic discount factor in (3) that correspond to the θ point estimates. To do this, I use the fact that, from (27), $b_1 = 2\theta_{1,t}$ and $b_2 = 3\theta_{2,t}$, with $\theta_{1,t}$ and $\theta_{2,t}$ defined as in (31) and (32). In the case of the unconditional 3M-CAPM, $\theta_{4,t} = 0$ and $\theta_{6,t} = 0$. Thus, for the model with no RA and NIARA constraints estimated by GMM, the elements of b_t are the following

$$b_{1,t} = 2\theta_{1,t} = 2(\theta_{3,t} + \theta_{4,t} z) = 2\theta_{3,t} = -3.90 \tag{45}$$

$$b_{2,t} = 3\theta_{2,t} = 3(\theta_{5,t} + \theta_{6,t} z) = 3\theta_{5,t} = 57.93 \tag{46}$$

Similarly, for the model estimated imposing RA and NIARA, the elements of b are:

$$b_{1,t} = 2\theta_{3,t} = -2.34 \tag{47}$$

$$b_{2,t} = 3\theta_{5,t} = 4.86 \tag{48}$$

[4] This is because I estimate using q_t as the second factor whereas Post, Levy and Van Vliet (2003) employ the squared market excess return. Using the squared market excess return instead of q_t results in a pricing kernel that is an approximation of the 3M-CAPM stochastic discount factor. This is because the conditionally risk free rate should be subtracted after taking the conditional expectation of the squared market return since marginal utility, in the 3M-CAPM, is defined over wealth and its square.

I do not report the corresponding elements of the b_t vector for the conditional models because these are time varying. However, the average value over the sample period of the elements of b_t is relatively close to the corresponding unconditional value[5]. Thus, comparing the SDF parameters in (45) and (46) with those implied by the 2-pass regression estimates reported in Table 1, the relative magnitude of the parameters that drive local risk aversion and attitude towards coskewness is similar. When RA and NIARA are imposed, (47) and (48) imply a lower degree of local risk aversion but also much less preference towards coskewness. Thus, both the 2-pass regression and the unconstrained GMM estimates imply too much coskewness preference for RA to hold over the entire sample range of market returns.

The differences between the OLS and GMM estimates can be ascribed to the estimation procedure and they effectively exemplify the relative merits and drawbacks of the two estimators. OLS typically weights the pricing errors of each asset equally, whereas GMM, in a quest for efficiency, typically under-weights the moments with the largest sampling error because the information that they carry is relatively unreliable. In doing this, however, it might place too much importance in pricing assets with little capitalization simply because of their low sampling error. My application of GMM, that uses a continuously updating optimal weighting matrix for the moment conditions, does precisely this. OLS (or GMM with the identity matrix in place of the weighting matrix) is instead a more robust estimation procedure.

Table 3
GMM-IV Estimates

Model	Constraints	*df*	*TJ*	θ_3	θ_4	θ_5	θ_6	δ_1 (%)	δ_2 (%)
3M-CAPM		58	58.79	-1.95		19.31		1.08	5.46
			[.446]	[.000]		[.000]			
3M-CCAPM		56	58.81	-1.90	-167.90	16.50	208.10	1.17	5.36
			[.372]	[.000]	[.000]	[.000]	[.114]		
3M-CAPM	NS, RA,	58	61.01	-1.17		1.62		5.91	0.63
	NIARA		[.368]	[.000]		[.000]			
3M-CCAPM	NS, RA,	56	60.25	-1.20	-0.00	1.65	-0.00	6.02	0.64
	NIARA		[.324]	[.000]	[.508]	[.000]	[.822]		

Notes. This table reports the GMM-IV estimation results for various sets of orthogonality conditions that correspond to the 3M-CAPM and to the 3M-CCAPM. The instruments in the GMM-IV estimation are a constant and the lagged market excess return. The symbol *df* denotes degrees of freedom (number of orthogonality conditions in excess of the number of parameters to be estimated). The expression *TJ* is *T* (the sample size) times Hansen's (1982) *J* statistic and it is distributed as a Chi-Squared with degrees of freedom equal to the number of over-identifying restrictions (*DF*). All the other variables are defined as in the text. The risk premia δ are annualised. Significance levels of t-statistics appear in brackets. The sample period is 1952-2002 and the data frequency is quarterly.

11. The Coskewness Puzzle

The considerable explanatory power of coskewness in the cross section of asset returns coupled with the ill-behaved shape of the representative investor's utility function that it implies, represent a puzzling conundrum. Numerous contributions from the literature on non standard utility theory and behavioural asset pricing, see for a review Shefrin (2005), admit a non linear pricing kernel that implies non concavity of the utility function over certain ranges of wealth. Similarly, active stock traders appear to play negative-sum games and their behavior can sometimes be interpreted as 'gambling' (see Statman (2002)). In addition, psychologists led by

[5] The time series of the b_t estimates for the conditional models is available from the author upon

Kahneman and Tversky (1979) find experimental evidence for local risk seeking behavior. Friedman and Savage (1948) and Markowitz (1952) argue that the willingness to purchase both insurance and lottery tickets implies that marginal utility is increasing over a range. See Hartley and Farrell (2001) and Post and Levy (2002) for a recent discussion.

However, if the utility function is non-concave, expected utility is not guaranteed to be quasi-concave and thus first order conditions like (1) are not guaranteed to pinpoint the maximum of investors' expected utility functions. The reason for this is the mathematical fact that the sum of concave functions is guaranteed to be quasi-concave[6] whereas the sum of quasi-concave functions is not guaranteed to be quasi-concave. Thus only a concave utility function can guarantee that expected utility is quasi-concave. In a constrained optimization problem, a stationary point is guaranteed to represent a maximum only when the objective function is quasi-concave. Lacking the quasi-concavity requirement, the parameters of marginal utility growth that satisfy the first order conditions $0 = E_t(m_{t+1}r_{i,t+1})$ for each investor are not guaranteed to represent the constrained maximum of his expected utility function.

It cannot be ruled out that, while the utility function is non-concave, expected utility does turn out to be quasi-concave (even though it is not guaranteed to do so). Then the parameters of the representative investor's stochastic discount factor that satisfy the first order conditions $0 = E_t(m_{t+1}r_{i,t+1})$ do represent the constrained maximum of

request.

his expected utility function. However, it is more natural to accept that b_{2t} is considerably smaller that the value implied by the 2-pass and unrestricted GMM estimates for the period 1952-2002 and that it is closer to the values implied either by the long-run 2-pass estimates for the period 1926-2002 or by the restricted GMM estimates for the period 1952-2002. This in turn implies that the coskewness premium is small relative to the covariance premium.

Once we accept to restrict the magnitude of b_{2t} to rule out risk seeking, two important issues arise. The first one involves understanding to what extent coskewness matters in asset pricing under this restriction and the second is to explain why coskewness successfully manages to capture a large portion of the cross-sectional variation of asset returns in spite of the small value of b_{2t} that is theoretically admissible under the 3M-CAPM. I will now address these two questions in turn.

12. The Role of Coskewness

If we restrict the elements of b_t to take the values implied by the restricted GMM estimates over the period 1952-2002, from (47) and (48) $b_{1,t} = -2.44$ and $b_{2,t} = 5.01$. From (15c), these imply the following risk premia:

[6] This is not the case for concave functions. The sum of concave functions is guaranteed to be concave. This is why the concavity of utility guarantees the concavity of the expected utility function.

$$\lambda_t = -\begin{bmatrix} Var_t(r_{m,t+1}) & Cov_t(r_{m,t+1}, q_{m,t+1}) \\ Cov_t(r_{m,t+1}, q_{m,t+1}) & Var_t(q_{m,t+1}) \end{bmatrix} \begin{bmatrix} -2.34 \\ 4.86 \end{bmatrix}$$
$$= -\begin{bmatrix} 0.0068 & 0.0002 \\ 0.0002 & 0.0002 \end{bmatrix} \begin{bmatrix} -2.34 \\ 4.86 \end{bmatrix} \tag{49}$$
$$= \begin{bmatrix} 0.0150 \\ -0.003 \end{bmatrix}$$

On an annualized basis, the market risk premium is 6.01 percent whereas the quadratic market factor premium is -0.13 percent. Thus the former is almost 48 times as large as the latter.

An alternative and more illuminating way of describing the importance of coskewness is to use represent expected asset returns, according to (14), as a function of the utility function parameters and of the asset covariance with the factors and rewrite it as follows:

$$\begin{aligned} E_t\left(r_{i,t+1}\right) &= -Cov_t(r_{i,t+1}, r_{m,t+1})b_{1,t} - Cov_t(r_{i,t+1}, q_{m,t+1})b_{2,t} \\ &= -Corr_t(r_{i,t+1}, r_{m,t+1})\sigma_{i,t}\sigma_{m,t}b_{1,t} - Corr_t(r_{i,t+1}, q_{m,t+1})\sigma_{i,t}\sigma_{q_m,t}b_{2,t} \end{aligned} \tag{50}$$

The standard deviation of $q_{m,t+1}$, about 2.5 percent on an annualized basis, is much smaller than market volatility, about 16.5 percent per annum in the 1952-2002 period. Thus the latter is almost 7 times as large as the former. Since the absolute value of $b_{2,t}$ is, even under NS, RA and NIARA, almost twice as large as the absolute value of $b_{1,t}$, for the asset coskewness premium to be as large, in absolute value, as the market coskewness premium the asset correlation with $q_{m,t+1}$ must be just over 3 times as large as the correlation with the market. For example, this would be the case if $Corr_t(r_{i,t+1}, r_{m,t+1}) = 0.3$ and $Corr_t(r_{i,t+1}, q_{m,t+1}) = 1$.

As this simple example shows, the coskewness premium can be important in explaining the expected return on certain assets and strategies even when $b_{2,t}$ and thus the price of coskewness risk is modest. Examples of such strategies are hedge funds. Amongst the industries in my sample, however, correlation with $q_{m,t+1}$ is much smaller than correlation with the market. The Mines industry index has the highest the ratio of these two correlation, just over 36 percent, followed by the Steel, Construction and Coal indices, just above 20 percent in all three cases. Thus, for the Mines industry index the covariance premium is just over 8 times as large as the coskewness premium, while it is 15 times as large for the Steel, Construction and Coal indices. Of course, for individual stocks within these indices the relative importance of the coskewness premium could be higher. Particular portfolios also display a high ratio, in absolute value, of correlation with $q_{m,t+1}$ to correlation with the market. For example, the SMB factor correlation with the market excess return is 25 percent while its correlation with $q_{m,t+1}$ is just above -10 percent. Thus the ratio of its correlation with $q_{m,t+1}$ to its correlation with the market is equal to -40 percent. More interestingly, the UMD momentum factor has a larger, in absolute value, correlation with $q_{m,t+1}$ than with the market, -11 and -6.74 percent respectively. Thus their ratio is 1.63 and the size of the UMD coskewness premium is, in absolute value, almost half the size of its covariance premium. From this perspective, it is not surprising that momentum strategies earn abnormal returns relative to those predicted by the CAPM.

13. Coskewness and the Cross-Section of Returns

From a 3M-CAPM point of view it is unlikely that coskewness proxies for the sensitivity to factors that predict future returns because including cay_t as a conditioning variable makes little difference. Thus it is unlikely that coskewness proxies for the sensitivity to changes to the investment opportunity set as in Merton's (1973), Campbell's (1993) and Chen's (2002) models. However, if stock market wealth is only a portion of the investors' overall wealth, a plausible reason for why coskewness explains such a large portion of the cross-section of average returns in the 1952-2002 period is that it proxies for omitted non market and background risk factors. The interesting issue is then the identification of these factors. I leave this however for future research.

14. An Alternative Solution

To tackle the coskewness puzzle, we can proceed along an alternative direction by accepting that it is difficult both to pinpoint the exact functional form of the utility function of traded wealth and to identify the factors for which the squared market excess return might proxy. Instead, we can focus on making the most, for pricing purposes, of the information conveyed by the squared market excess return to explain the cross-section of average returns while ruling out obvious arbitrage opportunities.

Such an approach recognises that coskewness is an asset characteristic that empirically is known, as shown in Table 2, to successfully explain a considerable portion of the cross-section of asset returns. Clearly, eliminating the requirement that the full set of NS, RA and NIARA conditions hold and, more generally, dispensing with the representative investor assumption means adopting a multi-factor, no-arbitrage perspective along the lines of Ross' (1976) APT rather than the CAPM and its extensions such as the 3M-CAPM.

The QMFM is a potential candidate as a multifactor APT-style model of asset returns. However, in the 1952-2002 period this model achieves its impressive performance thanks to an unduly high volatility of the corresponding m_{t+1}, denoted by $\sigma(m)$. The latter is estimated by taking the sample standard deviation of m_{t+1}, given the sample realizations of the factors and the point estimates of the model parameters. As shown in the last column of Table 2, the unconditional volatility of m_{t+1} for the 3M-CAPM and thus for the QMFM is three times as high as the $\sigma(m)$ of the 3F-FF model and more than twice as large as the market Sharpe ratio (henceforth, SR).

Even though the SR is not an exhaustive criterion to rank risky alternatives in a non mean-variance world, such a high SR does suggest that the QMFM is fitting noise rather than genuine asset pricing patterns. If this was not the case, it would be really surprising that so many investors actually hold the market portfolio instead of taking advantage of the much more attractive SR offered by portfolios perfectly correlated with the SDF. The very high absolute value of the correlation between the SDF and the market excess return suggests that it is unlikely that this difference in SR be due to exposure to extreme states of the world that the average investors particularly dislikes. It appears sensible therefore to bound the σ(m) of the QMFM. For example, we might wish to impose as the upper bound a somewhat higher threshold than the SR of the market portfolio to allow for the possibility of background risk that the representative investor wishes to shed. However, even an upper bound equal to the market SR would allow for an higher σ(m) than the 3F-FF model and thus for the possibility of a better explanatory power. This bound can be imposed as a moment condition in a GMM framework. Equivalently, and more simply, one could search directly for the values of the elements of b that satisfy the volatility bound and than, given the sample second moments of the factors, solve for the λ using (16). This is equivalent to estimating by GMM with restrictions. Once the λ that satisfy the restrictions have been computed, they can be multiplied by the asset beta coefficients estimated using the familiar time series regressions to give the expected return on the assets. Regressing actual returns on expected returns, it is then possible to evaluate the explanatory power of the restricted model.

Following this method, I start from the estimates of λ and of the sample moments of the $r_{m,t+1}$ market return and $q_{m,t+1}$ factors reported in (26b). These imply that the annualized risk premium associated with the squared market return factor, i.e. $\lambda_{2,t} = -2.2$ percent, is large in absolute value relative to the annualized price of the market risk premium, i.e. $\lambda_{1,t} = 3.8$ percent. Because of the ill-behaved shape of the utility function implied by the 3M-CAPM with these point estimates of the model parameters, I reduce σ(*m*) by decreasing b_2 from 38.04 to 6. The resulting σ(*m*) is 40.5 percent and thus approximately equal to the SR of the stock market portfolio (incidentally, the SDF is now well behaved and displays NS, RA and NIARA over the whole range of the sample market excess return realizations). However, the price of the coskewness factor is also largely reduced, i.e. $\lambda_{2,t} = -0.2$ percent annualized. Conversely, the market risk premium is now substantially higher, i.e. $\lambda_{1,t} = 6.3$ percent per annum, and it represents almost 100 percent of the sample average of the excess return on the market. With these values of the factors risk premia, the R^2 of the cross section of average excess returns is only 6.3 percent. Visual inspection of a scatter diagram of average excess returns against expected excess returns however reveals that the culprit for such a dismal performance is an outlier, corresponding to the average excess return on the Smoke industry index. Dropping this outlier increases the cross sectional fit to 25 percent.

Imposing an upper bound on σ(*m*) equal to the SR of the stock market portfolio is unfair on the QMFM in a non mean variance setting, since it is plausible that there are strategies that the average investor considers unappetizing because they pay off poorly in particularly bad states of the world. I therefore increase b_2 from 6 to 12. The resulting σ(*m*) is 45.8 percent and thus approximately just 5 percent higher than the SR of the stock market portfolio. The corresponding annualized factor risk premia are $\lambda_{1,t} = 5.8$ percent and $\lambda_{2,t} = -0.6$ percent. The correlation between expected asset returns, computed plugging these λ estimates in (10), and the sample averages of asset realized excess returns is 34.2 percent. The square of this figure, i.e. 11.7 percent, is the cross-sectional coefficient of determination, i.e. the R^2, of the QMFM with the volatility bound. Dropping the Smoke industry outlier increases cross sectional correlations of expected and average excess returns to 56 percent and the R^2 of the QMFM to 32 percent. Looser values of the upper bound would yield a better cross-sectional fit.

15. Main Findings and Conclusions

In this paper, I update the evidence provided by Harvey and Siddique (2000) and by Dittmar (2002) on the ability of the coskewness and gamma premia to explain the cross-section of US industry returns. My sample spans 50 years from 1952 to 2002, whereas the sample period of the studies of Harvey and Siddique (2000) and Dittmar (2002) stops, respectively, in 1993 and 1995. I also employ cay_t as a conditioning variable to model time variation in the parameters of the utility function. This variable had not been used by Dittmar (2002). Relative to Harvey and Siddique (2000), the main innovation of this study is an explicitly conditional empirical specification of the stochastic discount factor and the derivation of the beta-pricing representation of their model. More importantly, I show that this beta-pricing representation of the 3M-CAPM is a special case of a quadratic market factor model and I specify the restrictions that the former imposes on the latter. My beta pricing representation of Harvey and Siddique (2000) 3M-CCAPM is also different from Kraus and Litzenberger (1976) beta-gamma representation since neither their betas nor their gammas are regression coefficients. Their beta and gamma premium must be recovered estimating the utility function parameters.

While the parameter estimates of the unrestricted quadratic market factor model cannot be interpreted as the parameter estimates of the 3M-CAPM because they would imply some risk seeking over gains and an inverse S-shaped utility function, it is surprisingly successful at explaining the cross section of industry returns, with a coefficient of determination between 20 and 30 percent. These values are high for a model that does not include among the regressors portfolios returns that mimic additional and partially ad-hoc factors such as size and the book to market ratio.

However, even after constraining the utility function to display a well behaved shape and thus limiting the size of the coskewness premium, I show that the latter is still important in explaining the expected return of assets such as hedge funds and, more importantly, strategies that mimic the momentum factor. Thus, while the 3M-CAPM cannot explain a large portion of the cross section of returns, it does improve the pricing of particular, highly non-linear strategies relative to the simple 2M-CAPM and even relative to the Fama and French (1995) 3-Factor model.

Alternatively, in a state-preference framework, we can treat the market return and its square as factors that help explain the cross-section of asset returns and rule out arbitrage opportunity or unduly high Sharpe ratios by bounding the volatility of the implied stochastic discount factor. I show that the quadratic market factor model retains its cross-sectional explanatory power even after imposing this sensible constraint. Further research might suitably expand the set of conditioning variables to better model variation in the utility function parameters. This, while improving the fit of the model, might even lead to find a specification that requires a milder violation of RA.

Appendix I

The 3M-CAPM, in the formulation proposed by Kraus and Litzenberger (1976) and used, among others, by Post, Levy and van Vliet (2003), is an unconditional model based on the Euler equations for a cubic Taylor expansion of a standardised[7] admissible utility function $u(W) \cong W + \theta_1 W^2 + \theta_2 W^3$. In a conditional setting, the corresponding Euler equation for the determination of equilibrium expected rates of returns is the following:

$$E_t\left[u'(R_{m,t+1} \mid \theta_t) r_{i,t+1}\right] = 0 \tag{I.1}$$

where

$$u'(R_{m,t+1} \mid \theta_t) = 1 + 2\theta_{1,t} R_{m,t+1} + 3\theta_{2,t} R_{m,t+1}^2$$

Here, the parameters of the utility function are potentially time varying. The orthogonality conditions in (I.1) are equivalent to (2) with $a_1 = 1$, $b_1 = 2\theta_1$ and $b_2 = 3\theta_2$. This model imposes on the cross section of asset excess-returns a restriction similar to (10) but with time invariant utility function parameters. From (I.1):

$$E_t(r_{i,t+1}) = -\frac{Cov_t[u'(R_{m,t+1} \mid \theta_t),\ r_{i,t+1}]}{E_t[u'(R_{m,t+1} \mid \theta_t)]}$$

$$= -\frac{E_t\{[u'(R_{m,t+1} \mid \theta_t) - E_t(u'(R_{m,t+1} \mid \theta_t))][r_{i,t+1} - E_t(r_{i,t+1})]\}}{E_t[u'(R_{m,t+1} \mid \theta_t)]} \qquad \text{(I.2)}$$

Now, consider the following Taylor expansion of marginal utility around the point $u'[E_t(R_{m,t+1} \mid \theta_t)]$:

$$\begin{aligned} u'(R_{m,t+1} \mid \theta_t) - u'[E_t(R_{m,t+1} \mid \theta_t)] &\cong u''(R_{m,t+1} \mid \theta_t)[R_{m,t+1} - E_t(R_{m,t+1})] \\ &+ \frac{1}{2}u'''(R_{m,t+1} \mid \theta_t)[R_{m,t+1} - E_t(R_{m,t+1})]^2 \end{aligned} \qquad \text{(I.3)}$$

Using (I.3) and assuming that $u'[E_t(R_{m,t+1} \mid \theta_t)] \cong E_t[u'(R_{m,t+1} \mid \theta_t)]$, we can rewrite (I.2) as follows:

$$\begin{aligned} E_t(r_{i,t+1}) = &-\frac{E_t\{[u''(R_{m,t+1} \mid \theta_t)[R_{m,t+1} - E_t(R_{m,t+1})]][r_{i,t+1} - E_t(r_{i,t+1})]\}}{E_t[u'(R_{m,t+1} \mid \theta_t)]} \\ &-\frac{1}{2}\frac{E_t\{[u'''(R_{m,t+1} \mid \theta_t)[R_{m,t+1} - E_t(R_{m,t+1})]^2][r_{i,t+1} - E_t(r_{i,t+1})]\}}{E_t[u'(R_{m,t+1} \mid \theta_t)]} \end{aligned} \qquad \text{(I.4)}$$

[7] This utility function is standardized, following Post, Levy and Van Vliet (2003), such that $u(0|\theta) = 0$ and $u'(0 \mid \theta) = 1$. Since utility functions are unique up to a linear transformation, this standardization does not affect the results.

Finally, multiplying and dividing the first and second term on the right-hand side of this equation by $E_t\left[R_{m,t+1} - E_t(R_{m,t+1})\right]^2$ and $E_t\left[R_{m,t+1} - E_t(R_{m,t+1})\right]^3$, respectively, and re-arranging, we can write[8]:

$$E_t(r_{i,t+1}) \cong \delta_{t,1}\beta_{i,t} + \delta_{t,2}\gamma_{i,t} \tag{I.5}$$

Where,

$$\delta_{1,t} = \frac{-E_t\left[u''(R_{m,t+1} \mid \theta_t)\right]E_t\left[R_{m,t+1} - E_t(R_{m,t+1})\right]^2}{E_t\left[u'(R_{m,t+1} \mid \theta_t)\right]} \tag{I.6}$$

$$\delta_{2,t} = \frac{-\frac{1}{2}E_t\left[u'''(R_{m,t+1} \mid \theta_t)\right]E_t\left[R_{m,t+1} - E_t(R_{m,t+1})\right]^3}{E_t\left[u'(R_{m,t+1} \mid \theta_t)\right]} \tag{I.7}$$

Here, $\beta_{i,t}$ is the CAPM asset beta, the gamma coefficient $\gamma_{i,t}$ is standardized coskewness and it is defined as $\dfrac{E_t\left\{\left[R_{i,t+1} - E_t(R_{i,t+1})\right]\left[R_{m,t+1} - E_t(R_{m,t+1})\right]^2\right\}}{E_t\left[R_{m,t+1} - E_t(R_{m,t+1})\right]^3}$.

[8] I also assume that the second derivative of the utility function does not depend on the interaction between market and asset unexpected returns and that the third derivative does not depend on the interaction between squared market unexpected returns and asset unexpected returns, i.e. that

$$Cov_t\left\{u''(R_{m,t+1} \mid \theta), \left[r_{i,t+1} - E_t(r_{i+1})\right]\left[R_{m,t+1} - E_t(R_{m,t+1})\right]\right\} = 0$$

and

$$Cov_t\left\{u'''(R_{m,t+1} \mid \theta), \left[r_{i,t+1} - E_t(r_{i,t+1})\right]\left[R_{m,t+1} - E_t(R_{m,t+1})\right]^2\right\}$$

These are very useful and reasonable simplifications that, intuitively, correspond to the requirement that absolute risk aversion and preference towards skewness do not depend on the relation between a single asset and the market portfolio or its square (rather, they should depend only on the latter, i.e. the market return and its square). Essentially, only changes in overall wealth and in its volatility should determine moves along the utility function and, therefore, changes in the point at which its derivatives are evaluated.

Bibliography

Arditti, F.D., 1967, Risk and the Required Return on Equity, Journal of Finance 22, 19-36.

Arrow, K. J., 1971, Essays in the Theory of Risk Bearing, Markham Publishing, Chicago.

Barone Adesi, G., Gagliardini, P. and G. Urga, 2004, Testing Asset Pricing Models with Coskewness, Journal of Business and Economic Statistics 22, 474-485.

Campbell, J.Y., 1993, Intertemporal Asset Pricing Without Consumption Data, American Economic Review 83, 487-512.

Carhart, M., 1997, On Persistence on Mutual Fund Performance, Journal of Finance 52(1), 57-82.

Chamberlain, G. and M. Rothschild, 1983, Arbitrage, Factor Structure and Mean Variance Analysis in Large Asset Markets", Econometrica 51, 1281-1301.

Chen, J., 2002, Intertemporal CAPM and the Cross-Section of Stocks Returns, unpublished manuscript.

Cochrane, J.H., 2001, Asset Pricing, Princeton University Press, Princeton.

Dittmar, R.F., 2002, Non Linear pricing Kernels, Kurtosis Preference, and Evidence from the Cross-Section of Equity Returns, Journal of Finance 57, 369-402.

Fama E. and K. French, 1995, Size and Book-to-Market Factors in Earnings and Returns, Journal of Finance 50, 131-155.

Fama E. and J. MacBeth, 1973, Risk, Return, and Equilibrium: Empirical Tests, Journal of Political Economy 81, 607-636.

Friedman, M., and L. Savage, 1948. The Utility Analysis of Choices Involving Risk, Journal of Political Economy 56, 279-304.

Friend, I. and R. Westerfield, 1980. Coskewness and Capital Asset Pricing. Journal of Finance 35, 897-913.

Goetzmann, J., Ingersoll, J., Spiegel, M. and I. Welch (2004), Sharpening Sharpe Ratios, Yale School of Management ICF Working Paper 02-08.

Hansen, L.P., 1982, Large Sample Properties of Generalized Method of Moments Estimators, Econometrica 50, 1029-1054.

Hansen, L.P. and S.F. Richard, 1987, The Role of Conditioning Information in Deducting Testable Restrictions Implied by Asset Pricing Models, Econometrica 55, 587-614.

Hartley, R. and L. Farrell, 2001, Can Friedman-Savage Utility Functions Explain Gambling?, forthcoming in American Economic Review.

Harrison, J.M., and D.M. Kreps, 1979, Martingales and Arbitrage in Multiperiod Securities Markets, Journal of Economic Theory 20, 381–408.

Harvey, C. R. and A. Siddique, 2000, Conditional Skewness in Asset Pricing Tests, Journal of Finance 55, 1263-1295.

Kahneman, D. and A. Tversky, 1979, Prospect Theory: An Analysis of Decision under Risk, Econometrica 47, 263-291.

Kimball, M.S. (1990), Precautionary Savings in the Small and in the Large, Econometrica 58, 589-611.

Kraus, A. and R. H. Litzenberger, 1976, Skewness Preference and the Valuation of Risky Assets, Journal of Finance 31, 1085-1100.

Lettau M. and S. Ludvigson, 2001, Resurrecting the (C)CAPM: A Cross-Sectional Test When Risk Premia are Time-Varying, Journal of Political Economy 109, 1238-1287.

Lettau, M. and S. Ludvigson, 2001, Consumption, Aggregate Wealth and Expected Stock Returns, Journal of Finance 56, 815-849.

Levy, H., 1969, A Utility Function Depending on the First Three Moments, Journal of Finance (September).

Luengerger, D.G. (1993), A Preference Foundation for Log Mean-Variance Criteria in Portfolio Choice Problems", Journal of Economic Dynamics and Control 17, 887-906.

Markowitz, H., 1952, The Utility of Wealth, Journal of Political Economy 60, 151-158.

Post, T. and H. Levy, 2002, Does Risk Seeking Drive Asset Prices?, Timbergen Institute Discussion Paper, TI 2002-070/2.

Post, T., H. Levy and P. van Vliet, 2003, Risk Aversion and Skewness Preference, ERIM Report Series, No ERS-2003-009-F&A, http://ssrn.com/abstract=411660.

Ross, S.A., 1976, Arbitrage Theory of Capital Asset Pricing, Journal of Economic Theory 13, 341-60.

Shefrin, H., 2005, A Behavioral Approach to Asset Pricing, Elsevier Academic Press.

Statman, M., 2002, Lottery Players/Stock Traders, Financial Analysts Journal January/February, 14-21.

Treynor, J.L. and Mazuy, K., 1996, Can Mutual Funds Outguess the Market, Harvard Business Review 43(1), 63-75.

Tsiang, S., 1972, The Rational of the Mean-Standard Deviation Analysis, Skewness Preference and the Demand for Money, American Economic Review (June).