A Three-Moment Intertemporal Capital Asset Pricing Model: Theory and Evidence*

Omrane Guedhami
Assistant Professor of Finance
Faculty of Business Administration
Memorial University of Newfoundland
St. John's, NF Canada A1B 3X5
Phone: 709-737-4747
Fax: 709-737-7680
E-mail: guedhami@mun.ca

Oumar Sy♦
Assistant Professor of Finance
Dalhousie University
6100 University Avenue
Halifax, NS Canada B3H 3J5
Phone: 902-494-3849
Fax 902- 494-1107
E-mail: oumar.sy@dal.ca

* The authors would like to thank Kodjovi Assoé, Narjess Boubakri, Ines Chaieb, Sadok El Ghoul, Stefano Mazzotta, and Jeffrey Pittman. The usual disclaimer applies. The authors acknowledge financial support from the Social Sciences and Humanities Research Council (SSHRC).

♦ Presenting author.

A Three-Moment Intertemporal Capital Asset Pricing Model: Theory and Evidence

ABSTRACT

In this paper, the aggregate consumption function is characterized as a nonlinear function of the market and hedge factors to derive a three-moment intertemporal capital asset pricing model that prices coskewness and embeds the classical ICAPM and three-moment CAPM as special cases. The model is applied to examine the time-series behavior of the market, bond, size, value, and momentum premia as well as the cross-section of size and book-to-market sorted portfolio returns. We find strong evidence that the model performs better in explaining the variation of returns than the ICAPM and three-moment CAPM.

KEYWORDS: intertemporal risk; conditional coskewness

EFM classification code: 310

1. Introduction

Due to its practical and intuitive nature, the Capital Asset Pricing Model (CAPM) of Sharpe (1964) and Lintner (1965) has aroused the interest of both practitioners and academics. However, the model continues to attract substantial debate, mainly because the empirical literature has documented numerous violations of its predictions. For instance, several studies have demonstrated that firm characteristics such as size and book-to-market ratio do explain expected returns (Banz, 1981; Rosenberg et al., 1985) while market beta alone appears to be irrelevant (Fama and French, 1992).

Given the need for an adequate asset-pricing model in financial economics, the shortcomings of the CAPM have prompted new research on alternative ways to price securities. This literature generally comprises two broad strands. The first strand—which targets the CAPM's implication that market covariance suffices to explain the expected returns—investigates whether linear multifactor models explain the risk-return relationship better than the CAPM does (see, among others, Chen et al., 1986; Fama and French, 1993, 1996; Turtle et al., 1994; He et al., 1996; Jagannathan and Wang, 1996; Scruggs, 1998). Building on Merton's (1973) Intertemporal CAPM (ICAPM) or Ross's (1986) Asset Pricing Theory (APT), these studies generally find that non-market factors are priced. The second strand of the literature extends the mean-variance framework of the CAPM to take into account higher-order moments of returns in the pricing of securities. For instance, Harvey and Siddique (2000a) derive a conditional version of the Three-Moment CAPM (TM-CAPM) of Kraus and Litzenberger (1976) to find that the next conditional moment after the variance, conditional skewness, explains expected returns in the United States.[1]

Although these studies have made their cases that the risk-return relationship is potentially multivariate or nonlinear, to the best of our knowledge, no *formal* asset-pricing model exists in the literature that integrates both extensions of the CAPM. For instance, no study has evaluated the potential nonlinear relation between asset returns and hedge factors.

Most extant research that investigates the potential effect of the higher moments of returns (see, e.g., Harvey and Siddique (2000a) and the numerous references therein), albeit conditional, are static because they overlook the intertemporal dimension of the risk-return relationship. According to Dumas and Solnik (1995, p.474), "a conditional static asset-pricing model is internally inconsistent." The authors point out that if a model is conditional, then it should also be intertemporal since investors anticipate shifts in investment opportunities and hedge them over their lifetime.

In Section 2 of this paper, we attempt to fill this gap in the literature by deriving a nonlinear-multifactor model that expands the consumption function of a representative investor to allow additional coskewness terms to enter the pricing process, resulting in the Three-Moment ICAPM (TM-ICAPM). According to the TM-ICAPM, returns are not influenced solely by their conditional covariance with the market and hedge factors, but also by their conditional coskewness with the market and hedge factors. The derived model is theoretically sound, and embeds the CAPM, TM-CAPM, and ICAPM as special cases.

In Section 3, the data and methodology used to test the TM-ICAPM are described. In particular, we provide both unconditional and conditional descriptive statistics on the behavior of the first three moments of the market, bond, size, value, and momentum premia observed in the United States, with the specific aim of showing that conditional skewness is a rather nontrivial and dynamic phenomenon. Next, we outline the conditional methodology used for testing the models. We build on the instrumental variables approach of Campbell (1987) and Harvey (1989) to model the dynamic nature of the conditional comoments (covariance and coskewness). The method involves extracting the first moment of returns from a vector of instruments, and then using the resulting forecasting errors to compute the comoments and to test the asset-pricing models.

In Section 4, the simplest variant of the TM-ICAPM, which particularizes the long-term government bond excess return as the single hedge term, is used to investigate whether both covariance and coskewness help explain the time-series and cross-section of U.S. returns from

July 1963 through December 2004. The model is an extension of the two-factor ICAPM estimated by Turtle et al. (1994) and Scruggs (1998) to the pricing of coskewness risk and holds the classical two- and three-moment CAPMs as special cases.

We find that market covariance and coskewness risks are reliably priced in all specifications. Moreover, the long-term government bond excess return, taken as a hedge term, significantly and nonlinearly influences the time-series and cross-section of the portfolio returns. This is important because many of the proposed extensions to the CAPM are given as first-order approximations without duly investigating the role of possible nonlinearities. More importantly, all tests consistently show coskewness risk to be more significant than covariance risk in explaining the variation of returns. Overall, the TM-ICAPM is remarkably successful in capturing the variation of returns of various securities, and performs much better than the ICAPM and TM-CAPM.

In Section 5, we conclude that a multivariate and nonlinear risk-return relationship exists in the United States and recommend the use of a nonlinear intertemporal model to price securities.

There are at least three main contributions in this study. The first is to extend Merton's (1973) ICAPM to the pricing of coskewness risk. Since under Merton's framework—with the assumptions of continuous trading and the Markov structure of the stochastic processes—the higher moments become redundant, we directly expand the intertemporal marginal rate of substitution (IMRS) as a nonlinear function of the market and hedge factors, using the indirect relationship between these state variables and the level of aggregate consumption. The second contribution is to clarify the conditions under which Kimball's (1990) measure of *absolute prudence* determines the reward-to-conditional coskewness risk. Prior research finds that the price of coskewness risk will be negative as long as the absolute prudence is positive in the following three particular cases: (*a*) when a two-period economy is assumed (Harvey and Siddique, 2000a); (*b*) when investment opportunities are constant (Dittmar, 2002); or (*c*) when the marginal propensity to consume is independent of wealth (Errunza and Sy, 2005). We show that the price of coskewness risk cannot be signed in general. The third contribution, using a new empirical specification that does not need to stipulate the coskewness dynamics, is to document that the TM-ICAPM performs better than the nested linear (CAPM and ICAPM) and nonlinear (TM-CAPM) models in explaining returns.

This study adds to a rapidly growing strand of the financial economics literature that is aimed at improving the performance of the CAPM. On methodological grounds, the closest study to this work is perhaps that of Harvey and Siddique (2000b). The authors use a maximum likelihood framework to estimate simultaneously the conditional moments of the U.S. and world market portfolios and their equilibrium prices. However, Harvey and Siddique focus on skewness rather than coskewness, and assume a particular functional form for this moment, while we use conditional coskewness risk without having to make *a priori* assumptions about its dynamic. Using a bivariate EGARCH specification, Scruggs (1998) investigates the significance of the assumption of constant investment opportunity set on the market risk premium, and shows that the puzzling negative market risk-return relationship documented in the literature is due to the omission of the hedge term. This study goes beyond the covariance risk to include the coskewness dimension, which has also been proven to explain some of the negative episodes of expected market risk premia (Harvey and Siddique, 2000b; Guedhami and Sy, 2005).

2. The model

The first order condition describing the investor's optimal consumption and investment decisions (Hansen and Jagannathan, 1991) offers a framework that unifies the standard asset-pricing models:

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

where E_t is the conditional expectations operator conditioning on time t information, R_{pt+1} is the return on portfolio p at time $t+1$, and $m_{t+1} = u'(c_{t+1})/u'(c_t)$ represents the investor's

IMRS. Eq. (1) conceptualizes the idea that, in equilibrium, the marginal cost of postponing consumption at time t to buy an extra unit of portfolio p, $u'(c_t)$, must equate with the expected marginal utility of selling the holdings and consuming the proceeds at $t+1$, $E_t[(1+\mathrm{R}_{pt+1})u'(c_{t+1})]$.

Expanding Eq. (1) and assuming the existence of a conditionally riskless asset (with return R_{ft+1}), the expected portfolio excess return can be written as:

$$E_t[r_{pt+1}] = -(1+E_t[\mathrm{R}_{ft+1}])Cov_t[r_{pt+1}, m_{t+1}], \tag{2}$$

where $r_{pt+1} = \mathrm{R}_{pt+1} - \mathrm{R}_{ft+1}$ is the portfolio excess return at $t+1$ and Cov_t stands for the conditional covariance operator. It follows from Eq. (2) that systematic risk ensues from the covariation with the IMRS. A risk premium is required if the portfolio is negatively correlated with the IMRS, because the portfolio is then expected to earn less than average when it is needed the most.

It is apparent that Eq. (2) embraces various asset-pricing models, which differ by the proxies they use for the IMRS. Assuming the existence of a representative agent (Lucas, 1978; Breeden, 1979) allows us to express the IMRS as a function of aggregate consumption, C_{t+1}. The point of the CAPM and most of the existing asset-pricing models is to avoid the use of consumption data, and to use wealth or the rate of return on wealth instead. This substitution of wealth for consumption can be achieved by assuming: (*a*) a static setting and allow Eq. (1) to hold conditionally, so that consumption and wealth will be equivalent (Dittmar, 2002); (*b*) a two-period economy, so that consumption equals wealth in the second period (Harvey and Siddique, 2000); or (*c*) a log utility function, so that consumption will be proportional to wealth (Cochrane, 2001).

Rather than restricting the setting or particularizing the utility function, we use an alternative and more direct approach that consists of modeling the aggregate consumption as a function of state variables. As an illustration of this new approach, it can be shown that the conditional CAPM can be obtained by assuming that the classical Keynesian consumption function holds; i.e., $C_{t+1} = \phi_{0t} + \phi_{m1t} W_{mt+1}$, where ϕ_{0t} is the autonomous consumption and ϕ_{1t} is the marginal propensity to consume out of wealth.[2] To see this, use this consumption function and expand the IMRS as Taylor series around $W_{mt+1} = W_{mt}$, and obtain: $m_{t+1} = 1 + (W_{mt+1} - W_{mt})\phi_{m1t} u''(C_t)/u'(C_t)$. Using the definition of the return on aggregate wealth, the IMRS simplifies to $m_{t+1} = 1 - a_{m1t} r_{mt+1}$, where $a_{m1t} = -\phi_{m1t} W_{mt} u''(C_t)/u'(C_t)$ measures the investor's *relative risk aversion*. The CAPM is obtained by substituting the IMRS in Eq. (2).

According to Merton (1973), when the investment opportunity set is dynamic, the investor's optimal choice of consumption and investment will be a function of the state variables that help predict shifts in investment opportunities. Merton's ICAPM can be derived if we assume that the aggregate consumption is linear in both the aggregate wealth portfolio (W_{mt+1}) and K (hedge) portfolios (W_{kt+1}) that are most closely correlated with the state variables. However, since the true consumption function is unobservable, the hypothesis of a linear functional form relation can be restrictive (see, for example, Carroll and Kimball, 1996). Therefore, as an alternative to a linear specification, we assume that the consumption function is nonlinear in the state variables. That is to say $C_{t+1} = C(\boldsymbol{W}_{t+1})$, where $\boldsymbol{W}_{t+1} \equiv (W_{mt+1}, W_{1t+1}, ..., W_{Kt+1})$. Using the more general consumption function and expanding the IMRS as second-order series around $\boldsymbol{W}_t$ gives:[3]

$$m_{t+1} = a_{0t+1} - a_{m1t} r_{mt+1} - \sum_{k=1}^{K} a_{k1t} r_{kt+1} - a_{m2t} r_{mt+1}^2 - \sum_{k=1}^{K} a_{k2t} r_{kt+1}^2 - 2\sum_{k=1}^{K}\left(a_{mkt} r_{mt+1} + \sum_{j>k}^{K} a_{jkt} r_{jt+1} \right) r_{kt+1}, \qquad (3)$$

with the following economic values for the parameters:

$$
\begin{aligned}
a_{0t+1} &= 1+\varphi_2(\mathbf{W}_{t+1}) \\
a_{m1t} &= -C_m(\mathbf{W}_t)W_{mt}u''(C_t)/u'(C_t) \\
a_{k1t} &= -C_k(\mathbf{W}_t)W_{kt}u''(C_t)/u'(C_t) \\
a_{m2t} &= -\tfrac{1}{2}\{C_{mm}(\mathbf{W}_t)u''(C_t)+C_m(\mathbf{W}_t)^2u'''(C_t)\}(W_{mt})^2/u'(C_t) \\
a_{k2t} &= -\tfrac{1}{2}\{C_{kk}(\mathbf{W}_t)u''(C_t)+C_k(\mathbf{W}_t)^2u'''(C_t)\}(W_{kt})^2/u'(C_t) \\
a_{mkt} &= -\tfrac{1}{2}\{C_{mk}(\mathbf{W}_t)u''(C_t)+C_m(\mathbf{W}_t)C_k(\mathbf{W}_t)u'''(C_t)\}W_{mt}W_{kt}/u'(C_t) \\
a_{jkt} &= -\tfrac{1}{2}\{C_{jk}(\mathbf{W}_t)u''(C_t)+C_j(\mathbf{W}_t)C_k(\mathbf{W}_t)u'''(C_t)\}W_{jt}W_{kt}/u'(C_t)
\end{aligned}
\quad , \qquad (4)
$$

where r_{mt+1} and r_{kt+1} are respectively the returns on the market and the k^{th} hedge portfolios at $t+1$, φ_2 denotes the second-order Lagrange remainder of the expansion, subscripts of C denote partial derivatives with respect to the aggregate wealth and the hedge portfolios.

Substituting Eq. (3) into Eq. (2) produces the TM-ICAPM which states that, in equilibrium, the expected portfolio excess return is a function of the conditional covariance with the market and hedge factors, the conditional coskewness with the market and hedge factors, and the conditional covariance with the various cross-terms:

$$
\begin{aligned}
E_t[r_{pt+1}] = \lambda_{p0t} &+ \lambda_{m1t}Cov_t[r_{pt+1},r_{mt+1}] + \sum_{k=1}^{K}\lambda_{k1t}Cov_t[r_{pt+1},r_{kt+1}] \\
&+\lambda_{m2t}Cov_t[r_{pt+1},r_{mt+1}^2] + \sum_{k=1}^{K}\lambda_{k2t}Cov_t[r_{pt+1},r_{kt+1}^2] \\
&+\sum_{k=1}^{K}\lambda_{mkt}Cov_t[r_{pt+1},r_{mt+1}r_{kt+1}] + \sum_{k=1}^{K}\sum_{j>k}^{K}\lambda_{jkt}Cov_t[r_{pt+1},r_{jt+1}r_{kt+1}]
\end{aligned}
\quad , \qquad (5)
$$

where $\lambda_{p0t}=(1+E_t[\mathrm{R}_{ft+1}])Cov_t[r_{pt+1},a_{0t+1}]$ is the model's intercept, which represents the aggregated effect of the terms of order higher than skewness in the approximation.[4] The parameters $\lambda_{m1t}=(1+E_t[\mathrm{R}_{ft+1}])a_{m1t}$ and $\lambda_{k1t}=(1+E_t[\mathrm{R}_{ft+1}])a_{k1t}$ are, respectively, the equilibrium prices of market and hedge covariance, $\lambda_{m2t}=(1+E_t[\mathrm{R}_{ft+1}])a_{m2t}$ and $\lambda_{k2t}=(1+E_t[\mathrm{R}_{ft+1}])a_{k2t}$ are the prices of market and hedge coskewness, and $\lambda_{mkt}=2(1+E_t[\mathrm{R}_{ft+1}])a_{mkt}$ and

$\lambda_{jkt} = 2(1+E_t[R_{ft+1}])a_{jkt}$ are the prices associated with the various interactions between the market and/or the hedge factors.

Since $1+E_t[R_{ft+1}]$ is always positive, it follows that the signs of these rewards-to-risk will depend on the shapes of the investor's marginal utility and marginal propensity to consume functions. Given that C_m and C_k are *marginal propensities to consume*, the parameters a_{m1t} and a_{k1t} can be interpreted as relative risk aversion measures. If the principle of nonsatiety with respect to consumption holds ($u' > 0$) and the investor is risk averse ($u'' < 0$), then C_m and C_k will suffice to determine the signs of λ_{m1t} and λ_{k1t}, respectively. Since the aggregate wealth is a risky portfolio, it follows that $C_m > 0$ so that λ_{m1t} is positive; i.e., the risk-averse investor will require a premium to bear market covariance risk. In contrast, λ_{k1t} will be negative when investing on the hedge portfolio W_k smoothes the level of consumption; i.e., when $C_k < 0$.

The higher-order parameters are more interesting to analyze. Notice that their expressions in Eq. (4) contain the third-order derivative of the investor utility function (u'''). According to Pratt (1964), a good utility function must exhibit a decreasing *absolute risk aversion*, which is sufficient to restrict this third-order derivative to be positive (Arditti, 1967).[5] In the particular case of a two-period economy, the investor consumes everything at the end ($C_m = 1$ and $C_{mm} = 0$) so that λ_{m2t} will be negative as long as the absolute prudence of the risk-averse investor is positive, consistent with Harvey and Siddique (2000a).[6] The same relationship also holds when it is instead assumed that investment opportunities are constant (Dittmar, 2002) or that the marginal propensity to consume C_m is independent of wealth (Errunza and Sy, 2005). Another interesting particular case is observed under Keynes's (1936) *absolute income hypothesis*; i.e., when the investor consumes a smaller proportion of the aggregate wealth as it increases. In such a case, we have $C_{mm} < 0$ so that λ_{m2t} will be negative when the investor's level of absolute prudence is positive.

Still, in general, the signs of the various higher-order parameters are not solely determined by the level of the investor's absolute prudence, since they also depend on the marginal propensities to consume. From Eq. (4), we easily verify that λ_{m2t} will be negative if and only if the absolute prudence is higher than the ratio of the partial derivative of the marginal propensity to consume with respect to W_m to its squared value, $-u''' / u'' > C_{mm} / (C_m)^2$. Therefore, λ_{m2t} can be positive even when the investor's coefficient of absolute prudence is positive; viz., when $C_{mm} / (C_m)^2 > -u''' / u'' > 0$. The signs of λ_{k2}s are similarly determined by the *locus* of $C_{kk} / (C_k)^2$ relative to the coefficient of absolute prudence.

From Eq. (5), it is clear that the TM-ICAPM embeds most of the existing asset-pricing models. For example, the ICAPM is the particular case obtained when coskewness risk is not priced. In contrast to the ICAPM, the TM-ICAPM implies that the expected return on a risky asset will differ from the riskless rate even when it has zero covariance with the market and hedge factors. The CAPM makes a further restriction that covariation with any hedge term does not matter. In addition, the TM-ICAPM extends the TM-CAPM of Harvey and Siddique (2000b) to the pricing of intertemporal risk; i.e., the TM-ICAPM collapses to the TM-CAPM when a constant investment opportunity set is assumed.

3. Data and methodology

3.1. Data and summary statistics

This study considers U.S. monthly data that run from July 1963 to December 2004 (498 observations). The beginning of the sample is set to coincide with the beginning of the period examined by Fama and French (1992, 1993). Two types of variables are used: portfolios and instruments. Five instruments, motivated by prior research (Fama and Schwert, 1977; Keim and Stambaugh, 1986; Fama and French, 1988, 1989), are used to predict returns. These include: (*a*) a constant; (*b*) TB3M: the three-month Treasury bill yield; (*c*) DIV: the dividend yield of the S&P composite common stock; (*d*) DEFAULT: the spread between Moody's Baa and Aaa yields; and (*e*) TERM: the spread between the ten-year Treasury bond and the three-month Treasury bill yields.

Two sets of portfolios are used. The first set comprises five premia: (*a*) the market premium, which is the value-weight return on all NYSE, AMEX, and NASDAQ stocks minus the one-month Treasury bill rate; (*b*) the long-term government bond excess return, which is the return on the 30-year government bond minus the one-month Treasury bill rate; (*c*) the size premium (small minus big, SMB); (*d*) the value premium (high minus low, HML); and (*e*) the momentum premium (up minus down, UMD). These premia will be used in the time-series analysis. The second set of portfolios comprises the 25 Fama-French size and book-to-market sorted portfolios, which will be used to investigate the cross-sectional performance of the TM-ICAPM.[7]

Table 1 provides summary descriptive statistics on the five portfolios for 1963:7-2004:12. All the variables are expressed as continuously compounded monthly returns. Panel A shows the unconditional moments. We find that the market, value, and momentum premia are significant while the unconditional means of the size and bond premia are statistically indistinguishable from zero at the 5% level. The most interesting results are observed on the third and fourth moments, which are generally significant. Indeed, the unconditional skewness is statistically significant at the 1% level for all of the series except the value premium while the unconditional kurtosis is significant for all of the premia, to the extent that the null of normality of these series is always rejected at the conventional levels of statistical significance.

TABLE 1 ABOUT HERE

Several authors have documented that excess returns on stocks (Ferson and Harvey, 1991, 1993) and bonds (Campbell, 1987; Ilmanen, 1995) are partially predictable from available information. In light of this evidence, non-exploitation of the predictability could bias the tests. We follow Harvey and Siddique (2000b) and test whether the instruments help to explain the conditional mean, variance, and skewness of the portfolios using the following system of three equations:

$$\begin{pmatrix} E_t[r_{pt+1}] = Z_t\gamma_{p1} + \rho_{p1}r_{pt} \\ E_t[r^2_{pt+1}] = Z_t\gamma_{p2} + \rho_{p2}r^2_{pt} \\ E_t[r^3_{pt+1}] = Z_t\gamma_{p3} + \rho_{p3}r^3_{pt} \end{pmatrix}, \quad (6)$$

where r_p denotes the returns on the market, bond, size, value, or momentum premium, Z represents the instruments, and the γ's and ρ's stand for the parameters used to predict the moments of returns. Panel B of Table 1 investigates the dynamic nature of the first three moments of returns. For all the premia, the null hypothesis that the mean, variance, and skewness are constant over time is rejected at the 1% level. This result justifies our use of the conditional method to conduct the empirical tests.

3.2. Methodology

Hitherto, the empirical literature has adopted three approaches to model conditional covariance and/or test asset-pricing models. The first approach (Harvey, 1989, 1991) consists in modeling the first moment of returns, and then using the forecasting errors to construct the covariances and test the model's prediction using the Generalized Method of Moments (GMM) of Hansen (1982). The main advantage of this approach is that it allows the retrieving of various components of risk premia, but the downside is that it becomes unsustainable when the model's dimension broadens (given the proliferation of orthogonality conditions).

The second approach (Dumas and Solnik, 1995; Dittmar, 2002) consists in parameterizing solely the IMRS and then solving the resulting system using GMM. This has the advantage of limiting the number of parameters to estimate, and hence can easily accommodate larger models. One important drawback of this approach, however, is that it does not allow gauging of the various components of risk premia. Finally, the third approach (De Santis and Gerard, 1998) consists of modeling the covariances using ARCH-like specifications and then estimating the resulting system using Maximum Likelihood. The main advantage of this approach is its robustness. However, as argued by Harvey (1989, p.290), the main disadvantage of this approach is the a priori assumption made on the functional form of covariance.

In this paper, given our particular interest to gauge the various sources of systematic risk, we build on the instrumental approach derived by Harvey (1989, 1991). We begin by modeling the first moment of returns and then use the forecasting errors to construct the conditional covariance and coskewness terms without assuming a particular functional form. Doing so, though, dramatically increases the number of parameters, which ultimately renders the GMM unusable. As discussed below, we solve this problem by using Aitken's (1935) Generalized Least Squares (GLS). The price we pay in terms of efficiency for precluding GMM, however, is small compared to the gains in terms of robustness of the estimates, in addition to being able to assess the relative importance of the systematic sources of returns.

To derive the econometric system, we follow Harvey (1989) and begin by modeling the first moments of returns using the following linear filters:

$$x_{pt+1} = r_{pt+1} - (Z_t \delta_p + \kappa_p r_{pt}), \tag{7}$$

where the x's are zero-conditional mean forecasting errors on the various premia, and the δ's and κ's are parameters used by investors to derive the expected returns. With the forecasting errors defined in Eq. (7), the second and third comoment terms in the TM-ICAPM follow naturally:

$$\begin{aligned}
Cov_t[r_{pt+1}, r_{mt+1}] &= E_t[x_{pt+1} r_{mt+1}] \\
Cov_t[r_{pt+1}, r_{kt+1}] &= E_t[x_{pt+1} r_{kt+1}] \\
Cov_t[r_{pt+1}, r^2_{mt+1}] &= E_t[x_{pt+1} r^2_{mt+1}] \\
Cov_t[r_{pt+1}, r^2_{kt+1}] &= E_t[x_{pt+1} r^2_{kt+1}] \\
Cov_t[r_{pt+1}, r_{mt+1} r_{kt+1}] &= E_t[x_{pt+1} r_{mt+1} r_{kt+1}] \\
Cov_t[r_{pt+1}, r_{jt+1} r_{kt+1}] &= E_t[x_{pt+1} r_{jt+1} r_{kt+1}]
\end{aligned} . \tag{8}$$

Substituting Eq. (8) into Eq. (5) and using the linear property of the expectation operator yields the following expression for the prediction errors of the TM-ICAPM:

$$e_{pt+1} = r_{pt+1} - \lambda_{p0t} - \left(\begin{array}{c} \lambda_{m1t} x_{pt+1} r_{mt+1} + \sum_{k=1}^{K} \lambda_{k1t} x_{pt+1} r_{kt+1} \\ + \lambda_{m2t} x_{pt+1} r_{mt+1}^2 + \sum_{k=1}^{K} \lambda_{k2t} x_{pt+1} r_{kt+1}^2 \\ + \sum_{k=1}^{K} \lambda_{mkt} x_{pt+1} r_{mt+1} r_{kt+1} + \sum_{k=1}^{K} \sum_{j>k}^{K} \lambda_{jkt} x_{pt+1} r_{jt+1} r_{kt+1} \end{array} \right), \tag{9}$$

where $E_t[e_{pt+1}] = 0$.

Merton (1973, p.789) suggests that interest rates can capture changes in investment opportunities. Following Turtle et al. (1994) and Scruggs (1998), we use the excess return on the long-term government bond as a single hedge factor ($k \equiv b$). Note that in Eq. (9) we have not specified the functional form of the prices of risk. Several authors have found that the change in reward-to-risk has a greater impact on returns than the change in risk (Ferson and Harvey, 1991). Consequently, we assume that the prices of risk vary linearly with the instruments: $\lambda_{p0t} = Z_t \ell_{p0}$, $\lambda_{m1t} = Z_t \ell_{m1}$, $\lambda_{b1t} = Z_t \ell_{b1}$, $\lambda_{m2t} = Z_t \ell_{m2}$, $\lambda_{b2t} = Z_t \ell_{b2}$, and $\lambda_{mbt} = Z_t \ell_{mb}$; where the ℓ's are weighting vectors.

Given the prominence of the size, value, and momentum effects in explaining the cross-section of returns, the current litmus test for any candidate asset-pricing model is its ability to explain these premia.[8] Therefore, we apply the TM-ICAPM and its various special cases to not only explain the market and bond excess returns but, more challenging, to shed some light on the size, value, and momentum premia.[9] The following system, which stacks together the forecasting errors (Eq. (7)) and the model's prediction errors (Eq. (9)) on the five premia, is used to estimate the various parameters of the models:

$$
\varepsilon_{pt+1} = (\varkappa_{pt+1} \quad e_{pt+1}) = \begin{pmatrix}
\left[r_{mt+1} - (Z_t\delta_m + \boldsymbol{\kappa}_m r_{mt})\right]' \\
\left[r_{bt+1} - (Z_t\delta_b + \boldsymbol{\kappa}_b r_{bt})\right]' \\
\left[r_{SMBt+1} - (Z_t\delta_s + \boldsymbol{\kappa}_s r_{SMBt})\right]' \\
\left[r_{HMLt+1} - (Z_t\delta_h + \boldsymbol{\kappa}_h r_{HMLt})\right]' \\
\left[r_{UMDt+1} - (Z_t\delta_u + \boldsymbol{\kappa}_u r_{UMDt})\right]' \\
\begin{bmatrix} r_{mt+1} - Z_t\ell_{m0} - (Z_t\ell_{m1}\varkappa_{mt+1}r_{mt+1} + Z_t\ell_{b1}\varkappa_{mt+1}r_{bt+1} \\ +Z_t\ell_{m2}\varkappa_{mt+1}r_{mt+1}^2 + Z_t\ell_{b2}\varkappa_{mt+1}r_{bt+1}^2 + Z_t\ell_{mb}\varkappa_{mt+1}r_{mt+1}r_{bt+1}) \end{bmatrix}' \\
\begin{bmatrix} r_{bt+1} - Z_t\ell_{b0} - (Z_t\ell_{m1}\varkappa_{bt+1}r_{mt+1} + Z_t\ell_{b1}\varkappa_{bt+1}r_{bt+1} \\ +Z_t\ell_{m2}\varkappa_{bt+1}r_{mt+1}^2 + Z_t\ell_{b2}\varkappa_{bt+1}r_{bt+1}^2 + Z_t\ell_{mb}\varkappa_{bt+1}r_{mt+1}r_{bt+1}) \end{bmatrix}' \\
\begin{bmatrix} r_{SMBt+1} - Z_t\ell_{s0} - (Z_t\ell_{m1}\varkappa_{st+1}r_{mt+1} + Z_t\ell_{b1}\varkappa_{st+1}r_{bt+1} \\ +Z_t\ell_{m2}\varkappa_{st+1}r_{mt+1}^2 + Z_t\ell_{b2}\varkappa_{st+1}r_{bt+1}^2 + Z_t\ell_{mb}\varkappa_{st+1}r_{mt+1}r_{bt+1}) \end{bmatrix}' \\
\begin{bmatrix} r_{HMLt+1} - Z_t\ell_{h0} - (Z_t\ell_{m1}\varkappa_{ht+1}r_{mt+1} + Z_t\ell_{b1}\varkappa_{ht+1}r_{bt+1} \\ +Z_t\ell_{m2}\varkappa_{ht+1}r_{mt+1}^2 + Z_t\ell_{b2}\varkappa_{ht+1}r_{bt+1}^2 + Z_t\ell_{mb}\varkappa_{ht+1}r_{mt+1}r_{bt+1}) \end{bmatrix}' \\
\begin{bmatrix} r_{UMDt+1} - Z_t\ell_{u0} - (Z_t\ell_{m1}\varkappa_{ut+1}r_{mt+1} + Z_t\ell_{b1}\varkappa_{ut+1}r_{bt+1} \\ +Z_t\ell_{m2}\varkappa_{ut+1}r_{mt+1}^2 + Z_t\ell_{b2}\varkappa_{ut+1}r_{bt+1}^2 + Z_t\ell_{mb}\varkappa_{ut+1}r_{mt+1}r_{bt+1}) \end{bmatrix}'
\end{pmatrix}' \tag{10}
$$

Because the error terms should be unpredictable from available information, $E_t[\varepsilon_{pt+1}] = 0$, the GMM is the natural way to estimate Eq. (10). The GMM estimates are obtained by setting the expected crossproducts of the error terms with the instruments as close to zero as possible. However, in the GMM system, the parameterization of the first moment of returns dramatically increases the number of orthogonality conditions, because we need at least as many instruments as the maximum number of parameters in any equation. Consequently, Eq. (10) yields highly volatile pricing errors when it is estimated with the GMM.

One way to solve this problem is to restrict the GMM objective function to minimize solely the squared pricing errors. This is equivalent to estimating an Ordinary Least Squares (OLS).[10] Because of the potential correlations between the errors of the different equations, we further improve the efficiency of the estimation by making GLS corrections. Specifically, we estimate the parameters by minimizing the following quadratic form: $J_{GLS} = T^{-1}\sum_{t=0}^{T-1}[\varepsilon_{pt+1}]'\hat{\Omega}^{-1}[\varepsilon_{pt+1}]$, where the cross-equation covariance matrix Ω is estimated with Zellner's (1962) SUR effects from the first stage OLS residuals.[11]

4. Empirical Illustration

While Harvey and Siddique (2000a) generalize the conditional CAPM to the three-moment framework, we consider intertemporal risk and theoretically extend the ICAPM to the pricing of coskewness risk. We estimate the simplest form of the TM-ICAPM by assuming that changes in investment opportunities are proxied by a single state variable: the long-term government bond.

4.1. Goodness-of-Fit

Table 2 investigates the goodness-of-fit of the TM-ICAPM as well as the validity of the restrictions implied by the nested asset-pricing models (CAPM, ICAPM, and TM-CAPM). In particular, we report for each premium equation the Root Mean Squared Errors (RMSE) and the adjusted *pseudo*-R^2.

Ghysels (1998) suggests the use of the RMSE to compare models. A virtual look at the RMSEs for the market, size, value, and momentum premia in Panel B of Table 2 confirms the economic significance of conditional coskewness. Among all the asset-pricing models investigated, the TM-ICAPM yields the lowest pricing errors for all the premia. The superiority of the TM-ICAPM over the CAPM, ICAPM, and TM-CAPM is even more striking when we look at the adjusted R^2s. For all premium equations, the TM-ICAPM explains more than 42% of the time-series variation (the adjusted R^2s range from 42.97% for the value premium to 62.11% for the market premium). This result is remarkable given that the linear models (CAPM and ICAPM) explain less than 7% of the premia variances. Further, the relative success of the TM-ICAPM in capturing the time series of Fama-French's portfolios confirms the recent conclusions reached by Liew and Vassalou (2000) and Petkova (2005) that these portfolios are related to risk.

It is interesting to note that the TM-CAPM performs better than the ICAPM. This evidence suggests that it is more important to ease the two-moment restriction of the CAPM than to account for intertemporal risk. However, the TM-CAPM is still outclassed by the TM-ICAPM (at the 1% level), confirming the intertemporal nature of the risk-return relationship. Overall, because of the importance of coskewness risk, the CAPM, ICAPM, and TM-CAPM—all of which overlook, to an extent, the coskewness with the market or the hedge factors—are rejected at the 1% level relative to the TM-ICAPM (see Panel A of Table 2).

TABLE 2 ABOUT HERE

4.2. Prices of Risk

Table 3 reports the GLS estimates of the prices of risk in Eq. (10). The average price of conditional market covariance is positive and significant at the 1% level ($\overline{Z_t \ell_{m1}} = 1.30$, $t = 26.23$). The price of market covariance varies with the instruments (p-value = 0.00). As can be observed from Figure 1, the fitted price of market covariance is time varying and positive in nearly 92% of cases. This result is consistent with the theory, which predicts a positive price of conditional market covariance. [12]

Extant studies suggest that the CAPM may be misspecified due to the omission of hedge terms that describe the investment opportunity set. Our results confirm this hypothesis since the long-term government bond covariance is significantly priced at the 1% level ($\overline{Z_t \ell_{b1}} = -1.50$, $t = -32.56$). The price of bond covariance—albeit unpredictable from the instruments at the 5% level—is negative most of the time (see Figure 2).

As can be seen from Figures 3 and 4, the prices of market and bond coskewness risks are time varying and, most of the time, positive. This result supports the findings of Harvey and Siddique (2000b) that the price of skewness is primarily positive (in 72.80% and 87.18% of cases) using two different specifications. The positive nature of prices of coskewness implies that constraining this reward-to-risk to be always negative could be restrictive. As alluded to above, information on the shape of the investor's utility function—in particular, absolute prudence—alone is insufficient to determine the sigh of the price of coskewness risk. In other words, the price of coskewness risk depends also on the behavior of the marginal propensity to consume and cannot, in general, be signed. We find that both the market and bond coskewness risks are significantly priced at the 1% level, implying that a nonlinear risk-return relationship exists in the United States.

TABLE 3 AND FIGURES 1 TO 4 ABOUT HERE

4.3. Variance Decomposition

The empirical approach used in this study has the important advantage of allowing us to assess the sources of systematic variation of the market, bond, size, value, and momentum premia. Using Eq. (10), we decompose each of the five premia into an intercept ($Z_t \ell_{p0}$), a cross-term ($Z_t \ell_{mb} x_{pt+1} r_{mt+1} r_{bt+1}$), and four pure systematic effects: market covariance ($Z_t \ell_{m1} x_{pt+1} r_{mt+1}$), bond covariance ($Z_t \ell_{b1} x_{pt+1} r_{bt+1}$), market coskewness ($Z_t \ell_{m2} x_{pt+1} r_{mt+1}^2$), and bond coskewness ($Z_t \ell_{b2} x_{pt+1} r_{bt+1}^2$).

Table 4 shows descriptive statistics on the components of expected risk premium. The first reported statistic is the Mean Absolute Deviation (MAD), which tells us how much each component deviates from zero (Rouwenhorst, 1999). For the market premium, we find that the component due market coskewness is the most significant in terms of risk management, with a MAD of about 0.91% per month. Market coskewness also has the most significant MAD of all premia but the bond premium, for which bond coskewness is the most important source of risk.

The importance of coskewness is further confirmed when we consider the ratio of the variance of each component to the combined sum of the components variances (%VAR).[13] Market coskewness accounts for about 69.21% of the combined variances of the market premium, followed by the cross (16.30%) and bond coskewness (6.34%) terms. Similar results are observed for the size, value, and momentum premia. For the bond premium, market coskewness continues to be important, albeit dominated by bond coskewness.

In contrast, conditional covariance has little explanatory power on the systematic variation of the market, size, value, and momentum premia. Indeed, neither covariation with the market nor with bond factor is able to explain more than 7% of the combined variances.

TABLE 4 ABOUT HERE

4.4. *Cross-Sectional Test*

The results presented so far are based on the estimation of a system that considers a set of five portfolios to be explained, namely the market, bond, size, value, and momentum premia. This section evaluates the performance of the TM-ICAPM using a larger cross-section of assets: the 25 Fama-French portfolios. These portfolios are formed from the intersection of five size portfolios and five book-to-market portfolios. Using these portfolios, we obtain the following pooled time-series cross-sectional system of equations:

$$\varepsilon_{pt+1} = (\varkappa_{pt+1} \quad e_{pt+1}) = \begin{pmatrix} \left[r_{pt+1} - Z_t \delta_p \right]' \\ \begin{bmatrix} r_{pt+1} - Z_t \ell_0 - (Z_t \ell_{m1} \varkappa_{pt+1} r_{mt+1} + Z_t \ell_{b1} \varkappa_{pt+1} r_{bt+1} + \\ Z_t \ell_{m2} \varkappa_{pt+1} r_{mt+1}^2 + Z_t \ell_{b2} \varkappa_{pt+1} r_{bt+1}^2 + Z_t \ell_{mb} \varkappa_{pt+1} r_{mt+1} r_{bt+1}) \end{bmatrix}' \end{pmatrix}', \tag{11}$$

were the time-varying intercept and prices of risk are the same for all portfolios.

Table 5 shows the results from the GLS estimation of Eq. (11). None of the previous results on the prices of risk is materially changed by the use of the 25 portfolios. In particular, conditional coskewness continues to be significantly priced, and the most important source of risk. Moreover, consistent with our previous evidence, all the nested models (CAPM, ICAPM, and TM-CAPM) are reliably rejected relative to the TM-ICAPM.[14] Furthermore, the explanatory power of the TM-ICAPM remains high with a 52.12% average adjusted R^2s across the 25 portfolios. In comparison, the average adjusted-R^2 across the 25 portfolios is about

5.24% for the CAPM, 6.08% for the ICAPM, and 35.63% for the TM-CAPM. Overall, this evidence supports that the TM-ICAPM is capable of explaining the cross-section of stock returns, and performs no worse, but better than the ICAPM or the TM-CAPM.

TABLE 5 ABOUT HERE

5. Conclusion

A Three-Moment ICAPM (TM-ICAPM) has been derived from the expansion of the intertemporal marginal rate of substitution of a representative investor as a nonlinear function of the market and hedge factors. The derived model is theoretically sound, and implies that returns are not influenced solely by their conditional covariance with the market and hedge factors, but also by their conditional coskewness with the market and hedge factors. It is shown that the equilibrium risk-return relationship, as predicted by the classic ICAPM, holds only under the very special case in which coskewness risk is not priced. When coskewness risk is priced, expected returns on risky assets may differ from the riskless rate even when they are constructed to have zero covariance with the market and the hedge factors. The TM-ICAPM is used to investigate the time-series of the market, bond, size, value, and momentum premia as well as the cross-section of the Fama-French portfolio returns. We find that the TM-ICAPM is remarkably successful in capturing the variation of returns, and performs much better than the ICAPM and two- and three-moment CAPM. In particular, the results show that conditional covariance is priced, but taken alone, this component of risk premia is unable to describe expected returns because conditional coskewness risk is also priced, and appears to be the most significant source of risk. Overall, our research implies that the existing linear asset-pricing models are misspecified and suggest the use of a nonlinear intertemporal model to price securities.

6. Appendix: Derivation of the TM-ICAPM

This section details the derivation of Eq. (3). The starting point is the assumption that the IMRS is nonlinearly influenced by the aggregate wealth and hedge portfolios. In other words, $C_{t+1} = g(\mathbf{W}_{t+1})$ and $m_{t+1} = h(C_{t+1}) = h\left[g(\mathbf{W}_{t+1})\right]$, where $h(.) = u'(.)/u'(C_t)$ and $g(.)$ are two functions that are assumed to be (at least two times) differentiable in the domains that contain the points $g(\mathbf{W}_t)$ and $\mathbf{W}_t$, respectively. Applying Lagrange-Cauchy's theorem to expand the IMRS around $\mathbf{W}_{t+1} = \mathbf{W}_t$ gives:

$$
\begin{aligned}
m_{t+1} &= h(C_t) + (W_{mt+1} - W_{mt}) \left.\frac{\partial h\left[g(\mathbf{W}_{t+1})\right]}{\partial W_{mt+1}}\right|_{\mathbf{W}_t} + \sum_k (W_{kt+1} - W_{kt}) \left.\frac{\partial h\left[g(\mathbf{W}_{t+1})\right]}{\partial W_{kt+1}}\right|_{\mathbf{W}_t} \\
&+ \frac{1}{2}(W_{mt+1} - W_{mt})^2 \left.\frac{\partial^2 h\left[g(\mathbf{W}_{t+1})\right]}{(\partial W_{mt+1})^2}\right|_{\mathbf{W}_t} + \sum_k (W_{kt+1} - W_{kt})^2 \left.\frac{\partial^2 h\left[g(\mathbf{W}_{t+1})\right]}{(\partial W_{kt+1})^2}\right|_{\mathbf{W}_t} \\
&+ \sum_k \frac{1}{2}\left\{ 2(W_{mt+1} - W_{mt}) \left.\frac{\partial^2 h\left[g(\mathbf{W}_{t+1})\right]}{\partial W_{mt+1}\partial W_{kt+1}}\right|_{\mathbf{W}_t} + \sum_{j\neq k} (W_{jt+1} - W_{jt}) \left.\frac{\partial^2 h\left[g(\mathbf{W}_{t+1})\right]}{\partial W_{jt+1}\partial W_{kt+1}}\right|_{\mathbf{W}_t} \right\}(W_{kt+1} - W_{kt}) \\
&+ \varphi_2(\mathbf{W}_{t+1})
\end{aligned} \quad . \text{(A1)}
$$

From the definition of the returns on the aggregate wealth and hedge portfolios, we have $W_{mt+1} - W_{mt} = W_{mt} r_{mt+1}$ and $W_{kt+1} - W_{kt} = W_{kt} r_{kt+1}$. Substituting the latter identities into (A1) and applying the chain rule yields:

$$
\begin{aligned}
m_{t+1} &= h(C_t) + W_{mt} r_{kt+1} \left.\frac{\partial h\left[g(\mathbf{W}_{t+1})\right]}{\partial g(\mathbf{W}_{t+1})} \frac{\partial g(\mathbf{W}_{t+1})}{\partial W_{mt+1}}\right|_{\mathbf{W}_t} + \sum_k W_{kt} r_{kt+1} \left.\frac{\partial h\left[g(\mathbf{W}_{t+1})\right]}{\partial g(\mathbf{W}_{t+1})} \frac{\partial g(\mathbf{W}_{t+1})}{\partial W_{kt+1}}\right|_{\mathbf{W}_t} \\
&+ \frac{1}{2}(W_{mt} r_{mt+1})^2 \left.\left[\frac{\partial^2 h\left[g(\mathbf{W}_{t+1})\right]}{(\partial g(\mathbf{W}_{t+1}))^2} \left(\frac{\partial g(\mathbf{W}_{t+1})}{\partial W_{mt+1}} \right)^2 + \frac{\partial h\left[g(\mathbf{W}_{t+1})\right]}{\partial g(\mathbf{W}_{t+1})} \frac{\partial^2 g(\mathbf{W}_{t+1})}{(\partial W_{mt+1})^2} \right]\right|_{\mathbf{W}_t} \\
&+ \sum_k \frac{1}{2}(W_{kt} r_{kt+1})^2 \left.\left[\frac{\partial^2 h\left[g(\mathbf{W}_{t+1})\right]}{(\partial g(\mathbf{W}_{t+1}))^2} \left(\frac{\partial g(\mathbf{W}_{t+1})}{\partial W_{kt+1}} \right)^2 + \frac{\partial h\left[g(\mathbf{W}_{t+1})\right]}{\partial g(\mathbf{W}_{t+1})} \frac{\partial^2 g(\mathbf{W}_{t+1})}{(\partial W_{kt+1})^2} \right]\right|_{\mathbf{W}_t} \\
&+ \sum_k 2 \left(\begin{aligned} &\frac{1}{2} W_{mt} r_{mt+1} \left.\left[\frac{\partial^2 h\left[g(\mathbf{W}_{t+1})\right]}{(\partial g(\mathbf{W}_{t+1}))^2} \frac{\partial g(\mathbf{W}_{t+1})}{\partial W_{mt+1}} \frac{\partial g(\mathbf{W}_{t+1})}{\partial W_{kt+1}} + \frac{\partial h\left[g(\mathbf{W}_{t+1})\right]}{\partial g(\mathbf{W}_{t+1})} \frac{\partial^2 g(\mathbf{W}_{t+1})}{\partial W_{mt+1}\partial W_{kt+1}} \right]\right|_{\mathbf{W}_t} + \\ &\frac{1}{2} \sum_{j>k} W_{jt} r_{jt+1} \left.\left[\frac{\partial^2 h\left[g(\mathbf{W}_{t+1})\right]}{(\partial g(\mathbf{W}_{t+1}))^2} \frac{\partial g(\mathbf{W}_{t+1})}{\partial W_{jt+1}} \frac{\partial g(\mathbf{W}_{t+1})}{\partial W_{kt+1}} + \frac{\partial h\left[g(\mathbf{W}_{t+1})\right]}{\partial g(\mathbf{W}_{t+1})} \frac{\partial^2 g(\mathbf{W}_{t+1})}{\partial W_{jt+1}\partial W_{kt+1}} \right]\right|_{\mathbf{W}_t} \end{aligned} \right) W_{kt} r_{kt+1} + \varphi_2(\mathbf{W}_{t+1})
\end{aligned} \quad .
$$

(A2)

Since we always have: $h(C_t)=1$; $\left.\frac{\partial h\left[g(\mathbf{W}_{t+1})\right]}{\partial g(\mathbf{W}_{t+1})}\right|_{\mathbf{W}_t}=\frac{u''(C_t)}{u'(C_t)}$; $\left.\left[\frac{\partial^2 h\left[g(\mathbf{W}_{t+1})\right]}{(\partial g(\mathbf{W}_{t+1}))^2}\right]\right|_{\mathbf{W}_t}=\frac{u'''(C_t)}{u'(C_t)}$; $\left.\frac{\partial g(\mathbf{W}_{t+1})}{\partial W_{mt+1}}\right|_{\mathbf{W}_t}=C_m(\mathbf{W}_t)$; $\left.\frac{\partial g(\mathbf{W}_{t+1})}{\partial W_{kt+1}}\right|_{\mathbf{W}_t}=C_k(\mathbf{W}_t)$; $\left.\frac{\partial^2 g(\mathbf{W}_{t+1})}{(\partial W_{mt+1})^2}\right|_{\mathbf{W}_t}=C_{mm}(\mathbf{W}_t)$; $\left.\frac{\partial^2 g(\mathbf{W}_{t+1})}{(\partial W_{kt+1})^2}\right|_{\mathbf{W}_t}=C_{kk}(\mathbf{W}_t)$; $\left.\frac{\partial^2 g(\mathbf{W}_{t+1})}{\partial W_{mt+1}\partial W_{kt+1}}\right|_{\mathbf{W}_t}=C_{mk}(\mathbf{W}_t)$; and $\left.\frac{\partial^2 g(\mathbf{W}_{t+1})}{\partial W_{jt+1}\partial W_{kt+1}}\right|_{\mathbf{W}_t}=C_{jk}(\mathbf{W}_t)$, Eq. (A2) simplifies to:

$$
\begin{aligned}
m_{t+1} = {} & 1+\varphi_2(\mathbf{W}_{t+1}) \\
& -\left[-C_m(\mathbf{W}_t)W_{mt}u''(C_t)/u'(C_t)\right]r_{mt+1} \\
& -\sum_k\left[-C_k(\mathbf{W}_t)W_{kt}u''(C_t)/u'(C_t)\right]r_{kt+1} \\
& -\left[-\tfrac{1}{2}\left\{C_{mm}(\mathbf{W}_t)u''(C_t)+C_m(\mathbf{W}_t)^2u'''(C_t)\right\}(W_{mt})^2/u'(C_t)\right](r_{mt+1})^2 \\
& -\sum_k\left[-\tfrac{1}{2}\left\{C_{kk}(\mathbf{W}_t)u''(C_t)+C_k(\mathbf{W}_t)^2u'''(C_t)\right\}(W_{kt})^2/u'(C_t)\right](r_{kt+1})^2 \\
& -2\sum_k\left[-\tfrac{1}{2}\left\{C_{mk}(\mathbf{W}_t)u''(C_t)+C_m(\mathbf{W}_t)C_k(\mathbf{W}_t)u'''(C_t)\right\}W_{mt}W_{kt}/u'(C_t)\right]r_{mt+1}r_{kt+1} \\
& -2\sum_k\sum_{j>k}\left[-\tfrac{1}{2}\left\{C_{jk}(\mathbf{W}_t)u''(C_t)+C_j(\mathbf{W}_t)C_k(\mathbf{W}_t)u'''(C_t)\right\}W_{jt}W_{kt}/u'(C_t)\right]r_{jt+1}r_{kt+1}
\end{aligned}
\quad , \qquad \text{(A3)}
$$

which is Eq. (3). QED■

References

Aitken, A. C., 1935, On least squares and linear combinations of observations, *Proceedings of the Royal Society of Edinburgh* 55, 42-48.

Arditti, F., 1967, Risk and the required return on equity, *Journal of Finance* 22, 19-36.

Bansal, R., and S. Viswanathan, 1993, No arbitrage and arbitrage pricing: A new approach, *Journal of Finance* 48, 1231-1262.

Banz, R. W., 1981, The relationship between return and market value of common stocks, *Journal of Financial Economics* 9, 3-18.

Black, F., 1976, Studies of stock price volatility changes, *Proceedings of the 1976 Meeting of American Statistical Association*, 177-181.

Blanchard, O. J., and M. W. Watson, 1982, Bubbles, rational expectations, and financial markets, in CRISES IN THE ECONOMIC AND FINANCIAL STRUCTURE, P. Wachtel, ed., Lexington Books, 295-316.

Breeden, D., 1979, An intertemporal asset pricing model with stochastic consumption and investment opportunities, *Journal of Financial Economics* 7, 265-296.

Brennan, M. J., 1993, Agency and asset pricing, *Unpublished Manuscript*, UCLA and London Business School.

Campbell, J. Y., 1987, Stock returns and the term structure, *Journal of Financial Economics* 18, 373-399.

Cao, H. H., J. D. Coval, and D. Hirshleifer, 2002, Sidelined investors, trading-generated news, and security returns, *Review of Financial Studies* 15, 651-648.

Carroll, C. D., and M. S. Kimball, 1996, On the concavity of the consumption function, *Econometrica* 64, 981-992.

Chapman, D., 1997, Approximating the asset pricing kernel, *Journal of Finance* 52, 1383-1410.

Chen, N., R. Roll, and S. A. Ross, 1986, Economic forces and the stock market, *Journal of Business* 59, 383-403.

Cochrane, J., 2001, Asset pricing, Princeton University Press, Princeton, New Jersey.

De Santis G., and B. Gerard, 1998, How big is the premium for currency risk? *Journal of Financial Economics* 49, 375-412.

Dittmar, R., 2002, Nonlinear pricing kernels, kurtosis preference, and evidence from the cross section of equity returns, *Journal of Finance* 57, 369-403.

Dumas, B., and B. Solnik, 1995, The world price of foreign exchange risk, *Journal of Finance* 50, 445-479.

Errunza, V., and O. Sy, 2005, A three-moment international asset-pricing model: Theory and evidence, *Working paper*, McGill University.

Fama, E. F., and K. R. French, 1988, Dividend yields and expected stock returns, *Journal of Financial Economics* 22, 3-26.

Fama, E. F., and K. R. French, 1989, Business conditions and expected returns on stocks and bonds, *Journal of Financial Economics* 25, 23-50.

Fama, E. F., and K. R. French, 1992, The cross-section of expected stock returns, *Journal of Finance* 47, 427-465.

Fama, E. F., and K. R. French, 1993, Common risk factors in the returns on stocks and bonds, *Journal of Financial Economics* 33, 3-56.

Fama, E. F., and K. R. French, 1996, Multifactor explanations of asset pricing anomalies, *Journal of Finance* 51, 55-84.

Fama, E. F., and G. W. Schwert, 1977, Asset returns and inflation, *Journal of Financial Economics* 5, 115-146.

Fang, H., and T. Y. Lai, 1997, Cokurtosis and capital asset pricing, *Financial Review* 32, 293-307.

Ferson, W. E., and C. R. Harvey, 1991, The variation of economic risk premiums, *Journal of Political Economy* 99, 385-415.

Ferson, W. E., and C. R. Harvey, 1993, The risk and predictability of international equity returns, *Review of Financial Studies* 6, 527-566.

Friend, I., and R. Westerfield, 1980, Coskewness and capital asset pricing, *Journal of Finance* 35, 897-913.

Ghysels, E., 1998, On stable factor structures in the pricing of risk: Do time-varying betas help or hurt? *Journal of Finance* 53, 549-573.

Green, W. H., 2000, ECONOMETRIC ANALYSIS, fourth edition, Prentice Hall, New Jersey.

Guedhami, O., and O. Sy, 2005, Does conditional market skewness resolve the puzzling market risk-return relationship*? Quarterly Review of Economics and Finance* 45, 582-598.

Hansen, L. P., 1982, Large sample properties of generalized method of moments estimators, *Econometrica* 50, 1029-1054.

Hansen, L. P., and R. Jagannathan, 1991, Implications of security market data for models of dynamic economies, *Journal of Political Economy* 99, 225-262.

Harvey, C. R., 1989, Time-varying conditional covariances in tests of asset pricing models, *Journal of Financial Economics* 24, 289-317.

Harvey, C. R., 1991, The world price of covariance risk, *Journal of Finance* 46, 111-157.

Harvey, C. R., and A. Siddique, 2000a, Conditional skewness in asset pricing tests, *Journal of Finance* 55, 1263-1295.

Harvey, C. R., and A. Siddique, 2000b, Time-varying conditional skewness and the market risk premium, *Research in Banking and Finance* 1, 27-60.

He, J., R. Kan, L. Ng, and C. Zhang, 1996, Tests of the relations among marketwide factors, firm-specific variables, and stock returns using a conditional asset pricing model, *Journal of Finance* 51, 1891-1908.

Hong, H., and J. C. Stein, 2003, Differences of opinion, short-sales constraints and market crashes, *Review of Financial Studies* 16, 487-525.

Ilmanen, A., 1995, Time-varying expected returns in international bond markets, *Journal of Finance* 50, 481-506.

Jagannathan, R., and Z. Wang, 1996, The conditional CAPM and the cross-section of expected returns, *Journal of Finance* 51, 3-53.

Jegadeesh, N., and S. Titman, 1993, Returns to buying winners and selling losers: Implications for stock market efficiency, *Journal of Finance* 48, 65-91.

Keim, D. B., and R. F. Stambaugh, 1986, Predicting returns in the bond and stock market, *Journal of Financial Economics* 17, 357-390.

Keynes, J. M., 1936, The general theory of employment, interest, and money, London, Macmillan.

Kimball, M. S., 1990, Precautionary saving in the small and in the large, *Econometrica* 58, 53-73.

Kraus, A., and R. H. Litzenberger, 1976, Skewness preference and the valuation of risky assets, *Journal of Finance* 31, 1085-1099.

Liew, J., and M. Vassalou, 2000, Can book-to-market, size and momentum be risk factors that predict economic growth? *Journal of Financial Economics* 57, 221-245.

Lim, K., 1989, A new test of the three-moment capital asset pricing model, *Journal of Financial and Quantitative Analysis* 24, 205-216.

Lintner, J., 1965, The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets, *Review of Economics and Statistics* 47, 13-37.

Lucas, R. E. Jr., 1978, Asset prices in an exchange economy, *Econometrica* 46, 1429-45.

Merton, R. C., 1973, An intertemporal asset pricing model, *Econometrica* 41, 867-888.

Petkova, R., 2005, Do the Fama-French factors proxy for innovations in predictive variables? *Journal of Finance,* Forthcoming.

Pindyck, R. S., and D. L. Rubinfeld, 1981, ECONOMETRIC MODELS AND ECONOMIC FORECASTS, Second Edition, McGraw-Hill Book Co, New York.

Pindyck, R. S., 1984, Risk, inflation, and the stock market, *American Economic Review* 74, 335-351.

Pratt, J. W., 1964, Risk aversion in the small and in the large, *Econometrica* 32, 122-136.

Rosenberg, B., K. Reid, and R. Lanstein, 1985, Persuasive evidence of market inefficiency, *Journal of Portfolio Management* 11, 9-16.

Ross, S. A., 1976, The arbitrage theory of capital asset pricing, *Journal of Economic Theory* 13: 341-360.

Rouwenhorst, K. G., 1999, European equity markets and the EMU, *Financial Analysts Journal* 55, 57-64.

Rubinstein, M., 1973, The fundamental theorem of parameter-preference security valuation, *Journal of Financial and Quantitative Analysis* 8, 61-69.

Scruggs, J. T., 1998, Resolving the puzzling intertemporal relation between the market risk premium and conditional market variance: A two-factor approach, *Journal of Finance* 53, 575-603.

Sears, S. R., and J. K. C. Wei, 1985, Asset pricing, higher moments, and the market risk premium: A note, *Journal of Finance* 40, 1251-1253.

Sharpe, W. F., 1964, Capital asset prices: A theory of market equilibrium under conditions of risk, *Journal of Finance* 19, 425-442.

Thomas, J. J., 1989, The early econometric history of the consumption function, *Oxford Economic Papers* 41, 131-149.

Turtle, H., A. Buse, and B. Korkie, 1994, Tests of conditional asset pricing with time-varying moments and risk prices, *Journal of Financial and Quantitative Analysis* 29, 15-29.

Veronesi, P., 1999, Stock market overreaction to bad news in good times: A rational expectations equilibrium model, *Review of Financial Studies* 12, 975-1007.

Zellner, A., 1962, An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias, *Journal of American Statistical Association* 57, 348-368.

Notes

1. See also the studies by Rubinstein (1973), Friend and Westerfield (1980), Sears and Wei (1985), Lim (1989), Chapman (1997), Fang and Lai (1997), Harvey and Siddique (2000b), Dittmar (2002), and Errunza and Sy (2005). Ghysels (1998) provides evidence that nonlinear models work better empirically than linear factor models. Furthermore, Bansal and Viswanathan (1993) emphasize the fact that the linear factor-pricing model does not hold if security payoffs are nonlinearly related to the factors. The presence of limited liability (Black, 1976), bubbles (Blanchard and Watson, 1982), volatility feedback effects (Pindyck, 1984), agency problems (Brennan, 1993), overreactions to news (Veronesi, 1999), transaction costs (Cao, Coval, and Hirshleifer, 2002), or short-sale restrictions (Hong and Stein, 2003) is also consistent with nonlinear payoffs and therefore incompatible with a linear risk-return relationship.

2. Because both parameters are allowed to vary over time, this consumption function embeds as particular cases most of the neo-classical models (see the appendix in Thomas (1989) for a detailed description of the consumption functions used in the literature).

3. The detailed proof of this derivation is in the Appendix. See Harvey and Siddique (2000a) for the discussion of the rationale behind the choice of a second-order approximation.

4. The fact that the intercept is not necessarily trivial in theory could explain why extant studies (e.g., Harvey, 1989, p.302; Scruggs, 1998, p.589) have empirically documented that non-inclusion of the intercept in tests can produce misleading estimates.

5. Note, however, that nonlinearity does not exclusively result from the behavior of the marginal utility of consumption. Indeed, if the marginal propensity to consume is function of W_m or W_k, the risk-return relationship will be nonlinear even if $u''' = 0$.

6. The absolute prudence, measured by $-u'''/u''$, aims to appraise the investor's precautionary saving motives (Kimball, 1990).

7. The market, Treasury bill, and long-term government bond data are from the Center for Research in Security Prices (CRSP). The data on SMB, HML, and UMD factors as well as the 25 size and book-to-market portfolios are obtained from Professor Ken French's website. DIV is taken from DRI Basic Economics while data for TB3M, DEFAULT, and TERM originate from the Federal Reserve Statistical Release (H15).

8. The size effect is the propensity of small firms to outperform large firms (Banz, 1981), the value effect is the tendency of value stocks (stocks with a high book-to-market ratio) to outperform growth stocks (Rosenberg et al., 1985), while the momentum effect is the tendency of winners (stocks that performed well in the past) to outperform losers over the medium term (Jegadeesh and Titman, 1993).

9. We present in Section 4.4 additional evidence based on a larger cross-section of returns.

10. Note, however, that in all our systems the forecasting and prediction errors are orthogonal to the instruments by construction. This follows from the modeling of the intercepts of the prediction equations.

11. The GLS estimation can be seen as a tradeoff between the robustness of the OLS and the efficiency of the GMM; see Cochrane (2001, Sections 10.2 and 11.5) for a discussion of this issue. See also Pindyck and Rubinfeld (1981, p.331-333) for a detailed discussion of SUR effects, and Green (2000, p.615-616) for the efficiency gain accruing to GLS relative to OLS.

12. In comparison, the prices of market covariance (not reported) estimated from the CAPM and the ICAPM are positive in 41.57% and 65.66% of the cases, respectively.

13. This decomposition of variance ignores the covariances of the components to focus on the pure effects.

14. For the sake of brevity these test results as well as the goodness-of-fit results are not reported, but are available from the authors upon request.

TABLE 1

Descriptive Statistics, 1963:7–2004:12

Variable	A. Unconditional moments						
	Mean	Std Dev	Skewness	Exc. Kurtosis	Minimum	Maximum	Bera-Jarque
r_m	0.470*	4.452	-0.500**	2.033**	-23.090	16.050	106.511**
r_b	0.144	2.816	0.396**	2.283**	-9.289	13.959	121.218**
r_{SMB}	0.245	3.258	0.479**	5.211**	-16.690	21.490	582.498**
r_{HML}	0.442**	2.955	0.102	2.348**	-12.030	13.750	115.213**
r_{UMD}	0.842**	4.073	-0.638**	5.335**	-25.000	18.380	624.486**

	B. Conditional moments							
	Mean		Variance		Skewness		All moments	
	χ^2	p-value	χ^2	p-value	χ^2	p-value	χ^2	p-value
r_m	20.58**	0.0010	14.32*	0.0137	15.83**	0.0073	55.56**	<.0001
r_b	9.05	0.1070	69.13**	<.0001	12.84*	0.0250	98.8**	<.0001
r_{SMB}	21.25**	0.0007	218.75**	<.0001	175.73**	<.0001	680.99**	<.0001
r_{HML}	17.12**	0.0043	84.89**	<.0001	7.08	0.2144	102.85**	<.0001
r_{UMD}	1.90	0.8634	75.62**	<.0001	2.51	0.7754	89.34**	<.0001

This table reports summary descriptive statistics for the market, bond, size, value, and momentum premia. The market premium (r_m) is the value-weight return on all NYSE, AMEX, and NASDAQ stocks in excess of the one-month Treasury bill return; the bond premium (r_b) is the return on the long-term government bond of 30-year maturity in excess of the one-month Treasury bill return; and the size, value, and momentum premia are the returns on the SMB (r_{SMB}), HML (r_{HML}), and UMD (r_{UMD}) portfolios. The statistics presented in Panel A include the first four unconditional moments of the portfolios, as well as Bera-Jarque's χ^2 statistic for the test of normality. In Panel B, the results from the Wald tests for time-variation in conditional mean, variance, or/and skewness of the various portfolios are presented. The tests are based on the GLS estimation of the following system of three equations:

$$\left(E_t[r_{pt+1}^n] = Z_t\gamma_{pn} + \rho_{pn}r_{pt}^n\right) \qquad \text{for} \qquad n = 1, 2, 3, \tag{6}$$

where r_p is r_m, r_b, r_{SMB}, r_{HML}, or r_{UMD} and Z are the five instruments, which include: a constant, TB3M: the lagged three-month Treasury bill yields, DIV: the lagged dividend yield of the world market index, DEFAULT: the spread between Moody's Baa and Aaa yields, and TERM: the spread between the ten-year Treasury bonds and the three-month Treasury bill yields. The figures reported are the χ^2 statistics along with the associated robust p-value [in brackets]. Two and one asterisks denote statistical significance at the 1% and 5% levels, respectively. The mean, standard deviation, minimum, and maximum are reported in percent per month. The sample covers 498 monthly observations (from July 1963 to December 2004).

TABLE 2

Goodness-of-Fit of the Three-Moment ICAPM and the Nested Models, 1963:2-2004:12

A. Test of the restrictions implied by the nested models on the TM-ICAPM

Null hypothesis	χ^2	*df*	*p*-value
CAPM ($H_0 : \ell_{b1} = \ell_{m2} = \ell_{b2} = \ell_{mb} = \mathbf{0}$)	2177.1**	20	<.0001
ICAPM ($H_0 : \ell_{m2} = \ell_{b2} = \ell_{mb} = \mathbf{0}$)	1962.8**	15	<.0001
TM-CAPM ($H_0 : \ell_{b1} = \ell_{b2} = \ell_{mb} = \mathbf{0}$)	714.1**	15	<.0001

B. Goodness-of-fit

Model	r_m	r_b	r_{SMB}	r_{HML}	r_{UMD}
	Root mean squared errors (RMSE)				
CAPM	4.36	2.81	3.21	2.93	4.08
ICAPM	4.31	2.78	3.18	2.90	4.03
TM-CAPM	3.20	2.52	2.76	2.46	3.36
TM-ICAPM	2.74	1.74	2.40	2.23	2.98
	Adjusted pseudo-R^2				
CAPM	4.29	0.80	2.62	1.42	-0.11
ICAPM	6.10	2.80	4.82	3.70	2.22
TM-CAPM	48.25	19.65	28.39	30.94	32.06
TM-ICAPM	62.11	61.76	45.59	42.97	46.53

Panel A tests the restrictions implied by various nested models on the TM-ICAPM using a Wald test (the χ^2 statistic along with the degrees of freedom and the robust *p*-value are reported). We test whether the prices associated with (*a*) bond covariance and all coskewness risks are jointly zero (CAPM); (*b*) all coskewness risks are jointly zero (ICAPM); and (*c*) bond covariance and coskewness risks are jointly zero (TM-CAPM). Panel B compares the TM-ICAPM with the nested special cases. For each premium equation, the root mean squared errors (RMSE) and the adjusted *pseudo*-R^2 obtained from each model are reported. All estimates are obtained from the estimation of Eq. (10). The RMSEs are reported in percent per month while the R^2s are in percent. Two asterisks denote statistical significance at the 1%. The sample runs from July 1963 to December 2004.

TABLE 3

GLS Estimation of the Three-Moment ICAPM, 1963:7-2004:12

Price of risk		Estimates					Descriptive statistics		
		Intercept	TB3M	DIV	DEFAULT	TERM	Mean (%)	%Negative	χ^2
		A. Intercepts of the premium equations							
Market	$Z_t\ell_{m0}$	-0.004 (-0.768)	-0.001** (-4.875)	0.001** (4.371)	0.002** (2.781)	-0.000 (-1.221)	0.002** (5.513)	38.76	39.49** [<.0001]
Bond	$Z_t\ell_{b0}$	-0.007* (-2.071)	0.000* (2.563)	-0.000 (-0.708)	-0.000 (-1.016)	0.001** (3.870)	0.002** (11.778)	29.92	17.24** [0.0017]
Size	$Z_t\ell_{s0}$	-0.003 (-0.690)	-0.001** (-4.877)	0.001** (4.460)	0.001** (2.724)	-0.001 (-1.772)	0.002** (5.748)	38.55	36.08** [<.0001]
Value	$Z_t\ell_{h0}$	0.003 (0.721)	0.000** (4.387)	-0.001** (-3.173)	-0.001* (-2.043)	0.001** (2.872)	0.006** (27.652)	9.84	20.64** [0.0004]
Momentum	$Z_t\ell_{u0}$	0.010* (1.981)	0.000 (0.802)	-0.000 (-0.036)	-0.001 (-1.220)	0.000 (0.629)	0.010** (128.645)	0.00	1.60 [0.8088]
		Prices of risk							
Market covariance	$Z_t\ell_{m1}$	2.523** (3.035)	0.076** (3.342)	-0.104** (-4.047)	-0.328** (-2.772)	0.075 (1.466)	1.296** (26.234)	8.03	29.32** [<.0001]
Bond covariance	$Z_t\ell_{b1}$	-0.469 (-0.339)	-0.064 (-1.612)	-0.021 (-0.314)	0.459* (2.392)	-0.142* (-1.970)	-1.504** (-32.558)	92.17	6.44 [0.1685]
Market coskewness	$Z_t\ell_{m2}$	121.927** (12.918)	-1.006** (-6.968)	-0.715* (-2.558)	2.337* (1.985)	-2.175** (-7.438)	58.318** (72.310)	0.00	105.97** [<.0001]
Bond coskewness	$Z_t\ell_{b2}$	224.798** (8.312)	-2.470** (-3.916)	0.876 (0.522)	2.718 (0.744)	-1.043 (-1.197)	145.652** (85.717)	0.20	107.11** [<.0001]
The cross-term	$Z_t\ell_{mb}$	-7.255 (-0.296)	5.188** (8.229)	-9.964** (-8.045)	0.862 (0.255)	4.624** (3.969)	7.118* (2.322)	43.17	95.83** [<.0001]

The hypothesis that conditional covariance and/or coskewness explain the market, bond, size, value, and momentum premia in the United States is tested by estimating the following system:

$$\varepsilon_{pt+1} = (x_{pt+1} \quad e_{pt+1}) = \begin{pmatrix} \left[r_{mt+1} - (Z_t\delta_m + \kappa_m r_{mt})\right]' \\ \left[r_{bt+1} - (Z_t\delta_b + \kappa_b r_{bt})\right]' \\ \left[r_{SMBt+1} - (Z_t\delta_s + \kappa_s r_{SMBt})\right]' \\ \left[r_{HMLt+1} - (Z_t\delta_h + \kappa_h r_{HMLt})\right]' \\ \left[r_{UMDt+1} - (Z_t\delta_u + \kappa_u r_{UMDt})\right]' \\ \left[r_{mt+1} - Z_t\ell_{m0} - (Z_t\ell_{m1}x_{mt+1}r_{mt+1} + Z_t\ell_{b1}x_{mt+1}r_{bt+1} + Z_t\ell_{m2}x_{mt+1}r_{mt+1}^2 + Z_t\ell_{b2}x_{mt+1}r_{bt+1}^2 + Z_t\ell_{mb}x_{mt+1}r_{mt+1}r_{bt+1})\right]' \\ \left[r_{bt+1} - Z_t\ell_{b0} - (Z_t\ell_{m1}x_{bt+1}r_{mt+1} + Z_t\ell_{b1}x_{bt+1}r_{bt+1} + Z_t\ell_{m2}x_{bt+1}r_{mt+1}^2 + Z_t\ell_{b2}x_{bt+1}r_{bt+1}^2 + Z_t\ell_{mb}x_{bt+1}r_{mt+1}r_{bt+1})\right]' \\ \left[r_{SMBt+1} - Z_t\ell_{s0} - (Z_t\ell_{m1}x_{st+1}r_{mt+1} + Z_t\ell_{b1}x_{st+1}r_{bt+1} + Z_t\ell_{m2}x_{st+1}r_{mt+1}^2 + Z_t\ell_{b2}x_{st+1}r_{bt+1}^2 + Z_t\ell_{mb}x_{st+1}r_{mt+1}r_{bt+1})\right]' \\ \left[r_{HMLt+1} - Z_t\ell_{h0} - (Z_t\ell_{m1}x_{ht+1}r_{mt+1} + Z_t\ell_{b1}x_{ht+1}r_{bt+1} + Z_t\ell_{m2}x_{ht+1}r_{mt+1}^2 + Z_t\ell_{b2}x_{ht+1}r_{bt+1}^2 + Z_t\ell_{mb}x_{ht+1}r_{mt+1}r_{bt+1})\right]' \\ \left[r_{UMDt+1} - Z_t\ell_{u0} - (Z_t\ell_{m1}x_{ut+1}r_{mt+1} + Z_t\ell_{b1}x_{ut+1}r_{bt+1} + Z_t\ell_{m2}x_{ut+1}r_{mt+1}^2 + Z_t\ell_{b2}x_{ut+1}r_{bt+1}^2 + Z_t\ell_{mb}x_{ut+1}r_{mt+1}r_{bt+1})\right]' \end{pmatrix}' \qquad (10)$$

where r_m, r_b, r_{SMB}, r_{HML}, and r_{UMD} are respectively the market, bond, size, value, and momentum premia, and Z are our five instruments. All the variables are defined in Table 1. The system is estimated with GLS. For each price of risk, the figures reported are the points estimates of the coefficient associated with each instrument and the mean value of the equilibrium price of risk with the robust t-statistics (in parentheses), the percentage of negative values, and the χ^2 statistic associated with the Wald test of the time variation in the price of risk with the robust p-value [in brackets]. Two and one asterisks denote statistical significance at the 1% and 5% levels, respectively. The sample runs from July 1963 to December 2004.

TABLE 4

Decomposition of Systematic Risk Premia, 1963:2-2004:12

Components of risk premium		r_m	r_b	r_{SMB}	r_{HML}	r_{UMD}
		A. Mean absolute deviation (MAD)				
Intercept	$Z\ell_{p0}$	0.716	0.332	0.332	0.332	0.332
Market covariance	$Z\ell_{m1}\varkappa_p r_m$	0.297	0.123	0.152	0.158	0.197
Bond covariance	$Z\ell_{b1}\varkappa_p r_b$	0.125	0.118	0.073	0.069	0.088
Market coskewness	$Z\ell_{m2}\varkappa_p r_m^2$	**0.913**	0.294	**0.388**	**0.400**	**0.525**
Bond coskewness	$Z\ell_{b2}\varkappa_p r_b^2$	0.405	**0.445**	0.235	0.199	0.278
Cross-term	$Z\ell_{mb}\varkappa_p r_m r_b$	0.340	0.180	0.151	0.153	0.179
		B. Proportion of the variance (%VAR)				
Intercept	$Z\ell_{p0}$	4.147	4.288	14.427	7.547	0.705
Market covariance	$Z\ell_{m1}\varkappa_p r_m$	3.590	2.278	4.889	5.385	6.302
Bond covariance	$Z\ell_{b1}\varkappa_p r_b$	0.421	2.540	0.690	0.630	0.758
Market coskewness	$Z\ell_{m2}\varkappa_p r_m^2$	**69.209**	25.604	**57.586**	**66.510**	**70.149**
Bond coskewness	$Z\ell_{b2}\varkappa_p r_b^2$	6.337	**53.961**	9.999	6.876	10.350
Cross-term	$Z\ell_{mb}\varkappa_p r_m r_b$	16.296	11.328	12.409	13.051	11.736

This table uses Eq. (10) to decompose the expected market, bond, size, value, and momentum premia into an intercept ($Z_t\ell_{p0}$), a cross-term ($Z_t\ell_{mb}\varkappa_{pt+1}r_{mt+1}r_{bt+1}$), and four pure systematic effects: market covariance ($Z_t\ell_{m1}\varkappa_{pt+1}r_{mt+1}$), bond covariance ($Z_t\ell_{b1}\varkappa_{pt+1}r_{bt+1}$), market coskewness ($Z_t\ell_{m2}\varkappa_{pt+1}r_{mt+1}^2$), and bond coskewness ($Z_t\ell_{b2}\varkappa_{pt+1}r_{bt+1}^2$). For each component of risk premia, we report the mean absolute deviation (MAD) and the variance relative to the sum of the variances (%VAR). The MADs are reported in percent per month while the %VARs are in percent. The maximums are in boldface. The sample runs from July 1963 to December 2004.

TABLE 5

Cross-sectional Estimation of the Three-Moment ICAPM Portfolio, 1963:7-2004:12

Price of risk		Estimates					Descriptive statistics		
		Intercept	TB3M	DIV	DEFAULT	TERM	Mean (%)	%Negative	χ^2
Intercept	$Z_t \ell_0$	0.005 (1.173)	-0.000** (-3.424)	0.000 (1.645)	0.001** (2.998)	-0.000 (-0.519)	0.006** (24.421)	10.64	56.03** [<.0001]
Market covariance	$Z_t \ell_{m1}$	2.840** (8.689)	0.022* (2.353)	-0.034** (-3.197)	-0.239** (-5.088)	0.012 (0.594)	1.191** (35.655)	7.03	69.12** [<.0001]
Bond covariance	$Z_t \ell_{b1}$	-1.170* (-2.273)	0.023 (1.592)	-0.145** (-5.858)	0.413** (5.855)	0.004 (0.130)	-0.780** (-13.703)	72.29	61.27** [<.0001]
Market coskewness	$Z_t \ell_{m2}$	107.438** (25.899)	-1.068** (-14.483)	0.055 (0.405)	0.753 (1.449)	-1.883** (-13.007)	51.456** (70.486)	0.00	342.69** [<.0001]
Bond coskewness	$Z_t \ell_{b2}$	217.026** (17.915)	-2.462** (-8.784)	0.678 (0.967)	2.156 (1.394)	-1.299** (-3.035)	126.846** (73.738)	0.60	458.25** [<.0001]
The cross-term	$Z_t \ell_{mb}$	-2.688 (-0.240)	5.102** (18.352)	-11.387** (-21.481)	5.284** (3.436)	3.407** (6.128)	-5.023 (-1.478)	51.41	567.36** [<.0001]

The hypothesis that conditional covariance and/or coskewness explain the 25 Fama-French portfolios (ranked by size and book-to-market) is tested by estimating the following system:

$$\varepsilon_{pt+1} = (x_{pt+1} \quad e_{pt+1}) = \begin{pmatrix} \left[r_{pt+1} - Z_t \delta_p \right]' \\ \left[r_{pt+1} - Z_t \ell_0 - (Z_t \ell_{m1} x_{pt+1} r_{mt+1} + Z_t \ell_{b1} x_{pt+1} r_{bt+1} + Z_t \ell_{m2} x_{pt+1} r_{mt+1}^2 + Z_t \ell_{b2} x_{pt+1} r_{bt+1}^2 + Z_t \ell_{mb} x_{pt+1} r_{mt+1} r_{bt+1}) \right]' \end{pmatrix} \quad (11)$$

Where r_p, r_m, and r_b are respectively the portfolio excess return, and the market and bond premia, and Z are our five instruments. The system is estimated with GLS. For each price of risk, the figures reported are the points estimates of the coefficient associated with each instrument and the mean value of the equilibrium price of risk with the robust t-statistics (in parentheses), the percentage of negative values, and the χ^2 statistic associated with the Wald test of the time variation in the price of risk with the robust p-value [in brackets]. Two and one asterisks denote statistical significance at the 1% and 5% levels, respectively. The sample runs from July 1963 to December 2004.

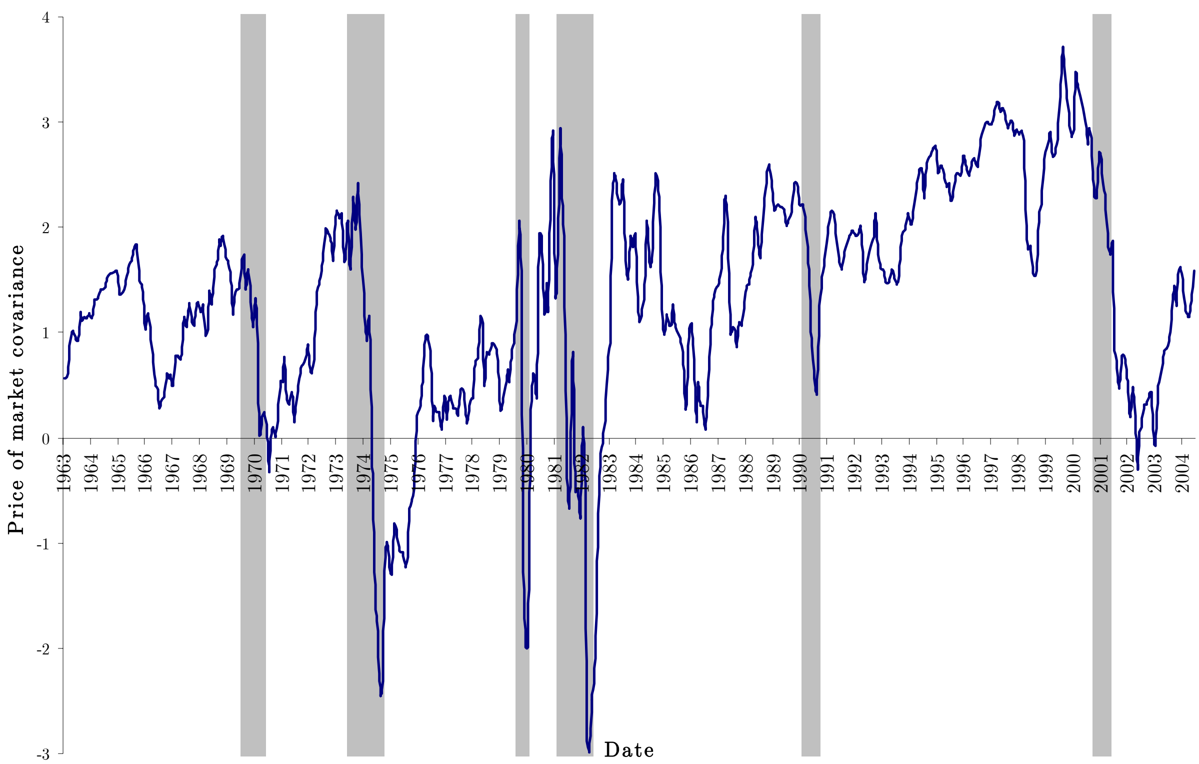

Figure 1. Time variation of the prices of market covariance, 1963:7–2004:12. This graph plots the time series of the estimated prices of market covariance, which is specified as a linear function of the lagged instruments. The estimates are obtained from Eq. (10). NBER recession dates are plotted in the shadowed area. The sample covers 498 monthly observations (from July 1963 to December 2004).

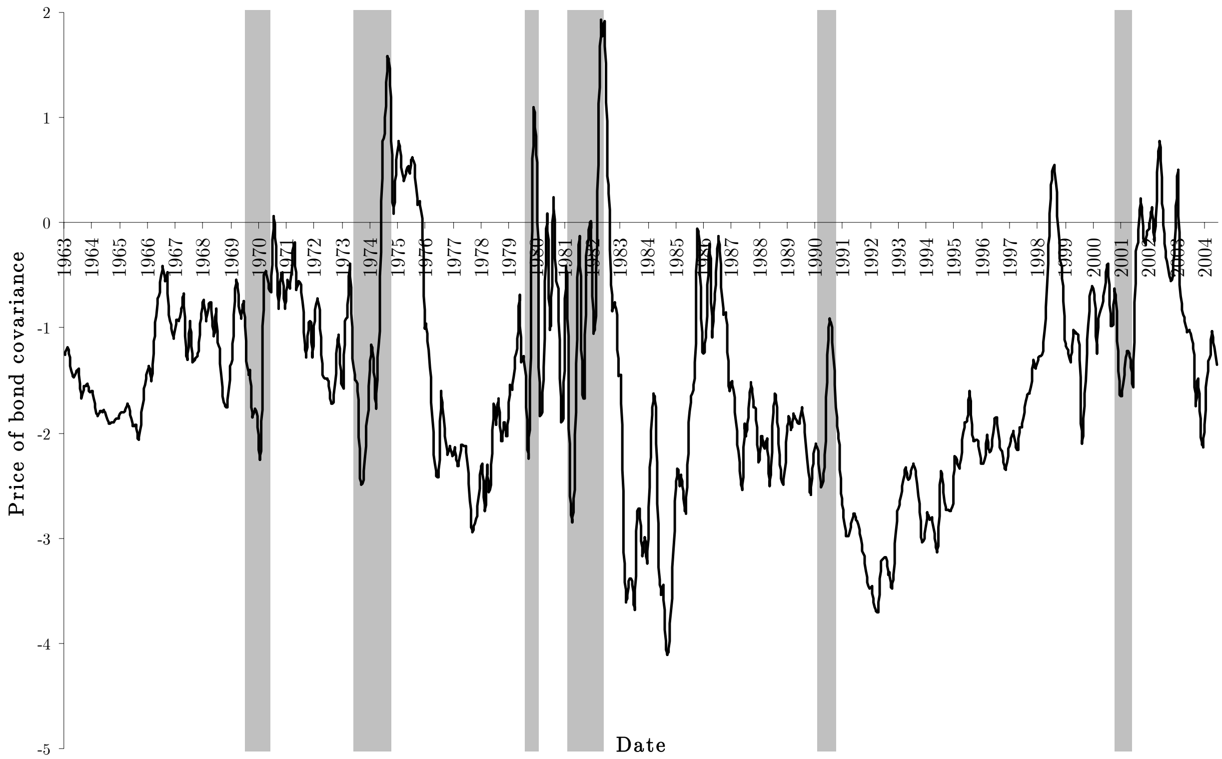

Figure 2. Time variation of the prices of bond covariance, 1963:7–2004:12. This graph plots the time series of the estimated prices of bond covariance, which is specified as a linear function of the lagged instruments. The estimates are obtained from Eq. (10). NBER recession dates are plotted in the shadowed area. The sample covers 498 monthly observations (from July 1963 to December 2004).

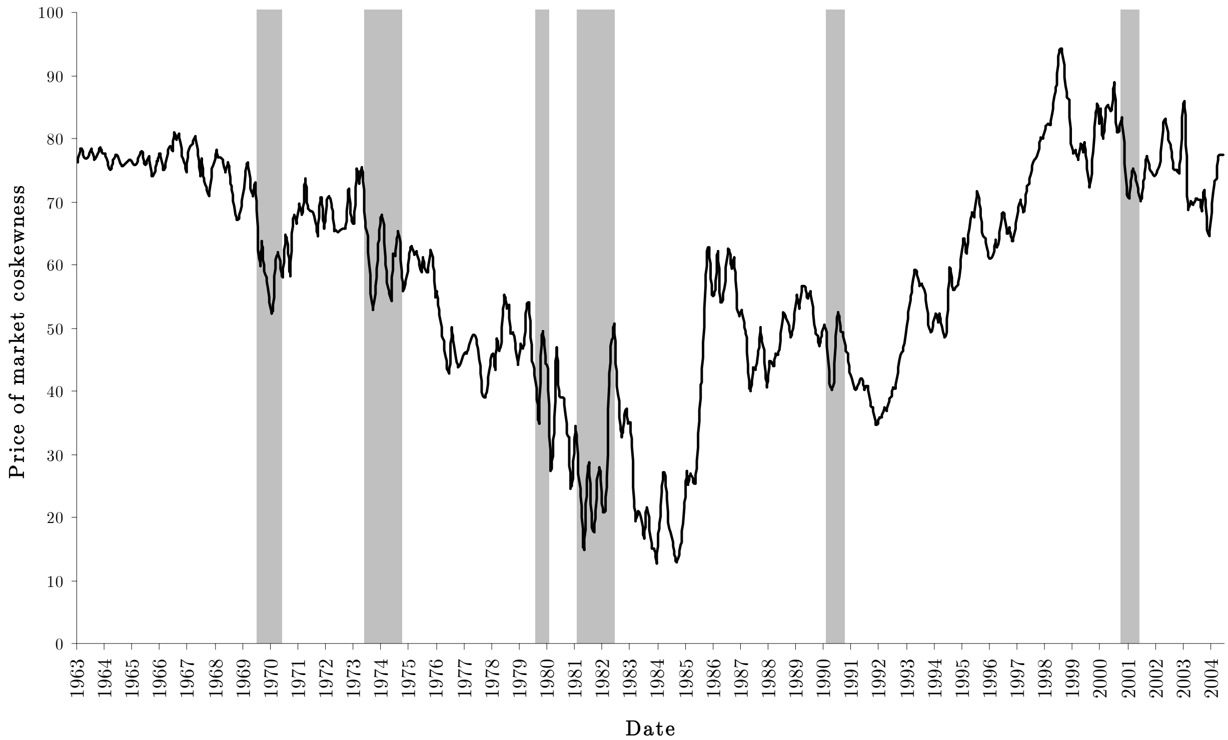

Figure 3. Time variation of the prices of market coskewness, 1963:7–2004:12. This graph plots the time series of the estimated prices of market coskewness, which is specified as a linear function of the lagged instruments. The estimates are obtained from Eq. (10). NBER recession dates are plotted in the shadowed area. The sample covers 498 monthly observations (from July 1963 to December 2004).

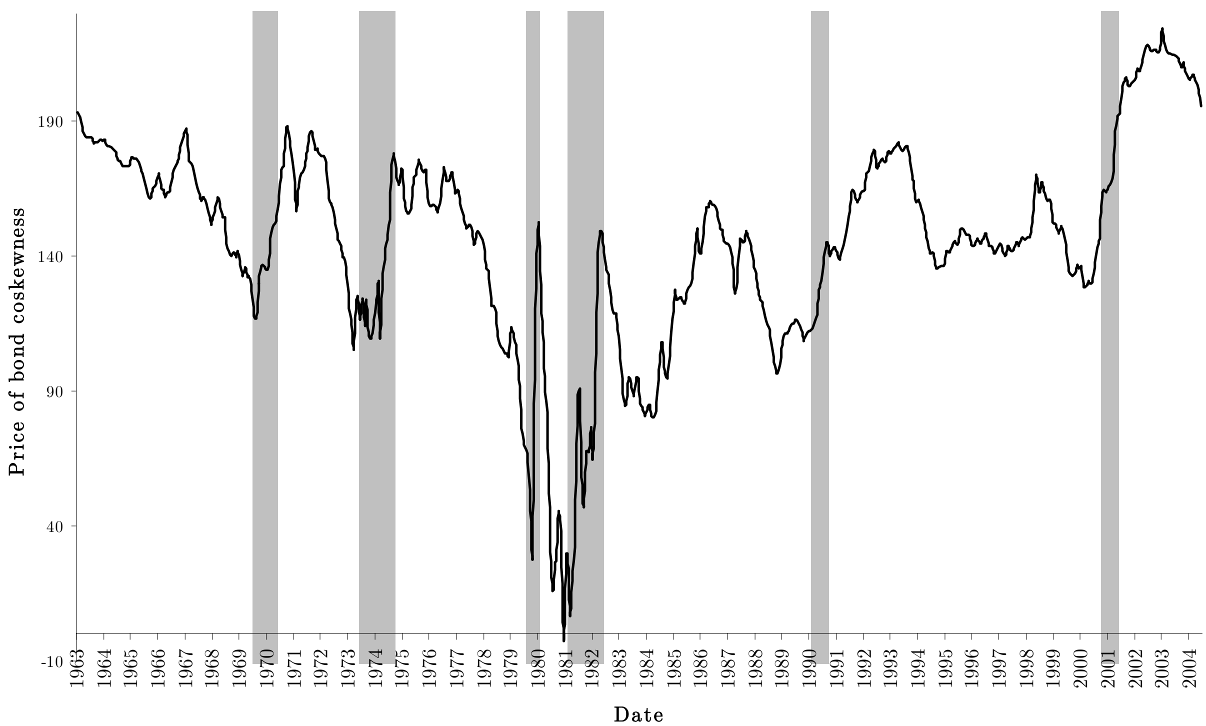

Figure 4. Time variation of the prices of bond coskewness, 1963:7–2004:12. This graph plots the time series of the estimated prices of bond coskewness, which is specified as a linear function of the lagged instruments. The estimates are obtained from Eq. (10). NBER recession dates are plotted in the shadowed area. The sample covers 498 monthly observations (from July 1963 to December 2004).