TIME-VARYING CONDITIONAL SKEWNESS AND THE MARKET RISK PREMIUM

Campbell R. Harvey and Akhtar Siddique

ABSTRACT

Single factor asset pricing models face two major hurdles: the problematic time-series properties of the ex ante market risk premium and the inability of the risk measure to account for a substantial degree of the cross-sectional variation of expected excess returns. We provide an explanation for the first failure using the following intuition: if investors know that the asset returns have conditional skewness at time t, *the expected excess returns should include rewards for accepting skewness. We formalize this intuition with an asset pricing model which incorporates conditional skewness. We decompose the expected excess returns into components due to conditional variance and skewness. Our results show that conditional skewness is important and, when combined with the economy-wide reward for skewness, helps explain the time-variation of the ex ante market risk premiums. Conditional skewness has greater success in explaining the ex ante risk premium for the world portfolio than for the U.S. portfolio.*

1. INTRODUCTION

The behavior of the market risk premium in the Capital Asset Pricing Model (CAPM) of Sharpe (1964) and Lintner (1965) has come under scrutiny from

two fronts. First, the evidence in Fama & French (1992) suggests that the estimated market risk premium is not different from zero. This implies that the systematic risk of the CAPM is not rewarded. Second, Boudoukh, Richardson & Smith (1993) present time-series evidence that the expected market risk premium is, at times, significantly less than zero. This implies that the market portfolio is on the negatively sloped portion of the mean-variance frontier – a violation of one of the CAPM's restrictions.

Aside from the criticism related to the market risk premium, results from cross-sectional tests of the single factor asset pricing model seem to indicate that the cross-asset variation in expected returns can not be explained by the market beta alone. For example, a number of recent papers find that 'fundamental' factors, which are idiosyncratic in nature, account for a sizable portion of the cross-sectional variation in expected returns [see Fama & French (1992); Chan, Hamao & Lakonishok (1991)]. Harvey & Siddique (2000) find that using conditional coskewness with the market can substantially mitigate the shortcomings of the single factor asset pricing model in explaining the cross-sectional variations in expected returns.

This study focuses on the time-series behavior of the risk premium. Our explanation for the time-series behavior of the market risk premium relies upon a single factor asset pricing model incorporating conditional skewness. This framework complements the cross-sectional results in Harvey & Siddique (2000).

Several important distinctions exist between our results and other recent work. In contrast to the usual beta-risk premium decomposition, we identify the risk factors in terms of the higher conditional moments such as the conditional covariance with the market and the conditional coskewness with the market. The second difference is in our use of a conditional methodology. We explicitly assume that the investor's information set changes over time. Thus, we allow time-varying risk and prices of risk and capture the variation with economically meaningful instrumental variables.

The following is our intuition for including skewness in the asset pricing framework. In the usual setup, investors have preferences over the mean and the variance of portfolio returns. The systematic risk of a security is measured as the contribution to the variance of a well-diversified portfolio. However, there is considerable evidence that the returns distributions cannot be adequately characterized by mean and variance alone. This leads us to next moment – skewness. Given the statistical evidence of skewness in returns, it is reasonable to assume that investors have preferences for skewness. With a large positive skewness (high probability of a large positive return), the investors may be willing to hold a portfolio even if its expected return is negative. As we

show later, this is still fully consistent with the Arrow-Pratt notion of risk aversion. Similarly, variation in skewness risk should also be important for the cross-section of assets.

Skewness may be important in investment decisions because of induced asymmetries in ex-post (realized) returns. At least two factors may induce asymmetries. First, the presence of limited liability in all equity investments may induce option-like asymmetries in returns [see Black (1976), Christie (1982) and Nelson (1991)]. Second, the agency problem may induce asymmetries in index returns [see Brennan (1993)]. That is, a manager has a call option with respect to the outcome of his investment strategies. Managers may prefer portfolios with high positive skewness.

We present an asset pricing model where skewness is priced. Our formulation is related to the seminal work of Kraus & Litzenberger (1976) and more recently, to the nonlinear factor model presented in Bansal & Viswanathan (1993). Our evidence documents significant time-variation in conditional skewness measures for both the U.S. stock market and, a broader world market portfolio. We estimate the price of skewness risk and show that this asset pricing model can account for much of the time-series variation in the expected market risk premium. We also find that this model helps explain many of the episodes of negative ex ante market risk premiums.

Our chapter is organized as follows. In the second Section, we use a general stochastic discount factor pricing framework to show how skewness can affect the expected market risk premia. Our econometric methodology and tests are detailed in the third Section. The data are described in this Section as well. The empirical results for the market risk premium are described in the fourth Section. The final part offers some concluding remarks.

2. SKEWNESS IN ASSET PRICING THEORY

The first-order condition for an investor holding a risky asset for one period is:

$$E[(1 + R_{i,t+1})m_{t+1}|\Omega_t] = 1 \tag{1}$$

where $(1 + R_{i,t+1})$ is the total return on asset i, m_{t+1} is the marginal rate of substitution of the investor between periods t and $t+1$, and Ω_t is the information set available to the investor at time t. m_{t+1} can be viewed as a pricing kernel or a stochastic discount factor that prices all risky asset payoffs.

As shown in Harvey & Siddique (2000), assuming a linear functional form for the marginal rate of substitution

$$m_{t+1} = a_t + b_t R_{M,t+1}, \tag{2}$$

and the existence of a riskfree asset, we get the standard CAPM

$$E_t[r_{i,t+1}] = \frac{E_t[r_{M,t+1}]}{\text{Var}_t[r_{M,t+1}]}\text{Cov}_t[r_{i,t+1}, r_{M,t+1}] \tag{3}$$

where lower case r represents returns of a conditionally riskfree return. In such a model, the expected excess returns of the risky assets are independent of the spanning weights a_t and b_t.

The expression for the market risk premium, however, does incorporate, the spanning weights since

$$E_t[r_{M,t+1}] = -b_t R_{f,t+1}\text{Var}_t[r_{M,t+1}] \tag{4}$$

where $R_{f,t+1}$ is one plus the conditionally riskfree rate of return. Thus, the expected market risk premium equals the conditional variance of the market return multiplied by the price of variance risk. The price of market variance risk is simply the spanning weight b_t inflated by $-R_{f,t+1}$. Temporal variation in the price of variance risk comes from both $R_{f,t+1}$ and b_t.

If we relate the discount factor to the marginal rate of substitution between periods t and $t+1$, a Taylor's series expansion allows us to make the following identification:

$$m_{t+1} = 1 + \frac{W_t U''(W_t)}{U'(W_t)}R_{M,t+1} + o(W_t)$$

where $o(W_t)$ is the remainder in the expansion and $\dfrac{W_t U''(W_t)}{U'(W_t)}$ (which is $-b_t$ in (2)) is relative risk aversion. Then $a_t = 1 + o(W_t)$ and $b_t < 0$. A negative b_t implies that with an increase in next period's market return, the marginal rate of substitution declines. This decline in the marginal rate of substitution is consistent with decreasing marginal utility. This restriction implies that the expected market risk premium is positive. Even if we only observe a proxy for the market, say portfolio $r_{m,t}^*$, a positive conditional covariance with the true market portfolio implies that the expected excess return on the proxy should also be positive.

Departing from the standard approach and assuming that the stochastic discount factor is quadratic in the market return:

$$m_{t+1} = a_t + b_t R_{M,t+1} + c_t R_{M,t+1}^2 \tag{5}$$

gives us a model where the expected excess return on the asset is determined by both its conditional covariance with the market return and with the square of the market return (conditional coskewness). Again assuming the existence of a conditionally riskfree asset:

$$E_t[r_{i,t+1}] = \frac{\text{Var}_t[r_{M,t+1}^2]E_t[r_{M,t+1}] - \text{Skew}_t[r_{M,t+1}]E_t[r_{M,t+1}^2]}{\text{Var}_t[r_{M,t+1}]\text{Var}_t[r_{M,t+1}^2] - \text{Skew}_t[r_{M,t+1}])^2}\text{Cov}_t[r_{i,t+1},r_{M,t+1}],$$
$$+ \frac{\text{Var}_t[r_{M,t+1}]E_t[r_{M,t+1}^2 - \text{Skew}_t[r_{M,t+1}]E_t[r_{M,t+1}]}{\text{Var}_t[r_{M,t+1}]\text{Var}_t[r_{M,t+1}^2] - (\text{Skew}_t[r_{M,t+1}])^2}\text{Cov}_t[r_{i,t+1},r_{M,t+1}^2]. \quad (6)$$

This is the conditional version of the three-moment CAPM first proposed by Kraus & Litzenberger (1976) who use a utility function defined over the unconditional mean, standard deviation and the third root of skewness. Once again, the expected excess returns of the risky assets are independent of the spanning weights, a_t, b_t, and c_t. Thus, if we test (5) for a cross-section of risky assets with an explicitly identified market portfolio, our tests cannot distinguish between different candidate pricing kernels with different spanning weights. Testing (5) alone also does not say anything about the expected market risk premium.

The expression for the expected market risk premium is

$$E_t[r_{M,t+1}] = -b_t R_{f,t+1}\text{Var}_t[r_{M,t+1}] - c_t R_{f,t+1}\text{Skew}_t[r_{M,t+1}] \quad (7)$$

Expanding, as before, the marginal rate of substitution in a power series gives

$$m_{t+1} = 1 + \frac{W_t U''(W_t)}{U'(W_t)}R_{M,t+1} + \frac{W_t^2 U'''(W_t)}{2U'(W_t)}R_{M,t+1}^2 + o(W_t). \quad (8)$$

Then $b_t < 0$ and $c_t > 0$ since non-increasing absolute risk aversion implies $U''' > 0$. According to Arrow (1964), non-increasing absolute risk aversion is one of the essential properties for a risk-averse individual.

In the standard CAPM, the expected market risk premium is the product of the conditional variance and the price of variance. In the three-moment CAPM, the market return is also a function of the conditional skewness and the price of skewness. The intuition of a positive conditional covariance of the proxy and the true portfolio ensuring that the expected excess return on the proxy is positive no longer follows. The expected excess return on the proxy can be positive or negative. The sign will depend on the magnitude of conditional skewness and the time-series behavior of the price of skewness.

We are able to decompose the contributions of conditional variance and skewness to the expected market risk premium and the contributions of conditional covariance and coskewness to the expected excess return of a specific asset. Alternative nonlinear frameworks such as Bansal & Viswanathan (1993) are unable to provide this decomposition. In addition, we are able to compute the prices of the various dimensions of risk represented by variance and skewness of the market return. This decomposition permits us to explain the time-series variation of the expected market risk premium and the cross-sectional variation in asset returns.

3. ECONOMETRIC METHODOLOGY

The formulation of the asset pricing model is very general in that it permits temporal variation in prices of variance and skewness risk as well as in the conditional moments themselves. The empirical estimation and tests of the model confront us with two problems. First, we need to distinguish between the time-varying prices of risk and the time-varying conditional moments. Second, we would like to avoid distributional assumptions about the conditional moments that do not come from the theory and may in fact conflict with it. Indeed, research has shown that the relation found between the conditional market risk premium and conditional variance may be largely a function of the specification chosen for the conditional moments. For example, Glosten, Jagannathan & Runkle (1993) report that using an asymmetric GARCH-M specification for conditional variance results in a positive relation between conditional risk premium and conditional variance whereas a regular GARCH-M yields a negative relation. We first document time-variation in the conditional moments using an explicitly chosen functional form for the conditional expectations. We also provide statistical tests. For example, we test whether conditional skewness is evident in the data.

   For the test of the model itself, we pursue two econometric formulations. In the first, we utilize the idea of Campbell (1987), Harvey (1989) and Dumas & Solnik (1995), where asset pricing restrictions can be tested without modelling the conditional higher moments. We use Hansen's (1982) generalized method of moments for the tests. We begin by testing the restrictions on the time-series variation in the expected market risk premium. We then add other assets and use this method in cross-sectional analysis. Our technique allows us to recover the fitted prices of variance and skewness. To explain the temporal variation in the expected market risk premium, we also need the corresponding higher moments. Therefore, we then estimate conditional skewness and variance in a non-parametric framework that imposes very few distributional assumptions. We combine the prices of risk with the conditional skewness and variance to get the expected market risk premium implied by the asset pricing model. We evaluate the relation between the statistically fitted expected market risk premium and theoretically implied expected market risk premium. We also determine whether the addition of the conditional skewness and its time-varying price helps explain any of the negative ex ante risk premiums.

   Ironically, the advantage of this first formulation (avoiding moment specification) is also its disadvantage. Expected returns implied by asset pricing theory can only be obtained by combining the prices of risk with fitted values for the higher moments from an ancillary, separate estimation. This motivates

our second formulation. We jointly estimate the conditional mean, variance and skewness as well as the prices of variance and skewness in a conditional maximum likelihood framework. This requires us to choose explicit functional forms for conditional variance and skewness. While heavily parameterized, this model allows us to directly test whether the addition of skewness helps explain the negative risk premiums and avoids the two-step estimation problem. However, this estimation method is impractical for the large group of assets included in cross-sectional analysis. Hence, this method is used only for understanding the time-series variation in the market risk premium.

3.1. Prices of variance and skewness risk

Campbell (1987), Harvey (1989) and Dumas & Solnik (1995) propose models where restrictions can be tested without specifying the variance dynamics. We extend this idea to skewness. Following Dumas & Solnik (1995), define the unexpected relative shock to the marginal rate of substitution:

$$u_{t+1} = - \frac{m_{t+1}}{E[m_{t+1} \mid \Omega]}, \tag{9}$$

where $E[u_{t+1} \mid \Omega_t] = 0$. Using u_{t+1} in (1) to substitute for m_{t+1} and using the conditionally riskfree rate of return to obtain the excess return:

$$E[r_{M,t+1}(1 - u_{t+1}) \mid \Omega_t] = 0$$
$$\Rightarrow E[r_{M,t+1} \mid \Omega_t] = E[R_{M,t+1} u_{t+1} \mid \Omega_t] \tag{10}$$
$$\Rightarrow E[r_{M,t+1} \mid \Omega_t] = \mathrm{Cov}[R_{M,t+1}, u_{t+1} \mid \Omega_t]$$

where the lower case $r_{M,t+1}$ is the excess market return (market risk premium) and upper case $R_{M,t+1}$ is the total market return. (9) and (10) impose restrictions on the conditional moments of the market risk premium. (10) should hold for excess returns of all other risky assets as well. We assume that the unexpected component of m_{t+1} is spanned by a quadratic function of the market return $R_{M,t+1}$:

$$u_{t+1} = \lambda_{0,t} + \lambda_{1,t} R_{M,t+1} + \lambda_{2,t} R_{M,t+1}^2, \tag{11}$$

where $\lambda_{1,t}$ and $\lambda_{2,t}$ are functions of period-t information set. Replacing u_{t+1} with this function in (10) gives us:

$$E[r_{M,t+1} \mid \Omega_t] = \lambda_{1,t} \mathrm{Var}[R_{M,t+1} \mid \Omega_t] + \lambda_{2,t} \mathrm{Skew}[R_{M,t+1} \mid \Omega_t]$$
$$= \frac{\lambda_{0,t}}{(1 - \lambda_{0,t})} R_{f,t+1} + \frac{\lambda_{1,t}}{(1 - \lambda_{0,t})} E[R_{M,t+1}^2 \mid \Omega_t] \tag{12}$$
$$+ \frac{\lambda_{2,t}}{(1 - \lambda_{0,t})} E[R_{M,t+1}^3 \mid \Omega_t].$$

Comparing (12) to (7) tells us that $\lambda_{1,t}$ and $\lambda_{2,t}$ are the time-varying market prices for variance and skewness risks, respectively. These two formulations for the expected market risk premium are equivalent. They are respectively in terms of the central and non-central conditional moments of the total market return.

For assets other than the market portfolio, equation (12) becomes:

$$E[r_{j,t+1}|\Omega_t] = \lambda_{1,t}\text{Cov}[R_{M,t+1},R_{j,t+1}|\Omega_t] + \lambda_{2,t}\text{Coskew}[R_{j,t+1},R_{M,t+1}|\Omega_t] \quad (13)$$

The coskewness between the market and the asset j is measured as the covariance between $R_{j,t+1}$ and $R_{M,t+1}^2$. This cross-sectional restriction impose the same prices of risk for all the assets.

We assume that $\lambda_{i,t}, i = 0, 1, 2$ are functions of Z_t where $Z_t \in \Omega_t$. Z_t are instruments in the information set available to investors at time t. We use the formulation of the model with non-central moments of the total market return and iterate the conditional expectations. Thus, the unconditional moment restriction for the market return that follows (11) is

$$E[u_{t+1}|Z_t] = 0 \Rightarrow E[(f_0(Z_t) + f_1(Z_t)R_{M,t+1}$$
$$+ f_2(Z_t)R_{M,t+1}^2) \otimes Z_t] = 0,$$

where f_i are the functional forms for the prices of risk, $\lambda_{i,t}$, since $f_i(Zt) = \lambda_{i,t}$.

Assuming that the f_i are linear in Z_t and using (10) give us the following two restrictions

$$E[(\delta_0'Z_t + \delta_1'Z_tR_{M,t+1} + \delta_2'Z_tR_{M,t+1}^2 \otimes Z_t] = 0.$$
$$E[(r_{M,t+1} - R_{M,t+1}[\delta_0'Z_t + \delta_1'Z_tR_{M,t+1} + \delta_2'Z_tR_{M,t+1}^2]) \otimes Z_t]. \quad (14)$$

where δ_i are parameters in the prices of risk. These moment restrictions do not include any parameters for the conditional moments themselves.

The inequality restrictions on the prices of risk are $\lambda_{1,t} \geq 0$ and $\lambda_{2,t} \leq 0$. Non-negativity of m_{t+1} requires that u_{t+1} be less than 1. We test the unconditional moment restrictions using Hansen's generalized method of moments (GMM). We estimate (14) with and without the inequality restrictions on the prices of risk. To impose the inequality restrictions, we use a quadratic specification for f_1 as square of a linear function, $(\delta_1'Z_t)^2$, and f_2 as -1 multiplied by a quadratic specification, i.e. $-(\delta_2'Z_t)^2$. In all cases we use a heteroskedasticity consistent variance-covariance matrix with a Parzen kernel. The minimized GMM criterion function multiplied by the number of observations is distributed as a χ^2 with degrees of freedom equal to the number of overidentifying restrictions which equals the number of orthogonality conditions less the number of parameters. This is a specification test of the model.

3.2. *Non-parametric estimation of conditional moments*

The second stage of our estimation involves the conditional moments. For each of the returns, we need three moments. For the market portfolio proxy, we need the conditional mean, variance and skewness. For the other asset returns, we need the conditional mean, covariance with the market, and coskewness with the market.

We first examine the market risk premium. We document that the conditional moments of the market risk premium vary over time using linear specifications. We then estimate the three conditional moments, mean, variance, and skewness without imposing a functional form. We use non-parametric kernels to compute these conditional moments. The kernel method does not impose any distributional assumptions on the market risk premium or the instruments. The method locally approximates the unknown underlying conditional density (of the market risk premium conditioned on the instruments) using a weighted sum of the market risk premia. The function chosen for the weighting scheme is the kernel or basis function. The expressions for the three moments using the kernel method are:

$$\widehat{E[r_{M,t+1} \mid \Omega_t]} = \sum_{j=1}^{T-1} r_{M,j+1} W_{t,j}$$

$$\widehat{\mathrm{Var}[r_{M,t+1} \mid \Omega_t]} = \sum_{j=1}^{T-1} (r_{M,j+1} - \widehat{E[r_{M,j+1} \mid \Omega_j]})^2 W_{t,j} \tag{15}$$

$$\widehat{\mathrm{Skew}[r_{M,t+1} \mid \Omega_j]} = \sum_{j=1}^{T-1} (r_{M,j+1} - \widehat{E[r_{M,j+1} \mid \Omega_t]})^3 W_{t,j}$$

where

$$W_{t,j} = \frac{K\left[\dfrac{Z_t - Z_j}{h}\right]}{\sum_{j=1}^{T-1} K\left[\dfrac{Z_t - Z_j}{h}\right]}$$

where $K[\cdot]$ is the multivariate kernel function and h is the *bandwidth*, different for each conditional moment. The bandwidth determines the number of

observations that get a non-negligible weight. We use the multivariate Gaussian kernel function with bandwidth chosen to be asymptotically optimal for a mean squared error criterion. Thus, the conditional mean of the market risk premium is formed as the weighted sum of all the market risk premia with the bandwidth determining which observations have non-negligible weights.

For conditional variance and skewness, we augment the instrument set, Z_t, to include ε_t^2 and ε_{t-1}^2 for variance and ε_t^3 and ε_{t-1}^3 for skewness where ε_t is the residual from the conditional mean estimation. Thus, for conditional variance and skewness, our specification is in the spirit of a non-parametric ARCH(2) specification.

3.3. Conditional maximum likelihood estimation

As an alternative to the two-stage methodology, we use conditional maximum likelihood estimation. Here, we assume explicit functional forms for the prices of risk and the dynamics of the conditional moment evolutions. Then we impose the restrictions on the moments implied by the theory and estimate the parameters in a conditional maximum likelihood framework. This method allows us to avoid the problems of multistage estimation. There are more parameters to estimate in the likelihood approach. However, when the unique elements of the weighting matrix are considered, the number of parameters to estimate in the GMM approach is in fact greater. This approach is feasible only for the market risk premium.

The asset pricing model implies

$$r_{M,t+1} = \lambda_{1,t}\mathrm{Var}[R_{M,t+1}|\Omega_t] + \lambda_{2,t}\mathrm{Skew}[R_{M,t+1}|\Omega_t] + \varepsilon_{t+1} \qquad (16)$$

To estimate the higher-order moments in (16), we assume that the expected market risk premium is linear in the instrumental variables. Conditional variance needs to be strictly positive. To ensure the positivity, we compute the conditional standard deviation using the absolute residuals. For conditional standard deviations, we impose a GARCH(2,2) specification with instruments. The advantage of the GARCH specification is that it allows dependence on past conditional variances. For skewness (which in our definition is not normalized by the standard deviation), we also choose a GARCH(2,2) specification with instruments. However, we do not impose any restrictions on the signs of the parameters. We assume the prices of variance and skewness risk are linear in the instruments. Define $h_{t+1} = \mathrm{Var}_t[r_{M,t+1}]$ and $s_{t+1} = \mathrm{Skew}_t[r_{M,t+1}]$ Thus:

$$\sqrt{h_{t+1}} = \beta_0 + \sum_{i=1}^{2} \beta_i |\varepsilon_{t+1-i}| + \sum_{i=1}^{2} \alpha_i \sqrt{h_{t+1-i}} + \phi' Z_t$$

$$s_{t+1} = \gamma_0 + \sum_{i=1}^{2} \gamma_i \varepsilon_{t+1-i}^3 + \sum_{i=1}^{2} \upsilon_i s_{t+1-i} + \delta' Z_t \qquad (17)$$

$$\lambda_{1,t} = \lambda_1' Z_t$$

$$\lambda_{2,t} = \lambda_2' Z_t$$

Assuming that the errors, ε_{t+1}, have a conditional t distribution we can write the sample log-likelihood function conditional on the first m observations as:

$$\sum_{t=1}^{T} \ln f(\varepsilon_{t+1}|Z_t, \Theta) = T \ln \left\{ \frac{\Gamma[\nu+1)/2]}{\pi^{1/2}\Gamma[\nu/2]}(\nu-2)^{-1/2} \right\} - \frac{1}{2} \sum_{t=1}^{T} \ln(h_{t+1})$$

$$- [(\nu+1)/2] \sum_{t=1}^{T} \ln \left\{ 1 + \frac{\varepsilon_{t+1}^2}{h_t(\nu-2)} \right\} \qquad (18)$$

where Γ is the *gamma* function and ν is the degrees of freedom of the t distribution. The choice of t distribution is motivated by the evidence that even after assuming a GARCH specification for conditional variance, the distribution of the residuals displays thick tails.

We also need to estimate the initial conditional variances and skewnesses, h_1, h_2, s_1, and s_2. Thus, the parameters to estimate are:

$$\Theta = [\delta \gamma \alpha \delta \upsilon \phi \lambda_1 \lambda_2 \nu h_1 h_2 s_1 s_2]'$$

The parameters are obtained by maximizing the sample log-likelihood function.

We estimate the model with the inequality constraints (positive price of variance risk and negative price of skewness risk) as well as without the constraints. As in the GMM-based methodology, to impose positivity on the price of variance risk, we use the square of the linear function, $(\lambda_1' Z_t)^2$. To ensure negativity of the price of skewness risk, we use -1 multiplied by the square of the linear function, $-(\lambda_2' Z_t)^2$.

To test whether skewness enters the asset pricing model, we also estimate the likelihood function (19) without skewness. This formulation of the model is equivalent to the standard CAPM. Twice the difference of the two sample log-likelihoods (with and without skewness) is approximately distributed as a $\chi^2(q)$

where q is the number of parameters to estimate in the price of skewness risk and conditional skewness.

3.4. Data and Summary Statistics

We use several different data sets. For understanding the time-series behavior of the market risk premium, we analyze three data sets. The first is studied by Boudoukh, Richardson & Smith (1993). These annual data include the historical U.S. stock market premium from 1802–1990. Second, we use the monthly U.S. data analyzed in Harvey (1989). This data set is from September 1941 to December 1987 which we then update to September 1991. Finally, we examine the data presented in Harvey (1991) and updated by Ostdiek (1994). These data measure world stock market returns at a monthly frequency from 1970–1992.

The Boudoukh, Richardson & Smith (1993) annual data derives from the market returns presented in Siegel (1990). Three instrumental variables are constructed: the lagged short-term interest rate, the lagged dividend yield, and the lagged slope of the term structure (as measured by the difference between long and short-term interest rates).

In the Harvey (1989) data, the market portfolio return is the value-weighted NYSE index return from CRSP files. The instruments include: the lagged return on the equally-weighted NYSE index, the lagged yield spread between Moody's Aaa and Baa rated bonds, the lagged excess return on a three month Treasury bill, and the lagged excess U.S. dividend yield.

For the world data, the market return is Morgan Stanley Capital International world index. The instruments are: lagged world excess return, the lagged yield spread between Moody's Aaa and Baa rated bonds, the lagged excess return on a three month U.S. Treasury bill, and the lagged U.S. dividend yield.

Table 1 presents the summary statistics for the three proxies for the market risk premium. The results show, that for both the U.S. and MSCI world portfolios, the ex-post market risk premium has been negative much more often than the statistically fitted market risk premium. Using the a model that is linear in the instrumental variables, 3.2%, 25.8% and 32.7% of the statistically fitted market risk premia are negative for the U.S. annual, U.S. monthly and world monthly returns respectively. The R^2s measuring predictability of the market risk premia are 6.3% for the annual U.S. data, 8.1% for the monthly U.S. and 9.8% for the world. Using the cross-validated adaptive kernel this measure of predictability increases to 11.5% for the monthly U.S. and 12.4% for the world. However, for the annual series, the cross-validated adaptive kernel shows only a trace amount of predictability. With the non-parametric measure, the

Table 1.   A. Summary Statistics for the Market Risk Premia

Summary statistics are provided for the market risk premium for the US and MSCI world portfolios. The mean and standard deviation are for annualized monthly returns. The skewness reported is the third central moment normalized by standard deviation cubed, $\dfrac{E[\varepsilon_{i,t}^3]}{\sqrt{E[\varepsilon_{i,t}^2]^3}}$

	Mean	Standard deviation	Skewness	Autocorrelation
U.S. 1941.09–1993.12	8.37	14.52	–0.40	0.08
World 1970.02 – 1992.12	5.10	14.82	–0.34	0.11

B. Tests for Time-variation in conditional moments of market returns

GMM-based Wald tests for time-variation in the conditional mean, variance and skewness of the U.S. and MSCI world portfolio returns are presented with a linear specification for for the time-varying moments:

$$E[r_{M,t+1}|\Omega_t] = \beta_0 + \sum_{i=1}^{4}\beta_i Z_{it}$$

$$E[R_{M,t+1}^2|\Omega_t] = \beta_0 + \sum_{i=1}^{4}\beta_i Z_{it} + \sum_{i=1}^{2}\gamma_i R_{M,t+1-i}^2$$

$$E[R_{M,t+1}^3|\Omega_t] = \beta_0 + \sum_{i=1}^{4}\beta_i Z_{it} + \sum_{i=1}^{2}\gamma_i R_{M,t+1-i}^3$$

where Z_i are the four instruments Baa-Aaa yield spread, market return, excess holding period return for 3 month Treasury bill, and riskfree rate, R_M is the market return and r_M is the market risk premium. The variance and skewness are defined as the non-central moments of the total return. The statistics reported are the χ^2-values and the numbers in parentheses are significances or the probability of observing a larger χ^2-statistic under the null hypothesis of no time-variation. The R^2 is presented for the risk premium.

		Market Risk Premium			Variance		Skewness	
		R^2	χ^2	*p*-value	χ^2	*p*-value	χ^2	*p*-value
U.S.	1941.09–1987.12	8.1%	32.87	(0.000)	47.17	(0.000)	23.52	(0.000)
	1941.09–1970.01	3.2%	9.47	(0.050)	21.90	(0.000)	10.27	(0.114)
	1970.01–1987.12	14.9%	41.74	(0.000)	23.88	(0.000)	32.89	(0.000)
World	1970.02–1992.12	9.8%	30.55	(0.000)	12.74	(0.047)	17.64	(0.007)
	1970.02–1979.12	15.6%	18.08	(0.001)	12.38	(0.054)	28.64	(0.000)
	1979.12–1992.12	7.2%	16.07	(0.003)	15.90	(0.014)	15.12	(0.019)

proportion of the statistically fitted market risk premia that is negative declines. Nevertheless, 10.3% of the U.S. and 21.1% of the world statistically fitted monthly market risk premia are still negative. This suggests that a potentially misspecified linear specification can not be blamed for the negativity of expected market risk premia.

For both of the monthly U.S. and world portfolios, the unconditional skewness is negative over the sample period. In contrast, the annual market risk premium displays positive unconditional skewness over the period 1803–1990. However, consistent with the monthly data, in the subperiod 1941–1990, the annual U.S. market risk premium has negative skewness.

Figure 1 presents the fitted market risk premiums from both the OLS and non-parametric models. Generally, the OLS and non-parametric fitted values look similar with a correlation between the two of 90% for the world portfolio and 84% for the U.S. portfolio. For the non-parametric analysis, we also present two standard error confidence bands for the fitted values. For the monthly U.S. and world returns, the standard error bands confirm that a number of the fitted values are negative. However, the standard error bands on the annual data are very large implying that there is little or no predictability in the returns given these instruments. This may be a result of the early data being approximated, poor quality of the instruments or a fundamental lack of predictability. Nevertheless, it does not make much sense to proceed with a model that attempts to explain the negative ex ante market risk premia with such large standard error bands. As a result, the rest of the paper concentrates on the evidence using monthly U.S. and world equity market returns. Previous studies documenting negativity of the expected market risk premia have largely used an explicit linear specification for the conditional mean. Hence, in the rest of the paper we use the the OLS-based mean as our statistical fitted market risk premium.

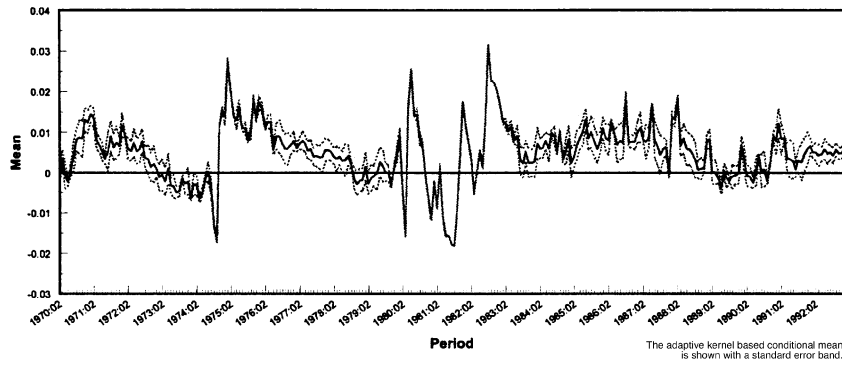## 4. RESULTS: MARKET RISK PREMIUM

4.1. Prices of risk and conditional moments: GMM results

The asset pricing model suggests that the variation in conditional market risk premium can be explained by variation in the prices of risk and the conditional variance and skewness with the conditional skewness potentially accounting for the negative market risk premia. Thus, documenting the variation in the conditional moments needs to be the first step. We propose the following tests for the variation of the conditional moments:

A. US Conditional Mean: Monthly returns 1941.09-1987.12

The adaptive kernel based conditional mean
is shown with a standard error band.
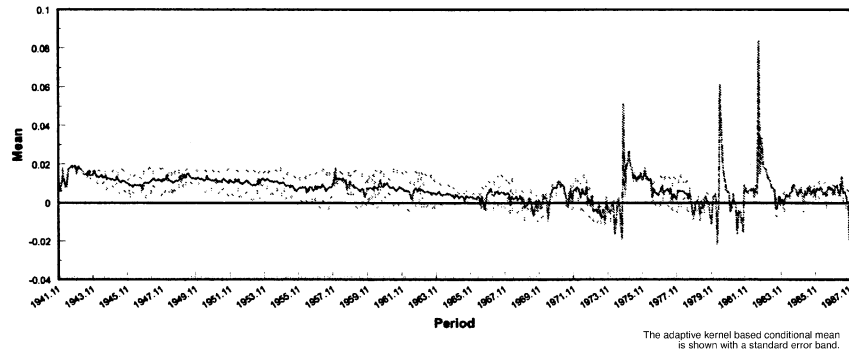
B. World Conditional Mean: Monthly returns 1970.04-1992.12

The adaptive kernel based conditional mean
is shown with a standard error band.

C. US Conditional Mean: Annual returns 1803-1990

The adaptive kernel based conditional mean
is shown with a standard error band.

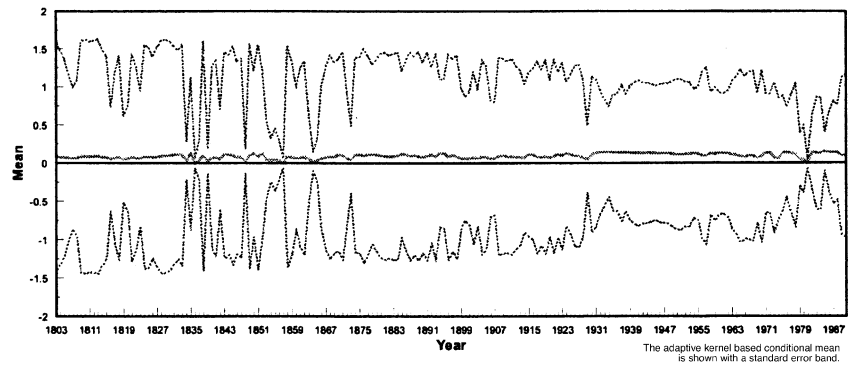*Fig. 1.*

$$E[r_{M,t+1}|\Omega_t] = \beta_0 + \sum_{i=1}^{N} \beta_i Z_{it}$$

$$E[R_{M,t+1}^2|\Omega_t] = \beta_0 + \sum_{i=1}^{4} \beta_i Z_{it} + \sum_{i=1}^{2} \gamma_i R_{M,t+1-i}^2 \qquad (19)$$

$$E[R_{M,t+1}^3|\Omega_t] = \beta_0 + \sum_{i=1}^{4} \beta_i Z_{it} + \sum_{i=1}^{2} \gamma_i R_{M,t+1-i}^3$$

where as before the lower case represents the market risk premium (market return in excess of the riskfree rate of return) and the upper case is the total market return. We separately estimate each conditional moment in (19) using generalized methods of moments and present Wald tests of their time-variation. The results are shown in panel B of Table I. For both U.S. and world portfolios, the tests reject the null hypotheses that the market risk premium, variance and skewness are constant over time. When we consider subperiods for the U.S. portfolio, variabilities of both conditional skewness and conditional market risk premium have become more significant after 1970. For the world portfolio, skewness appears to be more significant than variance over the entire sample (which is from 1970) as well as in subperiods.

Table 2 presents tests of the model in (14) using generalized method of moments. For each of the data sets, we have 18 orthogonality conditions and 15 parameters which produce 3 overidentifying restrictions. With 171 unique elements in the weighting matrix, the *saturation ratio* is 2.98 for the U.S. and 1.48 for the world. Under the null hypothesis, that the model is correctly specified, the objective function multiplied by the number of observations should be distributed as a χ^2 with 3 degrees of freedom. For the U.S. portfolio without constraints on prices of risk, the objective function of the model has a suspiciously low value of 1.29 (p-value 0.73). For the constrained model, with the price of variance risk constrained to be positive and the price of skewness risk constrained to be negative, the objective function has a value of 1.46 (p-value 0.69). Therefore, the null hypothesis cannot be rejected in either case at 5% significance level. For the world portfolio, the unconstrained model with a χ^2 statistic of 0.44 produces a p-value of 0.93. The constrained model has a larger χ^2 statistic of 5.35 resulting in a p-value of 0.15. Thus, we do not reject the null hypothesis for the world portfolio either. In all the cases, our Wald tests on time-varying conditional skewness shows it to be significant.

As a diagnostic of the model, we also estimate a specification consistent with the standard CAPM, i.e. (14) with the price of variance risk alone. For both the

Table 2. GMM-based Asset Pricing Model Tests on Market Risk Premium Without Parameterizing Variance & Skewness

The hypothesis that variance and skewness risk jointly explain the risk premium is tested by estimating:

$$E[\delta_0' Z_t + \delta_1' Z_t R_{M,t+1} + \delta_2' Z_t R_{M,t+1}^2 \otimes Z_t] = 0.$$
$$E[[r_{M,t+1} - R_{M,t+1}(\delta_0' Z_t + \delta_1' Z_t R_{M,t+1} + \delta_2' Z_t R_{M,t+1}^2)] \otimes Z_t] = 0.$$

using generalized method of moments where r_M is the market risk premium, R_M is the total market return and Z are the instruments. For both the U.S. and World portfolios 18 orthogonality conditions and 15 parameters produce 3 overidentifying restrictions. The overidentifying χ^2 statistic tests these restrictions. In addition, the skewness χ^2 statistic tests the significance of time-varying skewness by using a Wald test statistic. *p*-value, reported in parentheses, is the significance level at which the null hypothesis will be rejected. The CAPM χ^2 is for the test of overidentifying restrictions implied by the CAPM. $\bar{u}$ represents the average estimated relative shock to the marginal rate of substitution using both variance and skewness in the *u*-specification.

A. Unconstrained

	Overidentifying χ^2	χ^2 for Skewness $= 0$	$\bar{u}$	$\sigma(u)$	CAPM χ^2
U.S.	1.29	48.27	0.000	0.389	7.46
	(0.732)	(0.000)			(0.488)
World	0.44	20.17	0.422	3.518	0.76
	(0.932)	(0.091)			(0.999)

B. Constrained

	Overidentifying χ^2	χ^2 for Skewness $= 0$	$\bar{u}$	$\sigma(u)$	CAPM χ^2
U.S.	1.46	40.07	13.390	64.365	6.01
	(0.692)	(0.000)			(0.646)
World	5.23	22.99	0.201	0.422	10.34
	(0.156)	(0.042)			(0.242)

portfolios, the χ^2 for the overidentifying restrictions with variance risk alone has a substantially larger value than the χ^2 for overidentifying restrictions with both skewness and variance. These results, in Table II, confirm the importance of incorporating skewness in the asset pricing model.

Another diagnostic is provided by the estimates of u_{t+1}, the relative shock (or forecast error) to the marginal rate of substitution. For a proper, i.e. strictly positive, marginal rate of substitution, u_{t+1} should be less than 1. Thus, estimates of u_{t+1} can gauge the economic reasonableness of different market proxies used for the marginal rates of substitution. For the U.S. portfolio without constraints on the prices of risk, the average estimated shock is very small. Imposition of the constraints causes the estimated shocks to often exceed 1. For the world portfolio, the average estimated shock is quite small with and without constraints. The constraints actually reduce the magnitude and variability of estimated shocks. Thus, constraints on prices of risk appear reasonable for the world portfolio but untenable for the U.S. portfolio in light of the unreasonably large estimates for the shocks to the marginal rate of substitution. So constrained estimation for the U.S. portfolio is not that useful.

Estimation of the model in (14) also yields the market prices of variance and skewness risk, $\lambda_{1,t}$ and $\lambda_{2,t}$ respectively. We test whether those prices of risk are constant over time with a likelihood ratio statistic. The statistic is $T(g_T(\delta_R) - g_T(\delta_F))$ where $g_T(\delta)_F$ and $g_T(\delta)_R$ are respectively the GMM objective functions for the full and nested models. The nested model is estimated with the same weighting matrix as the full model. For both the U.S. and world portfolios, the tests presented in Table 3 reject the null hypothesis that the prices of variance and skewness risk are constant over time.

The price of variance risk is the market risk premium economic agents demand for an increase in conditional variance. Similarly, the price of skewness risk is the risk premium economic agents are willing to give up for an increase in conditional skewness. We examine how the prices of risk from the model behave in relation to the statistically fitted market risk premium. For the constrained model using the U.S. portfolio, the statistically fitted market risk premium has a correlation of 0.62 with price of variance risk and a correlation of –0.37 with the price of skewness risk. Thus, when expected market risk premium increases, price of variance risk rises but the price of skewness risk becomes more negative. For the constrained model using the world portfolio, correlation of the statistically fitted market risk premium with the prices of variance and skewness risks are 0.83 and –0.08 respectively. For both the portfolios, imposition of constraints required by theory reduce the variability of the prices of risk. This effect is particularly pronounced for the world portfolio.

Table 3. Behavior of GMM-based Market Prices of Risk Without Parameterizing Variance & Skewness

The prices of variance and skewness risk are computed for the U. S. and MSCI world portfolios according to (16) using GMM. The standard deviation shown is the variation of the estimated price of risk over time. For the constrained case, the price of variance risk is restricted to be positive and the price of skewness risk is restricted to be negative. The unconstrained case does not impose any restrictions on the sign of the prices. The mean and standard deviation are multiplied by 10^3. To test time-variation in the prices a likelihood ratio test is used where the nested model forces the prices of variance and skewness to be constants. The numbers in the parentheses report the probability of a larger χ^2 under the null hypothesis of no time-variation. The degrees of freedom for the test are 8 for both the portfolios.

A. Unconstrained

	Time-varying price	Mean	Standard deviation	Negative %	χ^2 for $\lambda_{i,t} = \lambda_i$
U.S.	$\lambda_{0,t}$	0.35	8.32		
Monthly	$\lambda_{i,t}$(Price of Variance Risk)	4.01	7.23	27.6%	21.49
1941.09–1987.12	$\lambda_{2,t}$(Price of Skewness Risk)	50.60	108.05	27.2%	(0.005)
World	$\lambda_{0,t}$	–0.11	1.20		
Monthly	$\lambda_{1,t}$(Price of Variance Risk)	38.75	66.35	22.7%	23.20
1970.02–1992.12	$\lambda_{2,t}$(Price of Skewness Risk)	188.07	190.96	9.5%	(0.003)

B. Unconstrained

	Time-varying price	Mean	Standard deviation
U.S.	$\lambda_{0,t}$	0.09	1.42
Monthly	$\lambda_{1,t}$(Price of Variance Risk)	5.88	7.44
1941.09–1987.12	$\lambda_{2,t}$(Price of Skewness Risk)	–12.17	27.99
World	$\lambda_{0,t}$	0.18	1.30
Monthly	$\lambda_{1,t}$(Price of Variance Risk)	5.20	7.45
1970.02–1992.12	$\lambda_{2,t}$(Price of Skewness Risk)	–10.63	18.82

Next we estimate the conditional variance and skewness using the nonparametric kernel approach described in Section 3.2. We then combine the prices of risk and the conditional moments to obtain the expected market risk premium implied by the asset pricing model. Figure 2 shows the implied market risk premium and the statistically fitted market risk premium for the U.S. and the world portfolios. If the model is true and if we could ignore estimation and measurement errors, then the implied expected market risk premium should be identical to the statistically fitted market risk premium. Of course, this intuition also assumes that the statistically fitted market risk premium is the true conditional mean of the market risk premium. We can not realistically expect these assumptions to hold. Nevertheless, the degree to which the dotted line (implied expected market risk premium) approximates the solid line (statistically fitted expected market risk premium) heuristically indicates the relative success or failure of the asset pricing model in explaining the variation of the expected market risk premium for the two portfolios.

We find that for both the portfolios, imposition of the constraints on the prices of risk implied by the asset pricing theory improves the degree of fit between the statistically fitted market risk premium and the implied expected market risk premium. The correlation increases from 0.23 to 0.33 for the U.S. and from 0.20 to 0.41 for the world portfolio. For the world portfolio in particular, the reduction of the large gyrations in the implied expected market risk premium because of the constraints is very noticeable in the panels C and D of Fig. 2. With the constraints imposed, the average annualized expected market risk premium for U.S. implied by the asset pricing model is 9.8% compared to the average annualized statistically fitted market risk premium of 8.4%. For the world portfolio, the comparable numbers are 8.4% versus 5.1%. Thus, the asset pricing model has some success in explaining the variation of the expected market risk premia. The success is greater for the world portfolio. To check if the differing sample periods are responsible for the difference in performance, we compare the performance over the common period 1970.04–1987.12. In this period, the correlation between statistically fitted expected market risk premium and the theoretically implied expected market risk premium is 0.41 for the world and only 0.06 for the U.S. portfolio.

4.2. Prices of risk and conditional moments: Maximum likelihood results

The generalized method of moments methodology has two drawbacks. First, we are compelled to use a two-step estimation procedure. Initially, we compute the prices of risk without having to make assumptions about the conditional moment dynamics. But when we compute the conditional moments, we can not
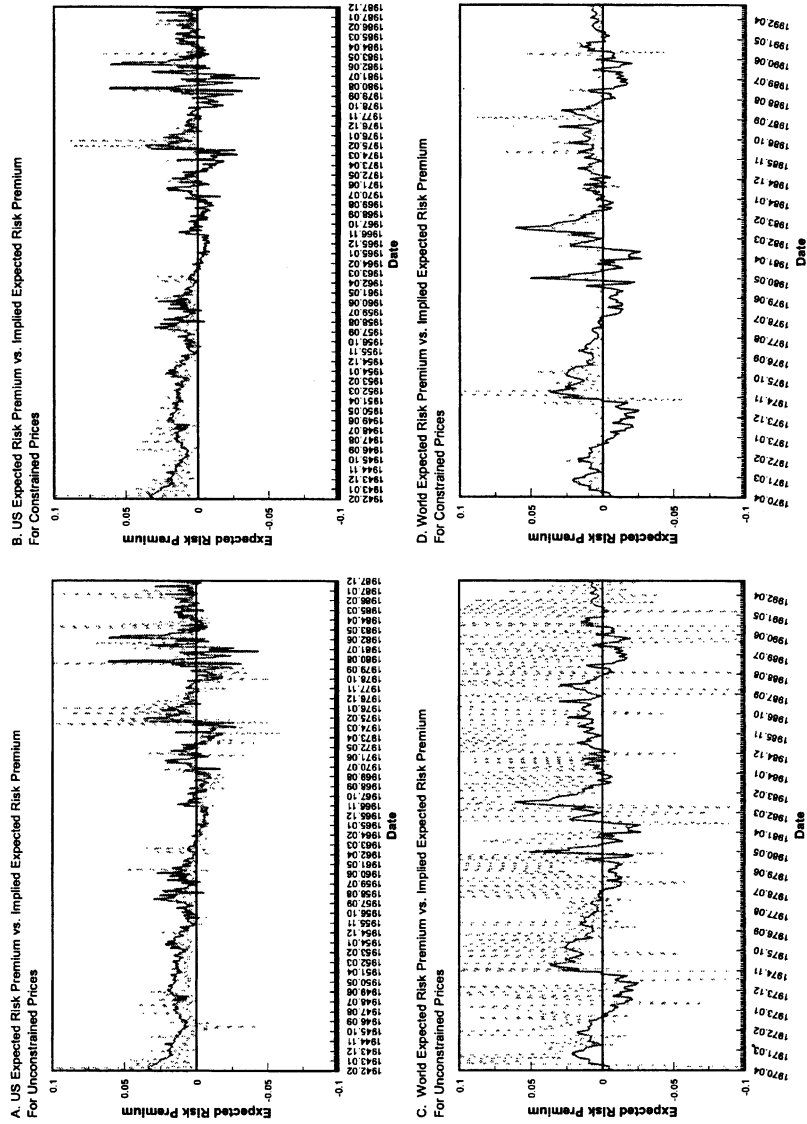
Fig. 2.

impose the restrictions on them (the moments) implied by asset pricing theory. Second, the saturation ratios for the two portfolios are low. These drawbacks motivate our alternative estimation method. We model the conditional mean of the market risk premium using a linear specification. Then, we estimate the prices of risk and the higher conditional moments simultaneously in a conditional maximum likelihood framework using the residuals from the mean. Thus, the prices of risk and the corresponding higher conditional moments are estimated under the restrictions that arise from the theory.

The maximum likelihood results are presented in Table 4. For each of the portfolios, there are 31 parameters to estimate. The estimated degrees of freedom for the t-distribution in all four cases are approximately 3. This confirms the presence of thick tails in the residuals.

The average log-likelihood for each of the models summarizes its relative success in capturing the variation in the expected market risk premium. For both the portfolios, the average likelihood declines when constraints are imposed on the prices of risk. However, the decline is relatively small for the world portfolio. We also see that the average likelihood for the world portfolio is larger than for the U.S. portfolio, though the sample size is smaller for the world. Thus, it appears that the asset pricing model incorporating skewness is a better description for the world portfolio than for the U.S. portfolio. These results are consistent with what we found using GMM.

Similar to our diagnostics for the GMM methodology, we also estimate a model with variance alone that is equivalent to the standard CAPM. The results, also presented in Table 4, show that conditional skewness is significant for explaining the variation in the expected market risk premium for the U.S. and world portfolios. Those conclusions are not affected by the imposition of positivity constraints. We also see that with variance alone the average likelihood declines much more precipitously for the world than for the U.S. This again confirms the greater importance of conditional skewness for the world portfolio.

The prices of risk along with estimated conditional variance and skewness are summarized in Table 5. The U.S. portfolio displays substantially greater average skewness than the world portfolio. For the world portfolio, the imposition of constraints on the prices of risk reduces the variability of the conditional moments. The absolute magnitudes of the prices of risk are substantially less than those obtained from the GMM methodology. For the U.S. portfolio, the mean of the unconstrained price of variance risk is negative and mean of the price of skewness risk is positive. However, for the world portfolio, the average prices of risk for variance and skewness are respectively

Table 4.   Asset Pricing Model Tests on Market Risk Premium Using Conditional Maximum Likelihood

The average estimated conditional maximum likelihood values are presented for the monthly U.S. and world portfolio returns. The *Skewness & Variance* model assumes that the expected market risk premium is explained by both conditional variance and skewness of market returns and has 33 parameters. The *Variance Alone* model assumes that the expected market risk premium is explained by conditional variance of market returns alone and has 17 parameters. This model is nested in the *Skewness & Variance* model and the significance of skewness is tested by the likelihood ratio statistic

$$2(\mathfrak{L}(\Theta_{Skewness\,\&\,Variance}) - (\mathfrak{L}(\Theta_{Variance})) \sim \chi^2_{15}$$

The constrained models force the price of variance risk to be strictly positive and the price of skewness risk to be strictly negative.

	Unconstrained			Constrained		
	Variance Alone	Skewness & Variance	χ^2 for Skewness $= 0$	Variance Alone	Skewness & Variance	χ^2 for Skewness $= 0$
U.S.	1.64	1.79	173.90 (0.000)	1.58	1.72	156.78 (0.000)
World	1.56	1.92	198.14 (0.000)	1.55	1.92	202.56 (0.000)

Table 5. Behavior of Conditional Maximum Likelihood-based Market Prices of Risk

The prices of variance and skewness risk are computed for the U.S. and MSCI world portfolios according to (19) using conditional maximum likelihood assuming that the errors have a conditional t distribution. The standard deviation shown is the variation of the prices of risk and moments over time. For the constrained case, the price of variance risk is restricted to be positive and the price of skewness risk is restricted to be negative. The unconstrained case does not impose any restrictions on the sign of the prices. The mean and standard deviation are multiplied by 10^3.

A. Unconstrained

	Time-varying price	Mean	Standard deviation	Negative %
	Price of Variance	−8.122	45.668	59.02%
U.S.	Variance	1.442	0.926	
1941.09–1987.12	Price of Skewness	3.550	3.121	12.82%
	Skewness	−2369.494	1215.883	
	Price of Variance	2235.900	774.140	0.00%
World	Variance	2.173	3.792	
1970.02–1992.12	Price of Skewness	−0.629	33.413	54.91%
	Skewness	−91.007	210.920	

B. Constrained

	Time-varying price	Mean	Standard deviation	Negative %
	Price of Variance	3.351	7.302	
U.S.	Variance	1.639	0.520	
1941.09–1987.12	Price of Skewness	−0.012	0.023	
	Skewness	−3225.457	7025.516	
	Price of Variance	360.970	270.210	
World	Variance	13.755	4.227	
1970.02–1992.12	Price of Skewness	−5.359	9.845	
	Skewness	−145.150	64.559	

positive and negative. Thus, even without constraints, prices of risk for the world portfolio display the right signs.

Combining the prices of variance and skewness risk with conditional variance and skewness, we obtain expected market risk premium implied by the model for the two market portfolios. Figure 3 plots the expected market risk premium implied by the higher conditional moments (with constraints on the prices of risk) along with the statistically fitted market risk premium. The plots show that for the model incorporating skewness is a much better fit for the world portfolio than for the U.S. portfolio. With constraints on the prices of risk, the annualized average implied expected market risk premium for the U.S. is only 1.0% versus average statistically fitted market risk premium of 8.4%. For the world the comparable numbers are 6.0% versus 5.1%. Thus, the constraints lead to unreasonable estimates for the U.S. portfolio. The figures also show that the implied expected market risk premium appears to approximate the statistically fitted risk premium for the world much better than for the U.S. In fact, with constraints on prices of risk, the correlation between the two is 47% for the world and 0% for the U.S. This corroborates the results in Table 4 that the average likelihood for the world portfolio return is higher than for the U.S. portfolio return. Thus, as we found in the GMM-based methodology, the asset pricing model is more successful in explaining the variation of the expected market risk premium for the world portfolio. As before, to check if the differing sample periods are responsible for the difference in performance, we compare the performance over the common period 1970.02–1987.12. Using the constrained prices of risk, the correlation between statistically fitted expected market risk premium and the theoretically implied expected market risk premium over this period is 47% for the world and 6% for the U.S. portfolio.

4.3. Can the negative ex ante market risk premium be explained

To understand whether the incidences of ex-ante negative market risk premium can be explained by skewness, we need to analyze the interactions of the prices of risk and the conditional moments. We consider two sets of periods, (1) when the statistically fitted market risk premium is negative and (2) when the statistically fitted market risk premium minus the expected market risk premium implied by conditional variance is negative. These are the periods of interest to us given that we are challenging the model to explain negative ex ante market risk premia. We can not expect any model to explain all these periods. However, the relative success for different models and different portfolios permits us to draw some conclusions. For these periods we compute
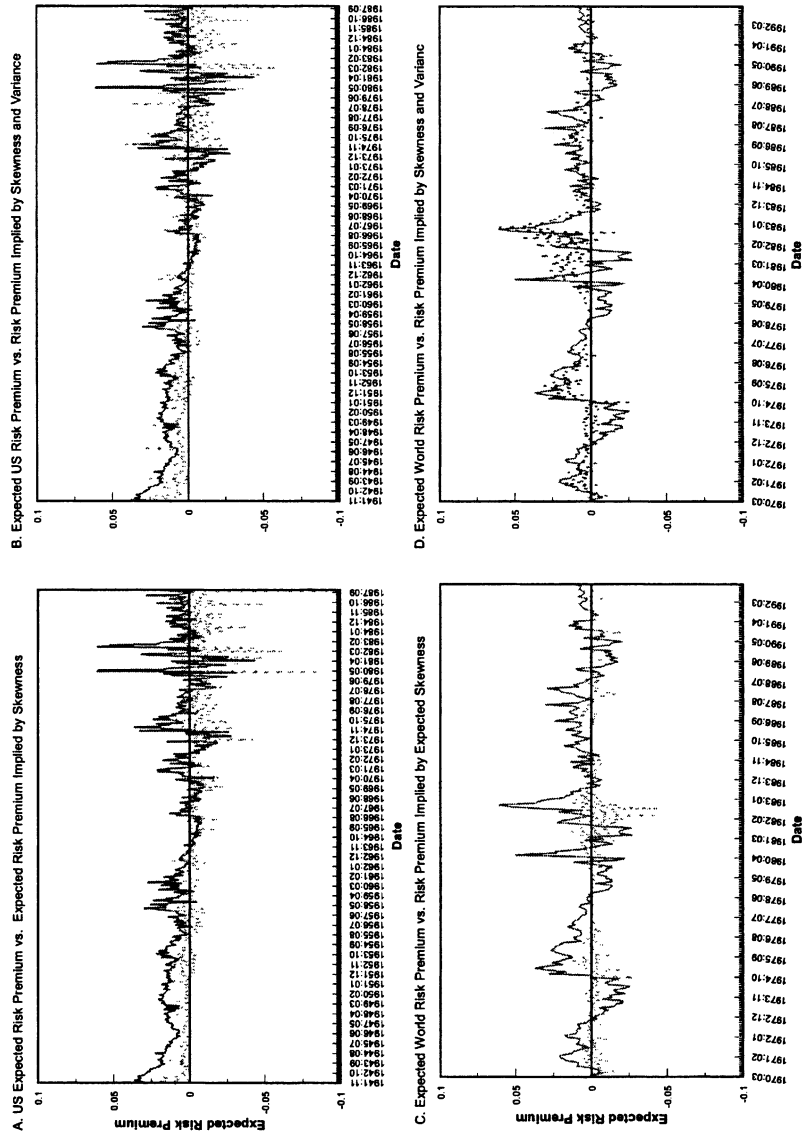
Fig. 3.

the expected market risk premium implied by conditional skewness for both the U.S. and world portfolios with the two econometric methodologies, GMM and conditional maximum likelihood.

Of the 554 months in the U.S. market portfolio sample, the statistically fitted market risk premium is negative 143 times (25.8% of the months). For the world portfolio, we have 275 months of data, of which 89 months (32.1%) show negative statistically fitted market risk premiums. When we constrain the prices of risk in maximum likelihood estimation in accordance with the asset pricing theory, the market risk premium implied by skewness is negative for 71 of the 143 negative months for U.S. and 87 of the 89 negative months for the world. Using GMM methodology and constraints, these proportions are less impressive.

The difference between the statistically fitted market risk premium and the expected market risk premium implied by conditional variance can be viewed as the part of the market risk premium attributable to skewness risk. Using maximum likelihood and constraints on prices of risk, the U.S. portfolio has negative difference between statistically fitted market risk premium and expected market risk premium implied by variance for 143 months (25.8% of the entire 554 month sample). For the world, 128 months (46.6% of the 275 month sample) show such negativity. For the world, the expected market risk premium implied by skewness is negative for 125 or 99.2% of the 128 months. For the U.S. the comparable number is 71 or 49.6% months out of 143. Using GMM, the performance of the asset pricing model in explaining the episodes of negativity is much less impressive. But again, the performance is better for the world portfolio than the U.S. portfolio.

Thus, the results show that the hypothesis that conditional skewness explains the negativity of the expected market risk premium has greater validity for the world portfolio than the U.S. portfolio. Furthermore, when we constrain the prices of risk in accordance with asset pricing theory, the model is more successful. These results are consistent with our findings in the previous sections showing that skewness itself is much more important for the world portfolio than the U.S. portfolio.

The dichotomy between the U.S. and world portfolios may be due to the different periods over which the data is available. But this hypothesis is untenable in light of the significantly poorer performance of the model with conditional skewness in explaining the time-series variation of the expected market risk premium for the U.S. than for the world portfolio when data over the same period are used. Thus, there could be a fundamental difference between the two portfolios. Our tests should hold true for a valid proxy for the market portfolio. Given the degree of integration in international capital

Table 6. Explanation of Negative Expected Risk Premia

The results in explaining the negative expected risk premia are summarized. The mean is modeled linearly in the instruments. The price of variance risk, $\lambda_{1,t}$, and the price of skewness risk, $\lambda_{2,t}$ are estimated as linear functions in both maximum likelihood and GMM. For maximum likelihood variance and skewness are modeled as GARCH(2,2). For GMM, variance and skewness are estimated with non-parametric kernels.

| | | Maximum likelihood | | GMM | |
		US	World	US	World
Number of observations		554	275	554	275
Number of negative ex ante risk premia		143	89	143	89
$E_t[r_{M,t+1}] < 0$		25.8%	32.1%	25.8%	32.1%
$\lambda_{2,t}\text{Skew}[r_{M,t+1}] < 0$	Unconstrained	39.2%	63.3%	64.4%	60.7%
when $E_t[r_{M,t+1}] < 0$	Constrained	49.6%	98.9%	27.3%	33.7%
$E_t[r_{M,t+1} - \lambda_{1,t}\text{Var}[r_{M,t+1}] < 0$	Unconstrained	25.8%	45.1%	34.3%	72.2%
	Constrained	25.8%	46.6%	47.6%	50.9%
$\lambda_{2,t}\text{Skew}[r_{M,t+1}] < 0$	Unconstrained	39.2%	66.9%	59.3%	46.7%
when $E_t[r_{M,t+1}] - \lambda_{1,t}\text{Var}[r_{M,t+1}] < 0$	Constrained	49.7%	99.2%	34.4%	41.0%
$E_t[r_{M,t+1}]$ and $E_t[r_{M,t+1}] - \lambda_{1,t}\text{Var}[r_{M,t+1}] < 0$	Unconstrained	25.8%	32.1%	18.0%	24.9%
	Constrained	25.8%	32.1%	25.8%	32.1%
$\lambda_{2,t}\text{Skew}[r_{M,t+1}] < 0$	Unconstrained	39.2%	63.3%	60.6%	60.3%
when $E_t[r_{M,t+1}]$ and $E_t[r_{M,t+1}] - \lambda_{1,t}\text{Var}[r_{M,t+1}] < 0$	Constrained	49.6%	98.9%	27.3%	33.7%

markets and increasingly unimpeded flow of capital across borders, the world portfolio should be a better proxy for the market. The U.S. portfolio is only a component of the world portfolio.

4.4. Extensions

The model and results presented in this paper have two interesting extensions. For both U.S. and world portfolios, the unconditional skewness has been negative over the periods considered. Since the price of skewness should be negative, the implied risk premium from skewness should be positive as well. Hence, the implied market risk premium from variance and skewness should be higher than from variance alone for a large number of the periods. This has the potential of explaining the equity market risk premium puzzle – the fact that unconditionally the equity market risk premium is higher than what one should expect from variance alone. Finally, the model and results of this paper will also have implications for asset allocation and portfolio analysis. Instead of analyzing asset returns in a conditional mean-variance framework, a richer conditional mean-variance-skewness framework may be employed.

5. CONCLUSIONS

Recent research has found evidence against the single factor asset pricing model. The evidence on the time-series properties of the market risk premium seems to suggest that the expected market risk premium is significantly negative in some states of the world. Cross-sectional tests of the single factor asset pricing model have shown that systematic risk as measured by the covariance (or the beta) with the market does not explain all of the cross-sectional variation in expected excess returns. We provide a possible explanation for these departures from the single factor asset pricing model. Our intuition is that if investors know that the asset returns have conditional skewness at time t, excess asset returns should include a component attributable to conditional skewness. Our asset pricing model formalizes this intuition by incorporating conditional skewness. This model can explain much of the time-series variation in the expected market risk premium.

We estimate this model for the U.S. and world portfolios using two methodologies. The first is a generalized method of moments methodology that does not require moment specifications for the portfolio returns. The second is a conditional maximum likelihood methodology that fully specifies the time-varying prices of risk and the dynamics of the conditional moments. Our results show that conditional skewness is important and, when combined with the time

varying price of skewness, explains some of the negative ex ante market risk premiums. Skewness is more important for the world market risk premium. Thus, for a model that requires us to specify a market portfolio, the world portfolio appears to be a better proxy for the market. Consistent with this finding, the asset pricing model with conditional skewness has greater success in explaining the negativity of the expected market risk premium for the world portfolio.

NOTES

1. Kane (1977) shows that if the asset returns follow diffusion processes, then mean-variance criterion is adequate only if continuous rebalancing is permitted. Therefore, without continuous rebalancing the moments of the discrete returns should include skewness as well.

2. Also see Ingersoll (1990, pp. 199–201) for an alternative derivation for the unconditional three-moment CAPM. Sears & Wei (1985) also derive the unconditional market risk premium as the sum of the market prices of variance and skewness multiplied by the moments. Since the riskfree rate does not possess variance or skewness, we can use variance and skewness of either the total returns or excess returns.

3. Non-increasing absolute risk aversion implies that its derivative should be less than or equal to 0. $U''' \geq 0$ is a necessary condition to satisfy this. Also see Scott & Horvath (1980) for a discussion of the preference of moments beyond variance.

4. As an alternative, we also estimated conditional variance and skewness using a GARCH(2,2) specification similar to the specifications used in the conditional maximum likelihood estimation and used in Harvey & Siddique (1999). We do not report the results but they do not change substantively using these fitted moments.

5. Pagan & Hong (1989) use a non-parametric ARCH specification for computing conditional variance. Also see Pagan & Schwert (1990) and Harvey (1991).

6. The linearity of the expected market risk premium follows Gibbons & Ferson (1985) and Campbell (1987).

7. Our approach extends the generalized autoregressive conditional heteroskedasticity, GARCH, to include autoregressive conditional skewness, and autoregressive conditional kurtosis. For additional details on this methodology see Harvey & Siddique (1995).

8. Boudoukh, Richardson & Smith (1993) also use lagged volatility for a part of the sample. But it is not available for the whole sample.

9. This reflects the crashes in the historical returns.

10. Boudoukh, Richardson & Smith (1993) find that they can not reject negativity of the expected market risk premia using the same data. Their results do not rely on predictability and they do not explicitly compute the expected market risk premia. However, one of our objectives is to explain the negativity of the expected market risk premia. For this purpose, we need to compute the expected market risk premia.

11. We have computed all of the results using the non-parametric risk premiums which are available on request.

12. *Saturation ratio* is the number of observations divided by the number of parameters to estimate and the number of unique elements in the weighting matrix.

13. We have also analyzed the properties of the conditional moments such as smoothness as diagnostics for our estimation.

14. In an economy with single consumption good, the expected market risk premium can not be negative. For the existence of a negative expected market risk premium in a single factor model, non-traded sector of the economy is important.

15. Results using GMM and unconstrained risk prices are not presented but are available.

REFERENCES

Andrews, D. W. K. (1991). Heteroscedasticity and autocorrelation consistent covariance matrix estimation. *Econometrica*, *59*, 817–858.

Arrow, K. J. (1964). Aspects of the theory of risk-bearing. *lectures*. Helsinki, Finland.

Bansal, R., & Viswanathan, S. (1993). No arbitrage and arbitrage pricing. *Journal of Finance*, *48*, 1231–1262.

Black, F. (1972). Capital market equilibrium with restricted borrowing. *Journal of Business*, *44*, 444–455.

Boudoukh, J., Richardson, M., & Smith, T. M. (1993). Is the ex ante risk premium always positive? *Journal of Financial Economics*, *34*, 387–408.

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*, 373–400.

Carhart, M. M., & Krail, R. J. (1993). Testing the conditional CAPM. *Unpublished manuscript*. University of Chicago.

Carhart, M. M., Krail, R. J., Stevens, R. I., & Welch, K. E. (1994). Testing the conditional CAPM. *Unpublished manuscript*. University of Chicago.

Chan, L., Hamao, Y., & Lakonishok, J. (1991). Fundamentals and stock returns in Japan. *Journal of Finance*, *46*, 1739–1764.

Cochrane, J. (1993). Unpublished lecture notes, University of Chicago.

Christie, A. A. (1982). The stochastic behavior of common stock variances: value, leverage, and interest rate effects. *Journal of Financial Economics*, *23*, 407–432.

Dumas, B., & Solnik, B. (1995). The world price of foreign exchange risk. *Journal of Finance*, *50*, 445–479.

Fama, E. F., & French, K. R. (1989). Business conditions and expected returns on stocks and bonds, *Journal of Financial Economics*, *59*, 23–49.

Fama, E., & French, K. (1992). The cross-section of expected stock returns. *Journal of Finance*, *59*, 427–465.

French, K., Schwert, W., & Stambaugh, R. (1995). Expected stock returns and volatility. *Journal of Financial Economics*, *19*, 3–29.

Friend, I., & Westerfield, R. (1980). Co-skewness and capital asset pricing. *Journal of Finance*, *35*, 897–914.

Gibbons, M. R., & Ferson, W. E. (1985). Tests of asset pricing models with changing expectations and an unobservable market portfolio. *Journal of Financial Economics*, *10*, 3–27.

Glosten, L. R., Jagannathan, R., & Runkle, D. E. (1993). On the relation between expected value and the volatility of the nominal excess return on stocks. *Journal of Finance*, *48*, 1779–1801.

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

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

Hansen, L. P., & Richard, S. F. (1987). The role of conditioning information in deducing testable restrictions implied by dynamic asset pricing models. *Econometrica*, *55*, 587–613.

Harrison, M., & Kreps, D. (1979). Martingales and arbitrage in multi-period securities markets. *Journal of Economic Theory*, *20*, 381–408.

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., & Siddique, A. (1999). Autoregressive conditional skewness. *Journal of Financial and Quantitative Analysis*, *34*, 465–487.

Harvey, C. R., & Siddique, A. (2000). Conditional skewness in asset pricing tests. *Journal of Finance*. forthcoming

Ingersoll J. E. Jr. (1990). *Theory of financial decision making*. Totowa, New Jersey.

Kandel, S., & Stambaugh, R. F. (1995). Bayesian inference and portfolio efficiency. *Review of Financial Studies*, *8*, 1–53.

Kane, A. (1977). Skewness preference and portfolio choice. *Journal of Financial and Quantative Analysis*, *17*, 15–25.

Kraus, A., & Litzenberger, R. H. (1976). Skewness preference and the valuation of risk assets. *Journal of Finance*, *31*, 1085–1100.

Lim, K-G. (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.

Nelson, D. B. (1991). Conditional heteroskedasticity in asset returns. *Econometrica*, *59*, 347–370.

Nummelin, K. (1994). Expected asset returns and financial risks. *Dissertation*. Swedish School of Economics and Business Administration (Svenska Handelshögskolan), Helskinki, Finland.

Ostdiek, B. (1994). The world ex ante risk premium: an empirical investigation. *Unpublished manuscipt*. Fuqua School of Business, Duke University.

Pagan, A. R., & Hong, Y. S. (1991). Non-parametric estimation and the risk premium. In: *Non-parametric and semiparametric methods in econometrics and statistics*, (51–75). Cambridge University Press, Cambridge.

Pagan, A. R., & Schwert, G. W. (1990). Alternative models for conditional stock volatility. *Journal of Econometrics*, *45*, 267–290.

Roll, R., & Ross, R. (1994). On the cross-sectional relaion between expected returns and betas. *Journal of Finance*, *49*, 101–121.

Roll, R. (1977). A critique of the asset pricing theory's tests, Part I: On past and potential testability of theory. *Journal of Financial Economics*, *4*, 129–176.

Ross, R. (1977). The capital asset pricing model (CAPM), short-sale restrictions and related issues. *Journal of Finance*, *32*, 177–183.

Scott, R. C., & Horvath, P. A. (1980). On the direction of preference for moments of higher order than the variance. *Journal of Finance*, *35*, 915–919.

Sears, R. S., & Wei, K. C. J. (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.