MANAGEMENT SCIENCE
Vol. 56, No. 11, November 2010, pp. 2031–2049
ISSN 0025-1909 | EISSN 1526-5501 | 10 | 5611 | 2031

informs®

DOI 10.1287/mnsc.1100.1237

Conditional Coskewness in Stock and Bond Markets: Time-Series Evidence

Jian Yang
Business School, University of Colorado Denver, Denver, Colorado 80217, jian.yang@ucdenver.edu

Yinggang Zhou
Faculty of Business Administration, Chinese University of Hong Kong, Hong Kong, People's Republic of China, ygzhou@baf.cuhk.edu.hk

Zijun Wang
Private Enterprise Research Center, Texas A&M University, College Station, Texas 77843, z-wang@neo.tamu.edu

In the context of a three-moment intertemporal capital asset pricing model specification, we characterize conditional coskewness between stock and bond excess returns using a bivariate regime-switching model. We find that both conditional U.S. stock coskewness (the relation between stock return and bond volatility) and bond coskewness (the relation between bond return and stock volatility) command statistically and economically significant negative ex ante risk premiums. The impacts of stock and bond coskewness on the conditional stock and bond premiums are quite robust to various model specifications and various sample periods, and also hold in another major developed country (the United Kingdom). The findings also carry important implications for portfolio management.

Key words: regime switching; conditional coskewness; intertemporal asset pricing; stock and bond comovements
History: Received October 6, 2009; accepted July 17, 2010, by Wei Xiong, finance. Published online in *Articles in Advance* October 11, 2010.

1. Introduction

The most fundamental investment decision is probably the allocation of funds across different asset classes, particularly between the two most important ones: stocks and bonds. It is widely recognized that the comovement between stock and bond markets is crucial for portfolio management. The comovement also carries important theoretical implications for asset pricing, because stock market returns can capture market risk, whereas bond market returns may represent the time-varying investment opportunity in Merton's (1973) intertemporal capital asset pricing model (ICAPM) (e.g., Scruggs 1998, Scruggs and Glabadanidis 2003, Gerard and Wu 2006).[1] Nevertheless, the stock–bond relation is often explored in the context of the correlation between stock and bond returns (e.g., Scruggs and Glabadanidis 2003; Connolly et al. 2005, 2007). Unfortunately, stock–bond comovement may not be fully captured by correlation, because correlation is a linear measure of dependence that is applicable only when investors display mean-variance preferences or, equivalently, when returns follow a multivariate normal distribution. It is well known, however, that asset returns usually have long tails. Similarly, many existing ICAPM studies have limitations in that they typically focus on the mean-variance framework.

Various attempts have thus been made to seek appropriate measures beyond such a linear measure as correlation and the mean-variance asset pricing

[1] Merton (1973) first proposed to use bond market returns as a proxy for the changing investment opportunities. Several recent studies (e.g., Guo et al. 2009) use different proxies for the investment opportunities and provide further support for Merton's (1973) ICAPM in the time-series context. Certainly, one could also consider both stock and bond markets as major components of a very broadly defined combined stock–bond market portfolio. Nevertheless, no such data on the combined portfolio are available throughout the entire sample period of 150 years, nor are necessary inputs (such as respective market capitalizations of the stock and bond markets) available for self-construction. More importantly, as discussed below, Merton's (1973) ICAPM framework is apparently more fitting for this study and also well in line with some most recent studies on coskewness.

framework.[2] In particular, although it has been long demonstrated theoretically that investors may seek higher return and lower volatility as well as higher (positive) skewness (Rubinstein 1973, Kraus and Litzenberger 1976), the importance of the preference for skewness generated much renewed interest recently (e.g., Chiu 2005, Briec et al. 2007, Mitton and Vorkink 2007, Brunnermeier et al. 2007, Barberis and Huang 2008). Although asset return skewness is well documented in the literature (e.g., Prakash et al. 2003, Charoenrook and Daouk 2005, Xu 2007, Bali et al. 2008, Boyer et al. 2010), different economic mechanisms have been proposed to explain why asset return skewness may matter for an investor's portfolio choice and asset prices. Menezes et al. (1980) provide a choice-theoretic foundation to the preference for a higher third moment. A representative investor with the skewness preference may be decreasingly risk averse with his wealth, and he would appreciate a decrease in downside risk defined as more dispersion above a specific wealth level (e.g., mean), which implies higher positive skewness. Kimball (1990) uses the term "prudence" to point to the skewness preference, which is meant to suggest the propensity to prepare and forearm oneself in the face of uncertainty, in contrast to "risk aversion," which is how much one dislikes uncertainty and would turn away from uncertainty if possible. Under the skewness preference, the precautionary saving motive may be stronger than risk aversion because the precautionary saving would endogenously decrease risk aversion and lead to more willingness to bear risk. Chiu (2005) further provides a general choice-theoretical characterization of representative investors with different degrees of prudence, and suggests that more prudent investors are more willing to accept a higher variance in exchange for a higher skewness. Mitton and Vorkink (2007) also recognize that in an economy with two types of agents, investors with a strong desire for upside potential may sacrifice the mean-variance efficiency for higher skewness exposure, which leads to a lack of diversification and a pricing effect of total skewness (including coskewness and idiosyncratic skewness) in equilibrium.[3]

Because skewness of a portfolio matters to investors, an asset's contribution to the skewness of a broadly diversified portfolio, referred to as "coskewness" with the portfolio, may also be rewarded. Skewness preference further suggests that the representative investor may adjust his diversified portfolio such that an individual security's contribution to the skewness of the market portfolio may become a component of the security's expected returns. Mathematically, as demonstrated in Conine and Tamarkin (1981), both individual assets' skewness (the first term in their Equation (4)) and coskewness between assets (the second term) contribute to the skewness of the portfolio that is composed of these assets (whereas the third item disappears in the case of two assets). In the case that individual assets (e.g., individual stocks) are all components of a reference portfolio (e.g., a stock market portfolio), coskewness of assets with a reference portfolio points to the asset's contribution to the skewness of the reference portfolio (e.g., Harvey and Siddique 2000). Intuitively, as positive (negative) skewness implies a probability of obtaining a large positive (negative) return (relative to a benchmark such as the normal distribution), a positive coskewness of an asset with the market portfolio (or another asset) means when the market portfolio (or another asset) price volatility goes up, the return of this asset also goes up, which partially offsets the large downward movement of the portfolio value and thus increase the portfolio skewness. Thus, individual assets that increase (decrease) the overall portfolio's skewness should have higher (lower) coskewness, and they may be more (less) desirable and should command lower (higher) expected returns, ceteris paribus. Empirically, the conditional coskewnesses of individual stocks and style (e.g., size and book-to-market) portfolios in the U.S. stock market are found to be priced (e.g., Harvey and Siddique 2000, Chung et al. 2006). Vanden (2006) showed that stock coskewness with the option market is also important for asset pricing.

Using a data set spanning one and a half centuries, and thus including a large number of extreme return observations in the United States and the United Kingdom, we find new time-series evidence on coskewness across stock and government bond markets at the aggregate level and their pricing effects. This study makes the following contributions

[2] See, among others, the extreme correlation approach of Longin and Solnik (2001) and the coexceedance approach of Bae et al. (2003).

[3] Another strand of economic theories based on behavioral biases also predict that skewness matters. Brunnermeier and Parker (2005) and Brunnermeier et al. (2007) show that idiosyncratic skewness can be priced if investors form subjective beliefs that optimally trade off the ex ante benefits of optimism against the ex post costs of basing investment on distorted expectations. An investor with optimal expectations biases upward his subjective beliefs about the likelihood of the states associated with the most positively skewed returns, and is willing to pay a higher price and earn a lower average return. The overpricing of positive skewness can even manifest itself in equilibrium because of heterogeneous beliefs. Barberis and Huang (2008) give similar predictions under the cumulative prospect theory. Through a weighting function, investors overweight the tails of the distribution and highly value a positively skewed asset or portfolio. As a result, the skewed asset or portfolio can become overpriced and earn a negative average excess return.

to the literature. First, to the best of our knowledge, this is the first study to explore the extreme stock–bond comovement beyond correlation that is captured by (cross-market) coskewness.[4] It is well known that skewness is about the long tail of a return distribution, and thus coskewness is about the comovement in the long tail. We find that the conditional U.S. bond coskewness (with stocks) is mostly positive over time and higher than the U.S. stock coskewness (with bonds) in the recent post–World War II (WWII) subsample of 1951–2004 and particularly the entire sample of 1855–2004. The higher bond coskewness during these periods suggests that government bond markets are also a good hedge against the volatility of stock market returns. Hence, bonds offer more diversification opportunities than what we think, based on the conventional wisdom that bonds are a good hedge against stock returns because of the relatively low correlation between their returns.

Second, extending the earlier literature (e.g., Harvey and Siddique 2000, Prakash et al. 2003, Vanden 2006, Hashmi and Tay 2007, Bali et al. 2008), we document that cross-market coskewness between stocks and bonds is priced in both the United States and the United Kingdom. Such consideration of pricing effects of the cross-market coskewness can be easily motivated by the arguments that bond market returns should be included as an additional factor in intertemporal asset pricing (Merton 1973), and coskewness of assets with additional factors useful in explaining risky asset returns should also be considered (Vanden 2006).[5] Our framework is also a natural extension of existing ICAPM time-series studies in the mean-variance framework (e.g., Scruggs 1998, Scruggs and Glabadanidis 2003, Gerard and Wu 2006). During the 1951–2004 period, the U.S. bond coskewness is estimated to lead to, on average, a 0.8% decrease in the expected bond premium during the post-war period, the magnitude of which is quite comparable to the average of corresponding unconditional premium on the bond market (1.1%). The main results on pricing effects hold even better during the entire 150-year sample period. Based on the estimates of the very long sample, the U.S. stock coskewness, on average, induces a 4.3% increase in the conditional expected stock premium and the U.S. bond coskewness, on average, leads to a 0.46% decrease in the expected bond premium, the magnitude of which is almost as large as the average of corresponding unconditional premium on these markets. As underscored by Schwert (1990) and Goetzmann et al. (2005), the availability of very long financial time series from both countries helps uncover longer term phenomena and mitigates the "data snooping" problem plaguing the literature focusing on a single country (e.g., the United States) or a rather short sample period (e.g., the more recent period since 1970 when financial market data are more easily available).[6]

Finally, we propose a new approach that helps examine the conditional comovement on a moment-by-moment basis. Although multivariate regime-switching models have become increasingly popular in the literature (e.g., Guidolin and Timmermann 2005; Connolly et al. 2005, 2007), so far few studies have characterized in detail the conditional comoments from multivariate Markov switching models. Specifically, extending earlier works on univariate higher moments (e.g., Timmermann 2000, Perez-Quiros and Timmermann 2001) and unconditional comoments (e.g., Chung et al. 2006), we present a bivariate regime-switching framework for estimating conditional mean, correlation, and coskewness, as well as other conditional moments in general.[7] Such model-based estimates are typically determined

[4] In the case of two different assets such as stocks and bonds, the cross-market coskewness can be interpreted in a different manner from the conventional one as discussed above, because bonds are not part of the stock market portfolio. Given the difficulty of finding a good proxy for the true market portfolio, we attempt to interpret the bond coskewness (with stock) and the stock coskewness (with bonds) as analogue to the correlation (and indeed for this purpose we use standardized coskewness). Specifically, the intuition of coskewness in this context goes as follows: the positive bond coskewness (with stock) means when the stock market is volatile, the bond market returns goes up. In face of stock market uncertainty, investors may flight to the safer Treasury bond market, to prepare and forearm themselves, as suggested by the precautionary saving motive in Kimball (1990), among others. Thus, this desirable characteristic of the bond market to risk-averse and prudent investors would push up the bond price and push down the expected bond return. Such an interpretation of cross-market coskewness extends the existing literature and makes it easy to use in many other contexts, which overcomes the hurdle of using an appropraite proxy for an all-inclusive market portfolio.

[5] As pointed out by Vanden (2006), such additional factors should have an effect on the functional form of the stochastic discount factor. The specific additional factor under consideration is the option returns in Vanden (2006), whereas it is the bond market return in this study. Also, because we focus on time-series analysis, both stock and bond market returns serve as two test assets.

[6] One might be concerned about the possibility that the earlier part of the long data may come from a time when economic systems differed. Somewhat surprisingly, according to Goetzmann et al. (2005), the modern era of global investing has parallels to the pre-WWI era, and the period of 1870 to 1913 was, in some ways, the golden era of global capitalism. This observation is confirmed by our basic results on pricing effects of the cross-market coskewness.

[7] As a noticeable exception, Guidolin and Timmermann (2008) also made a similar (and more general) extension in the context of international CAPM but not intertemporal CAPM (as in this study). Furthermore, our derivation differs from Guidolin and Timmermann (2008) in an important aspect, as the transition probability is time varying in our case, whereas it is time invariant in Guidolin and Timmermann (2008) (see their Equation (7)), which arguably may help better capture time variations of conditional moments and

Table 1 Summary Statistics

Period	Name	Nobs	Mean	SD	Skew	Kurt	Lag1	Lag3	Lag6
				Panel A: United States					
January 1855	STOCK	1,792	0.042	0.606	−0.566	6.264	0.083	−0.051	−0.006
—	BOND	1,792	0.004	0.177	0.607	12.552	0.131	−0.032	0.031
August 2004	SHORT	1,792	0.044	0.030	2.877	30.589	0.838	0.735	0.689
April 1951	STOCK	641	0.056	0.512	−0.701	2.558	0.061	0.011	−0.018
—	BOND	641	0.011	0.269	0.279	3.753	0.124	−0.045	0.025
August 2004	SHORT	641	0.050	0.028	0.989	1.283	0.986	0.951	0.906
				Panel B: United Kingdom					
January 1855	STOCK	1,775	0.001	0.441	0.283	18.603	0.064	0.065	−0.003
—	BOND	1,775	0.004	0.263	0.315	3.125	0.154	0.000	0.020
November 2005	SHORT	1,775	0.044	0.033	1.265	1.374	0.983	0.942	0.900
October 1945	STOCK	722	0.034	0.597	0.209	10.462	0.085	0.067	−0.022
—	BOND	722	0.003	0.339	0.354	0.692	0.220	−0.015	0.016
November 2005	SHORT	722	0.065	0.038	0.417	−0.466	0.990	0.965	0.927

Notes. This table reports summary statistics for U.S. and UK stock and bond excess returns and short rates. STOCK and BOND denote stock and bond excess returns, which are their monthly returns in excess of short rates labeled SHORT. The U.S. stock returns are the log-differences of the Schwert's index (prior to 1926) and the CRSP value-weighted index with dividend (since 1926). The U.S. bond returns are the log-differences of the U.S. 10-year government bond total return index from the GFD. The U.S. short rate is the log-difference of the U.S. commercial/T-bill total return from the GFD. The UK stock returns are the log-differences of the British stock price index (prior to 1970) with dividend added and the MSCI UK total return index (since 1970). The UK bond returns are the log-differences of the UK consol bond total return index from the GFD. The UK short rate is from three sources: (1) the "open market rate of discount at London" from the NBER (1855∼1939), (2) the discount rate on Treasury bills collected by Capie and Webber (1985) (1940∼1974), and (3) the discount rate on three-month Treasury bills (since 1975). All measures are annualized. "Nobs" is the numbers of monthly observations for the series after deleting missing data. For the United States, the data from August 1914 to November 1914 are deleted because stock returns are missing. For the United Kingdom, the data for some months in 1914, 1915, 1916, 1920, 1921, 1936, and 1937 are deleted because stock returns are missing. "SD" denotes standard deviation. "Lag *x*" denotes autocorrelation at *x* period lag.

with considerably more accuracy than estimates of the higher moments obtained directly from realized returns (Guidolin and Timmermann 2008, p. 896).

The rest of this paper is organized as follows. Section 2 describes the data. Section 3 discusses the regime-switching models and derives their conditional moments. Section 4 presents the empirical results. Section 5 makes concluding remarks.

2. Data

The U.S. stock and bond markets are the largest in the world. They are also the best documented and most heavily researched. The United Kingdom had the world's largest financial markets in the 19th and early 20th centuries and remains among the most important in the world. We will focus on the United States and use the United Kingdom as a robustness check. The samples begin in January 1855 and end by August 2004 for the United States and November 2005 for the United Kingdom.

The data are from a variety of sources. Monthly stock returns are calculated using composite indices. For the United States, the index prior to 1926 is from Schwert (1990) (with observations missing during August 1914 to November 1914 because of the closure of the New York Stock Exchange), whereas the later data series (since 1926) is from the Center for Research in Security Prices (CRSP) value-weighted index with dividend. For the United Kingdom, the index prior to 1970 is the monthly British stock price index with dividend yield added, whereas the later data series is the Morgan Stanley Capital International (MSCI) UK total return index from Datastream.[8]

Government bond data are used because they are the dominant segment of the bond markets. The longest available monthly data for government bond returns are from Global Financial Data (GFD). The U.S. Bond returns are calculated using the U.S. 10-year government bond total return index. For the United Kingdom, bond returns are calculated using the UK consol total return index. The UK consol is perpetually funded government debt.

Stock and bond returns are excess returns over and above short rates. As a proxy for the U.S. short rate, bill returns are calculated using the Treasury-bill (T-bill) total return index from the GFD. The UK short rate comes from three sources: (1) the National Bureau of Economic Research (NBER) provides the "open market rate of discount at London" from 1824

comoments. See Connolly et al. (2005) for an application of regime-switching models with time-varying transition probability on the stock–bond return relation.

[8] The series is available at http://fisher.osu.edu/fin/resources_data/data/britmon.dat. There are also some missing observations because the market was closed during WWI and data were not available for some months in 1920, 1921, and 1936. No dividend yield data were available through 1923, so a yield of 5% per annum is assumed, following Dimson et al. (2001). From 1924 to 1969, the Financial Times Actuaries dividend yields are added to the return series.

through 1939; (2) the discount rate on Treasury bills is filled in between 1940 and 1975 using the data by Capie and Webber (1985); (3) Datastream provides the discount rate on three-month Treasury bills from 1975. All monthly log returns are annualized.

Table 1 reports summary statistics of the stock and bond excess returns as well as short rates for both the entire sample and the subsample period of 1951 through 2004. The subsample starts in 1951 following the Fed-Treasury Accord. The U.S. stocks typically yielded a higher premium than bonds but were more volatile than bonds. Skewnesses of U.S. stocks are lower than those of fixed income securities. In particular, the skewness of the U.S. stock premium is negative, whereas the skewness of the U.S. bond premium is positive. Both are significantly different from zero, the expected value of skewness for a normal distribution (the sample errors are 0.058 and 0.097 when the sample sizes are 1,792 and 641, respectively). As for serial correlations, the stock and bond excess returns have a small autocorrelation at lag one. The short rates are highly persistent. A similar pattern also applies to the United Kingdom, except that the mean is lower for the stock premium than for the bond premium, which may be a reflection of UK stock data problems as discussed above.

3. Empirical Methodology

Measuring conditional skewness/coskewness is challenging, and various attempts have been made in the literature. Harvey and Siddique (1999) directly model time-varying dynamics of conditional skewness in a generalized autoregressive conditional heteroskedasticity (GARCH)-like framework. However, past skewness is probably not a good predictor of future skewness because skewness generally is not persistent over time. With the focus on the cross-sectional setting, Harvey and Siddique (2000) model conditional coskewness of an asset with the market portfolio as the conditional variance between the asset return and the square of the market return, whereas the conditional expectation is formed based on the autoregression with two lags. Dittmar (2002) also follows a similar approach to model conditional coskewness, whereas the conditional expectation is addressed in the way that relevant coefficients are linear functions of time t information variables. Boyer et al. (2010) directly compute idiosyncratic skewness and form the expectation based on predictive regressions using a number of cross-sectional firm characteristic variables. Xing et al. (2010) exploit the risk-neutral measure of conditional skewness from option data and measure the conditional skewness of the aggregate market based on the slope of the volatility smile for the S&P 500 index options. Our use of regime-switching models in this context is well in line with Timmermann (2000) and Guidolin and Timmermann (2008). As pointed out by Guidolin and Timmermann (2008, p. 896), such a regime-switching model allows us to obtain precise conditional estimates in a flexible manner, because it captures skewness and other higher moments as a function of the mean, variance, and persistence parameters of the underlying states. Thus, it may provide considerably more accuracy than estimates of the higher moments obtained directly from realized returns.

The empirical methodology used in this study follows a three-pass procedure. First, we estimate a bivariate regime-switching model for stock and bond excess returns. Second, we derive the conditional moments and comoments of stock and bond excess returns using the bivariate regime-switching model. From an econometric perspective, regime-switching models belong to a general class of mixture distributions and can generate higher-order and time-varying moments through the use of a simple distribution, such as the normal distribution. Finally, we examine the pricing behavior of estimated coskewness series by conducting time-series regressions of the expected stock (or bond) premium on the expected stock (or bond) coskewness, controlling for the expected correlation, volatility, and skewness and correcting for the errors-in-variables problem.

3.1. Regime-Switching Models

It is widely recognized that there is substantial time variation and regime dependence in the relation between stock and bond returns. Following the success of the univariate regime-switching model, a variety of multivariate regime-switching models have become increasingly popular in asset allocation between stocks and bonds.

The basic bivariate regime-switching model has the following general form:

$$\begin{gathered}\mathbf{r}_t = \boldsymbol{\mu}_{it} + \boldsymbol{\varepsilon}_{it},\\ \boldsymbol{\mu}_{it} = E(\mathbf{r}_t \mid s_t = i, \mathbf{F}_{t-1}),\\ \boldsymbol{\varepsilon}_{it} \mid \mathbf{F}_{t-1} \sim (\mathbf{0}, \mathbf{H}_{it}),\end{gathered} \tag{1}$$

where $\mathbf{r}_t = (r_t^s, r_t^b)'$ is a 2×1 vector of stock and bond excess returns at time t; s_t is the unobserved regime at time t; and $\boldsymbol{\mu}_{it} = (\mu_{it}^s, \mu_{it}^b)'$ is a 2×1 vector of means given regime i conditioned on the past information set $\mathbf{F}_{t-1}$. Note that $\mathbf{F}_{t-1}$ does not contain s_t or lagged values of s_t. In this study, s_t takes one of two values (1 or 2); $\boldsymbol{\varepsilon}_{it} = (\varepsilon_{it}^s, \varepsilon_{it}^b)'$ is a 2×1 vector of innovations given regime i; and $\mathbf{H}_{it}$ is the conditional variance–covariance matrix of r_t given regime i because

$$\operatorname{var}(\mathbf{r}_t \mid s_t = i, \mathbf{F}_{t-1}) = \operatorname{var}(\boldsymbol{\varepsilon}_{it} \mid \mathbf{F}_{t-1}) = \mathbf{H}_{it}, \quad i \in \{1, 2\}. \tag{2}$$

The latent regime s_t is usually parameterized as a first-order Markov chain. The simplest model assumes that the state transitions are constant over time. However, recent research (e.g., Gray 1996) suggests that the flexibility gained by allowing time-varying transition probabilities can be very substantial. The time-varying transition probabilities conditional on $\mathbf{F}_{t-1}$ can be written as

$$\begin{aligned} &\Pr(S_t = j \mid s_{t-1} = i, \mathbf{F}_{t-1}) = p_{ij,t}, \quad i, j \in \{1, 2\}, \\ &0 \le p_{ij,t} \le 1, \quad \sum_{j=1}^{2} p_{ij,t} = 1 \qquad \text{for all } i. \end{aligned} \tag{3}$$

Let us collect population parameters in conditional means, variance–covariance matrices, and transition probabilities in a vector $\boldsymbol{\theta}$. If the process is governed by regime i at date t, then the conditional density of $\mathbf{r}_t$ for regime i is given by

$$g_{it} = f(\mathbf{r}_t \mid s_t = i, \mathbf{F}_{t-1}, \theta). \tag{4}$$

The mixed conditional density g_t is obtained by summing up the density functions conditional on the state i, using the respective state probabilities as weights:

$$g_t = \sum_{i=1}^{2} g_{it} p_{it}, \tag{5}$$

where $p_{it} = \Pr(s_t = i \mid \mathbf{F}_{t-1}, \boldsymbol{\theta})$ is the conditional state probability, which measures how likely the process is to be under regime i on date t given the information through date $t-1$. The conditional state probability can be derived recursively:

$$\begin{aligned} p_{1t} &= (1 - p_{22,t}) q_{2,t-1} + p_{11,t} q_{1,t-1}, \\ p_{2t} &= (1 - p_{11,t}) q_{1,t-1} + p_{22,t} q_{2,t-1}, \end{aligned} \tag{6}$$

where $q_{it} = \Pr(S_t = i \mid \mathbf{F}_t, \boldsymbol{\theta})$ denotes Bayes' inference about the probability of regime i based on data obtained through date t and based on knowledge of $\boldsymbol{\theta}$.

$$q_{1t} = \frac{g_{1t} p_{1t}}{g_{1t} p_{1t} + g_{2t} p_{2t}}, \qquad q_{2t} = \frac{g_{2t} p_{2t}}{g_{1t} p_{1t} + g_{2t} p_{2t}}. \tag{7}$$

Then, the log-likelihood function can be written as $l = \sum_{t=1}^{T} g_t$. The quasi-maximum likelihood estimates (QMLEs) can be obtained by maximizing l with respect to the parameters. See Hamilton (1989) for details on regime-switching model specification and estimation. Later we are also interested in the smoothed inferences about the regime probabilities for date t based on whole sample information, which can be calculated using an algorithm of backward iteration developed by Kim (1993).

3.2. Conditional Moments of Regime-Switching Models

Without normality for the distribution of $\mathbf{r}_t$, correlation is an insufficient measure of comovement. Alternative measures for higher-order comovement are needed. This subsection derives the conditional moments of the bivariate regime-switching model in general, and the conditional volatility, correlation, skewness, and coskewness in particular. All proofs are given in the appendix.

Proposition 1. *Suppose $\mathbf{r}_t$ follows the bivariate Markov switching process $\mathbf{r}_t = \boldsymbol{\mu}_{it} + \boldsymbol{\varepsilon}_{it}$, $i \in \{1, 2\}$; then, the centered conditional moments of the process are given by*

$$\begin{aligned} E[(r_t^s - \mu_t^s)^k (r_t^b - \mu_t^b)^l \mid \mathbf{F}_{t-1}, \boldsymbol{\theta}] &= \sum_{i=1}^{2} p_{it} \left\{ \sum_{m=0}^{k} \sum_{n=0}^{l} [C_k^m C_l^n (\mu_{it}^s - \mu_t^s)^{k-m} \right. \\ &\quad \left. \cdot (\mu_{it}^b - \mu_t^b)^{l-n} \varphi(m, n; i)] \right\}, \end{aligned} \tag{8}$$

where $C_k^m = k!/((k-m)!\,m!)$ and $C_l^n = l!/((l-n)!\,n!)$. And $\varphi(m, n; i) = E[(\varepsilon_{it}^s)^m (\varepsilon_{it}^b)^n]$ follows the two-dimensional recursion below:

$$\begin{aligned} &\varphi(0, 0; i) = 1, \\ &\varphi(1, 0; i) = \varphi(0, 1; i) = 0, \\ &\varphi(1, 1; i) = \sqrt{h_i^s h_i^b}\, \rho_i, \\ &\varphi(m, 0; i) = (m-1) h_i^s \varphi(m-2, 0; i), \quad \text{if } m \ge 2, \\ &\varphi(0, n; i) = (n-1) h_i^b \varphi(0, n-2; i), \quad \text{if } n \ge 2, \\ &\varphi(m, n; i) = n \sqrt{h_i^s h_i^b}\, \rho_i \varphi(m-1, n-1; i) \\ &\qquad + (m-1) h_i^s \varphi(m-2, n; i), \quad \text{if } m, n \ge 2. \end{aligned}$$

This proposition extends the counterpart argument of Timmermann (2000) to a bivariate and conditional context. Because researchers are often particularly interested in conditional volatility, skewness, correlation, and coskewness, we characterize these moments more explicitly for the specified mixed normal model.

Corollary 1 (Conditional Means, Volatilities, and Skewnesses). *The conditional means of stock and bond excess returns are*

$$\mu_t^{s(b)} = \sum_{i=1}^{2} p_{it} \mu_{it}^{s(b)}, \tag{9}$$

where "s" stands for stock and "b" for bond, and $p_{2t} = 1 - p_{1t}$. As a proxy for volatilities, the conditional standard deviations of stock and bond excess returns are given by (10),

$$sd_t^{s(b)} = \sqrt{E[(r_t^{s(b)} - \mu_t^{s(b)})^2 \mid \mathbf{F}_{t-1}, \boldsymbol{\theta}]}, \tag{10}$$

where $E[(r_t^{s(b)} - \mu_t^{s(b)})^2 \mid \mathbf{F}_{t-1}, \boldsymbol{\theta}] = p_{1t} h_1^{s(b)} + p_{2t} h_2^{s(b)} + p_{1t} p_{2t} (\mu_{1t}^{s(b)} - \mu_{2t}^{s(b)})^2$, and $h_1^{s(b)}$ and $h_2^{s(b)}$ are conditional variances at states 1 *and* 2, *respectively, as defined in*

Equation (2). *The conditional skewnesses of stock and bond excess returns are given by* (11)

$$sk_t^{s(b)} = \frac{E[(r_t^{s(b)} - \mu_t^{s(b)})^3 \mid \mathbf{F}_{t-1}, \boldsymbol{\theta}]}{(sd_t^{s(b)})^3}, \tag{11}$$

where

$$\begin{aligned} &E\big[(r_t^{s(b)} - \mu_t^{s(b)})^3 \mid \mathbf{F}_{t-1}, \boldsymbol{\theta}\big] \\ &\quad = p_{1t}p_{2t}(\mu_{1t}^{s(b)} - \mu_{2t}^{s(b)})\big\{3[(h_1^{s(b)})^2 - (h_2^{s(b)})^2] \\ &\qquad + (p_{2t} - p_{1t})(\mu_{1t}^{s(b)} - \mu_{2t}^{s(b)})^2\big\}. \end{aligned}$$

Noteworthy is that conditional skewness in Equation (11) is essentially skewness from a mixture of normal distributions. Similar equations have been derived in the univariate case by Perez-Quiros and Timmermann (2001). Note that conditional means are better proxies for expected stock and bond excess returns compared to realized excess returns. Also note that the conditional state probabilities are sufficient for time-varying volatilities but not for time-varying skewnesses. More specifically, if the means in states 1 and 2 are identical, then time-varying volatilities are the weighted average of two constant variances with conditional state probabilities as weights, whereas the conditional skewnesses may be zeroes.

Corollary 2 (Conditional Covariance and Correlation). *The conditional covariance between stock and bond excess returns is given by*

$$\begin{aligned} &E[(r_t^s - \mu_t^s)(r_t^b - \mu_t^b) \mid \mathbf{F}_{t-1}, \boldsymbol{\theta}] \\ &\quad = p_{1t}\mathrm{cov}_1 + p_{2t}\mathrm{cov}_2 + p_{1t}p_{2t}[(\mu_{1t}^s - \mu_{2t}^s)(\mu_{1t}^b - \mu_{2t}^b)], \end{aligned} \tag{12}$$

where $\mathrm{cov}_i = \rho_i\sqrt{h_i^s h_i^b}$ *is the covariance in regime i. Thus, the conditional correlation is given by*

$$\begin{aligned} \rho_t = {} & \big(p_{1t}\mathrm{cov}_1 + p_{2t}\mathrm{cov}_2 + p_{1t}p_{2t}[(\mu_{1t}^s - \mu_{2t}^s)(\mu_{1t}^b - \mu_{2t}^b)]\big) \\ & \cdot \big\{[p_{1t}h_1^s + p_{2t}h_2^s + p_{1t}p_{2t}(\mu_{1t}^s - \mu_{2t}^s)^2]^{1/2} \\ & \cdot [p_{1t}h_1^b + p_{2t}h_2^b + p_{1t}p_{2t}(\mu_{1t}^b - \mu_{2t}^b)^2]^{1/2}\big\}^{-1}. \end{aligned} \tag{13}$$

Similar to conditional volatilities, if the means of stock (or bond) premium in states 1 and 2 are identical, then time-varying covariance is the weighted average of two constant covariances with conditional state probabilities as weights. If the discrepancy between stock means has a sign opposite to that of the discrepancy between bond means, their cross product is negative. This implies a better opportunity for diversification. Also, as argued in Guidolin and Timmermann (2008), the conditional correlation from regime-switching models is consistent with the evidence of asymmetric correlations in the literature (e.g., Longin and Solnik 2001).

Corollary 3 (Standardized Conditional Coskewness). *The standardized conditional coskewness between stock and bond excess returns is given by the following two measures together:*

the stock coskewness,

$$\gamma_t^s = \frac{E[(r_t^s - \mu_t^s)(r_t^b - \mu_t^b)^2 \mid \mathbf{F}_{t-1}, \boldsymbol{\theta}]}{\{E[(r_t^s - \mu_t^s)^2 \mid F_{t-1}, \theta]\}^{1/2} E[(r_t^b - \mu_t^b)^2 \mid \mathbf{F}_{t-1}, \theta]}, \tag{14}$$

where

$$\begin{aligned} &E[(r_t^s - \mu_t^s)(r_t^b - \mu_t^b)^2 \mid \mathbf{F}_{t-1}, \theta] \\ &\quad = p_{1t}p_{2t}[(p_{2t} - p_{1t})(\mu_{1t}^s - \mu_{2t}^s)(\mu_{1t}^b - \mu_{2t}^b)^2 + (\mu_{1t}^s - \mu_{2t}^s) \\ &\qquad \cdot (h_1^b - h_2^b) + 2(\mu_{1t}^b - \mu_{2t}^b)(\mathrm{cov}_1 - \mathrm{cov}_2)], \quad \textit{and} \end{aligned}$$

the bond coskewness,

$$\gamma_t^b = \frac{E[(r_t^s - \mu_t^s)^2(r_t^b - \mu_t^b) \mid \mathbf{F}_{t-1}, \boldsymbol{\theta}]}{E[(r_t^s - \mu_t^s)^2 \mid F_{t-1}, \theta]\{E[(r_t^b - \mu_t^b)^2 \mid \mathbf{F}_{t-1}, \boldsymbol{\theta}]\}^{1/2}}, \tag{15}$$

where

$$\begin{aligned} &E[(r_t^s - \mu_t^s)^2(r_t^b - \mu_t^b) \mid \mathbf{F}_{t-1}, \boldsymbol{\theta}] \\ &\quad = p_{1t}p_{2t}[(p_{2t} - p_{1t})(\mu_{1t}^s - \mu_{2t}^s)^2(\mu_{1t}^b - \mu_{2t}^b) \\ &\qquad + (\mu_{1t}^b - \mu_{2t}^b)(h_1^s - h_2^s) + 2(\mu_{1t}^s - \mu_{2t}^s)(\mathrm{cov}_1 - \mathrm{cov}_2)]. \end{aligned}$$

A few remarks are in order. First, Equations (14) and (15) clearly show that regimes play a key role in generating time-varying (co)skewness of stock and bond returns. Second, Equations (14) and (15) also show that conditional coskewness is proportional to the conditional covariance between one asset's return and another asset's volatility. More specifically, conditional stock coskewness captures the conditional relation between stock return and bond volatility, whereas conditional bond coskewness captures the conditional relation between bond return and stock volatility. Also, the two coskewness measures are not necessarily zeroes if the covariance in Equation (13) is zero. This suggests that coskewness captures certain extreme comovement that correlation does not, because coskewness is about the comovement in the long tail.

Finally, a high (low) cross-market coskewness indicates that an investor receives a high (low) return from one asset when the other is more volatile. Hence, the concept of the cross-market coskewness is naturally appealing in many other contexts, including measurements of hedging effectiveness and diversification potentials across various markets.

4. Empirical Results

4.1. Results on Regime-Switching Model Estimation

Estimating the model requires fully specifying the conditional means, conditional variance–covariance matrices, and transition probabilities. Because of the difficulty in estimating the model with the many parameters involved, we consider the following parsimonious specifications in line with Ang and Bekaert (2002). Specifically, the conditional means are specified as

$$\boldsymbol{\mu}_{it} = \boldsymbol{\mu}_i + \boldsymbol{\lambda}_i SHORT_{t-1}, \quad i \in \{1, 2\}, \tag{16}$$

where $\boldsymbol{\mu}_i = (\mu_i^s, \mu_i^b)'$ is a 2×1 vector of the constant means given regime i, $SHORT_{t-1}$ is the first lagged short rate, and $\boldsymbol{\lambda}_i = (\lambda_i^s, \lambda_i^b)'$ is a 2×1 vector of regression coefficients given regime i. Short rates are closely attuned to discount rates and have significant predictive power for future asset returns. Such predictive power of nominal short rates has a long tradition in finance, and Ang et al. (2006) showed recently that the short rate has more predictive power than any term spread in forecasting GDP out of sample. Of course, other predictor variables, such as dividend yield, could also be studied. In general, the conditional means may not just linearly depend on the first lag of the instrument. The specification adopted here thus represents a trade-off between flexibility and parsimony.

For the conditional variance–covariance matrices, we assume that $\boldsymbol{\varepsilon}_{it}$ follows an independent and identically distributed (i.i.d.) bivariate normal distribution. Then the conditional distribution of $\mathbf{r}_t$ is a mixture of two i.i.d. bivariate normal distributions as follows:

$$\mathbf{r}_t \mid \mathbf{F}_{t-1} \sim \begin{cases} \mathcal{N}(\boldsymbol{\mu}_{1t}, \mathbf{H}_{1t}), \text{ w.p. } p_{1t}; \\ \mathcal{N}(\boldsymbol{\mu}_{2t}, \mathbf{H}_{2t}), \text{ w.p. } p_{2t}. \end{cases} \tag{17}$$

Because mixtures of normals can approximate a very broad set of density families, this assumption is not very restrictive. Moreover, the variances and correlation are assumed to be constant within each regime, and conditional heteroskedasticity can be generated by switches between regimes. There is evidence that further inclusion of time-varying conditional volatilities and correlation in each regime is statistically insignificant (Guidolin and Timmermann 2005). This insignificance motivates the following parsimonious specification for the conditional variance–covariance matrices as follows:

$$\mathbf{H}_{it} = \mathbf{D}_{it}\mathbf{R}_{it}\mathbf{D}_{it}, \quad \mathbf{D}_{it} = \begin{pmatrix} \sqrt{h_i^s} & 0 \\ 0 & \sqrt{h_i^b} \end{pmatrix}, \quad \mathbf{R}_{it} = \begin{pmatrix} 1 & 0 \\ \rho_i & 1 \end{pmatrix}, \quad i \in \{1, 2\}, \tag{18}$$

where h_i^s and h_i^b are the constant conditional volatilities of stocks and bonds given regime i; and ρ_i is the constant conditional stock–bond correlation given regime i. Nevertheless, we also address below the possibility that the estimated correlations between stock and bond excess returns may vary across two regimes.

Furthermore, motivated by empirical evidence that short rates help predict transitions in the regime, we specify the transition probabilities to be a function of the lagged instrument:

$$\begin{aligned} p_{ii,t} &= p(S_t = i \mid S_{t-1} = i, \mathbf{F}_{t-1}) \\ &= \Phi(a_i + b_i SHORT_{t-1}), \quad i \in \{1, 2\}, \end{aligned} \tag{19}$$

where a_i and b_i are unknown parameters, and Φ is the cumulative normal distribution function, which ensures that $0 < p_{ii,t} < 1$. This specification makes transition probabilities monotonic in the instrument, thus facilitating the interpretations of the parameters (Connolly et al. (2005) also estimated a time-varying transition probability model using market volatility as the instrument). For comparison, we also consider a single-regime model where there are no transition probabilities and other parameters are not allowed to vary across regimes.

The analysis will proceed as follows. First, we estimate the single-regime model as a benchmark (Model III). Next, we introduce regimes and investigate Model II assuming constant conditional correlation. Having ensured the presence of regimes, we further examine time-varying correlations in Model I. The estimation results of various regime-switching models for the U.S. data (1951–2004) are reported in Table 2.[9] Based on estimated likelihood functions and the resulting likelihood ratio (LR) tests, the regime-switching models fit much better than the single-regime model. Thus, we focus attention on the regime-switching models, Models I and II. Following much of the literature (e.g., Gray 1996; Ang and Bekaert 2002; Connolly et al. 2005, 2007; Guidolin and Timmermann 2008), we also only focus on the two-state regime-switching model, which has an intuitively appealing interpretation.[10]

First, examining the parameters for the variances of the shocks to the stock and bond markets, it is clear

[9] To our best knowledge, none of the studies estimating regime-switching models calculate and report the fitness statistic of R^2, possibly because the predicted values or predicted residuals are not defined in the usual sense. To compute it, we first generate predicted values of the dependent variable as sum of the predicted values in two regimes weighted by the associated regime probabilities. Given this predicted values, the estimated residuals are obtained as the difference between real and predicted values. The R^2 values of the single-regime Model III are 1.01% and 0.00% for stock and bond premiums, respectively. The corresponding statistics of the two-regime Model I increase to 1.63% and 1.38%.

[10] Guidolin and Timmermann (2005) report that a three-state regime-switching model fits well the monthly UK stock and bond data during 1976–2000. We also tried such a model and encountered difficulty in estimation convergence for our data set. It is certainly possible that the two-state regime model may only capture most but not all of salient aspects of the stock–bond comovement. Nevertheless, as shown below, the model seems to be reasonably adequate, because the conditional means and standard deviations of stock and bond premiums based on the regime-switching model are, on average, very close to those of historical premiums.

Table 2 Estimation Results of the Regime-Switching Models for the United States

	Model I		Model II		
	Regime 1	Regime 2	Regime 1	Regime 2	Model III
μ_i^s	0.248***	0.114***	0.249***	0.111**	0.149***
t-statistic	(6.978)	(3.208)	(4.003)	(2.513)	(5.653)
μ_i^b	0.016	0.048**	0.019	0.048**	0.007
t-statistic	(1.122)	(2.083)	(0.725)	(2.474)	(0.408)
λ_i^s	−4.677***	−1.246**	−4.851**	−1.178	−1.847***
t-statistic	(−4.472)	(−2.237)	(−2.493)	(−1.604)	(−4.138)
λ_i^b	−0.980**	−0.265	−1.102	−0.241	0.085
t-statistic	(−2.076)	(−0.788)	(−1.189)	(−0.767)	(0.304)
h_i^s	0.154***	0.329***	0.155***	0.327***	0.259***
t-statistic	(11.483)	(14.633)	(11.099)	(14.382)	(27.072)
h_i^b	0.010***	0.114***	0.010***	0.114***	0.072***
t-statistic	(13.935)	(15.796)	(6.758)	(13.194)	(23.584)
ρ_i	0.084	0.207***	0.163***		0.186***
t-statistic	(1.094)	(4.614)	(4.602)		(5.051)
a_i	4.068***	−1.555***	4.039***	−1.574***	
t-statistic	(7.142)	(−4.972)	(5.254)	(−3.815)	
b_i	−55.855***	−1.984	−56.227***	−0.528	
t-statistic	(−4.385)	(−0.276)	(−3.297)	(−0.048)	
Likelihood	793.454		792.451		644.666
LR	2.006		295.571		

Notes. This table estimates three nested models for stock and bond excess returns using U.S. monthly data from April 1951 to August 2004. Model I is unrestricted:

$$\begin{pmatrix} r_t^s \\ r_t^b \end{pmatrix} = \begin{pmatrix} \mu_i^s \\ \mu_i^b \end{pmatrix} + \begin{pmatrix} \lambda_i^s \\ \lambda_i^b \end{pmatrix} SHORT_{t-1} + \begin{pmatrix} \varepsilon_{it}^s \\ \varepsilon_{it}^b \end{pmatrix}, \quad \text{where } \begin{pmatrix} \varepsilon_{it}^s \\ \varepsilon_{it}^b \end{pmatrix} \sim \mathcal{N}(\mathbf{0}, \mathbf{H}_{it}),$$

$$\mathbf{H}_{it} = \mathbf{D}_{it}\mathbf{R}_{it}\mathbf{D}_{it}, \quad \mathbf{D}_{it} = \begin{pmatrix} \sqrt{h_i^s} & 0 \\ 0 & \sqrt{h_i^b} \end{pmatrix}, \quad \mathbf{R}_{it} = \begin{pmatrix} 1 & 0 \\ \rho_i & 1 \end{pmatrix}, \quad i \in \{1, 2\},$$

and transition probabilities

$$p_{ii,t} = p(s_t = i \mid s_{t-1} = i, \mathbf{F}_{t-1}) = \Phi(a_i + b_i SHORT_{t-1}), \quad i \in \{1, 2\},$$

where $SHORT_{t-1}$ is the first lagged short rate, s_t is the unobserved regime at time t, $\mathbf{F}_{t-1}$ is the past information set, and Φ is the cumulative normal distribution function. Model II is a restricted model with the constraint $\rho_1 = \rho_2$. Model III is the single-regime model with the following constraints: $\mu_1 = \mu_2$, $\lambda_1 = \lambda_2$, $h_1^s = h_2^s$, $h_1^b = h_2^b$, and $\rho_1 = \rho_2$. The parameter estimates are the QMLE. The t-statistics are reported in parentheses. The log-likelihood values are also reported. The likelihood ratio test is a test of the unrestricted model against the restricted model.

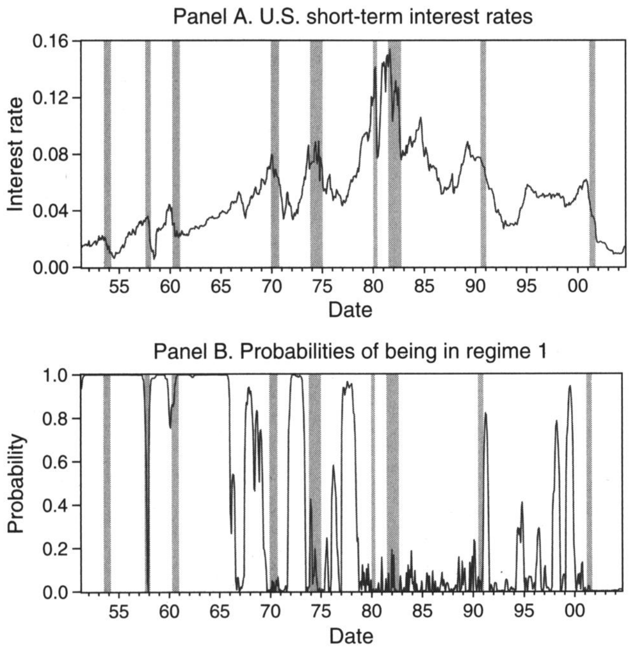
**5% and ***1% significance.

that regime 1 is a low volatility regime and regime 2 is a high-volatility regime in which diversification might be particularly useful. The stock market under regime 2 is about twice as volatile as under regime 1 (conditional on short rates). The contrast is much more striking for the bond market. These results are true in both Models I and II. Furthermore, the estimate for the correlation between shocks to stock and bond markets under regime 2 is also about twice as large as under regime 1 in Model I. Nevertheless, the estimate is not precisely estimated in regime 1. The LR test of Model I against Model II also suggests that stock–bond correlation is not significantly different across regimes (the associated p-value is 0.16). The finding is consistent with Longin and Solnik (2001), who also report that asymmetric correlation may not depend on the level of market volatility. If this is the case, no better diversification opportunities are available in the regime of high volatility based on correlations alone.[11]

Next, turn to the parameters for the conditional mean. Consistent with the existing literature, low short rates tend to predict high stock returns. Moreover, this negative association is more significant in

[11] Nevertheless, likelihood ratio tests are known to have low power in regime-switching models. Furthermore, using the Chicago Board Option Exchange's Volatility Index from 1986 to 2000, Connolly et al. (2005) find that correlation between stock and bond daily returns is not constant and is related to lagged market uncertainty (although the relation is negative). Parameter estimates from Model I are similar to those from Model II. Therefore, from this point on we will concentrate on the more flexible Model I.

Figure 1 Smoothed Regime Probabilities



Notes. These plots show U.S. short-term interest rates (Panel A) (April–July 2004) and the smoothed probability of being in regime 1 (Panel B) (May 1951–August 2004). They are based on parameter estimates from Model I. Shaded areas are NBER-dated economic recessions.

the low-volatility regime 1. The pattern is similar for bond returns. However, the relationship is insignificantly different from zero in regime 2. Last, we take a look at the parameters about the transition probabilities. The coefficient b_i measures the dependence of the probability of staying in regime i on the interest rate. Both b_1 and b_2 are negative, but only b_2 is significant. This implies that as the interest rate declines the probability of staying in the high-volatility regime increases, which may characterize periods of economic recessions with lower interest rates and high asset price volatility.

Figure 1 plots the smoothed probabilities of being in regime 1 along with short-term interest rates (shaded areas indicate NBER-dated economic recessions). All observations can be categorized as belonging to a particular regime if the associated smoothed probability is larger than 0.80 based on parameter estimates of Model I (but there are a total of 75 observations that fall outside the two categories). It is clear that low or declining short rates are more likely to be associated with regime 2 (the correlation between the probability of being in regime 1 and the one-period lagged short-term rate is −0.52). Not surprisingly, these are also the periods of economic recession. The noticeable exception is that the July 1953–May 1954 recession is estimated to be in regime 1. Figure 1 also provides some intuitive evidence of possible unstable correlation between stock and bond markets. Earlier empirical studies (e.g., Ilmanen 2003) found that the relation has tended to be positive but has dipped below zero in two episodes: 1956–1965 and 1998–2001. Panel B of Figure 1 shows that the first period is mostly falling in regime 1, which is associated with a lower correlation between the two assets. However, in the four-year period of 1998–2001, our model identifies 31 months to be in high-volatility regime 2. Only five (June–October 1999) are in regime 1 (we will discuss more about this at the end of §4.3).

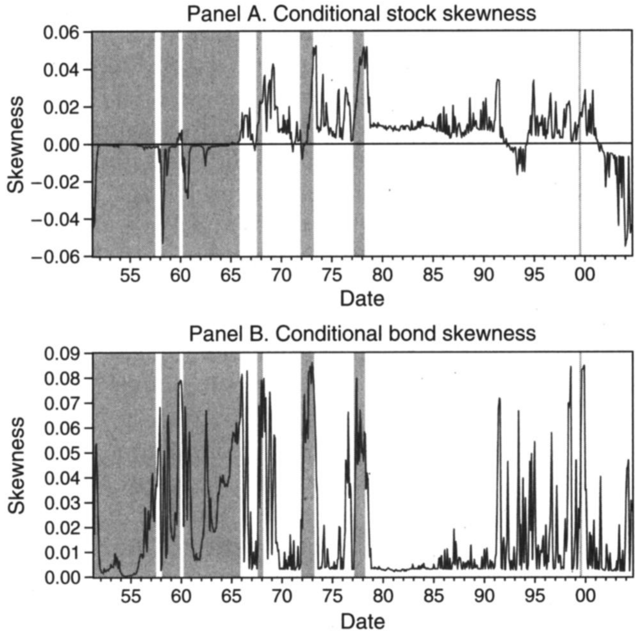
On average, from 1951 to 2004, the conditional expected U.S. stock premium is 5.6% per year, whereas the expected U.S. bond premium is only 1.1% per year. They are very close to means of realized stock and bond risk premiums as shown in Table 1. The estimated conditional correlation has a mean of 0.17, close to the unconditional mean of 0.18. In particular, conditional correlation is quite persistent with an autocorrelation larger than 0.70 after one year, matching an important feature of the data. Together, these results imply the adequacy of the two-state regime-switching model in modeling the data. Furthermore, conditionally or unconditionally, the stock premium is higher in regime 1 than in regime 2, whereas the reverse is true for the bond premium. The measures of skewess and coskewness also differ across two regimes for both stocks and bonds. For example, unconditional coskewness of bond returns is 0.034 and 0.183 in regimes 1 and 2, respectively. This means that when stock volatility is high (regime 2), bond tends to earn higher returns than when stock volatility is low, which is consistent with the finding reported by Connolly et al. (2005). Nevertheless, the differences of conditional (co)skewness measures between regimes 1 and 2 are generally less striking than among unconditional measures, to which we now turn our attention in the next subsection.

4.2. Results on Conditional Skewness and Coskewness

As indicated in Table 1, U.S. stock premium has negative unconditional skewness, whereas U.S. bond premium has positive unconditional skewness. Figure 2 plots the conditional skewnesses derived from Equation (11), with regime 1 indicated by shaded areas.[12] The conditional bond skewness remains to be mostly

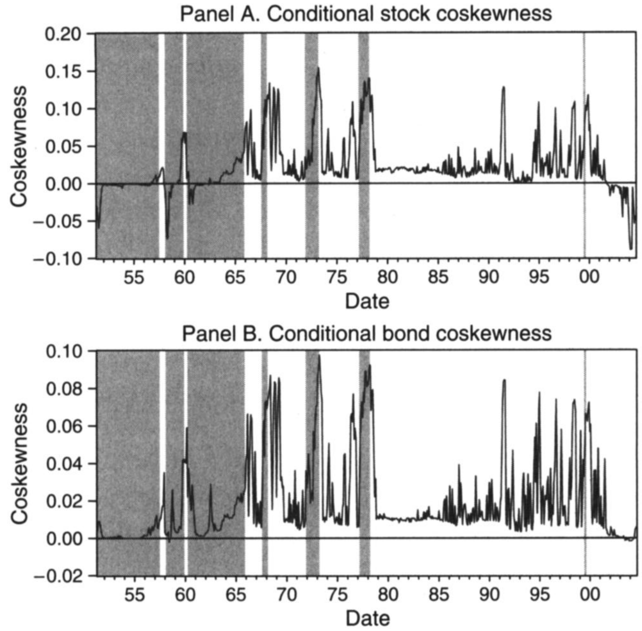
[12] A change in the skewness estimate of −0.095 tends to be large for the U.S. stock market based on the 150-year sample. Based on the null of a normal distribution or the empirical distribution from the bootstrapping, the skewness estimate of −0.095 is also statistically significant different from zero at the 5% level based on one-sided tests. The periods with the skewness estimates more negative than this number correspond to the Great Depression and various recessions during the 150-year period. Understandably, such a threshold level defining a large change in skewness might be sample specific, and more generally, large (or small) skewness might be most meaningfully defined relative to certain context-specific benchmarks.

Figure 2 U.S. Conditional Standardized Skewness



Notes. These plots show the monthly conditional standardized skewnesses for U.S. stock and bond premiums from May 1951 to August 2004. The time series are derived from the results of bivariate regime-switching Model I in Table 2 and Equation (11). An observation is categorized as belonging to regime 1 if the associated smoothed probability is larger than 0.80 based on parameter estimates from Model I. Shaded areas indicate regime 1.

Figure 3 U.S. Conditional Standardized Coskewness



Notes. These plots show the monthly conditional coskewnesses for U.S. stock and bond premiums from May 1951 to August 2004. The time series are derived from the results of bivariate regime-switching Model I in Table 2 and Equations (14) and (15). An observation is categorized as belonging to regime 1 if the associated smoothed probability is larger than 0.80 based on parameter estimates from Model I. Shaded areas indicate regime 1.

positive, whereas the evidence for conditional stock skewness is mixed. There are three episodes when conditional stock skewness was mostly negative or close to zero (also in regime 2 in general): from the beginning of the sample until the mid-1960s, May 1992–April 1994, and June 2001 through the end of the sample. These results largely confirm the eprevious literature (e.g., Prakash et al. 2003, Guidolin and Timmermann 2008) in a more robust manner that the U.S. stock market may exhibit negative skewness over an extended period of time. However, they present an important new finding that the U.S. bond market typically exhibits positive skewness with an average of 0.020 (higher than that of the stock market's 0.006). The results would be even much stronger based on the entire 150-year sample (to be discussed below). With the exception of the period of 1979–1984, both conditional stock skewness and bond skewness show rather dramatic time variations, which highlights the danger for investors of drawing an inference based on the unconditional skewness, particularly from an insufficiently long sample period.

Figure 3 plots the conditional stock and bond coskewness series derived from Equations (14) and (15). Note that Harvey and Siddique (2000) formulated coskewness as the relation between individual stock return and stock market volatility. The coskewness measure in their study comes from a hidden assumption: that stocks coskew in a same direction because they belong to the same asset class. This assumption should be relaxed for a stock–bond portfolio because stocks and bonds are two different asset classes.

Figure 3 reveals another new finding that U.S. conditional bond cosdkewness is, in general, positive during the sample period of 1951–2004, whereas the stock market shows both positive and negative coskewness. As measures of skewness in Figure 2, both stock and bond coskewnesses were small and stable during the period of 1979 and 1984. In contrast, they were both large and volatile from 1966 to 1978 and from 1993 to 1999. Nevertheless, in the last four years of the sample period, bond coskewness was essentially zero, and stock coskewness turned negative. That means that if investors add bonds into their stock portfolio, it may reduce the negative skewness of the portfolio return, even after controlling for the impact on the portfolio variance. Hence, investors may bid up the bond price and accept a lower risk premium for the skewness. In contrast, adding stocks will decrease a bond portfolio's skewness, and thus investors command a higher expected risk premium. To end this subsection, we note from Figures 2 and 3 that both conditional skewness and coskewness may change significantly over time. This is because they are cubic functions of conditional mean returns and conditional state probabilities, both of which depend on time-varying short rates.

4.3. Results on Coskewness Pricing Effects

An important question is whether stock and bond coskewnesses earn ex ante risk premiums.[13] Therefore, we estimate the following time-series regressions of expected excess returns:

$$\hat{\mu}_t^j = \alpha^j + \boldsymbol{\beta}^j \hat{\mathbf{f}}_t^j + e_t^j, \qquad (20)$$

where $\hat{\mu}_t^j$, $j = s, b$ is the estimate of conditional risk premium for a stock or bond in Equation (9), $\hat{\mathbf{f}}_t^j$, $j = s, b$ is a (sub)set of the risk factors for stock or bond returns, and $\boldsymbol{\beta}^j$ is the corresponding vector of factor loadings to be estimated. The candidate factors include expected coskewness, correlation, volatility, and skewness estimated from the above two-state regime-switching model.

In general, if the factors are important for explaining risk premiums, the estimated betas should be statistically significant. In particular, our goal is to investigate how aggregate coskewnesses are priced across stock and bond markets. If investors display skewness preferences and returns are not normally distributed, the slope coefficient on standardized coskewness should be significantly negative. To see this, we first reduce the number of factors in Equation (20) to just the expected coskewness estimate, $\hat{\gamma}_t^j$. However, if investors display mean-variance preferences or returns follow a normal distribution, the systematic risk can be measured as the correlation (or standardized covariance) alone. If this were the case, the addition of $\hat{\rho}_t$ in the second specification could make coskewness insignificant. The third specification further includes the expected standard deviation, $\widehat{sd}_t^j$, to control for potential pricing effect due to volatility. Finally, the fourth specification also controls for potential pricing effect due to skewness. Both Mitton and Vorkink (2007) and Barberis and Huang (2008) argue that a positively skewed asset should earn negative average excess returns.

The time-series regression is estimated individually for stocks or bonds in the sample. Each regression produces an estimate of the risk exposure vector, $\boldsymbol{\beta}^j$, $j = s, b$ for stocks or bonds. For each regression, we perform a χ^2 test for the (joint) significance of the beta coefficient(s). As the regressions are conducted using estimates from the regime-switching model, the variables are measured with noises. To deal with the errors-in-variables problem, coefficient estimates are adjusted for a serial correlation of 12 lags and heteroskedasticity following Newey and West (1987).[14]

The estimation results are reported in Table 3. In Panel A for U.S. stocks and Panel B for the U.S. bonds, the slope estimates of stock or bond coskewness are negative in all specifications. They are also statistically significant at lower than the 1% level except the coefficient of bond skewness in the third specification.[15] The relation of conditional coskewnesses and expected excess returns is also economically significant. For example, in Specification 2, which also controls for pricing effects due to correlation, the regression coefficients on the U.S. stock and bond coskewnesses are about −0.731 and −0.423, respectively. Thus, during the period of October 2001 through August 2004, the negative stock coskewness (−0.022 on average) induces a 1.6% increase in the expected stock premium, which is about one-third of the average unconditional stock premium (4.7%). For the U.S. bond market, the bond coskewness, which is 0.020 over the 1951–2004 sample period, leads to a 0.8% decrease in the expected bond premium (recall from Table 1 that the average of unconditional bond premium is 1.1%). Hence, the pricing of both stock and bond coskewnesses is also economically significant.

Finally, as shown in Specification 4 of Table 3, after allowing for the pricing effect of coskewness, the skewness on the bond market is also priced with the expected negative sign. The result is in line with both Mitton and Vorkink (2007) and Barberis and Huang (2008). The other parameter of interest is that of volatility. It is priced in both stock and bond markets with expected negative signs. Table 3 also shows that conditional stock–bond correlation is also priced

[13] We use conditional risk premiums because, as discussed earlier, they contain less noise and are a better proxy for expected stock and bond excess returns than raw premiums. Nevertheless, we also estimate Equation (20) using raw premiums as dependent variables and using conditional moments derived from Equations (10), (11), and (13)–(15) without conditioning on lagged short-term rates. Both stock and bond coskewnesses have expected negative signs, but the coefficients are less precisely estimated. Based on the subsample of 1951–2004, the coefficients for bond and stock coskewnesses are −2.534 and −1.026, respectively, with p-values of 0.07 and 0.47. If the entire 150-year sample is used, the coefficients are −2.366 and −1.226 (the associated p-values are 0.00 and 0.11). If the raw premiums are used as dependent variables and we also allow transition probabilities vary with short-term rates, the coefficients of factor loadings are more precisely estimated, particularly for bond coskewness. We thank an anonymous referee for motivating us to study the robustness of the results using raw premiums.

[14] To further account for the errors-in-variables problem due to the use of generated regressors, we compute standard errors of coefficient estimates using sampling-with-replacement bootstrapping with 1,000 repetitions. The bootstrapping results are qualitatively the same as those reported here based on Newey and West's (1987) robust standard errors. There are other important test procedures to deal with the errors-in-variables problem such as Shanken (1985) and studies that follow it.

[15] There is some evidence of collinearity among the control variables, which may explain the insignificance of the coskewness in the third specification for bond returns (Panel B) and the skewness in the fourth specification for stock returns (Panel A). Both estimates become highly significant in Table 4 using the full sample. This is what we would expect when the sample size increases and if the underlying factors are correlated.

Table 3 Conditional Coskewness and Ex Ante Premiums in the United States

Panel A: *Conditional stock premium* as the dependent variable						
Model	α^s	β_1^s	β_2^s	β_3^s	β_4^s	p-value of χ^2 test
1	0.076***	−0.888***				0.000
	(0.000)	(0.000)				
2	0.207***	−0.731***	−0.796***			0.000
	(0.000)	(0.000)	(0.000)			
3	0.554***	−1.174***	1.175***	−1.334**		0.000
	(0.000)	(0.000)	(0.000)	(0.017)		
4	0.559***	−1.192***	1.200*	−1.353***	0.037	0.000
	(0.000)	(0.000)	(0.064)	(0.003)	(0.944)	
Panel B: *Conditional bond premium* as the dependent variable						
Model	α^b	β_1^b	β_2^b	β_3^b	β_4^b	p-value of χ^2 test
1	0.018***	−0.347***				0.000
	(0.000)	(0.000)				
2	−0.031***	−0.423***	0.300***			0.000
	(0.000)	(0.000)	(0.000)			
3	−0.015	−0.287	−0.182***	0.251		0.000
	(0.620)	(0.298)	(0.844)	(0.598)		
4	0.196***	−0.756***	1.459***	−0.819***	−0.233***	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	

Notes. This table presents results of the following regressions for the U.S. stock and bond premiums using monthly data from April 1951 to August 2004:

$$\hat{\mu}_t^j = \alpha^j + \beta_1^j \hat{\gamma}_t^j + \beta_2^j \hat{\rho}_t + \beta_3^j \widehat{sd}_t^j + \beta_4^j \widehat{sk}_t^j + e_t^j,$$

where $\hat{\mu}_t^j$, $j = s, b$ is the estimated conditional risk premium for stock or bond, respectively; $\hat{\rho}_t$ is the estimated conditional stock–bond correlation; $\hat{\gamma}_t^j$, $\widehat{sd}_t^j$, and $\widehat{sk}_t^j$ $j = s, b$ are the estimated conditional stock or bond coskewness, volatility, and skewness, respectively. The p-values reported beneath each coefficient estimate are adjusted for a serial correlation of 12 lags and heteroskedasticity.

*10%, **5%, and ***1% significance; p-values of χ^2 tests for the (joint) significance of beta coefficient(s) are also reported.

in the fully specified Model 4. Nevertheless, because of possible collinearity among the control variables, coefficient estimates in Specification 4 should be read with caution when compared to relatively simple Specifications 1 and 2.

So far we have used short-term interest rates as our conditioning/state variable in estimating regime-switching models and deriving conditional moments for studying pricing effects of coskewness in stock and bond markets. To study the robustness of our results summarized in Tables 2 and 4, we also use inflation rates as the conditioning variable in place of short-term rates (results are available on request). The motivation of using this variable is that although stock–bond correlation is apparently different across the two regimes (0.084 versus 0.207) in the benchmark Model I, the standard LR test fails to reject the model against the constant correlation Model II (Table 2). According to Campbell and Ammer (1993), the only factor that might induce the unstable correlation is a differential response to inflation expectations. The pricing effects of stock and bond coskewnesses using inflation rates as the instrument are similar to those reported in Table 3 using short rates as the instrument. The relevant coefficients are −1.471 and −0.542, respectively, both of which are statistically significant at less than the 1% level. The estimates of the correlation in two regimes are also similar (0.073 versus 0.204). Nevertheless, again, the correlation in regime 1 is not precisely estimated with a p-value of 0.27 only. Similar to Connolly et al. (2005), this result suggests that factors other than inflation rates may be important in understanding the observed unstable bond–stock correlation.

The notions of cross-market hedging and flight to quality have also been often used (e.g., Ilmanen 2003) to explain the unstable stock–bond correlation (which may also help answer why the coskewness is a priced factor). It has been argued that the observed decoupling between the stock and bond returns (for example, in the late 1990s) may be related to the high stock volatility and uncertain economic times (e.g., Connolly et al. 2005). We estimated a modified version of Model I using one-period lagged realized stock volatility as the conditioning variable for time-varying transition probability (the monthly realized volatility is constructed as the sum of squared daily returns normalized to 22 trading days). The pricing effect of stock coskewness remains largely the same (−1.361), whereas the effect of bond coskewness is smaller (−0.283). The smoothed probabilities (available on request) are also similar to the pattern in Figure 1 with a few noticeable exceptions. Not surprisingly, the new model assigns higher probabilities to high-volatility regime 2 following the October 1987 crash and in most of the end of sample period when the market volatility is high. Nevertheless, the identified regimes are still similar to those in Table 2 using short rates as the instrument. Regime 1, on average, still has a lower volatility and stock–bond correlation than regime 2, meaning that the regime-switching model does not explain well one important feature of the late 1990s data featuring high volatility and negative stock–bond correlation. Similar failure to model the negative correlation is also reported in two recent studies by Baele et al. (2010) and David and Veronesi (2008). One possible reason is that further structural break(s) might have occurred during this period.

Furthermore, motivated by the finding of Piazzesi (2005) about the U.S. monetary policy impact on the bond market dynamics, we also use the effective federal funds rate as the conditioning variable for time-varying transition probability. Both the coskewness measures and their pricing effects are similar to those based on the short rate. The result perhaps is not surprising, given the fact that the federal funds rate is correlated with the short rate with a coefficient of 0.98 during the sample period. Finally, following Ludvigson and Ng (2009), we use as the conditioning variable the first component of four most popular state variables in asset pricing (the short rate, dividend yield, term spread, and default spread) (see, e.g., Bali et al. 2009) as well as the first component of 132 macroeconomic variables as used in Ludvigson and Ng (2009). In both cases, the robustness of our main results is confirmed, and the pricing effects of both stock and bond coskewnesses are even somewhat larger in magnitude in the latter case. Nevertheless, they are less precisely estimated, perhaps because the principal component itself is also estimated.

4.4. Further Robustness Checks

In this subsection, we provide further evidence on the robustness of cross-market coskewness effects. We first extend our above analysis to the entire sample period of 1855–2004. There is a trade-off in utilizing such a long sample. Substantial changes in capital market structure over the sample period may complicate modeling efforts and confound parameter estimates. However, with a long sample over one and a half centuries, we can test whether our findings are a result of data mining or reflect something fundamental in the market that is robust to structural changes in financial institutions or macroeconomic environment. Various regime-switching models using the whole sample are estimated but not reported in details to conserve the space. The result (available on request) provides even stronger evidence that during past 150 years the conditional stock skewness is mostly negative, whereas the conditional bond skewness is mostly positive. Furthermore, the U.S. conditional stock skewness was most negative during the period of 1930–1960, whereas it was least negative during the 1980s (except a small peak around the 1987 crash). By contrast, the U.S. conditional bond skewness was much higher before the 1930s than after, and hit almost the lowest level during the 1980s (again with the exception of a peak around the 1987 crash). Further relevant to investors, the conditional stock skewness became more negative, whereas the conditional bond skewness has become more positive, since the 1990s. Confirming the new finding that the bond coskewness is in general higher than the stock coskewness during past 150 years, the result also reveals that the U.S. conditional stock coskewness was also most negative during the period of 1930–1960, and least negative during the 1980s. Interestingly, the U.S. conditional bond coskewness was also negative during most of the same period, which differs somewhat from the pattern in the conditional bond skewness. Nevertheless, since the 1960s, the bond coskewness is mostly positive whereas the stock coskewness is mostly negative.

Table 4 summarizes the most relevant results of coskewness and other conditional moments pricing effects using Equation (20). Obviously, as the sample size increases, all parameters are more precisely estimated in Table 4. More importantly, our main results in Table 3 based on the recent subsample of 1951–2004 still hold. The estimated betas for stock and bond coskewnesses have expected negative signs in all four model specifications. In Specification 4, which controls for pricing effects due to correlation, volatility, and skewness, the coefficients on the stock and bond coskewnesses are −0.92 and −0.66, respectively. Both are reasonably close to the subsample estimates of −1.19 and −0.76. Coefficients of other factors generally have similar signs as their subsample counterparts but may differ in magnitude. Furthermore, in Tables 3 and 4 (and to some extent Table 5 as well), the slope magnitudes of the coskewness effect on risk premiums are typically greater than or at least comparable to those of the volatility effect, especially in the longer 150-year sample. It suggests that unit price of coskewness is greater in magnitude than unit price of volatility, where coskewness can be interpreted to be due to prudence, and volatility may be a proxy for risk. This is roughly consistent with Kimball's (1990)

Table 4 Conditional Coskewness and Ex Ante Premiums in the United States (Full Sample)

Panel A: *Conditional stock premium* as the dependent variable						
Model	α^s	β_1^s	β_2^s	β_3^s	β_4^s	p-value of χ^2 test
1	−0.009***	−0.112***				0.000
	(0.000)	(0.000)				
2	−0.122***	−1.145***	0.803***			0.000
	(0.000)	(0.000)	(0.000)			
3	−0.095***	−1.078***	0.825***	−0.046***		0.000
	(0.000)	(0.000)	(0.000)	(0.000)		
4	−0.272***	−0.924***	2.175***	−0.072***	0.431***	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	
Panel B: *Conditional bond premium* as the dependent variable						
Model	α^b	β_1^b	β_2^b	β_3^b	β_4^b	p-value of χ^2 test
1	0.011***	−1.085***				0.000
	(0.000)	(0.000)				
2	0.118***	−1.140***	−0.764***			0.000
	(0.000)	(0.000)	(0.000)			
3	0.125***	−1.082***	−0.850***	0.026***		0.000
	(0.000)	(0.000)	(0.000)	(0.000)		
4	0.126***	−0.662***	−0.474***	−0.062***	−0.664***	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	

Notes. This table presents results of the following regressions for the UK stocks and bonds using monthly data from January 1855 to August 2004:

$$\hat{\mu}_t^j = \alpha^j + \beta_1^j \hat{\gamma}_t^j + \beta_2^j \hat{\rho}_t + \beta_3^j \widehat{sd}_t^j + \beta_4^j \widehat{sk}_t^j + e_t^j,$$

where $\hat{\mu}_t^j$, $j = s, b$ is the estimated conditional risk premium for stock or bond, respectively; $\hat{\rho}_t$ is the estimated conditional stock–bond correlation; and $\hat{\gamma}_t^j$, $\widehat{sd}_t^j$, and $\widehat{sk}_t^j$ $j = s, b$ are the estimated conditional stock or bond coskewness, volatility, and skewness, respectively. The p-values reported beneath each coefficient estimate are adjusted for a serial correlation of 12 lags and heteroskedasticity.

***1% significance; p-values of χ^2 tests for the (joint) significance of beta coefficient(s) are also reported.

argument: the precautionary saving motive may be stronger than risk aversion under the skewness preference. Furthermore, during the 150-year U.S. sample period, the average estimate of stock coskewness is −0.047, which may induce a 4.3% increase in the average stock premium and as big as the average unconditional stock premium during the sample period (4.3%). For the U.S. bond market, the bond coskewness (with the mean of 0.007) on average leads to a 0.46% decrease in the expected bond premium, which is again comparable to the average of unconditional bond premium during the sample period (0.4%). Hence, again in line with Kimball (1990), economic significance of the pricing effects of U.S. stock and bond coskewnesses is indeed striking from a very long-term perspective. Interestingly, Bali et al. (2009, p. 908) also make a similar argument in a different context that rare large moves in the stock market can be interpreted as signals, whereas the frequent small fluctuations can be viewed as noise, which does not have the power to explain time-series variation in excess stock market returns.

A natural question is whether coskewnesses are priced only in the U.S. stock and bond markets, but not in other countries. Thus, we extend a similar analysis to the UK data for the sample period of 1855 through 2005. There is strong evidence of more than one regime using either the whole sample or the postwar subsample, which is generally in line with the findings on the U.S. market. For example, in the postwar sample, conditional variances of stock and bond returns are estimated to be 0.088 and 0.058, respectively, in regime 1. They are 0.741 and 0.201 in regime 2, respectively. The stock–bond correlation is 0.164 in regime 1 and 0.271 in regime 2. Both estimates are significant at the 1% level.

The estimation results for the time-series regressions of expected premium on expected coskewness are further reported in Table 5. Note that only fully specified model estimation results are reported here (with more detailed results available on request). All the slope estimates of the UK stock or bond coskewness are significantly negative at the lower than 1% level, meaning that the coskewness is also priced across the UK stock and bond markets, confirming the findings from the U.S. data. Nevertheless, the magnitudes are relatively small especially in the full sample, compared to the U.S. counterpart. Based on Specification 4, the full sample coefficients are −0.078 and

Table 5 Conditional Coskewness and Ex Ante Premiums in the United Kingdom

Panel A: *Conditional stock premium* as the dependent variable						
Period	α^s	β_1^s	β_2^s	β_3^s	β_4^s	p-value of χ^2 test
January 1855–November 2005	0.038**	−0.078***	−0.145***	−0.046***	−0.238***	0.000
	(0.030)	(0.000)	(0.000)	(0.000)	(0.000)	
October 1945–November 2005	0.164**	−0.495***	−1.132***	−0.109***	−0.154***	0.000
	(0.030)	(0.000)	(0.000)	(0.000)	(0.000)	
Panel B: *Conditional stock premium* as the dependent variable						
Period	α^s	β_1^b	β_2^b	β_3^b	β_4^b	p-value of χ^2 test
January 1855–November 2005	0.105**	−0.251***	−0.355***	−0.059***	−0.253***	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.027)	
October 1945–November 2005	−0.003***	−0.425***	0.508***	0.171***	0.176**	0.000
	(0.790)	(0.000)	(0.000)	(0.000)	(0.000)	

Notes. This table presents results of the following regressions for the UK stocks and bonds:

$$\hat{\mu}_t^j = \alpha^j + \beta_1^j \hat{\gamma}_t^j + \beta_2^j \hat{\rho}_t + \beta_3^j \widehat{sd}_t^j + \beta_4^j \widehat{sk}_t^j + e_t^j,$$

where $\hat{\mu}_t^j$, $j = s, b$ is the estimated conditional risk premium for stock or bond, respectively; $\hat{\rho}_t$ is the estimated conditional stock–bond correlation; and $\hat{\gamma}_t^j$, $\widehat{sd}_t^j$, and $\widehat{sk}_t^j$ $j = s, b$ are the estimated conditional stock or bond coskewness, volatility, and skewness, respectively. The p-values reported beneath each coefficient estimate are adjusted for a serial correlation of 12 lags and heteroskedasticity.

**5% and ***1% significance.

−0.251 for the stock and bond coskewnesses, respectively, which might be affected by the UK data quality problems in the early years as discussed previously. Indeed, the postwar sample estimates are closer to the U.S. estimates especially for the bond estimates (−0.495 and −0.425).

5. Concluding Remarks

Using a bivariate regime-switching model, this study examines time-varying ex ante conditional skewness and particularly conditional coskewness on stock and government bond markets in the United States and the United Kingdom in the last five decades as well as over the past 150 years. We report new evidence that U.S. government bonds exhibit positive, rather than negative, conditional skewness over most of the sample period, whereas U.S. stocks exhibited negative conditional skewness for extended periods of time. Hence, bonds are more likely to yield larger gains, and stocks are more likely to yield larger losses than suggested by a normal distribution, when investors buy and hold them during these time periods.

Further extending the previous literature (e.g., Harvey and Siddique 2000, Vanden 2006), we show that both stock coskewness and, particularly, bond coskewness command statistically and economically significant negative ex ante risk premiums on the stock and bond markets, which has not yet been documented. Hence, the desirable positive skewness of bonds can systematically offset much of the undesirable negative skewness of stocks if investors hold them together in a portfolio, and investors are willing to pay a higher price for such bond coskewness property. These findings are quite robust with respect to various model specifications and sample periods and in another major country (United Kingdom). We also find that in addition to coskewness, skewness on the U.S. stock and bond markets may also be priced, as suggested by the recent literature (Mitton and Vorkink 2007, Barberis and Huang 2008). The findings of this study carry an important asset pricing implication in that the lower expected return on bonds can be partially attributable to its desirable property of positive coskewness. Long-term investors may be willing to demand less compensation or lower returns to hold bonds that have positive coskewness with stocks (given all other things equal), whereas they demand higher returns to hold stocks for their negative coskewness. The findings also have another important implication for asset allocation: bonds may offer more diversification opportunities than we normally think based on the stock–bond correlation or the conventional mean-variance framework, supporting a higher strategic portfolio weight on bonds. In other words, bonds are not only a good hedge against stock market decline, but also a good hedge against the stock market volatility. The finding also suggests that coskewnesses between stock and bond returns should be accommodated for in future ICAPM studies, because existing studies mainly focus on covariance between stock and bond returns and might be inadequate.

Finally, this study can be extended in several ways. Obviously, because this study is limited to only two aggregate asset markets (stock and government bond), it would be interesting to investors to investigate whether stock and government bond coskewness helps explain the cross-section of stock returns and cross-section of bond returns. It would also be interesting to extend the research to emerging markets or other asset classes such as hedge funds and real estate, which might have positive rather than negative skewness (e.g., Bae et al. 2006). Investing in these markets or asset classes thus could reduce the negative skewness of a U.S. equity portfolio, suggesting that their coskewness might also be priced. We leave these topics for future research.

Acknowledgments
An earlier version of this paper was in circulation under the title "Regime Switching and Stock-Bond Co-Skewness." The authors are thankful to Warren Bailey and Sydney Ludvigson for providing some data used in this study. They gratefully acknowledge helpful comments on earlier versions from Warren Bailey, Turan G. Bali (discussant), Ethan Chiang (discussant), Hazem Daouk, Hui Guo, Yongmiao Hong, David Ng, Chunchi Wu, and seminar participants at Cornell University, the Chinese University of Hong Kong, Peking University, Rensselaer Polytechnic Institute, Singapore Management University, Sun Yat-Sen University, Southwestern University of Finance and Economics, Vanguard, the 2008 China International Conference in Finance, and the 2008 Financial Management Association annual meetings. Extensive comments from two anonymous referees, the editor (Wei Xiong), and the former editor (David A. Hsieh) have greatly improved this paper. Jian Yang also acknowledges partial financial support of the University of Colorado Denver CIBER research grant.

Appendix
This appendix contains the proofs of the proposition and the corollaries.

Proof of Proposition 1. From the law of iterated expectations we have

$$\begin{aligned}
&E[(r_t^s-\mu_t^s)^k(r_t^b-\mu_t^b)^l \mid \mathbf{F}_{t-1},\boldsymbol{\theta}]\\
&\quad =E\{E[(r_t^s-\mu_t^b)^k(r_t^b-\mu_t^b)^l \mid \mathbf{F}_{t-1},\boldsymbol{\theta},S_t]\}\\
&\quad =\sum_{i=1}^{2}p_{it}E[(\mu_{it}^s+\varepsilon_{it}^s-\mu_t^s)^k(\mu_{it}^b+\varepsilon_{it}^b-\mu_t^b)^l]\\
&\quad =\sum_{i=1}^{2}p_{it}\Bigg\{\sum_{m=0}^{k}\sum_{n=0}^{l}C_k^mC_l^n(\mu_{it}^s-\mu_t^s)^{k-m}\\
&\qquad\qquad \cdot(\mu_{it}^b-\mu_t^b)^{l-n}E[(\varepsilon_{it}^s)^m(\varepsilon_{it}^b)^n]\Bigg\},
\end{aligned}$$

where we use Newton's binomial formula and the assumption that the ex ante probabilities apply.

By assumption, $\boldsymbol{\varepsilon}_{it}=(\varepsilon_{it}^s,\varepsilon_{it}^b)'$ follows an i.i.d. bivariate normal distribution with zero means, the constant conditional volatilities and correlation, h_i^s, h_i^b, and ρ_i, given regime i. Then the moment-generating function of $\boldsymbol{\varepsilon}_{it}=(\varepsilon_{it}^s,\varepsilon_{it}^b)'$ can be expressed as

$$M(e_i^s,e_i^b)=\exp\Big[\tfrac{1}{2}(h_i^s(e_i^s)^2+2\sqrt{h_i^sh_i^b}\rho_ie_i^se_i^b+h_i^b(e_i^b)^2)\Big]. \qquad (21)$$

Repeated partial differentiation of Equation (21) and evaluation of the resultant expressions at $e_i^s=e_i^b=0$ yields Equation (8). Q.E.D.

Proof of Corollary 1. For the conditional mean of stock, $k=1$ and $l=0$. Following Proposition 1, we have

$$\begin{aligned}
&E[(r_t^s-\mu_t^s)\mid \mathbf{F}_{t-1},\boldsymbol{\theta}]\\
&\quad =\sum_{i=1}^{2}p_{it}\Bigg\{\sum_{m=0}^{1}[C_1^m(\mu_{it}^s-\mu_t^s)^{1-m}\varphi(m,0;i)]\Bigg\}\\
&\quad =\sum_{i=1}^{2}p_{it}(\mu_{it}^s-\mu_t^s)\varphi(0,0;i)\\
&\quad =\sum_{i=1}^{2}p_{it}(\mu_{it}^s-\mu_t^s),
\end{aligned}$$

where $\varphi(0,0;i)=1$ applies.

Note that $E[(r_t^s-\mu_t^s)\mid \mathbf{F}_{t-1},\boldsymbol{\theta}]=0$. This implies Equation (9).

For the conditional volatility of stock, $k=2$ and $l=0$. Following Proposition 1, we have

$$\begin{aligned}
&E[(r_t^s-\mu_t^s)^2\mid \mathbf{F}_{t-1},\boldsymbol{\theta}]\\
&\quad =\sum_{i=1}^{2}p_{it}\Bigg\{\sum_{m=0}^{2}[C_2^m(\mu_{it}^s-\mu_t^s)^{2-m}\varphi(m,0;i)]\Bigg\}\\
&\quad =\sum_{i=1}^{2}p_{it}\{(\mu_{it}^s-\mu_t^s)^2\varphi(0,0;i)\\
&\qquad\qquad +2(\mu_{it}^s-\mu_t^s)\varphi(1,0;i)+\varphi(2,0;i)\}\\
&\quad =\sum_{i=1}^{2}p_{it}\{(\mu_{it}^s-\mu_t^s)^2+h_i^s\}\\
&\quad =p_{1t}h_1^s+p_{2t}h_2^s+p_{1t}(\mu_{1t}^s-\mu_t^s)^2+p_{2t}(\mu_{2t}^s-\mu_t^s)^2\\
&\quad =p_{1t}h_1^s+p_{2t}h_2^s+p_{1t}(p_{2t})^2(\mu_{1t}^s-\mu_{2t}^s)^2+p_{2t}(p_{1t})^2(\mu_{2t}^s-\mu_{1t}^s)^2\\
&\quad =p_{1t}h_1^s+p_{2t}h_2^s+p_{1t}p_{2t}(\mu_{1t}^s-\mu_{2t}^s)^2,
\end{aligned}$$

where $\mu_t^s=\sum_{i=1}^2p_{it}\mu_{it}^s$, $\sum_{i=1}^2p_{it}=1$, $\varphi(0,0;i)=1$, $\varphi(1,0;i)=0$ and $\varphi(2,0;i)=(2-1)h_i^s\varphi(2-2,0;i)=h_i^s\varphi(0,0;i)=h_i^s$ apply.

Similarly, we can derive the formula for the conditional bond mean and volatility. Q.E.D.

Proof of Corollary 2. For the conditional covariance, $k=l=1$. Following from Proposition 1, we have

$$\begin{aligned}
&E\big[(r_t^s-\mu^s)(r_t^b-\mu^b)\mid \mathbf{F}_{t-1},\boldsymbol{\theta}\big]\\
&\quad =\sum_{i=1}^{2}p_{it}\Bigg\{\sum_{m=0}^{1}\sum_{n=0}^{1}\big[C_1^mC_1^n(\mu_{it}^s-\mu_t^s)^{1-m}(\mu_{it}^b-\mu_t^b)^{1-n}\varphi(m,n;i)\big]\Bigg\}\\
&\quad =\sum_{i=1}^{2}p_{it}\{(\mu_{it}^s-\mu_t^s)(\mu_{it}^b-\mu_t^b)\varphi(0,0;i)+(\mu_{it}^s-\mu_t^s)\varphi(0,1;i)\\
&\qquad\qquad +(\mu_{it}^b-\mu_t^b)\varphi(1,0;i)+\varphi(1,1;i)\}\\
&\quad =\sum_{i=1}^{2}p_{it}\Big\{(\mu_{it}^s-\mu_t^s)(\mu_{it}^b-\mu_t^b)+\sqrt{h_i^sh_i^b}\rho_i\Big\}
\end{aligned}$$

$$
\begin{aligned}
&= p_{1t}\sqrt{h_1^s h_1^b}\rho_1 + p_{2t}\sqrt{h_2^s h_2^b}\rho_2 + p_{1t}(\mu_{1t}^s-\mu_t^s)(\mu_{1t}^b-\mu_t^b) \\
&\quad + p_{2t}(\mu_{2t}^s-\mu_t^s)(\mu_{2t}^b-\mu_t^b) \\
&= p_{1t}\sqrt{h_1^s h_1^b}\rho_1 + p_{2t}\sqrt{h_2^s h_2^b}\rho_2 + p_{1t}(p_{2t})^2(\mu_{1t}^s-\mu_{2t}^s)(\mu_{1t}^b-\mu_{2t}^b) \\
&\quad + p_{2t}(p_{1t})^2(\mu_{2t}^s-\mu_{1t}^s)(\mu_{2t}^b-\mu_{1t}^b) \\
&= p_{1t}\sqrt{h_1^s h_1^b}\rho_1 + p_{2t}\sqrt{h_2^s h_2^b}\rho_2 + p_{1t}p_{2t}(\mu_{1t}^s-\mu_{2t}^s)(\mu_{1t}^b-\mu_{2t}^b) \\
&= p_{1t}\mathrm{cov}_1 + p_{2t}\mathrm{cov}_2 + p_{1t}p_{2t}(\mu_{1t}^s-\mu_{2t}^s)(\mu_{1t}^b-\mu_{2t}^b),
\end{aligned}
$$

where $\mu_t^s = \sum_{i=1}^2 p_{it}\mu_{it}^s$, $\mu_t^b = \sum_{i=1}^2 p_{it}\mu_{it}^b$, $\sum_{i=1}^2 p_{it} = 1$, $\varphi(0,0;i) = 1$, $\varphi(1,0;i) = \varphi(0,1;i) = 0$, and $\varphi(1,1;i) = \sqrt{h_i^s h_i^b}\rho_i = \mathrm{cov}_i$ apply.

The derivation of the conditional correlation is just by definition. Q.E.D.

Proof of Corollary 3. For the conditional coskewness, $k=1$ and $l=2$. Following from Proposition 1, we have

$$
\begin{aligned}
&E[(r_t^s-\mu^s)(r_t^b-\mu^b)^2 \mid \mathbf{F}_{t-1},\boldsymbol{\theta}] \\
&= \sum_{i=1}^2 p_{it}\left\{\sum_{m=0}^{1}\sum_{n=0}^{2}\left[C_1^m C_2^n(\mu_{it}^s-\mu_t^s)^{1-m}(\mu_{it}^b-\mu_t^b)^{2-n}\varphi(m,n;i)\right]\right\} \\
&= \sum_{i=1}^2 p_{it}\{(\mu_{it}^s-\mu_t^s)(\mu_{it}^b-\mu_t^b)^2\varphi(0,0;i) \\
&\qquad + 2(\mu_{it}^s-\mu_t^s)(\mu_{it}^b-\mu_t^b)\varphi(0,1;i) + (\mu_{it}^s-\mu_t^s)\varphi(0,2;i) \\
&\qquad + (\mu_{it}^b-\mu_t^b)^2\varphi(1,0;i) + 2(\mu_{it}^b-\mu_t^b)\varphi(1,1;i) + \varphi(1,2;i)\} \\
&= \sum_{i=1}^2 p_{it}\Big\{(\mu_{it}^s-\mu_t^s)(\mu_{it}^b-\mu_t^b)^2 \\
&\qquad + (\mu_{it}^s-\mu_t^s)h_i^b + 2(\mu_{it}^b-\mu_t^b)\sqrt{h_i^s h_i^b}\rho_i\Big\} \\
&= p_{1t}p_{2t}\Big[(p_{2t})^2(\mu_{1t}^s-\mu_{2t}^s)(\mu_{1t}^b-\mu_{2t}^b)^2 \\
&\qquad + (\mu_{1t}^s-\mu_{2t}^s)h_1^b + 2(\mu_{1t}^b-\mu_{2t}^b)\sqrt{h_1^s h_1^b}\rho_1\Big] \\
&\quad + p_{2t}p_{1t}\Big[(p_{1t})^2(\mu_{2t}^s-\mu_{1t}^s)(\mu_{2t}^b-\mu_{1t}^b)^2 \\
&\qquad + (\mu_{2t}^s-\mu_{1t}^s)h_2^b + 2(\mu_{2t}^b-\mu_{1t}^b)\sqrt{h_2^s h_2^b}\rho_2\Big] \\
&= p_{1t}p_{2t}[(p_{2t}-p_{1t})(\mu_{1t}^s-\mu_{2t}^s)(\mu_{1t}^b-\mu_{2t}^b)^2 \\
&\qquad + (\mu_{1t}^s-\mu_{2t}^s)(h_1^b-h_2^b) + 2(\mu_{1t}^b-\mu_{2t}^b)(\mathrm{cov}_1-\mathrm{cov}_2)],
\end{aligned}
$$

where

$$\mu_t^s = \sum_{i=1}^2 p_{it}\mu_{it}^s, \quad \mu_t^b = \sum_{i=1}^2 p_{it}\mu_{it}^b, \quad \sum_{i=1}^2 p_{it} = 1,$$

$$\varphi(0,0;i) = 1, \quad \varphi(1,0;i) = \varphi(0,1;i) = 0,$$

$$\varphi(0,2;i) = (2-1)h_i^b\varphi(0,2-2;i) = h_i^b\varphi(0,0;i) = h_i^b,$$

$$\varphi(1,1;i) = \sqrt{h_i^s h_i^b}\rho_i = \mathrm{cov}_i,$$

and $\varphi(1,2;i) = 2\sqrt{h_i^s h_i^b}\rho_i\varphi(0,1;i) + 0 = 0$ apply.

Then the stock coskewness in Equation (14) is derived by definition.

Similarly, we can derive the formula for the conditional bond coskewness. Q.E.D.

References

Ang, A., G. Bekaert. 2002. International asset allocation with regime shifts. *Rev. Financial Stud.* **15**(4) 1137–1187.

Ang, A., M. Piazzesi, M. Wei. 2006. What does the yield curve tell us about GDP growth? *J. Econometrics* **131**(1–2) 359–403.

Bae, K. H., A. Karolyi, R. Stulz. 2003. A new approach to measuring financial contagion. *Rev. Financial Stud.* **16**(3) 717–763.

Bae, K. H., C. Lim, K. C. J. Wei. 2006. Corporate governance and conditional skewness in the world's stock markets. *J. Bus.* **79**(6) 2999–3028.

Baele, L., G. Bekaert, K. Inghelbrecht. 2010. The determinants of Stock and bond return comovements. *Rev. Financial Stud.* **23**(6) 2374–2428.

Bali, T. G., K. O. Demirtas, H. Levy. 2009. Is there an intertemporal relation between downside risk and expected returns? *J. Financial Quant. Anal.* **44**(4) 883–909.

Bali, T. G., H. Mo, Y. Tang. 2008. The role of autoregressive conditional skewness and kurtosis in the estimation of conditional VaR. *J. Banking Finance* **32**(2) 269–282.

Barberis, N., M. Huang. 2008. Stocks as lotteries: The implications of probability weighting for security prices. *Amer. Econom. Rev.* **98**(5) 2066–2100.

Boyer, B., T. Mitton, K. Vorkink. 2010. Expected idiosyncratic skewness. *Rev. Financial Stud.* **23**(1) 169–202.

Briec, W., K. Kerstens, O. Jokung. 2007. Mean-variance-skewness portfolio performance gauging: A general shortage function and dual approach. *Management Sci.* **53**(1) 135–149.

Brunnermeier, M. K., J. A. Parker. 2005. Optimal expectations. *Amer. Econom. Rev.* **95**(4) 1092–1118.

Brunnermeier, M. K., C. Gollier, A. P. Jonathan. 2007. Optimal beliefs, asset prices, the preference for skewed returns. *Amer. Econom. Rev.* **97**(2) 159–165.

Campbell, J. Y., J. Ammer. 1993. What moves the stock and bond markets? A variance decompositions for long-term asset returns. *J. Finance* **48**(1) 3–37.

Capie, F., A. Webber. 1985. *A Monetary History of the United Kingdom, 1870–1982*. George Allen & Unwin, London.

Charoenrook, A., H. Daouk. 2005. Conditional skewness of aggregate market returns. Working paper, Cornell University, Ithaca, NY.

Chiu, W. H. 2005. Skewness preference, risk aversion, and the precedence relations on stochastic changes. *Management Sci.* **51**(12) 1816–1828.

Chung, Y. P., H. Johnson, M. J. Schill. 2006. Asset pricing when returns are nonnormal: Fama-French factors versus higher-order systematic comoments. *J. Bus.* **79**(2) 923–940.

Conine, T. E., M. J. Tamarkin. 1981. On diversification given asymmetry in returns. *J. Finance* **36**(5) 1143–1155.

Connolly, R., C. Stivers, L. Sun. 2005. Stock market uncertainty and the stock–bond return relation. *J. Financial Quant. Anal.* **40**(1) 161–194.

Connolly, R., C. Stivers, L. Sun. 2007. Commonality in the time variation of stock–bond and stock–stock return comovements. *J. Financial Markets* **10**(2) 192–218.

David, A., P. Veronesi. 2008. Inflation and earnings uncertainty and volatility forecasts: A structural form approach. Working paper, Booth School of Business, University of Chicago, Chicago.

Dimson, E., P. Marsh, M. Staunton. 2001. *Triumph of the Optimists: 101 Years of GlobalInvestment Returns*. Princeton University Press, Princeton, NJ.

Dittmar, R. F. 2002. Nonlinear pricing kernels, kurtosis preference, and evidence from the cross section of equity returns. *J. Finance* **57**(1) 369–403.

Gerard, B., G. Wu. 2006. How important is intertemporal risk for asset allocation? *J. Bus.* **79**(4) 2203–2241.

Goetzmann, W. N., L. Li, K. G. Rouwenhorst. 2005. Long-term global market correlations. *J. Bus.* **78**(1) 1–38.

Gray, S. F. 1996. Modeling the conditional distribution of interest rates as a regime-switching process. *J. Financial Econom.* **42**(1) 27–62.

Guidolin, M., A. Timmermann. 2005. Economic implications of bull and bear regimes in UK stock and bond returns. *Econom. J.* **115**(500) 111–143.

Guidolin, M., A. Timmermann. 2008. International asset allocation under regime switching, skew, and kurtosis preferences. *Rev. Financial Stud.* **21**(2) 889–935.

Guo, H., R. Savickas, Z. Wang, J. Yang. 2009. Is value premium a proxy for time-varying investment opportunities: Some time series evidence. *J. Financial Quant. Anal.* **44**(1) 133–154.

Hamilton, J. D. 1989. A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica* **57**(2) 357–384.

Hashmi, A. R., A. S. Tay. 2007. Global regional sources of risk in equity markets: Evidence from factor models with time-varying conditional skewness. *J. Internat. Money Finance* **26**(3) 430–453.

Harvey, C., A. Siddique. 1999. Autoregressive conditional skewness. *J. Financial Quant. Anal.* **34**(3) 465–487.

Harvey, C., A. Siddique. 2000. Conditional skewness in asset pricing tests. *J. Finance* **54**(3) 1263–1295.

Ilmanen, A. 2003. Stock-bond correlations. *J. Fixed Income* **13**(2) 55–66.

Kim, C.-J. 1993. Dynamic linear models with Markov-switching. *J. Econometrics* **60**(1) 1–22.

Kimball, M. S. 1990. Precautionary saving in the small and in the large. *Econometrica* **58**(1) 53–73.

Kraus, A., R. Litzenberger. 1976. On the distributional conditions for a consumption-oriented three moment CAPM. *J. Finance* **38**(5) 1381–1391.

Longin, F., B. Solnik. 2001. Extreme correlation of international equity markets. *J. Finance* **56**(2) 649–676.

Ludvigson, S. C., S. Ng. 2009. Macro factors in bond risk premia. *Rev. Financial Stud.* **22**(12) 5027–5067.

Menezes, C. F., C. Geiss, J. Tressler. 1980. Increasing downside risk. *Amer. Econom. Rev.* **70**(5) 921–932.

Merton, R. C. 1973. An intertemporal capital asset pricing model. *Econometrica* **41**(5) 867–887.

Mitton, T., K. Vorkink. 2007. Equilibrium underdiversification and the preference for skewness. *Rev. Financial Stud.* **20**(4) 1255–1288.

Newey, W., K. West. 1987. A simple, positive semi-definite, heteroskedastcity and autocorrelation consistent covariance matrix. *Econometrica* **55**(3) 703–708.

Perez-Quiros, G., A. Timmermann. 2001. Business cycle asymmetries in stock returns: Evidence from higher order moments and conditional densities. *J. Econometrics* **103**(1–2) 259–306.

Piazzesi, M. 2005. Bond yields and the federal reserve. *J. Political Econom.* **113**(2) 311–344.

Prakash, A. J., C. H. Chang, T. Pactwa. 2003. Selecting a portfolio with skewness: Recent evidence from U.S., European, and Latin American stock markets. *J. Banking Finance* **27**(7) 1375–1390.

Rubinstein, M. 1973. The fundamental theorem for parameter preference security valuation. *J. Financial Quant. Anal.* **8**(1) 61–69.

Schwert, W. G. 1990. Indexes of united states stock prices from 1802 to 1987. *J. Bus.* **63**(3) 399–426.

Scruggs, J. 1998. Resolving the puzzling intertemporal relation between the market risk premium and conditional market variance: A two factor approach. *J. Finance* **53**(2) 575–603.

Scruggs, J. T., P. Glabadanidis. 2003. Risk premia and the dynamic covariance between stock and bond returns. *J. Financial Quant. Anal.* **38**(2) 295–316.

Shanken, J. 1985. Multivariate tests of the zero-beta CAPM. *J. Financial Econom.* **14**(3) 327–348.

Timmermann, A. 2000. Moments of Markov switching models. *J. Econometrics* **96**(1) 75–111.

Xing, Y., X. Zhang, R. Zhao. 2010. What does individual option volatility smile tell us about future equity returns. *J. Financial Quant. Anal.* **45**(3) 641–662.

Xu, J. 2007. Price convexity and skewness. *J. Finance* **62**(5) 2521–2552.

Vanden, J. M. 2006. Option coskewness and capital asset pricing. *Rev. Financial Stud.* **19**(4) 1279–1320.