The Tradability Premium on the S&P 500 Index

Christian Gourieroux

CREST and University of Toronto

Email: christian.gourieroux@ensae.fr.

Joann Jasiak

York University

Email: jasiakj@yorku.ca

Peng Xu

ESSEC Business School, Asia-Pacific Campus

2 One-North Gateway, Singapore, 138502

Email: xup@essec.edu. Tel.: +65 68357471

Abstract

We derive a coherent multi-factor model for pricing various derivatives written on the same underlying (potentially non-tradable) asset. We show the difference between a case in which the underlying asset is self-financed and tradable and a case in which it is not. In the first case, an additional arbitrage condition must be introduced, which implies nontrivial parameter restrictions. These restrictions can be empirically tested to check whether the derivatives are priced as if the underlying were self-financed and tradable. This methodology also allows us to define the tradability premium. As an illustration, we compute the daily tradability premium for the S&P 500.

Keywords: Index Derivatives, Non-tradable Index, Generalized Method of Moments, Mispricing, Tradability Premium, Liquidity Premium

JEL Classification: C13, C51, G12, G13

The S&P 500 Index is an artificial number constructed to reflect the evolution of the market. The index is not traded in the financial market and it does not represent the price of a tradable portfolio. As explained in Xu (2014), it cannot be replaced with a mimicking portfolio, such as the SPDR, due to the way in which the S&P 500 Index is calculated and maintained, the lack of perfect foresight, the illiquidity of some index component stocks, and because of differences in the ways in which dividends and transaction costs are (not) accounted for. In other words, we cannot buy and hold a portfolio in the financial market equivalent to the index's current value and be guaranteed that this portfolio can be sold in the future for the index's value at that point.

The non-tradability of the S&P 500 Index has significant implications for risk hedging and pricing constraints. For example, the well-known Black-Scholes model [See Black and Scholes (1973) and Merton (1973)] assumes that the underlying asset is tradable and follows a geometric Brownian motion process with constant volatility. Therefore, the market is completed by the underlying asset itself and the underlying asset can be used to fully hedge against the risk involved. Under the no-arbitrage condition, the market price of risk is determined uniquely by the price of the underlying asset. All derivatives written on the underlying asset can be evaluated uniquely using this market price of risk combined with the terminal condition of the respective derivatives. If the underlying asset is non-tradable, then the underlying asset cannot be used as part of the arbitrage strategy and the value of the underlying asset does not need to satisfy the no-arbitrage condition. The risk associated with the underlying asset is not hedged by itself and the expected return of the underlying asset under the risk-neutral probability is not necessarily equal to the risk-free rate. The knowledge of the value of the underlying asset does not fully reveal the price of the risk. Therefore, the prices of options written on a non-traded underlying asset with a price that follows a geometric Brownian motion process do not have to be evaluated using the Black-Scholes formula. Similar ideas apply to other models. For instance, the stochastic volatility models in Heston (1993) and Ball and Roma (1994) assume that the expected return of the underlying asset is equal to the risk-free rate under the risk-neutral probability. In other words, those models assume that the underlying asset is tradable and that the risk associated with the underlying asset is hedged by itself. In general, because of the non-tradability of the S&P 500 Index, the prices of its options do not have to satisfy the restrictions imposed by pricing models that are based on the assumption that the underlying asset is a security traded in the market.

In this paper, we introduce a coherent multi-factor model for pricing various derivatives such as forwards, futures, and European options, written on the non-tradable S&P 500 Index. The model illustrates the relationship between the index and its futures, and the relationship between the index and its put and call options when the underlying asset is non-tradable. We also consider what the prices of the derivatives would be if the index were self-financed and tradable. The model explains why the prices of derivatives written on a tradable asset can differ from those written on a non-tradable asset. When the underlying asset is self-financed and tradable, it also needs to satisfy the no-arbitrage condition, which implies additional nontrivial parameter restrictions. This setup allows us to compute the premium of tradability for each day, which is defined as the difference between the market risk premium implied by the unrestricted model and the market risk premium implied by the restricted model. We use simulated data to illustrate how to estimate the tradability premium and its impact on derivative pricing. We find that certain derivatives, such as options, will be significantly mispriced if the tradability premium is ignored .

In this framework, we are also able to test whether the S&P 500 derivatives are priced by investors as if the index were self-financed and tradable. Our factor models are estimated by combining the spot, futures, and options data, and using the unscented Kalman Filter (UKF) method. The Wald tests strongly reject the null hypothesis that the derivatives are priced as if the index were self-financed and tradable. Our diagnostic analysis also shows that the multi-factor models are superior to the one-factor model. To test the robustness of our results, we investigate the restricted version of our model by assuming that the S&P 500 Index is self-financed and tradable. Both in-sample and out-of-sample pricing errors show that the unrestricted model performs statistically and economically better than the restricted model. The estimated tradability premium is significantly different from zero, which means that the tradability of the underlying asset is an important factor in derivative pricing.

Our model can easily be extended to pricing derivatives written on other non-tradable indices, such as a retail price index, a meteorological index, an index summarizing the results of a set of insurance companies, a population mortality index, or the VIX.

The rest of the paper is organized as follows. In Section 1, we present a coherent model for pricing derivatives written on the S&P 500. In Section 2, we derive the parameter restrictions that would characterize the derivative pricing if the index were tradable. This allows us to explain how derivative prices can differ for tradable and non-tradable underlying assets. We discuss the importance of tradability premium in derivative pricing in Section 3. In Section 4, we undertake a Monte Carlo simulation with realistically calibrated parameters to illustrate how the tradability premium is measured and how it affects the pricing of derivatives. In Section 5, we discuss the estimation method and testing procedure. In Section 6, we report the empirical results and compute the daily tradability premium. We conclude in Section 7. The technical results and details are gathered in the appendices.

1 The Pricing Model

Under the absence of arbitrage opportunity (AAO), market prices have to be compatible with a valuation system based on stochastic discounting [Harrison and Kreps (1979)]. The pricing formulas can be written in either discrete time or continuous time, according to the assumptions of discrete or continuous trading (and information sets). The modern pricing methodology requires a joint, coherent specification of the historical and risk-neutral distributions. For this purpose, we follow the practice initially introduced by Constantinides (1992), which specifies a parametric historical distribution and a parametric stochastic discount factor.

1.1 Assumptions

1.1.1 Historical Dynamics of the Index

The value of the index at date t is denoted by I_t. We assume that the log index satisfies a diffusion equation with affine drift and volatility functions of K underlying factors $\{x_{k,t}\}$, $k = 1, \cdots, K$.

Assumption 1.

$$d \log I_t = (\mu_0 + \sum_{k=1}^{K} \mu_k x_{k,t})dt + (\gamma_0 + \sum_{k=1}^{K} \gamma_k x_{k,t})^{1/2} dw_t, \tag{1.1}$$

where $\{\mu_k\}$ *and* $\{\gamma_k\}$, $k = 0, \cdots, K$ *are constants, and* $\{w_t\}$ *is a Brownian motion.*

The underlying factors summarize the dynamic features of the index. As seen in Equation (1.2), they are assumed to be independent Cox, Ingersoll, and Ross (CIR) processes that are independent of the standard Brownian motion $\{w_t\}$. As the CIR processes are nonnegative, the volatility of the log index is positive whenever parameters $\{\gamma_k\}$, $k = 0, \cdots, K$ are positive. This positive parameter restriction is imposed throughout the rest of the paper.

Assumption 2. *The CIR processes* $\{x_{k,t}\}$, $k = 1, \cdots, K$ *satisfy the stochastic differential equations:*

$$dx_{k,t} = \xi_k(\zeta_k - x_{k,t})dt + \nu_k \sqrt{x_{k,t}} dw_{k,t}, \qquad k = 1, \cdots, K, \tag{1.2}$$

where ξ_k, ζ_k *and* ν_k *are positive constants, and* $\{w_{k,t}\}$, $k = 1, \cdots, K$ *are standard independent Brownian motions that are independent of* $\{w_t\}$.

The condition $\xi_k \zeta_k > 0$ ensures the nonnegativity of the CIR process (for a positive initial value $x_0 > 0$), while the conditions $\xi_k > 0$ and $\zeta_k > 0$ imply the stationarity of the CIR process. The condition $\nu_k > 0$ can always be assumed for identifiability reason.

This general specification of the index dynamics includes the Black-Scholes model [Black and Scholes (1973)], when $\mu_k = \gamma_k = 0,\ k = 1, \cdots, K$; the stochastic volatility model considered by Heston (1993) and Ball and Roma (1994), when $K = 1$ and x_1 is interpreted as a stochastic volatility; or the model with stochastic dividend yield [see, for example, Schwartz (1997)], when $K = 1$ and x_1 appears only in the drift.

The transition distribution of the integrated CIR process is required for derivative pricing. This distribution is characterized by the conditional Laplace transform $E_t[\exp(-z \int_t^{t+h} x_{k,\tau} d\tau)]$, where E_t denotes the conditional expectation given the past values of the process and z is a nonnegative constant (or, more generally, a complex number), which belongs to the domain of the existence of the conditional Laplace transform. This domain does not depend on past factor realizations (i.e., on the information set). The conditional Laplace transform of the integrated CIR process allows for a closed-form expression [see, e.g., Cox, Ingersoll, and Ross (1985b)] and is an exponential affine function of the current factor value. It is given by:

$$E_t[\exp(-z \int_t^{t+h} x_{k,\tau} d\tau)] = \exp[-H_1^k(h, z) x_{k,t} - H_2^k(h, z)], \tag{1.3}$$

where

$$\begin{aligned} H_1^k(h, z) &= \frac{2z(exp[\varepsilon_k(z)h] - 1)}{(\varepsilon_k(z) + \xi_k)(exp[\varepsilon_k(z)h] - 1) + 2\varepsilon_k(z)}, \\ H_2^k(h, z) &= \frac{-2\xi_k \zeta_k}{\nu_k^2} \{\log[2\varepsilon_k(z)] + \frac{h}{2}[\varepsilon_k(z) + \xi_k] \\ &\quad - \log[(\varepsilon_k(z) + \xi_k)(\exp(\varepsilon_k(z)h) - 1) + 2\varepsilon_k(z)]\}, \\ \varepsilon_k(z) &= \sqrt{\xi_k^2 + 2z\nu_k^2}. \end{aligned} \tag{1.4}$$

This formula also holds for a complex number $z = u + iv$ whenever $u > -1$ and $v \in \mathbf{R}$.

The joint dynamics of factors and log index can be represented by means of the stochastic differential system in which both the drift vector and the volatility-covolatility matrix are affine functions of the current values of the joint process $(x_{1,t}, \cdots, x_{K,t}, \log I_t)'$. Thus, the stacked process $(x_{1,t}, \cdots, x_{K,t}, \log I_t)'$ is an affine process [see Duffie and Kan (1996)], and the conditional Laplace transform of the integrated process $E_t[\exp\int_t^{t+h}(z_1 x_{1,\tau} + \cdots + z_K x_{K,\tau} + z\log I_\tau)d\tau]$ will also allow for an exponential affine closed-form expression.

1.1.2 Specification of the Stochastic Discount Factor

The model is completed with a specification of a stochastic discount factor (SDF), which we later use to price all derivatives written on the index.

Assumption 3. *The stochastic discount factor (SDF) for period (t, t+dt) is*

$$\begin{aligned} M_{t,t+dt} &= \exp(dm_t) = \exp[(\alpha_0 + \sum_{k=1}^{K} \alpha_k x_{k,t})dt + \beta d\log I_t] \\ &= \exp\{[\alpha_0 + \beta\mu_0 + \sum_{k=1}^{K}(\alpha_k + \beta\mu_k)x_{k,t}]dt + \beta(\gamma_0 + \sum_{k=1}^{K}\gamma_k x_{k,t})^{1/2} dw_t\}. \end{aligned} \tag{1.5}$$

This SDF explains how to correct for risk when pricing derivatives. The "risk premia" depend on the factors and index values, whereas the sensitivities of this correction with respect to those risk variables are represented by the α and β parameters. The market risk premium associated with w_t is $-\beta(\gamma_0 + \sum_{k=1}^{K}\gamma_k x_{k,t})$. This specification of the SDF implicitly assumes that the market prices of the risk factors $\{w_{k,t}\}$, $k = 1, \cdots, K$ are 0. Equivalently, Equation (1.2) also describes the risk-neutral distribution of $\{x_{k,t}\}$, $k = 1, \cdots, K$. Under the risk-neutral probability, the joint dynamics of the underlying factors and log index can be represented by means of the stochastic differential system:

$$d\begin{bmatrix} x_{1,t} \\ \vdots \\ x_{K,t} \\ \log I_t \end{bmatrix} = \begin{bmatrix} \xi_1(\zeta_1 - x_{1,t}) \\ \vdots \\ \xi_K(\zeta_K - x_{K,t}) \\ \mu_0 + \beta\gamma_0 + \sum(\mu_k + \beta\gamma_k)x_{k,t} \end{bmatrix} dt \tag{1.6}$$

$$+ \begin{bmatrix} \nu_1\sqrt{x_{1,t}} & 0 & \cdots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & \nu_K\sqrt{x_{K,t}} & 0 \\ 0 & \cdots & 0 & \sqrt{\gamma_0 + \sum\gamma_k x_{k,t}} \end{bmatrix} \begin{bmatrix} dw_{1,t} \\ \vdots \\ dw_{K,t} \\ dw_t^* \end{bmatrix},$$

where $\{w_{k,t}\}$, $k = 1, \cdots, K$, and $\{w_t^*\}$ are standard independent Brownian motions under the risk-neutral probability. Thus, only the last row is corrected for risk. This differential stochastic system is still an affine process.

1.2 Pricing Formulas for European Derivatives Written on the Index

As mentioned above, the arbitrage pricing proposes a valuation approach, which is compatible with observed market prices and proposes coherent quotes for non-highly traded derivatives. More precisely, the value (price) at time t of a European derivative paying $g(x_{1,t+h}, \cdots, x_{K,t+h}, I_{t+h})$ at time $t + h$ is:

$$c(t, t+h, g) = E_t[\exp(\int_t^{t+h} dm_\tau) g(x_{1,t+h}, \cdots, x_{K,t+h}, I_{t+h})]. \tag{1.7}$$

The aim of this section is to derive explicit valuation formulas for European index derivatives[1]. All of the formulas are derived from the valuation of European index derivatives with power payoff. Such derivatives are not traded or, more generally, quoted. However, these basic computations are used to derive:

- The risk-free term structure of interest rates,
- The forward and futures prices of the index, and
- The prices of European options written on the index.

1.2.1 Power Derivatives Written on the Index

The proof for the following proposition is provided in Appendix A.

Proposition 1. *The value at time t of the European derivative paying* $\exp[u\log(I_{t+h})] = (I_{t+h})^u$ *at maturity t+h is:*

$$C(t,t+h,u) = E_t[(I_{t+h})^u \exp(\int_t^{t+h} dm_\tau)]$$
$$= \exp(u\log I_t)\exp[-hz_0(u) - \sum_{k=1}^{K} H_1^k(h, z_k(u))x_{k,t} - \sum_{k=1}^{K} H_2^k(h, z_k(u))], \quad (1.8)$$

where

$$z_k(u) = -\alpha_k - (\beta+u)\mu_k - \frac{\gamma_k}{2}(\beta+u)^2, \quad \forall\, k = 0, \cdots, K, \quad (1.9)$$

and $H_1^k(\cdot,\cdot)$ and $H_2^k(\cdot,\cdot)$ are given in equation(1.4).

Proposition 1 holds if and only if $z_k(u) > -1$, $\forall\ k = 1, \cdots, K$. When we apply this formula to different traded derivatives (i.e., different values of u), the inequalities above imply restrictions on parameters α, β, and γ.

1.2.2 The Risk-free Term Structure

The zero-coupon bonds correspond to a unitary payoff and their prices $B(t,t+h)$ correspond to the special case of $C(t,t+h,u)$ in which $u = 0$. The continuously compounded risk-free interest rates are defined by $r(t,t+h) = -\frac{1}{h}\log B(t,t+h)$. We arrive at the following proposition:

Proposition 2. *The prices of the zero-coupon bonds are:*

$$B(t,t+h) = \exp[-hz_0(0) - \sum_{k=1}^{K} H_1^k(h, z_k(0))x_{k,t} - \sum_{k=1}^{K} H_2^k(h, z_k(0))], \quad (1.10)$$

where $z_k(\cdot)$ is defined in Equation (1.9), and $H_1^k(\cdot,\cdot)$ and $H_2^k(\cdot,\cdot)$ are given in equation(1.4). We deduce the expressions of the interest rates:

$$r(t,t+h) = -\frac{1}{h}\log B(t,t+h)$$
$$= z_0(0) + \frac{1}{h}\sum_{k=1}^{K} H_1^k(h, z_k(0))x_{k,t} + \frac{1}{h}\sum_{k=1}^{K} H_2^k(h, z_k(0)). \quad (1.11)$$

The risk-free interest rates are affine functions of the CIR risk factors. This specification is the standard affine term structure model introduced in Duffie and Kan (1996) [see also Dai and Singleton (2000)]. It includes the one-factor CIR model [Cox, Ingersoll, and Ross (1985b)] as well as the multi-factor term structure model found in Chen and Scott (1993).

As explained in subsection 1.2.1, the following restrictions are imposed on the parameters:

$$z_k(0) = -\alpha_k - \beta\mu_k - \frac{\gamma_k}{2}\beta^2 > -1, \quad \forall\, k = 1, \cdots, K. \tag{1.12}$$

The short rate is defined by $r(t) = \lim_{h\to 0} -\frac{1}{h}\log B(t, t+h)$. The proof of the following proposition is provided in Appendix B.

Proposition 3. *The short rate is given by:*

$$\begin{aligned} r(t) &= \lim_{h\to 0} -\frac{1}{h}\log B(t,t+h) = \frac{d[-\log B(t,t+h)]}{dh}\,|_{h=0} \\ &= z_0(0) + \sum_{k=1}^{K} z_k(0)x_{k,t}. \end{aligned} \tag{1.13}$$

1.2.3 Forward Prices for the S&P 500 Index

A forward contract is an agreement to deliver or receive a specified amount of the underlying asset (or the equivalent cash value) at a specified price and date. A forward contract always has zero value when it is initiated. No money is exchanged initially or during the life of the contract, except at the maturity date when the price paid is equal to the specified forward price. The proof of the following proposition is provided in Appendix C.

Proposition 4. *The forward prices are given by:*

$$\begin{aligned} f(t,t+h) &= \frac{C(t,t+h,1)}{C(t,t+h,0)} \\ &= I_t \exp\{-hl_0 - \sum_{k=1}^{K} H_1^k(h, z_k(1))x_{k,t} + \sum_{k=1}^{K} H_1^k(h, z_k(0))x_{k,t} \\ &\quad - \sum_{k=1}^{K} H_2^k(h, z_k(1)) + \sum_{k=1}^{K} H_2^k(h, z_k(0))\}, \end{aligned} \tag{1.14}$$

where

$$l_k = -\mu_k - \frac{1+2\beta}{2}\gamma_k, \quad \forall\, k = 0, \cdots, K, \tag{1.15}$$

$z_k(\cdot)$ is defined in Equation (1.9), l_0 is defined in Equation (1.15), and $H_1^k(\cdot,\cdot)$ and $H_2^k(\cdot,\cdot)$ are given in equation(1.4).

In addition to the restrictions in (1.12), the following restrictions are imposed on the parameters:

$$z_k(1) = -\alpha_k - (\beta+1)\mu_k - \frac{\gamma_k}{2}(\beta+1)^2 > -1, \quad \forall\, k = 1, \cdots, K. \tag{1.16}$$

1.2.4 Futures Prices

Let us now consider the price at time t of a futures contract written on I_{t+h}. The major difference between a futures contract and a forward contract is the mark-to-market practice for the futures. A futures contract has also zero value when it is issued and no money is exchanged initially. However, at the end of each trading day during the life of the contract, the party against whose favor the price changes must pay the amount of change to the winning party. In other words, a futures contract always has zero value at the end of each trading day throughout the life of the contract. If the interest rate is stochastic, the forward price and the futures price are generally not the same [see Cox, Ingersoll, and Ross (1981) and French (1983)]. The proof of the following proposition is provided in Appendix D.

Proposition 5. *The prices at time t of futures written on I_{t+h} are given by:*

$$\begin{aligned} F_{t,t+h} &= E_t[\exp(\int_t^{t+h} dm_\tau)\exp(\int_t^{t+h} r_\tau d\tau) I_{t+h}] \\ &= I_t \exp[-hl_0 - \sum_{k=1}^{K} H_1^k(h, l_k)x_{k,t} - \sum_{k=1}^{K} H_2^k(h, l_k)], \end{aligned} \tag{1.17}$$

where l_k is defined in Equation (1.15), and $H_1^k(\cdot,\cdot)$ and $H_2^k(\cdot,\cdot)$ are given in equation(1.4).

As explained earlier, in addition to the restrictions in (1.12), the following restrictions are imposed on the parameters:

$$l_k = -\mu_k - \frac{1+2\beta}{2}\gamma_k > -1 \quad \forall\, k = 1, \cdots, K. \tag{1.18}$$

Propositions 4 and 5 show that

$$f(t,t+h) = E_t[\exp(\int_t^{t+h} dm_\tau)\exp(r(t,t+h)h)I_{t+h}]$$

and

$$F_{t,t+h} = E_t[\exp(\int_t^{t+h} dm_\tau)\exp(\int_t^{t+h} r_\tau d\tau)I_{t+h}].$$

As the short rate is stochastic, the forward and futures prices are not equal in general. A sufficient condition for the forward and futures prices to be identical is: $z_k(0) = 0,\ \forall\ k = 1,\cdots,K$, i.e., the interest rates are non-stochastic. This is Proposition 3 in Cox, Ingersoll, and Ross (1981).

1.2.5 European Call and Put Options Written on the Index

The prices of the European options are deduced by applying a transform analysis to function $C(t,t+h,u)$ computed for pure imaginary argument u [see Duffie, Pan, and Singleton (2000) and Appendix E].

Proposition 6.

i) The European call prices are given by:

$$G(t,t+h,X) = E_t\{\exp(\int_t^{t+h} dm_\tau)[\exp(\log I_{t+h}) - X]^+\} \tag{1.19}$$

$$= \frac{C(t,t+h,1)}{2} - \frac{1}{\pi}\int_0^\infty \frac{Im[C(t,t+h,1-iv)\exp(iv\log X)]}{v}dv - X\{\frac{C(t,t+h,0)}{2} - \frac{1}{\pi}\int_0^\infty \frac{Im[C(t,t+h,-iv)\exp(iv\log X)]}{v}dv\}, \tag{1.20}$$

where X is the strike price, h is the time to maturity, i denotes the pure imaginary number, and $Im(\cdot)$ is the imaginary part of a complex number.

ii) The European put prices are given by:

$$H(t,t+h,X) = E_t\{\exp(\int_t^{t+h} dm_\tau)[X - \exp(\log I_{t+h})]^+\} \tag{1.21}$$

$$= -\frac{C(t,t+h,1)}{2} + \frac{1}{\pi}\int_0^\infty \frac{Im[C(t,t+h,1+iv)\exp(-iv\log X)]}{v}dv + X\{\frac{C(t,t+h,0)}{2} - \frac{1}{\pi}\int_0^\infty \frac{Im[C(t,t+h,iv)\exp(-iv\log X)]}{v}dv\}. \tag{1.22}$$

Again, the restrictions in (1.12) and (1.16) are imposed[2].

2 Parameter Restrictions for a Tradable Index

In Section 1, the pricing formulas are valid for tradable and non-tradable indexes. In this section, we derive the restrictions implied by the tradability of the underlying index.

When the benchmark index is a self-financed and tradable asset, the pricing formula is valid for the index itself. In that case, we have an additional condition:

$$I_t = E_t[\exp(\int_t^{t+h} dm_\tau) I_{t+h}] = C(t, t+h, 1),$$

such that

$$C(t, t+h, 1) = I_t \exp[-hz_0(1) - \sum_{k=1}^{K} H_1^k(h, z_k(1)) x_{k,t} - \sum_{k=1}^{K} H_2^k(h, z_k(1))], \tag{2.1}$$

where $z_k(\cdot)$ is defined in Equation (1.9), and $H_1^k(\cdot,\cdot)$ and $H_2^k(\cdot,\cdot)$ are given in equation (1.4). $z_k(1) > -1$ is imposed, $\forall\, k = 1, \cdots, K$.

This additional pricing condition has to be satisfied in any environment. This implies a continuum number of restrictions. However, due to the linearity of these restrictions, they can be reduced to a finite number of restrictions, so we do not need to use the advanced GMM method found in Carrasco, Chernov, Florens, and Ghysels (2000). If we consider the expression $C(t, t+h, 1)$ and identify the different terms in the decomposition, we see that the dynamic parameters are constrained by:

$$\begin{cases} H_1^k(h, z_k(1)) = 0, & \forall\, k = 1, \cdots, K, \quad \forall\, h, \\ -hz_0(1) - \sum_{k=1}^{K} H_2^k(h, z_k(1)) = 0, & \forall\, h, \end{cases} \tag{2.2}$$

or equivalently by the conditions shown in Proposition 7 (see the proof in Appendix F).

Proposition 7. *When the benchmark index is a self-financed and tradable asset, the dynamic parameters are constrained by:*

$$z_k(1) = \alpha_k + (\beta+1)\mu_k + \frac{\gamma_k}{2}(\beta+1)^2 = 0, \quad \forall\, k = 0, \cdots, K. \tag{2.3}$$

These restrictions fix the parameters $\{\alpha_k\}$, $k = 0, \cdots, K$, and β of the SDF as functions of the parameters of the index dynamics.

The risk-neutral dynamics of $\log I_t$ is given in Equation (1.6). When the benchmark index is tradable, the risk-neutral dynamics of $\log I_t$ can also be written as:

$$d \log I_t = [r(t) - \frac{\gamma_0 + \sum \gamma_k x_{k,t}}{2}]dt + \sqrt{\gamma_0 + \sum \gamma_k x_{k,t}} dw_t^*, \tag{2.4}$$

or

$$\frac{dI_t}{I_t} = r(t)dt + \sqrt{\gamma_0 + \sum \gamma_k x_{k,t}} dw_t^*. \tag{2.5}$$

In other words, if the index is tradable, the no-arbitrage condition requires its risk neutral drift to be equal to the short rate. Therefore, conditional on the underlying factors, the risk w_t can be hedged by the index and the short rate if the index is tradable. Equivalently, if the index is tradable, given the underlying factors, the market risk premium associated with w_t is solely determined by the index's expected return and the short rate.

If the benchmark index is tradable, the formulas of derivative prices can be simplified. In particular, the forward price derived in Proposition 4 can be simplified to the standard formula:

$$f(t, t+h) = \frac{I_t}{B(t, t+h)}, \tag{2.6}$$

and the spot-futures parity will hold for the index and its forward price. Similarly, if the benchmark index is tradable, the relationship between the prices of European call and put options can be written as:

$$G(t, t+h, X) - I_t = H(t, t+h, X) - XC(t, t+h, 0), \tag{2.7}$$

which is the standard put-call parity.

3 Tradability Premium

In this framework, we can also measure the "tradability premium", which is defined as the difference between the market price of risk implied by the unrestricted model and the market price of risk implied by the restricted model in which the underlying asset is assumed to be self-financed and tradable.

The market price of risk is a key factor in determining derivative prices. When the underlying asset is tradable, its price should include a tradability premium, which should be reflected in the market price of risk associated with the underlying asset. When the underlying asset is not tradable, the market risk premium does not reflect the tradability premium. The difference between these two market risk premia is the tradability premium. In other words, if an investor estimates a pricing model assuming that the underlying asset is tradable when, in fact, it is not, s/he is essentially ignoring the tradability premium. This practice can lead to significant errors in derivative pricing. In this paper, the market price of risk associated with w_t is $-\beta(\gamma_0 + \sum_{k=1}^{K} \gamma_k x_{k,t})$. By looking at this market risk premium, which is estimated from the unrestricted model and from the restricted model with derivatives data, we can estimate the tradability premium. This is illustrated in Section 4.

4 Monte Carlo Simulation and Illustration

In this section, we use the one-factor affine model as an example. We implement a Monte Carlo study to illustrate how the tradability premium is estimated and how it affects derivative pricing.

To mirror the empirical estimation in the next section, we simulate the daily (1/252 years) value of the spot index; the prices of two index futures with 30 days and 120 days until maturity; and the prices of three index call options maturing in 30 days, 90 days, and 250 days with moneyness (S/X) of 1.02, 0.97, and 0.93, respectively. These simulated data are used to estimate parameters. The sample size is $T = 1,500$ and the total number of simulations is 50. The underlying factor $x_{1,t}$ is simulated based on Equation (1.2) using a fine discretion scheme ($dt = \frac{1}{30\times 252}$) to ensure that the data are generated from the true continuous-time model. The spot index value is also simulated using the fine discretion scheme based on Equation (1.1), but only the *daily* log difference of the spot index is used for estimation purpose. The observed prices of the futures and options are generated by adding white noise to the true prices computed based on Equations (1.17) and (1.19). The white noise is assumed to be normally distributed with different variances for the futures and the options. The parameters to be estimated are μ_0, μ_1, γ_0, γ_1, ξ_1, ζ_1, α_0, α_1, and β. We use θ to denote the vector of all the parameters. As explained in the next section, ν_1 is normalized to 1. The true values of the parameters are assumed to be θ^*=(0.015, -0.02, 0.005, 0.09, 3, 0.1, 0.28, 0.12, -0.42), which are close to the empirical estimates in Section 6. The model is estimated using the maximum likelihood estimation (MLE) method.

The first part of Table 1 reports the mean and standard deviation of the parameter estimates across 50 simulations. The second and third columns present the estimation results for the unrestricted model, while the last two columns present the results for the restricted model when the tradability restrictions in Equation (2.3) are imposed. As the observed prices are generated without imposing the tradability restrictions in Equation (2.3), we are not surprised to see that the parameter estimates in the unrestricted model are very close to the true values and show small standard deviations, while the parameter estimates in the restricted model are quite different from their true values. Therefore, if an investor incorrectly imposes the tradability restriction on the prices of derivatives for which the underlying asset is not actually tradable, the results of the estimation will be biased and inconsistent.

The second part of Table 1 reports the mean and standard deviation of the daily average in-sample pricing errors across 50 simulations. The pricing errors are measured as the absolute difference between the model-implied prices using the estimated parameters and the observed prices as a percentage of the observed prices. As the error terms are assumed to be normally distributed, the sum of the squared residuals between the observed prices and the model-implied prices are minimized in the MLE estimation. As a result, the in-sample pricing errors in both the unrestricted model and the restricted model are quite small, although the pricing errors in the restricted model are slightly higher for the options.

The market price of risk is crucial in derivative pricing. The tradability premium is measured as the difference between the market risk premium computed with parameters estimated from the unrestricted model and the market risk premium computed with parameters estimated from the restricted model. In other words, the tradability premium is the difference between the "right" market price of risk and the "wrong" market price of risk. In this paper, the market risk premium associated with w_t is $-\beta(\gamma_0 + \sum_{k=1}^{K} \gamma_k x_{k,t})$. Therefore, the tradability premium is a stochastic process. Figure 1 shows two samples. As shown in the figure, the tradability premium is significantly different from zero. On average, the tradability premium is 40 percent of the market price of risk in the unrestricted model. This implies that the tradability premium is very important in determining the prices of the derivatives.

To demonstrate the impact of the tradability premium on derivative pricing, we also simulated the daily prices of call options with various times to maturity (i.e., 15 days, 60 days, 100 days, 160 days, 240 days, and 320 days) and different levels of moneyness (i.e., 0.85, 0.92, 0.96, 0.99, 1.02, 1.05). Table 2 presents the out-of-sample option-pricing errors. Here, the pricing errors are measured as the absolute difference between the model-implied prices using the estimated parameters and the true prices as a percentage of the true prices. The options are divided into 36 groups based on maturity and moneyness. For each group, we compute the daily pricing errors with parameters estimated in the unrestricted model and in the restricted model when the tradability restrictions in Equation (2.3) are imposed. We report the mean and standard deviation of daily average pricing errors across 50 simulations. As shown in the table, the out-of-sample pricing errors with parameters estimated in the unrestricted model are very small and most of the them are not significantly different from zero. This is not surprising because the parameter estimates from the unrestricted model are close to the true values. On the other hand, the out-of-sample pricing errors with parameters estimated in the restricted model are quite large and are all significantly different from zero.

The largest pricing error (55.95 percent) occurs for options with a time to maturity of 100 days and moneyness of 0.85. The smallest pricing error (0.38 percent) occurs for options with a time to maturity of 160 days and moneyness of 0.96. Roughly speaking, the pricing errors tend to be larger for call options with low levels of moneyess (out-of-the-money) and long maturities.

The Monte Carlo simulation illustrates the importance of tradability in derivative pricing. In practice, a failure to recognize the non-tradability of the underlying asset and the tradability premium can cause significant pricing errors.

5 Estimation Method and Testing Procedure

In this section, we empirically check whether the tradability restrictions in Equation (2.3) are satisfied by the S&P 500 Index and its derivatives. If they are satisfied, then the market prices the derivatives written on the index as if the index were self-financed and tradable. To test the restrictions, we need an estimator of $z_k(1)$ for all $k = 0, \cdots, K$. This can be obtained by combining the spot, futures, and options data[3]. The pricing model in Section 1 can be estimated using the Unscented Kalman Filter (UKF) method found in Wan and Van Der Merwe (2000).

Equation (1.3) shows that we cannot identify ς_k, ν_k^2, and z separately. Only $\frac{\varsigma_k}{\nu_k^2}$ and $z\nu_k^2$ can be identified. Equation (1.8) also shows that we can only identify $\mu_k\nu_k^2$, $\gamma_k\nu_k^2$, $\alpha_k\nu_k^2$ from the pricing formulae. Therefore, in line with Dai and Singleton (2000), we normalize the model by setting $\nu_k = 1$ for $k = 1, \cdots, K$.

As the underlying factors $\{x_{k,t}\}$ are not observable, we transform the model into a dynamic state-space form and estimate it using a filtering method. The state equations and measurement equations are specified as follows.

Equation (1.2) is discretized daily (1/252 years) to generate the state equation for all factors. We consider one, two, and three factors, respectively. We use daily data from the spot index, two index futures, and three index options to estimate the model. Therefore, we have six measurement equations. Equation (1.1) is discretized daily to generate the first measurement equation for the spot index. The annualized log of futures spot ratio (ALFSR) is defined as[4] $\frac{1}{h}\log\frac{F_{t,t+h}}{I_t}$, where $F_{t,t+h}$ is given by Equation (1.17). The measurement equations for the futures are generated by adding an error term to the ALFSR. The call-options price in Equation (1.19) is normalized by dividing the corresponding Black-Scholes *vega*[5], which we denote as the option price-vega ratio (OPVR). The measurement equations for the options are generated by adding an error term to the OPVR. All of the error terms for the futures and options are assumed to be independently and normally distributed with a mean of 0 and a variance of σ_i^2, $i = 1, \cdots, 5$. The parameters to be estimated are μ_k, γ_k, α_k, $k = 0, \cdots, K$, ξ_k, ζ_k, $k = 1, \cdots, K$, β and σ_i^2, $i = 1, \cdots, 5$. We use θ to denote the vector of all of the parameters. The total number of parameters is $5K + 9$, where K is the number of latent factors.

The above state-space model is Gaussian but nonlinear. We therefore implement the Unscented Kalman Filter (UKF) method found in Wan and Van Der Merwe (2000), which linearizes the model and removes the need to explicitly calculate Jacobians or Hessians without sacrificing accuracy. As all of the error terms are assumed to be normally distributed in the state-space model, the log likelihood function of the observed variables $y_t = [d\log I_t, ALFSR_1, ALFSR_2, OPVR_1, OPVR_2, OPVR_3]'$ is given by:

$$\log L_t(\theta) = -\frac{6}{2}ln2\pi - \frac{1}{2}ln|P_{y_t}^-| - \frac{1}{2}(y_t - \hat{y}_t^-)'(P_{y_t}^-)^{-1}(y_t - \hat{y}_t^-), \tag{5.1}$$

where $\hat{y}_t^-$ is the predicted value of y_t based on earlier observations and $P_{y_t}^-$ is the predicted covariance matrix. The maximum likelihood estimator is obtained as:

$$\hat{\theta} = argmax_\theta \sum_{t=1}^{T} \log L_t(\theta), \tag{5.2}$$

where T is the number of days in the dataset. Under standard regularity conditions, these estimators are asymptotically normal.

The "sandwich" formula is used to estimate the covariance of the quasi-maximum likelihood estimator (see White (1982) and Gourieroux, Monfort, and Trognon (1984)) in order to take account of the possibility that the model is misspecified:

$$\hat{\Sigma}_{\hat{\theta}} = \frac{1}{T} I(\hat{\theta})^{-1} J(\hat{\theta}) I(\hat{\theta})^{-1} \tag{5.3}$$

where

$$I(\hat{\theta}) = -\frac{1}{T}\sum_{t=1}^{T} \frac{\partial^2 \log L_t(\hat{\theta})}{\partial\theta\partial\theta'} \quad and \quad J(\hat{\theta}) = \frac{1}{T}\sum_{t=1}^{T} \frac{\partial \log L_t(\hat{\theta})}{\partial\theta} \frac{\partial\ log L_t(\hat{\theta})}{\partial\theta'}.$$

We deduce that the estimated functions of the parameters characterizing the restrictions for index tradability

$$\widehat{z}_k(1) = \widehat{\alpha}_k + (\widehat{\beta}+1)\widehat{\mu}_k + \frac{\widehat{\gamma}_k}{2}(\widehat{\beta}+1)^2, \quad \forall\, k = 0, \cdots, K.$$

are also asymptotically normal with an estimated variance-covariance matrix $\widehat{\mathbf{\Omega}}$, which is computed using the δ method.

A Wald test statistic of the null hypothesis that the derivatives are priced as if the index were self-financed and tradable is:

$$\xi_{\mathbf{w}} = \widehat{z}(1)'\widehat{\mathbf{\Omega}}^{\mathbf{-1}}\widehat{z}(1). \tag{5.4}$$

Under the null hypothesis, this statistic asymptotically follows a chi-square distribution with $K+1$ degrees of freedom.

6 Empirical Results

In this section, we present the empirical results. Subsection 6.1 introduces the data. The model is estimated using the quasi-maximum likelihood estimation method discussed in Section 5. We analyze the estimation and testing results for the unrestricted model in Subsection 6.2, and those for the restricted model in Subsection 6.3. Subsection 6.4 presents the estimated tradability premium, while Subsection 6.5 investigates the robustness of the empirical results.

6.1 Description of Data

In order to identify all of the parameters in the model, we combine the data for the S&P 500 index, index futures, and index options in the estimation. These data were obtained from Optionmetrics. The spot index and the futures are only used for estimation, while the options data are divided into two parts: one for estimation and the other for the out-of-sample test.

The dataset covers the period from January 3, 2001, to December 29, 2006. There are a total of 1,506 days. The summary statistics are presented in Table 3 and the plots of the in-sample observations are shown in Figure 2.

The index level is computed using the last transaction prices of the component stocks. As seen in Table 3 and Figure 2, the average daily change in the S&P 500 Index is very small, while its standard deviation is relatively larger. The variable $dlogI_t$ is positively skewed and has a fatter tail than the normal distribution.

The S&P 500 futures contracts traded on the Chicago Mercantile Exchange (CME) are among the most actively traded financial derivatives in the world. Each day, there are eight futures contracts with different maturity dates. The maturity dates are the third Friday of the eight months in the following March's quarterly cycle (March, June, September, and December). The futures contracts are ranked by their maturities, and we selected two futures for estimation each day. The first futures (Fu1) have the shortest maturity. However, in the March cycle months and before the maturity date, the futures with the second-shortest maturity are used. The second futures (Fu2) expire a quarter later than Fu1. Fu1 has times to maturity ranging from 15 to 112 days, while the maturities of Fu2 range from 105 to 204 days. These two futures usually have the highest trading volumes with open interest[6] greater than 1,700 contracts. The futures prices are quoted in terms of index points, and the contract size is 250 times CME S&P 500 futures price. The prices vary considerably during our sample period. The ALFSR, which is equal to $\frac{1}{h}\log\frac{F_{t,t+h}}{I_t}$, is more stable as shown in Figure 2. The ALFSR for the first futures has a mean of 0.009 and standard deviation of

0.018, while the ALFSR for the second futures has a mean of 0.011 and standard deviation of 0.015.

The S&P 500 Index options traded on the Chicago Board Options Exchange (CBOE) are European options. They are among the most liquid exchange-traded options and are extensively used for testing option-pricing models. The exchange-traded S&P 500 Index options differ from over-the-counter options and have deterministic issuing dates, maturity dates, and strikes to enhance liquidity. The expiration months are the three near-term months followed by three additional months from the March quarterly cycle, plus two additional months from June and December. The expiration date is the Saturday following the third Friday of the expiration month. The underlying asset is the index level multiplied by 100. Strike price intervals are 5 points and 25 points for long-term contracts.

In this paper, the options data are filtered as follows. First, only call options are included. Second, to alleviate the liquidity concern, we only consider call options with open interest greater than 100, trading volume greater than 0, maturity between 7 and 540 days, and moneyness (defined as the underlying price divided by the strike price, i.e., I/X) between 0.85 and 1.06. Third, to mitigate the market-microstructure problem, we eliminate options with best bid prices of less than 3/8 dollars. The filtered dataset contains 1,506 days, 77,224 options, and an average of 51 observations per day. A similar filtering approach is used in Li (2012). For each day, we select three call options for estimation and the rest are used for the out-of-sample test. We try to use a variety of options with distinct moneyness and times to maturity in the estimation. We categorize three sets of options. Options in the first set have times to maturity of less than 60 days and moneyness between 0.97 and 1.03. These at-the-money, short time-to-maturity (ATM-SM) options are among the most liquid products in the market. Options with medium and long maturity are generally more liquid when they are out-of-the-money. Therefore, options in the second set (OTM-MM) have times to maturity between 60 to 180 days and moneyness of less than 1, and options in the third set (OTM-LM) has times to maturity of more than 180 days and moneyness of less than 1. For estimation purposes, we choose the option with the highest trading volume each day from each set.

Figure 2 shows that the price-vega ratio for the options is high when the underlying index is volatile and low when the index is relatively stable. Options used for the out-of-sample test have a wide range of times to maturity and moneyness, as shown in Table 3. The time to maturity varies from 10 days to 540 days, while the moneyness ranges from 0.85 to 1.06.

6.2 Estimation Results for the Unrestricted Model

Table 4 summarizes the estimation and testing results for the one-, two-, and three-factor models. The table presents the parameter estimates and the standard deviations. The maximized log likelihood for each model and the test statistics $\xi_{\mathbf{w}}$ for the null hypothesis of a tradable index computed in equation (5.4) are presented at the end of the table.

In Table 4, we see that the standard errors of the parameter estimates in the one-factor and two-factor models are high. As a result, the estimates for most parameters in θ are not statistically significant. The positive estimates for ξ_1 and ξ_2 imply that the underlying factors are mean reverting and not very persistent. The high Wald test statistic, $\xi_{\mathbf{w}}$, suggests that the null hypothesis that the derivatives are priced as if the index were self-financed and tradable is strongly rejected in the one-factor and two-factor models.

For the three-factor model, the estimates for most parameters become statistically significant, which means that a third CIR factor process is required. The estimates of 2.23, 2.23, and 1.72 for ξ_1, ξ_2, and ξ_3, respectively, imply that the underlying factors are mean reverting with mean half lives of 0.31, 0.31, and 0.40 years, respectively. All of the factors in this model have higher autocorrelations than in the one-factor and two-factor models, although the third factor has a longer run effect than the first two factors. This suggests that the three-factor model will do better in forecasting the price of derivatives with long maturities, which is indeed the case when we look at derivative-pricing errors. The market price of risk associated with w_t is $-\beta(\gamma_0 + \sum_{k=1}^{K} \gamma_k x_{k,t})^{1/2}$. All γ estimates are positive, and the estimate of -5.66 for β implies a positive market risk premium. The Wald test statistic, $\xi_{\mathbf{w}}$, is equal to 18,232. Therefore, the null hypothesis that the derivatives are priced as if the index were self-financed and tradable is strongly rejected in the three-factor model despite the rather low forecasting error on I_t. In other words, significant mispricing of options and futures would appear if the three-factor model were estimated with the parameters constrained by the tradable restrictions (2.3), as is usually done in the literature.

The likelihood-ratio test of the one-factor model against the two-factor model using the estimated log likelihood rejects the one-factor model. Similarly, the two-factor model is rejected against the three-factor model.

Table 5 reports the in-sample pricing (forecasting) errors for the three models. The pricing error is measured as the absolute difference between the model-implied price and the observed price as a percentage of the observed price. For the index that is not tradable, the interpretation is in terms of forecasting error. For the index, we report the log index value, while the actual prices are examined for the futures and options. There are 1,506 days in the data. The in-sample pricing (forecasting) errors for spot index and index futures in all three models are very low due to the linearity of the pricing functions. In general, the pricing errors are higher for the options, although all of the mean values are less than 10%.

Table 6 reports the out-of-sample option-pricing errors for the three models. There are 1,506 days and 72,706 out-of-sample observations in total. The options are divided into 15 groups based on maturity and moneyness. The means of the absolute percentage of option-pricing errors for each group are reported. As shown in the table, all three models do poorly in estimating the deep out-of-the-money options, especially those with short-term maturities. However, the models produce more accurate estimates for the other options, and the pricing errors for these options are fairly low. The two- and three-factor models perform better than the one-factor model in forecasting the prices of the medium- and long-term derivatives. The three-factor model yields the smallest mean absolute percentage pricing errors for options with long maturities, and the average pricing error is around 5 percent.

6.3 Estimation Results for the Restricted Model

In Subsection 6.2, we only looked at the unrestricted versions of the factor models and rejected the tradability restriction in Equation (2.3). To check the economic significance of our empirical results, we now present the estimation and test results for the restricted model in which the underlying asset is assumed to be self-financed and tradable. In other words, we now impose the tradability restrictions in Equation (2.3).

The last row of Table 4 reports the log likelihoods of the restricted models. They are all clearly rejected against the unrestricted models.

Tables 7 and 8 report the in-sample and out-of-sample pricing (forecasting) errors for the restricted versions of the three models. Due to the linearity of the pricing functions, the in-sample pricing (forecasting) errors for spot index and index futures are very low and not much different from the pricing errors of the unrestricted model. However, for the options, the restricted model yields much higher pricing errors than the unrestricted model. For the in-sample pricing-error estimates, the difference averages 4 percent. For the out-of-sample forecast, the difference is even greater. These comparisons are confirmed in Tables 9 and 10, in which we report the absolute difference of the in-sample and out-of-sample pricing errors between the unrestricted and the restricted models, respectively. In comparing the sample mean of the absolute difference with the standard error of the sample mean, we see that the restricted model generates significantly worse results than the unrestricted model. Therefore, the tradability of the underlying asset must be taken into account in derivative pricing.

6.4 Estimation of the Tradability Premium

Given the estimates of the parameters and the filtered factors in each factor model, we can look at the market risk premium estimated from the unrestricted model and from the restricted model, and estimate the tradability premium for each day.

The daily tradability premium is plotted in Figure 3. The sample mean of the premium and the standard error of the sample mean are also reported for each model. For the two-factor model, the tradability premium is approximately zero for some days, while it can be positive or negative for other days. For the one-factor and three-factor models, the tradability premium is quite different from zero and always remains positive during the sample period. The sample means of the daily premium are 130 basis points and 831 basis points, respectively. Therefore, the effect of tradability on derivative pricing cannot be ignored. Moreover, the tradability premium tends to be high and volatile when the daily change in the underlying index is high and volatile. This suggests that when the financial market is volatile, it is even more imperative for investors to have the tradable underlying asset to hedge the risk associated with it. In summary, the tradability of the underlying asset is important for derivative pricing.

6.5 Robustness Analysis and Caveats

Our empirical results strongly reject the tradability restrictions in Equation (2.3). However, as this is not a model-free test, we need to address the issue of possible model misspecification.

In Equation (1.5), we propose a simple specification of SDF, which implicitly assumes that the market prices of the risk factors $\{w_{k,t}\}$, $k = 1, \cdots, K$ are 0, and only the risk associated with w_t is corrected. We intend to use this parsimonious specification to emphasize the point that, conditional on the underlying factors, whether the index is tradable or not will solely affect the market risk premium associated with w_t. If the index is tradable, the market risk premium is determined by the price of the underlying asset and the short rate. If the index is not tradable, the value of the index does not provide enough information on the market risk premium. To check whether a more flexible specification of SDF will change our main conclusion, we consider a standard specification of SDF used in affine models (see, e.g., Duffie, Pan, and Singleton (2000)). In this specification, the SDF is assumed to be an exponential affine function of all of the risk factors:

$$M_{t,t+dt} = \exp(dm_t) = \exp(\alpha_0 dt + \sum_{k=1}^{K} \alpha_k dx_{k,t} + \beta d \log I_t).$$

With this more general specification, it is not surprising that the models produce lower pricing errors. However, we still strongly reject the null hypothesis that the derivatives are priced as if the index were self-financed and tradable in all three factor models. Therefore, we keep the simple model for ease of illustration.

The derivative-pricing literature recognizes that we need a distribution with a fatter tail and more negative skewness than the Black-Scholes model to more accurately calibrate option prices. A popular method of generating a distribution with these properties is to introduce stochastic volatility to account for the fat tail in the long run, a negative correlation between stochastic volatility and asset return to account for the negative skewness ("leverage effect"), and a negative or asymmetric jump (with stochastic jump intensity) to generate high kurtosis and negative asymmetry in the short run. The model in this paper only introduces stochastic volatility, while it lacks the other necessary features. Ideally, we should extend the model to cover all requisite complications, compute derivative prices, estimate all of the dynamics jointly, and derive the corresponding tradability constraints. However, this task is beyond the current computing capabilities. Moreover, a complicated model would deny the possibility of deriving simple analytical tradability restrictions, such as those in Equation (2.3), which can be tested empirically. We choose the simple model specification to demonstrate the importance of tradability in derivatives pricing. The possible misspecification of our simple model could explain some of the option-pricing errors presented above, especially the large out-of-sample pricing errors for the short-term out-of-the-money options. Therefore, the empirical results in this section should be approached with some caution. In addition, the tradability premium plotted in Figure 3 could partially reflect the model misspecification. However, as we have seen in the Monte Carlo simulation in Section 4, the tradability premium is not simply a result of model misspecification. It is an important factor that should be taken into account in derivative pricing.

7 Conclusion

In this paper, we consider a coherent multi-factor affine model for pricing various derivatives, such as forwards, futures, and European options, written on the non-tradable S&P 500 Index.

We consider cases in which the underlying index is self-financed and tradable, and cases in which it is not, and we show the difference between the two pricing models. When the underlying asset is self-financed and tradable, an additional arbitrage condition must be introduced, which implies additional parameter restrictions. These restrictions can be tested in practice to check whether the derivatives are priced as if the underlying index were self-financed and tradable.

More importantly, we are able to define and compute the "tradability premium" in this framework and show the impact of the tradability of the underlying asset on derivative pricing. In the Monte Carlo simulation study, we illustrate that whether the underlying asset is tradable makes a nontrivial difference and that ignoring the tradability premium could cause significant mispricing of the derivatives.

To empirically test the restrictions, we consider three nested factor models. The models are estimated by combining the spot, futures, and options data, and using the Unscented Kalman Filter (UKF) method. The Wald tests strongly reject the null hypothesis that the derivatives are priced as if the index were self-financed and tradable in all three models. The robustness test shows that the unrestricted model performs significantly better than the restricted model in which the underlying asset is constrained to be self-financed and tradable.

The daily tradability premium of the S&P 500 Index is clearly different from zero during our sample period. In other words, a tradability premium exists in the price of the tradable asset and the tradability of the index is an important factor in derivative pricing. Moreover, significant mispricing of options and futures contracts would be observed if the factor models were estimated with parameters constrained by the index tradability restriction, as is usually done in the literature. In addition, it is impossible to reproduce the index by means of a self-financed mimicking portfolio without significant errors.

The S&P 500 Index is not the only non-tradable index on which derivatives are written. Our model can easily be extended to price derivatives written on other non-tradable indices, such as a retail price index, a meteorological index, an index summarizing the results of a set of insurance companies, a population mortality index, or the VIX. These other applications are even more appealling as no liquid mimicking portfolio for these indices is generally proposed on the market.

Appendices: Proofs of Propositions

A Proof of Proposition 1

The price of the call option is:

$$
\begin{aligned}
&C(t,t+h,u)\\
=&E_t[\exp(\int_t^{t+h} dm_\tau + u\log I_{t+h})]\\
=&E_t\{\exp[\int_t^{t+h} dm_\tau + u(\log I_t + \int_t^{t+h} d\log I_\tau)]\}\\
=&\exp(u\log I_t)E_t\{\exp\int_t^{t+h}([\alpha_0+\beta\mu_0+\sum_{k=1}^{K}(\alpha_k+\beta\mu_k)x_{k,\tau}]d\tau\\
&\quad+\beta(\gamma_0+\sum_{k=1}^{K}\gamma_k x_{k,\tau})^{1/2}dw_\tau + u[(\mu_0+\sum_{k=1}^{K}\mu_k x_{k,\tau})d\tau+(\gamma_0+\sum_{k=1}^{K}\gamma_k x_{k,\tau})^{1/2}dw_\tau])\}\\
=&\exp(u\log I_t)E_t\{\exp\int_t^{t+h}([\alpha_0+(\beta+u)\mu_0+\sum_{k=1}^{K}(\alpha_k+(\beta+u)\mu_k)x_{k,\tau}]d\tau\\
&\quad+(\beta+u)(\gamma_0+\sum_{k=1}^{K}\gamma_k x_{k,\tau})^{1/2}dw_\tau)\}\\
=&\exp(u\log I_t)E_t\{\exp\int_t^{t+h}[\alpha_0+(\beta+u)\mu_0+\sum_{k=1}^{K}(\alpha_k+(\beta+u)\mu_k)x_{k,\tau}]d\tau\\
&\quad\times E_t(\exp[(\beta+u)\int_t^{t+h}(\gamma_0+\sum_{k=1}^{K}\gamma_k x_{k,\tau})^{1/2}dw_\tau]\mid \mathbb{X}_{k,\tau})\},
\end{aligned}
$$

where $\mathbb{X}_{k,\tau}$ denotes the set $\{x_{k,\tau}\}_{\tau=t\cdots t+h}^{k=1\cdots K}$.

As $\exp[(\beta+u)\int_t^{t+h}(\gamma_0+\sum_{k=1}^{K}\gamma_k x_{k,\tau})^{1/2}dw_\tau]\mid \mathbb{X}_{k,\tau}$

$$
\sim LN(0,(\beta+u)^2\int_t^{t+h}(\gamma_0+\sum_{k=1}^{K}\gamma_k x_{k,\tau})d\tau),
$$

we deduce that:

$$
\begin{aligned}
&C(t,t+h,u)\\
=&\exp(u\log I_t)E_t\{\exp\int_t^{t+h}[\alpha_0+(\beta+u)\mu_0+\sum_{k=1}^{K}(\alpha_k+(\beta+u)\mu_k)x_{k,\tau}]d\tau
\end{aligned}
$$

$$\times \exp[\frac{(\beta+u)^2}{2}\int_t^{t+h}(\gamma_0+\sum_{k=1}^K \gamma_k x_{k,\tau})d\tau]\}$$

$$= \exp(u\log I_t)\exp\int_t^{t+h}[\alpha_0+(\beta+u)\mu_0+\frac{(\beta+u)^2}{2}\gamma_0]d\tau$$

$$\times E_t\{\exp\int_t^{t+h}\sum_{k=1}^K[(\alpha_k+(\beta+u)\mu_k+\frac{(\beta+u)^2}{2}\gamma_k)x_{k,\tau}]d\tau\}$$

$$= \exp(u\log I_t)\exp\{h[\alpha_0+(\beta+u)\mu_0+\frac{(\beta+u)^2}{2}\gamma_0]\}$$

$$\times \prod_{k=1}^K E_t\{\exp-\int_t^{t+h}-[\alpha_k+(\beta+u)\mu_k+\frac{(\beta+u)^2}{2}\gamma_k]x_{k,\tau}d\tau\},$$

as factors $\{x_{k,t}\},\ k=1,\cdots,K$are independent,

$$= \exp(u\log I_t)\exp\{h[\alpha_0+(\beta+u)\mu_0+\frac{\gamma_0}{2}(\beta+u)^2]$$

$$-\sum_{k=1}^K H_1^k(h,z_k(u))x_{k,t}-\sum_{k=1}^K H_2^k(h,z_k(u))\},$$

where

$$z_k(u) = -\alpha_k-(\beta+u)\mu_k-\frac{\gamma_k}{2}(\beta+u)^2,$$

and $H_1^k(\cdot,\cdot)$ and $H_2^k(\cdot,\cdot)$ are given in equation (1.4)

B Proof of Proposition 3

The instantaneous interest rate is defined by:

$$r(t) = \lim_{h\to 0} -\frac{1}{h}\log B(t,t+h) = \frac{d[-\log B(t,t+h)]}{dh}\ |_{h=0}\ .$$

We have:

$$-\log B(t,t+h) = -h(\alpha_0+\beta\mu_0+\frac{\gamma_0}{2}\beta^2)+\sum_{k=1}^K H_1^k(h,z_k(0))x_{k,t}+\sum_{k=1}^K H_2^k(h,z_k(0)).$$

We deduce that:

$$\frac{d[-\log B(t,t+h)]}{dh} = -\alpha_0-\beta\mu_0-\frac{\gamma_0}{2}\beta^2+\sum_{k=1}^K\frac{dH_1^k(h,z_k(0))}{dh}x_{k,t}+\sum_{k=1}^K\frac{dH_2^k(h,z_k(0))}{dh},$$

where

$$
\begin{aligned}
H_1^k(h, z_k(0)) &= \frac{2z_k(0)(\exp[\varepsilon_k(z_k(0))h] - 1)}{(\varepsilon_k(z_k(0)) + \xi_k)(\exp[\varepsilon_k(z_k(0))h] - 1) + 2\varepsilon_k(z_k(0))}, \\
H_2^k(h, z_k(0)) &= \frac{-2\xi_k\zeta_k}{\nu_k^2}\{\log[2\varepsilon_k(z_k(0))] + \frac{h}{2}[\varepsilon_k(z_k(0)) + \xi_k] \\
&\qquad - \log[(\varepsilon_k(z_k(0)) + \xi_k)(\exp[\varepsilon_k(z_k(0))h] - 1) + 2\varepsilon_k(z_k(0))]\}.
\end{aligned}
$$

It is straightforward to show that:

$$
\frac{dH_1^k(h, z_k(0))}{dh} |_{h=0} = z_k(0), \text{and } \frac{dH_2^k(h, z_k(0))}{dh} |_{h=0} = 0.
$$

We deduce:

$$
r(t) = -\alpha_0 - \beta\mu_0 - \frac{\gamma_0}{2}\beta^2 + \sum_{k=1}^{K} z_k(0)x_{k,t}.
$$

C Proof of Proposition 4

As

$$
E[\exp(\int_t^{t+h} dm_\tau)(f(t, t+h) - I_{t+h})] = 0,
$$

we get:

$$
B(t, t+h)f(t, t+h) = E[\exp(\int_t^{t+h} dm_\tau)I_{t+h}],
$$

and

$$
f(t, t+h) = \frac{C(t, t+h, 1)}{C(t, t+h, 0)}.
$$

D Proof of Proposition 5

As $E_t[\int_t^{t+h}(\exp\int_t^{t+\tau} dm_s)dF_\tau] = 0,$ we get:

$$
\begin{aligned}
&F_{t,t+h} \\
=&E_t[\exp(\int_t^{t+h} dm_\tau)\exp(\int_t^{t+h} r_\tau d\tau)I_{t+h}] \\
=&I_t E_t[\exp(\int_t^{t+h}(dm_\tau + r_\tau d\tau + d\log I_\tau)]
\end{aligned}
$$

$$
\begin{aligned}
=& I_t E_t\{\exp \int_t^{t+h} [(\alpha_0 + \beta\mu_0 + \sum_{k=1}^{K} (\alpha_k + \beta\mu_k) x_{k,\tau}) d\tau \\
&\quad + \beta(\gamma_0 + \sum_{k=1}^{K} \gamma_k x_{k,\tau})^{1/2} dw_\tau + (\mu_0 + \sum_{k=1}^{K} \mu_k x_{k,\tau}) d\tau \\
&\quad + (\gamma_0 + \sum_{k=1}^{K} \gamma_k x_{k,\tau})^{1/2} dw_\tau + (-\alpha_0 - \beta\mu_0 - \frac{\gamma_0}{2}\beta^2 + \sum_{k=1}^{K} z_k(0) x_{k,\tau}) d\tau]\} \\
=& I_t \exp[h(\mu_0 - \frac{\gamma_0}{2}\beta^2)] E_t\{\exp \int_t^{t+h} \sum_{k=1}^{K} (\alpha_k + (\beta+1)\mu_k + z_k(0)) x_{k,\tau} d\tau \\
&\quad \times E_t[\exp \int_t^{t+h} (\beta+1)(\gamma_0 + \sum_{k=1}^{K} \gamma_k x_{k,\tau})^{1/2} dw_\tau \mid \mathbb{X}_{k,\tau}]\}.
\end{aligned}
$$

As $\exp \int_t^{t+h} (\beta+1)(\gamma_0 + \sum_{k=1}^{K} \gamma_k x_{k,\tau})^{1/2} dw_\tau \mid \mathbb{X}_{k,\tau}$

$$
\sim LN(0, (\beta+1)^2 \int_t^{t+h} (\gamma_0 + \sum_{k=1}^{K} \gamma_k x_{k,\tau}) d\tau),
$$

we get:

$$
\begin{aligned}
& F_{t,t+h} \\
=& I_t \exp[h(\mu_0 - \frac{\gamma_0}{2}\beta^2)] \\
&\quad \times E_t\{\exp \int_t^{t+h} [\sum_{k=1}^{K} (\alpha_k + (\beta+1)\mu_k + z_k(0)) x_{k,\tau} d\tau + \frac{(\beta+1)^2}{2} (\gamma_0 + \sum_{k=1}^{K} \gamma_k x_{k,\tau}) d\tau]\} \\
=& I_t \exp[h(\mu_0 - \frac{\gamma_0}{2}\beta^2 + \frac{(\beta+1)^2}{2}\gamma_0)] \\
&\quad \times \prod_{k=1}^{K} E_t\{\exp - \int_t^{t+h} -[\alpha_k + (\beta+1)\mu_k + z_k(0) + \frac{(\beta+1)^2}{2}\gamma_k] x_{k,\tau} d\tau\} \\
=& I_t \exp[h(\mu_0 + \frac{1+2\beta}{2}\gamma_0) - \sum_{k=1}^{K} H_1^k(h, l_k) x_{k,t} - \sum_{k=1}^{K} H_2^k(h, l_k)],
\end{aligned}
$$

where:

$$
\begin{aligned}
l_k &= -\alpha_k - (\beta+1)\mu_k - z_k(0) - \frac{\gamma_k}{2}(\beta+1)^2 \\
&= -\mu_k - \frac{1+2\beta}{2}\gamma_k.
\end{aligned}
$$

E Proof of Proposition 6

Let us first consider the call option with price $G(t, t+h, X)$. Its price is given by:

$$
\begin{aligned}
& G(t, t+h, X) \\
&= E_t\{\exp(\int_t^{t+h} dm_\tau)[\exp(\log I_{t+h}) - X]^+\} \\
&= E_t\{\exp(\int_t^{t+h} dm_\tau)[\exp(\log I_{t+h}) - X]\mathbf{1}_{-\log I_{t+h} \leq -\log X}\} \\
&= A_{1,-1}(-\log X; x_{1,t}, \cdots, x_{K,t}, \log I_t, h) - X A_{0,-1}(-\log X; x_{1,t}, \cdots, x_{K,t}, \log I_t, h),
\end{aligned}
$$

where $A_{a,b}(y; x_{1,t}, \cdots, x_{K,t}, \log I_t, h) = E_t[\exp(\int_t^{t+h} dm_\tau) \exp(a \log I_{t+h}) 1_{b \log I_{t+h} \leq y}]$.

The Fourier-Stieltjes transform of $A_{a,b}(y; x_{1,t}, \cdots, x_{K,t}, \log I_t, h)$ is:

$$
\begin{aligned}
& \int_{\Re} \exp(ivy) dA_{a,b}(y; x_{1,t}, \cdots, x_{K,t}, \log I_t, h) \\
&= E_t\{\exp(\int_t^{t+h} dm_\tau) \exp[(a + ivb) \log I_{t+h}]\} = C(t, t+h, a + ivb).
\end{aligned}
$$

We deduce that:

$$
\begin{aligned}
& A_{a,b}(y; x_{1,t}, \cdots, x_{K,t}, \log I_t, h) \\
&= \frac{C(t, t+h, a)}{2} - \frac{1}{\pi} \int_0^\infty \frac{Im[C(t, t+h, a + ivb) \exp(-ivy)]}{v} dv
\end{aligned}
$$

[see Duffie, Pan, and Singleton (2000), p.1352].

By substitution, we get the call price:

$$
\begin{aligned}
G(t, t+h, X) = & \frac{C(t, t+h, 1)}{2} - \frac{1}{\pi} \int_0^\infty \frac{Im[C(t, t+h, 1 - iv) \exp(iv \log X)]}{v} dv \\
& - X\{\frac{C(t, t+h, 0)}{2} - \frac{1}{\pi} \int_0^\infty \frac{Im[C(t, t+h, -iv) \exp(iv \log X)]}{v} dv\}.
\end{aligned}
$$

Similarly, for the put option with price $H(t, t+h, X)$, we have:

$$
\begin{aligned}
& H(t, t+h, X) \\
&= E_t\{\exp(\int_t^{t+h} dm_\tau)[X - \exp(\log I_{t+h})]^+\}
\end{aligned}
$$

$$
\begin{aligned}
&= E_t\{\exp(\int_t^{t+h} dm_\tau)[X - \exp(\log I_{t+h})]\mathbf{1}_{\log I_{t+h} \le \log X}\} \\
&= -A_{1,1}(\log X; x_{1,t}, \cdots, x_{K,t}, \log I_t, h) + XA_{0,1}(\log X; x_{1,t}, \cdots, x_{K,t}, \log I_t, h) \\
&= -\frac{C(t, t+h, 1)}{2} + \frac{1}{\pi}\int_0^\infty \frac{Im[C(t, t+h, 1+iv)\exp(-iv\log X)]}{v} dv \\
&\quad + X\{\frac{C(t, t+h, 0)}{2} - \frac{1}{\pi}\int_0^\infty \frac{Im[C(t, t+h, iv)\exp(-iv\log X)]}{v} dv\}.
\end{aligned}
$$

F Proof of Proposition 7

The first restriction in Equation (2.2) holds if, and only if:

$$
z_k(1) = -\alpha_k - (\beta+1)\mu_k - \frac{\gamma_k}{2}(\beta+1)^2 = 0, \ \forall k = 1, \cdots, K.
$$

This implies $\varepsilon_k(z_k(1)) = |\xi_k|$, $\forall k = 1, \cdots, K$, and

$$
\begin{aligned}
H_2^k(h, z_k(1)) &= \frac{-2\xi_k\zeta_k}{\nu_k^2}\{\log|2\xi_k| + \frac{h}{2}(|\xi_k| + \xi_k) - \log[(|\xi_k| + \xi_k)(\exp(|\xi_k|h) - 1) + 2|\xi_k|]\} \\
&= 0, \ \text{ regardless of whether } \ \xi_k > 0 \ \text{ or } \ \xi_k < 0, \ \forall k = 1, \cdots, K.
\end{aligned}
$$

This, together with the second restriction in Equation (2.2), implies that:

$$
\alpha_0 + (\beta+1)\mu_0 + \frac{\gamma_0}{2}(\beta_1)^2 = 0.
$$

Therefore, Equation (2.2) is equivalent to:

$$
\alpha_k + (\beta+1)\mu_k + \frac{\gamma_k}{2}(\beta_1)^2 = 0, \quad \forall k = 0, \cdots, K.
$$

G Figures and Tables

Table 1: Estimation in the Simulation

The first part of this table reports the mean and standard deviation of the parameter estimates across 50 simulations in Section 4. The model is normalized by setting $\nu_1 = 1$. It is estimated using the simulated data of the underlying asset, two futures, and three call options using the maximum likelihood estimation (MLE) method. The first column reports the parameters of the one-factor model and their true values as assumed in the simulation. The second and third columns present the estimation results for the unrestricted model, and the last two columns present the results for the restricted model when the tradability restrictions in Equation (2.3) are imposed. The second part of this table reports the mean and standard deviation of the in-sample pricing errors. The pricing errors are measured as the absolute difference between the model-implied prices using the estimated parameters and the observed prices as a percentage of the observed prices.

	Unrestricted Model		Restricted Model	
Parameter (true value)	Estimate	Std.Dev.	Estimate	Std.Dev.
$\mu_0(0.015)$	0.014975	0.000153	0.014649	0.001045
$\mu_1(-0.02)$	-0.020457	0.002801	-0.022989	0.016267
$\gamma_0(0.005)$	0.004995	0.000057	0.005347	0.000161
$\gamma_1(0.09)$	0.090017	0.000653	0.084409	0.002632
$\xi_1(3)$	2.996987	0.090782	1.230851	0.101356
$\zeta_1(0.1)$	0.100085	0.002416	0.292409	0.026611
$\alpha_0(0.28)$	0.280841	0.009660	-0.010594	0.002842
$\alpha_1(0.12)$	0.114479	0.061099	-0.001384	0.005172
$\beta(-0.42)$	-0.414964	0.030824	-0.347832	0.193869
In-Sample Pricing Errors	Mean	Std.Dev.	Mean	Std.Dev.
$d\log I_t$	0.0008	0.0001	0.0008	0.0001
$Fu_1(\tau = 30\ days)$	0.0001	0.0000	0.0001	0.0000
$Fu_2(\tau = 120\ days)$	0.0004	0.0000	0.0004	0.0000
$Op_1(\tau = 30\ days, S/X = 1.02)$	0.0211	0.0038	0.0266	0.0027
$Op_2(\tau = 90\ days, S/X = 0.97)$	0.0358	0.007	0.0436	0.0061
$Op_3(\tau = 250\ days, S/X = 0.93)$	0.0202	0.0038	0.0206	0.0037

Table 2: Out-of-Sample Option-Pricing Errors in the Simulation

This table reports the out-of-sample option-pricing errors in the Monte Carlo simulation in Section 4. The pricing errors are measured as the absolute difference between the model-implied prices using the estimated parameters and the true prices as a percentage of the true prices. The options are divided into 36 groups based on maturity and moneyness. For each group, we compute the daily pricing errors with parameters estimated in the unrestricted model (Unres) and the restricted model (Res) when the tradability restrictions in Equation (2.3) are imposed. We report the average of daily mean pricing errors across 50 simulations. The corresponding standard deviations are reported in parenthesis.

Maturity	Moneyness(I/K)											
	0.85		0.92		0.96		0.99		1.02		1.05	
	Unres	Res	Unres	Res	Unres	Res	Unres	Res	Unres	Res	Unres	Res
15 days	0.0098	0.1249	0.0078	0.0863	0.003	0.0235	0.0032	0.0278	0.0005	0.0123	0.0004	0.0169
	(0.0061)	(0.0192)	(0.0053)	(0.0269)	(0.0018)	(0.0031)	(0.0021)	(0.013)	(0.0002)	(0.0006)	(0.0002)	(0.0001)
60 days	0.0111	0.496	0.0056	0.2084	0.0026	0.0682	0.0011	0.0043	0.0009	0.035	0.0012	0.054
	(0.0084)	(0.0349)	(0.004)	(0.0153)	(0.0019)	(0.0062)	(0.0007)	(0.0019)	(0.0005)	(0.0012)	(0.0008)	(0.0014)
100 days	0.0108	0.5595	0.0042	0.1783	0.0016	0.0453	0.0008	0.0182	0.0013	0.0585	0.0018	0.0827
	(0.0077)	(0.0233)	(0.003)	(0.0076)	(0.0011)	(0.0018)	(0.0004)	(0.0015)	(0.0009)	(0.0023)	(0.0013)	(0.0022)
160 days	0.009	0.4857	0.0027	0.121	0.0007	0.0038	0.0012	0.0581	0.0021	0.0988	0.0027	0.126
	(0.0067)	(0.0159)	(0.0019)	(0.0029)	(0.0004)	(0.0011)	(0.001)	(0.0025)	(0.0016)	(0.0028)	(0.0021)	(0.0027)
240 days	0.0064	0.3458	0.001	0.0357	0.0015	0.0664	0.0026	0.1195	0.0035	0.1579	0.0041	0.1854
	(0.005)	(0.0113)	(0.0006)	(0.0012)	(0.0012)	(0.0015)	(0.0021)	(0.0023)	(0.0028)	(0.0025)	(0.0032)	(0.0024)
320 days	0.0039	0.2061	0.0014	0.0508	0.0032	0.1376	0.0042	0.1842	0.005	0.2191	0.0056	0.2452
	(0.0031)	(0.0104)	(0.0011)	(0.0022)	(0.0025)	(0.0014)	(0.0034)	(0.0016)	(0.004)	(0.0018)	(0.0045)	(0.0018)

Table 3: Summary Statistics for the Data

This table summarizes statistics for the daily data ranging from January 3, 2001, to December 29, 2006. It covers the spot S&P 500 Index (SPX), the two index futures (Fu1 and Fu2), and the three index options (Op1 to Op3) used for estimation, and the options used for the out-of-sample test. The daily log difference of the S&P 500 Index is reported in the first row. For the futures, statistics for time-to-maturity (T2M), the annualized log of futures spot ratio (ALFSR), and the total number of outstanding futures contracts (open interest) are presented. Fu1 refers to the short-term futures with the shortest maturity and Fu2 refers to the medium-term futures expiring a quarter later than Fu1. For the options, statistics for time-to-maturity (T2M), moneyness, the price-vega ratio, trading volume, and open interest are shown. Op1 represents at-the-money options with short maturity (ATM-SM). Op2 stands for out-of-the-money options with medium maturity (OTM-MM), and Op3 represents out-of-the-money options with long maturity (OTM-LM). The last panel summarizes the out-of-the sample options.

Variable	Mean	Std.Dev.	Skewness	Kurtosis	Min.	Median	Max.
In Sample							
dlogSPX	0.000	0.011	0.161	5.844	-0.050	0.000	0.056
In Sample							
Fu1 T2M (years)	0.172	0.0721	0.0142	1.817	0.041	0.173	0.307
Fu1 ALFSR	0.009	0.018	0.474	3.910	-0.051	0.005	0.105
Fu1 Open Int	520,907	128,143	-1.520	4.905	71,345	564,230	669,216
In Sample							
Fu2 T2M (years)	0.423	0.072	0.017	1.820	0.288	0.425	0.559
Fu2 ALFSR	0.011	0.015	0.481	2.218	-0.022	0.008	0.058

(continued on next page)

Table 3: (continued)

Variable	Mean	Std.Dev.	Skewness	Kurtosis	Min.	Median	Max.
Fu2 Open Int	85,164	140,107	2.440	8.105	1,744	25,647	665,176
In Sample							
Op1 T2M (years)	0.081	0.034	0.398	2.476	0.027	0.082	0.162
Op1 Moneyness	0.994	0.012	0.149	2.677	0.970	0.995	1.030
Op1 Price/Vega	0.160	0.093	1.173	5.109	0.018	0.143	0.665
Op1 Trading vol	7,467	6,353	2.840	15.921	239	5,853	59,702
Op1 Open int	31,716	27,922	1.696	6.587	143	24,269	191,634
In Sample							
Op2 T2M (years)	0.266	0.076	0.675	2.468	0.164	0.244	0.490
Op2 Moneyness	0.959	0.036	-0.812	2.703	0.851	0.968	1.000
Op2 Price/Vega	0.110	0.064	0.811	3.418	0.014	0.104	0.375
Op2 Trading vol	3,963	3,921	3.130	25.210	6	2,700	54,275
Op2 Open int	17,892	16,759	1.546	5.638	100	12,752	94,890
In Sample							
Op3 T2M (years)	0.810	0.235	0.875	3.056	0.501	0.753	1.479
Op3 Moneyness	0.931	0.043	-0.046	1.832	0.850	0.930	1.000
Op3 Price/Vega	0.111	0.056	0.557	2.492	0.017	0.103	0.278
Op3 Trading vol	1,627	1,820	3.556	25.661	1	1,050	21,000
Op3 Open int	11,766	10,091.15	1.535	6.421	100	8,925	75,069

(continued on next page)

Table 3: (continued)

Variable	Mean	Std.Dev.	Skewness	Kurtosis	Min.	Median	Max.
Out of Sample							
Op T2M	0.295	0.303	1.669	5.235	0.027	0.164	1.479
Op Moneyness	0.972	0.047	-0.422	2.614	0.850	0.977	1.060
Op Price/Vega	0.189	0.642	82.578	9,353	0.013	0.109	90.927
Op Trading vol	811	1,977	20.398	1,232	1	178	157,542
Op Open int	12,856	16,113	3.421	22.919	100	7,722	214,048

Table 4: Parameter Estimate

This table reports parameter estimates and standard deviations for three nested factor models. The models are normalized by setting $\nu_k = 1$, for $k = 1, \cdots, K$. The one-, two- and three-factor models derived in the paper are estimated using the daily data for the S&P 500 Index, two index futures, and three index options, as described in Section 6. The log likelihood for each model and test statistics $\xi_{\mathbf{w}}$ computed in equation (5.4) are presented at the end of the table. The last row of the table reports the log likelihood when the tradability restrictions in Equation (2.3) are imposed.

Parameter	One Factor		Two Factors		Three Factors	
	Estimate	Std.Err.	Estimate	Std.Err.	Estimate	Std.Err.
μ_0	0.014156	0.062649	0.001751	0.490664	0.020641	0.000306
μ_1	-0.019612	1.202200	-0.120038	3.543429	0.326886	0.001813
μ_2	-	-	0.501187	0.000786	0.241127	0.011074
μ_3	-	-	-	-	0.238038	0.000913
γ_0	0.005084	0.021024	0.005787	0.060267	0.005612	7.60e-05
γ_1	0.091091	0.226156	0.115763	0.453376	0.006901	0.000180
γ_2	-	-	1.72e-05	1.21e-07	0.003152	5.24e-05
γ_3	-	-	-	-	0.067161	0.004048
ξ_1	3.364535	6.302816	2.667334	11.54964	2.230514	0.006613
ξ_2	-	-	3.409025	0.474828	2.230497	0.032053
ξ_3	-	-	-	-	1.719731	0.079500
ζ_1	0.103510	0.442547	0.052931	0.450232	0.093168	0.000992
ζ_2	-	-	0.078183	0.000923	0.013873	0.000221
ζ_3	-	-	-	-	0.084855	0.002533

(continued on next page)

Table 4: (continued)

Parameter	One Factor		Two Factors		Three Factors	
	Estimate	Std.Err.	Estimate	Std.Err.	Estimate	Std.Err.
α_0	0.279563	0.710415	0.351444	0.043135	0.376432	0.017518
α_1	0.115896	3.850134	0.356801	6.467518	1.500325	0.013313
α_2	-	-	-0.476146	0.362013	1.181081	0.008354
α_3	-	-	-	-	0.059578	0.002488
β	-0.421196	12.37646	-1.860492	0.960758	-5.662848	0.415460
σ_{Fu1}	0.022757	0.001289	0.009532	0.003030	0.009446	0.000655
σ_{Fu2}	0.017257	0.001578	0.000290	0.065585	0.000790	3.36e-05
σ_{Op1}	0.012955	0.006663	0.015889	0.025387	0.016573	0.000553
σ_{Op2}	0.010243	0.008514	0.009056	0.066390	0.010097	0.000281
σ_{Op3}	0.013427	0.017725	0.010930	0.009359	0.008439	0.000746
Unrestricted						
Log likelihood	25,525.29		28,491.26		28,853.48	
$\xi_{\mathbf{w}}$	845.9943		997.2979		18,231.71	
$\chi^2_{99\%}(K+1)$	9.21		11.35		13.28	
Restricted						
Log likelihood	24,838.73		27,431.55		27,791.21	

Table 5: In-Sample Absolute Percentage Pricing (Forecasting) Errors

This table reports the in-sample pricing errors, which are measured as the absolute difference between model-implied price and the observed price as a percentage of the observed price. For the index that is not tradable, the interpretation is in terms of forecasting error. For the index, we report the log index value, while the actual prices are examined for the futures and options. Fu1 refers to the short-term futures, and Fu2 refers to the medium-term futures. Op1 represents at-the-money options with short maturity (ATM-SM). Op2 stands for out-of-the-money options with medium maturity (OTM-MM), and Op3 represents out-of-the-money options with long maturity (OTM-LM). There are 1,506 days in the sample.

Securities	One Factor		Two Factors		Three Factors	
	Mean	Std.Dev.	Mean	Std.Dev.	Mean	Std.Dev.
$\log I_t$	0.001108	0.001086	0.001109	0.001088	0.001112	0.001095
Fu1	0.002400	0.001722	0.001199	0.001045	0.001084	0.001049
Fu2	0.005233	0.003227	4.86e-06	4.70e-06	3.98e-05	4.12e-05
Op1	0.089782	0.164550	0.105236	0.178787	0.107083	0.179544
Op2	0.097314	0.152559	0.082150	0.113209	0.095510	0.128891
Op3	0.104998	0.142086	0.087099	0.109224	0.067479	0.084363

Table 6: Out-of-Sample Absolute Percentage Option-Pricing Errors

This table reports the out-of-sample option-pricing errors, which are measured as the absolute difference between the model-implied price and the observed price as a percentage of the observed price. There are 1,506 days and 72,706 out-of-the-sample observations in total. The options are divided into 15 groups based on maturity and moneyness. The mean of absolute percentage option-pricing errors for each group is reported.

Maturity	Model	Moneyness(I/K)				
		<0.94	0.94-0.97	0.97-1.00	1.00-1.03	>1.03
<60 days	One Factor	0.545435	0.466165	0.154574	0.048615	0.028152
	Two Factors	0.529769	0.466957	0.165987	0.044368	0.025586
	Three Factors	0.519731	0.436678	0.163203	0.045412	0.028491
60-180	One Factor	0.360111	0.122878	0.061926	0.053554	0.033464
	Two Factors	0.329843	0.111317	0.057041	0.046736	0.025826
	Three Factors	0.336643	0.106182	0.057058	0.049048	0.033831
>180	One Factor	0.158253	0.076913	0.072129	0.075685	0.087410
	Two Factors	0.127078	0.063742	0.054993	0.058140	0.080321
	Three Factors	0.109486	0.056811	0.051816	0.054455	0.068869

Table 7: In-Sample Absolute Percentage Pricing (Forecasting) Errors for the Restricted Model

This table reports the in-sample pricing errors, which are measured as the absolute difference between the model-implied price and the observed price as a percentage of the observed price, when the tradability restrictions in Equation (2.3) are imposed. For the index that is not tradable, the interpretation is in terms of forecasting error. For the index, we report the log index value, while the actual prices are examined for the futures and options. Fu1 refers to the short-term futures, and Fu2 refers to the medium-term futures. Op1 represents at-the-money options with short maturity (ATM-SM). Op2 stands for out-of-the-money options with medium maturity (OTM-MM), and Op3 represents out-of-the-money options with long maturity (OTM-LM). There are 1,506 days in the sample.

Securities	One Factor		Two Factors		Three Factors	
	Mean	Std.Dev.	Mean	Std.Dev.	Mean	Std.Dev.
$\log I_t$	0.001108	0.001086	0.001109	0.001088	0.001110	0.001089
Fu1	0.002483	0.001764	0.001029	0.000943	0.001012	0.000945
Fu2	0.005281	0.003247	7.91e-07	8.25e-07	6.91e-05	7.63e-05
Op1	0.088729	0.158748	0.106400	0.173483	0.105036	0.170920
Op2	0.148157	0.231590	0.133920	0.193963	0.135280	0.194014
Op3	0.147164	0.205395	0.127986	0.160872	0.126119	0.176783

Table 8: Out-of-Sample Absolute Percentage Option-Pricing Errors for the Restricted Model

This table reports the out-of-sample option-pricing errors, which are measured as the absolute difference between the model-implied price and the observed price as a percentage of the observed price, when the tradability restrictions in Equation (2.3) are imposed. There are 1,506 days and 72,706 out-of-sample observations in total. The options are divided into 15 groups based on maturity and moneyness. The mean of absolute percentage option-pricing errors for each group is reported.

Maturity	Model	Moneyness(I/K)				
		<0.94	0.94-0.97	0.97-1.00	1.00-1.03	>1.03
<60 days	one-factor	0.652439	0.513506	0.161487	0.057565	0.041584
	two-factor	0.692845	0.554652	0.183325	0.054476	0.040922
	three-factor	0.584405	0.480827	0.166922	0.059981	0.046241
60-180	one-factor	0.557312	0.152612	0.073373	0.083271	0.074360
	two-factor	0.499041	0.141435	0.074932	0.083613	0.074989
	three-factor	0.482627	0.136235	0.071584	0.084853	0.079741
>180	one-factor	0.244917	0.092458	0.106750	0.104697	0.093184
	two-factor	0.196711	0.091136	0.108988	0.104538	0.087812
	three-factor	0.220030	0.069068	0.084222	0.086135	0.079803

Table 9: Difference in the In-Sample Absolute Percentage Pricing (Forecasting) Errors between the Unrestricted and Restricted Models

This table reports the absolute difference in the in-sample pricing errors between the unrestricted and restricted models. The pricing errors are measured as the absolute difference between the model-implied price and the observed price as a percentage of the observed price. There are 1,506 days. We report both the sample mean of the absolute difference between the pricing errors and the standard error of the sample mean.

Securities	One Factor		Two Factors		Three Factors	
	Mean	Std.Err.	Mean	Std.Err.	Mean	Std.Err.
$\log I_t$	6.10e-06	1.63e-07	2.77e-06	6.34e-08	5.08e-05	9.68e-07
Fu1	0.000252	3.34e-06	0.000291	6.85e-06	0.000507	8.20e-06
Fu2	0.000523	5.85e-06	4.09e-06	1.04e-07	3.54e-05	1.14e-06
Op1	0.015049	0.000376	0.018334	0.000560	0.023515	0.000717
Op2	0.057128	0.002218	0.057749	0.002241	0.049668	0.002066
Op3	0.073896	0.002096	0.063799	0.001790	0.076463	0.002617

Table 10: Difference in the Out-of-Sample Absolute Percentage Option-Pricing Error between the Unrestricted and Restricted Models

This table reports the absolute difference in the out-of-sample option-pricing errors between the unresticted and restricted models. The pricing errors are measured as the absolute difference between the model-implied price and the observed price as a percentage of the observed price. There are 1,506 days and 72,706 out-of-sample observations in total. The options are divided into 15 groups based on maturity and moneyness. We report both the sample mean of the absolute difference and the standard error (in parentheses) of the sample mean for each group.

Maturity	Model	Moneyness(I/K)				
		<0.94	0.94-0.97	0.97-1.00	1.00-1.03	>1.03
<60 days	one-factor	0.1246	0.0668	0.0200	0.0126	0.0170
		(0.0028)	(0.0018)	(0.0005)	(0.0002)	(0.0002)
	two-factor	0.1790	0.1020	0.0311	0.0153	0.0183
		(0.0036)	(0.0028)	(0.0009)	(0.0003)	(0.0003)
	three-factor	0.0953	0.0777	0.0309	0.0217	0.0210
		(0.0028)	(0.0025)	(0.0008)	(0.0004)	(0.0003)
60-180	one-factor	0.2036	0.0423	0.0163	0.0309	0.0441
		(0.0044)	(0.0011)	(0.0003)	(0.0004)	(0.0005)
	two-factor	0.1750	0.0403	0.0229	0.0375	0.0495
		(0.0041)	(0.0013)	(0.0005)	(0.0005)	(0.0005)
	three-factor	0.1523	0.0457	0.0229	0.0375	0.0473
		(0.0043)	(0.0015)	(0.0005)	(0.0005)	(0.0006)
>180	one-factor	0.1139	0.0489	0.0789	0.0920	0.0947
		(0.0036)	(0.0013)	(0.0014)	(0.0015)	(0.0017)
	two-factor	0.0949	0.0471	0.0737	0.0781	0.0706
		(0.0030)	(0.0011)	(0.0011)	(0.0012)	(0.0014)
	three-factor	0.1282	0.0332	0.0534	0.0668	0.0703
		(0.0041)	(0.0008)	(0.0009)	(0.0010)	(0.0013)

Figure 1: Plots of the Tradability Premium in the Simulation

This figure plots the tradability premium and the tradability premium as a percentage of the market price of risk in the two simulations in Section 4. The tradability premium is measured as the difference between the market risk premium computed using parameters estimated from the unrestricted model and the market risk premium computed using parameters estimated from the restricted model in which the tradability restrictions in Equation (2.3) are imposed. In this paper, the market risk premium associated with w_t is $-\beta(\gamma_0 + \sum_{k=1}^{K} \gamma_k x_{k,t})$. Therefore, the tradability premium is a stochastic process. This figure shows two samples.

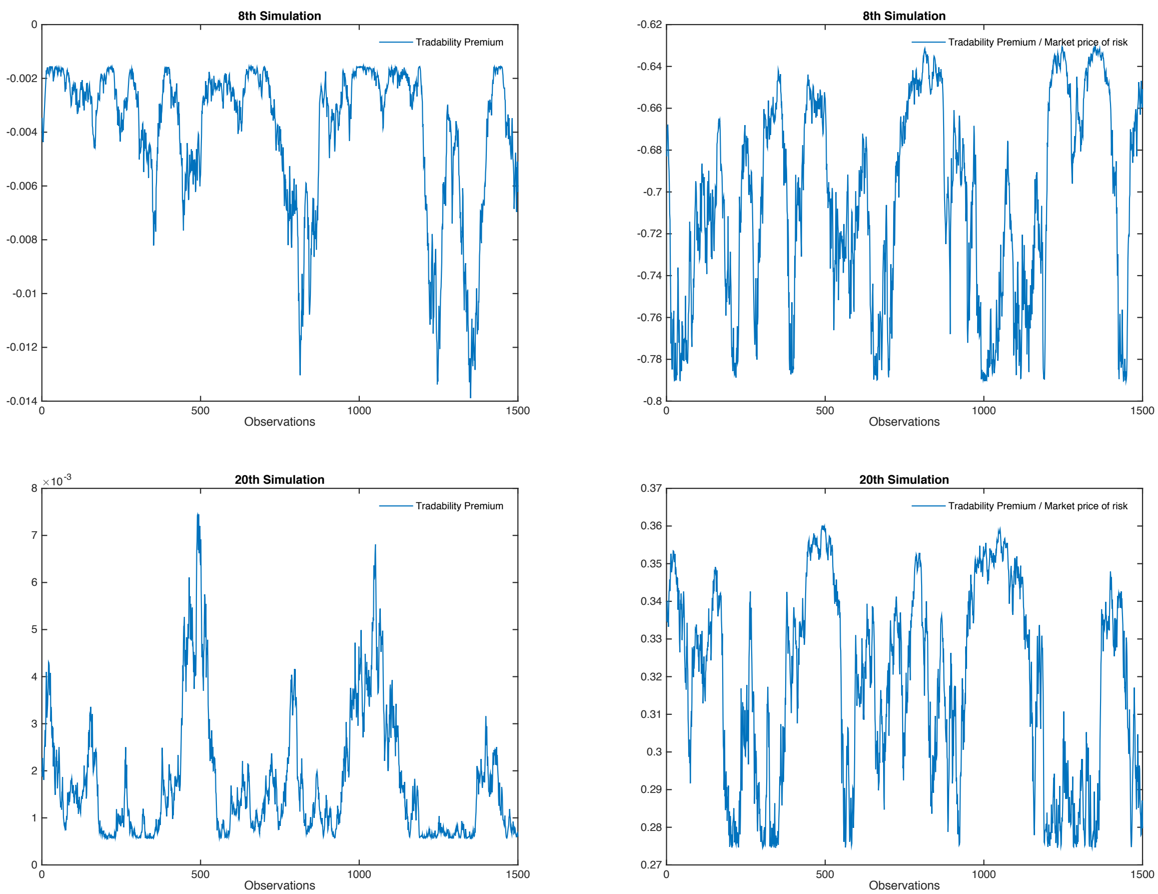


Figure 2: Plots of Observations

This figure plots the observations for estimation in this paper. The data covers the period from January 3, 2001, to December 29, 2006, and there are 1,506 days in total. For each day, there are six observations, including the daily log difference of the S&P 500 Index (dlogSPXspot), the annualized log of futures spot ratio (ALFSR) for two S&P 500 futures, and price-vega ratios for three S&P 500 options.

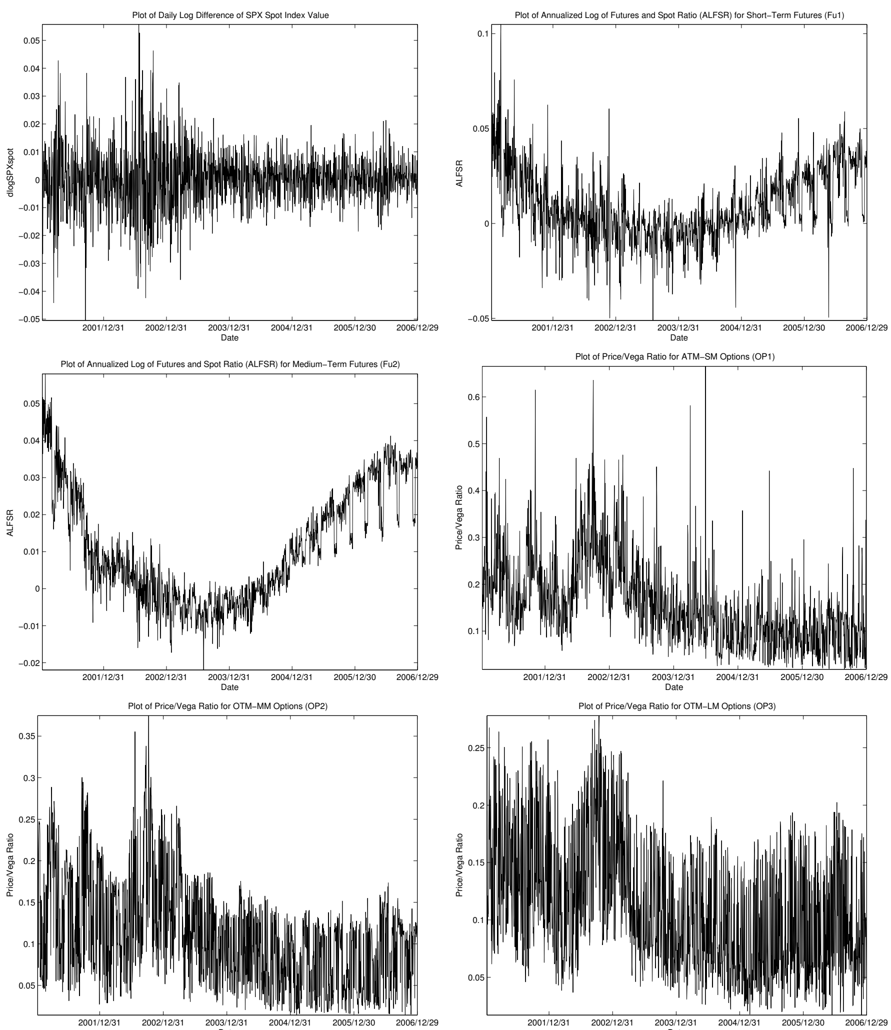
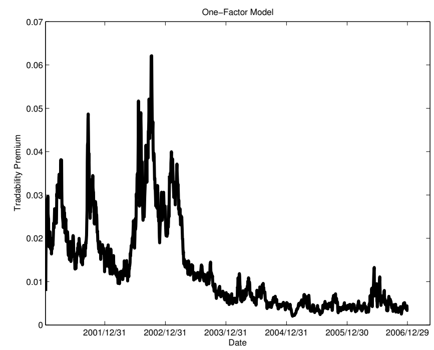


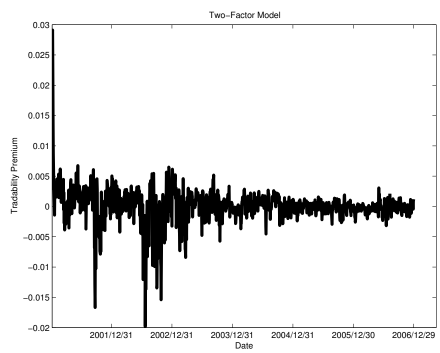
Figure 3: Tradability Premium

This figure plots the tradability premium for each model based on the parameter estimate. The tradability premium is defined as the difference between the market risk premium implied by the unrestricted model and the market risk premium implied by the restricted model when the tradability restrictions in Equation (2.3) are imposed. There are 1,506 days.

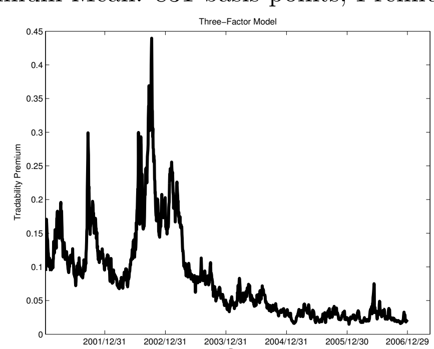
One-Factor Model (Premium Mean: 130 basis points; Premium Std Dev: 108 basis points)



Two-Factor Model (Premium Mean: 1 basis points; Premium Std Dev: 27 basis points)



Three-Factor Model (Premium Mean: 831 basis points; Premium Std Dev: 672 basis points)



References

[1] Ball, C., and A., Roma (1994), "Stochastic Volatility Option Pricing", Journal of Financial and Quantitative Analysis, 29, 589-607.

[2] Black, F., and M., Scholes (1973), "The Pricing of Options and Corporate Liabilities", Journal of Political Economy, 81, 637-659.

[3] Brennan, M., and E., Schwartz (1990), "Arbitrage in Stock Index Futures", Journal of Business, Vol. 63, 7-31.

[4] Carrasco, M., Chernov, M., Florens, J., and E., Ghysels (2000), "Efficient Estimation of Jump Diffusions and General Dynamic Models with a Continuum of Moment Conditions", IDEI Working Paper, n. 116.

[5] Chen, R., and L., Scott (1993), "Maximum Likelihood Estimation of a Multi-Factor Equilibrium Model of the Term Structure of Interest Rates", Journal of Fixed Income, December, 1993, 14-31.

[6] Constantinides, G. (1992), "A Theory of the Nominal Term Structure of Interest Rates", Review of Financial Studies, 5, 531-552.

[7] Cox, J., Ingersoll, J., and S., Ross (1981), "The Relation between Forward Prices and Future Prices", Journal of Financial Economics, 9, 321-346.

[8] Cox, J., Ingersoll, J., and S., Ross (1985a), "An Intertemporal General Equilibrium Model of Asset Prices", Econometrica, 53, 363-384.

[9] Cox, J., Ingersoll, J., and S., Ross (1985b), "A Theory of the Term Structure of Interest Rates", Econometrica, 53, 385-407.

[10] Dai, Q., and K., Singleton (2000), "Specification Analysis of Affine Term Structure Models", Journal of Finance, Vol 55, 1943-1978.

[11] Dai, Q., and K., Singleton (2003), "Term Structure Dynamics in Theory and Reality", The Review of Financial Studies, 16, 631-678.

[12] Duffie, D., and R., Kan (1996), "A Yield-Factor Model of Interest Rates", Mathematical Finance, 6, 379-406.

[13] Duffie, D., Pan, J., and K., Singleton (2000), "Transform Analysis and Asset Pricing for Affine Jump-Diffusions", Econometrica, 68, 1343-1376.

[14] French, K. R. (1983), "A Comparison of Futures and Forward Prices", Journal of Financial Economics, Vol. 12, 311-342.

[15] Gourieroux, C., Monfort, A., and A., Trognon (1984), Pseudo-maximum Likelihood Methods: Theory, Econometrica, 42, 681-700.

[16] Harrison, J., and D., Kreps (1979), "Martingales and Arbitrage in Multiperiod Securities Markets", Journal of Economic Theory, 20, 381-408.

[17] Heston, S.(1993) "A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options", The Review of Financial Studies, 6, 327-343.

[18] Li, J.(2012) "Option-Implied Volatility Factors and the Cross-Section of Market Risk Premia", Journal of Banking and Finance, Jan 2012, Vol 36, Issue 1, p. 249-260 .

[19] Merton, R. (1973), "The Theory of Rational Option Pricing", Bell Journal of Economics and Management Science, 4, 141-183.

[20] Roll, R. (1977), "A Critique of the Asset Pricing Theory's Tests Part I: On Past and Potential Testability of the Theory", Journal of Financial Economics, 4 (2), 129176.

[21] Schwartz, E. (1997), "The Stochastic Behavior of Commodity Prices: Implications for Valuation and Hedging", Journal of Finance, vol 52, 923-973.

[22] Wan, E. A., and R., Van Der Merwe (2000), "The Unscented Kalman Filter for Nonlinear Estimation", Adaptive Systems for Signal Processing, Communications, and Control Symposium 2000. AS-SPCC. The IEEE 2000, 153-158.

[23] White, H. (1982), "Maximum Likelihood Estimation of Misspecified Models", Econometrica, 50, 1-25.

[24] Xu, P. (2014), "Is the S&P 500 Index Tradable?", The Journal of Index Investing, Winter 2014, Volume 5, Number 3, 10-20.

[25] Xu, P. (2007), "S&P 500 Index Options and Put-Call Parity", Manuscript.

Notes

[1]The pricing formulas for European derivatives written on the index futures can also be explicitly derived in this framework. The details are given in this paper's additional material.

[2]$Re(z_k(1-iv)) > -1$, $Re(z_k(-iv)) > -1$, $Re(z_k(1+iv)) > -1$, and $Re(z_k(iv)) > -1, \forall\ k = 1, \cdots, K$, where $Re(\cdot)$ denotes the real part of a complex number, should also hold in order for the pricing formulas to exist. $Re(z_k(1-iv)) = Re(z_k(1+iv)) = -\alpha_k - (\beta+1)\mu_k - \frac{\gamma_k}{2}(\beta+1)^2 + \frac{\gamma_k}{2}v^2 = z_k(1) + \frac{\gamma_k}{2}v^2$ and $Re(z_k(-iv)) = Re(z_k(iv)) = -\alpha_k - \beta\mu_k - \frac{\gamma_k}{2}\beta^2 + \frac{\gamma_k}{2}v^2 = z_k(0) + \frac{\gamma_k}{2}v^2$. Therefore, the restrictions (1.12) and (1.16) are sufficient.

[3]Note that only futures and options data correspond to tradable assets.

[4]If the spot-futures parity, $F_{t,t+h} = I_t e^{(r-q)h}$, in which r is the annually continuously compounded riskfree interest rate and q is the dividend yield, holds for the index and its futures, then ALFSR simply equals the riskfree interest rate minus the dividend yield.

[5]The Black-Scholes *vega* is the derivative of the Black-Scholes options price with respect to the volatility.

[6]Open interest refers to the total number of long (short) positions outstanding in a derivative contract.