ASIAN JOURNAL OF STATISTICS AND APPLICATIONS

VOL NO. 2, ISSUE 1, PAGE NO. 50- 72, YEAR 2025

ISSN: 3048-7455

https://DOI:10.47509/AJSA.2025.v02i01.04

Minimum Contrast Estimation in Fractional Ornstein-Uhlenbeck Driven by Fractional Ornstein-Uhlenbeck Process

Jaya P. N. Bishwal

Department of Mathematics and Statistics, University of North Carolina at Charlotte, 376 Fretwell Buildings, 9201 University City Boulevard, Charlotte, NC 28223-0001

ABSTRACT

We generalize fractional Ornstein-Uhlenbeck process whose driving term is another fractional Ornstein-Uhlenbeck process. The motivation is related to stochastic volatility model. We estimate the parameters of both processes by maximum likelihood method and minimum contrast method. We obtain strong consistency and asymptotic normality of the estimators as the time length of observation becomes large.

KEYWORDS

Stochastic differential equation, fractional Brownian motion, fractional Ornstein-Uhlenbeck process, correlation, volatility, maximum likelihood estimator, minimum contrast estimator, Durbin-Watson statistic.

1. Introduction and Preliminaries

Ornstein-Uhlenbeck processes driven by Levy processes have received a lot of attention in finance, see Barndorff-Neilsen and Shephard (2001). Levy processes are processes with stationary independent increments. Levy Ornstein-Uhlenbeck (LOU) process generalizes the Ornstein-Uhlenbeck process to include jumps. The correlation phenomena of volatility exists in the driving process. First we consider fractional Ornstein-Uhlenbeck process as the driving process.

The fractional Brownian motion (fBm, in short), which provides a suitable generalization of the Brownian motion, is one of the simplest stochastic processes exhibiting long range-dependence. Consider a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ on which all random variables and processes below are defined. A normalized fractional Brownian motion $\{W_t^H, t \geq 0\}$ with Hurst parameter $H \in (0,1)$ is a centered Gaussian process with continuous sample paths whose covariance kernel is given by

$$E(W_t^H W_s^H) = \frac{1}{2}(s^{2H} + t^{2H} - |t-s|^{2H}), \quad s, t \geq 0. \tag{1.1}$$

The process is self similar (scale invariant) and it can be represented as a stochastic

CONTACT Jaya P. N. Bishwal. Email: J.Bishwal@charlotte.edu

integral with respect to standard Brownian motion. For $H = \frac{1}{2}$, the process is a standard Brownian motion. For $H \neq \frac{1}{2}$, the fBm is not a semimartingale and not a Markov process, but a Dirichlet process. The increments of the fBm are negatively correlated for $H < \frac{1}{2}$ and positively correlated for $H > \frac{1}{2}$ and in this case they display long-range dependence (persistence). The parameter H which is also called the self similarity parameter, measures the intensity of the long range dependence.

Consider the fractional SDEs

$$dX_t = \theta X_t dt + dV_t, \ t \geq 0, \ X_0 = 0 \tag{1.2}$$

$$dV_t = \rho V_t dt + dW_t^H, \ t \geq 0, \ V_0 = 0 \tag{1.3}$$

where $\theta < 0$ and $\rho < 0$ are the two unknown parameters to be estimated based on observations of the process $\{X_t, t \geq 0\}$. Thus we have the integro fractional SDE

$$dX_t = (\theta + \rho) X_t dt - \theta\rho \int_0^t X_s ds + dW_t^H, \ t \geq 0. \tag{1.4}$$

If $\rho = 0$, we have the standard fractional Ornstein-Uhlenbeck (FOU) model. First, we consider the case $H = 1/2$. We denote $W_t^{1/2} = W_t$. Observe that

$$\int_0^T X_t dX_t = \theta \int_0^T X_t^2 dt + \rho \int_0^T X_t V_t dt + \int_0^T X_t dW_t, \tag{1.5}$$

$$\frac{1}{T} \int_0^T X_t^2 dt \to -\frac{1}{2(\rho + \theta)} \quad \text{a.s.}, \tag{1.6}$$

$$\frac{1}{T} \int_0^T X_t V_t dt \to -\frac{1}{2(\rho + \theta)} \quad \text{a.s.}, \tag{1.7}$$

From the SLLN for continuous martingales,

$$\frac{1}{T} \int_0^T X_t dW_t \to 0 \quad \text{a.s.} \tag{1.8}$$

Hence

$$\frac{1}{T} \int_0^T X_t dX_t = -\frac{\theta}{2(\rho + \theta)} - \frac{\rho}{2(\rho + \theta)} = -\frac{1}{2} \quad \text{a.s.} \tag{1.9}$$

Observe that

$$\int_0^T V_t dV_t = \rho \int_0^T V_t^2 dt + \int_0^T V_t dW_t. \tag{1.10}$$

From Feigin (1976), it follows that

$$\frac{1}{T}\int_0^T V_t^2 dt \to -\frac{1}{2\rho} \quad \text{a.s.} \tag{1.11}$$

and from SLLN for continuous martingales

$$\frac{1}{T}\int_0^T V_t dW_t \to 0 \ \text{ a.s.} \tag{1.12}$$

Hence

$$\frac{1}{T}\int_0^T V_t dV_t \to -\frac{1}{2} \ \text{ a.s.} \tag{1.13}$$

By Itô's formula,

$$\frac{1}{T}\int_0^T X_t dX_t = \frac{1}{2}\left(\frac{X_T^2}{T} - 1\right), \ \frac{1}{T}\int_0^T V_t dV_t = \frac{1}{2}\left(\frac{V_T^2}{T} - 1\right). \tag{1.14}$$

From (1.9) and (1.13), it follows that

$$\lim_{T\to\infty}\frac{X_T^2}{T} = 0, \ \lim_{T\to\infty}\frac{V_T^2}{T} = 0. \tag{1.15}$$

Since $X_t = \theta\Sigma_T + V_T$, where $\Sigma_T = \int_0^T X_s ds$, hence from (1.25), we have

$$\lim_{T\to\infty}\frac{1}{2}\left(\frac{\widehat{V}_T^2}{T} - 1\right) = -\frac{1}{2} \text{ a.s.} \tag{1.16}$$

where the residuals generated by the estimation of θ at stage T

$$\widehat{V}_T = X_T - \frac{\int_0^T X_t dX_t}{\int_0^T X_t^2 dt}\int_0^T X_t dt \text{ a.s.} \tag{1.17}$$

We have the decomposition

$$\widehat{\rho}_T = \frac{T}{2\widehat{\Lambda}_T}\left(\frac{\widehat{V}_T^2}{T} - 1\right) \tag{1.18}$$

where

$$\widehat{\Lambda}_T = \int_0^T \widehat{V}_t^2 dt. \tag{1.19}$$

Further,

$$\lim_{T\to\infty}\frac{1}{T}\widehat{\Lambda}_T = -\frac{1}{2\rho^*} \text{ a.s.} \tag{1.20}$$

where ρ^* is defined in (1.36)
From the self-similarity of Brownian motion, we have

$$\Lambda_T = \int_0^T W_t^2 dt =^{\mathcal{L}} T \int_0^T W_{t/T}^2 dt = T^2 \int_0^1 W_s^2 ds = T^2 \Lambda_1.$$

Consequently, for any $0 < a < 2$,

$$\lim_{T\to\infty} \frac{1}{T^a}\Lambda_T = \infty \quad \text{a.s.} \tag{1.21}$$

From Liptser and Shiryayev (1978), we have

$$E\left[\exp(-\frac{1}{T^a}\Lambda_T)\right] = E\left[\exp(-\frac{T^2}{T^a}\Lambda_1)\right] = \frac{1}{\sqrt{\cosh(r_T(a))}} \tag{1.22}$$

where $r_T(a) = \sqrt{2T^{2-a}}$ for $0 < a < 2$. Hence

$$\lim_{T\to\infty} E\left[\exp(-\frac{1}{T^a}\Lambda_T)\right] = 0 \tag{1.23}$$

which also gives (1.21). Further,

$$\widehat{\Lambda}_T = \Lambda_T(1 + o(1)) \text{ a.s.} \tag{1.24}$$

The maximum likelihood estimators of θ and ρ are given by

$$\widehat{\theta}_T := \frac{X_T^2 - T}{2\int_0^T X_t^2 dt}, \quad \widehat{\rho}_T := \frac{\widehat{V}_T^2 - T}{2\int_0^T \widehat{V}_t^2 dt} \tag{1.24}$$

respectively where

$$\widehat{V}_t := X_t - \widehat{\theta}_T \int_0^t X_s ds \quad \text{a.s.} \tag{1.25}$$

$$= X_t - \frac{\int_0^T X_t dX_t}{\int_0^T X_t^2 dt} \int_0^t X_s ds \quad \text{a.s.} \tag{1.26}$$

The Durbin-Watson (Durbin and Watson (1950, 1951, 1971)) statistic is given by

$$\widehat{D}_T := \frac{2\int_0^T \widehat{V}_t^2 dt - \widehat{V}_T^2 + T}{\int_0^T \widehat{V}_t^2 dt} = 2(1 - \widehat{\rho}_T). \tag{1.28}$$

When $H = 1/2$, $\theta < 0$ and $\rho < 0$, using SLLN for continous martingales, Bercu *et al.* (2014) showed that

$$(\widehat{\theta}_T - \theta^*)^2 = O\left(\frac{\log T}{T}\right) \quad \text{a.s.} \tag{1.29}$$

and hence

$$\widehat{\theta}_T \to \theta^* \quad \text{a.s. as } T \to \infty. \tag{1.30}$$

$$\sqrt{T}(\widehat{\theta}_T - \theta^*) \to^{\mathcal{D}} \mathcal{N}(0, \sigma_\theta^2) \tag{1.31}$$

where

$$\theta^* := \theta + \rho \tag{1.32}$$

and

$$\sigma_\theta^2 = -2\theta. \tag{1.33}$$

Further

$$\widehat{\rho}_T \to \rho^* \quad \text{a.s. as } T \to \infty, \tag{1.34}$$

and

$$\sqrt{T}(\widehat{\rho}_T - \rho^*) \to^{\mathcal{D}} \mathcal{N}(0, \sigma_\rho^2) \text{ as } T \to \infty \tag{1.35}$$

where

$$\rho^* := \frac{\theta\rho(\theta + \rho)}{(\theta + \rho)^2 + \theta\rho}, \tag{1.36}$$

and

$$\sigma_\rho^2 := -\frac{2\rho^*((\theta^*)^6 + \theta\rho((\theta^*)^4 - \theta\rho(2(\theta^*)^2 - \theta\rho)))}{((\theta^*)^2 + \rho\theta)^3}. \tag{1.37}$$

2. Minimum Contrast Estimation

Minimum contrast estimator is known to be robust and efficient. First we consider the case with $H = 1/2$. The MCE's of θ and ρ are defined respectively as

$$\widetilde{\theta}_T := \frac{-T}{2\int_0^T X_t^2 dt}, \quad \widetilde{\rho}_T := \frac{-T}{2\int_0^T \widetilde{V}_t^2 dt}$$

where

$$\widetilde{V}_t := X_t - \widetilde{\theta}_T \int_0^t X_s ds = X_t + \frac{T \int_0^t X_s ds}{2 \int_0^T X_s^2 ds}.$$

We have

$$\widetilde{V}_T^2 - T = o(\Lambda_T) \ \text{ a.s.}$$

We have the following asymptotic properties of the MCEs:

Theorem 2.1 When $H = 1/2$, when $\theta < 0$ and $\rho < 0$

$$\widetilde{\theta}_T \to \theta^* \ \text{ a.s. as } T \to \infty.$$

$$\sqrt{T}(\widetilde{\theta}_T - \theta^*) \to^{\mathcal{D}} \mathcal{N}(0,\ \sigma_\theta^2) \text{ as } T \to \infty.$$

Further,

$$\widetilde{\rho}_T \to \rho^* \ \text{ a.s. as } T \to \infty,$$

and

$$\sqrt{T}(\widetilde{\rho}_T - \rho^*) \to^{\mathcal{D}} \mathcal{N}(0,\ \sigma_\rho^2) \text{ as } T \to \infty.$$

We consider the long memory case with $H \neq 1/2$. We turn to the equivalent semi-martingale representation of the model. Let $H > 1/2$. Define

$$\begin{aligned}
\kappa_H &:= 2H\Gamma(3/2 - H)\Gamma(H + 1/2),\\
k_H(t,s) &:= \kappa_H^{-1}(s(t-s))^{\frac{1}{2}-H}, \ \ \lambda_H = \frac{2H\Gamma(3-2H)\Gamma(H+\frac{1}{2})}{\Gamma(3/2-H)}\\
v_t \equiv v_t^H &:= \lambda_H^{-1} t^{2-2H}, \ \ M_t^H = \int_0^t k_H(t,s) dW_s^H.
\end{aligned}$$

From Norros *et al.* (1999) it is well known that M_t^H is a Gaussian martingale, called the *fundamental martingale* whose variance function $\langle M^H \rangle_t$ is v_t^H. The natural filtration of the martingale M^H coincides with the natural filtration of the fBm W^H since

$$W_t^H := \int_0^t K(t,s) dM_s^H \tag{2.1}$$

holds for $H \in (1/2, 1)$ where

$$K_H(t,s) := H(2H-1)\int_s^t r^{h-\frac{1}{2}}(r-s)^{H-\frac{3}{2}}, \quad 0 \le s \le t \tag{2.2}$$

and for $H = 1/2$, the convention $K_{1/2} \equiv 1$ is used.

Define

$$Q_t := \frac{d}{dv_t}\int_0^t k_H(t,s)X_s ds. \tag{2.3}$$

It is easy to see that

$$Q_t = \frac{\lambda_H}{2(2-2H)}\left\{t^{2H-1}Z_t + \int_0^t r^{2H-1}dZ_s\right\}. \tag{2.4}$$

X admits the representation

$$X_t = \int_0^t K_H(t,s)dZ_s. \tag{2.5}$$

The natural filtration generated by the fundamental semimartingale process

$$Z_t = \theta\int_0^t Q_s dv_s + M_t^H \tag{2.6}$$

and the process X coincide, see Kleptsyna and Le Breton (2002). The available information for X and Z are strictly equivalent.

Let the realization $\{X_t, 0 \le t \le T\}$ or equivalently $\{Z_t, 0 \le t \le T\}$ be denoted by Z_0^T. Let P_θ^T be the measure generated on the space (C_T, B_T) of continuous functions on $[0, T]$ with the associated Borel σ-algebra B_T generated under the supremum norm by the process X_0^T and P_0^T be the standard Wiener measure. Applying fractional Girsanov formula, when θ is the true value of the parameter, P_θ^T is absolutely continuous with respect to P_0^T and the Radon-Nikodym derivative (likelihood) of P_θ^T with respect to P_0^T based on Z_0^T is given by

$$L_T(\theta) := \frac{dP_\theta^T}{dP_0^T}(Z_0^T) = \exp\left\{\theta\int_0^T Q_t dZ_t - \frac{\theta^2}{2}\int_0^T Q_t^2 dv_t\right\}. \tag{2.7}$$

Consider the score function, the derivative of the log-likelihood function $l_t(\theta)$, which is given by

$$l_T'(\theta) := \int_0^T Q_t dZ_t - \theta\int_0^T Q_t^2 dv_t. \tag{2.8}$$

The maximum likelihood estimators of θ and ρ are defined as

$$\widehat{\theta}_T := \frac{\int_0^T Q_t dZ_t}{2\int_0^T Q_t^2 dv_t}, \quad \widehat{\rho}_T := \frac{\int_0^T \widehat{V}_t dZ_t}{2\int_0^T \widehat{V}_t^2 dv_t} \tag{2.9}$$

respectively, where

$$\widehat{V}_t := Q_t - \widehat{\theta}_T \int_0^t Q_s dv_s. \tag{2.10}$$

When $\rho = 0$, using the fractional Itô formula, the score function $l'_T(\theta)$ can be written as

$$l'_T(\theta) = \frac{1}{2}\left[\frac{\lambda_H}{(2-2H)} Z_T \int_0^T t^{2H-1} dZ_t - T\right] - \theta \int_0^T Q_t^2 dv_t. \tag{2.11}$$

Consider the contrast function

$$M_T(\theta) = -\frac{T}{2} - \theta \int_0^T Q_t^2 dv_t \tag{2.12}$$

and the minimum contrast estimate (MCE)

$$\widetilde{\theta}_T := \frac{-T}{2\int_0^T Q_t^2 dv_t.} \tag{2.13}$$

M-estimator is reduced to the minimum contrast estimator. Note that the MCE does not involve stochastic integral unlike the MLE.

Observe that

$$\left(\frac{T}{-2\theta}\right)^{1-H} (\widetilde{\theta}_T - \theta) = \frac{\left(\frac{-2\theta}{T}\right)^{1-H} N_T}{\left(\frac{2\theta}{T}\right)^{2-2H} I_T} \tag{2.14}$$

where

$$N_T := \theta I_T - \frac{T}{2} \quad \text{and} \quad I_T := \int_0^T Q_t^2 dv_t. \tag{2.15}$$

Kleptsyna and Le Breton (2002) proved the following Cameron-Martin type formula for $a > 0$

$$E \exp(-aI_T) = \left\{\frac{4 \sin \pi H \sqrt{\theta^2 + 2a}\, e^{-\theta T}}{\pi T \mathcal{D}_T^H(\theta; \sqrt{\theta^2 + 2a}}\right\}^{1/2} \tag{2.16}$$

where

$$\begin{aligned} \mathcal{D}_T^H(\theta;\alpha) \quad := \quad & [\alpha \cosh(\tfrac{\alpha}{2}T) - \theta \sinh(\tfrac{\alpha}{2}T)]^2 J_{-H}(\tfrac{\alpha}{2}T) J_{H-1}(\tfrac{\alpha}{2}T) \\ & -[\alpha \sinh(\tfrac{\alpha}{2}T) - \theta \cosh(\tfrac{\alpha}{2}T)]^2 J_{1-H}(\tfrac{\alpha}{2}T) J_H(\tfrac{\alpha}{2}T) \end{aligned} \tag{2.17}$$

for $\alpha > 0$ and J_ν is the modified Bessel function of first kind of order ν. For $H = 1/2$ this formula reduces to the *Novikov's formula* given in Liptser and Shiryayev (1978).

By analytic continuation, this formula can be extended to the complex plane $z_1 \in \mathbb{C}$. We have the following result from Bishwal (2011a).

Let $\phi_T(z_1) := E\exp(z_1 I_T),\ z_1 \in \mathbb{C}$. Then $\phi_T(z_1)$ exists for $|z_1| \le \delta$, for some $\delta > 0$ and is given by

$$\phi_T(z_1) = \left\{ \frac{4\sin\pi H\sqrt{\theta^2 - 2z_1}e^{-\theta T}}{\pi T\mathcal{D}_T^H(\theta; \sqrt{\theta^2 - 2z_1}} \right\}^{1/2} \tag{2.18}$$

and we choose the principal branch of the square root.

The Durbin-Watson (Durbin and Watson (1950, 1951, 1971)) statistic is given by

$$\widehat{D}_T = \frac{2\int_0^T \widehat{V}_t^2 dt - \widehat{V}_T^2 + T}{\int_0^T \widehat{V}_t^2 dt} = 2(1 - \widehat{\rho}_T) \tag{2.19}$$

where

$$\widetilde{V}_t := Q_t - \widetilde{\theta}_T \int_0^t Q_s dv_s. \tag{2.20}$$

From Bercu, Coutin and Savy (2010), when $\rho = 0$, we have

$$\lim_{T\to\infty} \frac{I_T}{T} = -\frac{1}{2\theta}$$

$$\frac{1}{\sqrt{T}}(I_T + \frac{T}{2\theta}) \to^{\mathcal{D}} \mathcal{N}(0,\ -\frac{1}{2\theta^3}).$$

We have the following asymptotic properties of the MCEs:

Theorem 2.2 When $H > 1/2$, when $\theta < 0$ and $\rho < 0$,

$$\widehat{\theta}_T \to \theta^* \ \text{ a.s. } \ \text{as } T \to \infty$$

and

$$\sqrt{T}(\widehat{\theta}_T - \theta^*) \to^{\mathcal{D}} \mathcal{N}(0, \sigma^2_{\theta,H}) \text{ as } T \to \infty$$

where

$$\theta^* := \theta + \rho$$

and

$$\sigma^2_{\theta,H} = -2\theta\lambda_H^{-1}.$$

Further,

$$\widehat{\rho}_T \to \rho^* \ \text{ a.s. } \ \text{as } T \to \infty,$$

and

$$\sqrt{T}(\widehat{\rho}_T - \rho^*) \to^{\mathcal{D}} \mathcal{N}(0, \sigma^2_{\rho,H}) \text{ as } T \to \infty$$

where

$$\rho^* := \frac{\theta\rho(\theta+\rho)}{(\theta+\rho)^2 + \theta\rho},$$

and

$$\sigma^2_{\rho,H} := \lambda_H^{-1} \frac{2\rho^*((\theta^*)^6 + \theta\rho((\theta^*)^4 - \theta\rho(2(\theta^*)^2 - \theta\rho)))}{((\theta^*)^2 + \rho\theta)^3}.$$

The minimum contrast estimators (MCEs) of θ and ρ are defined as

$$\widetilde{\theta}_T := \frac{-T}{2\int_0^T Q_t^2 dv_t}, \quad \widetilde{\rho}_T := \frac{-T}{2\int_0^T \widetilde{V}_t^2 dv_t}$$

respectively where

$$\widetilde{V}_t := Q_t - \widetilde{\theta}_T \int_0^t Q_s dv_s.$$

The Durbin-Watson statistic is given by

$$\widetilde{D}_T := 2(1 - \widetilde{\rho}_T).$$

Let

$$D^* := 2(1 - \rho^*).$$

Theorem 2.3 When $H > 1/2$, when $\theta < 0$ and $\rho < 0$

$$\widetilde{\theta}_T \to \theta^* \quad \text{a.s. as } T \to \infty$$

and

$$\sqrt{T}(\widetilde{\theta}_T - \theta^*) \to^{\mathcal{D}} \mathcal{N}(0,\ \sigma^2_{\theta,H}) \text{ as } T \to \infty.$$

Further

$$\widetilde{\rho}_T \to \rho^* \quad \text{a.s. as } T \to \infty,$$

and

$$\sqrt{T}(\widetilde{\rho}_T - \rho^*) \to^{\mathcal{D}} \mathcal{N}(0,\ \sigma^2_{\rho,H}) \text{ as } T \to \infty.$$

Also

$$\sqrt{T}(\widetilde{D}_T - D^*) \to^{\mathcal{D}} \mathcal{N}(0,\ \sigma^2_{D,H}) \text{ as } T \to \infty$$

where, $\sigma^2_{D,H} = 4\sigma^2_{\rho,H}$.

Theorem 2.4 If $\rho = 0$,

$$T\widetilde{\rho}_T \to^{\mathcal{D}} \mathcal{W} \quad \text{as } T \to \infty$$

where

$$\mathcal{W} = \frac{W_1^2 - 1}{2\int_0^1 W_s^2 ds.}$$

Also

$$T(\widetilde{D}_T - 2) \to^{\mathcal{D}} -2\mathcal{W} \quad \text{as } T \to \infty$$

where

$$\mathcal{W} := \frac{\mathcal{T}^2 - 1}{2\mathcal{S}}$$

where $\mathcal{T}$ and $\mathcal{S}$ are given by the Karhunen-Loeve expansion where

$$\mathcal{T} := \sqrt{2}\sum_{n=1}^{\infty} \gamma_n Z_n, \quad \mathcal{S} := \sum_{n=1}^{\infty} \gamma_n^2 Z_n^2$$

with $\gamma_n := 2(-1)^n/((2n-1)\pi)$ and Z_n are i.i.d. standard normal.

3. Approximate Minimum Contrast Estimation

The process Q depends continuously on X and therefore, the discrete observations of X does not allow one to obtain the discrete observations of Q. We consider discrete observations of Q. The process Q can be approximated by

$$\widetilde{Q}_n = \kappa_H^{-1}\eta_H n^{2H-1}\sum_{j=0}^{n-1} j^{1/2-H}(n-j)^{-1/2-H}X_j. \tag{3.1}$$

It is easy to show that $\widetilde{Q}_n \to Q_t$ almost surely (a.s.) as $n \to \infty$, see Tudor and Viens (2007). Define a new partition $0 \le r_1 < r_2 < r_3 < \cdots < r_{m_k} = t_k, \quad k = 1, 2, \cdots, n.$

Define

$$\widetilde{Q}_{t_k} = \kappa_H^{-1}\eta_H t_k^{2H-1}\sum_{j=1}^{m_k} r_j^{1/2-H}(r_{m_k}-r_j)^{-1/2-H}X_{r_j}(r_j-r_{j-1}), \quad k=1,2,\cdots,n. \tag{3.2}$$

It is easy to show that $\widetilde{Q}_{t_k} \to Q_t$ almost surely (a.s.) as $m_k \to \infty$ for each $k=1,2,\cdots,n$. We use this approximation of observations in the calculation of our estimators.

For $H=0.5$,

$$v_{t_i}-v_{t_{i-1}} = \lambda_H^{-1}\left(\frac{T}{n}\right)^{2-2H}[i^{2-2H}-(i-1)^{2-2H}] = \frac{T}{n}, \quad i=1,2,\ldots,n.$$

Define

$$\widetilde{V}_{t_{i-1}} = \widetilde{Q}_{t_{i-1}} - \widetilde{\theta}_{n,T,F}\sum_{i=1}^{n}\widetilde{Q}_{t_{i-1}}(v_{t_i}-v_{t_{i-1}}) = \widetilde{Q}_{t_{i-1}} + \frac{T\sum_{i=1}^{n}\widetilde{Q}_{t_{i-1}}(v_{t_i}-v_{t_{i-1}})}{2\sum_{i=1}^{n}\widetilde{Q}^2_{t_{i-1}}(v_{t_i}-v_{t_{i-1}})}. \tag{3.3}$$

Now we define the weighted approximate minimum contrast estimators (AMCEs). Define the weighted sum of squares

$$G_{n,T} := \left\{\sum_{i=1}^{n} w_{t_i}\widetilde{Q}^2_{t_{i-1}}(v_{t_i}-v_{t_{i-1}}) + \sum_{i=2}^{n+1} w_{t_i}\widetilde{Q}^2_{t_{i-1}}(v_{t_i}-v_{t_{i-1}})\right\}, \tag{3.4}$$

$$G_{1,n,T} := \left\{\sum_{i=1}^{n} w_{t_i}\widetilde{V}^2_{t_{i-1}}(v_{t_i}-v_{t_{i-1}}) + \sum_{i=2}^{n+1} w_{t_i}\widetilde{V}^2_{t_{i-1}}(v_{t_i}-v_{t_{i-1}})\right\} \tag{3.5}$$

where $w_{t_i} \geq 0$ is a weight function with $\sum_{i=1}^{n} w_{t_i} = 1$. Define the discrete increasing functions

$$I_{n,T} := \sum_{i=1}^{n}\widetilde{Q}^2_{t_{i-1}}(v_{t_i}-v_{t_{i-1}}), \tag{3.6}$$

$$J_{n,T} := \sum_{i=2}^{n+1}\widetilde{Q}^2_{t_{i-1}}(v_{t_i}-v_{t_{i-1}}) = \sum_{i=1}^{n}\widetilde{Q}^2_{t_i}(v_{t_i}-v_{t_{i-1}}), \tag{3.7}$$

$$I_{1,n,T} := \sum_{i=1}^{n}\widetilde{V}^2_{t_{i-1}}(v_{t_i}-v_{t_{i-1}}), \tag{3.8}$$

$$J_{1,n,T} := \sum_{i=2}^{n+1}\widetilde{V}^2_{t_{i-1}}(v_{t_i}-v_{t_{i-1}}) = \sum_{i=1}^{n}\widetilde{V}^2_{t_i}(v_{t_i}-v_{t_{i-1}}). \tag{3.9}$$

General weighted AMCE of θ is defined as

$$\widetilde{\theta}_{n,T} := -\left\{\frac{2}{T}G_{n,T}\right\}^{-1}. \tag{3.10}$$

General weighted AMCE of ρ is defined as

$$\widetilde{\rho}_{n,T} := -\left\{\frac{2}{T}G_{1,n,T}\right\}^{-1}. \tag{3.11}$$

With $w_{t_i} = 1$, we obtain the forward AMCE of θ as

$$\widetilde{\theta}_{n,T,F} := -\left\{\frac{2}{T}I_{n,T}\right\}^{-1}. \tag{3.12}$$

With $w_{t_i} = 1$, we obtain the forward AMCE of ρ as

$$\widetilde{\rho}_{n,T,F} := -\left\{\frac{2}{T}I_{1,n,T}\right\}^{-1}. \tag{3.13}$$

With $w_{t_i} = 0$, we obtain the backward AMCE of θ as

$$\widetilde{\theta}_{n,T,B} := -\left\{\frac{2}{T}J_{n,T}\right\}^{-1}. \tag{3.14}$$

With $w_{t_i} = 0$, we obtain the backward AMCE of ρ as

$$\widetilde{\rho}_{n,T,B} := -\left\{\frac{2}{T}J_{1,n,T}\right\}^{-1}. \tag{3.15}$$

With $w_{t_i} = 0.5$, the simple symmetric AMCE of θ is defined as

$$\widetilde{\theta}_{n,T,s} := -\left\{\frac{1}{T}[I_{n,T} + J_{n,T}]\right\}^{-1}. \tag{3.16}$$

Note that

$$\widetilde{\theta}_{n,T,s} = -\left\{\frac{2}{T}\sum_{i=2}^{n}\widetilde{Q}^2_{t_{i-1}}(v_{t_i} - v_{t_{i-1}}) + 0.5(\widetilde{Q}^2_{t_0} + \widetilde{Q}^2_{t_n})(v_{t_i} - v_{t_{i-1}})\right\}^{-1}. \tag{3.17}$$

With $w_{t_i} = 0.5$, the simple symmetric AMCE of ρ is defined as

$$\widetilde{\rho}_{n,T,s} := -\left\{\frac{1}{T}[I_{1,n,T} + J_{1,n,T}]\right\}^{-1} \tag{3.18}$$

Note that

$$\widetilde{\rho}_{n,T,s} = -\left\{\frac{2}{T}\sum_{i=2}^{n}\widetilde{V}_{t_{i-1}}^2(v_{t_i}-v_{t_{i-1}}) + 0.5(\widetilde{V}_{t_0}^2+\widetilde{V}_{t_n}^2)(v_{t_i}-v_{t_{i-1}})\right\}^{-1}. \tag{3.19}$$

With the weight function

$$w_{t_i} = \begin{cases} 0 & : \ i=1 \\ \frac{i-1}{n} & : \ i=2,3,\cdots,n \\ 1 & : \ i=n+1 \end{cases} \tag{3.20}$$

the weighted symmetric AMCE of θ is defined as

$$\widetilde{\theta}_{n,T,w} := -\left\{\frac{2}{T}\sum_{i=2}^{n}\widetilde{Q}_{t_{i-1}}^2(v_{t_i}-v_{t_{i-1}}) + \frac{1}{n}\sum_{i=1}^{n}\widetilde{Q}_{t_{i-1}}^2(v_{t_i}-v_{t_{i-1}})\right\}^{-1}. \tag{3.21}$$

and the weighted symmetric AMCE of ρ is defined as

$$\widetilde{\rho}_{n,T,w} := -\left\{\frac{2}{T}\sum_{i=2}^{n}\widetilde{V}_{t_{i-1}}^2(v_{t_i}-v_{t_{i-1}}) + \frac{1}{n}\sum_{i=1}^{n}\widetilde{V}_{t_{i-1}}^2(v_{t_i}-v_{t_{i-1}})\right\}^{-1}. \tag{3.22}$$

Note that estimators (3.21) and (3.22) are analogous to the trapezoidal rule in numerical analysis. One can instead use the midpoint rule to define the estimators

$$\widetilde{\theta}_{n,T,A} := -\left\{\frac{2}{T}\sum_{i=1}^{n}\left(\frac{\widetilde{Q}_{t_{i-1}}+\widetilde{Q}_{t_i}}{2}\right)^2(v_{t_i}-v_{t_{i-1}})\right\}^{-1}. \tag{3.23}$$

$$\widetilde{\rho}_{n,T,A} := -\left\{\frac{2}{T}\sum_{i=1}^{n}\left(\frac{\widetilde{V}_{t_{i-1}}+\widetilde{V}_{t_i}}{2}\right)^2(v_{t_i}-v_{t_{i-1}})\right\}^{-1}. \tag{3.24}$$

One can use the Simpson's rule to define the following estimators where the denominator is a convex combination of the midpoint and the trapezoidal rule:

$$\widetilde{\theta}_{n,T,S} := -\left\{\frac{1}{3T}\sum_{i=1}^{n}\left\{\widetilde{Q}_{t_{i-1}}^2 + 4\left(\frac{\widetilde{Q}_{t_{i-1}}+\widetilde{Q}_{t_i}}{2}\right)^2 + \widetilde{Q}_{t_i}^2\right\}(v_{t_i}-v_{t_{i-1}})\right\}^{-1}. \tag{3.25}$$

$$\widetilde{\rho}_{n,T,S} := -\left\{\frac{1}{3T}\sum_{i=1}^{n}\left\{\widetilde{V}_{t_{i-1}}^2 + 4\left(\frac{\widetilde{V}_{t_{i-1}}+\widetilde{V}_{t_i}}{2}\right)^2 + \widetilde{V}_{t_i}^2\right\}(v_{t_i}-v_{t_{i-1}})\right\}^{-1}. \tag{3.26}$$

Finally, the Durbin-Watson statistics are given by

$$\widetilde{D}_{n,T,F} := 2(1-\widetilde{\rho}_{n,T,F}) \tag{3.27}$$

and

$$\widetilde{D}_{n,T,S} := 2(1 - \widetilde{\rho}_{n,T,S}). \tag{3.28}$$

The following theorem shows that asymptotic normality of the AMCEs need $T \to \infty$ and $\frac{T}{\sqrt{n}} \to 0$.

Theorem 3.1 Denote $r_{n,T} := T^{-1/2}(\log T)^{1/2} \vee (\frac{T^4}{n^2})(\log T)^{-1})$.

$$\text{(a)} \quad \sup_{x\in\mathbb{R}} \left| P\left\{ \left(-\frac{T}{2\theta}\right)^{1/2} (\widetilde{\theta}_{n,T,F} - \theta^*) \le x \right\} - \Phi(x) \right| = O(r_{n,T}),$$

$$\text{(b)} \quad \sup_{x\in\mathbb{R}} \left| P\left\{ I_{n,T}^{1/2}(\widetilde{\theta}_{n,T,F} - \theta^*) \le x \right\} - \Phi(x) \right| = O(r_{n,T}),$$

$$\text{(c)} \quad \sup_{x\in\mathbb{R}} \left| P\left\{ \left(\frac{T}{2|\widetilde{\theta}_{n,T,F}|}\right)^{1/2} (\widetilde{\theta}_{n,T,F} - \theta^*) \le x \right\} - \Phi(x) \right| = O(r_{n,T}),$$

$$\text{(d)} \quad \sup_{x\in\mathbb{R}} \left| P\left\{ \left(\frac{T}{\sigma^2_{\rho,H}}\right)^{1/2} (\widetilde{\rho}_{n,T,F} - \rho^*) \le x \right\} - \Phi(x) \right| = O(r_{n,T}),$$

$$\text{(e)} \quad \sup_{x\in\mathbb{R}} \left| P\left\{ I_{1,n,T}^{1/2}(\widetilde{\rho}_{n,T,F} - \rho^*) \le x \right\} - \Phi(x) \right| = O(r_{n,T}),$$

$$\text{(f)} \quad \sup_{x\in\mathbb{R}} \left| P\left\{ \left(\frac{T}{2|\widetilde{\rho}_{n,T,F}|}\right)^{1/2} (\widetilde{\rho}_{n,T,F} - \rho^*) \le x \right\} - \Phi(x) \right| = O(r_{n,T}),$$

$$\text{(g)} \quad \sup_{x\in\mathbb{R}} \left| P\left\{ \left(\frac{T}{2\sigma^2_{D,H}}\right)^{1/2} (\widetilde{D}_{n,T,F} - D^*) \le x \right\} - \Phi(x) \right| = O(r_{n,T}).$$

In the following theorem, we improve the bound on the error of normal approximation using a mixture of random and nonrandom normings. Thus asymptotic normality of the AMCEs need $T \to \infty$ and $\frac{T}{n^{2/3}} \to 0$ which are sharper than the bound in Theorem 3.1.

Theorem 3.2

$$\text{(a)} \quad \sup_{x\in\mathbb{R}} \left| P\left\{ I_{n,T}\left(-\frac{2\theta}{T}\right)^{1/2} (\widetilde{\theta}_{n,T,F} - \theta^*) \le x \right\} - \Phi(x) \right| = O\left(T^{-1/2} \vee \left(\frac{T^3}{n^2}\right)^{1/3} \right).$$

$$\text{(b)} \quad \sup_{x\in\mathbb{R}} \left| P\left\{ I_{1,n,T}\left(\frac{\sigma^2_{\rho,H}}{T}\right)^{1/2} (\widetilde{\rho}_{n,T,F} - \rho^*) \le x \right\} - \Phi(x) \right| = O\left(T^{-1/2} \vee \left(\frac{T^3}{n^2}\right)^{1/3} \right).$$

The following theorem gives stochastic bound on the error of approximation of the continuous MCE by AMCEs.

Theorem 3.3

$$\text{(a)}\ \ |\widetilde{\theta}_{n,T} - \widetilde{\theta}_T| = O_P\left(\frac{T}{n}\right)^{1/2}, \quad \text{(b)}\ \ |\widetilde{\theta}_{n,T,s} - \widetilde{\theta}_T| = O_P\left(\frac{T^2}{n^2}\right)^{1/2},$$

$$\text{(c)}\ \ |\widetilde{\rho}_{n,T} - \widetilde{\rho}_T| = O_P\left(\frac{T}{n}\right)^{1/2}, \text{(d)}\ \ |\widetilde{\rho}_{n,T,s} - \widetilde{\rho}_T| = O_P\left(\frac{T^2}{n^2}\right)^{1/2}.$$

The proofs of Theorem 3.1, 3.2 and 3.3 are analogous to the respective proofs of Theorem 2.1, 2.2 and 2.3 in Bishwal (2011a) and hence we omit the details.

4. Examples: Factor Models

4.1. *Discrete Factor Models*

The Capital Asset Pricing Model (CAPM), introduced by Sharpe (1964), uses various assumptions about the markets and investors' behavior to establish a set of equilibrium conditions that allow investors to predict the return of an asset for its level of systematic risk. Systematic risk is also known as the aggregate, non-diversifiable, or market risk. The CAPM uses a measure of systematic risk that can be compared with other assets in the market. This measure of risk allows investors to improve their portfolios and to find their required rate of return.

The Capital Asset Pricing Model is based on the following assumptions:

A1) The Market prices are in equilibrium, i.e., supply equals demand.

A2) All investors are rational and risk-averse.

A3) All investors have the same forecast of expected returns and risk.

A4) All investors are price takers, i.e., they cannot influence the prices.

A5) Markets are frictionless (the borrowing rate is equal to the lending rate).

A6) All information is available at the same time to all investors.

A7) Trade without transaction or taxation costs.

A8) All assets are perfectly divisible and liquid.

A9) All investors are broadly diversified across a range of investments.

Under an efficient market hypothesis (EMH), which states that financial markets are *informationally efficient*, all the previously mentioned assumptions are reasonable and valid. The validity of the CAPM can only be guaranteed if all of these assumptions are true.

The Capital Market Line (CML) is the tangent line drawn from the risk free point to the feasible region for risky assets. CML relates the excess expected return on an efficient portfolio to its risk. Excess expected return is the difference between the expected return from the risk-free rate of the return. It is also called the *risk premium.* Let R be the return on the given efficient portfolio, i.e., the mixture of the Market portfolio and the risk-free asset. The equation for the CML is given by:

$$\mu_R = \mu_f + (\mu_M - \mu_f)\frac{\sigma_R}{\sigma_M} \tag{4.1}$$

Here μ_R is the expected return of the portfolio, $\mu_f$is the risk-free rate of return and μ_M is the expected return of the market portfolio, σ_M is the standard deviation of the return on the market portfolio, and σ_R is the standard deviation of the return on the portfolio. The excess expected return of a portfolio R is $\mu_R - \mu_f$ and the expected return of the market portfolio is $\mu_M - \mu_f$. In (4.1) μ_f, μ_M, and σ_M are constants. What varies are σ_R and μ_R. These vary as we change the efficient portfolio R.

The slope of the CML is given by $(\mu_M - \mu_f)/\sigma_M$ which can be interpreted as the ratio of the risk premium to the standard deviation of the market portfolio. This is also known as *Sharpe's reward –to-risk ratio*. The equation (4.1) can be rewritten as

$$(\mu_R - \mu_f)/\sigma_R = (\mu_M - \mu_f)/\sigma_M$$

which says that the reward-to-risk ratio for any efficient portfolio equals to the reward-to-risk ratio for the market portfolio. The CML is easy to derive. Consider an efficient portfolio that assigns a proportion weight (ω) of its assets to the market portfolio and $(1 - \omega)$ to the risk-free assets. Then

$$R = \omega R_M + (1 - \omega)\mu_f = \mu_f + \omega(R_M - \mu_f) \tag{4.2}$$

Therefore, taking expectation both sides in (4.2), we obtain

$$\mu_R = \mu_f + \omega(\mu_M - \mu_f) \tag{4.3}$$

Also from (4.2)

$$\sigma_R = \omega\sigma_M \tag{4.4}$$

which gives

$$\omega = \sigma_R/\sigma_M. \tag{4.5}$$

Substituting (4.5) into (4.3) gives the CML. One can view $\omega = \sigma_R/\sigma_M$ as an index of the risk aversion of the investor. The smaller the value of ω, the more risk averse the investor. If an investor places zero weight ($\omega = 0$) on an asset, then that investor is 100% in risk-free assets. Similarly, if an investor places 100% weight ($\omega = 1$) on an asset, then that investor is totally invested in the tangency portfolio of risky assets. The plot of reward as quantified by expected return versus the risk as measured by standard deviation of the return is a parabola. The left most point on the parabola achieves the minimum value of the risk and is called the *minimum variance portfolio*. The Sharpe ratio is the ratio of the reward to the risk. A line joining the risk-free rate to the parabola of reward versus risk with a large slope gives a higher expected return for a given level of risk. Sharpe ratio is the slope of the line. Thus larger the Sharpe ratio, the better is the expected return regardless of what level of risk one is willing to accept. The portfolio with highest Sharpe ratio is called *Tangency Portfolio*. The *efficient frontier* is the part of the parabola that has an expected return at least as large as the minimum variance portfolio. This is the optimal or efficient portfolio for the purpose of mixing the tangency portfolio of two risky assets with the risk-free asset. Each efficient portfolio has two properties:

1) It has a higher expected return than any other portfolio with the same (or smaller) risk.

2) It has a smaller risk than any other portfolio with the same (or smaller) expected return.

Therefore, we could only improve (reduce) the risk of an efficient portfolio by accepting a worse (smaller) expected return, and we can only improve (increase) the expected return of an efficient portfolio by accepting worse (higher) risk.

The Security Market Line (SML) relates the excess return on an asset to the slope of its regression on the market portfolio. Suppose that there are many securities indexed by j. Define

$$\beta_j = \sigma_{jM}/\sigma_M^2 \tag{4.6}$$

where σ_{jM} is the covariance between the returns on the j-th security and the market portfolio. By using returns of the market portfolio as the predictor variable, β_j is the slope of the best linear predictor of the j-th security's return. Another way to interpret the significance of β_j is based on linear regression. Linear regression is a method for estimating the coefficients of the best linear predictor based upon data. To apply linear regression, suppose that we have a bivariate time series $(R_{j,t}, R_{M,t})_{t=1}^n$ of returns on the j-th asset and the market portfolio. Then the estimated slope of the linear regression of $R_{j,t}$ on $R_{M,t}$ is

$$\hat{\beta}_j = \left(\sum_{t=1}^{n}(R_{j,t} - \bar{R}_j)(R_{M,t} - \bar{R}_M)\right) / \sum_{t=1}^{n}(R_{M,t} - \bar{R}_M)^2. \tag{4.7}$$

After multiplying the numerator and the denominator by n^{-1}, $\hat{\beta}_j$ becomes an estimate of σ_{jM} divided by an estimate of σ_M^2 and therefore the expression (4.7) is an estimate of β_j. The Security Market Line (SML) is given by:

$$\mu_j - \mu_f = \beta_j(\mu_M - \mu_f) \tag{4.8}$$

where μ_j is the expected return on the j-th security and $\mu_j - \mu_f$ is the risk premium for that security. Here, β_j is the independent variable in the linear equation, not the slope; more precisely, μ_j is a linear function of β_j with slope $(\mu_M - \mu_f)$. The SML says that the risk premium of the j-th asset is the product of its beta and the risk premium of the market portfolio. Therefore, β_j measures both the riskiness of the j-th asset and the reward for assuming that risk. When $\beta_j = 0$, a risky asset that is uncorrelated with the market portfolio will have an expected rate of return equal to the risk free rate. There is no expected excess return over risk-free asset even the investor bears some risk in holding a risky asset with zero beta. When $\beta_j = 1$, a risky asset which is perfectly correlated with the market portfolio has the same expected rate of return as the market portfolio. When $\beta_j > 1$, the expected excess rate of return is higher than the market portfolio, also known as an aggressive asset. When $\beta_j < 1$, the asset is said to be non-aggressive or risk averse. When $\beta_j < 0$, a risky asset with negative beta reduces the variance of the portfolio. This risk reduction potential of asset with negative β_j is something like paying premium to reduce risk.

The primary difference between the SML and the CML is that the SML applies to any asset but the CML applies only to the return of an efficient portfolio. It can be arranged so as to relate the excess expected return of that portfolio to be excess expected return of the market portfolio:

$$\mu_R - \mu_f = (\sigma_R/\sigma_M)(\mu_M - \mu_f). \tag{4.9}$$

The SML applies to any asset like the CML relates its excess expected return to the excess expected return of the market portfolio:

$$\mu_j - \mu_f = \beta_j(\mu_M - \mu_f) \tag{4.10}$$

If we take an efficient portfolio and consider it as an asset, then μ_R and μ_j both denote the expected return on that portfolio/asset. Both (4.9) and (4.10) hold so that

$$\sigma_R/\sigma_M = \beta_R. \tag{4.11}$$

The Security Characteristic Line (SCL) is a representation of the SML in the form of a regression model. The characteristic line takes the form of a straight line with the y-axis intercept representing the return of a security in excess of the risk-free return and the x-axis representing that for a portfolio made up of all assets in the market. The values that make up the characteristic line are obtained by performing a statistical regression analysis.

The Security Characteristic Line (SCL) is given as

$$R_{j,t} = \mu_{f,t} + \beta_j(R_{M,t} - \mu_{f,t}) + \epsilon_{j,t} \tag{4.12}$$

where $\epsilon_{j,t}$ is $N(0, \sigma^2_{\epsilon,j})$. It is often assumed that the $\epsilon_{j,t}$ s are uncorrelated across assets, that is, that $\epsilon_{j,t}$ is uncorrelated with $\epsilon_{j',t}$, for $j \neq j'$. This assumption has important ramifications for risk reduction by portfolio diversification. The SML gives us information about expected returns, but not about the variance of the returns. The characteristic line gives us a probability model of the returns, not just a model of their expected values. Hence, the characteristic line is said to be a return generating process. The similarity between the SML and characteristic line is that the regression line $E(Y|X) = \beta_0 + \beta_1 X$ gives the expected value of Y given X but not the conditional probability distribution of Y given X. The regression model $Y_t = \beta_0 + \beta_1 X_t + \epsilon_t$, and $\epsilon_t \sim N(0, \sigma^2)$ gives us this conditional probability distribution. The characteristic line implies that

$$\sigma_j^2 = \beta_j^2\sigma_M^2 + \sigma^2_{\epsilon,j}. \tag{4.13}$$

The total risk of the j-th asset is

$$\sigma_j = (\beta_j^2\sigma_M^2 + \sigma^2_{\epsilon,j})^{1/2}. \tag{4.14}$$

The risk has two components: $\beta_j^2\sigma_M^2$ is called the market or systematic component of risk and $\sigma^2_{\epsilon,j}$ is called non-market, unsystematic, or unique component of risk.

The security characteristic line (4.12) is a regression model

$$R_{j,t} = \mu_{f,t} + \beta_j(R_{M,t} - \mu_{f,t}) + \epsilon_{j,t} \tag{4.15}$$

The variable $(R_{M,t} - \mu_{f,t})$ is called a factor, which is the excess return on the market. CAPM is a single-factor model with the market index being the factor. In CAPM, the market factor is the only source of risk besides the unsystematic risk of each asset. Factor models generalize CAPM by allowing more factors than simply the market risk and the unique risk of each asset. A multifactor model is

$$R_{j,t} = \mu_{f,t} + \beta_{0,j} + \beta_{1,j}F_{1,t} \ldots + \beta_{p,j}F_{p,t} + \epsilon_{j,t} \tag{4.16}$$

where $F_{1,t}, \ldots, F_{p,t}$ are the values of p factors at time t. An example of factors are the stock market average, dividend yield of S&P 500, a measure of the risk of corporate bonds, interest rate variables, employment rate, inflation rate, monthly growth rate in industrial production, etc. The factors are to be correlated among themselves and are meant to simplify and reduce the amount of randomness required in an analysis of the assets. The factors are chosen by the modeler and depend on the type of assets being considered. With enough factors all commonalities between assets should be accounted for in the models. Then the error term should represent factors truly unique to the individual assets and therefore should be uncorrelated across j, i.e., $cov(\epsilon_{j,t}, \epsilon_{j',t}) = 0, j \neq j'$. The multi-factor model is a multiple linear regression model. If a factor model includes the returns on the market portfolio and other factors, then according to the CAPM, the betas on the other factors should be zero, so testing that they are zero also tests the CAPM.

With $p = 2$, one has the two-factor model:

$$R_{j,t} = \mu_{f,t} + \beta_{0,j} + \beta_{1,j}(R_{M,t} - \mu_{f,t}) + \beta_{2,j}F_{2,t} + \epsilon_{j,t}. \tag{4.17}$$

Fama and French (1995) introduced the Three-Factor Model which is given by

$$R_{j,t} = \mu_{f,t} + \beta_{0,j} + \beta_{1,j}(R_{M,t} - \mu_{f,t}) + \beta_{2,j}F_{2,t} + \beta_{3,j}F_{3,t} + \epsilon_{j,t} \tag{4.18}$$

where the factors F_2 and F_3 are SML and HML, respectively. The factor SML (small minus large) is the difference in returns of portfolio of small stocks and portfolio of large stocks and the factor HML (high minus low) is the difference in returns on a portfolio of high book-to-market value (BE/ME) stocks and a portfolio of low BE/ME stocks. Multifactor models refer to model parsimony. A statistical model should have as few parameters as possible like the CAPM model. One reasonable way of having few parameters is good is that each unknown parameter is another quantity that must be estimated and each estimate is a source of estimation error. On the other hand, a model must have enough parameters to adequately describe the behavior of data. A model with too few parameters can create bias as the model does not fit the data well. A model without excess parameters is parsimonious. Large models have less bias but more variability. Models with too few parameters and sizable bias underfit, while model with too many parameters overfit. Thus CAPM model underfits and Fama-French model is parsimonious. Autoregressive process driven by Autoregressive process was studied in Bercu and Proia (2013).

4.2. *Continuous Factor Models*

Barndorff-Nielsen and Shephard (2002) showed the good performance of superpositions of Ornstein-Uhlenbeck models for modeling stochastic volatility in exchange rate data. Nicolato and Venardos (2003) focused on derivative pricing based on Ornstein-Uhlenbeck volatility models and Benth (2011) showed their applicability in commodity markets.

Ornstein-Uhlenbeck process driven by Ornstein-Uhlenbeck process is related to stochastic volatility models, see Barndorff-Nielsen and Veraart (2013). Simple Ornstein-Uhlenbeck process has an exponentially declining autocorrelation function. Ornstein-Uhlenbeck process driven by Ornstein-Uhlenbeck process has a slowly decaying autocorrelation function compared to simple Ornstein-Uhlenbeck process. This generates long memory in the Ornstein-Uhlenbeck process. Parameter estimation for the fractional Black-Karasinski model of term structure of interest rates was studied in Bishwal (2022a).

Recently, Bishwal (2022b) introduced the hybrid asset price model with stochastic volatility, stochastic elasticity, stochastic interest rate and stochastic leverage under the risk neutral measure which is given by the following seven equations

$$dS_t = \chi X_t dt + \sqrt{V_{t-}} S_t dW_t + \rho_{\lambda t} dL_{\tau_{\lambda t}}, \tag{4.19}$$

$$dV_t = -\lambda V_t dt + \upsilon_{\lambda t-} d\Lambda_{\tau_{\lambda t}}, \tag{4.20}$$

$$dX_t = \alpha(\beta - X_t)dt + \sigma X_t^{\gamma_t} dN_t^H, \tag{4.21}$$

$$d\rho_t = ((2\zeta - \eta) - \eta\rho_t)dt + \theta\sqrt{(1+\rho_t)(1-\rho_t)}dZ_t, \tag{4.22}$$

$$d\xi_t = \kappa(\mu - \xi_t)dt + \varsigma\sqrt{\xi_t}dB_t, \tag{4.23}$$

$$d\gamma_t = \varpi(\psi - \gamma_t)dt + \delta\sqrt{\gamma_t}dM_t, \tag{4.24}$$

$$d\tau_t = \xi_{t-}dt \tag{4.25}$$

where L_t is a Levy process, Λ_t is a fractional Levy process, N^H is a sub-fractional Brownian motion, W_t, B_t, Z_t and M_t are correlated standard Brownian motions. Here S_t is the asset price which a geometric jump-diffusion, V_t is the stochastic volatility which is a Levy Ornstein-Uhlenbeck process, X_t is the stochastic interest rate which is a sub-fractional Chan-Karloyi-Longstaff-Sanders (CKLS) process (see Chan *et al.* (1992)), ρ_t is the stochastic leverage Jacobi (Beta) process, ξ_t is a volatility modulation of the driving Levy subordinator which is a Cox-Ingersoll-Ross (CIR) process (see Cox *et al.* (1985)), τ_t is the stochastic time change of the driving Levy subordinator which is a time integral of the CIR process, γ_t is the stochastic elasticity model which is a Cox-Ingersoll-Ross process, υ is the volatility of volatility which is independent of L, and all the 15 parameters $\lambda, \alpha, \beta, \sigma, \zeta, \eta, \theta, \kappa, \mu, \varsigma, \varpi, \psi, \delta, \chi, H$ in the model are positive.

Estimating 15 parameters is a daunting task. In this paper, we made an attempt to estimate the mean reversion speed parameter and the correlation parameter for fractional Brownian motion driver without jump component. Estimation of the remaining 13 parameters in the hybrid model remains open at this stage.

Concluding Remarks

1) The MCEs are more robust and efficient in comparison to the MLEs.

2) Let $T/n = h$. The estimators $\widetilde{\theta}_{n,T}$ and $\widetilde{\rho}_{n,T}$ have error order $O_P(h^{1/2})$ which are similar to the Euler scheme. The estimators $\widetilde{\theta}_{n,T,s}$ and $\widetilde{\rho}_{n,T,s}$ have error order $O_P(h)$ which are similar to the Milstein scheme.

3) It would be interesting to extend the paper to fractional Levy Ornstein-Uhlenbeck process as the driving process, see Bishwal (2011b).

Acknowledgement

The author would like to thank the Editor-in-Chief and the anonymous referee for their comments, which have enhanced the quality of the work.

References

[1] Barndorff-Nielsen, O.E. and Shephard, N. (2001) : Non-Gaussian Ornstein–Uhlenbeck-based models and some of their uses in financial economics (with discussion), *Journal of the Royal Statistical Society, Series B,* **63** 167-241.

[2] Barndorff-Nielsen, O.E. and Shephard, N. (2002) : Econometric analysis of realised volatility and its use in estimating stochastic volatility models, *Journal of the Royal Statistical Society, Series B,* **64**, 253-280.

[3] Barndorff-Nielsen, O.E. and Veraart, A. (2013) : Stochastic volatility of volatility and variance risk premia, *Journal of Financial Econometrics* **11** (1), 1-46.

[4] Benth, F.E. (2011) : The stochastic volatility model of Barndorff-Nielsen, O.E. and Shephard in commodity markets, *Mathematical Finance* **21**, 595-625.

[5] Bercu, B., Coutin, L. and Savy, N. (2010) : Sharp large deviations for the fractional Ornstein-Uhlenbeck process, **55** (4), 575-610.

[6] Bercu, B. and Proia, F. (2013) : A sharp analysis on the asymptotic behavior of Durbin-Watson statistic for the first-order autoregressive process, *ESAIM Probab. Stat.* **17** (1), 500-530.

[7] Bercu, B., Proia, F. and Savy, N. (2014) : On Ornstein-Uhlenbeck driven by Ornstein-Uhlenbeck processes, *Statist. Probab. Letters* **85**, 36-44.

[8] Bishwal, J.P.N. (2008) : *Parameter Estimation in Stochastic Differential Equations*, Lecture Notes in Mathematics **1923**, Springer-Verlag, Berlin.

[9] Bishwal, J.P.N. (2011a) : Minimum contrast estimation in fractional Ornstein-Uhlenbeck process: continuous and discrete sampling, *Fractional Calculus and Applied Analysis* **14** (3), 375-410.

[10] Bishwal, J.P.N. (2011b) : Maximum quasi-likelihood estimation in fractional Levy stochastic volatility model, *Journal of Mathematical Finance* **1** (3), 12-15.

[11] Bishwal, J.P.N. (2022a) : Berry-Esseen inequalities for the fractional Black-Karasinski model of term structure of interest rates, *Monte Carlo Methods and Applications*, **28** (2), 111-124.

[12] Bishwal, J.P.N. (2022b) : *Parameter Estimation in Stochastic Volatility Models*, Springer Nature, Cham.

[13] Bishwal, J.P.N. (2023) : Hypothesis testing in nonergodic fractional Ornstein-Uhlenbeck models, *European Journal of Statistics* **3** (6), 1-15.

[14] Chan, K.C., Karloyi, G.A., Longstaff, F.A. Sanders, A.B. (1992) : An empirical compar-

ison of the short termm interest rates, *Journal of Finance* **47**, 1209-1227.

[15] Cox, J.C., Ingersoll, J.E. and Ross, S.A. (1985) : The theory of term structure of interest rates, *Econometrica* **53**, 363-384.

[16] Durbin, J. and Watson, G.S. (1950) : Testing for serial correlation in least squares regression I, *Biometrika* **37**, 409-428.

[17] Durbin, J. and Watson, G.S. (1951) : Testing for serial correlation in least squares regression II, *Biometrika* **38**, 159-178.

[18] Durbin, J. and Watson, G.S. (1971) : Testing for serial correlation in least squares regression III, *Biometrika* **58**, 1-19.

[19] Fama, E.F. and French, K.R. (1995) : Size and book-to-market factors in earnings and returns, *Journal of Finance* **50**, 131-155.

[20] Feigin, P.D. (1976) : Maximum likelihood estimation for continuous time stochastic processes, *Adv. Appl. Prob.* **8**, 712-736.

[21] Feigin, P.D. (1979) : Some comments concerning a curious singularity, *J. Appl. Probab.* bf 16 (2), 440-444.

[22] Kleptsyna, M.L. and Le Breton, A. (2002) : Statistical analysis of the fractional Ornstein-Uhlenbeck type process, *Statistical Inference for Stochastic Processes* **5** (3), 229-248.

[23] Liptser, R.S. and Shiryayev, A.N. (1978) : *Statistics of Random Processes II : Applications* Springer-Verlag, Berlin.

[24] Nicolato, E. and Venardos, E. (2003) : Option pricing in stochastic volatility models of Ornstein-Uhlenbeck type, *Mathematical Finance* **13**, 445-466.

[25] Norros, I., Valkeila, E. and Virtamo, J. (1999) : An elementary approach to a Girsanov formula and other analytical results on fractional Brownian motion, *Bernoulli* **5**, 571-587.

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

[27] Tudor, C.A. and Viens, F. (2007) : Statistical aspects of the fractional stochastic calculus, *Annals of Statistics*, **35** (3), 1183-1212.