1-2019

Local Lagged Adapted Generalized Method of Moments: An Innovative Estimation and Forecasting Approach and its Applications.

Olusegun Michael Otunuga  
*Marshall University*, otunuga@marshall.edu

Gandaram S. Ladde

Nathan G. Ladde

Follow this and additional works at: https://mds.marshall.edu/mathematics_faculty

Part of the *Dynamical Systems Commons*

Recommended Citation


This Article is brought to you for free and open access by the Mathematics at Marshall Digital Scholar. It has been accepted for inclusion in Mathematics Faculty Research by an authorized administrator of Marshall Digital Scholar. For more information, please contact zhangj@marshall.edu, beachgr@marshall.edu.
Local Lagged Adapted Generalized Method of Moments: An Innovative Estimation and Forecasting Approach and its Applications

OLUSEGUN M. OTUNUGA
Department of Mathematics and Statistics, Marshall University, One John Marshall Dr, Huntington, West Virginia, 25755, USA.

GANGARAM S. LADDE
Department of Mathematics and Statistics, University of South Florida, 4202 E. Fowler Avenue, Tampa, Florida, 33620, USA.

NATHAN G. LADDE
CSM, 5775 Glenridge Dr NE, Atlanta, Georgia, 30328, USA.

Abstract

In this work, an attempt is made to apply the Local Lagged Adapted Generalized Method of Moments (LLGMM) to estimate state and parameters in stochastic differential dynamic models. The development of LLGMM is motivated by parameter and state estimation problems in continuous-time nonlinear and non-stationary stochastic dynamic model validation problems in biological, chemical, engineering, energy commodity markets, financial, medical, physical and social sciences. The byproducts of this innovative approach (LLGMM) are the balance between model specification and model prescription of continuous-time dynamic process and the development of discrete-time interconnected dynamic model of local sample mean and variance statistic process (DTIDMLSMVSP). Moreover, LLGMM is a dynamic non-parametric method. The DTIDMLSMVSP is an alternative approach to the GARCH (1,1) model, and it provides an iterative scheme for updating statistic coefficients in a system of generalized method of moment/observation equation. Furthermore, applications of LLGMM to energy commodities price, U.S. Treasury Bill interest rate and the U.S.-U.K. foreign exchange rate data strongly exhibit its unique role, scope and performance, in particular, in forecasting and confidence-interval problems in applied statistics.

Keywords: Conceptual computational/theoretical parameter estimation scheme; Method of Moments; Nonparametric; Simulation; Forecasting; Mean Square Optimal Procedure; Reaction/response time delay.

Mathematics Classification Code : 37M05; 37M10; 62M10; 62G05; 62P05
1. Introduction

For the past 40 years, researchers [3, 9, 11, 15, 17, 18, 19, 42, 31, 32, 35, 36, 38, 39, 40, 41] have given a lot of attention to estimating continuous-time dynamic models from discrete time data sets. The Generalized Method of Moments (GMM) developed by Hansen [17] and its extensions [11, 18, 19] have played a significant role in literature related to the parameter and state estimation problems in linear and nonlinear stochastic dynamic processes.

Most of the existing parameter and state estimation techniques are centered around the usage of either overall data sets [11, 18, 19], batched data sets [7], or local data sets [39] drawn on an interval of finite length $T$. This leads to an overall parameter estimate on the interval.

In this paper, recently developed method referred to as an innovative method, called the "Local Lagged Adapted Generalized Method of Moments" (LLGMM) [30] is used to estimate state and parameters in stochastic differential dynamic models. The LLGMM approach [30] is composed of seven interconnected components: (1) development of stochastic mathematical model of continuous time dynamic process [23, 24], (2) development of the discrete-time interconnected dynamic model for statistic process, (3) utilizing the Euler-type discretized scheme [21] for nonlinear and non-stationary system of stochastic differential equations (1), (4) employing lagged adaptive expectation process [33] for developing generalized method of moment/observation equations, (5) introduction of the conceptual and computational parameter estimation problem, (6) formulation of the conceptual and computational state simulation scheme and (7) defining the mean square $\epsilon$-sub-optimal procedure.

In fact, the LLGMM approach [30] is also applicable to apparently different dynamic processes that are in actuality conceptually similar dynamic processes in biological, chemical, engineering, financial, medical, physical and social sciences. Moreover, one of the goals of the parameter and state estimation problems is for model validation rather than model misspecification [11]. For the continuous-time dynamic model validation, we utilize existing real world data. The real world time varying data is drawn/recorded at discrete-time on a time interval of finite length. Because of this, instead of using existing econometric specification/Euler-type numerical scheme, we construct the stochastic numerical approximation scheme [21] using continuous time stochastic differential equations. In real world dynamic modeling problems [23, 24, 25, 33, 34], future states of continuous time dynamic processes are influenced by the past state history. This is due to the influence of response/reaction time delay processes [25, 29, 30, 33]. The influence of state history, the concept of lagged adaptive expectation process [33], and the idea of a moving average [20] are incorporated in the development of the DTIDLMSMVSP (we refer readers to Lemma 2.1 of [30]). The presented approach is more suitable and robust for forecasting problems than existing methods. It also provides upper and lower bounds for the forecasted state of the system. Moreover, its computational aspect is a nested "two scale hierarchic" quadratic mean-square optimization process whereas the existing GMM and its extensions are "single-shot".
The organization of this paper is as follows.

In Section 2, we utilize the theoretical components (1)-(7) of the LLGMM method. For easy reference, we construct a local observation system from nonlinear stochastic functional differential equations. This is based on the Itô-Doob stochastic differential formula, Euler-type numerical scheme in the context of the original stochastic systems of differential equations and the given data. In addition, we briefly outline a procedure to estimate the state and parameters, locally. Using the LLGMM components (2), (3) and (4), conceptual computational iterative scheme, state and parameter estimation scheme, the simulation processes are coordinated with a real world data process in Section 3. This has led to generate the following concepts: (a) local admissible set of lagged sample/data/observation size, (b) local class of admissible lagged-adapted finite sequence of conditional sample/data, (c) local admissible sequence of parameter estimates and corresponding admissible sequence of simulated values, (d) $\epsilon$-best sub-optimal admissible subset of set of $m_k$-size local conditional samples at time $t_k$ in (a), (e) $\epsilon$-sub-optimal lagged-adapted finite sequence of conditional sample/data, and (f) finally, the $\epsilon$-best sub-optimal parameter estimates and simulated value at time $t_k$ for $k = 1, 2, ..., N$. These are summarized in Section 3 in a systematic way. Moreover, the local lagged adaptive process and DTIDMLSMVSP generate a finite chain of discrete-time admissible sets/sub-data and a corresponding chain described by the simulation algorithm. Furthermore, in Section 4, the usefulness of the conceptual computational LLGMM algorithm is illustrated by applying the algorithm to energy commodity’s price, U.S. Treasury Bill interest rate and the U.S.-U.K. foreign exchange rate data for the state and parameter estimation problems. The graphical, simulation and statistical results as well as the goodness-of-fit measures are also outlined. In Section 5, the LLGMM is applied to investigate the forecasting and confidence-interval problems in applied statistics. The presented results show the long-run prediction exhibiting a degree of confidence. Moreover, it exhibits a wider role and scope to play in the 21st century. Because of the knowledge of nanotechnology coupled with the usage of advancements in electronic communication systems/tools which exhibit that almost everything is dynamic, highly nonlinear, non-stationary and operating under endogenous and exogenous processes, a multitude of applications of the proposed model exists. A few by-products of LLGMM namely: (a) the development of the second component, DTIDMLSMVSP and its component with the GARCH as well as Ex Post Volatility work in Section 6, (b) the Aggregated Generalized Method of Moments(AGMM) (described in Section 7 and Appendix Appendix B) and (c) Orthogonal Condition Based Generalized Method of Moments (OCBGMM-Analytic) are compared in Section 7. Using the average of locally estimated parameters in LLGMM, an aggregated generalized method of moment (AGMM) is also developed and applied to six data sets in Appendix B. In fact, in Section 7, we summarize a comparative study between LLGMM and the existing parametric OCBGMM techniques. The details are outlined in Appendix C. The LLGMM method exhibits superior performance to the existing and newly developed OCBGMM methods. The LLGMM method is independent and
dynamic. On the other hand, the OCBGMM method is highly dependent and static.

2. Theoretical Parametric Estimation Procedure

In this section, for the sake of completeness and easy reference, we outline the theoretical components (1), (3) and (4) of the LLGMM [30]. The outline is based on a mathematically rigorous theoretical state and parameter estimation procedure for any general continuous-time nonlinear and non-stationary stochastic dynamic model described by a system of stochastic differential equations [24]. As stated before, this work is not only motivated by the continuous-time dynamic model validation problem [29, 30] in the context of real data energy commodities, but also motivated by any continuous-time nonlinear and non-stationary stochastic dynamic model validation problems in biological, chemical, engineering, financial, medical, physical and social sciences. For the sake of comparison of the presented results, we also sketch the existing OCBGMM procedure [9, 10, 18, 19] that uses the entire time series data set for single-shot parameter and state estimates. It lacks the usage of Itô-Doob calculus, properties of stochastic differential equations and its connectivity with the usage of econometric specification based discretization scheme, orthogonality conditional vector and the quadratic form.

We consider a general system of stochastic differential equations under the influence of hereditary effects in both the drift and diffusion coefficients described by

\[
dy = f(t, y_t)dt + \sigma(t, y_t)dW(t), \quad y_0 = \varphi_0,
\]

(2.1)

where \(y_t(\theta) = y(t + \theta), \theta \in [-\tau, 0]\), \(f, \sigma : [0, T] \times C \rightarrow \mathbb{R}^q\) are Lipschitz continuous bounded functionals; \(C\) is a Banach space of continuous functions defined on \([-\tau, 0]\) into \(\mathbb{R}^q\) equipped with the supremum norm; \(W(t)\) is standard Wiener process defined on a complete filtered probability space \((\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, P)\); \(\varphi_0 \in C\); and \(y_0(t_0 + \theta)\) is \((\mathcal{F}_t)_0\) measurable; the filtration function \((\mathcal{F}_t)_{t \geq 0}\) is right-continuous, and each \(\mathcal{F}_t\) with \(t \geq t_0\) contains all \(P\)-null events in \(\mathcal{F}\); the solution process \(y(t_0, \varphi_0)(t)\) of (2.1) is adapted and non-anticipating with respect to \((\mathcal{F}_t)_{t \geq 0}\).

2.1. Transformation of System of Stochastic Differential Equations (2.1) [30]

Let \(V \in C([-\tau, \infty] \times \mathbb{R}^q, \mathbb{R}^m)\), and its partial derivatives \(V_t, \frac{\partial V}{\partial y}, \frac{\partial V}{\partial y^2}\) exist and continuous. We apply Itô-Doob stochastic differential formula [24] to \(V\) and obtain

\[
dV(t, y) = LV(t, y, y_t)dt + V_y(t, y)\sigma(t, y)W(t),
\]

(2.2)
where the $L$ operator is defined by

\[
\begin{align*}
LV(t, y, y_i) &= V(t, y) + V_y(t, y)\mathbf{f}(t, y_i) + \frac{1}{2}tr(V_{yy}(t, y)b(t, y_i)), \\
b(t, y_i) &= \sigma(t, y_i)\sigma^T(t, y_i).
\end{align*}
\] (2.3)

2.2. Euler-type Discretization Scheme for (2.1) and (2.2) [30]

For (2.1) and (2.2), we present the Euler-type discretization scheme [21]:

\[
\begin{align*}
\Delta y_i &= \mathbf{f}(t_{i-1}, y_{i-1})\Delta t_i + \sigma(t_{i-1}, y_{i-1})\Delta W(t_i), \\
\Delta V(t_i, y(t_i)) &= LV(t_{i-1}, y(t_{i-1}), y_{i-1})\Delta t_i + V_y(t_{i-1}, y(t_{i-1}))\sigma(t_{i-1}, y_{i-1})\Delta W(t_i), \quad i = 1, 2, ..., n.
\end{align*}
\] (2.4)

Define $\mathcal{F}_{t_{i-1}} \equiv \mathcal{F}_{t_{i-1}}$ as the filtration process up to time $t_{i-1}$.

2.3. Formation of Generalized Moment Equations From (2.4)

With regard to the continuous time dynamic system (2.1) and its transformed system (2.2), the more general moments of $\Delta y(t_i)$ are as follows:

\[
\begin{align*}
E[\Delta y(t_i)|\mathcal{F}_{t_{i-1}}] &= \mathbf{f}(t_{i-1}, y_{i-1})\Delta t_i, \\
E[(\Delta y(t_i) - E[\Delta y(t_i)|\mathcal{F}_{t_{i-1}}])(\Delta y(t_i) - E[\Delta y(t_i)|\mathcal{F}_{t_{i-1}}])^T|\mathcal{F}_{t_{i-1}}] &= \sigma(t_{i-1}, y_{i-1})\sigma^T(t_{i-1}, y_{i-1})\Delta t_i, \\
E[\Delta V(t_i, y(t_i))|\mathcal{F}_{t_{i-1}}] &= LV(t_{i-1}, y(t_{i-1}), y_{i-1})\Delta t_i, \\
E[(\Delta V(t_i, y(t_i)) - E[\Delta V(t_i, y(t_i))|\mathcal{F}_{t_{i-1}}])(\Delta V(t_i, y(t_i)) - E[\Delta V(t_i, y(t_i))|\mathcal{F}_{t_{i-1}}])^T|\mathcal{F}_{t_{i-1}}] &= B(t_{i-1}, y(t_{i-1}), y_{i-1}),
\end{align*}
\] (2.5)

where $B(t_{i-1}, y(t_{i-1}), y_{i-1}) = V_y(t_{i-1}, y(t_{i-1}))b(t_{i-1}, y_{i-1})V_y(t_{i-1}, y(t_{i-1}))\Delta t_i$, and $T$ is the transpose of the matrix.

2.4. Basis for Local Lagged Adaptive Discrete-time Expectation Process

For $i = 1, 2, ..., n$, it follows from (2.4) and (2.5) that

\[
\begin{align*}
\Delta y_i &= E[\Delta y(t_i)|\mathcal{F}_{t_{i-1}}] + \sigma(t_{i-1}, y_{i-1})\Delta W(t_i), \\
\Delta V(t_i, y(t_i)) &= E[\Delta V(t_i, y(t_i))|\mathcal{F}_{t_{i-1}}] + V_y(t_{i-1}, y(t_{i-1}))\sigma(t_{i-1}, y_{i-1})\Delta W(t_i), \quad i = 1, 2, ..., n.
\end{align*}
\] (2.6)

This provides the basis for the development of the concept of lagged adaptive expectation process [33] with respect to continuous time stochastic dynamic systems (2.1) and (2.2). This indeed leads to a formulation of $m_k$-local generalized method of moments at time $t_k$. 


2.5. Block Orthogonality Condition Vector for (2.1) and (2.2) [30]

From (2.6), we note that one can define a block vector of orthogonality condition [10] as

\[
H(t_{i-1}, y(t_i), y(t_{i-1})) = \begin{pmatrix}
\Delta y(t_i) - f(t_{i-1}, y(t_{i-1})) \Delta t_i \\
\Delta V(t_{i-1}, y(t_i)) - L V(t_{i-1}, y(t_{i-1}), y(t_{i-1})) \Delta t_i
\end{pmatrix}.
\] (2.7)

**Remark 2.1.** Using the LLGMM method described in [30], we attempt to estimate the state and parameters of the stochastic differential dynamic model of the type (2.1). This involves the construction of Euler-type discretization scheme for (2.1) and (2.2) in Sub-section 2.2, the formation of generalized moment equations from (2.4) in Sub-section 2.3 and the basis for local lagged adaptive discrete-time expectation process in Sub-section 2.4. All of these are in the framework of correct mathematical reasoning. Further, it is logical and interconnected/interactive within the context of the continuous-time dynamic system (2.1). The theoretical parameter estimation procedure in this section adapts to and incorporates the continuous-time changes in the state and parameters of the system and moves into a discrete-time theoretical numerical schemes in (2.4) as a model validation of (2.1). It further successively moves in the local moment equations within the context of local lagged adaptive, local discrete-time statistic and computational processes in a natural, systematic and coherent manner.

We illustrate this by applying the LLGMM method [30] to estimate state and parameters in stochastic differential dynamic models for energy commodity price, U.S. Treasury Bill interest rate, and the U.S.-U.K. foreign exchange rate data.

2.6. Illustration 1: Dynamic Model for Energy Commodity Price [29, 30]

We consider the stochastic dynamic model for energy commodities [29, 30] described by the following nonlinear stochastic differential equation

\[
dy = a(t)(\mu(t) - y)dt + \sigma(t, y)ydW(t), \quad y_0 = \varphi_0,
\] (2.8)

where \( y(t) = y(t + \theta); \theta \in [-\tau, 0] \), \( \mu, a : [t_0, T] \to \mathbb{R} \) are continuous functions; the initial process \( \varphi_0 = [y(t_0 + \theta)]_{\theta \in [-\tau, 0]} \) is \( \mathcal{F}_{t_0} \)-measurable and independent of \( \{W(t), t \in [t_0, T]\} \); \( W(t) \) is a standard Wiener process defined in (2.1); \( \sigma : [t_0, T] \times C \to \mathbb{R}^+ \) is a Lipschitz continuous and bounded functional; \( C \) is the Banach space of continuous functions defined on \([-\tau, 0]\) into \( \mathbb{R} \) equipped with the supremum norm.
The solution \( y(t) \) of (2.8) satisfies
\[
y(t) - y(t_0) = \int_{t_0}^{t} a(s) y(s)(\mu(s) - y(s)) ds + \int_{t_0}^{t} \sigma(s, y(s)) dW(s),
\]
and
\[
E \left[ y(t) - y(t_0) | \mathcal{F}_{t_0} \right] = \int_{t_0}^{t} a(s) y(s)(\mu(s) - y(s)) ds.
\]

**Transformation of Stochastic Differential Equation (2.8):** We pick a Lyapunov function \( V(t, y) = \ln(y) \) in (2.2) for (2.8). Using Itô-differential formula [24], we have
\[
d(\ln(y)) = \left[ a(t)(\mu(t) - y) - \frac{1}{2} \sigma^2(t, y) \right] dt + \sigma(t, y) dW(t), \quad (2.9)
\]

**The Euler-type Discretization Schemes for (2.8) and (2.9) [21, 30]:** By setting \( \Delta t = t_i - t_{i-1}, a_i = a(t_i), \mu_i = \mu(t_i), \sigma_i = \sigma(t_i), \Delta y_i = y(t) - y(t_{i-1}), \) the combined Euler discretized scheme for (2.8) and (2.9) is
\[
\begin{cases}
\Delta y_i &= a_{i-1}(t_{i-1} - y_{i-1}) \Delta t + \sigma(t_{i-1}, y_{i-1}) \Delta W(t_i), \quad y_0 = \varphi_0, \\
\Delta (\ln(y_i)) &= \left[ a_{i-1}(t_{i-1} - y_{i-1}) - \frac{1}{2} \sigma^2(t_{i-1}, y_{i-1}) \right] \Delta t + \sigma(t_{i-1}, y_{i-1}) \Delta W(t_i), \quad y_0 = \varphi_0.
\end{cases} \quad (2.10)
\]

where \( \varphi_0 = \{y_i\}_{i=1}^0 \) is a given finite sequence of \( \mathcal{F}_0 \)-measurable random variables, and it is independent of \( \{\Delta W(t_i)\}_{i=0}^N \).

**Generalized Moment Equations [30]:** Applying conditional expectation to (2.10) with respect to \( \mathcal{F}_{t_{i-1}} = \mathcal{F}_{t_{i-1}} \), we obtain
\[
\begin{align*}
E[\Delta y_i | \mathcal{F}_{t_{i-1}}] &= a_{i-1}(t_{i-1} - y_{i-1}) \Delta t, \\
E[\Delta (\ln(y_i)) | \mathcal{F}_{t_{i-1}}] &= \left[ a_{i-1}(t_{i-1} - y_{i-1}) - \frac{1}{2} \sigma^2(t_{i-1}, y_{i-1}) \right] \Delta t, \\
E[\Delta (\ln(y_i)) - E[\Delta (\ln(y_i)) | \mathcal{F}_{t_{i-1}}]^2 | \mathcal{F}_{t_{i-1}}] &= \sigma^2(t_{i-1}, y_{i-1}) \Delta t. \tag{2.11}
\end{align*}
\]

**Basis for Lagged Adaptive Discrete-time Expectation Process:** From (2.11), (2.10) reduces to
\[
\begin{cases}
\Delta y_i &= E[\Delta y_i | \mathcal{F}_{t_{i-1}}] + \sigma(t_{i-1}, y_{i-1}) y_{i-1} \Delta W(t_i), \\
\Delta (\ln(y_i)) &= E[\Delta (\ln(y_i)) | \mathcal{F}_{t_{i-1}}] + \sigma(t_{i-1}, y_{i-1}) \Delta W(t_i). \tag{2.12}
\end{cases}
\]

(2.12) provides the basis for the development of the concept of lagged adaptive expectation process [33] with respect to continuous time stochastic dynamic systems (2.8) and (2.9).

**Remark 2.2. Orthogonality Condition Vector for (2.8) and (2.9)**
Using (2.10), (2.11) and (2.12), we further remark that the orthogonality condition vector [10] with respect to continuous-time stochastic dynamic model (2.8) is represented by

\[
H(t_{i-1}, y(t_i), y(t_{i-1})) = \begin{bmatrix}
\Delta y(t_i) - a(t_{i-1})y(t_{i-1})(\mu(t_{i-1}) - y(t_{i-1}))\Delta t_i \\
\Delta \ln(y(t_i)) - L \ln(y(t_{i-1}), y_{t_{i-1}})\Delta t_i \\
(\Delta \ln(y(t_i)) - L \ln(y(t_{i-1}), y_{t_{i-1}})\Delta t_i)^2 - \sigma^2(t_{i-1}, y_{t_{i-1}})\Delta t_i
\end{bmatrix}, \tag{2.13}
\]

where \( L \ln(y(t_{i-1}), y_{t_{i-1}})\Delta t_i = \left( a(t_{i-1})(\mu(t_{i-1}) - y(t_{i-1})) - \frac{1}{2}\sigma^2(t_{i-1}, y_{t_{i-1}}) \right)\Delta t_i \). Moreover, unlike the orthogonality condition vector defined in the literature [8, 10, 38], this orthogonality condition vector is based on the discretization scheme (2.10) associated with nonlinear continuous-time stochastic differential equations (2.8) and (2.9) and the Itô-Doob stochastic differential calculus [21, 24]

**Local Observation System of Algebraic Equations [30]:** Following definition for \( k \in I_0(N) \), applying the LLGMM method [30] and using Definitions 2.3-2.7 of [30] together with the discretized form (2.12), we formulate a local observation/measurement process at time \( t_k \) as algebraic functions of \( m_k \)-local functions of restriction of the overall finite sample sequence \( y_{t_i} \) to a subpartition \( P_k \) in Definition 2.2 of [30]. Let \( a_{t_i} \) and \( \mu_{t_i} \) be estimates of \( a_i \) and \( \mu_i \), respectively, at each time \( t_i \). We have

\[
\begin{aligned}
\frac{1}{m_k} \sum_{i=k}^{k-1} \mathbb{E} [\Delta y_i | F_{t_{i-1}}] &= a_{t_i} \left[ \frac{\mu_{t_{i-1}}}{m_{k}} \sum_{i=k}^{k-1} y_{t_{i-1}} - \frac{1}{m_{k}} \sum_{i=k}^{k-1} y_{t_{i-1}} \right] \Delta t, \\
\frac{1}{m_k} \sum_{i=k}^{k-1} \mathbb{E} [\Delta \ln(y_i) | F_{t_{i-1}}] &= a_{t_i} \left[ \mu_{t_{i-1}} - \frac{1}{m_{k}} \sum_{i=k}^{k-1} y_{t_{i-1}} \right] \Delta t - \frac{1}{2m_k} \sum_{i=k}^{k-1} \mathbb{E} \left[ (\Delta \ln(y_i) - \mathbb{E} [\Delta \ln(y_i) | F_{t_{i-1}}])^2 | F_{t_{i-1}} \right], \\
\hat{\sigma}^2_{m_k} &= \left\{ \begin{array}{ll}
\frac{1}{m_k \Delta t} \sum_{i=k}^{k-1} \mathbb{E} \left[ (\Delta \ln(y_i)) - \mathbb{E} [\Delta \ln(y_i) | F_{t_{i-1}}])^2 | F_{t_{i-1}} \right] & \text{if } m_k \text{ is small} \\
\frac{1}{(m_{k-1}) \Delta t} \sum_{i=k}^{k-1} \mathbb{E} \left[ (\Delta \ln(y_i)) - \mathbb{E} [\Delta \ln(y_i) | F_{t_{i-1}}])^2 | F_{t_{i-1}} \right] & \text{if } m_k \text{ is large},
\end{array} \right.
\end{aligned}
\tag{2.14}
\]

where \( m_k \in I_0(r+k-1) = \{2, 3, ..., r+k-1\} \) is defined as the local admissible sample/data observation size at time \( t_k \) (Definition 3.3 [30]). Following Definitions (2.5-2.7) in [30], we define \( s_{m_k}^r \) and \( s_{m_k}^2 \) as the \( m_k \)-local average/mean and \( m_k \)-local variance, respectively, corresponding to a local sequence \( S_{m_k} = \{x_i\}_{k-m_k}^{k-1} \). From the third equation in (2.14), it follows that the average volatility square \( \hat{\sigma}^2_{m_k} \) is given by

\[
\hat{\sigma}^2_{m_k} = \frac{s_{m_k}^2}{\Delta t}, \tag{2.15}
\]

where \( s_{m_k}^2 \) is the local sample variance statistics for volatility at time \( t_k \) in the context of \( x(t_i) = \Delta \ln(y_i) \) satisfying the following discrete-time interconnected dynamic model of local sample mean \( \hat{S}_{m_k} \) and variance \( s_{m_k}^2 \) processes (DTIDMLSMVSP)
where $p$ is the order of the system (2.16) and

\[
\begin{align*}
\eta_{m_k,k-p} &= \frac{1}{m_{k+p}} \left[ \sum_{j=m_k-m_{k+p}+1}^{m_k-1} F^j x_{k-p} - F^{-m_k-p+1} x_{k-p} - F^{-m_k-p} x_{k-p} + F^0 x_{k-p} \right], \\
\epsilon_{m_k,k-1} &= m_{k-1} m_k \left[ \sum_{j=m_k-1}^{m_k-2} \left( \sum_{i=1}^{m_k-1} \frac{F^{i-1} x_{k-1}}{m_k-i} \right)^2 - \sum_{i=1}^{m_k-2} \left( \sum_{j=1}^{m_k-i} \frac{F^{j-1} x_{k-1}}{m_k-j} \right)^2 \right] \\
&\quad + m_{k-1} m_k \left[ \sum_{j=m_k-1}^{m_k-2} \left( \sum_{i=1}^{m_k-1} \frac{F^{j-i} x_{k-1}}{j} \right)^2 \right] \\
&\quad - \frac{1}{m_k} \sum_{l=0}^{m_k-1} F^l x_{k-1} F^s x_{k-1}, \\
\epsilon_{m_k,k-1} &= \frac{1}{m_{k+1}} \left[ \sum_{i=1}^{m_k} \left( \sum_{j=m_k-i}^{m_k-1} \frac{F^{j-1} x_{k-1}}{j} \right)^2 \right] \\
&\quad - \frac{1}{m_k} \sum_{l=0}^{m_k-1} F^l x_{k-1} F^s x_{k-1}.
\end{align*}
\]

For details, see [30] (Lemma 2.1) and [29].

Thus, by the application of the Implicit Function Theorem [2], we conclude that for every non-constant $m_k$-local sequence $\{x(t)\}_{t=m_k-m_0}$, there exists a unique solution $\hat{a}_{m_k,k} \equiv a_k$ and $\hat{\mu}_{m_k,k} \equiv \mu_k$ of system of algebraic equations (2.14) as a point estimates of $a(t)$ and $\mu(t)$, respectively, at time $t = t_k$, given by

\[
\begin{align*}
\hat{a}_{m_k,k} &= \left[ \frac{1}{m_k} \sum_{i=m_k}^{m_k-1} \Delta y_{i+1} + \sum_{i=1}^{m_k-1} \frac{y_{i+1} - y_i}{m_k-i} \right] \Delta t, \\
\hat{\mu}_{m_k,k} &= \left[ \frac{1}{m_k} \sum_{i=m_k}^{m_k-1} \Delta y_{i+1} + \sum_{i=1}^{m_k-1} \frac{y_{i+1} - y_i}{m_k-i} \right] \Delta t, \\
\hat{\sigma}_{m_k,k}^2 &= \hat{\sigma}_{m_k,k}^2 \Delta t.
\end{align*}
\]
Remark 2.3. We note that without loss of generality, the discrete-time data set \( \{y_i - r + i : i \in I_{(r - 1)} \} \) is assumed to be close to the true values of the solution process of the continuous-time dynamic process. In fact, this assumption is feasible in view of the uniqueness and continuous dependence of the solution process for stochastic functional or ordinary differential equation with respect to the initial data [24].

Remark 2.4. If the sample \( \{y_i\}_{k-1}^{k-1} \) is a constant sequence, then it follows from (2.18) and the fact that \( \Delta(\ln y_i) = 0 \) and \( s_{y_i}^2 = 0 \), that \( \hat{a}_{m,k} \to \frac{1}{m_k} \sum_{i=k-m_k}^{k-1} y_{i-1} \). Also, it follows from (2.14) that \( \hat{a}_{m,k} = 0 \).

Remark 2.5. As we stated before, the theoretically estimated parameters \( \hat{a}, \hat{\mu}, \) and \( \hat{\sigma}^2 \) depend upon the time at which a data point is drawn. This is what we expected because of the fact that nonlinearity of the dynamic model together with environmental stochastic perturbations generates a non-stationary solution process. Using locally estimated parameters of the continuous-time dynamic system, we can find the average of the local parameters over the entire size of data set as follows:

\[
\begin{align*}
\bar{a} &= \frac{1}{N} \sum_{i=0}^{N} a_{\hat{a}_{i,j}}, \quad \bar{\mu} = \frac{1}{N} \sum_{i=0}^{N} \mu_{\hat{\mu}_{i,j}}, \quad \bar{\sigma}^2 = \frac{1}{N} \sum_{i=0}^{N} \sigma_{\hat{\sigma}^2_{i,j}},
\end{align*}
\] (2.19)

where \( \bar{a}, \bar{\mu}, \) and \( \bar{\sigma}^2 \) are referred to as aggregated parameter estimates of \( a, \mu, \) and \( \sigma^2 \) over the given entire finite interval of time, respectively. Further detailed statistical analysis is outlined in Appendix B.

Remark 2.6. The discrete-time interconnected dynamic model for statistic process (DTIDMLSMVSP) (Lemma 2.1 [30]) and its transformation of data are utilized in (2.14), (2.15), (2.18), and (2.19) for updating statistic coefficients of equations in (2.11). This indeed accelerates the computation process. Furthermore, DTIDMLSMVSP plays a very significant role in the local discretization and model validation process.

2.7. Illustration 2: Dynamic Model for U.S. Treasury Bill Interest Rate and the U.S.-U.K. Foreign Exchange Rate

We also apply the above presented scheme for estimating parameters of a continuous-time model for U.S. Treasury Bill interest rate [44] and U.S.-U.K. foreign exchange rate [45] processes. By employing dynamic modeling process [23, 24], a continuous time dynamic model of interest rate process under random environmental perturbations is described in [30] as follows:

\[
dy = (\beta y + \mu y^\delta)dt + \sigma y^\gamma dW(t), \quad y(t_0) = y_0,
\] (2.20)

where \( \beta, \mu, \delta, \sigma, \gamma \in \mathbb{R}; \ y(t, t_0, y_0) \) is adapted, non-anticipating solution process with respect to \( \mathcal{F} \); the initial process \( y_0 \) is \( \mathcal{F}_{t_0} \)-measurable and independent of \( \{W(t), t \in [t_0, T]\} \); \( W(t) \) is a standard Wiener process defined on a filtered probability space \( (\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P}) \).
Transformation of Stochastic Differential Equation (2.20): For (2.20), we consider the Lyapunov functions $V_1(t,y) = \frac{1}{2}y^2$ and $V_2(t,y) = \frac{1}{2}y^3$ as in (2.2). The Itô differentials of $V_i$, for $i = 1, 2$, are given by

\[
\begin{align*}
   dV_1 &= \left[ y(\beta y + \mu y^\delta) + \frac{1}{2}\sigma^2 y^2 \right] dt + \sigma y^r dW(t), \\
   dV_2 &= \left[ y^2(\beta y + \mu y^\delta) + \sigma^2 y^{2r+1} \right] dt + \sigma y^{r+2} dW(t).
\end{align*}
\]

The Euler-type Numerical Schemes for (2.20) and (2.21) [21, 30]: Following the approach in Section 3 and illustration 2.6, the Euler discretized scheme ($\Delta t = 1$) for (2.20) is defined by

\[
\begin{align*}
   \Delta y_i &= (\beta y_{i-1} + \mu y^\delta_{i-1}) + \sigma y^r_{i-1}\Delta W(t), \\
   \frac{1}{2}\Delta(y^2_i) &= y_{i-1}(\beta y_{i-1} + \mu y^\delta_{i-1}) + \frac{1}{2}\sigma^2 y^2_{i-1} + \sigma y^{r-1} y^r_{i-1}\Delta W(t), \\
   \frac{1}{2}\Delta(y^3_i) &= y^2_{i-1}(\beta y_{i-1} + \mu y^\delta_{i-1}) + \sigma^2 y^{2r+1}_{i-1} + \sigma y^{r-1} y^{r+2}_{i-1}\Delta W(t). 
\end{align*}
\]

Generalized Moment Equations: Applying conditional expectation to (2.22) with respect to $\mathcal{F}_{t-1}$, we obtain

\[
\begin{align*}
   \mathbb{E}\left[ \Delta y_i | \mathcal{F}_{t-1} \right] &= \beta y_{i-1} + \mu y^\delta_{i-1}, \\
   \frac{1}{2}\mathbb{E}\left[ \Delta(y^2_i) | \mathcal{F}_{t-1} \right] &= \beta y^2_{i-1} + \mu y^{\delta^2}_{i-1} + \frac{1}{2}\sigma^2 y^2_{i-1}, \\
   \frac{1}{2}\mathbb{E}\left[ \Delta(y^3_i) | \mathcal{F}_{t-1} \right] &= \beta y^3_{i-1} + \mu y^{\delta^3}_{i-1} + \sigma^2 y^{2r+1}_{i-1}, \\
   \mathbb{E}\left[ \Delta y_i - \mathbb{E}\left[ \Delta y_i | \mathcal{F}_{t-1} \right] \right]^2 | \mathcal{F}_{t-1} &= \sigma^2 y^{2r+1}_{i-1}, \\
   \frac{1}{2}\mathbb{E}\left[ \Delta(y^2_i) - \mathbb{E}\left[ \Delta(y^2_i) \right] \right]^2 | \mathcal{F}_{t-1} &= \sigma^2 y^{2r+2}_{i-1}.
\end{align*}
\]

Remark 2.7. Orthogonality Condition Vector for (2.20) and (2.21): Again, in the context of (2.20), (2.21), (2.22), and (2.23), the orthogonality condition vector [10, 30] with respect to continuous-time stochastic dynamic model (2.20) is as:

\[
H(t_{i-1},y(t_{i-1}),y(t_{i-1})) = \begin{pmatrix}
   \Delta y(t_i) - (\beta y(t_{i-1}) + \mu y^\delta(t_{i-1}))\Delta t_i \\
   \frac{1}{2}\Delta(y^2(t_i)) - L(y^2(t_{i-1}))\Delta t_i \\
   \frac{1}{2}\Delta(y^3(t_i)) - L(y^3(t_{i-1}))\Delta t_i \\
   \left( \frac{1}{2}\Delta(y^2(t_i)) - L(y^2(t_{i-1}))\Delta t_i \right)^2 - \sigma^2 y^{2r+1}(t_{i-1})\Delta t_i \\
   \left( \frac{1}{2}\Delta(y^3(t_i)) - L(y^3(t_{i-1}))\Delta t_i \right)^2 - \sigma^2 y^{2r+2}(t_{i-1})\Delta t_i
\end{pmatrix},
\]

where $L(y^2(t_{i-1}))\Delta t_i = \left( y(t_{i-1}) \left( \beta y(t_{i-1}) + \mu y^\delta(t_{i-1}) \right) + \frac{1}{2}\sigma^2 y^{2r}(t_{i-1}) \right) \Delta t_i$ and $L(y^3(t_{i-1}))\Delta t_i = \left( y^2(t_{i-1}) \left( \beta y(t_{i-1}) + \mu y^\delta(t_{i-1}) \right) + \sigma^2 y^{2r+1}(t_{i-1}) \right) \Delta t_i$. Moreover, unlike the orthogonality condition vector defined in the literature [8, 10, 38], this orthogonality condition vector is based on the discretization scheme (2.22).
associated with nonlinear continuous-time stochastic differential equations (2.20) and (2.21).

**Local Observation System of Algebraic Equations:** Following the argument used in (2.14), for \( k \in I_0(N) \), from Definitions 2.3-2.7 in [30] and using (2.23), we formulate a local observation/measurement process at \( t_k \) as an algebraic functions of \( m_k \)-local functions of restriction of the overall finite sample sequence \( \{y_i\}_{i=0}^N \) to \( m_k \)-point subpartition \( P_k := t_{k-m_k} < t_{k-m_k+1} < \ldots < t_{k-1} \) as follows:

\[
\begin{align*}
\frac{1}{m_k} \sum_{i=k-m_k}^{k-1} E \left[ \Delta y_i | \mathcal{F}_{t_i} \right] &= \beta \frac{\sum_{i=k-m_k}^{k-1} y_{i-1}}{m_k} + \mu \frac{\sum_{i=k-m_k}^{k-1} y_{i+1}}{m_k}, \\
\frac{1}{2m_k} \sum_{i=k-m_k}^{k-1} \left[ E \left( \Delta (\gamma_i^2) | \mathcal{F}_{t_i} \right) \right] - E \left[ \left( \Delta y_i - E \left[ \Delta y_i | \mathcal{F}_{t_i} \right] \right)^2 | \mathcal{F}_{t_i} \right] &= \beta \frac{\sum_{i=k-m_k}^{k-1} \gamma_{i+1}}{m_k} + \mu \frac{\sum_{i=k-m_k}^{k-1} \gamma_{i+2}}{m_k}, \\
\sum_{i=k-m_k}^{k-1} E \left( \Delta (\gamma_i) - E \left[ \Delta (\gamma_i) | \mathcal{F}_{t_i} \right] \right) \gamma_{i-1} &= \beta \frac{\sum_{i=k-m_k}^{k-1} \gamma_{i+1}}{m_k} + \mu \frac{\sum_{i=k-m_k}^{k-1} \gamma_{i+2}}{m_k}, \\
\sum_{i=k-m_k}^{k-1} \gamma_{i+2} &= \beta \frac{\sum_{i=k-m_k}^{k-1} \gamma_{i+1}}{m_k} + \mu \frac{\sum_{i=k-m_k}^{k-1} \gamma_{i+2}}{m_k}. 
\end{align*}
\]

(2.25)

The solution \( \sigma_{m_k} \) is given by

\[
\sigma_{m_k} = \left[ \frac{s_{m_k}^2 \sum_{i=k-m_k}^{k-1} \gamma_{i-1}}{\frac{1}{m_k} \sum_{i=k-m_k}^{k-1} \gamma_{i-1}} \right]^{1/2},
\]

and \( \gamma_{m_k} \) satisfies the following nonlinear algebraic equation

\[
s_{m_k}^2 \sum_{i=k-m_k}^{k-1} \gamma_{i-1}^2 + \frac{1}{4}s_{m_k}^2 \sum_{i=k-m_k}^{k-1} \gamma_{i+1}^2 = 0,
\]

(2.27)

where \( s_{m_k}^2 \) and \( s_{m_k}^2 \) denote the local moving variance of \( \Delta y_i \) and \( \Delta (\gamma_i^2) \), respectively.

By the application of the Implicit Function Theorem [2], we conclude that for every non-constant \( m_k \)-local sequence \( \{y_i\}_{i=k-m_k}^N \), \( \delta \neq 1 \), there exist solution \( \hat{\beta}_{m_k}, \hat{\mu}_{m_k}, \) and \( \hat{\delta}_{m_k} \) of system of algebraic equations (2.25) as a point estimates of \( \beta, \mu \) and \( \delta \) respectively, at time \( t_k \), given by

\[
\begin{align*}
\hat{\beta}_{m_k} &= \frac{\sum_{i=k-m_k}^{k-1} \Delta y_i \sum_{i=k-m_k}^{k-1} \gamma_{i-1}^2 + \frac{1}{2} \sum_{i=k-m_k}^{k-1} \Delta (\gamma_i^2) \gamma_{i-1} + \frac{1}{2} \sum_{i=k-m_k}^{k-1} \gamma_{i+1} \gamma_{i-1} + \frac{1}{2} \sum_{i=k-m_k}^{k-1} \gamma_{i+2} \gamma_{i-1}}{\frac{1}{m_k} \sum_{i=k-m_k}^{k-1} \gamma_{i-1}}, \\
\hat{\mu}_{m_k} &= \frac{\frac{1}{m_k} \sum_{i=k-m_k}^{k-1} \Delta y_i \sum_{i=k-m_k}^{k-1} \gamma_{i+1} + \frac{1}{2} \sum_{i=k-m_k}^{k-1} \Delta (\gamma_i^2) \gamma_{i+1}}{\sum_{i=k-m_k}^{k-1} \gamma_{i-1}}, \\
\hat{\delta}_{m_k} &= \frac{\frac{1}{2} \sum_{i=k-m_k}^{k-1} \Delta y_i \sum_{i=k-m_k}^{k-1} \gamma_{i+2} + \frac{1}{2} \sum_{i=k-m_k}^{k-1} \Delta (\gamma_i^2) \gamma_{i+2}}{\sum_{i=k-m_k}^{k-1} \gamma_{i-1}}.
\end{align*}
\]

(2.28)
where $\delta_{m,k}$ satisfies the third equation in (2.25) described by

$$
\frac{1}{3m_k} \sum_{i=k-m_k}^{k-1} \Delta(y_i^d) = \frac{\sigma^2_{m,k}}{m_k} \sum_{i=k-m_k}^{k-1} y_{t-1}^2 - \beta_{m,k} \frac{\sum_{i=k-m_k}^{k-1} y_{t-1}^2}{m_k} - \mu_{m,k} \frac{\sum_{i=k-m_k}^{k-1} \delta_{m,k+1}}{m_k} = 0.
$$

We further note that the parameters of continuous-time dynamic process (2.20) are time-varying functions. This justifies the modifications needed for the development of continuous-time models of dynamic processes.

Remark 2.8. The presented illustrations exhibit the important features of the theoretical parameter estimation procedure [30]. The illustrations further clearly differentiate the Itô-Doob differential formula [24] based formation of orthogonality condition vectors in Remarks 2.2 and 2.7 and the algebraic manipulation and discretized scheme using the econometric specification based orthogonality condition vectors in [9, 11, 17].

Remark 2.9. The DTIDMLSMVSP and its transformation of data are utilized in (2.25), (2.26), (2.27), (2.28) and (2.29) for updating statistic coefficient of equations in (2.23). This indeed accelerates the computation process. Furthermore, DTIDMLSMVSP plays a very significant role in the local discretization and model validation errors.

3. Computational Algorithm

In this section, we outline theoretical computational components (5), (6) and (7) of LLGMM [30]. Again, for easy reference, we review the definitions of terms and expressions regarding computational, data organizational and simulation schemes. We introduce the idea of iterative data process and data simulation process time schedules in relation to the real time data observation/collection schedule. For the computational estimation of continuous time stochastic dynamic system state and parameters, it is essential to determine an admissible set of local conditional sample average and sample variance of local conditional sample in the context of a partition of time interval $[-\tau, T]$. Moreover, the discrete time dynamic model of conditional sample mean and sample variance statistic processes in Section 2 of [30] and the theoretical parameter estimation scheme in Section 3 coupled with the lagged adaptive expectation process motivate to outline a computational scheme in a systematic and coherent manner. A brief summary of the conceptual computational and simulation scheme is shown below.

3.1. Coordination of data observation, Iterative process, and Simulation schedules.

For easy reference, we present definitions 2.1-2.7 of iterative process and simulation time schedules discussed in Otunuga et al. [30]. Without loss of generality, we assume that the real data observation/collection partition schedule
$P$ of $[−τ, T]$ is defined in [30] by

$$P := \{t_i = −τ + (r + i)\Delta t\}, \quad \text{for } i \in I_r(N),$$

(3.1)

where $I_r(k) = \{j \in \mathbb{Z} : i \leq j \leq k\}$, $r + N$ stands for the total size of data.

**Definition 3.1.** An iterative process time schedule in relation with a real data collection schedule is defined by

$$IP = \{F^{-r}t_i : \text{for } t_i \in P\},$$

(3.2)

where $F^{-r}t_i = t_{i-r}$, and $F^{-r}$ is a forward shift operator [6], and $r$ is a discrete version of time delays of $τ$ defined in [30] by $r = \left\lceil \frac{τ}{\Delta t} \right\rceil + 1$.

The simulation time depends on an order $p$ of the time series model (2.16) of $m_k$-local conditional sample mean and variance processes.

**Definition 3.2.** Let $P, F^{-r}$ and $p$ be as defined in Definition 3.1. A simulation process time schedule in relation with a real data observation schedule is defined by

$$SP = \begin{cases} \{F^{-r}t_i : \text{for } t_i \in P\}, & \text{if } p \leq r \\ \{F^{-p}t_i : \text{for } t_i \in P\}, & \text{if } p > r. \end{cases}$$

(3.3)

**Remark 3.1.** We note that initial times of iterative and simulation processes are equal to the real data times $t_r$ and $t_p$, respectively. Moreover, iterative and simulation processes time in (3.2) and (3.3), respectively, justify Remark 2.3. In short, $t_i$ is a scheduled time clock for a collection of $i$th observation of the state of the system under investigation. The iterative and simulation process times are $t_{i+r}$ and $t_{i+p}$, respectively.

### 3.2. Conceptual Computational Parameter Estimation Scheme [30]

For a conceptual computational dynamic system parameter estimation, we need to introduce a few concepts of local admissible sample/data observation size, $m_k$-local admissible conditional finite sequence at $t_k ∈ SP$, local finite sequence of parameter estimates at $t_k$.

**Definition 3.3.** For each $k \in I_0(N)$, we define a local admissible sample/data observation size $m_k$ at $t_k ∈ SP$ in (3.3) as $m_k ∈ OS_k$, where

$$OS_k = \begin{cases} I_2(r + k - 1), & \text{if } p \leq r, \\ I_2(p + k - 1), & \text{if } p > r, \end{cases}$$

(3.4)
Moreover, $OS_k$ is referred to as the local admissible set of lagged sample/data observation size at $t_k$; $OS_k \subseteq S_P$ for $k \in I_0(N)$ and $OS_k \subseteq OS_{k+1}$ for $k \in I_0(N - 1)$.

**Definition 3.4.** For each admissible $m_k \in OS_k$ in Definition 3.3, a $m_k$-local admissible lagged-adapted finite restriction sequence of conditional sample/data observation at $t_k$ to subpartition $P_k$ of $P$ is defined by $\{E[y_i|F_{t-1}]\}_{i=k-m_k}^{k-1}$. Moreover, a $m_k$-class of admissible lagged-adapted finite sequences of conditional sample/data observation of size $m_k$ at $t_k$ is defined by

$$\mathcal{AS}_k = \{\{E[y_i|F_{t-1}]\}_{i=k-m_k}^{k-1} : m_k \in OS_k\} = \{\{E[y_i|F_{t-1}]\}_{i=k-m_k}^{k-1} m_k \in OS_k\}, \quad (3.5)$$

for each $k \in I_0(N)$.

Without loss of generality, in the case of energy commodity model, for each $m_k \in OS_k$, we find corresponding $m_k$-local admissible adapted finite sequence of conditional sample/data observation at $t_k$. $\{E[y_i|F_{t-1}]\}_{i=k-m_k}^{k-1}$. Using this sequence and (2.18), we compute $\hat{a}_{m_k,k}, \hat{\beta}_{m_k,k}$ and $\hat{\sigma}^2_{m_k,k}$. This leads to a local admissible finite sequence of parameter estimates at $t_k$ defined on $OS_k$ as follows: $\{(\hat{a}_{m_k,k}, \hat{\beta}_{m_k,k}, \hat{\sigma}^2_{m_k,k})\}_{m_k \in OS_k} = \{(\hat{a}_{m_k,k}, \hat{\beta}_{m_k,k}, \hat{\sigma}^2_{m_k,k})\}_{m_k \in OS_k}^{p+1} \text{ or }$ $(\hat{a}_{m_k,k}, \hat{\beta}_{m_k,k}, \hat{\sigma}^2_{m_k,k})^{p+1}$. It is denoted by

$$(\mathcal{A}_k, \mathcal{M}_k, \mathcal{S}_k) = \left\{(\hat{a}_{m_k,k}, \hat{\beta}_{m_k,k}, \hat{\sigma}^2_{m_k,k})\right\}_{m_k \in OS_k}, \quad (3.6)$$

for $k \in I_0(N)$.

### 3.3. Conceptual Computation of State Simulation Scheme: Energy Commodity Model

For the development of a conceptual computational scheme, we need to employ the method of induction. The presented simulation scheme is based on the idea of lagged adaptive expectation process [33]. An autocorrelation function (ACF) analysis [6, 8] performed on $s_{m,k}^2$ suggests that the discrete time interconnected dynamic model of local conditional sample mean and sample variance statistic in (2.8) of [30] is of order $p = 2$. Because of this, we need to identify an initial data. We begin with a given initial data $y_{m_0}, (s_{m_0,0}^2)_{m_0 \in OS_{0}}, (s_{m_{-1},1}^2)_{m_0 \in OS_{-1}}, \text{ and } (\tilde{s}_{m_{-1},1}^2)_{m_0 \in OS_{-1}}$.

Let $y_{m_{-1},k}^x$ be a simulated value of $E[y_i|F_{t-1}]$ at time $t_k$ corresponding to a local admissible lagged-adapted finite sequence $\{E[y_i|F_{t-1}]\}_{i=k-m_k}^{k-1} \in \mathcal{AS}_k$ of conditional sample/data observation of size $m_k$ at $t_k$ defined in (3.5). This simulated value is derived from the discretized Euler scheme (2.10) by

$$y_{m_{-1},k}^x = y_{m_{-1},k-1}^x + \hat{a}_{m_{-1},k-1}(\hat{\beta}_{m_{-1},k-1} - y_{m_{-1},k-1}^x)\Delta t + \hat{\sigma}_{m_{-1},k-1}y_{m_{-1},k-1}^x\Delta W_{m_{-1},k}. \quad (3.7)$$
Let

\[ \{y_{m,k}^{s}\}_{m \in \text{OS}_k}, \quad k \in I_0(N), \]  

be a \( m_k \)-local admissible sequence of simulated values corresponding to \( m_k \)-class \( \mathcal{AS}_k \) of local admissible lagged-adapted finite sequences of conditional sample/data observation of size \( m_k \) at \( t_k \) in (3.5) for \( k \in I_0(N) \).

### 3.4. Mean-Square Sub-Optimal Procedure

For each \( k \in I_0(N) \), to find the best estimate of \( \mathbb{E}[y_k|F_{k-1}] \) at time \( t_k \) from a \( m_k \)-local admissible finite sequence \( \{y_{m,k}^{s}\}_{m \in \text{OS}_k} \) of a simulated value of \( \mathbb{E}[y_k|F_{k-1}] \) defined in (3.8), we need to compute a local admissible finite sequence of quadratic mean square error corresponding to \( \{y_{m,k}^{s}\}_{m \in \text{OS}_k} \). A quadratic mean square error is defined below.

**Definition 3.5.** A local quadratic mean square error of \( \mathbb{E}[y_k|F_{k-1}] \) relative to each member of the term of local admissible sequence \( \{y_{m,k}^{s}\}_{m \in \text{OS}_k} \) of simulated values in (3.8) is defined by

\[ \Xi_{m,k,y_k} = \left( \mathbb{E}[y_k|F_{k-1}] - y_{m,k}^{s} \right)^2, \]  

for \( k \in I_0(N) \).

For any arbitrary small positive number \( \epsilon \) and for each time \( t_k \), to find the best estimate from the \( m_k \)-local admissible sequence \( \{y_{m,k}^{s}\}_{m \in \text{OS}_k} \) of simulated values, we determine the following \( \epsilon \)-sub-optimal admissible subset of set of \( m_k \)-size local admissible lagged sample size \( m_k \) at \( t_k \) (OS_): \n
\[ \mathcal{M}_k = \{ m_k : \Xi_{m,k,y_k} < \epsilon \text{ for } m_k \in \text{OS}_k \}, \]  

for \( k \in I_0(N) \). There are three different cases that determine the \( \epsilon \)-best sub-optimal sample size \( \hat{m}_k \) at time \( t_k \).

**Case 1:** If \( m_k \in \mathcal{M}_k \) gives the minimum, then \( m_k \) is recorded as \( \hat{m}_k \).

**Case 2:** If more than one value of \( m_k \in \mathcal{M}_k \), then the largest of such \( m_k \)'s is recorded as \( \hat{m}_k \).

**Case 3:** If condition (3.10) is not met (the property that defines \( \mathcal{M}_k \) at time \( t_k \), that is, \( \mathcal{M}_k = \emptyset \)), then the value of \( m_k \) where the minimum \( \min_m \Xi_{m,k,y_k} \) is attained is recorded as \( \hat{m}_k \). The \( \epsilon \)-best sub-optimal estimates of the parameters \( \hat{a}_{m,k}, \hat{\mu}_{m,k} \) and \( \hat{\sigma}_{m,k}^2 \) at the \( \epsilon \)-best sub-optimal sample size \( \hat{m}_k \) are also recorded as \( a_{\hat{m}_k, k}, \mu_{\hat{m}_k, k} \) and \( \sigma_{\hat{m}_k, k}^2 \), respectively, for \( k \in I_0(N) \).

Finally, the simulated value \( y_{m,k}^{s} \) at time \( t_k \) with \( \hat{m}_k \) is now recorded as the \( \epsilon \)-best sub-optimal state estimate for \( \mathbb{E}[y_k|F_{k-1}] \) at time \( t_k \) and denoted by \( \hat{y}_{\hat{m}_k, k} \). Similar reasoning can be provided for the estimates of the parameters of the U.S. Treasury Bill interest rate and U.S.-U.K. foreign exchange rate model.
Remark 3.2. We augment a few more Conceptual Computational Comparison between the LLGMM and the existing OCBGMM (Appendix C) [30] as follows.

a: The LLGMM approach is focused on parameter and state estimation problems at each data collection/observation time $t_k$ using the local lagged adaptive expectation process. In fact, LLGMM is discrete-time dynamic process. On the other hand, OCBGMM is centered on the state and parameter estimates using entire data that is to the left of the final data collection time $T_N = T$. Implied weakness in forecasting, as seen in the next section, is explicitly shown with the OCBGMM approach and the ensuing results.

b: We note that Remark 2.1 exhibits the interactions/interdependence between the first three components of LLGMM [30], namely (1) development of stochastic model for continuous-time dynamic process, (2) development of the discrete-time interconnected dynamic model for statistic process, (3) utilizing the Euler-type discretized scheme for nonlinear and non-stationary system of stochastic differential equations and their interactions. On the other hand, the OCBGMM is partially connected.

c: From the development of the computational algorithm section, we remark that the interdependence/interconnectedness of the four remaining components of the LLGMM [30], ”(4) employing lagged adaptive expectation process for developing generalized method of moment equations, (5) introducing conceptual computational parameter estimation problem, (6) formulating conceptual computational state estimation scheme, and (7) defining conditional mean square $\epsilon$-sub optimal procedure” is clearly illustrated. Moreover, the above stated components as well as data are directly connected with the original continuous-time SDE. On the other hand, OCBGMM is ”single shot approach” and highly dependent on its second component rather than the first component.

d: The LLGMM method [30] is a discrete-time dynamic system composed of seven interactive interdependent components. In fact, it is a dynamic non-parametric applied statistics method. On the other hand, the OCBGMM is static dynamic process of five almost isolated components.

e: Furthermore, LLGMM is a ”two scale hierarchic” quadratic mean-square optimization process, but the optimization process of OCBGMM is ”single-shot”

f: Although LLGMM performs in discrete-time, it operates like the original continuous-time dynamic process. The performance of the LLGMM approach is superior to the OCBGGM and IRGMM approaches.
The LLGMM method does not require a large size data set. In addition, as time $t_k$ increases, it generates a larger size of lagged adapted data set thereby further stabilizing the state and parameter estimation procedure with finite size data set. On the other hand, the OCBGMM does not exhibit this flexibility.

Further comparative summary analysis is described in Sections 7 in the context of conceptual, computational and statistical settings, exhibiting the role, scope and performance of LLGMM.

**Remark 3.3.** We note that the choice of $p = 2$ was determined based on the statistical procedure known as the Autocorrelation Function Analysis (AFA) [6, 8].

### 4. LLGMM and Statistical Analysis

In this section, we apply the theoretical LLGMM algorithm [30] to four energy commodities, U.S. Treasury Bill interest rate and the U.S.-U.K. foreign exchange rate data sets. In addition, we investigate the parameter and state estimation, forecasting and confidence-interval problems in the context of stochastic dynamic models (2.8) and (2.20) for these data sets.

#### 4.1. Application to Four Energy Commodity Data Sets:

In this section, we apply the above conceptual computational algorithm to the real time data sets, namely, daily Henry Hub natural gas data set for the period 01/04/2000-09/30/2004, daily crude oil data set for the period 01/07/1997 – 06/02/2008, daily coal data set for the period of 01/03/2000 – 10/25/2013, and weekly ethanol data set for the period of 03/24/2005 – 09/26/2013, [12, 13, 14, 48] in the context of stochastic dynamic model (2.8).

The descriptive statistics of data for the daily Henry Hub natural gas, daily crude oil data, daily coal data, and weekly ethanol data are recorded in Table 1 below.

<table>
<thead>
<tr>
<th>Data Set</th>
<th>$N$</th>
<th>$\bar{Y}$</th>
<th>Std($Y$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nat. Gas</td>
<td>1184</td>
<td>4.5504</td>
<td>1.5090</td>
</tr>
<tr>
<td>Crude Oil</td>
<td>4165</td>
<td>54.0093</td>
<td>31.0248</td>
</tr>
<tr>
<td>Coal</td>
<td>3470</td>
<td>27.1441</td>
<td>17.8394</td>
</tr>
<tr>
<td>Ethanol</td>
<td>438</td>
<td>2.1391</td>
<td>0.4455</td>
</tr>
</tbody>
</table>

**Graphical, Simulation and Statistical Results-Case 1:** We consider three cases for initial discrete-time delay $r$. We then later show that as $r$ increases, the root mean square error reduces, significantly. Here, we pick $r = 5$, $\Delta t = 1$, $\epsilon = 0.001$, and $p = 2$. The $\epsilon$- best sub-optimal estimates of parameters $a, \mu$ and $\sigma^2$ at each real data times are exhibited in Table 2 below.
Table 2: Estimates $\hat{\theta}_k, \sigma_{\hat{\theta}_k}^2, \hat{\mu}_{\hat{\theta}_k}$ and $\mu_{\hat{\theta}_k}$ for initial delay $r = 5$.

| $t_k$ | Natural gas |  |  |  |  |  |  |  |  |  |  |  |  |  |
|-------|-------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| 5     | 3           | 0.0001 | 2.2321 | 0.0011 | 5     | 3           | 0.0001 | 2.4400 | 0.0031 | 5     | 3           | 0.0001 | 11.3554 | 0.0142 |
| 6     | 3           | 0.0002 | 2.2410 | 0.0012 | 6     | 3           | 0.0002 | 2.7418 | 0.0034 | 6     | 3           | 0.0002 | 11.2529 | 0.0109 |
| 7     | 3           | 0.0002 | 2.2513 | 0.0007 | 7     | 4           | 0.0003 | 25.5964 | 0.0037 | 7     | 3           | 0.0001 | 9.9161  | 0.0016 |
| 8     | 4           | 0.0002 | 2.2494 | 0.0028 | 8     | 5           | 0.0006 | 25.5550 | 0.0047 | 8     | 3           | 0.0002 | 11.4663 | -0.0043 |
| 9     | 4           | 0.0002 | 2.2658 | 0.0014 | 9     | 4           | 0.0006 | 25.6595 | 0.0049 | 9     | 3           | 0.0005 | 10.5922 | -0.0043 |
| 10    | 4           | 0.0003 | 2.1571 | 0.0009 | 10    | 4           | 0.0009 | 8.5739  | 0.0071 | 10    | 4           | 0.0004 | 1.1500  | 0.0062 |
| 11    | 4           | 0.0004 | 2.9701 | -0.2781 | 11    | 3           | 0.0003 | 23.7742 | 0.0030 | 11    | 4           | 0.0032 | 8.9091  | 0.1784 |
| 12    | 4           | 0.0000 | 2.2556 | 0.0045 | 12    | 3           | 0.0002 | 26.9477 | -0.0157 | 12    | 3           | 0.0015 | 9.0169  | 0.0055 |
| 13    | 4           | 0.0005 | 2.3122 | 0.0046 | 13    | 3           | 0.0003 | 23.8766 | -0.0112 | 13    | 3           | 0.0020 | 8.6231  | 0.0739 |
| 14    | 4           | 0.0015 | 2.4850 | 0.0064 | 14    | 2           | 0.0001 | 22.1834 | 0.0049 | 14    | 2           | 0.0007 | 1.1669  | -0.0289 |
| 15    | 4           | 0.0007 | 2.5578 | 0.0006 | 15    | 4           | 0.0004 | 23.5245 | 0.0010 | 15    | 4           | 0.0014 | 0.7459  | -0.0079 |
| 16    | 3           | 0.0007 | 2.7151 | 0.0073 | 16    | 4           | 0.0002 | 23.8500 | 0.0000 | 16    | 4           | 0.0058 | 6.8121  | 0.0069 |
| 17    | 5           | 0.0011 | 2.6868 | 0.0084 | 17    | 4           | 0.0015 | 23.8486 | 0.0052 | 17    | 4           | 0.0015 | 8.8087  | 0.0044 |
| 18    | 4           | 0.0010 | 2.3831 | 0.0042 | 18    | 3           | 0.0003 | 23.2913 | -0.0113 | 18    | 4           | 0.0035 | 9.6841  | 0.0062 |
| 19    | 5           | 0.0007 | 2.3083 | 0.0035 | 19    | 5           | 0.0000 | 24.4715 | 0.1282 | 19    | 3           | 0.0048 | 9.0752  | 0.1257 |
| 20    | 5           | 0.0006 | 2.1000 | 0.0063 | 20    | 3           | 0.0004 | 24.8787 | 0.0415 | 20    | 4           | 0.0049 | 8.9899  | 0.0160 |
| 21    | 5           | 0.0007 | 2.3171 | 0.0693 | 21    | 4           | 0.0003 | 24.3336 | 0.0267 | 21    | 3           | 0.0014 | 1.8090  | 0.0044 |
| 22    | 4           | 0.0015 | 2.7943 | 0.0093 | 22    | 4           | 0.0002 | 23.9993 | 0.0280 | 22    | 4           | 0.0054 | 8.9148  | 0.0036 |
| 23    | 3           | 0.0009 | 2.5900 | 0.0016 | 23    | 4           | 0.0001 | 24.1909 | -0.0094 | 23    | 4           | 0.0018 | 8.6771  | 0.0084 |
| 24    | 3           | 0.0010 | 2.4917 | 0.0022 | 24    | 3           | 0.0003 | 20.8812 | -0.0252 | 24    | 4           | 0.0035 | 7.5566  | 0.0095 |
| 25    | 4           | 0.0017 | 2.9620 | 0.0021 | 25    | 4           | 0.0002 | 22.2942 | 0.0064 | 25    | 5           | 0.0006 | 8.7799  | -0.1155 |

$\epsilon_{k}: \text{From } (3) \text{ and Definition 3.3 (OS)}_k \text{, at each time } t_k, \text{ for the four energy price data sets, the } \epsilon_{k} \text{-best sub-optimal local admissible sample size } \hat{m}_k \text{ is attained on the subset } \{2, 3, 4, 5\} \text{ of } (OS)_k \text{. Hence, the } \epsilon_{k} \text{-best sub-optimal local state and parameter estimates are obtained in at most four iterates rather than } k + r - 1.$

Table 2 shows the $\epsilon_{k}$-best sub-optimal local admissible sample size $\hat{m}_k$ and the parameters $a_{\hat{m}_k}, \sigma_{\hat{m}_k}^2, \mu_{\hat{m}_k}$ for four energy commodity price data at time $t_k$. This is based on $p \leq r = 5$. Here, the range of the $\epsilon_{k}$-best sub-optimal local admissible sample size $\hat{m}_k$ for any time $t_k \in [5, 25] \cup [1145, 1165], t_k \in [5, 25] \cup [2440, 2460], t_k \in [5, 25] \cup [2865, 2885]$, and $t_k \in [5, 25] \cup [375, 395]$ for natural gas, crude oil, coal and ethanol data, respectively, is

$$2 \leq \hat{m}_k \leq 5.$$  

(4.1)

Remark 4.1. From (4.1), we draw the following conclusions:

a. From (3.4) and Definition 3.3 (OS$_k$), at each time $t_k$, for the four energy price data sets, the $\epsilon_{k}$-best sub-optimal local admissible sample size $\hat{m}_k$ is attained on the subset $\{2, 3, 4, 5\}$ of (OS$_k$). Hence, the $\epsilon_{k}$-best sub-optimal local state and parameter estimates are obtained in at most four iterates rather than $k + r - 1$.

b. The basis for the conclusion (a) is due to the fact that the $\epsilon_{k}$-best sub-optimization process described in Subsec-

19
tions 3.3 and 3.4 stabilizes the local state and parameter estimations at each time $t_k$.

c: From (a) and (b), we further remark that, in practice, the entire local lagged admissible set $OS_k$ of size $m_k$ at time $t_k$ is not fully utilized. In fact, for any $m_k \in OS_k$ and $m_k > \hat{m}_k$ such that as $m_k$ approached $k + r - 1$, the corresponding state and parameters relative to $m_k$ approach the $\epsilon$-best sub-optimal local state and parameter estimates relative to the $\epsilon$-best sub-optimal local admissible sample size $\hat{m}_k$ at time $t_k$. This is not surprising because of the nature of the state hereditary process, that is, as the size of the time-delay $m_k$ increases, the influence of the past state history decreases.

d: From (c), we further conclude that the second (DTIDMLSMVSP) and the fourth (local lagged adaptive process) component of the LLGMM [30] are stabilizing agents. This justifies the introduction of the term, namely, conceptual computational state and parameter estimation scheme. In fact, these components play a role in not only the local $\epsilon$-best suboptimal quadratic error reduction, but also local error stabilization problem depending on the choice of $\epsilon > 0$.

e: The conclusions (a), (b), (c) and (d) are independent of "large" data size and stationary conditions.

In the following, the graphs of $a_{\hat{m}_k,k}$ for natural gas, crude oil, coal and ethanol are exhibited in Figures 1 (a), (b), (c) and (d), respectively.
Figure 1: The graphs of mean reverting rate $a_{0,k}$ with time $t_k$ using initial delay $r = 5$.

Figure 1: (a), (b), (c) and (d) are the graphs of $a_{0,k}$ (with respect to (2.6)) against time $t_k$ for the daily Henry Hub natural gas data [14], daily crude oil data [13], daily coal data [12], and weekly ethanol data [48], respectively. Each sketch exhibits the rate at which the data sets are reverting to the mean level.

Furthermore, we show the graphs of $\mu_{0,k}$ for each of the data set.
Figure 2: The graphs of mean level $\hat{\mu}_{n,k}$ with time $t_k$ using initial delay $r = 5$

Figure 2: (a), (b), (c) and (d) are the graphs of $\hat{\mu}_{n,k}$ (with respect to (2.8)) against time $t_k$ for the daily Henry Hub natural gas [14], daily crude oil [13], daily coal [12], and weekly ethanol data [48], respectively. The sample mean value of the real data $y_k$ for natural gas, crude oil, coal and ethanol data are given by 4.5385, 54.0093, 27.1441 and 2.1391, respectively. It can be seen from Figure 2: (a), (b), (c) and (d) that the graph of $\hat{\mu}_{n,k}$ for the Henry Hub natural gas, crude oil, coal and ethanol data moves around the mean value 4.5385, 54.0093, 27.1441 and 2.1391, respectively. This analysis shows that the parameter $\hat{\mu}_{n,k}$ is close to the respective mean value of the data at time $t_k$. We also note that $[\hat{\mu}_{n,k}]_{t=0}^N$ and $[\alpha_{n,k}]_{t=0}^N$ are discrete-time $\epsilon$-best sub-optimal simulated random samples generated by the scheme described at the beginning of Section 4.1

Figures 3 (a), (b), (c) and (d) show the graph of $\hat{s}_{n,k}^2$ for natural gas, crude oil, coal and ethanol, respectively.
Figure 3: Moving Variance \( s^2_{m,k} \) against \( t_k \) with initial delay \( r = 5 \).

Figure 3: (a), (b), (c) and (d) are graphs of \( s^2_{m,k} \) (with respect to (2.8)) against time \( t_k \) with initial delay \( r = 5 \) for the daily Henry Hub natural gas data [14], daily crude oil data [13], daily coal data [12], and weekly ethanol data [48], respectively. We found these estimates using the discrete time dynamic model (2.16) (see Lemma 2.1 in [30]) with \( p = 2 \). This is based on the autocorrelation and partial autocorrelation function described in [6, 8]. Using the third part of (2.18), the volatility square at time \( t_k \) can be calculated.

In Table 3, the real and LLGMM simulated price values for the four energy commodities: natural gas, crude oil, coal and ethanol are exhibited in columns 2-3, 6-7, 10-11, and 14-15, respectively. The absolute error of each of the energy commodity’s simulated value is shown in columns 4, 8, 12, and 16.
Table 3: Real, Simulation using LLGMM method, and absolute error of simulation with starting delay $\tau = 5$.

<table>
<thead>
<tr>
<th>$t_k$</th>
<th>Natural gas</th>
<th>Crude oil</th>
<th>Coal</th>
<th>Ethanol</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\gamma_k$</td>
<td>$\sigma^2_{\gamma_k}$</td>
<td>$\gamma_k$</td>
<td>$\sigma^2_{\gamma_k}$</td>
<td>$\gamma_k$</td>
</tr>
<tr>
<td>5</td>
<td>2.218</td>
<td>2.218</td>
<td>0.000</td>
<td>5</td>
</tr>
<tr>
<td>6</td>
<td>2.252</td>
<td>2.252</td>
<td>0.000</td>
<td>6</td>
</tr>
<tr>
<td>7</td>
<td>2.285</td>
<td>2.285</td>
<td>0.000</td>
<td>7</td>
</tr>
<tr>
<td>8</td>
<td>2.318</td>
<td>2.318</td>
<td>0.000</td>
<td>8</td>
</tr>
<tr>
<td>9</td>
<td>2.351</td>
<td>2.351</td>
<td>0.000</td>
<td>9</td>
</tr>
<tr>
<td>10</td>
<td>2.384</td>
<td>2.384</td>
<td>0.000</td>
<td>10</td>
</tr>
</tbody>
</table>

Remark 4.2. We have used the estimated parameters $\alpha_{0_{k,l}}, \alpha_{0_{k,l}}$, and $\alpha_{0_{k,l}}$, in Table 2 to simulate the daily prices of natural gas, crude oil, coal, and ethanol. For this purpose, we pick $\varepsilon = 0.001$. For each time $t_k$, we estimate the simulated prices $y_{0_{k,l}}$. Among the set of admissible set values $m_k$, the value that gives the minimum $M_k$ is recorded as $\hat{m}_k$. If condition (3.10) is not satisfied at time $t_k$, the value of $m_k$ where the minimum $\Sigma_{m_k}\lambda_k$ is attained, is recorded as $\hat{m}_k$. The $\varepsilon$- best sub-optimal estimates of the parameters $\hat{m}_k, \tilde{m}_k, \hat{m}_k$, and $\sigma^2_{m_k}$ at $\hat{m}_k$ are also recorded as $\alpha_{0_{k,l}}$, $\alpha_{0_{k,l}}$, and $\sigma^2_{m_k}$; the value of $\gamma_{0_{k,l}}$ at time $t_k$ corresponding to $\hat{m}_k, \tilde{m}_k, \hat{m}_k$, and $\sigma^2_{m_k}$ is also recorded as the $\varepsilon$- best sub-optimal simulated value $y_{0_{k,l}}$ of $y_k$.

Next, we show the graphs of the simulated data using the LLGMM method for each commodity in Figure 4.
Figures 4: (a), (b), (c) and (d) under the application of the LLGMM approach show the graphs of the real and simulated spot prices for the daily Henry Hub natural gas [14], daily crude oil [13], daily coal [12], and weekly ethanol data [48], respectively. The red line represents the real data $y_k$ while the blue line represents the simulated value $y_{\hat{s}, k}$.

**Graphical, Simulation and Statistical Results-Case 2:** For better simulation results in Figure 4, we increase the magnitude of time delay $r$. We pick $r = 10$, $\Delta t = 1$, $\epsilon = 0.001$, and $p = 2$. The $\epsilon-$ best sub-optimal estimates of parameters $a$, $\mu$ and $\sigma^2$ at each real data times are exhibited in Appendix A.1, Table A.13. In Table A.14, the real and LLGMM simulated price values of each of the four energy commodities: natural gas, crude oil, coal and ethanol are exhibited in columns 2-3, 6-7, 10-11, and 14-15, respectively. The absolute error of each of the energy commodity’s simulated value is shown in columns 4, 8, 12, 16, respectively.
The following graphs exhibit the simulation using the LLGMM approach for natural gas, crude oil, coal and ethanol data with an initial time delay $r = 10$:

![Graphs showing real and simulated spot prices for natural gas, crude oil, coal, and ethanol.](image)

Figure 5: Real and Simulated Prices for initial delay $r = 10$.

Figures 5: (a), (b), (c) and (d) show the graphs of the real and simulated spot prices for the daily Henry Hub natural gas data [14], daily crude oil data [13], daily coal data [12], and weekly ethanol data [48], respectively using $r = 10$. The red line represents the real data $y_k$ while the blue line represent the simulated value $y_{\hat{m},k}$. The root mean square error of the simulation for the natural gas, crude oil, coal and ethanol data are given by 0.1004, 0.5401, 0.8879 and 0.0618, respectively.

**Graphical, Simulation and Statistical Results-Case 3:** Again, we pick $r = 20$, $\Delta t = 1$, $\epsilon = 0.001$, and $p = 2$, the $\epsilon-$ best sub-optimal estimates of parameters $a$, $\mu$ and $\sigma^2$ at each real data times are recorded in Appendix A.2, Table A.15. In Appendix A.2, Table A.16, the real and the LLGMM simulated price values of natural gas, crude oil,
coal and ethanol are exhibited in columns 2-3, 6-7, 10-11, and 14-15, respectively. The absolute error of each of the energy commodity’s simulated value is shown in columns 4, 8, 12, 16, respectively. The following graphs exhibit the simulation using the LLGMM method for natural gas, crude oil, coal and ethanol data with an initial discrete-time delay $r = 20$:

![Real and Simulated Price Graphs](image)

**Figure 6: Real and Simulated Prices for initial delay $r = 20$.**

Figures 6: (a), (b), (c) and (d) for $r = 20$ show the graphs of the real and simulated spot prices for the daily Henry Hub natural gas data [14], daily crude oil data [13], daily coal data [12], and weekly ethanol data [48], respectively. The red line represents the real data $y_k$ while the blue line represent the simulated value $\hat{y}_{s,k}^d$.

**Goodness-of-fit Measures:** We find the goodness-of-fit measures for four energy commodities: natural gas, crude oil, coal and ethanol. This is achieved by using the goodness-of-fit measures described in [11]:

27
\[
\begin{align*}
\hat{\text{RAMSE}} &= \left[ \frac{1}{N} \sum_{t=1}^{N} \left( \frac{1}{S} \sum_{s=1}^{S} \left( y_{t}^{(s)} - y_{t} \right) \right)^2 \right]^{1/2}, \\
\hat{\text{AMAD}} &= \frac{1}{N} \sum_{t=1}^{N} \text{median} \left( |y_{t}^{(s)} - \text{median} \{y_{t}^{(s)}\}| \right), \\
\hat{\text{AMB}} &= \frac{1}{N} \sum_{t=1}^{N} \left( \text{median} \{y_{t}^{(s)}\} - y_{t} \right),
\end{align*}
\]  

where \( \{y_{t}^{(s)}\}_{r=1,2,...,N} \) is a double sequence of simulated values at the data collected/observed time \( t = 1, 2, ..., N \); \( \hat{\text{RAMSE}} \) is the root mean square error of the simulated path, \( \hat{\text{AMAD}} \) measures the variability and \( \hat{\text{AMB}} \) measures the average median bias. The goodness-of-fit measures are computed using \( S = 100 \) pseudo-data sets. The comparison of the goodness-of-fit measures \( \hat{\text{RAMSE}}, \hat{\text{AMAD}} \) and \( \hat{\text{AMB}} \) for the four energy commodities are recorded in Table 4.

**Remark 4.3.** As the \( \hat{\text{RAMSE}} \) decreases, the state estimates approach the true value of the state. As the value of \( \hat{\text{AMAD}} \) increases, the influence of the random environmental fluctuations on the state dynamic process increases. In addition, if the value of \( \hat{\text{RAMSE}} \) decreases and the value of \( \hat{\text{AMAD}} \) increases, then the method of study possesses a greater degree of ability for state and parameter estimation accuracy and greater degree of ability to measure the variability of random environmental perturbations on the state dynamic of system. Moreover, as the \( \hat{\text{RAMSE}} \) decreases, \( \hat{\text{AMAD}} \) increases and the \( \hat{\text{AMB}} \) decreases, the method of study increases its performance under the three goodness of fit measures in a coherent way. On the other hand, as the \( \hat{\text{RAMSE}} \) increases, the state estimates tend to move away from the true value of the state. As the value of \( \hat{\text{AMAD}} \) decreases, the influence of the random environmental fluctuations on state dynamic process decreases. In addition, if the value of \( \hat{\text{RAMSE}} \) increases and the value of \( \hat{\text{AMAD}} \) decreases, then the method of study possesses a lesser degree of ability for state and parameter estimation accuracy and lesser degree of ability to measure the variability of random environmental perturbations on the state dynamic of system. Moreover, as the \( \hat{\text{RAMSE}} \) increases, \( \hat{\text{AMAD}} \) decreases and the \( \hat{\text{AMB}} \) increases, the method of study decreases its performance under the three goodness-of-fit measures in a coherent manner.

**The Comparison of Goodness-of-fit Measures:** The following table exhibits the Goodness-of-fit Measures for the energy commodities natural data, crude oil, coal, and ethanol data using the initial delays \( r = 5, r = 10, \) and \( r = 20. \)

<table>
<thead>
<tr>
<th>Goodness of-fit Measure</th>
<th>( r = 5 )</th>
<th>( r = 10 )</th>
<th>( r = 20 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Natural gas</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Crude oil</td>
<td>0.1801</td>
<td>0.1001</td>
<td>0.0674</td>
</tr>
<tr>
<td>Coal</td>
<td>1.1122</td>
<td>0.5401</td>
<td>0.4625</td>
</tr>
<tr>
<td>Ethanol</td>
<td>1.2235</td>
<td>0.8879</td>
<td>0.4794</td>
</tr>
<tr>
<td>Natural gas</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Crude oil</td>
<td>1.1521</td>
<td>1.1330</td>
<td>1.1318</td>
</tr>
<tr>
<td>Coal</td>
<td>24.6476</td>
<td>24.5376</td>
<td>24.5010</td>
</tr>
<tr>
<td>Natural gas</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Crude oil</td>
<td>0.3409</td>
<td>0.3323</td>
<td>0.3213</td>
</tr>
<tr>
<td>Coal</td>
<td>1.3372</td>
<td>1.3371</td>
<td>1.3371</td>
</tr>
<tr>
<td>Ethanol</td>
<td>12.8370</td>
<td>12.8369</td>
<td>12.8370</td>
</tr>
</tbody>
</table>

Table 4: Goodness-of-fit Measures for the cases: \( r = 5, r = 10, \) and \( r = 20. \)
Remark 4.4. From Tables 3, A.14 and A.16, it is clear that as $r$ increases the absolute error decreases. Furthermore, the comparison of the Goodness-of-fit measures in Table 4 for the energy commodities using the initial delays $r = 5$, $r = 10$, and $r = 20$ shows that as the delay $r$ increases, the root mean square error decreases, significantly; $\overline{AMAD}$, the average median absolute deviation decreases very slowly, and $\overline{AMB}$, the average median bias remains unchanged.

Remark 4.5. All the codes for the parameter estimation, simulations and forecasting are written and tested using Matlab program. Due to the online control nature of $m_k$ in our model, it is worth mentioning that the execution times for each of the four commodities depend on the robustness of the data.

4.2. Application to U.S. Treasury Bill Interest Rate and U.S.-U.K. Foreign Exchange Rate Data Sets:

In this subsection, we apply the conceptual computational algorithm discussed in Sections 3 and 4 to estimate the parameters in equation (2.20) in the context of the U.S. Treasury Bill interest rate (U.S. TBYIR) [44] and the U.S.-U.K. foreign exchange rate (U.S.-U.K. FER) [45] data collected on Forex database.

Graphical, Simulation and Statistical Results: Using $\epsilon = 0.001$, $r = 20$, and $p = 2$, the $\epsilon$-best sub-optimal estimates of parameters $\beta$, $\mu$, $\delta$, $\sigma$ and $\gamma$ for each U.S. Treasury Bill and U.S.-U.K. Foreign Exchange rate data sets are exhibited in Tables 5 and 6, respectively.

Table 5: Estimates for $\hat{\beta}_1$, $\hat{\beta}_2$, $\hat{\beta}_3$, $\hat{\beta}_4$, $\hat{\beta}_5$, $\hat{\beta}_6$, $\hat{\beta}_7$, $\hat{\beta}_8$, $\hat{\beta}_9$, $\hat{\beta}_10$, $\hat{\beta}_11$, $\hat{\beta}_12$, $\hat{\beta}_13$, $\hat{\beta}_14$, $\hat{\beta}_15$, $\hat{\beta}_16$, $\hat{\beta}_17$, $\hat{\beta}_18$, $\hat{\beta}_19$, $\hat{\beta}_20$, $\hat{\beta}_21$, $\hat{\beta}_22$, $\hat{\beta}_23$, $\hat{\beta}_24$, $\hat{\beta}_25$, $\hat{\beta}_26$, $\hat{\beta}_27$, $\hat{\beta}_28$, $\hat{\beta}_29$, $\hat{\beta}_30$, $\hat{\beta}_31$, $\hat{\beta}_32$, $\hat{\beta}_33$, $\hat{\beta}_34$, $\hat{\beta}_35$, $\hat{\beta}_36$, $\hat{\beta}_37$, $\hat{\beta}_38$, $\hat{\beta}_39$, $\hat{\beta}_40$, $\hat{\beta}_41$, $\hat{\beta}_42$, $\hat{\beta}_43$, $\hat{\beta}_44$, $\hat{\beta}_45$, $\hat{\beta}_46$, $\hat{\beta}_47$, $\hat{\beta}_48$, $\hat{\beta}_49$, $\hat{\beta}_50$, $\hat{\beta}_51$, $\hat{\beta}_52$, $\hat{\beta}_53$, $\hat{\beta}_54$, $\hat{\beta}_55$, $\hat{\beta}_56$, $\hat{\beta}_57$, $\hat{\beta}_58$, $\hat{\beta}_59$, $\hat{\beta}_60$, $\hat{\beta}_61$, $\hat{\beta}_62$, $\hat{\beta}_63$, $\hat{\beta}_64$, $\hat{\beta}_65$, $\hat{\beta}_66$, $\hat{\beta}_67$, $\hat{\beta}_68$, $\hat{\beta}_69$, $\hat{\beta}_70$, $\hat{\beta}_71$, $\hat{\beta}_72$, $\hat{\beta}_73$, $\hat{\beta}_74$, $\hat{\beta}_75$, $\hat{\beta}_76$, $\hat{\beta}_77$, $\hat{\beta}_78$, $\hat{\beta}_79$, $\hat{\beta}_80$, $\hat{\beta}_81$, $\hat{\beta}_82$, $\hat{\beta}_83$, $\hat{\beta}_84$, $\hat{\beta}_85$, $\hat{\beta}_86$, $\hat{\beta}_87$, $\hat{\beta}_88$, $\hat{\beta}_89$, $\hat{\beta}_90$, $\hat{\beta}_91$, $\hat{\beta}_92$, $\hat{\beta}_93$, $\hat{\beta}_94$, $\hat{\beta}_95$, $\hat{\beta}_96$, $\hat{\beta}_97$, $\hat{\beta}_98$, $\hat{\beta}_99$, $\hat{\beta}_{100}$ for U.S. Treasury Bill interest rate data.
respectively. This is based on \( \hat{\sigma}_k \), \( \epsilon \) in the range of the approach.

Tables 5 and 6 show the \( \epsilon \)-best sub-optimal local admissible sample size \( \hat{m}_k \) and the corresponding parameter estimates \( \beta_{\hat{m}_k,k} \), \( \mu_{\hat{m}_k,k} \), \( \delta_{\hat{m}_k,k} \), \( \sigma_{\hat{m}_k,k} \), and \( \gamma_{\hat{m}_k,k} \) for the U.S.-U.K. foreign exchange rate data at time \( t_k \), respectively. This is based on \( p \leq r \), and the initial real data time-delay \( r = 20 \), that is, the data schedule time \( t_0 = t_0 \). Furthermore, note that the range of the \( \epsilon \)-best sub-optimal local admissible sample size for the U.S. TBYIR and U.S.-U.K. FER data for time \( t_k \in [21,45] \bigcup [155,180] \), respectively, is \( 2 \leq \hat{m}_k \leq 20 \). All comments (Remark 4.1) made with regard to Table 2 remains valid with regard to Tables 5 and 6 in the context of the U.S. Treasury Bill interest rate and the U.S.-U.K. foreign exchange rate data at time \( t_k \) and the LLGMM approach.

We show the graphs of \( \beta_{\hat{m}_k,k} \), \( \mu_{\hat{m}_k,k} \), \( \delta_{\hat{m}_k,k} \), \( \sigma_{\hat{m}_k,k} \), and \( \gamma_{\hat{m}_k,k} \) for both monthly U.S. TBYIR and U.S.-U.K. FER.
Figure 7: $\hat{\beta}_{m,k}$, $\hat{\mu}_{m,k}$, $\hat{\delta}_{m,k}$, $\hat{\sigma}_{m,k}$, and $\hat{\gamma}_{m,k}$ for U.S. TYBIR using $r = 20$. 

\[ \delta_{m,k}, k, \mu_{m,k}, k, \sigma_{m,k}, k, \delta_{m,k}, k, \sigma_{m,k}, k, \gamma_{m,k} \]
Figure 8: $\beta_{m,k}^\delta$, $\mu_{m,k}^\delta$, $\delta_{m,k}^\delta$, $\sigma_{m,k}^\delta$, and $\gamma_{m,k}^\delta$ for U.S.-U.K. FER using $r = 20$.

Figures 7-8: (a), (b), (c), (d) and (e) are graphs of parameters in model (2.20) for U.S. TYBIR and U.S.-U.K. FER, respectively.

The following graphs show simulated path for the U.S. Treasury Bill interest rate and U.S.-U.K. FER with $r = 20$. 
Figures 9(a) and (b) show the real and simulated path for U.S. TBYIR and U.S.-U.K. FER, respectively.

**Comparison of Goodness-of-fit Measures for U.S. TBYIR and U.S.-U.K. FER using** \( r = 20 \): The following table compares the Goodness-of-fit Measures for the U.S. TBYIR and U.S.-U.K. FER data using \( r = 20 \).

Table 7: Goodness-of-fit Measures for the U.S. TBYIR and U.S.-U.K. FER data using \( r = 20 \).

<table>
<thead>
<tr>
<th>Goodness of-fit Measure</th>
<th>U.S. TBYIR</th>
<th>U.S.-U.K. FER</th>
</tr>
</thead>
<tbody>
<tr>
<td>( RAMSE )</td>
<td>0.0024</td>
<td>0.0137</td>
</tr>
<tr>
<td>( AMAD )</td>
<td>0.0148</td>
<td>0.0718</td>
</tr>
<tr>
<td>( AMB )</td>
<td>0.0165</td>
<td>0.1033</td>
</tr>
</tbody>
</table>

5. **Forecasting**

In this section, we outline the application of the LLGMM approach to robust forecasting and the confidence interval problems. Moreover, this approach does not require a large data size as well as any type of stationary conditions [6]. First, we shall sketch an outline of the forecasting problem. The \( \epsilon \)-best sub-optimal simulated value \( y_{S_{\hat{m}_k,k}}^{\epsilon} \) at time \( t_k \) is used to define a forecast \( y_{f_{\hat{m}_k,k}}^{\epsilon} \) for \( y_k \) at a lead time \( t_k \) for each of the energy commodity model, the U.S. TBYIR and U.S.-U.K. FER.

5.1. **Forecasting for Energy Commodity Model**

In the context of Illustration 1 (Section 2.6), we begin forecasting from a lead time \( t_k \). Using the data up to time \( t_{k-1} \), we compute \( \hat{\mu}_i, \sigma^2_{\hat{m}_i}, a_{\hat{m}_i}, \) and \( \mu_{\hat{m}_i} \) for \( i \in I_0(k-1) \). We assume that we have no information about the real
data \( y_i \) for \( t \). Under these considerations, imitating the computational procedure outlined in Sections 3 and 4 and using (2.18), we find the estimate of the forecast \( y_{t_k}^f \) at time \( t_k \) by employing the following discrete-time iterative process

\[
y_{t_{k-1}+1}^f = y_{t_{k-1}+1}^f = y_{t_{k-1}+1}^f + a_{t_{k-1},k-1}y_{t_{k-1},k-1}^f (\mu_{t_{k-1},k-1} - y_{t_{k-1},k-1}^f)\Delta t + \sigma_{t_{k-1},k-1}y_{t_{k-1},k-1}^f \Delta W(t_k),
\]

where the estimates \( \sigma_{t_{k-1},k-1}^2 \), \( a_{t_{k-1},k-1} \) and \( \mu_{t_{k-1},k-1} \) are defined in (2.18) with respect to the known past data up to the time \( t_{k-1} \). We note that \( y_{t_{k-1+k}}^f \) is the \( \epsilon \)-sub-optimal estimate for \( y_{t} \) at time \( t_k \).

To determine \( y_{t_{k+1},k+1}^f \), we need \( \sigma_{t_{k+1},k}^2 \), \( a_{t_{k+1},k} \) and \( \mu_{t_{k+1},k} \). Since we only have information of real data up to time \( t_{k-1} \), we use the forecasted estimate \( y_{t_{k-1}}^f \) as the estimate of \( y_{t} \) at time \( t_k \), and to estimate \( \sigma_{t_{k+1},k}^2 \), \( a_{t_{k+1},k} \) and \( \mu_{t_{k+1},k} \). Hence, we can write \( a_{t_{k+1},k+1} \equiv a_{t_{k+1},k+1} \equiv a_{t_{k+1},k+1}(a_{t_{k+1},k+1},a_{t_{k+1},k+1}^2,a_{t_{k+1},k+1}^3,a_{t_{k+1},k+1}^4,a_{t_{k+1},k+1}^5,a_{t_{k+1},k+1}^6,a_{t_{k+1},k+1}^7,a_{t_{k+1},k+1}^8,a_{t_{k+1},k+1}^9,a_{t_{k+1},k+1}^{10})^T \). To find \( y_{t_{k+1},k+2} \), we use the estimates \( a_{t_{k+1},k+1} \equiv a_{t_{k+1},k+1}(a_{t_{k+1},k+1},a_{t_{k+1},k+1}^2,a_{t_{k+1},k+1}^3,a_{t_{k+1},k+1}^4,a_{t_{k+1},k+1}^5,a_{t_{k+1},k+1}^6,a_{t_{k+1},k+1}^7,a_{t_{k+1},k+1}^8,a_{t_{k+1},k+1}^9,a_{t_{k+1},k+1}^{10})^T \). Continuing this process in this manner (using principle of Mathematical Induction, [23]), we use the estimates \( a_{t_{k+1,k+1}} \equiv a_{t_{k+1,k+1}}(a_{t_{k+1,k+1}},a_{t_{k+1,k+1}^2},a_{t_{k+1,k+1}^3},a_{t_{k+1,k+1}^4},a_{t_{k+1,k+1}^5},a_{t_{k+1,k+1}^6},a_{t_{k+1,k+1}^7},a_{t_{k+1,k+1}^8},a_{t_{k+1,k+1}^9},a_{t_{k+1,k+1}^{10}})^T \) to estimate \( y_{t_{k+1,k+1}} \).

### 5.2. Prediction/Confidence Interval for Energy Commodities

In order to be able to assess the future certainty, we also discuss the prediction/confidence interval. We define the 100(1 - \( \alpha \))\% confidence interval for the forecast of the state \( y_{t_{k},j} \) at time \( t_i \), \( i \geq k \), as \( y_{t_{k},j} \pm z_{1-\alpha/2}(\sigma_{t_{k-1},j}^2)^{1/2} y_{t_{i+1},j} \), where \( (\sigma_{t_{k-1},j}^2)^{1/2} y_{t_{i+1},j} \) is the estimate for the sample standard deviation for the forecasted state derived from the following iterative process

\[
y_{t_{k},j} = y_{t_{k},j} + a_{t_{k},k-1}y_{t_{k},k-1}(\mu_{t_{k-1},k-1} - y_{t_{k-1},k-1}^f)\Delta t + \sigma_{t_{k-1},k-1}y_{t_{k-1},k-1}^f \Delta W(t_k). \]

It is clear that the 95% confidence interval for the forecast at time \( t_i \) is

\[
\{ y_{t_{k},j} - 1.96 y_{t_{i+1},j}^{1/2}, y_{t_{k},j} + 1.96 y_{t_{i+1},j}^{1/2} \}
\]

where the lower and upper end of the interval denotes the lower and upper bound of the state estimate, respectively.
Figures 10: (a), (b), (c) and (d) show the graphs of the forecast and 95% confidence limit for the daily Henry Hub natural gas [14], daily crude oil [13], daily coal [12], and weekly ethanol data [48], respectively. Moreover, Figure 10: (a), (b), (c) and (d) show two regions: the simulation region $S$ and the forecast region $F$. For the simulation region $S$, we plot the real price data in red color and the simulated price data in blue color. For the forecast region $F$, we plot the real price in red color, the forecast price in green and the 95% confidence estimate of the forecast (as explained in Section 5.1) in black dots. The upper and lower simulated sketches in Figures 10 (a), (b), (c) and (d) corresponds to the upper and lower ends of the 95% confidence interval, respectively.

5.2.1. Sample forecasting error for energy commodities: $r = 20$

In this subsection, utilizing the procedure described in Subsection 5.1, we consider ten samples of forecast for each four energy commodities for each lead time greater than or equal to corresponding initial lead times. We draw
scattered plots of deviations of the ten sample forecasts from corresponding real and 95% forecast price for each lead time. These plots would exhibit the degree of stability and reliability of LLGMM approach. In addition, it would shed a light on the behavior of long-range forecasting. For the four energy commodities, namely, natural gas, crude oil, coal and ethanol, we choose the following initial lead-time forecast times and the lead-time domains as: 08/01/2004 to 12/13/2004 for natural gas, from lead time 06/03/2008 to 08/12/2008 for crude oil, from lead time 10/28/2013 to 01/08/2014 for coal and from lead time 10/03/2013 to 09/11/2014 for ethanol data.

Figure 11: Out-Of-Sample Forecast deviation from future price for energy commodities using $r = 20$.

Figures 11 (a), (b), (c) and (d) exhibit scattered plots of 10 out-of-sample forecast errors/deviation from real/future price for the daily Henry Hub natural gas, daily crude oil, daily coal, and weekly ethanol data with lead times 08/01/2004, 06/03/2008, 10/28/2013 and 10/03/2013, respectively. The 10 scattered sample error points for each lead time are identified by blue, red, orange, purple circles ($\circ$), dots ($\bullet$), triangles ($\triangle$ and $\triangledown$), diamond ($\diamond$), and mark ($\times$). It is easily noticeable that at least 8/9 are within one unit absolute deviations with 70-80% lead time domain.
Figure 12: Ten samples forecast and 95% confidence forecast with bounds for energy commodities using $r = 20$.

Figures 12 (a), (b), (c) and (d) show the 95% confidence and 10 forecast sample path trajectories coupled with 95% confidence upper and lower sample bound trajectories for the daily Henry Hub natural gas, daily crude oil, daily coal, and weekly ethanol data with lead times 08/01/2004, 06/03/2008, 10/28/2013 and 10/03/2013, respectively. 10 sample and the 95% sample forecasts are represented by black, blue, red, etc. colors. It is obvious that all 11 sample forecast paths are within the 95% upper and lower bounds. In short, the 95% confidence interval is robust with respect to a sample of 10 forecast realizations. 95% interval estimate is highly reliable and stable in the sense of longer lead times.

5.2.2. Sensitivity of forecast estimates to perturbations in model parameters for energy commodity: $r = 20$

In this subsection, we demonstrate the influence of the state dynamic parametric variations with respect to a nominal forecasting dynamic process (5.1) for each of the four energy commodities. We introduce random parametric
perturbations in the $\epsilon$-sub-optimal parameters $a_{\hat{h}_i,k}, \mu_{\hat{h}_i,k}$ and $\sigma_{\hat{h}_i,k}$ as follows

$$a_{\hat{h}_i,k,\nu_1} \equiv a_{\hat{h}_i,k} = a_{\hat{h}_i,k} + \nu_1 e_1, \quad \mu_{\hat{h}_i,k,\nu_2} \equiv \mu_{\hat{h}_i,k} = \mu_{\hat{h}_i,k} + \nu_2 e_2,$$  \hspace{1cm} (5.3)

where $e_i$, $i = 1, 2, 3$ are independent standard normal random variables and $\nu_i$, $i = 1, 2, 3$ characterize the magnitudes (noise intensities) of the random fluctuations. A dynamic process corresponding to the $\epsilon$-sub-optimal parameters $a_{\hat{h}_i,k}$ and $\mu_{\hat{h}_i,k}$ is referred as nominal price dynamic process of (5.1) or (5.2). A price dynamic process corresponding to parameters described in (5.3) is referred to perturbation of (5.1) or (5.2). This, together with (5.1)-(5.2) are now used to estimate the forecast value.

![Sensitivity plot of forecast error to perturbation in mean level](image)

Figure 13: Forecast error sensitivity analysis of the mean level nominal parameter $\mu_{\hat{h}_i,k}$ for (a) natural gas, (b) crude oil, (c) coal and (d) ethanol data.

Figures 13 (a), (b), (c) and (d) exhibit the absolute value of the sample forecast error/deviation of parametric perturbation in the mean level $\mu_{\hat{h}_i,k,\nu_2}$ from the corresponding nominal parameter (the $\epsilon$-sub-optimal parameters $\mu_{\hat{h}_i,k}$) for the natural gas, crude oil, coal and ethanol data.
respectively. These plots show how the absolute value of the sample forecast error changes with respect to the sample nominal forecast and noise intensity introduced in the mean level $\hat{\mu}_{k,\nu}$ as described in (5.3) for each energy commodity. The lead time for each forecast is the same as that presented in Subsubsection 5.2.1. It is clear from the sample path trajectories’ plots that as the noise intensity in the mean level parameter $\hat{\mu}_{k,\nu}$ increases, the color of the graph changes from blue to yellow. Here, a dark blue color represents region with very low absolute value of the error; a light blue error represents region with low error; a light yellow color represents region with high magnitude error and a dark yellow color represents regions with very high error.

![Sensitivity plots](image)

**Figure 14:** Forecast error sensitivity analysis in the mean reverting rate $a_{\theta_k,\kappa_1}$ (the nominal parameter) using initial delay $\tau = 20$.

Figures 14 (a), (b), (c) and (d) show the sensitivity of the energy commodity price out-of-sample-forecast error to perturbation in the mean reverting rate $a_{\theta_k,\kappa_1}$ for the natural gas, crude oil, coal and ethanol data, respectively. By definition, the mean reverting rate, $a_{\theta_k,\kappa_1}$, is the rate by which price shocks dissipate and the variable reverts towards the mean. They show how the out-of-sample-forecast error changes with respect to out-of-sample-time and noise intensity introduced in the mean reverting rate $a_{\theta_k,\kappa_1}$ as described in (5.3) for each energy commodity.
5.3. Prediction/Confidence Interval for U.S. Treasury Bill Interest Rate and U.S.-U.K. Foreign Exchange Rate

Following the procedure presented in Section 5.1, we show the graph of the real, simulated, forecast and 95% confidence limit (with respect to (2.20)) for the U.S. TBYIR and U.S.-U.K. FER for the initial delay \( r = 20 \).

Figure 15: Real, Simulated, Forecast and 95% Confidence Limit for U.S. TBYIR and U.S.-U.K.-FER Data.

Figure 15(a) shows the real, simulated, forecast and 95% confidence limit for the Interest rate data and Figure 15(b) shows the real, simulated, forecast and 95% confidence level for the U.S.-U.K. FER.

5.3.1. Sensitivity of forecast estimates to perturbations in model parameters for USTBYIR and U.S.-U.K. FER: \( r = 20 \)

Similar to the approach discussed in Subsection 5.2.2, we give the sensitivity of forecast estimates to perturbation plot for the U.S. Treasury Bill Interest Rate and U.S.-U.K. FER.
Figure 16: Forecast error sensitivity analysis in $\beta_{\hat{m},k}$ and $\mu_{\hat{m},k}$ (the nominal parameter) for USTBYIR: $r = 20$.

Figures 16 (a1) and (b1) show both the sensitivity analysis of the absolute value of sample forecast errors corresponding to the perturbed and nominal parameter $\beta_{\hat{m},k}$ for the U.S. Treasury Bill interest rate and the U.S.-U.K.-FER, respectively. Figures 16 (a2) and (b2) show both the sensitivity analysis of the absolute value of sample forecast errors corresponding to the perturbed and nominal parameter $\mu_{\hat{m},k}$ for the U.S. Treasury Bill interest rate and the U.S.-U.K.-FER, respectively.

6. The byproduct of the LLGMM approach

The second component, DTIDMLSMVSP (2.16) of the LLGMM [30] in Section 2 plays role: (a) to initiate ideas for the usage of discrete-time interconnected dynamic approach parallel to the continuous-time dynamic process, (b) to speed-up the computation time, (c) to significantly reduce the state error estimates, and also (d) as an alternative approach to the GARCH(1,1) model [4, 5] as well as comparable results with ex post volatility results of Chan et al [9]. Furthermore, LLGMM directly generates (Remark 2.5, Section 3) a GMM based method. In this section, we briefly discuss these comparison in the context of four energy commodity, U.S. TBYIR and U.S.-U.K. FER data.

6.1. Comparison between DTIDMLSMVSP and GARCH Model

In this subsection, we briefly compare the applications of DTIDMLSMVSP (2.16) and GARCH in the context of four energy commodities. We compare the estimates $s^2_{\hat{m},k}$ with the estimate derived from the usage of a GARCH(1,1) model described in [4] and defined by

$$
\begin{align*}
\frac{z_t}{\mathcal{F}_{t-1}} & \sim \mathcal{N}(0, h_t), \\
h_t &= \alpha_0 + \alpha_1 h_{t-1} + \beta_1 z^2_{t-1}, \quad \alpha_0 > 0, \quad \alpha_1, \beta_1 \geq 0.
\end{align*}
$$

(6.1)
The parameters $\alpha_0$, $\alpha_1$, and $\beta_1$ of the GARCH(1,1) conditional variance model (6.1) for the four commodities natural gas, crude oil, coal, and ethanol are estimated. The estimates of the parameters are given in Table 8 below.

<table>
<thead>
<tr>
<th>Data Set</th>
<th>$\alpha_0$</th>
<th>$\alpha_1$</th>
<th>$\beta_1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Natural Gas</td>
<td>$6.863 \times 10^{-5}$</td>
<td>0.853</td>
<td>0.112</td>
</tr>
<tr>
<td>Crude Oil</td>
<td>$9.622 \times 10^{-5}$</td>
<td>0.917</td>
<td>0.069</td>
</tr>
<tr>
<td>Coal</td>
<td>$3.023 \times 10^{-5}$</td>
<td>0.903</td>
<td>0.081</td>
</tr>
<tr>
<td>Ethanol</td>
<td>$4.152 \times 10^{-4}$</td>
<td>0.815</td>
<td>0.019</td>
</tr>
</tbody>
</table>

We later show a side by side comparison of $s_{m,k}^2$ [30] and the volatility described by GARCH(1,1) model (6.1) with coefficients in Table 8.

![Graph of $s_{m,k}^2$ and Garch(1,1) for Natural Gas](image1.png)

![Graph of $s_{m,k}^2$ and Garch(1,1) for Crude Oil](image2.png)
Figure 17: $s^2_{\text{m},k}$ and GARCH(1,1) for Coal

![Graph for Coal](c)

Figure 17: $s^2_{\text{m},k}$ and GARCH(1,1) for Ethanol

![Graph for Ethanol](d)

Figure 17: (a), (b), (c) and (d) show comparison of the volatility graphs using $s^2_{\text{m},k}$ and GARCH(1,1) model (6.1) for the daily Henry Hub natural gas [14], daily crude oil [13], daily coal [12] and weekly ethanol data [48], respectively. The blue and red lines show the graphs of estimates using $s^2_{\text{m},k}$ and GARCH(1,1) model, respectively. The GARCH model does not estimate volatility as heavily evidenced in Figure 17 (d). In fact, it demonstrated insensitivity.

We also compare Figure 4 with simulations using the GARCH(1,1) model (6.1) as the conditional variance below.
Figure 18: Simulation derived by using $s_{m,k}^2$ and GARCH(1,1)

Figure 18: (a), (b), (c) and (d) are graphs of the simulations using $s_{m,k}^2$ and GARCH(1,1) model (6.1) to estimate the volatility process for the daily Henry Hub natural gas [14], daily crude oil [13], daily coal [12], and weekly ethanol data [48], respectively. The blue dotted line shows the graph of estimates for the simulations using GARCH(1,1) model to simulate the volatility, the green dotted line is simulated estimates described in Figure 4, and the red line shows the real data. It can be seen that the GARCH model fails to capture most of the spikes in the data. Moreover, the GARCH model creates significant errors by over-and-under estimating the variance. These spikes were perfectly captured in Figure 4 where we use the discrete-time dynamic model of local sample variance statistics process [30] to estimate the volatility process described by GARCH (1,1) model (6.1). This illustrates the significance of dynamic statistic model (2.16) (Lemma 2.1of [30]) that performs better than the GARCH (1,1) volatility model.

6.2. Comparison of DTIDMLSMVSP [30] with the work of Chan et al [9]:

In this subsection, using the U.S. TBYIR and U.S.-U.K. FER data, the comparison between the DTIDMLSMVSP and ex post volatility of Chan et al [9] is made.

According to the work of Chan et al [9], we define the ex post volatility by the absolute value of the change in U.S. TBYIR data. Likewise, we define simulated volatility by the square root of the conditional variance implied by the estimates of the model (2.20). Using (2.20), we calculate our simulated volatility as $\sigma_{m,k}(\tilde{y}_{m,k})^{\tilde{x}_{m,k}}$. Figure 19 below shows the comparison between ex post volatility and simulated volatility. We compare our result with the result of Chan et al [9] and conclude that our method performs better than their method.
Figure 19: Ex Post Volatility and Simulated Volatility for Interest Rate.

Figure 19 shows the Ex post volatility and simulated volatility for the U.S. Treasury Bill interest rate data [44]. We compare our work in Figure 19 (using DTIDMLSMVSP [30]) with Figure 1 of Chan et al [9]. Their model does not clearly estimate the volatility. It demonstrated insensitivity in the sense that it was unable to capture most of the spikes in the interest rate ex post volatility data.

7. Comparisons of LLGMM with Existing and Newly Developed OCBGMM

In this Section, we briefly compare LLGMM and OCBGMM in the frame-work of the conceptual, computational, mathematical and statistical results coupled with role, scope and applications. For this purpose, to better appreciate and understand the comparative work, we utilize the state and parameter estimation problems for the stochastic dynamic model of interest rate that has been studied extensively [9, 11] in the frame-work of orthogonality condition vector based generalized method of moments (OCBGMM). We recall that the LLGMM approach is based on seven interactive components (Section 1). On the other hand, the existing OCBGMM (GMM [9] and IRGMM [11]) approach and its extensions are based on five components (Section 3, [30]). We further remark that the basis for the formation of orthogonality condition parameter vectors (OCPV) in the LLGMM (Section 3) and OCBGMM (GMM/IRGMM) are different. In fact, in the existing OCBGMM (GMM/IRGMM [9, 11]), the orthogonality condition vectors are formed on the basis of algebraic manipulation coupled with econometric specification-based discretization scheme.
(OCPV-Algebraic) rather than stochastic calculus and a continuous-time stochastic dynamic model independent based (OCPV-Analytic). This motivates to extend a couple of OCBGMM-based state and parameter estimation methods.

In this section, using the stochastic calculus based formation of the OCPV-Analytic in the context of the continuous-time stochastic dynamic model (Section 2), we develop two new OCBGMM based methods for the state and parameter estimation problems. The proposed OCBGMM methods are direct extensions of the existing OCBGMM method and its extension IRGMM [9, 11] in the context of the OCPV. Because of this difference and for the sake of comparison, the newly developed OCBGMM and the existing OCBGMM methods are referred to as the OCBGMM-Analytic and OCBGMM-Algebraic, respectively. In particular, the GMM [9] and IRGMM [11] with OCPV-algebraic are denoted as GMM-Algebraic and IRGMM-Algebraic and corresponding extensions under the OCPV-Analytic as GMM-Analytic and IRGMM-Analytic, respectively. Furthermore, using LLGMM based method, the aggregated generalized method of moments (AGMM) (2.19) introduced in Subsection 2.6 and described in Appendix B is also compared along with the above stated methods, namely, GMM-Algebraic, GMM-Analytic, IRGMM-Algebraic, and IRGMM-Analytic.

A comparative analysis of the results of GMM-Algebraic, GMM-Analytic, IRGMM-Algebraic, IRGMM-Analytic and Aggregated Generalized Method of Moment (AGMM) methods with the LLGMM for the state and parameter estimation problems of the interest rate and energy commodities stochastic dynamic models are briefly outlined in Appendix B, Appendix C, and Appendix D. First, based on the material in Sections 1, 2, and 3, we briefly summarize the comparison between the LLGMM and OCBGMM methods.

7.1. Theoretical Comparison Between LLGMM and OCBGMM

Based on the foundations of the analytical, conceptual, computational, mathematical, practical, statistical and theoretical motivations in the context of the material in Sections 2, 3, 4, 5 and 6, we summarize the comparison between the applications of innovative LLGMM [30] approach with the existing and newly developed LLGMM based OCBGMM methods. The comparative results are presented in tables in Appendix C in a systematic manner.

7.2. Comparisons of LLGMM and existing methods: Stochastic Interest rate model

The continuous-time interest rate process is described by a nonlinear Itô-Doob-type stochastic differential equation:

\[ dy = (\alpha + \beta y)dt + \sigma y^\gamma dW(t). \] (7.1)

An energy commodity stochastic dynamic model described in (2.8) in Subsection 2.6 is a modified generalized version of (7.1). These models would be utilized to further compare the role, scope and merit of the LLGMM and
OCBGMM methods in the framework of the graphical, computational and statistical results and its applications to forecasting and prediction with certain degree of confidence.

**Remark 7.1.** The continuous-time interest rate model (7.1) was chosen so that we can compare our LLGMM method with OCBGMM method described in [9] and [11]. Our proposed model for the continuous-time interest rate model is described in (2.20). We will later compare the results derived using model (7.1) with the results using (2.20) in Subsections 2.7 and 4.2.

In order to fulfil the objectives of the comparison, we need to construct the necessary tools outlined in Section 2.

**Descriptive Statistic for Time-series Data Set:** First, we consider one-month risk free rates from the Monthly Interest rate data sets for the period of June 30, 1964 to December 31, 2004.

Table 9: Statistics for the interest rate data for June 30, 1964 to December 31, 2004

<table>
<thead>
<tr>
<th>Variable</th>
<th>N</th>
<th>Mean</th>
<th>Std dev</th>
<th>$\rho_1$</th>
<th>$\rho_2$</th>
<th>$\rho_3$</th>
<th>$\rho_4$</th>
<th>$\rho_5$</th>
<th>$\rho_6$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$y_t$</td>
<td>487</td>
<td>0.0592</td>
<td>0.0276</td>
<td>0.9809</td>
<td>0.9508</td>
<td>0.9234</td>
<td>0.8994</td>
<td>0.8764</td>
<td>0.8519</td>
</tr>
<tr>
<td>$\Delta y_t$</td>
<td>486</td>
<td>-0.00003</td>
<td>0.0050</td>
<td>-0.0919</td>
<td>-0.0351</td>
<td>0.0403</td>
<td>-0.1877</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Mean, standard deviations, and autocorrelations of monthly U.S. Treasury Bill interest rate data. $\rho_j$ denotes the autocorrelation coefficient of order $j$. $N$ represents the total number of observations used in this study.

**The Orthogonality Condition Vector for (7.1):** In the following, we first present the orthogonality condition parameter vectors (OCPV) for the GMM-Algebraic, GMM-Analytic, IRGMM-Algebraic, and IRGMM-Analytic methods. These orthogonality vectors are then used for the state and parameter estimation problems. For this, we need to follow the procedure (Section 2) for obtaining the analytic orthogonality condition parameter vector (OCPV-Analytic).

We consider the Lyapunov functions $V_1(t, y) = \frac{1}{2} y^2$ and $V_2(t, y) = \frac{1}{3} y^3$. The Itô differentials of $V_1$ and $V_2$ with respect to (7.1) are:

\[
\begin{align*}
\frac{d}{dt} \left( \frac{1}{2} y^2 \right) &= \left[ \alpha y + \beta y^2 + \frac{1}{2} \sigma^2 y^2 \gamma \right] dt + \sigma \gamma y dW(t) \quad \text{(7.2)} \\
\frac{d}{dt} \left( \frac{1}{3} y^3 \right) &= \left[ \alpha y^2 + \beta y^3 + \sigma^2 y^2 \gamma^2 \right] dt + \sigma \gamma y^2 dW(t). 
\end{align*}
\]

The components of orthogonality condition vector (OCPV-Analytic) are listed below:

\[
\begin{align*}
\Delta y_t &- (E [y_t | F_{t-1}] - y_{t-1}) \\
\frac{1}{2} \Delta \left( \frac{y_t^2}{y_{t-1}^2} \right) &- \frac{1}{2} \left( E \left[ \frac{y_t^2}{F_{t-1}} \right] - y_{t-1}^2 \right) \\
\frac{1}{3} \Delta \left( \frac{y_t^3}{y_{t-1}^3} \right) &- \frac{1}{3} \left( E \left[ \frac{y_t^3}{F_{t-1}} \right] - y_{t-1}^3 \right) \\
E \left( \left( \Delta y_t - E [\Delta y_t | F_{t-1}] \right)^2 | F_{t-1} \right) &- \sigma^2 y_{t-1}^2 \Delta t,
\end{align*}
\]

47
are recorded in the row of AGMM approach in Table 10. Similarly, following the computational procedure described in [11], the parameter estimates in (7.4) are determined under the IRGMM-algebraic and IRGMM-analytic approaches, respectively, in Table 10. In order to statistically compare the different estimation techniques, we estimate the statistics $\text{RAMSE}$, $\text{A MAD}$, and $\text{AMB}$ defined in (4.2).

On the other hand, using discrete-time econometric specifications coupled with algebraic manipulations, the components of orthogonality condition parameter vector (OCPV-Algebraic) [9, 18, 30] are as follows:

\[
\begin{cases}
\mathbb{E}[y_t | \mathcal{F}_{t-1} ] - y_{t-1} &= (\alpha + \beta y) \Delta t \\
\frac{1}{2} \left( \mathbb{E} [ y_t^2 | \mathcal{F}_{t-1} ] - y_{t-1}^2 \right) &= [\alpha y_{t-1} + \beta y_{t-1}^2 + \frac{1}{2} \sigma^2 y^2 y_{t-1}^2] \Delta t \\
\frac{3}{4} \left( \mathbb{E} [ y_t^3 | \mathcal{F}_{t-1} ] - y_{t-1}^3 \right) &= [\alpha y_{t-1}^2 + \beta y_{t-1}^3 + \sigma^2 y^2 y_{t-1}^3] \Delta t \\
\mathbb{E} [ (\Delta y_t - \mathbb{E} [\Delta y_t | \mathcal{F}_{t-1} ] )^2 | \mathcal{F}_{t-1} ] &= \sigma^2 y^2 y_{t-1}^2 \Delta t.
\end{cases}
\]

Now, we are ready to apply the GMM-Algebraic, IRGMM-Algebraic, GMM-Analytic and IRGMM-Analytic methods.

**Parameter Estimates of (7.1) using LLGMM and OCBGMM Methods:** Following the construction in Remark 2.5, we define the average $\bar{\alpha}, \bar{\beta}, \bar{\sigma},$ and $\bar{y}$ by

\[
\begin{cases}
\bar{\alpha} = \frac{1}{N} \sum_{k=1}^{N} \alpha_{t_k,k}, \quad \bar{\beta} = \frac{1}{N} \sum_{k=1}^{N} \beta_{t_k,k}, \quad \bar{\sigma} = \frac{1}{N} \sum_{k=1}^{N} \sigma_{t_k,k}, \quad \bar{y} = \frac{1}{N} \sum_{k=1}^{N} y_{t_k,k},
\end{cases}
\]

where the parameters $\alpha_{t_k,k}, \beta_{t_k,k}, \sigma_{t_k,k},$ and $y_{t_k,k}$ are estimated in Table C.27 at time $t_k$ using LLGMM method.

Imitating the argument used in Appendix B, the parameters and state are also estimated. These parameter estimates are recorded in the row of AGMM approach in Table 10.

We also estimate the parameters in (7.1) by following the computational procedure described in [9] and applying it to both the GMM-algebraic and GMM-analytic frame-work. Similarly, following the computational procedure described in [11], the parameter estimates in (7.1) are determined under the IRGMM-algebraic and IRGMM-analytic approaches. These parameter estimates are recorded in rows of GMM-algebraic, GMM-analytic, IRGMM-algebraic and IRGMM-analytic approaches, respectively, in Table 10.

**Comparison of Goodness-of-fit Measures:** In order to statistically compare the different estimation techniques, we estimate the statistics $\text{RAMSE}$, $\text{A MAD}$, and $\text{AMB}$ defined in (4.2).

The goodness-of-fit measures are computed using $S = 100$ pseudo-data sets of the same sample size, and the real
data set, \( N = 487 \) months. The \( t \)-statistics of each parameter estimate is in parenthesis in Table 10, the smallest value of \( \hat{RMS}E \) for all method is italicized. The goodness-of-fit measures \( \hat{RMS}E, \hat{AMAD} \) and \( \hat{AMB} \) are recorded under the columns 6, 7, and 8, respectively.

Table 10: Comparison of Parameter estimates of model (7.1) and the goodness-of-fit measures \( \hat{RMS}E, \hat{AMAD} \) and \( \hat{AMB} \) using GMM-Algebraic, GMM-Analytic, IRGMM-Algebraic, IRGMM-Analytic, AGMM and LLGMM methods.

<table>
<thead>
<tr>
<th>Method</th>
<th>( \alpha )</th>
<th>( \beta )</th>
<th>( \sigma )</th>
<th>( \gamma )</th>
<th>( \hat{RMS}E )</th>
<th>( \hat{AMAD} )</th>
<th>( \hat{AMB} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>GMM-Algebraic</td>
<td>0.0017</td>
<td>-0.0308</td>
<td>0.4032</td>
<td>1.5309</td>
<td>0.0424</td>
<td>0.0098</td>
<td>0.0195</td>
</tr>
<tr>
<td></td>
<td>(1.53)</td>
<td>(-1.33)</td>
<td>(1.55)</td>
<td>(3.21)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GMM-Analytic</td>
<td>0.0009</td>
<td>-0.0153</td>
<td>0.0184</td>
<td>0.4981</td>
<td>0.0315</td>
<td>0.0161</td>
<td>0.0190</td>
</tr>
<tr>
<td></td>
<td>(1.06)</td>
<td>(-0.90)</td>
<td>(1.25)</td>
<td>(1.73)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IRGMM-Algebraic</td>
<td>0.0020</td>
<td>-0.0410</td>
<td>0.207</td>
<td>1.3031</td>
<td>0.0316</td>
<td>0.00843</td>
<td>0.01972</td>
</tr>
<tr>
<td></td>
<td>(0.32)</td>
<td>(-0.21)</td>
<td>(0.25)</td>
<td>(1.02)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IRGMM-Analytic</td>
<td>0.0084</td>
<td>-0.1436</td>
<td>0.1075</td>
<td>1.3592</td>
<td>0.0278</td>
<td>0.0028</td>
<td>0.01968</td>
</tr>
<tr>
<td></td>
<td>(0.44)</td>
<td>(-0.40)</td>
<td>(0.22)</td>
<td>(1.01)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AGMM</td>
<td>0.0084</td>
<td>-0.1436</td>
<td>0.1075</td>
<td>1.3592</td>
<td>0.0288</td>
<td>0.0047</td>
<td>0.0207</td>
</tr>
<tr>
<td></td>
<td>(0.41)</td>
<td>(-0.33)</td>
<td>(0.25)</td>
<td>(0.98)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LLGMM</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.0027</td>
<td>0.0146</td>
<td>0.0178</td>
</tr>
</tbody>
</table>

Table 10 shows a comparison of parameter estimates of model (7.1) and the goodness-of-fit measures \( \hat{RMS}E, \hat{AMAD} \) and \( \hat{AMB} \) under the usage of GMM-Algebraic, GMM-Analytic, IRGMM-Algebraic, IRGMM-Analytic, AGMM and LLGMM methods. The LLGMM estimates are derived using initial discrete-time delay \( r = 20, p = 2 \) and \( \epsilon = 0.001 \). Among these stated methods, the LLGMM method generates the smallest \( \hat{RMS}E \) value. In fact, the \( \hat{RMS}E \) value for the LLGMM is smaller than one tenth of any other \( \hat{RMS}E \) values. Moreover, second, third and fourth smaller \( \hat{RMS}E \) values are due to the IRGMM-Analytic, AGMM and GMM-Analytic methods, respectively. This exhibits the superiority of the LLGMM method over all other methods. We further observe that the LLGMM approach yields the smallest \( \hat{AMB} \) in comparison with the OCBGMM approaches. The GMM-Analytic, IRGMM-Analytic and IRGMM-Algebraic rank the second, third and fourth smaller values, respectively. The high value of \( \hat{AMAD} \) for the LLGMM method signifies that LLGMM captures the influence of random environmental fluctuations on the dynamic of interest rate process. We further note that the first, second, third and fourth smaller \( \hat{AMB} \) values are due to the LLGMM, GMM-Analytic, GMM-Algebraic and IRGMM-Analytic methods, respectively. These statements can be confirmed by comparing fluctuations in LLGMM simulation in Figure C.22 with other simulations in Figures C.23-C.26. Again, from Remark 4.3, the smallest \( \hat{RMS}E \), higher \( \hat{AMAD} \) and smallest \( \hat{AMB} \) value under the LLGMM method exhibits superior performance under the three goodness-of-fit measures. We also notice that the performance of stochastic calculus based-OCPV-Analytic methods (GMM-Analytic, IRGMM-Analytic and AGMM) is better than the performance of OCPV-Algebraic based methods (GMM-Algebraic and IRGMM-Algebraic). In short, this suggests that the OCPV-Analytic based GMM methods are superior to the OCPV-Algebraic based GMM methods.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>GMM-Algebraic</td>
<td>0.0468</td>
<td>0.0377</td>
</tr>
<tr>
<td>GMM-Analytic</td>
<td>0.0315</td>
<td>0.0347</td>
</tr>
<tr>
<td>IRGMM-Algebraic</td>
<td>0.0307</td>
<td>0.0326</td>
</tr>
<tr>
<td>IRGMM-Analytic</td>
<td>0.0307</td>
<td>0.0326</td>
</tr>
<tr>
<td>LLGMM</td>
<td>0.0030*</td>
<td>0.0017*</td>
</tr>
</tbody>
</table>

Table 11 shows the goodness-of-fit measures $\hat{\text{RAMSE}}$ using GMM-Algebraic, GMM-Analytic, IRGMM-Algebraic, IRGMM-Analytic and LLGMM methods for two separate sub-periods: 06/1964-12/1981 and 01/1982-12/2004. Among all methods, the LLGMM method generates the smallest $\hat{\text{RAMSE}}$ value for each subperiod. Moreover, the goodness-of-fit measure $\hat{\text{RAMSE}}$ regarding the LLGMM method is less than one sixth, and one twelfth of any other $\hat{\text{RAMSE}}$ value, respectively. In fact, the ranking of IRGMM-Analytic, IRGMM-Algebraic, GMM-Analytic and GMM-Algebraic methods are second, third, fourth and fifth place, respectively.

In the following, using the LLGMM method and three goodness of fit measures, we validate dynamic models (2.20) and (7.1) in the context of real data set.

### 7.3. Comparison of Goodness of fit Measures of model (2.20) with model (7.1) using LLGMM method

As stated in Remark 7.1, we compare the Goodness of fit Measures $\hat{\text{RAMSE}}, \hat{\text{AMAD}}, \text{ and } \hat{\text{AMB}}$ using the U.S. Treasury Bill interest rate data and the LLGMM applied to the model validation problems of two proposed continuous-time dynamic models of U.S. Treasury Bill interest rate process described by (2.20) and (7.1). The LLGMM state estimates of (2.20) and (7.1) are computed under the same initial discrete-time delay $r = 20$, $p = 2$ and $\epsilon = 0.001$. The results are recorded in the following table.

Table 12: Comparison of Goodness of fit Measure of model (2.20) with model (7.1)

<table>
<thead>
<tr>
<th>LLGMM</th>
<th>$\hat{\text{RAMSE}}$</th>
<th>$\hat{\text{AMAD}}$</th>
<th>$\hat{\text{AMB}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model (2.20)</td>
<td>0.0024*</td>
<td>0.0145</td>
<td>0.0178</td>
</tr>
<tr>
<td>Model (7.1)</td>
<td>0.0027</td>
<td>0.0146</td>
<td>0.0178</td>
</tr>
</tbody>
</table>

Table 12 shows that the Goodness of fit Measures $\hat{\text{RAMSE}}, \hat{\text{AMAD}}, \text{ and } \hat{\text{AMB}}$ of the LLGMM method using both models (7.1) and (2.20) are very close. Model (2.20) appears to have the least $\hat{\text{RAMSE}}$ value. This shows that the model (2.20) is a better dynamic model for U.S. Treasury Bill interest rate prices than model (7.1) under the LLGMM application, since it has a lower root mean square error. The $\hat{\text{AMAD}}$ value using (7.1) is larger than the value using (2.20). This suggests that the influence of the random environmental fluctuations on state dynamic model (7.1) is higher than using the model (2.20). The $\hat{\text{AMB}}$ value derived using both models appeared to be the same, indicating that both models give the same average median bias estimates. Based on this statistical analysis, we conclude that (2.20) is most realistic continuous-time stochastic dynamic model for the
short-term riskless rate model. Model (2.20) includes many well-known interest rate models such as CKLS diffusion model [9] as a special case (with $\delta = 0$).
APPENDIX AND SUPPLEMENTAL MATERIALS

For $\Delta t = 1$, $\epsilon = 0.001$, $p = 2$, the $\epsilon$-best sub-optimal estimates of parameters $a$, $\mu$ and $\sigma^2$ in (2.8) for four energy commodity data sets using $r = 10$ and $r = 20$ are outlined in Appendix A. The AGMM approach generated by the idea in Remark 2.5 is fully outlined, applied and compared with four energy data sets in Appendix B. A detailed comparison regarding the theoretical, graphical and performance of the LLGMM and OCBGMM methods are presented in Appendix C. In addition, a comparison of LLGMM with a few nonparametric statistical methods is also outlined in Appendix D.

Appendix A. State and Parameter Estimates of daily Natural gas, daily Crude oil, daily Coal, and weekly Ethanol data for initial delays $r = 10$ and $r = 20$

An initial choice of $r$ and $p$ in Section 3 plays a very significant role in computational coordination, parameter and state estimation and state simulation. The ACF [6, 8] is used to determine the value $p$. An initial discrete-time delay $r$ is used based on the increasing sequential choice $r = 5, 10, 20$ for the best graphical simulation result. The results are outlined in Appendix A.1 and Appendix A.2.
### Appendix A.1. State and Parameter Estimates for daily Natural gas, Crude oil, and weekly Ethanol data using initial delay $r = 10$

Table A.13: Estimates $\hat{\mathbf{m}}_2$, $\sigma^2_{\hat{\mathbf{k}}_A}$, $\mu_{\hat{\mathbf{k}}_A}$, and $\sigma^2_{\hat{\mathbf{k}}_A}$ for initial delay $r = 10$.

<table>
<thead>
<tr>
<th>$t$</th>
<th>$\hat{\mathbf{m}}_2$</th>
<th>$\sigma^2_{\hat{\mathbf{k}}_A}$</th>
<th>$\mu_{\hat{\mathbf{k}}_A}$</th>
<th>$\sigma^2_{\hat{\mathbf{k}}_A}$</th>
<th>$\mu_{\hat{\mathbf{k}}_A}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>11</td>
<td>8 0.0003 2.9051</td>
<td>11 0.1726</td>
<td>6 0 24.3552</td>
<td>11 0.0015 5.9591</td>
<td>11 6 0.0009 1.1809</td>
</tr>
<tr>
<td>12</td>
<td>6 0.0003 2.1346</td>
<td>12 0.1311</td>
<td>6 0 23.8577</td>
<td>12 0.0011 2.9737</td>
<td>12 6 0.0007 1.2887</td>
</tr>
<tr>
<td>13</td>
<td>7 0 2.5701</td>
<td>0.0630</td>
<td>3 0.003 23.8786</td>
<td>-0.0152</td>
<td>13 2 0.0029 7.6663</td>
</tr>
<tr>
<td>14</td>
<td>9 0.0007 2.6746</td>
<td>0.0461</td>
<td>14 0.0100 24.0133</td>
<td>0.0264</td>
<td>14 5 0.0033 9.7962</td>
</tr>
<tr>
<td>15</td>
<td>7 0.0002 2.4425</td>
<td>0.4071</td>
<td>15 0.0100 22.7352</td>
<td>0.0065</td>
<td>15 10 0.0041 9.4047</td>
</tr>
<tr>
<td>16</td>
<td>3 0.0013 2.5459</td>
<td>0.4621</td>
<td>16 0.0002 23.6665</td>
<td>0.0463</td>
<td>16 5 0.0030 9.4866</td>
</tr>
<tr>
<td>17</td>
<td>8 0.0015 2.5576</td>
<td>0.1934</td>
<td>17 0.0005 24.0777</td>
<td>0.0194</td>
<td>17 10 0.0048 9.1894</td>
</tr>
<tr>
<td>18</td>
<td>8 0.0014 2.5626</td>
<td>0.2495</td>
<td>18 0.0008 24.2210</td>
<td>0.0138</td>
<td>18 0.0046 9.6061</td>
</tr>
<tr>
<td>19</td>
<td>7 0.0015 2.5705</td>
<td>0.3522</td>
<td>19 0.0006 24.1147</td>
<td>0.0268</td>
<td>19 4 0.0043 9.7055</td>
</tr>
<tr>
<td>20</td>
<td>9 0.0011 2.5943</td>
<td>0.2946</td>
<td>20 0.0004 24.2748</td>
<td>0.0256</td>
<td>20 3 0.0039 9.8081</td>
</tr>
<tr>
<td>21</td>
<td>9 0.0010 2.6047</td>
<td>0.0775</td>
<td>21 7 0.0005 24.2175</td>
<td>0.0258</td>
<td>21 3 0.0030 7.4241</td>
</tr>
<tr>
<td>22</td>
<td>9 0.0010 2.6046</td>
<td>0.1883</td>
<td>22 4 0.0002 23.9993</td>
<td>0.0137</td>
<td>22 8 0.0085 8.8553</td>
</tr>
<tr>
<td>23</td>
<td>3 0.0009 2.7119</td>
<td>0.6893</td>
<td>23 10 0.0008 23.6479</td>
<td>0.0130</td>
<td>23 5 0.0010 8.6669</td>
</tr>
<tr>
<td>24</td>
<td>10 0.0013 2.6442</td>
<td>0.2966</td>
<td>24 10 0.0009 24.7857</td>
<td>-0.0087</td>
<td>24 6 0.0060 8.7592</td>
</tr>
<tr>
<td>25</td>
<td>9 0.0018 2.6387</td>
<td>0.2382</td>
<td>25 0.0001 21.8890</td>
<td>0.0069</td>
<td>25 0.0064 8.8440</td>
</tr>
<tr>
<td>26</td>
<td>2 0.0015 2.3232</td>
<td>0.6595</td>
<td>26 4 0.0003 22.2071</td>
<td>0.0258</td>
<td>26 8 0.0067 8.8466</td>
</tr>
<tr>
<td>27</td>
<td>4 0.0014 2.3464</td>
<td>0.3474</td>
<td>27 10 0.0011 35.7200</td>
<td>-0.0010</td>
<td>27 3 0.0012 9.0667</td>
</tr>
<tr>
<td>28</td>
<td>3 0.0008 2.5780</td>
<td>0.2807</td>
<td>28 4 0.0003 22.1582</td>
<td>0.0391</td>
<td>28 8 0.0033 9.5557</td>
</tr>
<tr>
<td>29</td>
<td>2 0.0011 2.6588</td>
<td>-0.1271</td>
<td>29 6 0.0004 22.1948</td>
<td>0.0401</td>
<td>29 4 0.0007 9.0561</td>
</tr>
<tr>
<td>30</td>
<td>7 0.0001 2.5610</td>
<td>0.3718</td>
<td>30 7 0.0005 22.2596</td>
<td>0.0394</td>
<td>30 8 0.0041 8.9685</td>
</tr>
</tbody>
</table>

- Table A.13 shows the $e$- best sub-optimal local admissible sample size $\hat{\mathbf{n}}_2$ and the parameter estimates $\hat{\mathbf{n}}_{\hat{\mathbf{k}}_A}$, $\hat{\mathbf{A}}_{\hat{\mathbf{k}}_A}$, and $\hat{\mathbf{Q}}_{\hat{\mathbf{k}}_A}$ for four energy commodities price at time $t$. This is based on the value of $p$ and the initial real data time delay $r = 10$. We further note that the range of the $e$-best sub-optimal local admissible sample size $\hat{\mathbf{n}}_2$ for any time $t \in [11, 30] \cup [1145, 1165]$, $t_0 \in [11, 30] \cup [2440, 2460]$, $t_1 \in [11, 30] \cup [2865, 2885]$, and $t_2 \in [11, 30] \cup [375, 395]$ for natural gas, crude oil, and ethanol data, respectively, is $2 \leq \hat{\mathbf{n}}_2 \leq 10$. Moreover, all comments (Remark 4.1) that are made with regard to Table 2 regarding the four energy commodities remain valid with regard to Table A.13.
| \( t_i \) | Real Simulated (LLGMM) | \( |E_{|t_i}| \) | Real Simulated (LLGMM) | \( |E_{|t_i}| \) | Real Simulated (LLGMM) | \( |E_{|t_i}| \) | Real Simulated (LLGMM) | \( |E_{|t_i}| \) |
|-----|-----------------------|-----|-----------------------|-----|-----------------------|-----|-----------------------|-----|
| 10  | 2.3260 2.3200 0.0060 | 11  | 2.4170 2.4179 0.0009 | 12  | 2.3590 2.4035 0.0655 | 13  | 2.4850 2.4949 0.0099 | 14  | 2.5280 2.5213 0.0157 |
| 15  | 2.6160 2.6158 0.0002 | 16  | 2.5230 2.5233 0.0003 | 17  | 2.6180 2.6134 0.0214 | 18  | 2.6100 2.5852 0.0248 | 19  | 2.6100 2.6130 0.0030 |
| 20  | 2.6990 2.6728 0.0262 | 21  | 2.7590 2.7601 0.0111 | 22  | 2.6950 2.6427 0.0163 | 23  | 2.7420 2.7675 0.0055 | 24  | 2.5620 2.5610 0.0010 |
| 25  | 2.4950 2.3457 0.0050 | 26  | 2.5900 2.5245 0.0115 | 27  | 2.5920 2.5996 0.0076 | 28  | 2.5790 2.5849 0.0149 | 29  | 2.5410 2.5403 0.0007 |
| 30  | 2.6180 2.6151 0.0029 | 31  | 2.6290 2.8757 0.0012 | 32  | 2.9370 2.9352 0.0023 | 33  | 2.9370 3.9371 0.0191 | 34  | 2.9370 3.9352 0.0023 |

In Table A.14, the real and LLGMM simulated price values for each of the four energy commodities: natural gas, crude oil, coal and ethanol are recorded in columns 2-3, 6-7, 10-11, and 14-15, respectively. The natural gas, crude oil, coal and ethanol are recorded in columns 2-3, 6-7, 10-11, and 14-15, respectively. The absolute error of each of the energy commodity’s simulated value is shown in columns 4, 8, 12, 16, respectively.
# Appendix A.2. State and Parameter Estimates of daily Natural gas, daily Crude oil, daily Coal, and weekly Ethanol data for initial delay $r = 20$

Table A.15: Estimates $\hat{\mu}_t$, $\sigma^2_t$, $\hat{\mu}_{mk}$, $\hat{\sigma}_{mk}$ and $\hat{\theta}_{mk}$ for initial delay $r = 20$.

<table>
<thead>
<tr>
<th>$t_i$</th>
<th>Natural gas</th>
<th>Crude oil</th>
<th>Coal</th>
<th>Ethanol</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\hat{\mu}_t$</td>
<td>$\sigma^2_t$</td>
<td>$\hat{\mu}_{mk}$</td>
<td>$\hat{\sigma}_{mk}$</td>
<td>$\hat{\theta}_{mk}$</td>
</tr>
<tr>
<td>21</td>
<td>13 0.0011 2.5056 0.0816</td>
<td>21</td>
<td>13 0.0011 2.5056 0.0816</td>
<td>21</td>
</tr>
<tr>
<td>22</td>
<td>5 0.0009 2.6748 0.2351</td>
<td>22</td>
<td>7 0.0009 2.4353 0.2788</td>
<td>22</td>
</tr>
<tr>
<td>23</td>
<td>3 0.0013 2.7139 0.6983</td>
<td>23</td>
<td>2 0.0006 2.4031 0.3143</td>
<td>23</td>
</tr>
<tr>
<td>24</td>
<td>12 0.0021 2.6397 0.2119</td>
<td>24</td>
<td>15 0.0007 14.2468 0.0069</td>
<td>24</td>
</tr>
<tr>
<td>25</td>
<td>10 0.0022 2.6210 0.1999</td>
<td>25</td>
<td>19 0.0011 18.5423 0.0201</td>
<td>25</td>
</tr>
<tr>
<td>26</td>
<td>5 0.0015 2.5677 0.2085</td>
<td>26</td>
<td>19 0.001 21.7388 0.0331</td>
<td>26</td>
</tr>
<tr>
<td>27</td>
<td>9 0.0021 2.6285 0.1919</td>
<td>27</td>
<td>14 0.0007 20.0450 0.0155</td>
<td>27</td>
</tr>
<tr>
<td>28</td>
<td>11 0.0022 2.6099 0.1688</td>
<td>28</td>
<td>14 0.0007 22.0968 0.0334</td>
<td>28</td>
</tr>
<tr>
<td>29</td>
<td>10 0.0014 2.5821 0.2593</td>
<td>29</td>
<td>9 0.0004 22.2419 0.1544</td>
<td>29</td>
</tr>
<tr>
<td>30</td>
<td>7 0.0013 2.5660 0.3993</td>
<td>30</td>
<td>3 0.0022 2.7790 0.2020</td>
<td>30</td>
</tr>
<tr>
<td>31</td>
<td>9 0.0016 2.5738 0.3887</td>
<td>31</td>
<td>6 0.0004 2.2236 0.4272</td>
<td>31</td>
</tr>
<tr>
<td>32</td>
<td>11 0.0035 2.6195 0.2084</td>
<td>32</td>
<td>7 0.0005 22.0844 0.0296</td>
<td>32</td>
</tr>
<tr>
<td>33</td>
<td>20 0.0041 2.6678 0.2833</td>
<td>33</td>
<td>11 0.0011 21.6807 0.0138</td>
<td>33</td>
</tr>
<tr>
<td>34</td>
<td>16 0.0033 2.6031 0.2024</td>
<td>34</td>
<td>10 0.0009 20.4462 0.0041</td>
<td>34</td>
</tr>
<tr>
<td>35</td>
<td>9 0.0007 2.5797 0.2816</td>
<td>35</td>
<td>3 0 21.0276 0.0489</td>
<td>35</td>
</tr>
<tr>
<td>36</td>
<td>7 0.0013 2.5841 0.3453</td>
<td>36</td>
<td>4 0.0002 20.9620 0.0485</td>
<td>36</td>
</tr>
<tr>
<td>37</td>
<td>9 0.0013 2.5841 0.3453</td>
<td>37</td>
<td>4 0.0002 20.9620 0.0485</td>
<td>37</td>
</tr>
<tr>
<td>38</td>
<td>10 0.0014 2.5836 0.3371</td>
<td>38</td>
<td>3 0.0002 21.2875 0.0327</td>
<td>38</td>
</tr>
<tr>
<td>39</td>
<td>3 0.0015 2.6003 0.3925</td>
<td>39</td>
<td>13 0.0014 15.4853 0.0012</td>
<td>39</td>
</tr>
<tr>
<td>40</td>
<td>18 0.0048 2.6026 0.2551</td>
<td>40</td>
<td>5 0.0004 20.6174 0.2828</td>
<td>40</td>
</tr>
</tbody>
</table>

Table A.15 shows the $e$-best sub-optimal local admissible sample size $\hat{\mu}_t$ and the parameter estimates $\hat{\mu}_{mk}$, $\hat{\sigma}_{mk}$ and $\hat{\theta}_{mk}$ for four energy commodities price at time $t_i$. This was based on the value of $r$ and the initial real data delay time $r = 20$. We further note that the range of the $e$-best sub-optimal local admissible sample size $\hat{\mu}_t$ for any time $t_i \in [21, 40] \cup [1145, 1165], t_i \in [2400, 2460], t_i \in [2865, 2885]$, and $t_i \in [373, 395]$ for natural gas, crude oil, coal and ethanol data, respectively, is $3 \leq \hat{\mu}_t \leq 20$. Moreover, all comments (Remark 4.1) that are made with regard to Table 2 regarding the four energy commodities remain valid with regard to Table A.15.

In Table A.16, the real and the LLGMM simulated price values for each of the four energy commodities: natural gas, crude oil, coal and ethanol are exhibited in columns 2-3, 6-7, 10-11, and 14-15, respectively. The absolute error of each of the energy commodity’s simulated value is shown in columns 4, 8, 12, 16, respectively.

---

55
Appendix B. Formulation of Aggregated Generalized Method of Moment (AGMM):

In this section, using the theoretical basis of the LLGMM and Remark 2.5 (Section 2), we generated aggregated state and parameter estimates based on the method for state and parameter estimation problems. The generalized method is then applied to energy commodity dynamic model (2.8). The results are compared with the LLGMM method.

Appendix B.1. AGMM Method Applied to Energy Commodity:

In this Subsection, using the aggregated parameter estimates \( \hat{a}, \hat{\mu}, \) and \( \sigma^2 \) described by the mean value of the estimated samples \( \{a_{h_i,\theta_i}\}_{i=1}^N, \{\mu_{h_i,\theta_i}\}_{i=1}^N \) and \( \{\sigma^2_{h_i,\theta_i}\}_{i=1}^N \) (Remark 2.5), respectively, we discuss the simulated price values for the four energy commodities. \( \hat{a}, \hat{\mu}, \) and \( \sigma^2 \) defined in (2.19) are referred to as aggregated parameter estimates of \( a, \mu, \) and \( \sigma^2 \) over the given entire finite interval of time. These estimates are derived using the following discretized
system:

\[
y_{it}^{ag} = y_{i,t-1}^{ag} + \bar{a}(\bar{\mu} - y_{i,t-1}^{ag})y_{i,t-1}^{ag}\Delta t + \bar{\sigma}^{2}/2 y_{i,t-1}^{ag}\Delta W_i
\]

(B.1)

where \(y_{it}^{ag}\) denotes the simulated value for \(y_t\) at time \(t_k\). The overall descriptive data statistic regarding the four energy commodities price and estimated parameters are recorded in Table B.17 below.

Table B.17 shows the descriptive statistics for \(a\), \(\mu\) and \(\sigma^2\) with time delay \(r = 20\). Moreover, \(\bar{\mu}\) is approximately close to the overall descriptive statistics of the mean \(\bar{Y}\) of the real data for each of the energy commodity shown in column 2. Also, \(\bar{\sigma}^2\) is approximately close to the overall descriptive statistics of the variance of \(\Delta\ln(Y) = \ln(Y_t) - \ln(Y_{t-1})\) in Column 5. Moreover, column 12 shows that the mean of the actual data set in Column 2 falls within the 95% confidence interval of \(\bar{\mu}\). This exhibits that the parameter \(\mu_{\text{est}, k}\) is the mean level of \(y_t\) at time \(t_k\).

Using the aggregated parameter estimates \(\bar{a}\), \(\bar{\mu}\), and \(\bar{\sigma}^2\) in Table B.17 (Column 6, 8, and 10), the simulated price values for the four energy commodities are shown in columns 3, 6, 9 and 12 of Table B.18.
Table B.18: Real, Simulation value using AGMM with $r = 20$.

<table>
<thead>
<tr>
<th>Natural gas</th>
<th>Real</th>
<th>Simulated</th>
</tr>
</thead>
<tbody>
<tr>
<td>21 2.759 2.449</td>
<td>21 2.759 2.449</td>
<td></td>
</tr>
<tr>
<td>22 2.740 2.437</td>
<td>22 2.740 2.437</td>
<td></td>
</tr>
<tr>
<td>23 2.742 2.636</td>
<td>23 2.742 2.636</td>
<td></td>
</tr>
<tr>
<td>24 2.562 2.625</td>
<td>24 2.562 2.625</td>
<td></td>
</tr>
<tr>
<td>25 2.406 2.903</td>
<td>25 2.406 2.903</td>
<td></td>
</tr>
<tr>
<td>26 2.586 2.451</td>
<td>26 2.586 2.451</td>
<td></td>
</tr>
<tr>
<td>27 2.557 2.399</td>
<td>27 2.557 2.399</td>
<td></td>
</tr>
<tr>
<td>28 2.633 2.543</td>
<td>28 2.633 2.543</td>
<td></td>
</tr>
<tr>
<td>29 2.515 2.515</td>
<td>29 2.515 2.515</td>
<td></td>
</tr>
<tr>
<td>30 2.658 2.866</td>
<td>30 2.658 2.866</td>
<td></td>
</tr>
<tr>
<td>31 2.467 2.295</td>
<td>31 2.467 2.295</td>
<td></td>
</tr>
<tr>
<td>32 2.466 2.296</td>
<td>32 2.466 2.296</td>
<td></td>
</tr>
<tr>
<td>33 2.663 2.465</td>
<td>33 2.663 2.465</td>
<td></td>
</tr>
<tr>
<td>34 2.559 2.491</td>
<td>34 2.559 2.491</td>
<td></td>
</tr>
<tr>
<td>35 2.488 2.810</td>
<td>35 2.488 2.810</td>
<td></td>
</tr>
<tr>
<td>36 2.463 2.428</td>
<td>36 2.463 2.428</td>
<td></td>
</tr>
<tr>
<td>37 2.663 2.463</td>
<td>37 2.663 2.463</td>
<td></td>
</tr>
<tr>
<td>38 2.663 2.463</td>
<td>38 2.663 2.463</td>
<td></td>
</tr>
<tr>
<td>39 2.663 2.463</td>
<td>39 2.663 2.463</td>
<td></td>
</tr>
<tr>
<td>40 2.663 2.463</td>
<td>40 2.663 2.463</td>
<td></td>
</tr>
</tbody>
</table>

Figure B.20 below shows a comparison between the real data set, simulated price using LLGMM and AGMM methods.
Figure B.20: (a), (b), (c) and (d) show the graphs of the real, simulated prices using the local lagged adaptive generalized method (LLGMM), and the simulated price using the average of the parameters for Henry Hub natural gas data [14], daily crude oil data [13], daily coal data [12], and weekly ethanol data [48], respectively, for $r = 20$. The red line represents the real data set $y_t$, the blue line represents the simulated prices using the LLGMM method, while the black line represents the simulated price (AGMM) using the aggregated parameter estimates $\hat{\delta}$, $\hat{\mu}$, and $\hat{\sigma}^2$ in Table B.17, Columns 6, 8, and 10, respectively. From these simulated graphs, it is clear that the LLGMM simulation results are more realistic than the AGMM simulation results. This exhibits the superiority of LLGMM over AGMM.

Comparison of Goodness-of-fit Measures for the LLGMM and AGMM methods using initial delay $r = 20$. 

59
Table B.19: Comparison of Goodness-of-fit Measures for the LLGMM and AGMM methods using initial delay \( r = 20 \).

<table>
<thead>
<tr>
<th>Goodness-of-fit Measure</th>
<th>LLGMM</th>
<th>AGMM</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Natural gas</td>
<td>Crude oil</td>
</tr>
<tr>
<td>( \hat{RAMS}E )</td>
<td>0.0674</td>
<td>0.4625</td>
</tr>
<tr>
<td>( \hat{AMA}D )</td>
<td>1.1318</td>
<td>24.5010</td>
</tr>
<tr>
<td>( \hat{AMB} )</td>
<td>1.1371</td>
<td>27.2707</td>
</tr>
</tbody>
</table>


The overall descriptive statistics of data sets regarding U.S. Treasury Bill interest rate and U.S.-U.K. foreign exchange rate are recorded in the following table for initial delay \( r = 20 \).

Table B.20: Descriptive Statistics for \( \bar{\beta}, \bar{\mu}, \bar{\delta}, \bar{\sigma}, \) and \( \bar{\gamma} \) for Interest rate data using initial delay \( r = 20 \).

<table>
<thead>
<tr>
<th>( \bar{Y} )</th>
<th>Std(( \bar{Y} ))</th>
<th>( \bar{\beta} )</th>
<th>Std(( \bar{\beta} ))</th>
<th>( \bar{\mu} )</th>
<th>Std(( \bar{\mu} ))</th>
<th>( \bar{\delta} )</th>
<th>Std(( \bar{\delta} ))</th>
<th>( \bar{\sigma} )</th>
<th>Std(( \bar{\sigma} ))</th>
<th>( \bar{\gamma} )</th>
<th>Std(( \bar{\gamma} ))</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.05667</td>
<td>0.0268</td>
<td>0.8739</td>
<td>1.8129</td>
<td>-3.8555</td>
<td>8.7608</td>
<td>1.4600</td>
<td>0.00</td>
<td>0.3753</td>
<td>0.5197</td>
<td>1.4877</td>
<td>0.1357</td>
</tr>
</tbody>
</table>

Table B.21: Descriptive Statistics for \( \bar{\beta}, \bar{\mu}, \bar{\delta}, \bar{\sigma}, \) and \( \bar{\gamma} \) for U.S.-U.K. foreign exchange rate data using initial delay \( r = 20 \).

<table>
<thead>
<tr>
<th>( \bar{Y} )</th>
<th>Std(( \bar{Y} ))</th>
<th>( \bar{\beta} )</th>
<th>Std(( \bar{\beta} ))</th>
<th>( \bar{\mu} )</th>
<th>Std(( \bar{\mu} ))</th>
<th>( \bar{\delta} )</th>
<th>Std(( \bar{\delta} ))</th>
<th>( \bar{\sigma} )</th>
<th>Std(( \bar{\sigma} ))</th>
<th>( \bar{\gamma} )</th>
<th>Std(( \bar{\gamma} ))</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.6249</td>
<td>0.1337</td>
<td>1.5120</td>
<td>2.1259</td>
<td>-1.1973</td>
<td>1.6811</td>
<td>1.4892</td>
<td>0.00</td>
<td>0.0243</td>
<td>0.0180</td>
<td>1.08476</td>
<td>1.0050</td>
</tr>
</tbody>
</table>

Tables B.20 and B.21 show the descriptive statistics for \( \bar{\beta}, \bar{\mu}, \bar{\delta}, \bar{\sigma}, \) and \( \bar{\gamma} \) for the U.S. TBYIR and the U.S.-U.K. FER data, respectively.

In Table B.22, the real and the LLGMM simulated rates of the US-TBYIR and the U.S.-U.K. foreign exchange rate (US-UK FER) are exhibited in the first and second columns, respectively. Using the aggregated parameter estimates \( \bar{\beta}, \bar{\mu}, \bar{\delta}, \bar{\sigma}, \) and \( \bar{\gamma} \) in the respective Tables B.20 (columns 3, 5, 7, 9 and 11) and Table B.21 (columns 3, 5, 7, 9, and 11), the simulated rates for the U.S. TYBIR and the U.S.-U.K. FER are shown in column 3 of Table B.22. These estimates are derived using the following discretized system:

\[
y^\text{AGMM}_t = y^\text{AGMM}_{t-1} + (\bar{\beta}y^\text{AGMM}_{t-1} + \bar{\mu}(y^\text{AGMM}_{t-1})^2 + \bar{\sigma}(y^\text{AGMM}_{t-1})^2)\Delta W_t \tag{B.2}
\]

where AGMM, \( y^\text{AGMM}_k \), \( y^\text{AGMM}_k \) at time \( t_k \) are defined in (B.1).

<table>
<thead>
<tr>
<th>$t$</th>
<th>Real</th>
<th>Simulated LLGMM</th>
<th>Simulated AGMM</th>
</tr>
</thead>
<tbody>
<tr>
<td>23</td>
<td>1.662</td>
<td>1.6444</td>
<td>1.6681</td>
</tr>
<tr>
<td>24</td>
<td>1.5102</td>
<td>1.5028</td>
<td>1.5445</td>
</tr>
<tr>
<td>25</td>
<td>1.7982</td>
<td>1.8056</td>
<td>1.5845</td>
</tr>
<tr>
<td>26</td>
<td>0.0609</td>
<td>0.0601</td>
<td>0.04</td>
</tr>
<tr>
<td>27</td>
<td>1.4965</td>
<td>1.4854</td>
<td>1.6208</td>
</tr>
<tr>
<td>28</td>
<td>0.0456</td>
<td>0.0463</td>
<td>0.0365</td>
</tr>
<tr>
<td>29</td>
<td>1.5138</td>
<td>1.5017</td>
<td>1.5287</td>
</tr>
<tr>
<td>30</td>
<td>0.0472</td>
<td>0.0479</td>
<td>0.0548</td>
</tr>
<tr>
<td>31</td>
<td>0.0427</td>
<td>0.043</td>
<td>0.023</td>
</tr>
<tr>
<td>32</td>
<td>1.5581</td>
<td>1.5635</td>
<td>1.6326</td>
</tr>
<tr>
<td>33</td>
<td>0.0537</td>
<td>0.053</td>
<td>0.0538</td>
</tr>
<tr>
<td>34</td>
<td>0.0464</td>
<td>0.0463</td>
<td>0.034</td>
</tr>
<tr>
<td>35</td>
<td>37</td>
<td>1.4276</td>
<td>1.4353</td>
</tr>
<tr>
<td>36</td>
<td>1.7856</td>
<td>1.7785</td>
<td>1.6206</td>
</tr>
<tr>
<td>37</td>
<td>1.8562</td>
<td>1.8964</td>
<td>1.5417</td>
</tr>
<tr>
<td>38</td>
<td>1.8207</td>
<td>1.8214</td>
<td>1.6147</td>
</tr>
<tr>
<td>39</td>
<td>1.5579</td>
<td>1.5601</td>
<td>1.6759</td>
</tr>
<tr>
<td>40</td>
<td>0.0384</td>
<td>0.0413</td>
<td>0.0339</td>
</tr>
<tr>
<td>41</td>
<td>1.8985</td>
<td>1.9002</td>
<td>1.606</td>
</tr>
<tr>
<td>42</td>
<td>0.0596</td>
<td>0.0602</td>
<td>0.0393</td>
</tr>
<tr>
<td>43</td>
<td>1.7902</td>
<td>1.7971</td>
<td>1.6221</td>
</tr>
<tr>
<td>44</td>
<td>0.0569</td>
<td>0.0588</td>
<td>0.0404</td>
</tr>
<tr>
<td>45</td>
<td>1.6097</td>
<td>1.6195</td>
<td>1.606</td>
</tr>
<tr>
<td>46</td>
<td>0.0532</td>
<td>0.0539</td>
<td>0.029</td>
</tr>
<tr>
<td>47</td>
<td>1.7902</td>
<td>1.7971</td>
<td>1.6221</td>
</tr>
<tr>
<td>48</td>
<td>0.06</td>
<td>0.0601</td>
<td>0.0483</td>
</tr>
</tbody>
</table>

In Table B.22, we show a side by side comparison of the estimates for the simulated value using LLGMM and AGMM methods for U.S. Treasury Bill interest rate and U.S.-U.K. foreign exchange rate using initial delay $r = 20$.

Figure B.21: Real, Simulated paths using LLGMM and AGMM methods for U.S. Treasury Bill interest rate and U.S.-U.K. foreign exchange rate for initial delay $r = 20$. 

61
Appendix C. Comparative study of the LLGMM with OCBGMM Methods:

In this Appendix, an additional detailed comparisons regarding the theoretical, graphical and performance of the LLGMM and OCBGMM methods are presented in Appendix C.1, Appendix C.2, and Appendix C.3, respectively. In fact, by employing three statistical goodness-of-fit measures [11], a comparative performance analysis of forecasting and ranking of the LLGMM and OCBGMM based methods are presented in Appendix C.3.

Appendix C.1. Theoretical Comparison Between LLGMM and OCBGMM

Based on the foundations of the analytical, conceptual, computational, mathematical, practical, statistical and theoretical motivations and developments outlined in Sections 2, 3, 4, 5 and 6, we summarize the comparison between the innovative approach LLGMM with the existing and newly developed OCBGMM methods in separate tables in a systematic manner.

In the following, we state the differences between the LLGMM method and existing orthogonality condition based GMM/IRGMM-Algebraic and the newly formulated GMM/IRGMM-Analytic methods together with AGMM.

Table C.23: Mathematical Comparison Between the LLGMM and OCBGMM

<table>
<thead>
<tr>
<th>Feature</th>
<th>LLGMM</th>
<th>OCBGMM-Algebraic</th>
<th>OCBGMM-Analytic</th>
<th>Justifications</th>
</tr>
</thead>
<tbody>
<tr>
<td>Composition:</td>
<td>Seven components</td>
<td>Five components</td>
<td>Five components</td>
<td>Sections 1, 2</td>
</tr>
<tr>
<td>Model:</td>
<td>Development</td>
<td>Selection</td>
<td>Development/ Selection</td>
<td>Sections 1, 2</td>
</tr>
<tr>
<td>Goal:</td>
<td>Validation</td>
<td>Specification/Testing</td>
<td>Validation/Testing</td>
<td>Sections 1, 2</td>
</tr>
<tr>
<td>Discrete-Time Scheme:</td>
<td>Constructed from SDE</td>
<td>Using Econometric specification</td>
<td>Constructed from SDE</td>
<td>Remark 2.8</td>
</tr>
<tr>
<td>Formation of Orthogonality Vec-</td>
<td>Using stochastic calculus</td>
<td>Formed using algebraic manipulation</td>
<td>Using Stochastic calculus</td>
<td>Remarks 2.2, 2.7, 2.8</td>
</tr>
<tr>
<td>tor:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table C.24: Intercomponent Interaction Comparison Between LLGMM and OCBGMM

<table>
<thead>
<tr>
<th>Feature</th>
<th>LLGMM</th>
<th>OCBGMM-Algebraic</th>
<th>OCBGMM-Analytic</th>
<th>Justifications</th>
</tr>
</thead>
<tbody>
<tr>
<td>Moment Equations:</td>
<td>Local Lagged adaptive process</td>
<td>Single/ global system</td>
<td>Single/ global system</td>
<td>Remarks 3.2a, and 3.2b</td>
</tr>
<tr>
<td>Type of Moment Equations:</td>
<td>Local lagged adaptive process</td>
<td>Single-shot</td>
<td>Single-shot</td>
<td>Remarks 2.6, 2.8, and 2.9</td>
</tr>
<tr>
<td>Component Interconnections:</td>
<td>Strongly connected</td>
<td>Weakly connected</td>
<td>Weakly connected</td>
<td>Remarks 2.6, 2.7, 2.8, 2.9, and 3.2</td>
</tr>
<tr>
<td>Dynamic and Static:</td>
<td>Discrete-time Dynamic</td>
<td>Static</td>
<td>Static</td>
<td>Remarks 3.2 and Lemma 1 (Section 2)</td>
</tr>
</tbody>
</table>
### Table C.25: Conceptual Computational Comparison Between LLGMM and OCBGMM

<table>
<thead>
<tr>
<th>Feature</th>
<th>LLGMM</th>
<th>OCBGMM-Algebraic</th>
<th>OCBGMM-Analytic</th>
<th>Justifications</th>
</tr>
</thead>
<tbody>
<tr>
<td>Local admissible Lagged Data Size:</td>
<td>Multi-choice</td>
<td>Single-choice/data size</td>
<td>Single-choice/data size</td>
<td>Definition 3.3, Remark 3.2, Subsection 3.2</td>
</tr>
<tr>
<td>Local admissible class of lagged finite restriction sequences</td>
<td>Multi-choice</td>
<td>Single-choice data sequence</td>
<td>Single-choice data sequence</td>
<td>Adapted finite restricted sample data: Definition 3.4, Remark 3.2, Subsection 3.2</td>
</tr>
<tr>
<td>Local admissible finite sequence parameter estimates:</td>
<td>Multi-choice</td>
<td>Single-shot estimate</td>
<td>Single-shot estimates</td>
<td>Subsection 3.2</td>
</tr>
<tr>
<td>Local admissible sequence of finite state simulation values:</td>
<td>Multi-choice</td>
<td>Single-choice</td>
<td>Single-choice</td>
<td>Remark 3.2, Subsection 3.3</td>
</tr>
<tr>
<td>Quadratic Mean Square $\epsilon$-sub-optimal errors:</td>
<td>Multi-choice</td>
<td>Single-choice</td>
<td>Single-choice</td>
<td>Remark 3.2, Subsection 3.3</td>
</tr>
<tr>
<td>$\epsilon$-sub-optimal local lagged sample size:</td>
<td>Multi-choice</td>
<td>Single-choice</td>
<td>Single-choice</td>
<td>Definition 12, Remark 3.2, Subsection 3.3</td>
</tr>
<tr>
<td>$\epsilon$-best sub optimal sample size:</td>
<td>$\epsilon$-best sub optimal choice</td>
<td>No-choice</td>
<td>No-choice</td>
<td>Remark 3.2, Subsection 3.3</td>
</tr>
<tr>
<td>$\epsilon$-best sub optimal parameter estimated:</td>
<td>$\epsilon$-best estimators</td>
<td>No-choice</td>
<td>No-choice</td>
<td>Remark 3.2, Subsection 3.3</td>
</tr>
<tr>
<td>$\epsilon$-best sub optimal state estimate:</td>
<td>$\epsilon$-best sub optimal choice</td>
<td>No-choice</td>
<td>No-choice</td>
<td>Remark 3.2, Subsection 3.3</td>
</tr>
</tbody>
</table>

### Table C.26: Theoretical Performance Comparison Between LLGMM and OCBGMM

<table>
<thead>
<tr>
<th>Feature</th>
<th>LLGMM</th>
<th>OCBGMM-Algebraic</th>
<th>OCBGMM-Analytic</th>
<th>Justifications</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data Size:</td>
<td>Reasonable Size</td>
<td>Large Data Size</td>
<td>Large Size</td>
<td>For Respectable results</td>
</tr>
<tr>
<td>Stationary Condition:</td>
<td>Not required</td>
<td>Need Ergotic/Asymptotic stationarity</td>
<td>Need Ergodic/Asymptotic</td>
<td>For Reasonable results</td>
</tr>
<tr>
<td>Multi-level optimization:</td>
<td>At least 2 level hierarchical optimization</td>
<td>Single-shot</td>
<td>Single-shot</td>
<td>Not comparable</td>
</tr>
<tr>
<td>Computational Stability:</td>
<td>Algorithm converges in a single / double digit trials</td>
<td>Single-choice</td>
<td>Single-choice</td>
<td>Simulation results</td>
</tr>
<tr>
<td>Significance of lagged adaptive process:</td>
<td>Stabilizing agent</td>
<td>Non-existence of the feature</td>
<td>Non-existence</td>
<td>Not comparable</td>
</tr>
<tr>
<td>Operation:</td>
<td>Operates like Discrete time Dynamic Process</td>
<td>Operates like a static dynamic process</td>
<td>Operates like static process</td>
<td>Obvious, details see Sections 3, 4, 5, 6 and 7</td>
</tr>
</tbody>
</table>
Appendix C.2. Graphical Comparison of the LLGMM with OCBGMM Methods

Parameter Estimates of (7.1) using LLGMM and OCBGMM Methods: Using the LLGMM method, the parameter estimates $\hat{\alpha}_{m,k}$, $\hat{\beta}_{m,k}$, $\sigma_{m,k}$, and $\gamma_{m,k}$ of (7) in USTBIR are shown in Table C.27. Here, we use $\epsilon = 0.001$, $p = 2$, and initial delay $r = 20$.

Table C.27: Estimates for $\hat{\alpha}_{m,k}$, $\hat{\beta}_{m,k}$, $\sigma_{m,k}$, and $\gamma_{m,k}$ in the model (7.1) for U.S. Treasury Bill interest rate data.

<table>
<thead>
<tr>
<th>$t$</th>
<th>$\hat{\alpha}_{m,k}$</th>
<th>$\hat{\beta}_{m,k}$</th>
<th>$\hat{\sigma}_{m,k}$</th>
<th>$\hat{\gamma}_{m,k}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>21</td>
<td>-0.0334</td>
<td>-0.7143</td>
<td>0.0463</td>
<td>0.8464</td>
</tr>
<tr>
<td>22</td>
<td>-0.0427</td>
<td>-0.8254</td>
<td>0.0464</td>
<td>0.8464</td>
</tr>
<tr>
<td>23</td>
<td>-0.0413</td>
<td>-0.8379</td>
<td>0.0464</td>
<td>0.8464</td>
</tr>
<tr>
<td>24</td>
<td>-0.0425</td>
<td>-0.8109</td>
<td>0.0464</td>
<td>0.8464</td>
</tr>
<tr>
<td>25</td>
<td>-0.0413</td>
<td>-0.7833</td>
<td>0.0464</td>
<td>0.8464</td>
</tr>
<tr>
<td>26</td>
<td>-0.0425</td>
<td>-0.8171</td>
<td>0.0464</td>
<td>0.8464</td>
</tr>
<tr>
<td>27</td>
<td>-0.0413</td>
<td>-0.7833</td>
<td>0.0464</td>
<td>0.8464</td>
</tr>
</tbody>
</table>

Table C.27 shows the parameter estimates of $\hat{\alpha}_{m,k}$, $\hat{\beta}_{m,k}$, $\sigma_{m,k}$, and $\gamma_{m,k}$ in the model (7.1) for U.S. Treasury Bill interest rate data. As noted before, the range of the $\epsilon$-best sub-optimal local admissible sample size $\hat{m}_k$ for ant time $t_k \in [21, 45] \cup [420, 445]$ is $2 \leq \hat{m}_k \leq 20$. We also draw the similar conclusions (a) to (e) as outlined in Remark 4.1.

Figure C.22: Real and simulated path using LLGMM method

Figure C.22 shows the real and simulated path of the monthly interest rate data [44] using the LLGMM method. The root mean square error of the simulated value is 0.0027.
Figure C.23: Comparison of simulation result using GMM-Analytic and GMM-Algebraic methods

Figures C.23 (a) and (b) show the real and simulated value of the monthly interest rate data [44] using the GMM-Analytic and GMM-Algebraic methods, respectively. The root mean square errors of simulated values are shown in Table 10. Figure C.23(c) shows the comparison between the real and simulated values of GMM-Analytic and GMM-Algebraic methods. The red, green, and blue line represent the real data path [44], the simulated path using GMM-Algebraic, and the GMM-Analytic, respectively.
Figures C.24(a) and (b) show the real and simulated value of the monthly interest rate data [44] using the IRGMM-Analytic and IRGMM-Algebraic methods, respectively. The root mean square errors of simulated values are shown in Table 10. Figure C.24(c) shows the comparison between the real and simulated values of IRGMM-Analytic and IRGMM-Algebraic method. The red, green, and blue curve represents the real data path [44], simulated path using IRGMM-Algebraic, and GMM-Analytic, respectively.
Figure C.25: Comparison of simulation results of GMM-analytic, IRGMM-analytic, GMM-algebraic and IRGMM-algebraic methods together with AGMM method

Figure C.25(a) compares the simulation results using GMM-algebraic and IRGMM-algebraic. The blue denotes the GMM-algebraic simulation curve while the green line represents the IRGMM-algebraic simulation curve. Figure C.25(b) compares the simulation results using the GMM-Analytic, and IRGMM-Analytic represented by the black, and green lines, respectively. Figure C.25(c) compares the simulation results using the GMM-Algebra, IRGMM-Algebra, and LLGMM represented by the black, green, and blue lines, respectively.
Figure C.26: Comparison of simulation results for GMM-Analytic, IRGMM-analytic, GMM-Algebraic and IRGMM-Algebraic methods as well as the LLGMM and AGMM methods.

Figure C.26(a) compares the simulation results using GMM-analytic, IRGMM-analytic and the LLGMM methods. The GMM-analytic, IRGMM-analytic and the LLGMM simulation results are exhibited by the black, green and blue lines, respectively. Figure C.26(b) compares the simulation results using the GMM-Algebraic, IRGMM-Algebraic, LLGMM and AGMM methods represented by the black, green, blue and magenta lines, respectively. Figure C.26(c) compares the simulation results using the GMM-Analytic, IRGMM-Analytic, LLGMM and AGMM methods represented by the black, green, blue and magenta curves, respectively.

Comparative Analysis of Forecasting with 95% Confidence Intervals: Using data set from June 1964 to December 1989, the parameters of model (7.1) are estimated. Using these parameter estimates, we forecasted the monthly interest rate for January 1, 1990 to December 31, 2004. Table C.28 shows the parameter estimates in the context of the data from June 1964 to December 1989.

68
Table C.28: Parameter estimates in (7.1) in the context of data July 1964-December 1989

<table>
<thead>
<tr>
<th>Method</th>
<th>$\alpha$</th>
<th>$\beta$</th>
<th>$\sigma$</th>
<th>$\gamma$</th>
</tr>
</thead>
<tbody>
<tr>
<td>GMM-Algebraic</td>
<td>0.0033</td>
<td>-0.051</td>
<td>0.4121</td>
<td>1.5311</td>
</tr>
<tr>
<td>GMM-Analytic</td>
<td>0.0009</td>
<td>-0.0155</td>
<td>0.0197</td>
<td>0.4854</td>
</tr>
<tr>
<td>IRGMM-Algebraic</td>
<td>0.0023</td>
<td>-0.0421</td>
<td>0.3230</td>
<td>1.3112</td>
</tr>
<tr>
<td>IRGMM-Analytic</td>
<td>0.0084</td>
<td>-0.1436</td>
<td>0.1073</td>
<td>1.3641</td>
</tr>
<tr>
<td>AGMM</td>
<td>0.0154</td>
<td>-0.2497</td>
<td>0.2949</td>
<td>1.4414</td>
</tr>
</tbody>
</table>

Figure C.27: Real, Simulation and Forecast state estimates using LLGMM method

In Figure C.27, region $S$ shows the real, simulated value using the monthly interest rate data from June 30, 1964 to December 31, 1989 [44]. In the $F$ region (forecasting region), the estimated parameters in the context of the data set [44] are used to forecast interest rate from January 1, 1990 to December 31, 2004 using the LLGMM method.
Figure C.28: Real, Simulation and Forecast estimates using GMM-Algebraic, GMM-Analytic, IRGMM-Algebraic and IRGMM-Analytic methods

Figure C.28(a), (b), (c) and (d) exhibit the side-by-side comparison of the simulated forecasting results of the GMM-Analytic, GMM-Algebraic, IRGMM-Analytic, and IRGMM-Algebraic methods, respectively. The $S$ region represents the simulation region based on the real data while the $F$ region represents the forecasting region. In addition, the 95% confidence level of the simulation results are also shown (in black).

Appendix C.3. Performance Comparisons of LLGMM Method with Existing and Newly Introduced OCBGMM Methods Using Energy Commodity Stochastic Model

Using the stochastic dynamic model in (2.8) of energy commodity represented by stochastic differential equation:

$$dy = a(y - \mu)dt + \sigma(t, y)dw(t), \quad y(t_0) = y_0,$$

the orthogonality condition parameter vector (OCPV) is described in (2.13) in Remark (2.2). Based on discretized scheme using the econometric specification [9], the orthogonality condition parameter vector in the context of algebraic manipulation is as [9]: OCBGMM looks like

$$\begin{align*}
    y_t - y_{t-1} &= a y_{t-1} (\mu - y_{t-1}) \Delta t \\
    y_{t-1} (y_t - y_{t-1}) &= a y_{t-1} (\mu - y_{t-1}) \Delta t \\
    \left(y_t - y_{t-1} - a y_{t-1} (\mu - y_{t-1}) \Delta t\right)^2 - \sigma^2 y_{t-1}^2
\end{align*}$$

(C.2)

The goodness-of-fit measures are computed using pseudo-data sets of the same sample size as the real data set: (i) $N = 1184$ days for natural gas data, (ii) $N = 4165$ days for crude oil data, (iii) $N = 3470$ for coal data, and (iv) $N = 438$ weeks for ethanol data. The smallest value of $RAME$ for all method is italicized.
Table C.29: Comparison of Parameter estimates of model (C.1) using GMM-Algebraic, GMM-Analytic, IRGMM-Algebraic, IRGMM-Analytic and AGMM for natural Gas Data

<table>
<thead>
<tr>
<th>Method</th>
<th>( \alpha )</th>
<th>( \mu )</th>
<th>( \sigma^2 )</th>
<th>( \text{RAMS}E )</th>
<th>( \text{AMAD} )</th>
<th>( \text{AMB} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>GMM-Algebraic</td>
<td>0.0023</td>
<td>5.3312</td>
<td>0.0019</td>
<td>1.5119</td>
<td>0.0663</td>
<td>1.1488</td>
</tr>
<tr>
<td>GMM-Analytic</td>
<td>0.0018</td>
<td>5.4106</td>
<td>0.0015</td>
<td>1.5014</td>
<td>0.0538</td>
<td>1.1677</td>
</tr>
<tr>
<td>IRGMM-Algebraic</td>
<td>0.2000</td>
<td>4.9966</td>
<td>0.0010</td>
<td>1.4985</td>
<td>0.0050</td>
<td>1.2299</td>
</tr>
<tr>
<td>IRGMM-Analytic</td>
<td>0.1998</td>
<td>4.4917</td>
<td>0.0011</td>
<td>1.4901</td>
<td>0.0044</td>
<td>1.2329</td>
</tr>
<tr>
<td>AGMM</td>
<td>0.1867</td>
<td>4.5538</td>
<td>0.0013</td>
<td>1.4968</td>
<td>0.0068</td>
<td>1.2267</td>
</tr>
<tr>
<td>LLGMM</td>
<td>0.0674</td>
<td></td>
<td></td>
<td>1.1318</td>
<td>1.1371</td>
<td></td>
</tr>
</tbody>
</table>

Table C.30: Comparison of Parameter estimates of model (C.1) using GMM-Algebraic, GMM-Analytic, IRGMM-Algebraic, IRGMM-Analytic and AGMM for crude oil data

<table>
<thead>
<tr>
<th>Method</th>
<th>( \alpha )</th>
<th>( \mu )</th>
<th>( \sigma^2 )</th>
<th>( \text{RAMS}E )</th>
<th>( \text{AMAD} )</th>
<th>( \text{AMB} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>GMM-Algebraic</td>
<td>0.0023</td>
<td>54.4847</td>
<td>0.0005</td>
<td>39.2853</td>
<td>0.5577</td>
<td>29.1587</td>
</tr>
<tr>
<td>GMM-Analytic</td>
<td>0.0021</td>
<td>51.2145</td>
<td>0.0006</td>
<td>38.8007</td>
<td>0.5181</td>
<td>28.7414</td>
</tr>
<tr>
<td>IRGMM-Algebraic</td>
<td>0.0000</td>
<td>88.5951</td>
<td>0.0005</td>
<td>30.7511</td>
<td>0.0920</td>
<td>27.5791</td>
</tr>
<tr>
<td>IRGMM-Analytic</td>
<td>0.0021</td>
<td>51.2195</td>
<td>0.0005</td>
<td>28.9172</td>
<td>0.2496</td>
<td>27.3564</td>
</tr>
<tr>
<td>AGMM</td>
<td>0.0215</td>
<td>54.0307</td>
<td>0.0005</td>
<td>30.7776</td>
<td>0.0857</td>
<td>27.3050</td>
</tr>
<tr>
<td>LLGMM</td>
<td>0.4625</td>
<td></td>
<td></td>
<td>24.501</td>
<td>27.2707</td>
<td></td>
</tr>
</tbody>
</table>

Table C.31: Comparison of Parameter estimates of model (C.1) using GMM-Algebraic, GMM-Analytic, IRGMM-Algebraic, IRGMM-Analytic and AGMM for coal Data

<table>
<thead>
<tr>
<th>Method</th>
<th>( \alpha )</th>
<th>( \mu )</th>
<th>( \sigma^2 )</th>
<th>( \text{RAMS}E )</th>
<th>( \text{AMAD} )</th>
<th>( \text{AMB} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>GMM-Algebraic</td>
<td>0.0000</td>
<td>94.4847</td>
<td>0.0006</td>
<td>22.6866</td>
<td>0.2015</td>
<td>16.3444</td>
</tr>
<tr>
<td>GMM-Analytic</td>
<td>0.0000</td>
<td>94.4446</td>
<td>0.0006</td>
<td>21.6564</td>
<td>0.2121</td>
<td>16.3264</td>
</tr>
<tr>
<td>IRGMM-Algebraic</td>
<td>0.0027</td>
<td>34.4838</td>
<td>0.0013</td>
<td>17.6894</td>
<td>0.3438</td>
<td>13.4981</td>
</tr>
<tr>
<td>IRGMM-Analytic</td>
<td>0.0021</td>
<td>23.1151</td>
<td>0.0005</td>
<td>17.6869</td>
<td>0.3448</td>
<td>13.4989</td>
</tr>
<tr>
<td>AGMM</td>
<td>0.0464</td>
<td>27.0567</td>
<td>0.0014</td>
<td>17.7620</td>
<td>0.0833</td>
<td>13.106</td>
</tr>
<tr>
<td>LLGMM</td>
<td>0.4794</td>
<td></td>
<td></td>
<td>9.4009</td>
<td>12.8370</td>
<td></td>
</tr>
</tbody>
</table>

Table C.32: Comparison of Parameter estimates of model (C.1) using GMM-Algebraic, GMM-Analytic, IRGMM-Algebraic, IRGMM-Analytic and AGMM for Ethanol data

<table>
<thead>
<tr>
<th>Method</th>
<th>( \alpha )</th>
<th>( \mu )</th>
<th>( \sigma^2 )</th>
<th>( \text{RAMS}E )</th>
<th>( \text{AMAD} )</th>
<th>( \text{AMB} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>GMM-Algebraic</td>
<td>0.0000</td>
<td>94.4847</td>
<td>0.0006</td>
<td>22.6866</td>
<td>0.2015</td>
<td>16.3444</td>
</tr>
<tr>
<td>GMM-Analytic</td>
<td>0.0000</td>
<td>94.4446</td>
<td>0.0006</td>
<td>21.6564</td>
<td>0.2121</td>
<td>16.3264</td>
</tr>
<tr>
<td>IRGMM-Algebraic</td>
<td>0.0014</td>
<td>3.4506</td>
<td>0.0026</td>
<td>0.5844</td>
<td>0.0322</td>
<td>0.4346</td>
</tr>
<tr>
<td>IRGMM-Analytic</td>
<td>0.0015</td>
<td>3.4506</td>
<td>0.0026</td>
<td>0.5813</td>
<td>0.0336</td>
<td>0.4303</td>
</tr>
<tr>
<td>AGMM</td>
<td>0.3167</td>
<td>2.166</td>
<td>0.0018</td>
<td>0.4356</td>
<td>0.0035</td>
<td>0.3579</td>
</tr>
<tr>
<td>LLGMM</td>
<td>0.0375</td>
<td></td>
<td></td>
<td>0.3213</td>
<td>0.3566</td>
<td></td>
</tr>
</tbody>
</table>

71
Tables C.29, C.30, C.31, and C.32 show a comparison parameter estimates of model (C.1) and the goodness-of-fit measures $\hat{\text{RAMS}_E}$, $\hat{\text{AMAD}}$ and $\hat{\text{AMB}}$ using GMM-Algebraic, GMM-Analytic, IRGMM-Algebraic, IRGMM-Analytic, AGMM and LLGMM method for the daily natural gas data [14], daily crude oil data [13], daily coal data [12], and weekly ethanol data [48], respectively. The LLGMM estimates are derived using initial delay $r = 20$, $p = 2$ and $\epsilon = 0.001$. Among all methods under study, the LLGMM method generates the smallest $\hat{\text{RAMS}_E}$ value. In fact, the $\hat{\text{RAMS}_E}$ value is smaller than the $1/22$, $1/62$, $1/36$, and $1/10$ of any other $\hat{\text{RAMS}_E}$ values regarding the natural gas, crude oil, coal and ethanol, respectively. This exhibits the superiority of the LLGMM method over all other methods. We further observe that the LLGMM approach yields the smallest $\hat{\text{AMB}}$ and highest $\hat{\text{AMAD}}$ value regarding the natural gas, crude oil, coal and ethanol. The high value of $\hat{\text{AMAD}}$ for the LLGMM method signifies that LLGMM method captures the influence of random environmental fluctuations on the dynamic of energy commodity process. From Remark 4.3, the smallest $\hat{\text{RAMS}_E}$, highest $\hat{\text{AMAD}}$ and smallest $\hat{\text{AMB}}$ value under the LLGMM method exhibits the superior performance under the three goodness-of-fit measures.

### Ranking of Methods under Goodness of Fit Measure

<table>
<thead>
<tr>
<th>Method</th>
<th>Natural gas</th>
<th>Crude oil</th>
<th>Coal</th>
<th>Ethanol</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\hat{\text{RAMS}_E}$</td>
<td>AMAD</td>
<td>AMB</td>
<td>$\hat{\text{RAMS}_E}$</td>
</tr>
<tr>
<td>GMM-Algebraic</td>
<td>6</td>
<td>2</td>
<td>2</td>
<td>6</td>
</tr>
<tr>
<td>GMM-Analytic</td>
<td>5</td>
<td>3</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>IRGMM-Algebraic</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>IRGMM-Analytic</td>
<td>2</td>
<td>6</td>
<td>6</td>
<td>2</td>
</tr>
<tr>
<td>AGMM</td>
<td>3</td>
<td>4</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>LLGMM</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

Table C.33: Ranking result for Natural gas, Crude oil, Coal and Ethanol Under three Statistical measures.

**Remark Appendix C.1.** The ranking of LLGMM is top one in all three Goodness-of-fit statistical measures for all four energy commodity data sets. Moreover, one of the IRGMM-Analytic and AGMM is ranked either as top two or three under $\hat{\text{RAMS}_E}$ measure. This exhibits the influence of the usage of stochastic calculus based orthogonality condition parameter vectors (OCPV-Analytic).

### Appendix D. Comparative analysis of LLGMM with Existing Nonparametric Statistical Methods

In this section, we compare our LLGMM method with existing nonparametric methods. We consider the following existing nonparametric methods.

**Appendix D.1. Nonparametric estimation of nonlinear dynamics by metric-based local linear approximation (LLA)**

The LLA method [39] assumes no functional form of a given model but estimates from experimental data by approximating the curve implied by the function of the tangent plane around the neighborhood of a tangent point. Suppose the state of interest $x_t$ at time $t$ is differentiable with respect to $t$ and satisfies $dx_t = f(x_t)dt$, where $f: \mathbb{R}^n \to \mathbb{R}$. 
is a smooth map, \( x_t \in \mathbb{R}^k \). The approximation of the curve \( f(x_t) \) in a neighborhood \( U_\epsilon(x_0) = \{ x : d(x, x_0) < \epsilon \} \) is defined by a tangent plane at \( x_0 \):

\[
y_t = f(x_0) + \sum_{i=1}^{k} \frac{\partial f}{\partial x_i}(x_0)(x_i - x_0),
\]

where \( d \) is a metric on \( \mathbb{R}^k \). Allowing error in the equation and assigning a weight \( w(x_t) \) to each error term \( \epsilon_t \), the method reduces to estimating parameters \( \beta_i = \frac{\partial f}{\partial x_i}(x_0), \) \( i = 1, 2, ..., k \) in the equation

\[
w(t)y_t = \beta_0 \cdot w(x_t) + \sum_{i=1}^{k} \beta_i \cdot w(x_i)(x_i - x_0),
\]

Applying the standard linear regression approach, the least square estimate \( \hat{\mathbf{\beta}} \) is given by

\[
\hat{\mathbf{\beta}} = (\mathbf{\tilde{X}}^T \mathbf{\tilde{X}})^{-1} \mathbf{\tilde{X}}^T \mathbf{\tilde{Y}},
\]

where

\[
\begin{align*}
\mathbf{\tilde{x}_i} & = (w(x_{t_1})(x_{t_1,i} - x_{0,i}), ..., w(x_{t_n})(x_{t_n,i} - x_{0,i}))^T, \quad i = 1, ..., k \\
\mathbf{\tilde{w}} & = (w(x_{t_1}), ..., w(x_{t_n}))^T \\
\mathbf{\tilde{Y}} & = (w(x_{t_1})y_{t_1}, ..., w(x_{t_n})y_{t_n})^T \\
\mathbf{\tilde{X}} & = (\mathbf{\tilde{w}}, \mathbf{\tilde{x}_1}, ..., \mathbf{\tilde{x}_k})
\end{align*}
\]

Particularly, the trajectory \( f(x_t) \) is estimated by choosing \( x_0 = x_{t_i} \), for each \( i = 1, 2, ..., n \). As discussed in [39], we use \( d(x, x_0) = |x - x_0| \), where \(| \cdot | \) is the standard Euclidean metric on \( \mathbb{R}^k \), and \( w(x) = \phi(d(x, x_0)) \), where \( \phi(u) = K(u/\epsilon) \) and \( K \) is the Epanechnikov Kernel [39] \( K(x) = 0.75(1 - x^2)_+ \).

Appendix D.2. Risk Estimation and Adaptation after Coordinate Transformation (REACT) method

Given \( n \) pairs of observations \( (x_1, Y_1), ..., (x_n, Y_n) \). Using the REACT method [47], the response variable \( Y \) is related to the covariate \( x \) (called a feature) by the equation

\[
Y_i = r(x_i) + \sigma \epsilon_i
\]

73
where \( \epsilon_i \sim N(0,1) \) are identically independently distributed, and \( x_i = \frac{i}{n}, \ i = 1, 2, \ldots, n \) and the function \( r(x) \), approximated using orthogonal cosine basis \( \phi_j, \ i = 1, 2, 3, \ldots \) of \([0,1] \) described by

\[
\phi_1(x) \equiv 1, \quad \phi_j(x) = \sqrt{2} \cos((j-1)\pi x), \quad j \geq 2. \quad (D.3)
\]

is expanded as

\[
r(x) = \sum_{j=1}^{\infty} \theta_j \phi_j(x), \quad (D.4)
\]

where \( \theta_j = \int_0^1 \phi_j(x) r(x) dx \), is approximated. The function estimator \( \hat{r}(x) = \sum_{j=1}^{J} Z_j \phi_j(x) \), where

\[
J \leq n, \quad Z_j = \frac{1}{n} \sum_{i=1}^{n} Y_i \phi_j(x_i), \quad j = 1, 2, \ldots, n
\]

and \( \hat{J} \) is found so that the risk estimator \( \hat{R}(J) = \frac{\hat{\sigma}^2}{n} + \sum_{j=\hat{J}+1}^{n} \left( Z_j^2 - \frac{\hat{\sigma}^2}{n} \right) \) is minimized, \( \hat{\sigma}^2 \) is the estimator of variance of \( Z_j \).

Appendix D.3. Exponential Moving Average method (EMA)

The EMA [27] for an observation \( y_t \) at time \( t \) may be calculated recursively as

\[
S_t = \alpha y_t + (1-\alpha)S_{t-1}, \quad t = 1, 2, 3, \ldots, n \quad (D.5)
\]

where \( 0 < \alpha \leq 1 \) is a constant that determines the depth of memory of \( S_t \).

Appendix D.4. Goodness-of-fit Measures for the LLA, REACT, and EMA methods

In this subsection, we show the Goodness-of-fit Measures for the LLA, REACT, and EMA methods. We use \( \hat{J} = 183 \) for the REACT method and \( \alpha = 0.5 \) for the EMA method.

<table>
<thead>
<tr>
<th>Goodness of fit Measure</th>
<th>LLGMM method</th>
<th>LLA method</th>
<th>REACT method</th>
<th>EMA method</th>
</tr>
</thead>
<tbody>
<tr>
<td>Natural gas</td>
<td>Crude oil</td>
<td>Coal</td>
<td>Ethanol</td>
<td>Natural gas</td>
</tr>
<tr>
<td>( \hat{R}_{AMSE} )</td>
<td>0.0874</td>
<td>0.0425</td>
<td>0.4724</td>
<td>0.8755</td>
</tr>
<tr>
<td>( \hat{R}_{AMSE} )</td>
<td>1.1378</td>
<td>24.9160</td>
<td>9.4688</td>
<td>0.3513</td>
</tr>
<tr>
<td>( \hat{R}_{AMSE} )</td>
<td>1.1378</td>
<td>27.2797</td>
<td>12.4876</td>
<td>0.3866</td>
</tr>
</tbody>
</table>

Table D.3: Goodness-of-fit Measures for the LLGMM, REACT, and EMA methods

Comparison of the results derived using these non-parametric methods with the LLGMM method show that the results derived using the LLGMM method is far better than the results of the nonparametric methods.

Graphical Comparison of the LLGMM with LLA, REACT, and EMA Methods:
Figure D.29: Real and simulated curve using LLA method

Figure D.29: (a), (b), (c) and (d) show the graphs of the Real and Simulated Spot Prices for the daily Henry Hub natural gas data [14], daily crude oil data [13], daily coal data [12], and weekly ethanol data [48], respectively, using LLA method. The red line represents the real data $y_k$ while the blue line represents the simulated value.
Figure D.30: (a), (b), (c) and (d) show the graphs of the Real and Simulated Spot Prices for the daily Henry Hub natural gas data [14], daily crude oil data [13], daily coal data [12], and weekly ethanol data [48], respectively, using the REACT method. The red line represents the real data $y_k$ while the blue line represents the simulated value.
Figure D.31: Real and simulated curve using EMA method

Figure D.31: (a), (b), (c) and (d) show the graphs of the Real and Simulated Spot Prices for the daily Henry Hub natural gas data [14], daily crude oil data [13], daily coal data [12], and weekly ethanol data [48], respectively, using the EMA method. The red line represents the real data $y_k$ while the blue line represents the simulated value.
Figure D.32: Comparison of Real and simulated curve using LLGMM, LLA, REACT, and EMA method

Figure D.32: (a), (b), (c) and (d) show the graphs of the Real and Simulated Spot Prices for the daily Henry Hub natural gas data [14], daily crude oil data [13], daily coal data [12], and weekly ethanol data [48], respectively, using the LLGMM, LLA, REACT, and EMA method. The red line represents the real data.

Acknowledgements

This research is supported by the Mathematical Sciences Division, the U.S. Army Office, under Grant Numbers W911NF-12-1-0090 and W911NF-15-1-0182.
References


[34] Paathong, Arnut and Ladde, G. S. , Agent-Based Modeling Simulation under Local Network Externality, Economic Interaction and Coordination, 9 (2014), 1-26.


