A

Mean Reversion And the Volatility of Interest Rates: Essays on the Monte-Carlo Simulation And Empirical Studies Based on Different Sampling Frequencies

by

TANA TANARUGSACHOCK

A dissertation submitted to the Graduate Faculty in Economics in partial fulfillment of the requirements for the degree of Doctor of Philosophy,
The City University of New York

2002

UMI Number: 3047270

ProQuest Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346

Approval Page

This manuscript has been read and accepted for the Graduate Faculty in Economics in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.

04/04/02
Date

Chair of Examining Committee
Prof. Temisan Agbeyegbe

04/04/02
Date

Executive Officer
Prof. Thom Thurston

04/04/02
Date

Supervisory Committee
Prof. Salih Neftci

The City University of New York

Abstract

Mean Reversion And the Volatility of Interest Rates: Essays On the Monte-Carlo Simulation And Empirical Studies Based on Different Sampling Frequencies

by

Tana Tanarugsachock

Advisor: Professor Terence Agbeyegbe

This dissertation consists of five chapters. The first chapter is the introduction to this dissertation. It covers the importance of the term structure on financial markets as the price structure of most financial products are related directly to this interest rate. The second chapter is the literature review of various studies that have been done during the past twenty five years. As we know that the continuous-time finance has emerged as a very important tool gearing towards new facets of research challenged. We look at various estimation methods that were introduced to capture dynamics of the term structure. Researchers have introduced many new statistical methods, yet there is no concensus on the best fitting model. One important issue is the insignificance of the mean reversion parameter of the interest rate. In the third chapter, we discuss the Monte-Carlo study of the interest rate process, its mean reversion, and its volatility elasticity parameters. We examine the small sample properties of estimators of the mean reversion and the elasticity of the volatility with respect to the level of the interest rates. We show that the maximum likelihood estimation methodology is a very crucial process to estimate parameters and evaluate statistical inferences. Ignoring the information on data frequency can lead to potential serious issues of financial mismanagement. We know that short-term and long-term traders have different views in their strategies. We show how to estimate these parameters correctly and show the comparison between the correct and the incorrect specifications of the maximum likelihood estimators. By taking into consideration the aggregation of data, we can correct the biases and can obtain the correct inferences. Results show that the absolute size of the mean reversion parameter does not vary as the aggregation of data changes. The volatility parameters, however, depend on the level of aggregation. Chapter four is the empirical study using the Euro Dollar deposit rate data. Empirical results concur with the Monte-Carlo study. Chapter five shows the conclusion of this dissertation.

Dedication

To my brother, Tarn

Acknowledgements

Many people helped me finish this dissertation. First of all, I would like to express my sincere gratitude to my advisor, Terence Agbeyegbe, for every thing he has done for me during my tenure as a student under his supervision. He has been a wonderful mentor. His wisdom to patiently guide me has made this work possible. In him, I found not only a source of knowledge and encouragement but also a great friend.

I would also like to thank my thesis advisory committee, of Salih Neftci and Tom Thurston for their support and constructive suggestions. Their inputs resulted in several improvement on both methodology and empirical results.

I would like to especially thank Teh Tai Yu, Horace Codjoe, and Alison Bacchus for constant support, advice, and invaluable friendship.

A very special thanks to my parents, brothers, and sisters for their ever-present love, heartfelt support and for having faith in me. They made it possible for me to enjoy these years in a home far away from home. And at last, I would like to thank my beloved wife, Ann. Without her unconditional support, I would not have made it through.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

Term structure of the interest rate plays an important role to most financial markets since it is fundamental to the pricing structure of most financial assets. As a key economic variable, the short rate is a core mechanism in terms of price setting of derivative securities as evidenced by today's varieties of financial products offered throughout the market. Bonds, Futures, and Options are examples of those products, derived from the movement of the term structure of the interest rate. Researchers, over the past twenty five years, have tried to determine the mechanism that affects the form and the dynamic revolution of this short rate. Advances in the theory of continuous-time finance have made many complicated issues in pricing become relatively simple to understand, leading to a widespread of models and ideas of the estimation methods of the interest rate processes. Many of these developments sometimes presuppose knowledge of the parameters underlying the interest rate process while, in fact, these parameters must be estimated properly.

In his influential works, Merton (1973, 1975) pioneers a series of research on the continuous-time modelling in financial economics, introducing Brownian Motion into the process of the stochastic differential equation. Vasicek (1977) modifies Merton's work by adding mean reverting and constant variance features into the model. This is done to avoid 'excessive' dispersion of rates by making the deterministic term take on the burden of preventing rates from spreading too much. Cox, Ingersoll, and Ross (1985) (CIR, henceforth) introduce a general equilibrium model of the term structure of default-free securities which have led to the concept of risk-neutral pricing and the so-called arbitrage-free models of the term structure. Other contributors also include

Richard (1977), Dothan (1978), and Brennan and Schwartz (1979) who each have specified their own versions of different parameter specifications aiming to accurately estimate the dynamic of the short rate. Results of these adjustments have added many important economic fundamentals such as mean reverting, stationarity, and volatility processes into the model.

One of the most well-known research on one-factor parametric estimation of the short term interest rate is that of Chan, Karolyi, Longstaff, and Sanders (1992a, 1992b)(CKLS, hereafter), who develop a common framework that nests all other important models into one general model. Broze, Scaillet, and Zakoian (1995), Tse (1995), Ball and Torous (1996, 1999), Bliss and Smith (1998), Nowman (1997, 1998), Koutmos(1998), Atkins and Krehbiel (1999), and Episcopos (2000) use parametrical estimation methods to estimate the interest rate parameters. In general, they employ the generalized method of moment, the maximum likelihood estimation, and the efficient method of moment in their studies.

Ait-Sahalia (1998), and Stanton (1999), on the other hand, have introduced non-parametrical pricing methods, believing that matching kernel density to instantaneous drift and diffusion functions are more consistent than parametric estimation method. This is because functional density distributions of some forms are required in parametric estimation. The fact that a priori idea of the functional form of the instantaneous volatility in financial series is not observable would invalidate most of the parametric estimation methods.

While others study one-factor process, some consider multi-factor processes that take account of other stochastic state variables as separated equation. Andersen and Lund (1997), Longstaff and Schwartz (1992), Brennen, Harjes, and Kronner (1996), and Dai and Singleton (2000) have estimated parameters using two-, or three-factor processes in their studies. Although multi-factor processes are relatively difficult to implement, these theoretical approaches have proved to be quite an attractive way to conduct research, which have led to several important empirical research.

A lot of research conducted thus far addressed important issues of finding the common ground model that best fits the short rate movement and estimating the sensitivity of interest rate volatility with respect to the interest rate levels. Although many theoretical works in this area have led to several empirical tests, there is still no concensus on any one particular model to represent the best fitting model. The difficulties of developing models that are attractive from a theoretical perspective and perform satisfactorily in practical sense have become increasingly challenged. As we know, the performance of any participants in financial markets relies directly on his/her own ability to predict the movement of the market. An abrupt change in stock prices due to an increase in volatility, for example, has created stringent demand for both researchers and practitioners to produce new quantitative methods of estimating the most accurate model. Because the short rate is an important input for any business cycle analysis through its impact on the cost of credit, and because of its sensitivity to the stance of monetary policy and the inflationary expectation, the term structure of the interest rate is considered an extremely vital state variable in both dynamic term structure and full-fledged macro-economic models. It simply affects the economic system as a whole. Essentially, traders and investors look for opportunities to make profits in both short and long terms. Their concerns on the interest rate are during the period when they have open positions in the market. While short-term traders would only want to hold their positions on an intraday basis and close out their positions at the end of any trading day, long-term investors would prefer to hold their positions for longer period of time, say weekly, monthly, or even yearly. Given this fact, knowledge of the consequences of using data with different frequencies would, therefore, be of great value in predicting the behavior of the short rate. Time series data such as those of the financial assets fluctuates tremendously. An abrupt upswing followed by sharp nose-diving patterns are often experienced in a very short run, where random noise(s), either positive or negative, can cause unexpected shock(s) to the market. Thus financial participants bear a decent amount of risks from these movements in the market. With this thought, aggregation of data is considered very essential information for every financier in the market.

Despite the importance of the parametric estimation of the term structure on the interest rate, little is known about how the interest rate behaves using maximum likelihood estimation that takes into account the aggregation of data. This is an important process when one transforms the continuous-time process to the discrete-time process. Ignorance of using correct specification of the likelihood function can really lead to many potential serious issues as one deals with these volatile markets. Poor decision can be extremely costly.

The purpose of this dissertation is to study the behavior of the interest rate process by examining the small sample properties of estimators of the mean reversion and the volatility elasticity across data with different data frequencies, using Monte-Carlo simulation method. This dissertation contributes to the literature by focusing on the effect data frequencies have on statistical inferences. In this study, I employ the Monte-Carlo simulation to evaluate the effect of aggregation of data when using a mis-specified method of the maximum likelihood function versus a correct specification of the maximum likelihood function. In addition, I show the empirical results using Euro-Dollar deposit rate.

This dissertation is organized as follows. In this chapter, I state the objective of this research. Chapter 2 shows the summarization of the literature review. Chapter 3 describes the numerical set-up and the Monte-Carlo simulation method. In the research, we study issues involving the aggregation of data and its effect on a small sample data. The focus is on two important parameters that drive the stochastic differential equation, i.e., the mean reversion parameter and the volatility elasticity of the interest rate. In chapter 4, we report the empirical evidence, using the Euro-dollar Deposit data. Chapter 5 concludes this dissertation.

Chapter 2

Literature Review

I. Introduction

All interest rate processes are known to have stochastic behaviors in which they change randomly over time. There is no doubt that management of interest rate risk, by which we mean the control of changes in value of a stream of future cash flows resulting from changes in the interest rates, or more specifically, the pricing and hedging of the interest rate products, is an important and complex issue. Many researchers observe that although these interest rate processes are highly stochastic, the manner in which they behave can be modelled. However, the construction of a reliable model for stochastic behavior of the term structure of interest rate is a task of considerable complexity.

CKLS (1992a) develop a unifying framework that nests many single-factor short rate models together.[1] These models are classified into two important categories which belong to the homoscedasticity and the heteroscedasticity classes of the volatility. For example, by setting $\gamma = 0$, we obtain a homoscedastic model where the volatility is constant. These homoscedastic models are those of the Merton (1973, 1975), and Vasicek models (1977). The heteroscedasticity models, on the other hand, are models

[1]Table 1 shows eight important single-factor models of term structure of the interest rate. By setting parameter according to the specification given in the table, one can get any one factor term structure models. For example, by setting γ to equal to 0 then one can get the Vasicek model. One can also derive the Cox, Ross, Ingersoll model by setting the γ to equal to 0.5. In addition, if one sets γ to equal to 1.0, the Brennan and Schwartz can be derived. One can also get the Merton model if setting the β and γ parameters to equal to 0. The geometric Brownian motion can be derived by setting α and γ to equal to 0.

where $\gamma > 0$, and that the diffusion processes fluctuate through their level of interest rates. They are those of the Cox. et al (1985), Brennan and Schwartz (1978), Dothan (1977), and amongst other models.

The homoscedastic model assumes the price of interest rate risk, i.e., the ratio of the expected excess return on a bond to the standard deviation of the excess returns on the bond - is a constant that does not depend on the level of the short rate. Although Vasicek model deems appropriate for the real interest rate, it is less appropriate to use for nominal interest rate since it allows the interest rate to become negative in some occasion. More importantly, the short rate volatility has been observed to lack homoscedasticity - it varies with the absolute level of the interest rate themselves.

This chapter is organized as follows. The rest of this chapter contains discussions on a variety of interest rate modelling and estimation done in the past. Models, such as the Generalized Method of Moment, Maximum Likelihood Estimation, Non-parametric Estimation, and Efficient Method of Moment are discussed in this section. Both single-factor and multi-factor diffusion processes are presented here.

The functional form of these models is the stochastic differential equation, which is represented as follows:

$$dr(t) = [\alpha + \beta r(t)]dt + \sigma r(t)^{\gamma} dW(t), \forall t \in [0, T] \tag{2.1}$$

where $r(t)$ is the logarithm of the interest rate level at time t, t is time, $W(t)$ is a standard Brownian motion. The first term on the right hand side, $[\alpha + \beta\ r(t)]dt$, characterizes a linear drift function, where β represents a mean reverting parameter. From an economic perspective, this process makes a lot of sense, since very high values of rates, historically, tend to be followed by a decrease in the interest rates more frequently than by an increase in the interest rate. The converse is also observed to be true for an unusually low rates. The second term, $\sigma r(t)^{\gamma} dW(t)$, represents a diffusion process. It is a Brownian motion and represents the volatility shock that builds into the equation.

When taking logarithm and its second moment, the γ represents constant elasticity of the interest rate level volatility. This parameter plays an important role in the interest processes, e.g., if σ=0, then $dr(t)$ becomes nonstochastic and is a deterministic function of time. In addition, if $\gamma = 0$, the effect of the volatility on $dr(t)$ then becomes independent of the level of interest rates. As γ increases, the effect of the interest rate level on volatility increases as well.[2]

Equation (2.1) can also be discretized, using Euler approximation method[3], as

[2]One of the most frequently used models, for example, is a geometric Brownian motion which can be interpreted, using Ito's Lemma, as follows.

$$dr(t) = \beta r(t)dt + \sigma r(t)dW(t)$$

$$\frac{dr(t)}{r(t)} = \beta dt + \sigma dW(t)$$

Integrate both sides yield

$$\int_0^t \frac{dr(s)}{r(s)} = \beta \int_0^t ds + \sigma \int_0^t dW(s)$$

$$\int_0^t \frac{dr(s)}{r(s)} = \beta t + \sigma W_t, (W_0 = 0)$$

We evaluate the integral on the left hand side using Ito formula for the function $g(t, x) = ln(x)$; $x > 0$, obtaining

$$d(ln(r_t)) = \frac{dr_t}{r_t} + \frac{1}{2}(-\frac{1}{r_t^2})(dr_t)^2$$

$$d(ln(r_t)) = \frac{dr_t}{r_t} - \frac{1}{2r_t^2}\sigma^2 r_t^2 dt$$

$$\frac{dr_t}{r_t} = d(ln(r_t)) + \frac{1}{2}\sigma^2 dt$$

Hence,

$$ln(\frac{r_t}{r_0}) = (\beta - \frac{1}{2}\sigma^2)t + \sigma W_t$$

$$r_t = r_0 e^{((\beta - \frac{1}{2}\sigma^2)t + \sigma W_t)}$$

[3]The formulation of Euler approximation can be done as follow. Let X_t denote an L-dimensional diffusion process satisfying the stochastic differential equation

$$dx_t = \mu(x_t, \phi)dt + \sigma(x_t, \phi)dB_t$$

where $\mu(x, \phi) : \Re \times \Phi \to \Re^{\mathcal{L}}$ and $\sigma(x, \phi) : \Re \times \Phi \to \Re^{\mathcal{L}} \otimes \Re^{\mathcal{D}}$ satisfy growth and Lipschitz conditions, B_t is a D-dimensional standard Brownian motion, and ϕ is a vector of parameters. Kloeden, Planten, Schurz (1991) show that under regular condition, a variety of discretized approximations always converge weakly to Ito processes as the discretized time step $\to 0$. The Euler approximation, in general, has the following form.

$$X_{t+1,h} = X_{t,h} + h\mu(X_{t,h}, \phi) + \sqrt{h}\sigma(X_{t,h}, \phi)\epsilon_{t+1}$$

where ϵ_t is iid, $N(0, I_D)$, I_D is the $D-$dimensional identity matrix, and h is the discretized interval length.

follows.

$$r_{t+1} - r_t = \alpha + \beta r_t + \sigma r_t^{\gamma} \epsilon_{t+1}, t \in \aleph \tag{2.2}$$

$$r_{t+1} = \alpha + (\beta + 1) r_t + \sigma r_t^{\gamma} \epsilon_{t+1}, t \in \aleph \tag{2.3}$$

First, taking the conditional expectation of the equation above yields[4]

$$E[r_{t+1} \mid r_t] = E[\alpha + (\beta + 1) r_t + \sigma r_t^{\gamma} \epsilon_{t+1} \mid r_t] \tag{2.4}$$

which becomes

$$E[r_{t+1} \mid r_t] = \alpha + (\beta + 1) r_t \tag{2.5}$$

since

$$E[\sigma r_t^{\gamma} \epsilon_{t+1} \mid r_t] = \sigma r_t^{\gamma} E[\epsilon_{t+1} \mid r_t] = 0. \tag{2.6}$$

Taking the second moment yields

$$V[r_{t+1} \mid r_t] = \sigma_{t+1}^2 = \sigma^2 r_t^{2\gamma}. \tag{2.7}$$

Taking a natural logarithm of this variance yields

$$log(V) = 2log(\sigma_{t+1}) + 2\gamma log(r_t). \tag{2.8}$$

Taking the partial derivative yields

$$\frac{\partial log(V)}{\partial log(r_t)} = 2\gamma. \tag{2.9}$$

Equation (2.9) shows the higher the volatility elasticity, γ, the volatility in the market will be twice as high.

The following sections will review some of the past research of those statistical methods.

[4]Some of the basic properties of the Brownian motion can be found in general stochastic processes textbooks of Brzezniak and Zastawniak (1999), Karlin and Taylor (1975), Koa (1997), Oksendal (1998), Ross (1996), and a more advance level of Karatzas and Shreve (1998).

II. Generalized Method of Moment

Using the Generalized Method of Moment (GMM)[5], Chan et. al. (1992a) calibrate parameters using the US one-month Treasury bill rate from 1964 to 1989 and compare among many nested short-rate models. GMM, in general, is used to compare certain sample moments with the theoretical counterparts. Parameter values are sampled and estimated as the values of the sample moments converge to their theoretical values.

In general, the GMM estimators are based on the empirical counterpart of some functional form(s) that satisfies orthogonality conditions, i.e.,

$$E[Z_i K(y_t, z_t) - k((z_t); \theta_0)] = 0 \tag{2.10}$$

where E[.] is the expectation of the true distribution of (y, z), Z_i is a matrix function of z_i and θ_0 is the true value of the parameter. The moment conditions for the typical GMM estimation are normally chosen arbitrarily.

If Ω is a (K,K) symmetric positive semi-definite matrix, the estimator is then defined by:

$$\hat{\theta}(\Omega) = \arg\min_{\theta} \left[\sum_{t=1}^{T} Z_t[K(y_t, x_t) - k(x_t; \theta)] \right]' \Omega \left[\sum_{t=1}^{T} Z_t[K(y_t, x_t) - k(x_t; \theta)] \right] \tag{2.11}$$

where Z_i is a matrix function of z_i with size of (K, q). Under regular conditions (see Hansen 1982),

(i) $\hat{\theta}_n(\Omega)$ is a consistent estimator of the true value θ_0.

(ii) The GMM estimator is asymptotically normal, i.e., as follows:

$\sqrt{n}(\hat{\theta}_n(\Omega) - \theta_0) \rightarrow N(0, \Sigma^{-}1_1\Sigma_2\Sigma^{-}1_1)$,

where: $\Sigma_1 = D'\Omega D$, $\Sigma_2 = D'\Omega V_0(Z[K(y, z) - k(z, \theta_0)])\Omega D$, $D = E\left[Z\frac{\delta k}{\delta\theta'}(z; \theta_0)\right]$

In the study, CKLS specifies θ as parameter vector, consisting of α, β, σ, γ, where α, β are the drift parameters and σ, γ are the diffusion parameters. A vector of a function $f_t(\theta)$ is given by

[5] Hansen (1982) developed the GMM method and, later on, Hansen and Scheinkman (1994) modified such method and made them applicable to the unobserved state variables Markov processes of the stochastic volatility.

$$f_t(\theta) = \begin{bmatrix} \epsilon_{t+1} \\ \epsilon_{t+1} r_t \\ \epsilon_{t+1}^2 - \sigma^2 r_t^{2\gamma} \\ (\epsilon_{t+1}^2 - \sigma^2 r_t^{2\gamma}) r_t \end{bmatrix} \tag{2.12}$$

where $\epsilon_{t+1} = r_{t+1} - (\alpha + \beta r_t)$, for $t = [0,\dots, n-1]$.

The sample moments for the vector in equation (2.10) can be shown from equations (2.13) to (2.16), as follow.

$$f_1(\hat{\theta}) = \frac{1}{n}\sum_{t=1}^{n} [r_{t+1} - \alpha - \beta r_t] \tag{2.13}$$

$$f_2(\hat{\theta}) = \frac{1}{n}\sum_{t=1}^{n} [(r_{t+1} - \alpha - \beta r_t) r_t] \tag{2.14}$$

$$f_3(\hat{\theta}) = \frac{1}{n}\sum_{t=1}^{n} \left[(r_{t+1} - \alpha - \beta r_t)^2 - \sigma^2 r_t^{2\gamma} \Delta t\right] \tag{2.15}$$

$$f_4(\hat{\theta}) = \frac{1}{n}\sum_{t=1}^{n} \left[[(r_{t+1} - \alpha - \beta r_t)^2 - \sigma^2 r_t^{2\gamma} \Delta t] r_t\right] \tag{2.16}$$

Choosing $\hat{\alpha}, \hat{\beta}, \hat{\sigma}, \hat{\gamma}$ and set $f_1(\hat{\theta}), f_2(\hat{\theta}), f_3(\hat{\theta}), f_4(\hat{\theta})$ to zero by minimizing

$$J(\theta) = f_1^2(\hat{\theta}) + f_2^2(\hat{\theta}) + f_3^2(\hat{\theta}) + f_4^2(\hat{\theta}) \tag{2.17}$$

The GMM estimator, $\hat{\theta}$, is the value that minimizes the quadratic form of matrices

$$\arg\min_{\theta} J(\theta) = g'(\theta) X(\theta) g(\theta) \tag{2.18}$$

where $g(\theta) = \frac{1}{n}\sum_{t=1}^{n} f_t(\theta)$ and $X(\theta)$ is the weighting matrix given by

$$X(\theta) = S^{-1}(\theta) \tag{2.19}$$

and where $S(\theta) = E[f_t(\theta) f_t'(\theta)]$.

$$D(\theta) = \frac{\partial g(\theta)}{\partial \theta_0'} \tag{2.20}$$

$$\frac{1}{n}[D'(\hat{\theta}_0) S^{-1}(\hat{\theta}_0) D(\hat{\theta}_0)]^{-1} \tag{2.21}$$

CKLS compare eight interest rate models to gauge the performances of the interest rate process. They conclude that the interest rate process depends strongly on its volatility elasticity. Term structure models with volatilities more highly sensitive to the level of interest rates than generally used models have a closer empirical fit to the data.[6] As shown in their paper, CKLS find the parameter of the volatility elasticity, γ, to be quite high. They also find that the mean reversion parameter, β, is negligible and often statistically insignificant.

Tse (1995) also uses GMM to test the performance of the CKLS nested short rate models on different countries' interest rates. In his study, he looks at interest rate structures of different major industrialized countries by evaluating if CKLS' claims of high volatility elasticity can be universally used across many countries. Those interest rates are three-month money market rates of Australia, Belgium, Canada, France, Germany, Holland, Italy, Japan, Switzerland, UK, and the United States. Empirical results show a varieties of differences in the volatility elasticity with respect to the interest rate level, γ, across most nations. They range from -0.3600 to 1.7283 on unrestricted models across all countries.

Interestingly, when compared those of the US interest rate, CKLS find the volatility elasticity parameter, γ, to be 1.4935 while Tse finds the volatility elasticity parameter to be around 1.7283. One reason is due to the differences in the data set used in their studies. Another possible interpretation could be that the interest rate volatility might have become more sensitive in the later years which causes the volatility elasticity to increase. His final conclusion is that no single model can satisfactorily describe the stochastic structure of interest rates for all countries. However, with countries that have low volatility elasticity with respect to the interest rate, the Vasicek model may be appropriate to use. For countries that have moderate volatility, the CIR-SR model or BS model can be used.

Adkins and Krehbiel (1999) use three-month and six-month London Inter-Bank Offer Rate (LIBOR) data to estimate the dynamics of the short-term interest rate. Again, their results show little evidence of the mean reversion in either of the three-month or six-month LIBOR rates. For each model in which the mean reversion parameter, β, is not constrained to zero, its estimated value is negative and not statistically signif-

[6]CKLS's study shows the elasticity of the volatility of the unrestricted model is reported at 1.4935.

icant (except the Brennan and Schwartz models). However, the estimated values of the volatility elasticity, γ, in the unrestricted and the CEV models range below the unit root process and are significantly different from zero. The volatility elasticity is estimated to equal to γ = 0.7500 and 0.7800 on the unrestricted model. Adkins et. al., conclude that neither the three-month nor six-month LIBOR are mean reverting and the volatility increases disproportionally to the level of the interest rate since γ < 1. They also claim that the discrete-time approximations can account for no more than 20 percent of the volatility. The validity of a CIR-square root process for the three- and six-month LIBORs is not ratified by the data. Using LIBOR data, a t-test of the square root null hypothesis produces a t-ratio of 2.25 and 1.76 for 3-month and 6-month data.[7] Adkins and Krehbiel claim that the evidence does not support the use of the pricing estimation of CIR-square root medel (1985).

Bliss and Smith (1998)(BS henceforth) show result that is contrary to those of CKLS. While CKLS find no evidence of the regime switching in their studies, Bliss et. al. argue that there is a strong evidence of regime shift during the Federal Reserve Experiment period of October 1979 through September 1982, which might have caused the mis-specification in the conclusion of CKLS. Bliss and Smith use the one-month Treasury bill rate from June 1964 through December 1989 for ensuring conformity and comparability with CKLS. BS also use the GMM method to estimate the parameter and test the regime switching in the form of changes in the coefficients by setting up dummy parameters in the following equation:

$$r_t - r_{t-1} = (\alpha + \delta_1 D_t) + (\beta + \delta_2 D_t) r_{t-1} + \epsilon_t \tag{2.22}$$

$$\epsilon_t^2 = (\sigma^2 + \delta_3 D_t) r_{t-1}^{2(\gamma + \delta_4 D_t)} + \eta_t \tag{2.23}$$

where D_t is an indicator variable. The δ_js measure the shift in parameters during the alternative regime period. D_t is set to one so that the parameter is $\alpha + \delta_1$ from October 1979 through September 1982, which is the period the Federal Reserve conducted the monetary experiment, and to zero so that the parameter is α both before and after that period (the period prior to October 1979 and the period after September 1982).

[7]The hypothesis testing is set as follows.
$H_0 : \gamma = 0.5$
$H_1 : \gamma \neq 0.5$

BS apply the test for a structural shift to the general CKLS model on the ground that the other eight models are nested in this general model. If the test shows any existence of regime switching in the general model, then the other restricted versions of that models should be affected as well. While CKLS claim that there is no shift in the structure, BS, using the structural break dummy during the federal reserve's experimental period, find that the Merton, the Vasicek, and the CIR-SR are rejected on the no shift null hypothesis at the 5 percent level. They also examine if any of the restricted variants, of the CKLS model, are able to fit the data. They find that the CIR-SR and Brennan-Schwartz models are not rejected in their unrestricted forms, when the parameter values are permitted to change. In addition, examining the unrestricted CKLS model shows that the γ in the unrestricted model is equal to 0.95 outside the temporary structural shift period and γ is equal to 0.33 within the shift period. BS find strong evidence of a structural break in the interest rate data. They find that a moderate value of the volatility elasticity, γ, can capture the dependence of volatility on the level of interest rate better than the high value γ. Nevertheless, they show that the result is quite robust to changes in the short rate used and the treatment of the outliers. As a result, BS conclude that the CIR-SR (square root) model seems to fit the data the best (amongst the class of one-factor diffusion model).

Longstaff and Schwartz (1992) develop a two-factor general equilibrium model using the CIR-SR (1985) framework to price the discount bonds and other interest rate sensitive contingent claims. The factors are the short-term interest rate and the instantaneous variance of changes in the short-term interest rate. They claim that this would allow the contigent claims to reflect both the current level of interest rate and the current level of the interest rate volatility. The general equilibrium framework idea is that in an assumed economy, all physical investment produces a good that is either consumed or reinvested in production. The realized returns on physical investment are in the form

$$\frac{dQ(t)}{Q} = (\mu X + \theta Y_t)dt + \sigma\sqrt{Y}dz_{1,t} \tag{2.24}$$

where μ, θ, and σ are positive constants X, and Y are state variables, and $Z_{1,t}$ is a scalar Wiener process.

The short-term interest rate and the instantaneous variance of changes in this short-rate are the two-factor that correlated in determining the interest rate. There are two state variables, x_t and y_t and their dynamic processes are

$$dx(t) = (a - bx_t)dt + c\sqrt{x_t}dz_{2,t} \tag{2.25}$$

$$dy(t) = (d - ey_t)dt + f\sqrt{y_t}dz_{3,t} \tag{2.26}$$

where $a, b, c, d, e, f > 0$ and z_1 and z_2 are the Wiener processes.

They assume that there is a fixed number of identical investors with time-additive preferences of the form

$$E_t\left[\int_t^{\infty} exp(-\rho s)\log(C_s)ds\right] \tag{2.27}$$

where ρ is the utility discount factor, C_t represents consumption at time $= s$, and $E[.]$ is the conditional expectation operator. Each investor's decision is to maximize the consumption function with respect to the budget constraint (wealth) of

$$dW = W\frac{dQ}{Q} - Cdt \tag{2.28}$$

The equilibrium dynamic equation of wealth is then become

$$dW = (\mu X + \theta Y - \rho)Wdt + \sigma W\sqrt(Y)dZ_1 \tag{2.29}$$

The instantaneous short rate, r_t, and the instantaneous variance, ν_t, are linearly related to x_t and y_t in the following form

$$r_t = \alpha x_t + \beta y_t \tag{2.30}$$

$$V_t = \alpha^2 x_t + \beta^2 y_t \tag{2.31}$$

The dynamic equations of the instantaneous short rate and the instantaneous variance is obtained by the following equation.

$$dr = \left(\alpha\gamma + \beta\eta - \frac{\beta\delta - \alpha\xi}{\beta - \alpha}r - \frac{\xi - \delta}{\beta - \alpha}V\right)dt + \alpha\sqrt{\frac{\beta r - V}{\alpha(\beta - \alpha)}}dZ_2 + \beta\sqrt{\frac{V - \alpha r}{\beta(\beta - \alpha)}}dZ_3 \tag{2.32}$$

$$dV = \left(\alpha^2\gamma + \beta^2\eta - \frac{\alpha\beta(\delta - \xi)}{\beta - \alpha}r - \frac{\beta\xi - \alpha\delta}{\beta - \alpha}V\right)dt + \alpha^2\sqrt{\frac{\beta r - V}{\alpha(\beta - \alpha)}}dZ_2 + \beta^2\sqrt{\frac{V - \alpha r}{\beta(\beta - \alpha)}}dZ_3 \tag{2.33}$$

III. The Maximum Likelihood Estimation

For interest rate time series, r_t, $t = 1, ..., T$, with transition densities

$$p(r_t; r_{t-1}|\theta)$$

The joint density is

$$p(r_t, ..., r_{t-T}|\theta) = p_0(r_0|\theta) \prod_{i=1}^{T-1} p(r_t; r_{t-1}|\theta)$$

where p_0 is some prior density for r_0. The likelihood function becomes

$$\mathcal{L} = \prod_{i=1}^{T-1} p(r_t; r_{t-1}|\theta)$$

The maximum likelihood estimator of a parameter θ is defined as:

$$\hat{\theta}_T = \arg\max_{\theta} \sum_{t=1}^{T} \log p(r_t, r_{t-1}|\theta) \tag{2.34}$$

CKLS (1992b) also use an AutoRegression Conditional Heteroscedasticity (ARCH) framework, comparing a set of single-factor term structure models on the Japanese interest data. In general, Japanese rates are observed to be less volatile when compared to the US interest rates. CKLS use the maximum likelihood estimation to calibrate parameters and find significant evidence of time-varying volatility in the change of the Gensaki rate. Despite Japanese interest rate being less volatile, the dynamics of the interest rate volatility between the two countries (US *vs* Japan) are quite similar. As a result, the same conclusion has been drawn. They conclude that the best model to capture the dynamics of the interest rate is one that allows the conditional volatility to be highly sensitive to the level of the interest rate, i. e., the models with high γ. Results show that using Gensaki rate, the CEV model, and the CIR-VR model perform best. CKLS claim that the relation between interest-rate volatility and the level of r is a very important feature for term structure modeling. The mean reversion parameter, β, however, is much less important. As γ becomes increasingly higher, mean reversion parameter β become less statistically insignificant. CKLS also use the same methodology applying to a one-month US Treasury bill rate. Again the same result has been concluded.

Nowman (1997, 1998) presents an alternative approach to the nested model of CKLS (1992), employing Bergstrom's (1983,1985,1986,1990) studies. This approach is based on the Gaussian estimation methods for estimating parameters of open continuous-time from the discrete data, using an exact discrete model that considers exact restrictions on the distribution of the discrete data implied by the continuous-time model. Nowman's assumption is that the volatility of the interest rate only changes at the beginning of the unit observation period and then remains constant afterwards. The variance of the stochastic variables remains constant after the change at the beginning of the unit observation period between discrete observations. As a result, exact discrete model of Bergstrom (1984, Theorem 2) can be used to obtain the Gaussian estimates modified for heteroskedasticity. Mathematically, this can be written as follows.

$$dr(t) = [\alpha + \beta r(t)]dt + \sigma r(t)^{\gamma} dW(t) \tag{2.35}$$

Bergstrom's proposition generates the following equation.

$$dr(t) = [\alpha + \beta r(t)]dt + \sigma r(t'-1)^{\gamma} dW(t) \tag{2.36}$$

$$r(t) - r(t'-1) = \int_{t'-1}^{t} [\alpha + \beta r(s)]ds + \sigma[r(t'-1)]^{\gamma} \int_{t'-1}^{t} dW(s), \forall t \in [t'-1, t] \tag{2.37}$$

where $t'-1 < t \leq t'$ and $\int_{t'-1}^{t} dW(s) = W[t'-1, t]$.

This equation can also be discretized into

$$r_t = e^{\beta} r_{t-1} + \frac{\alpha}{\beta}(e^{\beta} - 1) + \epsilon_t, \forall t = [1, 2, ..., T] \tag{2.38}$$

where ϵ_t, $[t = 1, 2, ..., T]$ satisfies the conditions

$$E[\epsilon_s \epsilon_t] = 0, s \neq t \tag{2.39}$$

and

$$E[\epsilon_t^2] = \int_{t'-1}^{t} e^{2(t-\tau)\beta} \sigma^2 [r(t-1)]^{2\gamma} d\tau \equiv m_{tt}^2 = \frac{\sigma^2}{2\beta}[e^{2\beta} - 1][r(t-1)]^{2\gamma} \tag{2.40}$$

Bergstrom (1985, 1986) defines the following maximum likelihood.

$$\arg\max_{\theta} \mathcal{L}(\theta) = \sum_{t=1}^{T} \left[2\ln m_{tt} + \frac{[r(t) - e^{\beta} r(t-1) - \frac{\alpha}{\beta}(e^{\beta} - 1)]^2}{m_{tt}^2} \right] \tag{2.41}$$

where $m_{tt}^2 = \frac{\sigma^2}{2\beta}(e^{2\beta} - 1)(r_{t-1})^{2\gamma}$ and $\theta = (\alpha, \beta, \sigma^2, \gamma)$

$$\mathcal{L}(\theta) = \sum_{t=1}^{T} [2\ln m_{tt} + \eta_t^2] \tag{2.42}$$

where $\eta = [\eta_1, \eta_2, ..., \eta_T]$ is a vector whose elements can be computed from $m_{tt}\eta_t = \epsilon_t$.

Using a one-month sterling interbank rate and the US Treasury bills, Nowman (1997) finds that the Gaussian estimates show that the bias from using CKLS approximation is very small when compared to the exact Gaussian discrete estimation. Using the British data, he shows that the CIR-SR model performs the best followed by the CEV and Vasicek models. Nowman finds that the volatility elasticity in the unrestricted model, γ, is equal to 0.2898 and is insignificant. This result of the volatility elasticity is contrast to the result of CKLS. Using the Treasury bill data, Nowman also finds that the unrestricted model's estimate of γ is equal to 1.3610 and is highly significant while CKLS reports that γ in his unrestricted model is equal to 1.4935 and significant. Again, he claims that the difference might be due to the differences in estimation methodologies. However, the two estimations are quite similar, i.e., the interest rate structure is very sensitive to the level of the interest rate. The results show a very small bias using both methods (CKLS discrete approximation or Bergstrom's gaussian approximation). However, the results regarding the conditional volatility with respect to the level of the interest rates show strong contrast to those of CKLS.

Nowman (1998) also employs the Gaussian exact discrete estimation to fit the parameter using the Euro Currency rate data. The one-month Euro-currency rate on US, Japanese, French, Italian currencies, starting in February 1981 through March 1995. Using the Euro currency rate data, Nowman (1998) finds different results regarding the fit of the estimation on each currencies. For example, using Japanese Yen Rate and the US dollar rate, the unrestricted model performs the best followed by the Brennan-Schwartz model, CEV, GBM, and Dothan models. Using the French Franc rate, the unrestricted model performs the best followed by the CEV, CIR-VR, and Brennan-Schwartz models. For Italian Lira rate, the best performing model is unrestricted model, followed by CEV, CIR-VR and the Brennan-Schwartz models. Nowman's estimation indicates that conditional volatility on the level of the interest rate are marginally related, i.e., it has become less dependent on the rate level.

Episcopos (2000) employs the same methodology as Nowman (1997, 1998), testing the stochastic behavior of the one-month interbank rate of ten leading countries (Australia, Belgium, Germany, Japan, Netherlands, New Zealand, Singapore, Switzerland, the United Kingdom, and the United States). In this article, Episcopos also finds that the elasticity of the volatility parameter, γ, is very important and statistically significant. However, he finds that the γ differs widely from country to country. Seven out of ten countries show the γ of less than unity. Again, this estimation is drastically contrast to what TSE (1996) found but more in line with what Nowman (1997, 1998) found. However, the suitability of the restricted models are ranked differently from those of others. In other words, Episcopos finds that the CEV, in terms of fitting the data, perform relatively better than any restricted model and followed by the Generalized Brownian Motion and the CIR-SR (1985).

In their interesting article on stress testing, Bali and Neftci (BN)(2000) use extreme value theory to estimate the level of volatility by studying the dynamic relationships of the volatility associated with the extreme up and down movements in various short term interest rates. BN assume that a risk factor is distributed according to the logistic density due to the heavy concentration towards the tails. The extremes are defined as those more than two standard deviations away from the sample mean of daily interest rate changes. They compare two volatility measurement methods: the GARCH approach and the Extreme Value approach. First, the GARCH approach has the following form.

$$r_{t+1} - r_t = \alpha_0 + \alpha_1 r_t + \epsilon_{t+1} \tag{2.43}$$

$$\epsilon_{t+1}|\Omega_t \; f(\mu_{t+1}, \sigma_{t+1}, \nu; \epsilon_{t+1}) \tag{2.44}$$

$$E[\epsilon_{t+1}^2|\Omega_t] \equiv \sigma_{t+1}^2 \equiv \beta_0 + \beta_1\epsilon_t^2 + \beta_2\sigma_t^2 \tag{2.45}$$

where $E[\epsilon_{t+1}^2|\Omega_t]$ represents the GARCH(1,1) model. In words, the current volatility is a function of the last period's unexpected news and the last period's volatility. Second, the extreme value approach has the following form.

$$x_i = \frac{(X_{max(i)} - \mu_i)}{\sigma_i} \tag{2.46}$$

The extreme value approach transformed the variate, X_{max}, using a location parameter, μ_i, and a scale parameter, σ_i, so that it has the form as equation (4.2) as above. The generalized Pareto distribution of the following form is used:

$$H(x) = 1 - (1 + \phi X)^{\frac{-1}{\phi}}, \phi \neq 0,$$

$$H(x) = 1 - \exp(-X), \phi \equiv 0 \tag{2.47}$$

The ϕ determines the tail behavior of the distribution. For ϕ=0, the tail decreases exponentially. For ϕ<0, the distribution is short tailed. For ϕ>0, it has a polynomially decreasing tail. The maximum likelihood of the extreme value (generalized Pareto distribution) has the following form.

$$\arg\max_{\mu,\sigma,\phi} \log \mathcal{L}((\mu,\sigma,\phi); X_{max}) = -n\log\sigma - n\left[\frac{1+\phi}{\phi}\right]\sum_{i=1}^{n}\log\left[(1+\phi)\frac{X_{max}-\mu}{\sigma}\right] \tag{2.48}$$

Bali and Neftci show that the volatility is overestimated if the normal distribution is used as compared to those of the student t-Garch model. The generalized Pareto Distribution also yields better approximations on both the local maxima and minima. When the extremal theory is used, the estimates of the historical volatilities are significantly lower than other standard models.

Koutmos (1998) also shows that the the volatility elasticity of the interest rate process across different maturities and different frequencies is an important source of time variation in volatility. He claims that investors behave differently according to the holding period horizon one has had in his/her investment plan. For example, investors who would like to hold their positions on a monthly basis would be interested in monthly elasticity estimates, and, likewise, investors with daily holding horizons would look for the daily elasticity estimates. His findings show the mean reversion parameters, β, are not statistically significant. He also reports results showing that the magnitude of the estimate of the mean reversion becomes more noticeable as we move from daily to monthly data. He also claims that the volatility and the volatility elasticity are strongly significant in all cases.

Two-factor arbitrage models have been suggested by Brennan and Schwartz (1979) who use both the short-term and long-term interest rates, by Schaefer and Schwartz (1984) who use the long term interest rate and the spread between the long-term and short-term interest rates. Ball and Torous (1999) incorporate the stochastic volatility into the CKLS (1992) dynamic SDE framework. They assert that short-term interest rate dynamics are different from those of the stock return dynamics, and it depends on the source of the economic shocks. In contrast to estimated stock return dynamics, stochastic volatility of the interest rate is distinguished by a faster mean-reverting behavior. They are short-lived, transient, and less persistent which implies a faster movement towards its mean after a shock is introduced into the market. BT also show that while the innovations in interest rates are almost uncorrelated with innovations in the interest rate volatility, there is a very strong negative correlation between the stocks returns and the volatility (as evidenced by current stock market situation). Ball and Torous add the stochastic interest rate volatility to the CKLS general equation. The unobservable interest rate volatility, σ_t, evolves stochastically and has an AR(1) process with the speed of adjustment to its unconditional mean, μ governed by the parameter b and its volatility characterized by the parameter ξ.

$$r_t = \alpha + \beta r_{t-1} + \sigma r_{t-1}^{\gamma} \epsilon_{1,t} \tag{2.49}$$

$$r_t - [\alpha + \beta r_{t-1}] = \sigma r_{t-1}^{\gamma} \epsilon_{1,t} \tag{2.50}$$

square on both sides and take natural logarithm yield

$$\log[r_t - (\alpha + \beta r_{t-1})]^2 = \log[\sigma r_{t-1}^{\gamma} \epsilon_{1,t}]^2 \tag{2.51}$$

Letting $x_t = \log(\sigma_t^2)$, and $res_t = r_t - (\alpha + \beta r_{t-1})$ yield

$$\log(res_t^2) = \log[\sigma r_{t-1}^{\gamma} \epsilon_{1,t}]^2$$

$$\log(res_t^2) = \log(\sigma_{t-1}^2) + 2\gamma \log(r_{t-1}) + \log(\epsilon_{1,t}^2) \tag{2.52}$$

$$\log(res_t^2) = x_{t-1} + 2\gamma \log(r_{t-1}) + \log(\epsilon_{1,t}^2) \tag{2.53}$$

$$x_t - \mu = b(x_{t-1} - \mu) + \xi \epsilon_{2,t} \tag{2.54}$$

BT use the quasi-maximum likelihood of the iterative filtering procedure. Given a prior on the state $p(x_{t-1}|Y_{t-1})$, there are three steps: a projection to obtain $p(x_t|Y_{t-1})$, an integration to calculate the conditional likelihood, and an update to obtain $p(x_t|Y_t)$. The quasi-maximum likelihood then is implemented by maximizing the function with respect to the parameters, θ.

BT use a one-month Euro-currency rates and one-month UK interbank rates as proxies to the short term interest rate. The Euro-dollar, Euro-mark, Euro-sterling, and Euro-yen rates are used in this empirical study, dated from January 1986 to December 1995, with the exception of the UK interbank rates was from January 1975 to September 1995. Results show that there is clear evidence across all interest rate series of mean-reverting stochastic volatility. They also claim that although the interest rate volatility seems to be sensitive to the level of the interest rates, the relation is difficult to measure accurately. The dynamics of the interest rate across nations are broadly similar (results are in accordance with Nowman (1998)).

Another study of Ball and Tourous (1996) on the CIR's square root process shows an interesting result on the unit root problem that occurs in estimating the short term interest rate dynamics. BT show that the degree of mean reversion depends on a speed of an adjustment coefficient and the slower the speed of adjustment, the closer the interest rate's stochastic behavior is to a non-stationary process with an exact unit root. This would cause a significant upward bias in estimation of the speed of adjustment coefficient. BT (1996) shows that no matter which parametric estimation method one uses to estimate the parameters, namely, least square, general method of moment, or maximum likelihood, the estimation will be biased upward even the large sample sizes are employed.

Ball and Tourous perform a simulation study to test the statistical properties on most time series estimation procedures. They Generate five year and twenty year of monthly data and set the initial spot rates to range from 3 percent to 7 percent, with an increment of 1 percent. The result shows an upward bias of a sampling distribution of the adjustment coefficients. He also shows that the magnitude of the unit root problem depends on the shape of the term structure and on whether the long term yields are included.

$$dr(t) = (\alpha + \beta r(t))dt + \sigma r(t)^{\gamma} dW(t) \tag{2.55}$$

CIR model sets the γ to equal to 0.5 so we can rewrite the above formula in other form as follow. Dividing through the above equation by β yields

$$dr(t) = \kappa[\theta - r(t)]dt + \sigma\sqrt{r}(t)dW(t) \tag{2.56}$$

where κ = the speed of adjustment process of the mean reversion, and θ = the mean reversion parameter. CIR shows that

$$E[r(s) \mid r(t)] = r(t)e^{-\kappa(s-t)} + \theta(1 - e^{-\kappa(s-t)}) \tag{2.57}$$

Let time $(s-t) = \Delta$, the $e^{-\kappa(s-t)}$ is approximately equal to (1-$\kappa\Delta$), the above equation can then be rewritten as

$$E[r_{t+\Delta} \mid r_t] = r_t(1 - \kappa\Delta) + \theta(1 - (1 - \kappa\Delta)) \tag{2.58}$$

$$E[r_{t+\Delta} \mid r_t] = r_t(1 - \kappa\Delta) + \theta(\kappa\Delta) + O(\Delta^2) \tag{2.59}$$

He claims that the sampling distribution of κ is biased upward throughout, and the bias occurs most pronounced when the initial spot rate is close to θ. He concludes that it is difficult to precisely estimate the interest rate where it is very important to accurately estimate the mean reversion in the drift process of the underlying interest rate process.

Another important model that introduces both the interest rate levels and information shocks into the dynamics of interest rate modeling is that of Brenner, Harjes, and Kroner (1996) (BHK). BHK claim that both the CKLS generalized one-factor framework and the GARCH(1,1) type model lack features that reflects the movement of the interest rate in reality. For example, one-factor model set the volatility to depend only on its level of the interest rate, but not the news arrival process. The GARCH model does not permit volatility to be a function of the interest rate levels. Additionally, most empirical evidence of the interest rate process, using GARCH (1,1) model, follows the unit root process and possesses the non-stationarity. This would mean that the volatility shock is a random walk and would stay persistently into the future. BHK propose a new model that accounts for newly information arrival and permits that information to vary through time. In particular, BHK propose the followings.

$$E[r_t \mid r_{t-1}] = E[\alpha + (\beta + 1)r_{t-1} + \sigma r_{t-1}^{\gamma}\epsilon_t \mid r_{t-1}] \tag{2.60}$$

$$E[\epsilon_t \mid \mathcal{F}_{t-1}] = 0 \tag{2.61}$$

$$E[\epsilon_t^2 \mid \mathcal{F}_{t-1}] = 0, \sigma_t^2 = \psi_t^2 r_{t-1}^{2\gamma} \tag{2.62}$$

$$\psi_t^2 = a_0 + a_1\epsilon_{t-1}^2 + b\psi_{t-1}^2 \tag{2.63}$$

From the above equation, we can see that the sensitivity of volatility to levels is a function of the information flow. If setting $a_1 = b = 0$, then the time variation ψ_t disappears and we are back to the LEVELS model. Also, if γ=0, then the level effect disappears and we are back in the GARCH framework. In addition, to avoid the equal impact on volatility (negatively or positively), BHK also proposed the asymmetric TVP.

$$\psi_t^2 = a_0 + a_1\epsilon_{t-1}^2 + a_2\eta_{t-1}^2 + b\psi_{t-1}^2 \tag{2.64}$$

where $\eta_{t-1} = min(\epsilon_{t-1}, 0)$. If a_2 >0, then bad news (negative shocks) has a larger impact on volatility than good news (positive shock). An alternative model is to add a levels term directly to the GARCH(1,1) model as

$$\sigma_t^2 = a_0 + a_1\epsilon_{t-1}^2 + b\sigma_{t-1}^2 + a_3 r_{t-1}^{2\gamma} \tag{2.65}$$

BHK use two sets of data in his study. The first consists of 909 weekly observations on 13-week Treasury bill yields from February 9, 1973 to July 6, 1990. The second data set also consists of 407 monthly observations of the total return of 30-day Treasury bills, from January 1960 to December 1993. They conclude that while the relationship between the volatility elasticity and the level of interest rate is important, the modelling of the volatility parameter as a function of the unexpected news is also very important.

IV. Non-parametric Estimation

Ait-Sahalia (1996) and Stanton (1997) use nonparametric estimators applied to the short-term interest rate data and conclude that the drift function contains important nonlinearities. Ait-Sahalia (1996) argues that a parametrical estimation approach sometimes lead to mis-specification in terms of matching the new data to the sample. Instead he uses a non-parametric method to estimate the term structure of the interest rate process. Claiming that a priori information on exact functional forms of the drift function and diffusion function are unobserved, pricing derivatives using a parametrical method would, therefore, yield a mis-specified function. Matching density seems to be a better alternative since the instantaneous drift and diffusion functions of the short rate process are derived to be consistent with the observed distribution of the discrete data. Since the densities of the interest rate process can be obtained and estimated from the data on the short-term rate, Ait-Sahalia uses Kernel function to estimate a stationary distribution process.[8] Ait-Sahalia estimates the stationary density and develops seperately an explicit connection between the drift, the diffusion functions and the stationary density using the Kolmogorov forward equation (Fokker-Planck partial differential equation).

$$\frac{\partial f(\Delta, r_{t+\Delta}|r_t)}{\partial \Delta} + \frac{\partial(\mu(r_{t+\Delta}, \theta) f(\Delta, r_{t+\Delta}|r_t))}{\partial r_{t+\Delta}} - \frac{1}{2}\frac{\partial^2(\sigma^2(r_{t+\Delta}) f(\Delta, r_{t+\Delta}|r_t))}{\partial(r^2_{t+\Delta})} = 0 \quad (2.67)$$

$$\frac{\partial f(\Delta, r_{t+\Delta}|r_t)}{\partial \Delta} = -\frac{\partial(\mu(r_{t+\Delta}, \theta) f(\Delta, r_{t+\Delta}|r_t))}{\partial r_{t+\Delta}} + \frac{1}{2}\frac{\partial^2(\sigma^2(r_{t+\Delta}) f(\Delta, r_{t+\Delta}|r_t))}{\partial(r^2_{t+\Delta})}$$

where $f(\Delta, r_{t+\Delta}|r_t)$ is the transition density function between two successive observations corresponding to the Markov process of $dr(t) = (\alpha + \beta r(t))dt + \sigma r(t)^\gamma dW(t)$. In particular, if $\mu(x;\psi)$ and $\sigma(x;\psi)$ are specific functional forms for the drift and diffusion processes, using the parameter vector ψ, then

$$\pi(x;\psi) = \frac{\eta(\psi)}{\sigma^2(x;\psi)} \exp\left[\int^x \frac{2\mu(u;\psi)}{\sigma^2(u;\psi)} du\right], \quad (2.68)$$

where the lower limit of the integration and $\eta(\psi)$ is a constant that ensure the integration goes to one. If $\mu(x;\psi)$ and $\sigma(x;\psi)$ are adequate representations of the drift and diffusion parameters, then for some parameter choice ψ^*, the parameterized density $\pi(x;\psi^*)$ should be close to the nonparametric density estimated from the data.

Ait-Sahalia assumes a drift function and the diffusion process have the following forms, respectively.

$$\mu(x;\psi) = \alpha_0 + \alpha_1 x + \alpha_2 x^2 + \alpha_3 \frac{1}{x} \quad (2.69)$$

[8] A kernel density has the following form:

$$\hat{f}(r) = \frac{1}{Th^m} \sum_{t=1}^{T} K\left[\frac{r - r_t}{h}\right] \quad (2.66)$$

where $K(.)$ is a kernel function and h^m or h is the bandwidth. For an extensive discussion on Kernel density function, see Karlin and Taylor (1981) in Chapter 5.

$$\sigma(x;\psi) = \beta_0 + \beta_1 x + \beta_2 x^{\beta_3} \tag{2.70}$$

where ψ is a vector of $\psi \equiv [\alpha_0, \alpha_1, \alpha_2, \alpha_3, \beta_0, \beta_1, \beta_2, \beta_3]'$. This procedure requires a number of restrictions to ensure that the stochastic differential equation would have a unique solution and that the stationary density would be achieved. He then reconstructs the drift and diffusion of the continuous-time process by matching those densities estimates using the minimization of the mean square distance measure. The minimization equation has the following solution.

$$\psi^* \equiv \arg\min \frac{1}{T}\sum_{t=1}^{T}(\pi(x_t;\psi) - \hat{\pi}(x_t))^2, \tag{2.71}$$

In the empirical part, Ait-Sahalia uses the seven-day Euro-dollar deposit rate, bid-ask midpoint starting from June 1, 1973 to February 25, 1995 on the assumption that this Euro-dollar rate typically moves closely to other short-term interest rate.

Stanton (1997) also evaluates the finite sample using Monte-Carlo Simulation by arguing that the parametric estimation method is subjected to mis-specification since fitting historical data is not entirely guaranteed of matching the entire distribution and this can cause misspecified functional form. Also, since there are so many term structure models, there is no reason that one should prefer one functional form over another. He finds that the estimated drift drops sharply as the interest rate increases beyond 14 percent. The interest rate process is simulated using a non-parametric estimation method, employing the Taylor expansion series and approximates the discretized version of the two well-known interest rate models. To avoid having to specify functional form of μ and σ, a Kernel Density function is employed to estimate the model. A kernel estimator has the following form:

$$\hat{f}(r) = \frac{1}{Th^m}\sum_{t=1}^{T} K\left[\frac{r - r_t}{h}\right] \tag{2.72}$$

where $K(.)$ is a kernel function and h^m is the bandwidth.

A kernel function is normally a hump-shaped, with high density around zero and tailing off to 0 at both tails (± 1). Three necessary conditions for the kernel function are as follow:

$$Volume: \int_{-1}^{1} K(u)du = 1,$$

$$Variance: \int_{-1}^{1} u^2 K(u)du > 0,$$
$$Symmetry: K(u) = K(-u).$$

Kernel estimation is a non-parametric method for estimating the joint probability density of a set of the random variables. Statisticians as well as Econometricians have different opinions on the choices of which $K(.)$ and h to use. In general, h detemines how much the smoothing of the function is done. The choice of this bandwidth, however, are normally made through the trial-and-error process and even subjective to each personal preferences.[9] In this article, Stanton chooses $h = \hat{\sigma}_i T^{-\frac{1}{m+4}}$, where $\hat{\sigma}_i$ is the standard deviation, T is the total observations, and m is the dimension of the variables.

Epanechnikov (1969) shows that any reasonable kernel would result in almost optimal results. The most frequently used kernel functions are

Type	$K(u), u \in [-1,1]$
Uniform	$\frac{1}{2}$
Epanechnikov	$\frac{3}{4}(1-u^2)$
Quartic	$\frac{15}{16}(1-u^2)^2$
Gaussian	$\frac{1}{\sqrt{(2\pi)}} \exp^{-\frac{u^2}{2}}$

In term structure estimation, a Gaussian kernel is often used. Stanton uses the Taylor series expansion to approximate both the drift and the diffusion functions. These, however, were done seperately. From the differential equation

$$dr_t = \mu(r_t)dt + \sigma(r_t)dW_t \tag{2.74}$$

[9]A common selection method, for example, is to use a version of a data-dependent rule called "Least Square Cross Validation". This LSCV bandwidth is normally selected and to minimize the distance of

$$\arg\min \frac{1}{T}\sum_{t=1}^{T}[y_t - \hat{m}_{h_i}(x_t)]^2 \omega(x_t) \equiv (\frac{1}{T}h_i) \tag{2.73}$$

where $i = 1, ..., k$ $\hat{m}_{h_i}(x_t)$ is the fitted value of the kernel regression estimated at $x_t, \omega(x_t)$ is a function that weights the observation at x_t, and $\equiv (\frac{1}{T}h_i)$ is a function that penalizes small bandwidths. Also, see Hardle (1990) for a more detailed explanation of bandwidth, h, selection process.

taking the Taylor expansion yields

$$E_t[f(r_{t+\Delta}, t+\Delta)] = f(r_t,t)+\mathcal{L}f(r_t,t)\Delta+\frac{1}{2}\mathcal{L}^2 f(r_t,t)\Delta^2+...+\frac{1}{n!}\mathcal{L}^n f(r_t,t)\Delta^n+O(\Delta^{n+1}) \quad (2.75)$$

where $\mathcal{L}$ is the infinitesimal generator of the process (r_t).

$$\mathcal{L}f(r_t,t) = \frac{1}{\Delta}E_t[f(r_{t+\Delta}, t+\Delta) - f(r_t,t)] - \frac{1}{2}\mathcal{L}^2 f(r_t,t)\Delta - \frac{1}{6}\mathcal{L}^3 f(r_t,t)\Delta^2 - ... \quad (2.76)$$

Ignoring all terms but the first term, the first order approximation is

$$\mathcal{L}f(r_t,t) = \frac{1}{\Delta}E_t\left[f(r_{t+\Delta}, t+\Delta) - f(r_t,t)\right] + O(\Delta) \quad (2.77)$$

For higher order approximation, e.g., time step equals 2Δ,

$$\mathcal{L}f(r_t,t) = \frac{1}{2\Delta}E_t\left[f(r_{t+2\Delta}, t+2\Delta) - f(r_t,t)\right] - \frac{1}{2}\mathcal{L}^2 f(r_t,t)(2\Delta) - \frac{1}{6}\mathcal{L}^3 f(r_t,t)(2\Delta^2) - ... \quad (2.78)$$

Multiply the equation by 2 then subtract from the above equation yields the second order approximation which is

$$\mathcal{L}f(r_t,t) = \frac{1}{2\Delta}(4E_t[f(r_{t+\Delta}, t+\Delta) - f(r_t,t)] - E_t[f(r_{t+2\Delta}, t+2\Delta)] - f(r_t,t)) + O(\Delta^2) \quad (2.79)$$

In the same manner, the third order approximation can be generated as

$$\mathcal{L}f(r_t,t) = \frac{1}{6\Delta}(18E_t[f(r_{t+\Delta}, t+\Delta) - f(r_t,t)] - 9E_t[f(r_{t+2\Delta}, t+2\Delta) - f(r_t,t)]$$
$$+2E_t[f(r_{t+3\Delta}, t+3\Delta) - f(r_t,t)]) + O(\Delta^3) \quad (2.80)$$

The drift, μ_i, is approximated and it yields the first, second, and third order approximations, respectively, as

$$\mu_1(r_t) = \frac{1}{\Delta}E_t[r_{t+\Delta} - r_t|r_t] + O(\Delta) \quad (2.81)$$

$$\mu_2(r_t) = \frac{1}{2\Delta}(4E_t[r_{t+\Delta} - r_t|r_t] - E_t[r_{t+2\Delta} - r_t|r_t]) + O(\Delta^2) \quad (2.82)$$

$$\mu_3(r_t) = \frac{1}{6\Delta}(18E_t[r_{t+\Delta} - r_t|r_t] - 9E_t[r_{t+2\Delta} - r_t|r_t]$$

$$+2E_t[r_{t+3\Delta} - r_t|r_t]) + O(\Delta^3) \quad (2.83)$$

At the same time, the diffusion process is done in the same way, and the first, second, and third order approximations for σ_i^2 are

$$\sigma_1^2(r_t) = \frac{1}{\Delta}E_t[(r_{t+\Delta} - r_t)^2|r_t] + O(\Delta) \quad (2.84)$$

$$\sigma_2^2(r_t) = \frac{1}{2\Delta}(4E_t[(r_{t+\Delta} - r_t)^2|r_t] - E_t[(r_{t+2\Delta} - r_t)^2|r_t]) + O(\Delta^2) \quad (2.85)$$

$$\sigma_3^2(r_t) = \frac{1}{6\Delta}(18E_t[r_{t+\Delta} - r_t|r_t] - 9E_t[r_{t+2\Delta} - r_t|r_t] + 2E_t[r_{t+3\Delta} - r_t|r_t]) + O(\Delta^3) \quad (2.86)$$

and take the square root yields the diffusion process as

$$\sigma_1(r_t) = \sqrt{\frac{1}{\Delta}E_t[(r_{t+\Delta} - r_t)^2|r_t]} + O(\Delta) \quad (2.87)$$

$$\sigma_2(r_t) = \sqrt{\frac{1}{2\Delta}(4E_t[(r_{t+\Delta} - r_t)^2|r_t] - E_t[(r_{t+2\Delta} - r_t)^2|r_t])} + O(\Delta^2) \quad (2.88)$$

$$\sigma_3(r_t) = \sqrt{\frac{1}{6\Delta}(18E_t[r_{t+\Delta} - r_t|r_t] - 9E_t[r_{t+2\Delta} - r_t|r_t] + 2E_t[r_{t+3\Delta} - r_t|r_t])} + O(\Delta^3) \quad (2.89)$$

Normally, the higher the order of the approximation, the faster it will converge to the true drift and diffusion of the process at finer and finer time intervals. Stanton uses two interest rate models, i.e., the Cox, Ingersoll, and Ross-Square Root (1985) and the Black, Derman, and Toy (1990) models.[10] The simulation method is implemented to understand how the approximation methods perform over different sampling intervals. Results show that the approximation performed remarkably well. Setting the true parameter values $\kappa = 0.50$, $\theta = 0.07$, and $\sigma = 0.10$, he finds that regardless of the sampling interval used, the approximations on both models of any approximation order are all indistinguishable from the true drift and true diffusion processes for low values of r_t. However, as the sampling interval decreases, the performance of all of

[10]The CIR-SR (1985) and the BDT (1990) have the following forms, respectively:

$$dr(t) = \kappa(\theta - r_t)dt + \sigma\sqrt{r(t)}dW(t)$$

$$dr(t) = r[\kappa(\theta - \ln r_t) + \frac{1}{2}s^2]dt + \sigma r(t)dW(t)$$

the approximations deteriorates and the deterioration are worse for the lower order approximations. The errors introduced by using approximations method rather than the true drift and diffusion are extremely small.

Stanton also simulated empirical method using secondary market yields on three and six month Treasury Bills between January 1965 and July 1995. Using Nadaraya-Watson (N-W) kernel regression estimator to construct non-parametric estimates , the conditional expectation is estimated as

$$E[r_{t+\Delta} - r_t | r_t = r] = \frac{\sum_{t=1}^{T-1}(r_{t+\Delta} - r_t)K[(r - r_r)/h]}{\sum_{t=1}^{T-1} K[(r - r_t)/h]} \tag{2.90}$$

where $K(.)$ assumes a gaussian distribution, $\{r_n\}_{i=1}^{N}$ is a set of N points defining an equally spaced partition of a subset of the support of the stationary density. The $\frac{K[(r-r_r)/h]}{\sum_{t=1}^{T-1} K[(r-r_t)/h]}$ is the weighted average of the observed interest rate changes. It represents the influence that r_t has at the point r_n, as a proportion of the total influence from all points at r_n. Result shows that the first-, second-, and third-order approximations all yield very similar estimates. Further, he claims that the estimated drift does not look linear. The low and medium values of the interest rates, r_t, show only a slight mean reversion. Only when the interest rate, r_t, value is higher than 14 percent does the estimated drift show strong drops. The diffusion estimation, on the other hand, shows that the interest rate process is non-stationary. However, a swift nonlinear decline of the drift process at high interest rates helps reduce the explosiveness of the interest rate process eventhough a high volatility might have caused this nonstationarity.

V. Efficient Method of Moment

Dai and Singleton (1998), Gallant and Long (1997), and Anderson and Lund (1997) use the Efficient Method of Moment to estimate multi-factor interest rate processes. The basic idea of the Efficient Method of Moment estimation principle depends on a two-step process. The first step is to specify and estimate an auxiliary time series model, called the score generator. This step is comparable to the GMM. However, EMM's relative advantage over the GMM is that it offers considerable guidance on the choice of moment conditions.

The basic idea is that given a structural model for y_t, $t = 1, ..., T$, and parameters θ, we want to compute theoretical values of $E[f(y_t)|\theta]$. By simulating a long sample path $\hat{y}_t|\theta$, we can compute an estimate of

$$\hat{g}(\theta) = \frac{1}{N}\sum_{t=1}^{T} f(\hat{y}_t|\theta) \tag{2.91}$$

$$J(\theta) = [f - \hat{g}(\theta)]' W [f - \hat{g}(\theta)] \tag{2.92}$$

Andersen and Lund (1997) extends the CKLS model further by incorporating a stochastic volatility factor into the diffusion function. This specification involves both a stochastic volatility factor and a level effect, i.e., the specification of the short rate process implies mean reversion of the interest rate level as well as mean reversion of the (log) variance of interest rate changes.

$$dr_t = \kappa_1(\mu - r_t)dt + \sigma r_t^\gamma dW_{1,t} \tag{2.93}$$

$$d\log(\sigma_t^2) = \kappa_2(\alpha - \log\sigma_t^2)dt + \psi dW_{2,t} \tag{2.94}$$

where $W_{1,t}$ and $W_{2,t}$ are independent standard Brownian Motion Processes, κ_1, κ_2 represent the mean reversion of the interest rate and the mean reversion of the stochastic volatility, respectively. Specifically, the Koedijk's (1994) level-AR(s) and Nelson's E-GARCH(p, q) dynamic are selected as structural forms as follows.

$$\Delta r_t = \phi_0 + \phi_1 r_{t-1} + \sum_{i=1}^{s-1} \phi_{i+1}\Delta r_{t-i} + r^\gamma{}_{t-1}\sqrt{h_t} z_t \tag{2.95}$$

$$\log h_t = \omega + \sum_{i=1}^{p} \beta_i \log h_{t-i} + (1 + \alpha_1 L + ... + \alpha_q L^q)\left[\theta_1 z_{t-1} + \theta_2(b(z_{t-1}) - \sqrt{\frac{2}{\pi}}\right] \tag{2.96}$$

where L denotes the lag-operator, z_t is standard Gaussian, and b_z is a two times differentiable approximation to the absolute value function of $|z|$. The score generator is given by

$$f_k(r_t|x_t;\eta) = \frac{[P_k(z_t, x_t)]^2\phi(z_t)}{\int[P_k(u, x_t]^2\phi(u)du}\frac{1}{r^\gamma{}_{t-1}\sqrt{h_t}} \tag{2.97}$$

where $\phi(.)$ is the standard normal density, $z_t = \frac{(r_t - \mu_t)}{r^\gamma{}_{t-1}\sqrt{(h_t)}}$, $\mu_t = \phi_0 + (1 + \phi_1)r_{t-1} + \sum_{i=1}^{s-1}\phi_{i+1}\Delta r_{t-i}$, $P_k(z, x_t) = \sum_{i=0}^{K_z} a_i(x_t)z^i = \sum_{i=0}^{K_z}\left[\sum_{j=0}^{K_x} a_{ij}x_t^j\right] z^i, a_{00} = 1$, with h_t

given above and η is a vector containing the parameters of Level-AR(s)-EGARCH(p,q) terms. We then can span the score generators to get the quasi-maximum likelihood. The moment condition is the expectation of the score function of

$$m_N(\rho, \hat{\eta}) = \int \frac{\partial \log f_K(r|X;\hat{\eta})}{\partial \eta} dP(r, X; \rho) \tag{2.98}$$

The EMM estimator of ρ is obtained by minimizing a quadratic form in the vector of moment conditions, $m_N(\rho, \hat{\eta})$, using the weighting matrix W_t:

$$\hat{\rho} = \arg\min_{\rho} m_N(\rho, \hat{\eta})' W_T m_N(\rho, \hat{\eta}) \tag{2.99}$$

Using the treasury bills from January 1954 to April 1995, the result shows that the point estimates of γ are very close to the CIR-SR model. However, the mean reversion parameter is not significant. The stochastic volatility setting that is introduced into the model greatly enhances the model's ability to fit the data.

Earlier research show that many statistical methods are used to calibrate the term structure of the interest rates. Nevertheless, there are still a lot of unsettled debates on which model is the most accurate in terms of fitting the parameters. Some researchers conclude that a high value of the volatility elasticity, γ, fit the term structure model better while others argue that a moderate value of the volatility elasticity fit the model the best. However, the only agreed-upon study is that the drift function, especially the mean reversion parameter, β, is very hard to detect. In the following chapter, we would like to introduce the Monte-Carlo Simulation to show that the mean reversion parameter does exist, even when the aggregation of data changes.

Table 2.1: Parameter Restrictions Imposed by Models of One-Factor Short-Term Interest Rate.

Model		α	β	σ^2	γ
Merton (1973)	$dr(t) = (\alpha)dt + (\sigma)dW(t)$		0		0
Vasicek (1977)	$dr(t) = (\alpha + \beta r(t))dt + \sigma dW(t)$				0
Cox, Ingersoll, and Ross SR (1985)	$dr(t) = (\alpha + \beta r(t))dt + \sigma\sqrt{r}(t)dW(t)$				1/2
Dothan (1978)	$dr(t) = \sigma r(t)dW(t)$	0	0		1
Geometric Brownian Motion	$dr(t) = \beta r(t)dt + \sigma r(t)dW(t)$	0			1
Brennan-Schwartz (1980)	$dr(t) = (\alpha + \beta r(t))dt + \sigma r(t)dW(t)$				1
Cox, Ingersoll, and Ross VR (1980)	$dr(t) = \sigma r^{3/2}(t)dW(t)$	0	0		3/2
Constant Elasticity of Variance	$dr(t) = (\beta r(t))dt + \sigma r(t)^{\gamma}dW(t)$	0			
CKLS (1992)	$dr(t) = (\alpha + \beta r(t))dt + \sigma r(t)^{\gamma}dW(t)$				

Note: CKLS sets a unifying framework of the one-factor interest rate process. His unrestricted model can be modified into any one-factor process. For example, by letting β and $\gamma = 0$, we get the Merton model. Setting the $\gamma = 0$, we get the Vasicek model. Setting $\gamma = \frac{1}{2}$, we get the CIR model. Setting $\alpha = 0$, $\beta = 0$ and $\gamma = 1$, we get the Dothan model, and amongst others.

Chapter 3

Monte-Carlo Simulation on CKLS Nested Models

I. Introduction

Information on data frequency is a crucial resource for financial participants who are actively involved in the financial market, especially when one tries to estimate term structure of interest rate parameters with different maturities. Because the rate of return on investment in financial assets is different from one time span to another, the way the data affects statistical inferences based on different frequencies must be different. Participants in the financial markets usually have different trading strategies. Traders with open positions in the market are exposed to various degree of risks associated with changes in the interest rates. Short-term investors, for example, would look for a day-to-day fluctuation to make quick profits to fulfill their speculative objectives. Long-term investors, on the other hand, would buy and hold their positions for a certain length of time. At any calculated risk level, they try to estimate parameters related to the term structure in order to maximize profits. Because of the explosive growth in this area, a lot of research efforts have been devoted to the development of interest rate process methodology. This fact leads to the idea of this essay on estimating term structure using different sampling frequencies and measuring its effect on statistical inferences.

Knowledge of consequences of using data with different frequencies can be helpful in answering the important question of whether mean reversion is hard to detect when aggregation of data varies. Although research has been done on the consequences of using data with different frequencies, little is known about the way data based on different frequencies affect inference on mean reversion and the elasticity of volatility in the CKLS' (1992a) short-term interest rate model. In addition, given the controversy surrounding the magnitude of volatility estimates, it is important to investigate whether the size of the parameter estimates is related to the sampling frequencies used in various studies. In this dissertation, we examine the small sample properties of estimators of mean reversion and the elasticity of volatility with respect to the level of interest rates using data sampled at daily, weekly, and monthly intervals. This dissertation corroborates Koutmos' findings, who considered both the issues of the sensitivity of mean reversion and volatility parameter estimates across maturities and frequencies using actual treasury data. One of Koutmos' conclusions is that there is no statistical evidence that interest rates revert to its long-run mean. This follows his claim that the mean reversion parameter is statistically insignificant, observed at daily, weekly, and monthly frequencies data. However, in this dissertation, we will illustrate that the mean reversion parameter can be observed at any sampling interval regardless of the data frequency used. The question arises on how this interest rate dynamics based on different frequency can be understood and properly used both in academic and in practice. Agbeyegbe and Tanarugsachock (2000) (AT) have developed a Monte-Carlo simulation approach to study the way data affects statistical inferences on mean reversion and the volatility elasticity parameter estimation due to aggregation of data sampling frequencies.[1] Although many researchers have studied these term structures, there are no concensus which model is universally best fitting those parameters. Since controversies surround this topic is still a debate, it is very suitable to show that the Monte-Carlo simulation approach can shed some light on these important issues and investigate how data aggregation affects the small sample behavior of the maximum

[1]Studies on mean reversion of the interest rate have been done somewhat extensively. Bessembinder et. al (1995) looked at mean reversion from the Futures Term structure. They look at the term structure of futures prices to test whether investors anticipate mean reversion in spot asset prices. Mollick (1999) looks at Brazil's real exchange rate and finds mixed result of the mean reversion and random walk. Poterba and Summers (1988) predicted mean reversion in stock prices, focusing on the possibility that mean reversion results from temporary divergences of prices from fundamental value. Fama and French (1988) argue that negative comovement between prices and risk premia can generate mean reversion in equilibrium. However, we look at the mean reversion of the interest rate context.

likelihood estimators.[2] This gives a valid reason to investigate into the sampling frequency and its statistical inference. Here, a Monte-Carlo simulation approach is used to examine if a mean reversion is difficult to detect when the data is aggregated (as claimed by other researchers) and explore whether the sensitivity of the parameters is affected by the magnitude of the volatility. Since simulating sample paths is very important for validating any claimed estimation procedure, we can address the issue of estimator bias where a result of accuracy depends very highly on how data aggregation are formulated.

This analysis extends Koutmos' (1998) study further by examining the following important issues.

(1) How does aggregation of data affect the small sample behavior of the maximum likelihood estimators of the parameters of the models? Is the mean reversion hard to detect when data is aggregated? Are the magnitudes of the elasticity estimates sensitive to sampling frequencies?

(2) How does confidence regions of the parameter constructed from the small sample simulation perform relative to the confidence regions based on the usually constructed standard errors?

(3) How are small-sample biases of the maximum likelihood estimators affected by the choice of the volatility elasticity estimates?

Although Koutmos' conclusion shows no statistical evidence on the mean reversion to its long-run level due to its insignificance of statistical inferences, the findings here show otherwise. We divided the processes into two sections. The first section shows a methodological set-up of the Monte-Carlo simulation. The methodology is purposely set-up so that it would show both the proper maximum likelihood and the

[2]An interesting Monte-Carlo study that is parallelled to this work is that of Pristker (1998), who studies the finite sample distribution on Ait-Sahalia's (1996) non-parametric work. In addition, Stanton (1997) uses MC simulation to study the drift and diffusion processes separately on different interest rate models using different orders of Taylor approximation. Honore (1998) analyzes the differences between an estimator using the full distribution and the Generalized Method of Moment in the CKLS model. Again, he studies the method of bias reduction which is closely related to Gourieroux and Monfort's indirect inference (1996). However, none of them addresses the issues of different sampling frequencies that we look at in this paper.

mis-specification maximum likelihood estimations. The second section shows results of the Monte-Carlo estimation.

II. Methodology and Monte-Carlo Set Up

First, the simulation of interest rate sample paths is generated with known CKLS parameter values. The general framework of CKLS is given below.

$$dr(t) = [\alpha + \beta r(t)]dt + \sigma r(t)^{\gamma} dW(t), \forall t \in [0, T] \tag{3.1}$$

For estimation purpose, the stochastic differential equation as described in equation 3.1 is discretized to stochastic difference equation, using Euler-Maruyama method. This is the method employed by Marsh and Rosenfeld (1983), CKLS (1992a), and Dietrich-Campbell and Schwartz (1986), and others. We consider the following Euler Maruyama discrete-time approximation to the continuous-time model given in equation (3.4).[3] It then is discretized as the following equations.

$$r_{t+1} - r_t = \alpha + \beta r_t + \sigma r_t^{\gamma} \epsilon_{t+1}, t = [0, 1, 2, ..., T] \tag{3.2}$$

$$r_{t+1} = \alpha + (\beta + 1) r_t + \sigma r_t^{\gamma} \epsilon_{t+1}, t = [0, 1, 2, ..., T] \tag{3.3}$$

$$r_{t+1} = \alpha + (\beta + 1) r_t + \sigma(| r_t |^{\gamma}) \epsilon_{t+1}, t = [0, 1, 2, ..., T] \tag{3.4}$$

Three selective nested CKLS models are chosen and replicated to generate these interest rate paths. While not exhaustive of all one-factor diffusion models, it does include several of the most important diffusion processes used including the Vasicek Model (1977), the Cox, Ingersoll, and Ross - Square Root Model (1985), as well as

[3] Second-order discretization schemes are also available when using other approximation steps as in Milshtein (1974) and Pardoux and Talay (1985). In fact, it has been demonstrated that for some special cases, e.g., Vasicek (1977) and CIR (1985) models, that it is more efficient to work on exact discrete-time analogs of the underlying continuous-time models. Unfortunately, no correct proof exists for the exact discrete analog for the CKLS model, although Nowman (1997, 1998) in a series of interesting papers, has derived an exact discrete analog to a model which closely resembles the CKLS model. Nowman's proof uses the general methodology to deriving exact discrete model of first-order continuous-time models proposed by Bergstrom (1983, 1984). Agbeyegbe (1983, 1987, 1988) discusses the estimation and derivation of first-order mixed sample continuous-time systems.

the Brennan-Schwartz Model (1980).[4] Interestingly, these models are selected because they are widely used in practice and in academia. Moreover, they satisfy the ergodicity and stationarity conditions.

The vector of true parameters are listed below.

$$\theta = [\alpha, \beta, \sigma, \gamma]'$$

Model	Interest Rate Process	α	β	σ	γ
Vasicek	$dr(t) = [\alpha + \beta r(t)]dt + \sigma dW(t)$	0.200	-2.000	0.0225	0.000
CIR-SR	$dr(t) = [\alpha + \beta r(t)]dt + \sigma\sqrt{r(t)}dW(t)$	0.200	-2.000	0.0225	0.500
Brennan Schwartz	$dr(t) = [\alpha + \beta r(t)]dt + \sigma r(t)dW(t)$	0.200	-2.000	0.0225	1.000

These true parameter values are chosen, representing plausible values from previous empirical studies. Using maximum likelihood estimation method that accounts for a sampling interval, Broze, Scaillet, and Zakovian (1995) show that the interest rate process, r_t^h, is ergodic and stationary.[5] Broze, Scaillet and Zakovian (1995) prove that the models that have $\gamma \geq 1$ is non-stationary as shown in proposition 3 of his paper.

To identify statistical inference from the model, we use two methods of the maximum likelihood estimation: one that accounts for sampling intervals in which we will call them the correct specification of the maximum likelihood estimation henceforth, and the other which does not take into account the sampling interval of time horizon and we will call them the mis-specified maximum likelihood estimation. Equation 3.4 is thus transformed into the interest rate process that takes account of the sampling interval, as follow.

$$r_{t+1}^{(h)} = \alpha_h + (\beta_h + 1)r_t^{(h)} + \sigma_h(| r_t^{(h)} |^{\gamma})\epsilon_{t+1}, t = [0, 1, 2, ..., T] \tag{3.5}$$

where h is the sampling interval, $\alpha_h = \alpha h$, $\beta_h = \beta h$, $\sigma_h = \sigma\sqrt{h}$, and ϵ is i.i.d. sequence of standard Gaussian variables independent of the past time processes.

[4]Chapman, Long, and Pearson (1999), Stanton (1997), and among others also used these models in their research since they are very common and widely used.

[5]Broze, et al. show that if $\gamma \neq 1$, then the interest rate process , $r_t^{(h)}$, is ergodic and second-order stationary if and only if $|\beta_h+1| < 1$ and $0 \leq \gamma < 1$. He also show that if $\gamma = 1$, then a sufficient condition for second-order stationarity is $(\beta_h + 1)^2 + \sigma_{0,h}^2 < 1$. In addition, a sufficient condition for ergodicity is $E[log|(\beta + 1) + \gamma_{0,h}W_{t+1}|] < 0$.

A sequence of interest rate series, $r_t^h \equiv (r_0^{(h)}, \ldots, r_{T-1}^{(h)})$, is generated with the initial interest value, $r_0^{(h)}$, of 0.10. In a correct specification, data is generated as follows. Daily data, (r_t^d), is generated with total time points equal to 2400 and 4800, representing 10 and 20 years of daily data, respectively. The simulated sample paths are constructed under the assumption that the length of time between observations of the diffusion is $\Delta = 1/240$, corresponding to daily observations. The second and larger set of value (4800 obs.) allows us to evaluate the effect on the estimator of doubling the sample size. At the same time, weekly data, (r_t^w), is generated by skipping 5 daily time points, i.e., $\Delta = 5/240$, yielding either 720 and 1440 for 10 years and 20 years simulated data, respectively. Monthly data, (r_t^m), is generated by using 20 skipped time points, i.e., $\Delta = 20/240$, generating either 180 and 360 monthly observations for 10 years and 20 years, respectively. The mis-specification method of maximum likelihood, on the other hand, ignores the sampling frequency information, used $\Delta = 1/240$, regardless of the daily, weekly, or monthly observations. In both approaches, we assume that the approximated discrete-time process is a proxy of the true time-series model.

The standard normal distribution (Gaussian distribution) is used for the innovation process. Also, the Newton-Raphson algorithm is used in the process of numerical optimization in the maximum likelihood estimation.[6] A sample of 1000 simulations are generated and results are examined.

The maximum likelihood estimator of $\theta^{(h)}$ is

$$\max_{\theta^{(h)}} \mathcal{L}_T(r^{(h)}; \theta^{(h)}) = \sum_{t=1}^{T} log f_t$$

$$\max_{\theta^{(h)}} \mathcal{L}_T(r^{(h)}; \theta^{(h)}) = -\frac{T}{2} log(\sigma_{0,h}^2) - \frac{1}{2}\left[\sum_{t=1}^{T} log(\mid r_{t-1}^{(h)} \mid)^{2\gamma} + \sum_{t=1}^{T} \frac{(r_t^{(h)} - \alpha_h - (\beta_h + 1) r_{t-1}^{(h)})^2}{\sigma_{0,h}^2 (\mid r_{t-1}^{(h)} \mid)^{2\gamma}}\right] \tag{3.6}$$

[6]Preliminary analysis shows no difference in numerical optimization methods employed whether it is BHHH, Newton-Raphson, or Gauss-Newton methods. However, BHHH requires a much longer processing time to convergence than the Newton-Raphson and Gauss-Newton methods.

III. Monte-Carlo Simulation Results

Results of the simulations are reported in Table 3.1 to Table 3.18 at the end of this chapter. The mis-specification model is reported first from Table 3.1 to Table 3.9, and followed by the correct specification model which is reported in Table 3.10 to Table 3.18. Tables 3.1 and 3.2 report usual descriptive statistics where the average, the standard deviation, the minimum, and the maximum values of mis-specification models are shown. Table 3.1 and Table 3.2 show daily, weekly, and monthly sampling frequency results which were generated from ten-year (where T = 2,400) and from twenty-year (where T = 4,800) simulated sampling data, respectively. It is clear that while the mean reversion parameter, β, vary noticeably as the level of aggregation changes from daily to weekly, and to monthly intervals, the volatility elasticity, γ, does not change significantly as the aggregation of data varies. This is observable throughout every classes of the interest rate processes regardless of the models implemented.[7] The mean reversion parameter, β, using ten-year simulated data, ranges from the minimum of -3.7882 to the maximum of -0.3547 in case of daily data, from the minimum of -18.1049 to the maximum of -1.8660 in case of weekly data, and from the minimum of -49.9735 to the maximum of -8.2222 for monthly data.[8] At the same time, β, using twenty-year simulated data, ranges from the minimum of -3.1257 to the maximum of -1.0151 for daily data, ranges from -15.5262 to -4.9314 for weekly data, and ranges from -49.9714 to -20.2077 for monthly data.

The average values of the mean reversion parameter, $\hat{\beta}$, of daily frequencies are estimated at -2.0484, -2.0609, and -2.0035 for the Vasicek, the CIR-SR, and the Brennan-Schwartz models, respectively. The average weekly values of $\hat{\beta}$ are estimated at -10.0384, -10.0850, and -9.7975 while the average monthly estimated values of $\hat{\beta}$ are at -37.0151, -37.1751, and -36.2802. Evidently, using mis-specified maximum likelihood function creates strong distortion in estimating interest rate process as evidenced by

[7]To be specific, three interest rate models used here are the Vasicek, the CIR-SR, and the Brennan-Schwartz models.

[8]We compare sampling frequencies of day to week and to month for each interest rate model. For example, model that used daily data frequency is compared to model that used weekly frequencies and to model that used monthly frequencies. In this way, we look at how the interest rate behaves with different interval frequencies instead of comparing it to the day models with different volatility elasticity parameter, γ.

misrepresented average values of the simulations shown in Table 3.1. The ten-year simulated data shows higher fluctuation than those of the twenty-year simulated data. Also, notice that the range of the interest rate process is at the highest when the level of aggregation is changed to the monthly interval. As level of aggregation of data decreases, the average values of the mean reversion parameters become more distorted so severely that small sample cannot be used to reflect the true value. The longer data point would not help in eliminating this distortion. Average values do not change much even if the twenty-year simulated data is used. With average daily frequency values of $\hat{\beta}$ generated from the twenty-year simulated data (as shown in Table 3.2), they are estimated at -2.0598, -2.0579 and -2.0186 for daily frequencies. The average weekly frequency values of $\hat{\beta}$ are recorded at -10.1184, -10.1068, and -9.9039. At the same time, the average monthly frequency values of $\hat{\beta}$ are recorded at -37.9354, -37.8020, and -37.0255. Average weekly and monthly values on both ten-year and twenty-year simulated data, using the mis-specification likelihood, show strong distortions from true parameter values. In addition, the standard deviations of these mean reversion parameters, β, of each models increase almost 5 times to 20 times as the sampling frequencies change from daily to weekly, and to monthly intervals, respectively. This is also the case even if we use twenty-year simulated data. Longer sampling frequencies do not help curbing deviation if the mis-specification of the likelihood model is being implemented.

The average volatility elasticity parameters, $\hat{\gamma}$, behave much better and are closer to their true values in every interest rate models. Since true value of the volatility elasticity parameter has its unique value and which it depends on each interest rate model, we compare the estimates within the interest rate model itself. For the Vasicek model (where the true parameter value of $\gamma = 0.000$), the estimated average values of $\hat{\gamma}$ are reported at -0.0003, 0.0013, and 0.0050 for the daily, weekly, and monthly intervals. For CIR-SR model (where the true parameter value of $\gamma = 0.500$), the estimated average values of $\hat{\gamma}$ are reported at 0.4994, 0.4953, and 0.4664. For the Brennan-Schwartz model (where the true parameter value of $\gamma = 1.000$), the estimated average values of $\hat{\gamma}$ are 0.9980, 0.9989, and 0.9867. Notice that average $\hat{\gamma}$ in every interest models are very close to their true values. The average values of the volatility elasticity, $\hat{\gamma}$, are quite stable and do not vary from their true values as the level of aggregation changes. Standard deviations of the volatility elasticity parameters are also relatively small when compared to those of the standard deviations of the mean reversion parameters. Notice, however, that the model with true parameter of $\gamma = 1$ fluctuates the most among many models presented here. This holds true for both ten-year and twenty-year simulated data.

Table 3.3 shows the bias of the mean reversion parameter for the mis-specification model. Both ten-year and twenty-year simulated data are reported side by side. In every interest rate model, the absolute values of biases of $\hat{\beta}$ change from approximately -0.000 to -8.0000, and to -35.000 as the aggregation level changes to lower frequencies. Notice that the sample bias of the mean reversion parameter increases in magnitude with the level of aggregation. However, the bias for the volatility elasticity parameters, $\hat{\gamma}$, do not change as the aggregation of data decreases. Most interest rate models show roughly equal bias values regardless of the size of γ. While the bias of the mean reversion parameters is very sensitive to the true level of the volatility parameter, we find that the bias of the volatility parameter is not so sensitive to the level of the interest rate. Table 3.4 shows that the bias of the volatility elasticity parameters, $\hat{\gamma}$, of the mis-specification model are close to 0. This holds true even when the sampling frequency changes from daily to weekly and to monthly data. The bias is negligible even at the monthly frequency level. This result also holds when a longer data set is used. With twenty-year simulated data, the bias is very close to 0.

Tables 3.5 and 3.6 show the mean square error of the mean reversion parameter, β, and the mean square error of the volatility elasticity parameter, γ, of the mis-specification model. In general, the MSE is the overall measure of the size of the measurement error which is defined as

$$MSE = E[(\hat{\theta} - \theta)^2] = Var(\hat{\theta} - \theta) + [E(\hat{\theta} - \theta)]^2 = \sigma^2 + bias^2 \tag{3.7}$$

Like the bias of the mean reversion parameter, the MSE of the mean reversion parameter, $\hat{\beta}$, increases in magnitude as the level of aggregation changes from daily to weekly and to monthly frequencies. Again, daily data frequency yields the least mean square error and monthly frequency gives the highest mean square error in every interest rate model. Daily MSE ranges from 0.2173 to 0.2345 in three different interest rate models while the weekly MSE ranges from 66.1256 to 70.3585, and from 1238.949 to

1295.497 for the monthly frequencies. The twenty-year simulated data shows the same pattern. The MSE are worsen when the level of aggregation decreases. The percentage increase in the MSE of the mean reversion parameters increase more significantly than those of the volatility elasticity. The mean square error of the volatility elasticity, γ, is quite small.

The standard error bias of the estimators is the comparison of the standard deviation of the Monte-Carlo simulation coefficients to the mean standard errors from each sampling scheme. Essentially, the Monte-Carlo simulation standard errors are the standard deviations of the elements in each column of the parameters and the asymptotic standard errors are the average of n times of the standard deviations of the parameters of each sampling. These biases of the standard errors are used to evaluate if the estimated standard errors averaged across replications for each sampling experiment can reasonably approximate the asymptotic standard errors. Mathematically, it can be written as

$$\mathcal{VB} = \frac{\nu(\hat{\theta})}{\xi(\hat{\theta})} \tag{3.8}$$

where $\mathcal{VB}$ is the bias of standard error of the parameters estimated, $\nu(\hat{\theta})$ is the average standard errors across replications, and $\xi(\hat{\theta})$ is the asymptotic standard errors. The ratio of $\mathcal{VB}$ closest to 1 represents the Monte-Carlo standard errors can effectively approximate the asymptotic standard errors well. Table 3.7 shows the bias of the variances of the mis-specification model. Results show that the weekly and monthly frequencies of the Monte-Carlo standard errors underperformed those of the asymptotic standard errors for both ten-year and twenty-year daily data especially the mean reversion parameter, β. The Monte-Carlo standard errors only captures seventy percent of the times in most cases and the performances deteriorated even more as the frequency changes from daily to weekly and to monthly intervals. The γs bias of the variances, on the contrary, perform reasonably well in every cases.

The coverage probabilities ratio of the parameters are also estimated here and are shown in Table 3.8 and 3.9. These probabilities are computed by calculating the number of times in 1000 replications that the true value of the parameter falls within 90 percent, 95 percent, and 99 percent confidence interval of its estimator. This statistic is useful in evaluating whether there are significant size distortions when inference is made with respect to t-statistics. It has the following form:

$$\eta = \frac{\tau(\hat{\theta})}{\upsilon(\hat{\theta})} \tag{3.9}$$

where η is the coverage probabilities ratio of the parameters, $\tau(\hat{\theta})$ is the number of the times the simulation values fall within the confident regions of the estimaters, and $\upsilon(\hat{\theta})$ is the total numbers of simulation.[9] Results show that the coverage probabilities on the mean reversion parameter, β, perform poorly as the aggregation of data decreases. Varying the sampling scheme from daily, to weekly, and to monthly intervals greatly reduces the coverage probabilities, thus suggesting that tests will lead to over rejection of the null hypothesis. The twenty-year simulated data perform worse than those of the ten-year data. In most cases, the coverage probabilities fall to zero as the sampling frequencies change from daily to weekly and to monthly data.

Table 3.10 and 3.11 show the descriptive statistics for the correct specification models from the simulated data of ten years and twenty years, respectively. Both tables show averages of the mean reversion parameter, $\hat{\beta}$, that are very close to true parameter value. The average $\hat{\beta}$ are reported at -2.4068, -2.3649, and -2.2318 for the Vasicek model, at -2.4157, -2.3721, and -2.2232 for the CIR-SR model, and at -2.2907, -2.2521, and -2.1118 for the B-S model as sampling aggregation changes from daily to weekly, and to monthly data, respectively. Also, twenty-year data shows average $\hat{\beta}$s that are closer to the true values. For Vasicek model (where $\gamma = 0.0$), the average values are -2.1960, -2.1631, and -2.0409. The CIR-SR model (where $\gamma = 0.50$) shows the average values of -2.1789, -2.1439, and -2.0130 while the B-S model (where $\gamma = 1.0$) shows the average values of -2.1236, -2.0906, and -1.9606. It is clear that the mean reversion parameters, $\hat{\beta}$, do not become more noticeable as the level of aggregation decreases. The parameters are fairly stable across sampling frequencies eventhough there is a slight tendency of decreased magnitudes as the level of aggregation changes. As evidenced in Table 3.12 and 3.13, the biases of the estimates remains at an insignificant level throughout. This is also true when the twenty-year simulated data are used as well.

[9]The confidence region is $CI = [\hat{\theta} - Z_{\frac{\alpha}{2}}\hat{\sigma}, \hat{\theta} + Z_{\frac{\alpha}{2}}\hat{\sigma}]$, $\hat{\theta}$ is the parameter estimate, $Z_{\frac{\alpha}{2}}$ is the $\frac{\alpha}{2}$ critical value from the normal distribution table, and $\hat{\sigma}$ is the standard errors of the estimates.

Table 3.14 and 3.15 show the mean square error of the mean reversion parameter and elasticity volatility. The mean square errors show very slight bias in the correct specification model. This finding is important since it suggests that when the maximum likelihood function is correctly specified and appropriately adjusted, the estimates of both mean reversion parameters and volatility elasticity will be independent of any sampling interval. If mean reversion exists, one should be able to find it using daily, weekly or monthly data. The mean square error for the volatility elasticity parameter also shows numbers close to zero. Eventhough the mean square errors are worse with the level of aggregation, it is not significant.

Table 3.16 reports the bias of the standard error of the estimators using different sampling schemes. The ratio of the MCSE to the asymptotic standard errors are in the range of 0.97 to 1.7. This suggests that for most of the experiments, the MCSE can approximate the asymptotic standard errors well. Table 3.17 and 3.18 report the coverage probabilities of the correct specification of the likelihood function. In most entries, the coverage probabilities are close to the nominal sizes. It is clear that there is a substantial improvement in coverage probability when compared to the results based on the mis-specified maximum likelihood.

Figure 3.1 to 3.4 show Box-Plot graphs of mean reversion parameters, β. The correct specification were plotted against and parallel to the mis-specification models. Figure 3.1 shows a ten-year weekly simulated data of β of the proposed models, i.e., the Vacisek, CIR-SR, and Brennan-Schwartz models, respectively. In any cases, the correct specification models show averages of the mean reversion parameters close to the true parameters (where true parameters of β = -2). The mis-specified models, however, show averages of the mean reversion parameters of approximately at -10 in all three cases. Notice that in the correct specification models, the parameters estimated in each simulations are approximately close to one another. The observations of parameters estimated for the mis-specified models, on the other hand, are more scattered as the range of the simulations in the mis-specified models are from -4 to -18. Although there are some outliers in the correct-specification models, they account for only less than one percent of total simulations. The mis-specified models also shows a much fatter tail in all three cases. Figure 3.2 shows the box-plot graph of mean reversion parameters, β, on monthly sampling intervals. Again, correct specification shows average of the mean reversion parameters approximately at -2.0. The mis-specification models show strong distortions. The average of the mis-specification models are roughly at -38.00 and the range is much higher than those of the correct models. Figure 3.3 and Figure 3.4 shows a twenty-year weekly and monthly simulated data of βs, respectively. Results are similar to those of the ten-year data on both weekly and montly frequencies. The correct specification models are in line with their true parameters while the mis-specified models misrepresent the true parameters.

Table 3.1: Descriptive Statistics, Ten-year Simulated Data, n=1000, Normal Distribution, Mis-Specification Model

		$\gamma = 0.0$			$\gamma = 0.5$			$\gamma = 1.0$		
		Day	Week	Month	Day	Week	Month	Day	Week	Month
$\hat{x}$	α	0.2036	0.9968	3.6726	0.2063	1.0097	3.7243	0.2006	0.9810	3.6329
$\hat{x}$	β	-2.0484	-10.038	-37.015	-2.0609	-10.085	-37.175	-2.0035	-9.7975	-36.280
$\hat{x}$	σ	0.0225	0.1110	0.4213	0.0235	0.1347	0.9698	0.0352	0.6599	24.136
$\hat{x}$	γ	0.0000	0.0013	0.0076	0.5008	0.5001	0.4808	1.0000	1.0095	0.9868
SD	α	0.0687	0.3307	1.1759	0.0461	0.2232	0.7756	0.0484	0.2310	0.7989
SD	β	0.4637	2.2336	7.5997	0.4633	2.2341	7.6294	0.4843	2.3135	7.9885
SD	σ	0.0020	0.0205	0.1717	0.0071	0.1032	2.9822	0.0434	2.2568	89.682
SD	γ	0.0170	0.0348	0.0725	0.0618	0.1408	0.2859	0.2057	0.4221	0.7708
Min	α	0.0249	0.1349	0.3774	0.0796	0.4145	1.3600	0.0402	0.2085	0.9124
Min	β	-3.6916	-16.567	-49.966	-3.7882	-18.104	-49.973	-3.7063	-16.613	-49.952
Min	σ	0.0168	0.0601	0.0867	0.0100	0.0167	0.0031	0.0013	0.0004	0.0001
Min	γ	-0.0532	-0.1228	-0.3054	0.3310	0.0834	-0.5626	0.3971	-0.2473	-0.7768
Max	α	0.5220	2.3314	7.8205	0.3701	1.6326	5.7684	0.3770	1.6890	5.1261
Max	β	-0.8478	-4.0863	-13.219	-0.8219	-4.0480	-13.551	-0.3547	-1.8660	-8.2222
Max	σ	0.0304	0.1908	1.8107	0.0608	1.5197	47.599	0.5773	30.542	879.00
Max	γ	0.0582	0.1269	0.3904	0.7147	1.0962	1.6123	1.6976	2.2491	2.7144

Notes: The table reports the Monte-Carlo averages, standard deviations, minimum and maximum of the parameter estimates from the CKLS model. Data were generated from the model $r_{t+1}^{(h)} = \alpha_h + (\beta_h + 1)r_t^{(h)} + \sigma_h(\mid r_t^{(h)} \mid^\gamma)Z_{t+1}, t\epsilon\aleph$,
where Z_t is an i.i.d. sequence of standard Gaussian variables independent of the past and where: $\alpha_h = \alpha h$; $\beta_h = \beta h$; $\sigma_h = \sigma\sqrt{h}$; h being equal to 1, 5, or 20, representing daily, weekly or monthly frequencies; $\aleph$ is the set of natural numbers. The number of repetitions was 1000.

Table 3.2: Descriptive Statistics, Twenty-year Simulated Data, n=1000, Normal Distribution, Mis-Specification Model

		$\gamma = 0.0$			$\gamma = 0.5$			$\gamma = 1.0$		
		Day	Week	Month	Day	Week	Month	Day	Week	Month
$\hat{x}$	α	0.2056	1.0101	3.7885	0.2054	1.0092	3.7766	0.2020	0.9913	3.7066
$\hat{x}$	β	-2.0598	-10.118	-37.935	-2.0573	-10.106	-37.802	-2.0186	-9.9039	-37.025
$\hat{x}$	σ	0.0225	0.1104	0.4060	0.0229	0.1178	0.5050	0.0279	0.3048	10.241
$\hat{x}$	γ	-0.0003	0.0013	0.0050	0.4994	0.4953	0.4664	0.9980	0.9989	0.9867
SD	α	0.0505	0.2475	0.9029	0.0359	0.1767	0.6460	0.0366	0.1778	0.6415
SD	β	0.3653	1.7721	6.3352	0.3647	1.7865	6.4277	0.3687	1.7906	6.4426
SD	σ	0.0014	0.0155	0.1086	0.0049	0.0563	0.6001	0.0210	0.8361	51.937
SD	γ	0.0118	0.0262	0.0487	0.0449	0.0956	0.1952	0.1467	0.3022	0.5906
Min	α	0.0588	0.3166	1.0487	0.1152	0.5887	2.0836	0.1038	0.4961	2.0168
Min	β	-3.1144	-15.446	-49.955	-3.1218	-15.465	-49.971	-3.1257	-15.526	-49.932
Min	σ	0.0175	0.0711	0.1673	0.0082	0.0264	0.0088	0.0011	0.0017	0.0001
Min	γ	-0.0448	-0.0718	-0.1958	0.2892	0.1713	-2.2703	0.3500	0.1020	-0.7738
Max	α	0.3929	1.8788	6.7987	0.3152	1.5371	5.5362	0.3122	1.5514	5.1015
Max	β	-1.0151	-4.9722	-21.199	-1.1703	-5.7786	-21.005	-1.0269	-4.9314	-20.207
Max	σ	0.0269	0.2096	1.0459	0.0458	0.4239	8.2262	0.2127	13.2895	696.41
Max	γ	0.0394	0.1442	0.1900	0.6585	0.8158	1.1267	1.4936	2.0666	2.6350

Notes: The table reports the Monte-Carlo averages, standard deviations, minimum and maximum of the parameter estimates from the CKLS model. Data were generated from the model $r_{t+1}^{(h)} = \alpha_h + (\beta_h + 1)r_t^{(h)} + \sigma_h(\mid r_t^{(h)} \mid^{\gamma})Z_{t+1}, t\epsilon\aleph$,
where Z_t is an i.i.d. sequence of standard Gaussian variables independent of the past and where: $\alpha_h = \alpha h$; $\beta_h = \beta h$; $\sigma_h = \sigma\sqrt{h}$; h being equal to 1/240, 5/240, or 20/240, representing daily, weekly or monthly frequencies; $\aleph$ is the set of natural numbers. The number of repetitions was 1000.

Table 3.3: Bias of the Mean Reversion Parameters β, Normal Distribution, Mis-Specification Model

β	Bias (T=2400)			Bias (T=4800)		
Model	Day	Week	Month	Day	Week	Month
$\gamma = 0.0$	-0.0484	-8.0384	-35.0151	-0.0598	-8.1184	-35.9354
$\gamma = 0.5$	-0.0609	-8.0850	-35.1751	-0.0573	-8.1068	-35.8020
$\gamma = 1.0$	-0.0035	-7.7975	-34.2802	-0.0186	-7.9039	-35.0255

Notes: The table reports the biases of the mean reversion parameter estimates from the CKLS model. Data were generated from the model $r_{t+1}^{(h)} = \alpha_h + (\beta_h + 1)r_t^{(h)} + \sigma_h(| r_t^{(h)} |^{\gamma})Z_{t+1}, t\epsilon\aleph$,
where Z_t is an i.i.d. sequence of standard Gaussian variables independent of the past and where: $\alpha_h = \alpha h$; $\beta_h = \beta h$; $\sigma_h = \sigma\sqrt{h}$; h being equal to 1/240, 5/240, or 20/240, representing daily, weekly or monthly frequencies; $\aleph$ is the set of natural numbers. The number of repetitions was 1000.

Table 3.4: Bias of the Interest Rate Volatility Elasticity Parameters γ, Normal Distribution, Mis-specification Model

γ	Bias (T=2400)			Bias (T=4800)		
Model	Day	Week	Month	Day	Week	Month
$\gamma = 0.0$	0.0000	0.0013	0.0076	-0.0003	0.0013	0.0050
$\gamma = 0.5$	0.0008	0.0001	-0.0192	-0.0006	-0.0047	-0.0336
$\gamma = 1.0$	0.0000	0.0095	-0.0132	-0.0020	-0.0011	-0.0133

Notes: The table reports the biases of the interest rate volatility parameter estimates from the CKLS model. Data were generated from the model $r_{t+1}^{(h)} = \alpha_h + (\beta_h + 1)r_t^{(h)} + \sigma_h(| r_t^{(h)} |^{\gamma})Z_{t+1}, t\epsilon\aleph$,
where Z_t is an i.i.d. sequence of standard Gaussian variables independent of the past and where: $\alpha_h = \alpha h$; $\beta_h = \beta h$; $\sigma_h = \sigma\sqrt{h}$; h being equal to 1/240, 5/240, or 20/240, representing daily, weekly or monthly frequencies; $\aleph$ is the set of natural numbers. The number of repetitions was 1000.

Table 3.5: Mean Square Error of the Mean Reversion Parameters β, Normal Distribution, Mis-specification Model

β	(T=2400)			(T=4800)		
Model	Day	Week	Month	Day	Week	Month
$\gamma = 0.0$	0.2173	69.6042	1283.811	0.1370	69.0492	1331.487
$\gamma = 0.5$	0.2184	70.3585	1295.497	0.1363	68.9125	1323.100
$\gamma = 1.0$	0.2345	66.1526	1238.949	0.1363	65.6772	1268.295

Notes: The table reports the mean square error of the mean reversion parameter estimates from the CKLS model. Data were generated from the model $r_{t+1}^{(h)} = \alpha_h + (\beta_h + 1)r_t^{(h)} + \sigma_h(\mid r_t^{(h)} \mid^{\gamma})Z_{t+1}, t\epsilon\aleph$,
where Z_t is an i.i.d. sequence of standard Gaussian variables independent of the past and where: $\alpha_h = \alpha h$; $\beta_h = \beta h$; $\sigma_h = \sigma\sqrt{h}$; h being equal to 1/240, 5/240, or 20/240, representing daily, weekly or monthly frequencies; $\aleph$ is the set of natural numbers. The number of repetitions was 1000.

Table 3.6: Mean Square Error of the Volatility Elasticity Parameters γ, Normal Distribution, Mis-specification Model

γ	(T=2400)			(T=4800)		
Model	Day	Week	Month	Day	Week	Month
$\gamma = 0.0$	0.0003	0.0012	0.0053	0.0001	0.0007	0.0024
$\gamma = 0.5$	0.0038	0.0198	0.0821	0.0020	0.0092	0.0392
$\gamma = 1.0$	0.0423	0.1783	0.5943	0.0215	0.0913	0.3490

Notes: The table reports the mean square error of the interest rate volatility parameter estimates from the CKLS model. Data were generated from the model $r_{t+1}^{(h)} = \alpha_h + (\beta_h + 1)r_t^{(h)} + \sigma_h(\mid r_t^{(h)} \mid^{\gamma})Z_{t+1}, t\epsilon\aleph$,
where Z_t is an i.i.d. sequence of standard Gaussian variables independent of the past and where: $\alpha_h = \alpha h$; $\beta_h = \beta h$; $\sigma_h = \sigma\sqrt{h}$; h being equal to 1/240, 5/240, or 20/240, representing daily, weekly or monthly frequencies; $\aleph$ is the set of natural numbers. The number of repetitions was 1000.

Table 3.7: Bias of the Variances, Normal Distribution, Mis-specification Model, Monte Carlo SE/Asymptotic SE

$\gamma = 0.0$	Bias (T=2400)			Bias (T=4800)		
Param.	Day	Week	Month	Day	Week	Month
α	0.8515	0.8394	0.7962	0.8843	0.8875	0.8600
β	0.7169	0.7056	0.6363	0.7984	0.7921	0.7499
σ	0.9780	0.9598	1.0450	1.0953	0.9877	0.9572
γ	1.0119	0.9747	0.9927	1.1372	0.9929	0.9380
$\gamma = 0.5$	Bias (T=2400)			Bias (T=4800)		
Param.	Day	Week	Month	Day	Week	Month
α	0.7333	0.7206	0.6567	0.8148	0.8123	0.7778
β	0.7184	0.7028	0.6305	0.8031	0.7982	0.7549
σ	1.0287	1.1658	2.1428	1.0386	1.0474	1.2596
γ	0.9904	1.0132	0.9893	1.0229	0.9847	0.9583
$\gamma = 1.0$	Bias (T=2400)			Bias (T=4800)		
Param.	Day	Week	Month	Day	Week	Month
α	0.7577	0.7378	0.6672	0.8112	0.8021	0.7629
β	0.7529	0.7335	0.6626	0.8108	0.8019	0.7609
σ	1.1826	1.4657	0.8185	0.8158	1.7057	1.6130
γ	0.9721	0.8206	0.7567	0.9060	0.7855	0.8814

Notes: The table reports the bias of the variance of each parameter estimates from the CKLS model. Data were generated from the model $r_{t+1}^{(h)} = \alpha_h + (\beta_h + 1)r_t^{(h)} + \sigma_h(| r_t^{(h)} |^{\gamma})Z_{t+1}, t\epsilon\aleph$,
where Z_t is an i.i.d. sequence of standard Gaussian variables independent of the past and where: $\alpha_h = \alpha h$; $\beta_h = \beta h$; $\sigma_h = \sigma\sqrt{h}$; h being equal to 1/240, 5/240, or 20/240, representing daily, weekly or monthly frequencies; $\aleph$ is the set of natural numbers. The number of repetitions was 1000.

Table 3.8: Coverage Probabilities for T=2400, End of Period, Normal Distribution, Mis-specification Model

$\gamma = 0.0$	Day			Week			Month		
Param.	.90	.95	.99	.90	.95	.99	.90	.95	.99
α	0.9420	0.9760	0.9960	0.2680	0.4830	0.8780	0.1060	0.2330	0.6920
β	0.9810	0.9920	1.0000	0.0380	0.1430	0.5550	0.0030	0.0160	0.2080
σ	0.8410	0.9170	0.9850	0.0080	0.0130	0.0320	0.0230	0.0790	0.6270
γ	0.8290	0.9150	0.9820	0.8980	0.9580	0.9910	0.9030	0.9530	0.9880
$\gamma = 0.5$	Day			Week			Month		
Param.	.90	.95	.99	.90	.95	.99	.90	.95	.99
α	0.9720	0.9890	0.9970	0.0260	0.0970	0.4830	0.0030	0.0100	0.1780
β	0.9720	0.9910	0.9970	0.0460	0.1290	0.5570	0.0060	0.0170	0.2280
σ	0.8890	0.9440	0.9790	0.9730	1.0000	1.0000	0.9970	0.9970	0.9970
γ	0.8990	0.9460	0.9890	0.8970	0.9430	0.9920	0.9070	0.9430	0.9780
$\gamma = 1.0$	Day			Week			Month		
Param.	.90	.95	.99	.90	.95	.99	.90	.95	.99
α	0.9630	0.9810	1.0000	0.0580	0.1640	0.5910	0.0050	0.0240	0.2700
β	0.9660	0.9850	1.0000	0.0620	0.1780	0.6130	0.0050	0.0280	0.2870
σ	0.7830	0.8210	0.8690	0.9440	0.9520	0.9620	0.8740	0.8830	0.8940
γ	0.8110	0.8790	0.9550	0.9140	0.9510	0.9810	0.9430	0.9830	0.9880

Notes: The table reports coverage probabilities of the various parameter estimates. For example, β is computed from confidence intervals that are based on the asymptotic distribution of the MLE and take the form $\hat{\beta} \pm z_{\frac{\alpha}{2}}\hat{\sigma}$ where $\hat{\beta}$ is the parameter estimate, $z_{\frac{\alpha}{2}}$ is the $\frac{\alpha}{2}$ critical point from the normal distribution, and $\hat{\sigma}$ is the square root of the variance estimate. The approximate probability that the true parameter falls within the interval (nominal coverage) is equal to 1 - α. The number of repetitions was 1000.

Table 3.9: Coverage Probabilities for T=4800, End of Period, Normal Distribution, Mis-specification Model

$\gamma = 0.0$	Day			Week			Month		
Param.	.90	.95	.99	.90	.95	.99	.90	.95	.99
α	0.9370	0.9750	0.9970	0.0250	0.0810	0.3040	0.0030	0.0050	0.0730
β	0.9680	0.9890	0.9980	0.0000	0.0020	0.0190	0.0000	0.0010	0.0010
σ	0.7820	0.8670	0.9480	0.0080	0.0110	0.0220	0.0080	0.0100	0.0350
γ	0.7260	0.8130	0.9460	0.8800	0.9320	0.9820	0.9120	0.9470	0.9890
$\gamma = 0.5$	Day			Week			Month		
Param.	.90	.95	.99	.90	.95	.99	.90	.95	.99
α	0.9620	0.9890	0.9990	0.0000	0.0000	0.0140	0.0000	0.0000	0.0000
β	0.9680	0.9890	1.0000	0.0000	0.0000	0.0250	0.0000	0.0000	0.0010
σ	0.8830	0.9390	0.9760	0.3780	0.8230	1.0000	1.0000	1.0000	1.0000
γ	0.8970	0.9420	0.9880	0.9020	0.9580	0.9920	0.9090	0.9590	0.9900
$\gamma = 1.0$	Day			Week			Month		
Param.	.90	.95	.99	.90	.95	.99	.90	.95	.99
α	0.9570	0.9830	0.9980	0.0000	0.0010	0.0330	0.0000	0.0000	0.0010
β	0.9520	0.9820	0.9970	0.0000	0.0030	0.0360	0.0000	0.0000	0.0010
σ	0.7710	0.8210	0.8730	0.9790	0.9820	0.9880	0.9390	0.9420	0.9480
γ	0.7590	0.8280	0.9300	0.8990	0.9340	0.9770	0.9040	0.9510	0.9820

Notes: The table reports coverage probabilities of the various parameter estimates. For example, β is computed from confidence intervals that are based on the asymptotic distribution of the MLE and take the form $\hat{\beta} \pm z_{\frac{\alpha}{2}}\hat{\sigma}$ where $\hat{\beta}$ is the parameter estimate, $z_{\frac{\alpha}{2}}$ is the $\frac{\alpha}{2}$ critical point from the normal distribution, and $\hat{\sigma}$ is the square root of the variance estimate. The approximate probability that the true parameter falls within the interval (nominal coverage) is equal to 1 - α. The number of repetitions was 1000.

Table 3.10: Descriptive Statistics, Ten-year Simulated Data, n=1000, Normal Distribution, Correct-Specification Model

		$\gamma = 0.0$			$\gamma = 0.5$			$\gamma = 1.0$		
		Day	Week	Month	Day	Week	Month	Day	Week	Month
$\hat{x}$	α	0.2394	0.2354	0.2222	0.2395	0.2353	0.2206	0.2289	0.2251	0.2111
$\hat{x}$	β	-2.4068	-2.3649	-2.2318	-2.4157	-2.3721	-2.2232	-2.2907	-2.2521	-2.1118
$\hat{x}$	σ	0.0226	0.0222	0.0213	0.0235	0.0279	0.0570	0.0394	0.2760	2.7068
$\hat{x}$	γ	-0.0002	0.0015	0.0078	0.4992	0.5014	0.4708	1.0091	1.0643	1.1650
SD	α	0.0940	0.0929	0.0896	0.0706	0.0694	0.0679	0.0707	0.0696	0.0651
SD	β	0.7231	0.7094	0.6783	0.7138	0.7013	0.6833	0.7118	0.7005	0.6551
SD	σ	0.0022	0.0045	0.0096	0.0078	0.0220	0.1875	0.0486	2.6741	12.6045
SD	γ	0.0183	0.0373	0.0800	0.0677	0.1503	0.3168	0.2217	0.4267	0.6945
Min	α	0.0249	0.0270	0.0291	0.0796	0.0937	0.0680	0.0402	0.0417	0.0456
Min	β	-5.3502	-5.5750	-5.5040	-5.7363	-5.7571	-5.5588	-5.9566	-5.8340	-5.1924
Min	σ	0.0159	0.0121	0.0035	0.0087	0.0033	0.0001	0.0016	0.0001	0.0001
Min	γ	-0.0577	-0.1227	-0.3328	0.2990	0.0834	-0.5198	0.4241	-0.1279	-0.1316
Max	α	0.6807	0.6100	0.8099	0.6192	0.6217	0.6094	0.5948	0.5964	0.5336
Max	β	-0.8478	-0.9426	-0.6610	-0.8219	-0.9269	-0.6776	-0.3547	-0.3732	-0.4111
Max	σ	0.0304	0.0470	0.0905	0.0608	0.1774	3.1643	0.4472	81.6230	199.994
Max	γ	0.0591	0.1743	0.3904	0.7147	0.9270	1.6371	1.6566	2.8142	2.9986

Notes: The table reports the Monte-Carlo averages, standard deviations, minimum and maximum of the parameter estimates from the CKLS model. Data were generated from the model $r_{t+1}^{(h)} = \alpha_h + (\beta_h + 1)r_t^{(h)} + \sigma_h(|\, r_t^{(h)}\,|^{\gamma})Z_{t+1}, t\epsilon\aleph$,
where Z_t is an i.i.d. sequence of standard Gaussian variables independent of the past and where: $\alpha_h = \alpha h$; $\beta_h = \beta h$; $\sigma_h = \sigma\sqrt{h}$; h being equal to 1, 5, or 20, representing daily, weekly or monthly frequencies; $\aleph$ is the set of natural numbers. The number of repetitions was 1000.

Table 3.11: Descriptive Statistics, Twenty-year Simulated Data, n=1000, Normal Distribution, Correct-Specification Model

		$\gamma = 0.0$			$\gamma = 0.5$			$\gamma = 1.0$		
		Day	Week	Month	Day	Week	Month	Day	Week	Month
$\hat{x}$	α	0.2182	0.2150	0.2029	0.2167	0.2132	0.2003	0.2125	0.2092	0.1962
$\hat{x}$	β	-2.1960	-2.1631	-2.0409	-2.1789	-2.1439	-2.0130	-2.1236	-2.0906	-1.9606
$\hat{x}$	σ	0.0225	0.0221	0.0204	0.0230	0.0238	0.0245	0.0286	0.0596	0.7080
$\hat{x}$	γ	-0.0003	0.0011	0.0053	0.4993	0.4971	0.4613	0.9980	1.0051	1.0209
SD	α	0.0598	0.0593	0.0555	0.0460	0.0452	0.0427	0.0472	0.0470	0.0434
SD	β	0.4784	0.4704	0.4422	0.4750	0.4661	0.4373	0.4749	0.4721	0.4362
SD	σ	0.0015	0.0030	0.0057	0.0052	0.0115	0.0266	0.0239	0.1476	5.0575
SD	γ	0.0125	0.0254	0.0515	0.0470	0.0965	0.1975	0.1515	0.2971	0.5495
Min	α	0.0588	0.0633	0.0524	0.1152	0.1177	0.1042	0.1038	0.0992	0.1022
Min	β	-4.2155	-4.1525	-3.9515	-4.2062	-4.0380	-3.9771	-4.2110	-4.3170	-3.8547
Min	σ	0.0175	0.0143	0.0047	0.0082	0.0058	0.0004	0.0011	0.0006	0.0001
Min	γ	-0.0448	-0.0709	-0.2571	0.2892	0.2040	-0.2703	0.3500	0.2057	-0.1312
Max	α	0.4759	0.4669	0.4379	0.4316	0.4137	0.4049	0.4274	0.4193	0.3743
Max	β	-1.0151	-0.9944	-1.0600	-1.1703	-1.1557	-1.0503	-1.0269	-0.9863	-1.0104
Max	σ	0.0376	0.0332	0.0523	0.0475	0.1005	0.2098	0.2298	2.4820	132.454
Max	γ	0.1124	0.1029	0.1900	0.6661	0.8367	1.0157	1.5095	2.0250	2.9366

Notes: The table reports the Monte-Carlo averages, standard deviations, minimum and maximum of the parameter estimates from the CKLS model. Data were generated from the model $r_{t+1}^{(h)} = \alpha_h + (\beta_h + 1)r_t^{(h)} + \sigma_h(| r_t^{(h)} |^\gamma)Z_{t+1}, t\epsilon\aleph$,
where Z_t is an i.i.d. sequence of standard Gaussian variables independent of the past and where: $\alpha_h = \alpha h$; $\beta_h = \beta h$; $\sigma_h = \sigma\sqrt{h}$; h being equal to 1/240, 5/240, or 20/240, representing daily, weekly or monthly frequencies; $\aleph$ is the set of natural numbers. The number of repetitions was 1000.

Table 3.12: Bias of the Mean Reversion Parameters β, Normal Distribution, Correct Specification Model

β	Bias (T=2400)			Bias (T=4800)		
Model	Day	Week	Month	Day	Week	Month
$\gamma = 0.0$	-0.4068	-0.3649	-0.2318	-0.1960	-0.1631	-0.0409
$\gamma = 0.5$	-0.4157	-0.3721	-0.2232	-0.1789	-0.1439	-0.0130
$\gamma = 1.0$	-0.2907	-0.2521	-0.1118	-0.1236	-0.0906	0.0394

Notes: The table reports the biases of the mean reversion parameter estimates from the CKLS model. Data were generated from the model $r_{t+1}^{(h)} = \alpha_h + (\beta_h + 1)r_t^{(h)} + \sigma_h(|$ $r_t^{(h)}$ $|^\gamma)Z_{t+1}, t\epsilon\aleph$,
where Z_t is an i.i.d. sequence of standard Gaussian variables independent of the past and where: $\alpha_h = \alpha h$; $\beta_h = \beta h$; $\sigma_h = \sigma\sqrt{h}$; h being equal to 1/240, 5/240, or 20/240, representing daily, weekly or monthly frequencies; $\aleph$ is the set of natural numbers. The number of repetitions was 1000.

Table 3.13: Bias of the Interest Rate Volatility Elasticity Parameters γ, Normal Distribution, Correct Specification Model

γ	Bias (T=2400)			Bias (T=4800)		
Model	Day	Week	Month	Day	Week	Month
$\gamma = 0.0$	-0.0002	0.0015	0.0078	-0.0003	0.0011	0.0053
$\gamma = 0.5$	-0.0008	0.0014	0.0292	0.0007	0.0029	0.0387
$\gamma = 1.0$	0.0090	0.0643	0.1650	0.0020	0.0051	0.0209

Notes: The table reports the biases of the interest rate volatility parameter estimates from the CKLS model. Data were generated from the model $r_{t+1}^{(h)} = \alpha_h + (\beta_h + 1)r_t^{(h)} +$ $\sigma_h(|$ $r_t^{(h)}$ $|^\gamma)Z_{t+1}, t\epsilon\aleph$,
where Z_t is an i.i.d. sequence of standard Gaussian variables independent of the past and where: $\alpha_h = \alpha h$; $\beta_h = \beta h$; $\sigma_h = \sigma\sqrt{h}$; h being equal to 1/240, 5/240, or 20/240, representing daily, weekly or monthly frequencies; $\aleph$ is the set of natural numbers. The number of repetitions was 1000.

Table 3.14: Mean Square Error of the Mean Reversion Parameters β, Normal Distribution, Correct Specification Model

β	(T=2400)			(T=4800)		
Model	Day	Week	Month	Day	Week	Month
$\gamma = 0.0$	0.6883	0.6364	0.5139	0.2673	0.2478	0.1972
$\gamma = 0.5$	0.6823	0.6303	0.5168	0.2576	0.2380	0.1914
$\gamma = 1.0$	0.5911	0.5543	0.4416	0.2408	0.2311	0.1918

Notes: The table reports the mean square error of the mean reversion parameter estimates from the CKLS model. Data were generated from the model $r_{t+1}^{(h)} = \alpha_h + (\beta_h + 1)r_t^{(h)} + \sigma_h(|\, r_t^{(h)}\, |^{\gamma})Z_{t+1}, t\epsilon\aleph$,
where Z_t is an i.i.d. sequence of standard Gaussian variables independent of the past and where: $\alpha_h = \alpha h$; $\beta_h = \beta h$; $\sigma_h = \sigma\sqrt{h}$; h being equal to 1/240, 5/240, or 20/240, representing daily, weekly or monthly frequencies; $\aleph$ is the set of natural numbers. The number of repetitions was 1000.

Table 3.15: Mean Square Error of the Volatility Elasticity Parameters γ, Normal Distribution, Correct Specification Model

γ	(T=2400)			(T=4800)		
Model	Day	Week	Month	Day	Week	Month
$\gamma = 0.0$	0.0003	0.0014	0.0065	0.0002	0.0006	0.0027
$\gamma = 0.5$	0.0046	0.0226	0.1012	0.0022	0.0093	0.0405
$\gamma = 1.0$	0.0492	0.1862	0.5095	0.0230	0.0883	0.3023

Notes: The table reports the mean square errors of the volatility elasticity parameter estimates from the CKLS model. Data were generated from the model $r_{t+1}^{(h)} = \alpha_h + (\beta_h + 1)r_t^{(h)} + \sigma_h(|\, r_t^{(h)}\, |^{\gamma})Z_{t+1}, t\epsilon\aleph$,
where Z_t is an i.i.d. sequence of standard Gaussian variables independent of the past and where: $\alpha_h = \alpha h$; $\beta_h = \beta h$; $\sigma_h = \sigma\sqrt{h}$; h being equal to 1/240, 5/240, or 20/240, representing daily, weekly or monthly frequencies; $\aleph$ is the set of natural numbers. The number of repetitions was 1000.

Table 3.16: Bias of the Variances, Normal Distribution, Correct Specification Model, Monte Carlo SE/Asymptotic SE, $\gamma = 0.0$

$\gamma = 0.0$	Bias (T=2400)			Bias (T=4800)		
Param.	Day	Week	Month	Day	Week	Month
α	1.1108	1.1226	1.1555	1.0311	1.0458	1.0414
β	1.0384	1.0407	1.0549	1.0158	1.0206	1.0167
σ	1.0026	0.9853	1.1017	1.1912	0.9577	0.9982
γ	1.0428	0.9770	1.0457	1.2228	0.9541	0.9929
$\gamma = 0.5$	Bias (T=2400)			Bias (T=4800)		
Param.	Day	Week	Month	Day	Week	Month
α	1.0489	1.0458	1.0740	1.0175	1.0139	1.0050
β	1.0284	1.0246	1.0501	1.0184	1.0140	1.0027
σ	1.0403	1.1084	2.0992	1.0640	1.0276	1.1229
γ	0.9970	1.0042	1.0141	1.0394	0.9657	0.9571
$\gamma = 1.0$	Bias (T=2400)			Bias (T=4800)		
Param.	Day	Week	Month	Day	Week	Month
α	1.0429	1.0443	1.0262	1.0230	1.0360	1.0098
β	1.0416	1.0420	1.0228	1.0212	1.0338	1.0068
σ	1.1939	3.9871	1.1302	0.8986	1.4470	2.6932
γ	0.9999	0.8124	0.6647	0.9178	0.8352	0.8020

Notes: The table reports the bias of the variance of each parameter estimates from the CKLS model. Data were generated from the model $r_{t+1}^{(h)} = \alpha_h + (\beta_h + 1)r_t^{(h)} + \sigma_h(|r_t^{(h)}|^{\gamma})Z_{t+1}, t\epsilon\aleph$,
where Z_t is an i.i.d. sequence of standard Gaussian variables independent of the past and where: $\alpha_h = \alpha h$; $\beta_h = \beta h$; $\sigma_h = \sigma\sqrt{h}$; h being equal to 1/240, 5/240, or 20/240, representing daily, weekly or monthly frequencies; $\aleph$ is the set of natural numbers. The number of repetitions was 1000.

Table 3.17: Coverage Probabilities for T=2400, End of Period, Normal Distribution, Correct Specification Model

$\gamma = 0.0$	Day			Week			Month		
Param.	.90	.95	.99	.90	.95	.99	.90	.95	.99
α	0.8690	0.9380	0.9900	0.8700	0.9390	0.9880	0.8630	0.9200	0.9860
β	0.8820	0.9450	0.9910	0.8850	0.9400	0.9910	0.8860	0.9450	0.9900
σ	0.8280	0.9110	0.9810	0.8800	0.9400	0.9750	0.8200	0.8700	0.9220
γ	0.8190	0.9030	0.9840	0.8990	0.9500	0.9860	0.9000	0.9440	0.9860
$\gamma = 0.5$	Day			Week			Month		
Param.	.90	.95	.99	.90	.95	.99	.90	.95	.99
α	0.8800	0.9440	0.9870	0.8890	0.9450	0.9900	0.8990	0.9450	0.9870
β	0.8840	0.9410	0.9880	0.8960	0.9450	0.9930	0.9090	0.9540	0.9840
σ	0.8900	0.9390	0.9710	0.8500	0.8790	0.9300	0.7340	0.7640	0.8140
γ	0.8910	0.9440	0.9900	0.8950	0.9580	0.9940	0.8930	0.9310	0.9780
$\gamma = 1.0$	Day			Week			Month		
Param.	.90	.95	.99	.90	.95	.99	.90	.95	.99
α	0.8960	0.9490	0.9940	0.8980	0.9490	0.9920	0.8990	0.9570	0.9900
β	0.8970	0.9500	0.9940	0.8950	0.9500	0.9920	0.9080	0.9580	0.9900
σ	0.7850	0.8180	0.8630	0.8210	0.8340	0.8620	0.7730	0.7920	0.8120
γ	0.7970	0.8630	0.9440	0.9270	0.9670	0.9850	0.9430	0.9650	0.9780

Notes: The table reports the coverage probabilities of the various parameter estimates. For example, β is computed from confidence intervals that are based on the asymptotic distribution of the MLE and take the form $\hat{\beta} \pm z_{\frac{\alpha}{2}}\hat{\sigma}$ where $\hat{\beta}$ is the parameter estimate, $z_{\frac{\alpha}{2}}$ is the $\frac{\alpha}{2}$ critical point from the normal distribution, and $\hat{\sigma}$ is the square root of the variance estimate. The approximate probability that the true parameter falls within the interval (nominal coverage) is equal to 1 - α. The number of repetitions was 1000.

Figure 3.1: Box Plot Graph of Weekly Data of Mean Reversion Parameters, β, of both Correct and Mis-specified Models, Ten-Year Simulated Data

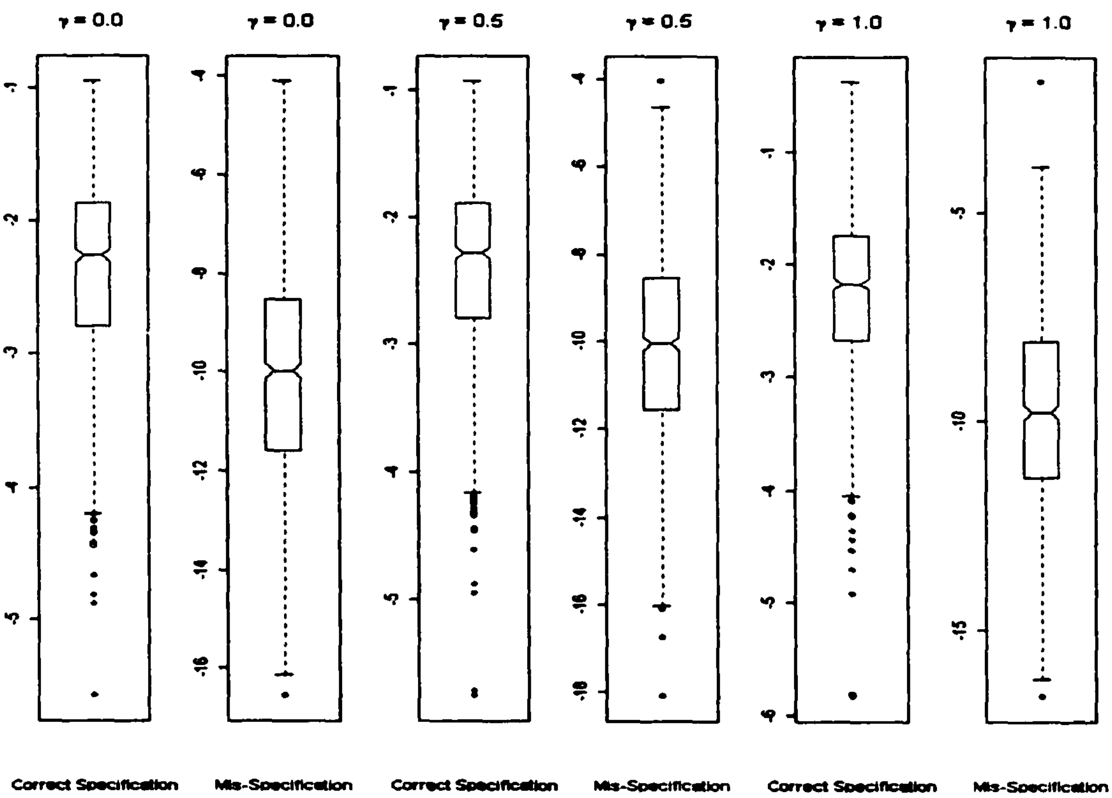


Table 3.18: Coverage Probabilities for T=4800, End of Period, Normal Distribution, Correct Specification Model

$\gamma = 0.0$	Day			Week			Month		
Param.	.90	.95	.99	.90	.95	.99	.90	.95	.99
α	0.8850	0.9490	0.9920	0.8780	0.9470	0.9890	0.8830	0.9460	0.9910
β	0.8990	0.9450	0.9900	0.9030	0.9470	0.9900	0.9000	0.9560	0.9880
σ	0.7660	0.8520	0.9490	0.8610	0.9210	0.9750	0.7920	0.8550	0.9350
γ	0.7120	0.7970	0.9440	0.8880	0.9400	0.9860	0.9010	0.9450	0.9880
$\gamma = 0.5$	Day			Week			Month		
Param.	.90	.95	.99	.90	.95	.99	.90	.95	.99
α	0.9010	0.9510	0.9910	0.9050	0.9550	0.9910	0.9140	0.9590	0.9870
β	0.9100	0.9470	0.9900	0.9100	0.9550	0.9890	0.9120	0.9530	0.9910
σ	0.8820	0.9350	0.9760	0.8870	0.9040	0.9430	0.7390	0.7720	0.8300
γ	0.8880	0.9390	0.9860	0.9140	0.9550	0.9890	0.9190	0.9700	0.9920
$\gamma = 1.0$	Day			Week			Month		
Param.	.90	.95	.99	.90	.95	.99	.90	.95	.99
α	0.9010	0.9480	0.9910	0.8980	0.9470	0.9910	0.8860	0.9500	0.9850
β	0.8990	0.9470	0.9920	0.9030	0.9480	0.9910	0.8920	0.9460	0.9870
σ	0.7550	0.8120	0.8670	0.8170	0.8380	0.8770	0.7310	0.7490	0.7810
γ	0.7480	0.8270	0.9260	0.9200	0.9610	0.9860	0.9320	0.9560	0.9800

Notes: The table reports the coverage probabilities of the various parameter estimates. For example, β is computed from confidence intervals that are based on the asymptotic distribution of the MLE and take the form $\hat{\beta} \pm z_{\frac{\alpha}{2}}\hat{\sigma}$ where $\hat{\beta}$ is the parameter estimate, $z_{\frac{\alpha}{2}}$ is the $\frac{\alpha}{2}$ critical point from the normal distribution, and $\hat{\sigma}$ is the square root of the variance estimate. The approximate probability that the true parameter falls within the interval (nominal coverage) is equal to 1 - α. The number of repetitions was 1000.

Figure 3.2: Box Plot Graph of Monthly Data of Mean Reversion Parameters, β, of both Correct and Mis-specified Models, Ten-Year Simulated Data

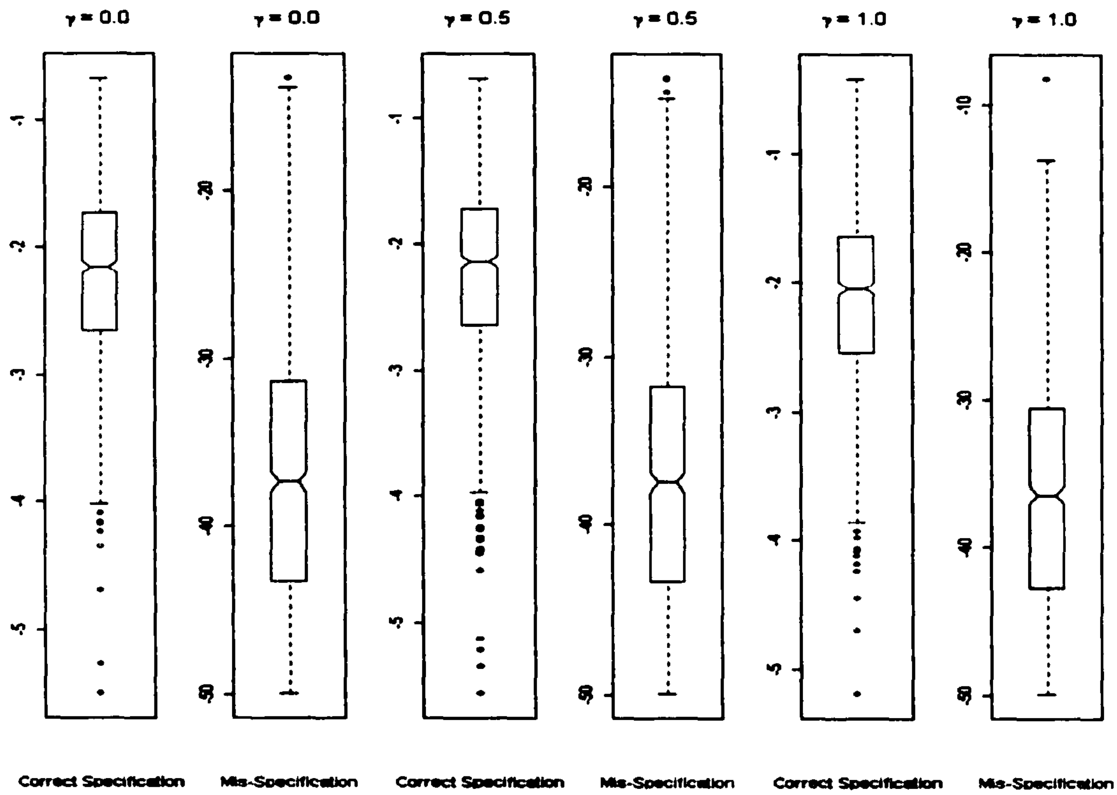


Figure 3.3: Box Plot Graph of Weekly Data of Mean Reversion Parameters, β, of both Correct and Mis-specified Models, Twenty-Year Simulated Data

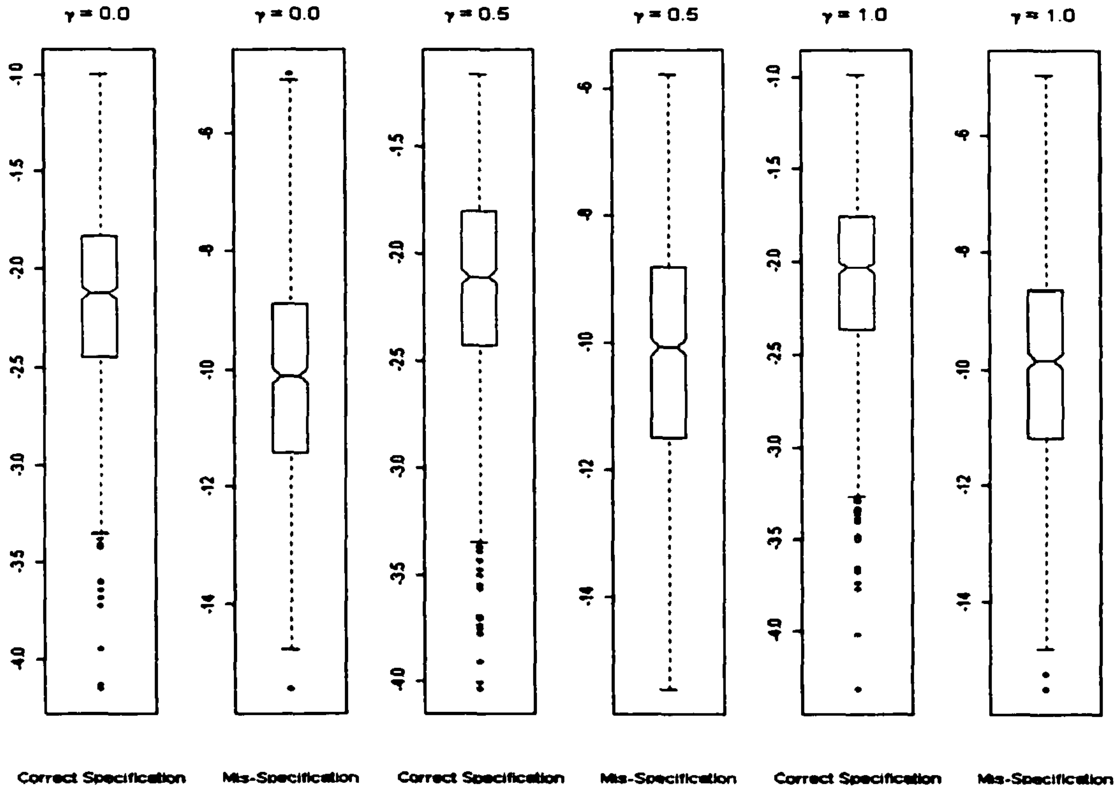
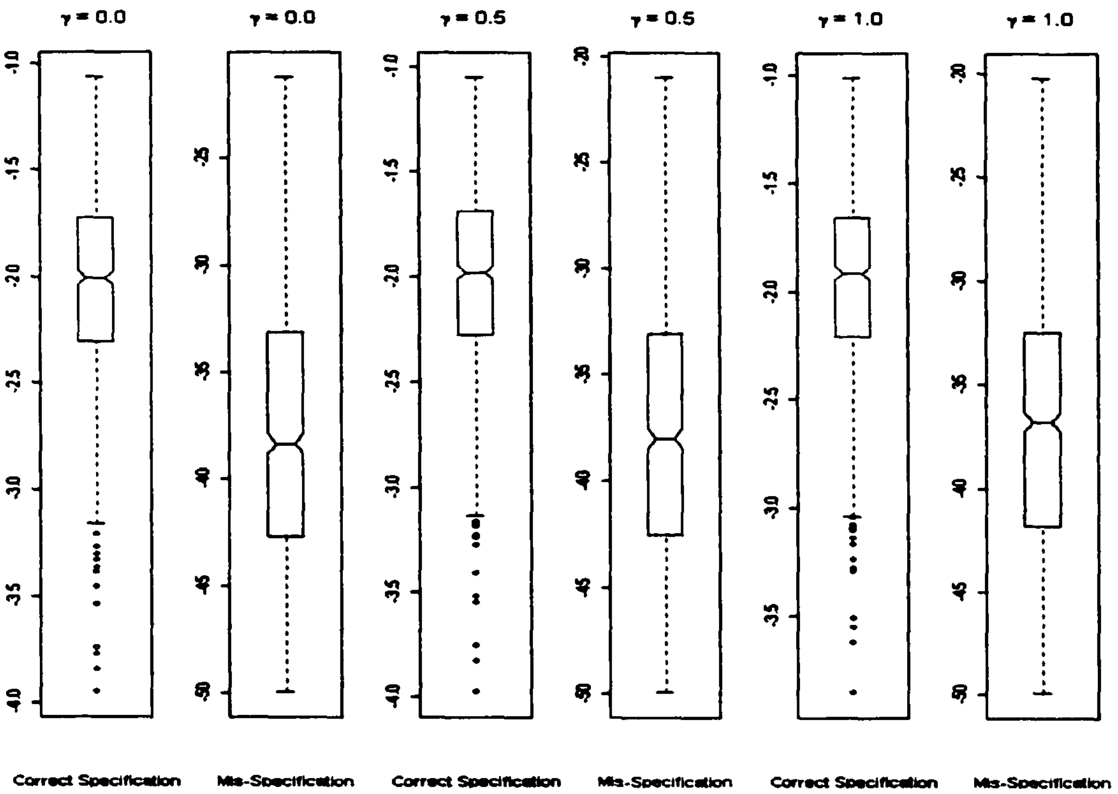


Figure 3.4: Box Plot Graph of Monthly Data of Mean Reversion Parameters, β, of both Correct and Mis-specified Models, Twenty-Year Simulated Data



Chapter 4

Empirical Evidence

I. Introduction

we have discussed extensively in the last chapter that interest rate parameter estimation can behave dramatically different if the method of maximum likelihood estimation is not properly used. The Monte-Carlo simulation result supports the claim that sampling frequency plays a critical role in estimating these parameters. Implementing the mis-specified model of maximum likelihood estimation leads to a significant increase in the magnitude of bias as the level of aggregation varies. These changes in the data aggregation can lead to strong distortion regardless of the interest rate model implemented. The further apart the sampling frequency, the higher the distortion becomes. As a result, statistical inference can be misleading if drawn from an improper usage of the maximum likelihood function.

In this chapter, we examine the magnitude of the biases associated with the use of actual data to ratify earlier Monte-Carlo hypothesis. To be specific, we implement the maximum likelihood estimation that accounts for the sampling interval and compare them to those that do not take into account the data aggregation. In this chapter, we report findings using empirical data to validate the Monte-Carlo simulation. As shown earlier, aggregation of data does affect a small sample behavior of the maximum likelihood estimators.

II. Data Sources

Data used in this dissertation is the Euro-Dollar Deposit Mid rate, starting from November 25, 1981 to December 31, 1989. The Euro-Dollar rate is based on the average of the bid and ask prices for the Euro-Dollar deposit interest rates. Interestingly, this data set has been used in Ait-Sahalia's (1996) non parametric estimation article. To make the data consistent with the Monte-Carlo simulation method, we use the available daily data information and skip that data to attain weekly data, and monthly data. For weekly data information, we use every Wednesday of the week. In a case that a Wednesday should fall on a holiday, Thursday will be used instead, and if both Wednesday and Thursday of the same week are holidays, then Tuesday of the same week will be selected as our data. Likewise, monthly data is chosen using the last Wednesday of every month. Again, if that Wednesday is a holiday, then Thursday will be picked. If both Wednesday and Thursday are holidays, Tuesday of the same week will be selected. The sample size for daily rate is 2046 observations, for the weekly rate it is 422 observations, and for the monthly rate it is 96 observations. All data are annualized percentage rates. A time series plot of the data is provided in Figure 3.1. The use of a long sample is particularly important in this context in order to facilitate the identification of the mean and volatility dynamics.

III. Empirical Methodology Setup

The stochastic differential equation of interest rate dynamic, as described in equation 4.1, is used here as our general framework, as follows,

$$dr(t) = [\alpha + \beta r(t)]dt + \sigma r(t)^{\gamma} dW(t), \forall t \in [0, T] \tag{4.1}$$

and which it is discretized into the following process,[1]

$$r_{t+1}^{(h)} = \alpha_h + (\beta_h + 1) r_t^{(h)} + \sigma_h(| r_t^{(h)} |^{\gamma})\epsilon_{t+1}, t = [0, 1, 2, ..., T] \tag{4.2}$$

where h is the sampling interval, $\alpha_h = \alpha h$, $\beta_h = \beta h$, $\sigma_h = \sigma\sqrt{h}$, and ϵ is i.i.d. sequence of standard Gaussian variables independent of the past time processes.

[1] For the derivation, please refer to chapter 3 from equation 3.1 to equation 3.5

The log-likelihood function that was introduced in the last chapter is re-introduced again here.

$$\max_{\theta^{(h)}} \mathcal{L}_T(r^{(h)};\theta^{(h)}) = \sum_{t=1}^{T} log f_t$$

$$\max_{\theta^{(h)}} \mathcal{L}_T(r^{(h)};\theta^{(h)}) = -\frac{T}{2} log(\sigma^2_{0,h}) - \frac{1}{2}\left[\sum_{t=1}^{T} log(|\, r^{(h)}_{t-1}\,|)^{2\gamma} + \sum_{t=1}^{T} \frac{(r^{(h)}_t - \alpha_h - (\beta_h+1) r^{(h)}_{t-1})^2}{\sigma^2_{0,h}(|\, r^{(h)}_{t-1}\,|)^{2\gamma}}\right] \tag{4.3}$$

IV. Empirical Tests on Term Structure of Interest Rate

Table 4.1 shows summary statistics where the average, standard deviation, variance, minimum, and maximum are reported. The averages of the interest rate are at 0.0905, 0.0904, and 0.0910, for daily, weekly, and monthly data, respectively. The standard deviations are recorded at 0.0218, 0.0217, and 0.0216. The minimum are recorded at 0.0589, 0.0596, and 0.0596, for the daily return, weekly return, and the monthly return. The maximum are recorded at 0.1679, 0.1679, and 0.1572 for all three cases.

Table 4.2 reports the parameter estimates of the CKLS discretized model as described in the equation 4.2. Interestingly, it shows consistent evidence supporting the Monte-Carlo simulation. The mean reversion parameter estimates, $\hat{\beta}$, are approximately equal and consistent at -1.0720, -0.9490, and -1.0540 for daily, weekly, and monthly data using correct specification of the maximum likelihood estimation. Although weekly parameter of $\hat{\beta}$ is slightly off, this is due to what we call a 'holiday effect'. As a matter of fact, the counting process is a big contributor to this over-estimation in weekly parameter estimation. There is one important notion, however. Note that both drift parameters α and β estimated here are statistically significant at 5 percent level.[2] Asymptotic t-statistics of the mean reversion parameter, $\hat{\beta}$, reject the null hypothesis

[2]Other findings (Engle and Lee (1995), Koutmos (1998), amongst others) show no statistical evidence of the reversion of the interest rate to its mean level in every cases, namely, day, week, and month intervals. This implies that the economic perspective of the mean reverting process is true and is supported strongly by our findings.

of no mean reverting. Our findings lend strong support to the fact that there is a strong tendency for interest rates to revert to its long-run mean. On the other hand, the estimated mean reversion parameter, $\hat{\beta}$, using the mis-specification of the maximum likelihood are reported at -4.7452, and -21.0790 for weekly and monthly data, respectively. The mis-specification model of the maximum likelihood under-estimates the coefficients by more than four times in weekly process and 20 times in monthly estimation. This ratifies supporting evidence of bias which we have seen earlier in the simulation results.[3] The magnitude of bias in mean reversion parameters decreases strongly if the proper specification of the maximum likelihood function is employed.

The standard errors of the mean reversion parameter are 0.5391, 0.5360, and 0.5430 in all three cases where the correct specification is used. With proper maximum likelihood estimation, the standard errors are in the same range in all cases. The standard errors of the mean reversion parameter, β, increase more than five times when the mis-specification of the maximum likelihood is used on weekly data and increase more than twenty times on monthly data. They are reported at 0.5391, 2.6746, and 10.8489, respectively. This shows that standard errors become extremely explosive as the mis-specified models are used.

The volatility elasticity parameter estimates, γ, are at 1.5696, 1.7819, and 1.4060 for daily, weekly, and monthly data for both correct and mis-specified maximum likelihood estimations. All of these parameters estimated are asymptotically significant at 1 percent level. The parameter estimates here also coincide with an early simulation result. As observed, the size of the volatility elasticity are not sensitive to the sampling frequencies. The estimator biases are at an insignificant level, regardless of the interest rate models used. Interestingly, the ranges of the elasticity parameter estimates are closely resemble those of the CKLS' findings (1992) in his unrestricted model.

The findings show evidence that the parameter estimates should yield approximately equal to one another regardless of the sampling frequencies used. From the monte-carlo simulation and the study on empirical data, it is certain that when the maximum likelihood function is correctly specified, the estimates of both mean rever-

[3] As shown in the simulation chapter, the sample means of the mean reversion parameter, $\hat{\beta}$, increases from approximately -2.000 to -10.000, and to -37.000 as the frequency intervals change from day, to week, and to month returns. Please refer to chapter four for more information

sion, and the volatility elasticity will be independent of any sampling intervals and the estimated bias can be significantly reduced. The next step that should also pursue is to estimate the short term interest rate using some other empirical data such as the US federal fund rate or the LIBOR rate.

Table 4.1: Descriptive Statistics of Euro Dollar Deposit Mid Rates.

Euro Dollar Rates	Obs.	Mean	Std. Dev.	Variance	Minimum	Maximum
Daily Rates	2049	0.0905	0.0218	0.0005	0.0589	0.1679
Weekly Rates	422	0.0904	0.0217	0.0005	0.0596	0.1679
Monthly Rates	96	0.0910	0.0216	0.0005	0.0596	0.1572

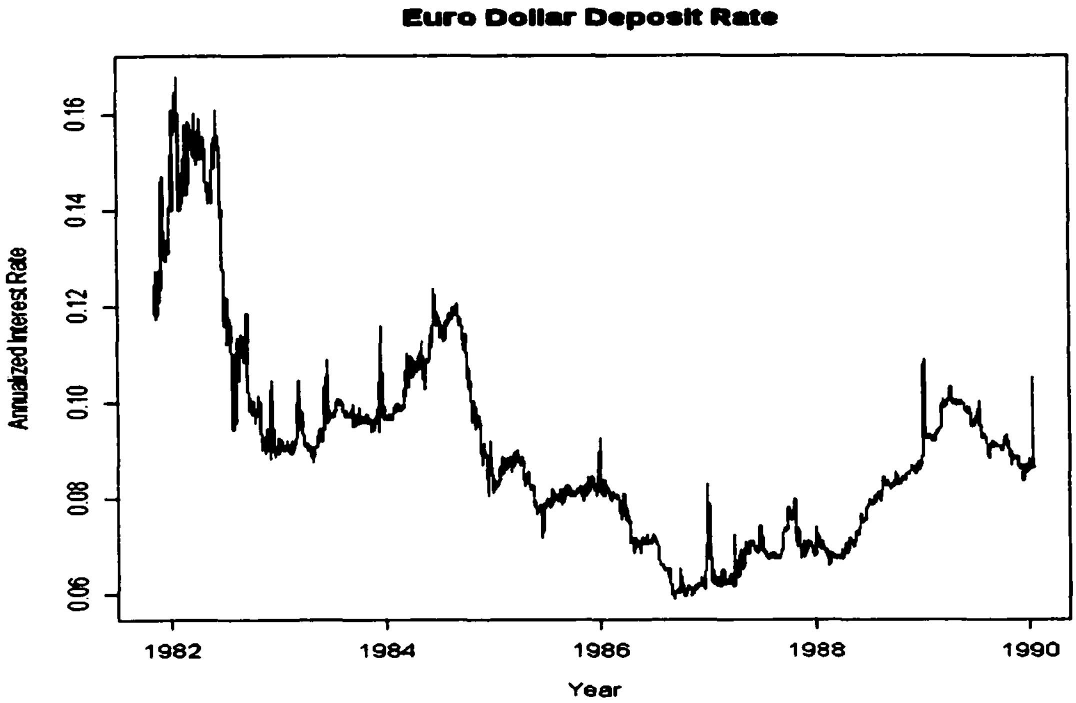
Notes: The data set extends from October 25, 1981 through December 31, 1989, for a total of 2046 observations in the daily, 422 observations in the weekly, and 96 observations in the monthly rates. All rates are annualized percentage returns.

Table 4.2: Comparison of Correct Specification VS. Mis-specification for the Short-Term Interest Rate.

Specification	α	β	σ	γ
CORRECT SPECIFICATION - DAY	0.0932	-1.0720	1.5589	1.5696
STANDARD ERRORS - DAY	0.0434	0.5391	0.7227	0.0949
CORRECT SPECIFICATION - WEEK	0.0827	-0.9490	3.9760	1.7819
STANDARD ERRORS - WEEK	0.0425	0.5360	2.9502	0.1520
MIS-SPECIFICATION - WEEK	0.4136	-4.7452	19.8798	1.7819
STANDARD ERRORS - WEEK	0.2119	2.6746	13.5731	0.1399
CORRECT SPECIFICATION - MONTH	0.0919	-1.0540	0.6309	1.4060
STANDARD ERRORS - MONTH	0.0445	0.5429	0.8749	0.2847
MIS-SPECIFICATION - MONTH	1.8389	-21.0791	12.6177	1.4060
STANDARD ERRORS - MONTH	0.8888	10.8489	17.7463	0.2887

Notes: The parameters are estimated by using correct specification and mis-specification models for comparison. The standard errors are reported below the parameters estimates.

Figure 4.1: Daily Euro Dollar Deposit Rate, from Nov. 25, 81 to Dec. 31, 89



Chapter 5

Conclusion

This dissertation shows that it is very crucial to use a proper maximum likelihood specification, that incorporates information about the sampling interval in studies based on different sampling frequencies. Results from both the Monte-Carlo simulation and the empirical estimation chapters show strong supporting evidence of biases and misrepresented parameters. Without proper adjustment for the sampling frequency in the likelihood specification, one faces the risk of estimation bias and misleading statistical inferences.

When implementing the correct specification of the maximum likelihood estimation, both Monte-Carlo simulation process and empirical estimation show conformed results. Here, we show some important findings, as follows:

1. The absolute value of the mean reversion parameter, β, does not vary with the level of aggregation.

2. The absolute value of the volatility elasticity parameter, γ, varies slightly with the level of aggregation.

3. Without proper usage of the maximum likelihood estimation, explosive standard errors are evidenced throughout (as shown both in the simulation process and in the empirical estimation).

4. The bias of the mean reversion parameter, β, increases more than four times in case of weekly data, and increases more than twenty times in case of monthly data if one does not exercise proper likelihood specification.

5. The volatility elasticity, γ, shows almost undetectable level of bias. This happens on both simulation and empirical tests.

6. The estimators considered tend to overstate the variance of the maximum likelihood estimator in finite sample.

Results of the Monte-Carlo simulation and the empirical tests on the Euro-Dollar Deposit rate indicate that it is very critical to use correct specification of the maximum likelihood estimation when examining the data with different frequencies. Our results show strong support that it is possible to estimate the true value of the mean reversion parameter, β, no matter which level of aggregation of data is used.

It is important to remember that the results obtained and the conclusion that is drawn from this research refer only to particular CKLS models selected for the Monte-Carlo study. Nevertheless, we feel confident that the conclusion reached will extend to other models, as long as the data generating process takes into account the aggregation of data and follows the skipped sampling mentioned above.

Bibliography

[Adkins 99] Adkins, L. C., T. Krehbiel, 1999, Mean Reversion and Volatility of Short-term London Interbank Offer Rates: An empirical comparison of competing models, *International Review of Economics and Finance*, 8, 45-54.

[Agbeyegbe 83] Agbeyegbe, T. D., 1983, Some Topics in Continuous-time Econometrics, Department of Economics, University of Essex, Unpublished Ph.D. dissertation.

[Agbeyegbe 87] Agbeyegbe, T. D., 1987, The exact discrete analog to a closed linear first-order system with mixed sample, *Econometric Theory*, 3, 142-9.

[Agbeyegbe 88] Agbeyegbe, T. D., 1988, An exact discrete analog of an open order linear nonstationary first-order continuous-time system with mixed sample, *Journal of Econometrics*, 39, 237-250.

[Agbeyegbe 00] Agbeyegbe, T. D, and T. Tanarugsachock, 2000, Mean Reversion and the Volatility of Interest Rates: A Monte-Carlo Study Based on Different Sampling Frequencies, *Working Papers*, CUNY, Graduate Center.

[Ait-Sahalia 96] Ait-Sahalia, Y. A., 1996, Nonparametric Pricing of Interest Rate Derivative Securities, *Econometrica*, 64:3, 527-560.

[Andersen 97] Andersen, T. T., and J. Lund, 1997, Estimating Continuous-Time Stochastic Volatility Models of the Short-Term Interest, *Journal of Econometrics* 77, 343-377.

[Andrew 93] Andrews, D. W. K., 1993, Exactly Median-Unbiased Estimation of First Order Autoregressive/Unit Root Models, *Econometrica* 61, 139-1651.

[Baillie 89] Baillie, R. T. and T. Bollerslev, 1989, The Message in Daily Exchange Rates:A Conditional-Variance Tales, *Journal of Business and Economic Statistics*, 7:3, 297-306.

[Ball 97] Ball, C. A. and W. N. Torous, 1997, Unit Roots and the Estimation of Interest Rate Dynamics, *Journal of Empirical Finance*, 3, 215-238.

[Ball 99] Ball, C. A. and W. N. Torous, 1999, The Stochastic Volatility of Short-term Interest rates: Some International Evidence, *Journal of Finance*, 54, 2339-2359.

[Bali 00] Bali, T. and S. Neftci, 2000, The Relativity of Volatility, Stress Testing, *Journal of Risk*, 54.

[Bergstrom 83] Bergstrom, A. R., 1983, Gaussian Estimation of Structural Parameters in Higher-Order Continuous-time Dynamic Models, *Econometrica*, 51, 117-152.

[Bergstrom 84] Bergstrom, A. R., 1984, Continuous-time Stochastic Models and Issues of Aggregation Over Time, Ch. 20 in *Handbook of Econometrics* Vol. II eds. Z. Griliches and M.D. Intriligator, (North-Holland, Amsterdam), 1145-1212.

[Bessembinder 95] Bessembinder, H., J. F. Coughenour, P. J. Seguin, and M. Smoller, 1995, Mean Reversion in Equilibrium Asset Prices: Evidence from the Futures Term Struture, *Journal of Finance*, L:1, 361-375.

[Black 90] Black, F., E. Derman, and W. Toy, 1990, A One-factor Model of Interest Rates and its Application to Treasury Bond Options, *Financial Analysts Journal*, Jan-Feb, 33-39.

[Bliss 98] Bliss, R. R., and D. C. Smith, 1998, The Elasticity of Interest rate Volatility: Chan, Karolyi, Longstaff and Sanders Revisisted, *Journal of Risk*, 1:1, 21-466.

[Brennan 80] Brennan, M. J. and E. S. Schwartz, 1980, Analyzing Convertible Bonds, *Journal of Financial and Quantitative Analysis*, 15, 907-929.

[Brenner 96] Brenner, R. J., R. H. Harjes and K. F. Kroner, 1996, Another Look at Models of the Short-Term Interest Rates, *Journal of Financial and Quantitative Analysis*, 31:1, 85-107.

[Broze 95] Broze, L. O. Scaillet and J. Zakoian, 1995, Testing for Continuous-Time Models of Short-term Interest Rate, *Journal of Empirical Finance*, 2, 199-223.

[Broze 98] Broze, L. O. Scaillet and J. Zakoian, 1998, Quasi-Indirect Inference for Diffusion Processes, *Econometric Theory*, 14:2, 161-186.

[Chan 92a] Chan, K. C., G. A. Karolyi, F. A. Longstaff and A. B. Sanders, 1992, An Empirical Comparison of Alternative Models of the Short-Term Interest Rate, *Journal of Finance*, 48, 1209-1227.

[Chan 92b] Chan, K. C., G. A. Karolyi, F. A. Longstaff and A. B. Sanders, 1992, The Volatility of Japanese Interest Rates: A Comparison of Alternative Term Structure Models *Pacific-Basin Capital Markets Research*, 3, 119-135.

[Cox 80] Cox, J., J. Ingersoll, and S. Ross, 1980, An Analysis of Variable Rate Loan Contracts, *Journal of Finance*, 35, 389-403.

[Cox 85] Cox, J., J. Ingersoll, and S. Ross, 1985, A Theory of the Term Structure of Interest Rates, *Econometrica*, 53, 385-407.

[Dai 00] Dai, Q and K. J. Singleton, 2000, Specification Analysis of Affine Term Structure Models, *Journal of Finance*, 5, 1943-1978..

[Davidson 85] Davidson, R and J. G. MacKinnon, 1993, *Estimation and Inference in Econometrics*, (Oxford University Press), (New-York).

[Dietrich 86] Dietrich-Campbell, B. and E. Schwartz, 1986, Valuing Debt Options: Empirical Evidence, *Journal of Financial Economics*, 16, 321-343.

[Dybvig 89] Dybvig, P. H. and S. A. Ross, 1989, Arbitrage in The New Palgrave, *Finance, ed. J. Eatwell, M. Milgate and P. Newman*, W. W. Norton, New york

[Dothan 78] Dothan M. U., 1978, On the Term Structure of Interest Rate, *Journal of Finance*, 6, 59-69.

[Engle 96] Engle, F. R. and G. G. J. Lee, 1996, Estimating Diffusion Models of Stochastic Volatility, *Modelling Stock Market Volatility, ed. P. E. Rossi*, Academic Press, California

[Episcopos 00] Episcopos, A, 2000, Further Evidence on Alternative Continuous Time Models of the Short Term Interest Rate, *Journal of International Financial Markets, Institutions and Money*, 10, 199-212.

[Gourieroux 96] Gourieroux, C and A. Monfort, 1996, *Simulation-based Econometric Methods*, (Oxford University Press), (New-York).

[Hendry 84] Hendry, D. F., 1984, Monte-Carlo Experimentation in Econometrics, Ch.16 in *Handbook of Econometrics* Vol. II eds. Z. Griliches and M.D. Intriligator, (North-Holland,Amsterdam), 937-976.

[Honore 98] Honore, Peter, 1998, Maximum Likelihood Estimation of Non-Linear Continuous-Time Term Structure Models, Department of Finance, The Aarhus School of Business, Unpublished Ph.D dissertation.

[Karatzas 91] Karatzas, I. and S. Shreve 1991, Brownian Motion and Stochastic Calculus, *Graduate Text in Mathematics*, Springer, 2nd. Edition.

[Karlin 68] Karlin, S. and T. Taylor 1968, Kernel Function, *A Second Course of Stochastic Processes*, Academic Press, Chapter 5.

[Koutmos 98] Koutmos, G., 1998, The Volatility of Interest Rates Across Maturities and Frequencies, *The Journal of Fixed Income*, 8, 27-32.

[Marsh 83] Marsh, T. and E. Rosenfeld., 1983, Stochastic Processes for Interest Rates and Equilibrium Bond Prices, *Journal of Finance*, 38, 635-646.

[Merton 73] Merton, R. C., 1973, Theory of Rational Option Pricing, *Bell Journal of Economics and Management Science*, 4, 141-183.

[Merton 75] Merton, R. C., 1975, Theory of Finance from the Perspective of Continuous Time, *Journal of Financial and Quantitative Analysis*, 10, 659-674.

[Milshtein 74] Milshtein, G. N., 1974, Approximate Integration of Stochastic Differential Equations, *Theory of Probability and Its Applications*, 19, 557-562.

[Mollick 99] Mollick, A. V., 1999, The Real Exchange Rate in Brazil Mean Reversion or Random Walk in the Long Run?, *International Review of Economics and Finance*, 8, 115-126.

[Nankervis 96] Nankervis, J. C and N. E. Savin., 1996, The Level and Power of the Bootstrap t Test in the AR(1) Model With Trend, *Journal of Business and Economic Statistics*, 14:2, 161-168.

[Nowman 97] Nowman, K. B., 1997, Gaussian Estimation of Single-Factor Continuous-Time Models of The Term Structure of Interest Rates, *Journal of Finance*, 52, 1695-1706.

[Nowman 98] Nowman, K. B., 1998, Continuous-time Short Term Interest Rate Models, *Applied Financial Economics*, 8, 401-407.

[Oksendal 98] Oksendal, B., 1998, Stochastic Differential Equations, An Introduction with Application, *Fifth Edition*, Springer.

[Pardoux 85] Pardoux, E and Talay, D., 1985, Discretization and Simulation of Stochastic Differential Equations, *Acta. Applicandae Mathematica*, 3, 23-47.

[Poterba 98] Poterba, J. and Summers, L., 1998, Mean Reversion in Stock Prices: Evidence and Implications, *Journal of Financial Economics*, 22, 27-59.

[Pritsker 98] Pritsker, M., 1998, Nonparametric Density Estimation and Tests of Continuous-Time Interest Rate Models, *The Review of Financial Studies*, 11:3, 449-487.

[Quandt 83] Quandt, R. E., 1983, Computational Problems and Methods, Ch. 12 in *Handbook of Econometrics* Vol. I eds. Z. Griliches and M.D. Intriligator, (North-Holland, Amsterdam), 699-764.

[Richard 77] Richard S., 1977, An Arbitrage Model of the Term Structure of Interest Rates, *Journal of Financial Economics*, 6, 33-57.

[Stanton 97] Stanton R., 1997, A Nonparametric Model of Term Structure Dynamics and the Market Price Of Interest Rate Risk, *Journal of Finance*, 52, 1145-1185.

[Shiller 85] Shiller, R. J. and P. Perron, 1985, Testing the random-walk hypothesis: power versus frequency of observation, *Economic Letters*, 18, 381-386.

[Tse 95] Tse, Y. K., 1995, Some International Evidence on the Stochastic Behavior of Interest Rate, *Journal of International Money and Finance*, 14, 721-738.

[Vasicek 77] Vasicek, O., 1977, An Equilibrium Characterization of The Term Structure, *Journal of Financial Economics*, 5, 177-188.