MLE Is Alive and Well in the Financial Markets

B. Sailesh Ramamurtie and Scott Ulman

Federal Reserve Bank of Atlanta
Working Paper 96-17
November 1996

Abstract: In this paper we specify the basic set of economic criteria that any diffusion-driven interest rate or FX rate process must satisfy. We also develop the methodology that is implementable to test the validity of a proposed process insofar as it satisfies the basic criteria as well as the actual estimation of the parameters of an acceptable candidate process. In this paper we focus on processes such as the overnight repo rate process or the FX rate process, each of which is directly observable. We develop what we call the marginal maximum-likelihood estimation (MMLE) technique to distinguish it from the joint maximum-likelihood estimation (JMLE) technique, which we present in a separate paper. We also present some preliminary empirical results for both the interest rate process and the FX rate process.

JEL classification: G12, F31, E43, C13, C52

Ramamurtie is a visiting scholar at the Federal Reserve Bank of Atlanta. The views expressed here are those of the authors and not necessarily those of the Federal Reserve Bank of Atlanta or the Federal Reserve System. Any remaining errors are the authors' responsibility.

Please address questions of substance to Scott Ulman, Financial Solutions Incorporated, 3800 East 26th Street, Minneapolis, Minnesota 55406, 612/729-2202, sulman@rfc.com (or ScottUlman@msn.com); or B. Sailesh Ramamurtie, Department of Finance, College of Business Administration, Georgia State University, Atlanta, Georgia 30303, 404/651-2710, 404/651-2630 (fax), sramamurtie@gsu.edu.

Questions regarding subscriptions to the Federal Reserve Bank of Atlanta working paper series should be addressed to the Public Affairs Department, Federal Reserve Bank of Atlanta, 104 Marietta Street, N.W., Atlanta, Georgia 30303-2713, 404/521-8020, http://www.frbatlanta.org.

1. Introduction

Following the seminal work by Black, Scholes, and Merton, academics have developed increasingly sophisticated models for pricing bonds or foreign exchange (FX), and their derivative assets. Standard pricing models rely on stochastic differential equations (SDE) to describe the dynamics of underlying sources of uncertainty like interest or FX rates. Practical application of stochastic pricing models and forecasting require estimation of the SDE's parameters using *daily* sampling. The method of maximum likelihood (ML) provides a highly intuitive solution to the estimation problem because the parameters defining the drift and diffusion functions of an SDE are the only parameters appearing in the corresponding transition density.[1] Since the same parameters are the only arguments in the likelihood function formed from the transition density, ML estimators represent viable coefficients for valuation and forecasting equations and may be interpreted as the parameter values which make the observed data sample most likely.

Unfortunately, many previous attempts to estimate parameters for rate-type processes (e.g. interest or FX rates) have failed due to inadequate care in computational analysis. Proper and practical MLE analysis in the financial markets for rate-type processes must deal with the following set of economic and mathematical restrictions.

i) Sampling should be daily.

ii) Proposed dynamics must insure that economic processes are nonnegative, return immediately to positive levels (or assume that the processes cease to exist) if zero is reached, and possess steady-state distributions to prevent explosive growth of future values. Additional economic restrictions may also be legitimately imposed.

iii) Although it is tempting to describe asset or rate-process dynamics in terms of SDEs, maximum likelihood estimation deals with transition densities. Unfortunately, SDEs do **not** map uniquely to transition densities: a single SDE can generate **several** transition densities depending upon the process' behavior at its boundaries. This **specification problem** has profound implications for the form of the likelihood equation and testing and interpreting statistical significance of parameter estimates.

iv) When the rate-process is a *price* (like FX), then ML estimators from a single transition density function are the only necessary inputs for the partial differential equation (PDE) models academics have developed to price derivative assets based on the underlying process. For convenience, we designate this type of analysis as Marginal Maximum Likelihood Estimation (MMLE). On the other hand, when the rate-process is *not* a price (like interest rates), then ML estimators from a joint transition density function are necessary to provide inputs for PDE models to price assets driven by the non-price process. We call this type of analysis Joint Maximum Likelihood Estimation (JMLE).

The prior set of restrictions induces computational problems which fall into two general categories: technical specification and computer algorithms. Both categories are relatively esoteric and imply problems which may never be discovered, let alone resolved, by using *canned* computer algorithms. Fortunately, by implementing an active set strategy to enforce constraints, providing analytic gradients and hessians, disdaining attractive asymptotic expansions, and paying careful attention to machine overflow or underflow problems, it is possible to craft computational

[1] A diffusion SDE has the form $dr(t) = \mu(r,t) + \sigma(r,t)dZ(t)$ where $\mu(r,t)$ is called the *drift* function and $\sigma(r,t)$ is called the diffusion function. Both functions depend upon at least the current process' level and time t.

modules which provide statistically significant parameter estimates satisfying desired economic restrictions for financial rates.

In the remainder of this paper we will discuss the impact of our restriction set on specifying appropriate transition densities for rate processes; on computational adjustments necessary ~~to derive ML~~ estimators subject to the restrictions; and on statistical significance and testing. We will report statistically-significant parameter estimates for a number of different foreign-exchange and interest-rate processes using daily sampling and MMLE. We will conclude by suggesting a similar methodology for **joint** maximum-likelihood estimation procedures. In a companion paper [19], we provide the corresponding theory to implement JMLE for the interest-rate and preference parameters necessary to price bonds and their derivative assets.

2. Using economic restrictions to specify likelihood functions

Academics have discovered that **mean-reverting** processes will generally insure the economic restrictions in (ii)[2]. Consequently, the ensuing discussion will concentrate on computational implications of mean-reverting, square-root (MRSR) dynamics for interest rates and mean-reverting logarithmic (MRL) processes for FX rates (insuring that both x and 1/x follow the same distribution and are invariant to changes in the unit of accounting)[3]. However, the same type of analysis is appropriate for other dynamics as well.

Nearly everyone is familiar with the mean-reverting SDE associated with Cox, Ingersoll, and Ross'[7] general equilibrium description of interest-rate dynamics:

1) $dr(t)=\kappa[\theta-r(t)]dt+\sigma\sqrt{r(t)}\,dZ(t)$

where κ represents the speed of adjustment, θ the long-run mean, and σ the proportional volatility parameter. In (1), the elastically pulled long-run mean prevents explosive growth and the nonconstant diffusion function (proportional to the current process' level) prevent negative values[4]

Unfortunately, an SDE like (1) does **not** completely describe MRSR dynamics. The *same SDE can generate two (or more) transition densities with radically different properties and even different state spaces*! A pair of drift and diffusion functions from an SDE does not define a unique transition density because the associated infinitesimal parameters fail to govern the interest rate's behavior near a boundary like r(t) = 0.[5] In one case, relationships among the parameters may automatically restrict boundary behavior by insuring that the process will never reach zero; in another case, the process will reach zero for certain and remain there forever; in still other cases, a boundary may be arbitrarily defined as an absorbing or reflecting barrier; alternative definitions will result in radically different transition densities.

[2] However, some mean-reverting processes like the Ornstein-Uhlenbeck do not preclude negative values.
[3] See Dothan, Ramamurtie, and Ulman [12] for a description of the MRL process applied to FX rates.
[4] It is easy to show that the long-run conditional mean $\lim_{T\to\infty} E(r_T|r_t)$ and variance $\lim_{T\to\infty} Var(r_T|r_t)$ of the process are finite. Hence the process does not move inexorably toward either zero or infinity like the lognormal process.

Recall that the purpose of maximum-likelihood estimation is to determine the parameter values which are the most likely to have generated an observed time series given a specification of the underlying transition density function. Mathematically, a likelihood function is defined as the product of the hypothesized transition density evaluated ate each observation in a time-series of rates viewed as independent drawings. Hence, for a sample of size n+1 taken at time intervals of Δt from the process r(t) with the parameter set $\{\kappa, \theta, \sigma\}$ and the transition density p($\bullet$), the likelihood function L may be written as

$$2)\ L_n(\kappa,\theta,\sigma,\Delta t \mid r_0, r_1, \ldots, r_n) \equiv \prod_{j=1}^{n} p(r_j \mid r_{j-1}; \kappa,\theta,\sigma,\Delta t).$$

Note that the product of the transition density on the right-hand side of (2) is interpreted as a function of the process *conditional* on values of the parameters κ,θ,σ. Alternatively, the likelihood function on the left-hand side of (2) is defined in terms of the arguments κ, θ, σ *conditional* on the observed sample of interest rates. Since maximum-likelihood estimators (MLE) are the set of arguments $\{\kappa,\theta,\sigma\}$ which maximize the function in (2), they may be interpreted as the parameter values which make the observed time-series most likely to occur (since the likelihood function is conditioned on the observed sample). Regardless of interpretation, the first step in MLE is resolving the **specification problem** (cited in the introduction as **(iii)**; only then can we turn to computational issues implied by the specification.

The SDE in (1) provides an especially nice framework to investigate the specification problem. To properly specify a likelihood function for MLE estimation, we must answer four questions:

(Q1) When the SDE describing the process' dynamics is consistent with multiple transition densities, which densities are legitimate to use in the likelihood function if economic restrictions are to hold?

(Q2) Are analytic representations known for the legitimate densities?[6]

(Q3) What constraints on the parameters must apply to the resulting likelihood function?

(Q4) What do the constraints imply about hypothesis testing of the parameters?

For the SDE in (1), different relationships among parameters imply at least two possible transition densities. To decide which results are economically feasible, examine the following cases.

Case 1. Suppose that $\kappa > 0$; then the process will attain value r(t) = 0 at some time t with probability one and remain there forever! Alternatively, if $\kappa,\theta > 0$ but $2\kappa\theta < \sigma^2$, then r(t) = 0 may be defined as an *absorbing barrier* and the same result will occur. The resulting transition density function is *defective* (i.e., *it sums to less than one* due to the positive probability of absorption at the origin) and was first identified by Feller [13].[7] This defective density (with its implicit parameter constraints) was first used in the finance literature by Cox and Ross [8] to describe the price behavior of a dividend-paying stock.

---

[5]In other words, the parameters do not govern the behavior of the interest rate as the level gets close to zero. See Karlin and Taylor [16] for technical details.

[6] This is not a trivial question. The transition density for $dr(t) = \kappa[\theta - r(t)]dt + \sigma r(t) dZ(t)$ (the mean-reverting lognormal process) is currently unknown. In general, mathematicians utilize Laplace transforms or continuous-time spectral representation methods (as described in Karlin and Taylor [16], Chapter 15) to identify transition densities corresponding to an SDE and its boundary conditions. See Prezas and Ulman [18] for the derivation of several such densities.

[7]Feller called his solutions *norm-reducing*. When the solutions are probability distributions, norm-reducing is synonymous with *defective*, i.e. the distribution function sums to less than 1!

Case 2. Suppose that $2\kappa\theta \geq \sigma^2$; then the process will always be positive and cannot reach zero. Alternatively, if $\kappa,\theta > 0$ but $2\kappa\theta < \sigma^2$, $r(t) = 0$ may also be arbitrarily described as a reflecting barrier causing the rate to immediately become positive if it ever reaches zero.[8] In either case, the resulting transition density (with its implicit parameter constraints) insures desired behavior for nominal rates and was used in the finance literature by CIR [7]to describe interest rates.

Case 1 has the following economic interpretation for FX rates: at some finite time in the future, the exchange rate will sink to 0 and stay there, i.e., the currency will become worthless relative to the reciprocal currency. Its implication for interest rates is equally startling: at some finite future date, investors will lend money with no recompense in nominal terms for all future time! Fortunately, case 2 appears to satisfy economic criteria as long as the actual ML computational process *enforces the constraints* $\kappa,\theta > 0$. Enforcing the constraints requires an active-set strategy during optimization. In addition, the constraints have a significant impact on hypothesis testing: they imply hypothesis tests like

H0: $\kappa > 0$

H1: $\kappa \leq 0$

should not be proposed since (H1) is inconsistent with the tested transition density.

3. Computational problems induced by proper economic specification

It is now possible to write the optimization problem using the properly restricted MRSR transition density and then examine the computational difficulties. The appropriate specification (using the standard approach of maximizing the log of the likelihood function) is:

$$\max_{\{\kappa,\theta,\sigma\}} \quad \log L_n = n\log c + (q/2)\left[\log(\frac{r_n}{r_0}) + n\kappa\Delta t\right] - \sum_{j=1}^{n} c[r_j + r_{j-1}e^{-\kappa\Delta t}] + \sum_{j=1}^{n}\log(I_q(z_j)) \tag{3}$$

subject to $\kappa > 0,\ 2\kappa\theta \geq \sigma^2$

where

$$c(\kappa,\sigma,\Delta t) \equiv \frac{2\kappa}{\sigma^2[1-\exp(-\kappa\Delta t)]} \quad ;$$

$$q(\kappa,\theta,\sigma) \equiv 2\kappa\theta/\sigma^2 - 1;$$

$$z_j(\kappa,\sigma,\Delta t,r_j,r_{j-1}) \equiv 2c(r_j r_{j-1}\exp(-\kappa\Delta t))^{1/2};$$

[8]Note that $\kappa,\theta > 0$ but $2\kappa\theta < \sigma^2$ is consistent with arbitrarily defining $r(t)=0$ as either a reflecting or absorbing barrier. Different densities result from the choice. The ML parameters for the absorbing-barrier density must satisfy the restrictions i) $\kappa,\theta > 0$ but $2\kappa\theta < \sigma^2$; or ii) $\kappa \leq 0$. On the other hand, ML estimators for a reflecting-barrier density need merely satisfy $\kappa,\theta > 0$ since the same density is consistent with $\kappa\theta > \sigma^2$. The important point is that both densities require a constrained optimization to obtain ML estimators.

$I_q(\bullet)$ is a modified Bessel function of the first kind of order q and Δt is the time interval between successive observations in the sample time series of size n + 1.

Equation (3) is a standard nonlinear optimization problem with respect *to continuously differentiable* functions. Hence, a(n approximate) global solution can be obtained numerically via computer using a Newton method modified with a backward line search to find the parameter set $\{\kappa, \theta, \sigma\}$ which makes the gradient vector zero and the *Hessian matrix* negative definite subject to the constraints $\kappa, \theta > 0$.[9] Although the computer approach seems straightforward, several technical details induced by *daily sampling* and the Bessel function in the objective equation seriously complicate the solution algorithm.

The following four factors required major revamping of the traditional solution algorithm.

i) Constrained optimization required use of an active-set strategy.[10]
ii) No single asymptotic expansion for the bessel function was appropriate because the relative magnitudes of the order and argument changed significantly with each Newton step. This forced usage of an infinite series representation for the bessel function.
iii) The objective function exhibited radically different sensitivities to changes in the parameters $\{\kappa, \theta, \sigma\}$. This required the use of analytic derivatives in place of numerical derivatives.
iv) Daily sampling ($\Delta t \equiv 1/252$) coupled with interest rates or FX rates less than 1 in value resulted in very large arguments for the bessel function. This caused problems with machine underflow/overflow in the convergence algorithms.

Since active-set strategies are well known, we will confine our discussion to factors (ii)-(iv) in the remainder of this section.

Asymptotic expansions and the bessel function

With daily sampling, the argument z of the Bessel function in (3) becomes very large, making it tempting to use Hankel's asymptotic expansion[11]

$$\text{4)} \quad I_q(z) \approx \left[\frac{e}{(2\pi z)^{1/2}}\right]\left[1 - \frac{(\mu-1)}{(8z)} + \frac{(\mu-1)(\mu-9)}{2!(8z)^2} - \dots\right]$$

where $\mu = 4q^2$.

Unfortunately, (4) is appropriate only for Bessel functions with fixed order q as the argument z becomes large. When solving (3) the order $q(\kappa, \theta, s)$ is not fixed. Instead, each Newton step induces a change in each of the parameters. As the speed of adjustment parameter κ rises, q also increases dramatically (while the argument z declines marginally). For large (but reasonable) values of κ (e.g., values greater than 10), the alternating series in (4) contains a succession of terms with magnitudes greater than 1.0 and hence fails to converge. Consequently, the asymptotic expansion can lead to nonsensical results for reasonable values of κ: derivatives of the bessel function

[9]See Dennis and Schnabel [10] for technical details.

[10] In active set strategies, when parameter estimates reach their constraint boundaries during the algorithm, their values are fixed at the boundary and the rank of the gradient and hessian are reduced accordingly to calculate the next Newton step. At each new step, the algorithm checks whether the new gradient and Hessian values would create a new Newton step moving the constrained variable away from its bound. If so the variable is reactivated; otherwise it is held at its bound.

[11] The argument becomes large since with daily sampling Δt = 1/252. See Abramowitz and Stegun [1], p.377, for Hankel's expansion.

with respect to order q apparently become positive, causing the log-likelihood function in (3) to increase without bound as κ rises.[12]

Consequently, the bessel function should be evaluated as the infinite series

$$5)\quad I_q(z_j)=\sum_{p=0}^{\infty}\frac{(z_j/2)^{2p+q}}{\Gamma(q+p+1)\Gamma(p+1)}$$

Analytical versus numerical derivatives

Canned optimization packages generally calculate the gradient and hessian by using numerical derivatives calculated at each Newton step using a uniform step size.[13] This *ostensibly* protects the user from making errors in the calculations of complicated mathematical functions. Unfortunately, some objective functions like (3) exhibit radically different sensitivities to changes in the parameters $\{\kappa,\theta,\sigma,\}$ at different levels. This implies that at many parameter levels, the uniform step size is too small to accurately measure change in the objective function for one or more variables. In such cases, the relative change measured by the numerical derivative is merely noise. Unfortunately, the algorithm then uses the noise to reflect a specious Newton step and the entire mechanism soon becomes stuck.

Even near an optimum, numerical derivatives may be highly unreliable. The magnitude of discrepancies can be clearly seen by comparing the (correct) analytic Hessian matrix to the (incorrect) numerical Hessian matrix near an optimum for the loglikelihood function in (3):

$$\nabla^2 L_n = \begin{bmatrix} -0.1143 & 5.0020 & -1.4376 \\ 5.0020 & 10.3699 & -53.3809 \\ -1.4376 & 53.3809 & 3525.2856 \end{bmatrix} \quad \textit{analytic Hessian}$$

$$H^2 L_n = \begin{bmatrix} 0.7418 & 10.1190 & 1.0518 \\ 8.6741 & 35.9514 & -34.6556 \\ -8.8423 & -95.7455 & 3489.5289 \end{bmatrix} \quad \begin{matrix}\textit{numerical Hessian} \\ \textit{step size of } 8.43 \times 10^{-8}\end{matrix}$$

Note that analytical values deviate substantially from their numerical counterparts. Furthermore, pairs of numerical cross-derivatives assume substantially different values. In fact, one pair of numerical cross-derivatives has values with opposite signs![14] Such discrepancies are never unearthed by *canned* optimization packages. Instead, they result in improper Newton steps and poor parameter estimates. In other words, a researcher may end up rejecting the model on the basis of a poor fit because undiscovered finite-precision arithmetic errors resulted in improper calculations for derivatives. Our analysis avoids this problem by using analytic equations for the gradient and Hessian (see Appendices A and B for the appropriate equations).

Algorithms for infinite-series solutions

[12]It can be proved that the Bessel function's derivative with respect to order is negative for all q. The nonsensical numerical results obtained from (4) occur because of the improper use of a nonconvergent series. (4) holds only for small, fixed values of q.

[13] Generally, packages set the step size as the square root of machine epsilon.

[14] The numerical derivative H13 is calculated by differentiating first with respect to parameter 1 and then with respect to parameter 3 whereas H31 is merely calculated in the reverse order.

The objective function (3) and its analytic gradient and hessian all include infinite series. This representation presents two significant computational problems: potential machine underflow/overflow[15] and excessive calculation times. Fortunately, there is a simple and elegant solution to both.[16] Merely rewrite the summand of the infinite series as an exponential function $e^{G(p,j,z)}$. Then to avoid underflow or overflow, multiply by an exponential scaling factor. To speed the calculations, solve for the index which (approximately) maximizes the summand and, beginning with the optimal index (rather than at 0), sum forward and backward until each term added is machine zero.

To illustrate the methodology, consider the bessel function in (5). Use Stirling's formula for the gamma functions to write

6a) $$2\pi e^{-z_j} I_q(z_j) = \sum_{p=0}^{\infty} \frac{e^{G1(p,j,z)}}{d(p+q+1)d(p+1)}$$

where

6b) $$G1(p,j,z) = (2p+q)\log\frac{z_j}{2} + (2p+q+2) - (p+q+\tfrac{1}{2})\log(p+q+1) - (p+\tfrac{1}{2})\log(p+1) - z_j$$

and

6c) $$d(x) = 1 + \frac{1}{12x} + \frac{1}{288x^2} - \frac{139}{51840x^3} - \frac{571}{2488320x^4} + \cdots$$

To avoid underflow and speed calculations, we must skip most of the negligible terms at the start of the infinite series in (6a)[17]. To do this, we find the value p^* which solves the subproblem

6d) $$\max_p G1(p,j,z)$$

and then actually compute the infinite series in (6a) by beginning the summation at p^* and summing forward and backward until all additional terms are machine zero.[18]

4) Empirical results

Assessment of any model's statistical accuracy generally requires a specification of the sampling distribution. Unfortunately, exact sampling distributions for estimators obtained by nonlinear methods are generally intractable and asymptotic distribution theory applies.[19]

---

[15] In (5) any or all of the individual terms in the summand may overflow prior to convergence. In ANSI C 3.1.2.5, the smallest double-precision exponential is $1.7e^{-308}$ and for long doubles e^{-4392}; otherwise, underflow occurs.

[16] Similar methods work for all gradient and Hessian functions presented in the appendices. We use this type of solution whenever any z_j value exceeds 1400. When $z_j < 1400$ for all j, we use slightly different algorithms. E.g. if $q > 1000$ and $z_j < 1400$, we utilize asymptotic expansion 9.7.7 from Abramowitz and Stegun [1] but scale the terms to prevent underflow.

[17] All computers have a smallest value, called machine zero. Any computation returning an absolute value less than machine zero produces an underflow which must be avoided. Any such terms are negligible, i.e. they add nothing to the sum.

[18] This method proves very efficacious since terms larger than machine epsilon frequently don't occur until p is larger than 5000. In some cases, we might also find $z\max = \max_j\{z_j\}$. and multiply both sides of (6a) by $e^{z\max}$ prior to calculate the infinite sum on the RHS of (6a). We would then scale the function down similarly when calculating the log of the bessel function in (3).

Prezas and Ulman [18] have proved that estimators for the norm-preserving MRSR process (associated with the loglikelihood function in (3)) are asymptotically normal with a covariance matrix equal to the inverse of the expected Hessian matrix calculated using the log likelihood function. Additionally, they have derived the Information Matrix (inverse of the expected Hessian) to calculate exact asymptotic precision measures. However, to simplify actual calculation of asymptotic standard errors for this paper, we employ a standard practice in nonlinear optimization suggested by Bard [2]: we replace the inverse of the expected Hessian matrix with the inverse of the Hessian matrix evaluated at the optimal parameter values.[20] In the attached panels, we display estimators and asymptotic standard errors for each parameter. We feel relatively comfortable with the old rule-of-thumb: an estimator is statistically significant if it is at least twice its (asymptotic) standard error.[21]

Our algorithms run as a Windows NT program using ODBC database access. The algorithms are currently single-threaded. Although time to convergence is a function of the sample time series, estimations generally take 10-90 seconds on a single-processor DEC Alpha AXP or powerful Intel Pentium.[22] Since many auxiliary functions used to calculate the Hessian are double summations (see Appendix B), it appears highly likely that multithreading to take advantage of multiple processors would provide significant performance gains. However, we leave multithreading as a topic for future research.

Interest rate processes

The norm-preserving MRSR process is ideal for describing interest-rate movements. Panel 1 displays overnight repo rates in the U.S. between 4/3/87 and 6/29/87 (60 observations). Note the (asymptotic) standard errors are less than half the estimators for all parameters and the value of the loglikelihood function is relatively high. Panel 2 shows similar results for 7-day LIBOR rates in a completely different time frame (6/22/92-9/15/92). By carefully crafting optimization algorithms rather than relying on canned packages, it is possible to obtain significant ML estimators for any interest-rate time series which displays mean-reverting behavior.

FX rate processes

If the only economic restrictions for FX rates are those described in the introduction (economic processes which are nonnegative, return immediately to positive levels if zero is reached, and possess steady-state distributions to prevent explosive growth of future values) then the norm-preserving MRSR density is adequate for rate dynamics. Panel 3 presents ML estimators for the U.S. Dollar versus British Pound exchange rate between September 21, 1992 and June 22, 1993 (197 observations) under norm-preserving MRSR dynamics. Panel 4 presents similar results for

[19]A few processes (like those with normal or lognormal increments) can be factored as a member of an exponential family. In such cases, joint sufficient statistics can be recognized in the argument of the exponential function and exact standard errors may be reported. Such cases are exceptions rather than the rule.

[20]Bard [2] rationalizes such approximations to the asymptotic covariance matrix (which are utilized by all *canned* computer packages) by claiming that "computation of the required expectation is very tedious, if not impossible". Hence we leave any discrepancy between the approximate and theoretical covariance matrices as a topic for future research. Quite interestingly, Bard has run simulations which indicate that nonlinear estimation procedures can provide very good data fits in small samples even if the resulting asymptotic standard errors are very large.

[21] Quite interestingly, practitioners routinely utilize the ML estimator for volatility from a lognormal process even though the mean estimate is poor. In fact, the asymptotic standard error for the mean from a lognormal process is proportional to $(1/n\Delta t)$ and hence is always large relative to the estimator unless the sample size is gargantuan.

[22] We have found some cases where convergence required 300 seconds.

the Deutschmark versus Pound. Note that in both cases, all parameter estimates are more than double the corresponding asymptotic standard errors.

In recent work, Dothan, Ramamuritie, and Ulman (DRU) [12] have suggested that additional economic restrictions (beyond (i)-(iv) listed in the introduction) should be applied to foreign exchange. Specifically, they argue that FX processes should satisfy two additional economic restrictions which we now add to our list:

v) (*Symmetry*) The distribution of any exchange rate and its inverse must be symmetric, i.e., the distribution functions for the FX rate and its inverse will be identical but the parameters may be different. In other words, it doesn't matter whether traders consider Deutschmarks to Dollars or vice versa: the distributional characteristics will be the same.

vi) *(Invariance)* The distribution of any exchange rate must be invariant to changes in the unit of currency. In other words, changing the domestic currency's name from dollars to ducats and defining ten ducats per old dollar should not change the characteristics of the exchange rate process. Mathematically, given the present rate (but subject to an adjustment of parameters) the conditional distribution must be homogenous of degree zero in the present and future rates of exchange.

The MRSR density fails to satisfy these properties: if Dollars vs. Pounds has MRSR dynamics, then Pounds vs. Dollars does not. Such a result is highly unintuitive.

To resolve these additional economic issues, DRU [12] developed a four-parameter mean-reverting logarithmic (MRL) process as an FX analogue to the MRSR process; it satisfies all six economic restrictions. Denote the four parameters as $\{n_1,n_2,n_3,n_4\}$ where n_1 is denoted the volatility parameter, n_2 is a critical floor where the home government intervenes to strengthen its currency, n_3 is a traditional speed-of-adjustment, and n_4 is a (logarithmic) long-run mean. Let x represent the exchange rate and let $y \equiv \log(x/n_2)$. Then the dynamics for both an exchange rate and its reciprocal can be represented as

7) $dy = n_3\left[\log(n_4/n_2) - y\right]dt + n_1 y^{1/2}\, dW.$

If we substitute $\theta \equiv \log(n_4/n_2)$ then (7) is a precise analogue of (1).

If the critical floor n_2 is defined as an exogenous variable specifying the level where traders believe the Central Bank will intervene to protect its currency, then the norm-preserving MRL process may be estimated using the same algorithms as the MRSR process subject to the proper transformations[23]. Alternatively, if n_2 is treated as an endogenous variable, then (3) must be respecified as an optimization across four variables and modified algorithms must be developed. For this paper, we have chosen to treat the critical floor as an exogenous variable. Panels 5-12 present results for the Dollar versus the British Pound and for the Deutschmark versus British Pound between September 1992 and June 1993 using daily sampling.

---

[23] It is easy to show that the MLE estimator of n_4 may then be calculated as $n_2 e^{\theta^*}$ where θ^* is the MLE for θ and $s(n_4) = n_4^{1/2} s(\theta^*)$ is the appropriate standard error.

Any optimizer occasionally becomes stuck at a local optimum. Panel 5 depicts a warning message displayed when the optimizer became stuck using the original guesses for Dollar versus Pound. A quick change in the initial guess dialog box resulted in convergence to the values shown in Panel 6 with a critical floor specified at 1.1. Panels 7 and 8 show the results of increasing the critical floor to 1.2 and 1.3 respectively. Note that the only significant difference is the sharp increase in the rate volatility parameter n_1. Panel 9 shows the conditional mean and variance for the Dollar versus Pound FX rate for the next three to seven weeks. Panels 10 and 11 show similar results for the Deutschmark to Pound FX rate in the same time interval at critical floor values of 1.96 and 2.00. Once again, the volatility parameter n_1 rose with the critical floor. Panel 12 depicts the parameter estimates resulting from selecting a subseries of approximately one-third the observations around the low point reached by the time series. These results may be contrasted to those in Panel 10. Finally, Panel 13 shows the time series over the September 1992 through June 1993 period for the Japenese Yen versus British Pound. Note that the graphed series moves inexorably lower, demonstrating no evidence of reversion toward any mean. The parameter estimate for the speed of adjustment is relatively low compared to its standard error and the value of the loglikelihood function is abysmal compared to results for the other two FX series. Hence the optimizer performs poorly (as expected) in the absence of mean-reversion. This time series actually looks like a lognormal process headed toward zero.

5) Implications for forecasting and pricing

Cox, Ingersoll, and Ross' [9] seminal article on general equilibrium pricing showed that *any* asset can be valued as the risk-adjusted, discounted expectation of future payouts. The SDE approach to uncertainty has been widely adopted because its continuous-time analogues of first-order autoregressive processes[24] with independent increments allow good mathematicians to determine transition densities necessary to determine corresponding pricing equations. Unfortunately, forecasts from the transition densities contain only drift-related terms. For example, a forecast from (1) could be written as

8) $E_t[r(s)] = \theta + [r(t) - \theta]e^{-\kappa(s-t)}$

which tends to the long-run mean θ as $s \to \infty$. Note , however, that forecasts using our ML estimators do not answer questions like, if the previous day's error was negative, will today's error be negative? Such queries might be better answered by ARIMA models which require information only from the autocovariance function rather than the entire transition density. Panel 9 depicts 3-7 week-ahead forecasts of the Dollar versus Pound FX rate estimated in Panel 8. Note the movement toward the long-run mean and the increasing variance of the forecasts farther out in time.

The ML estimators for the FX series depicted in Panels 3-8 and 10-12 are sufficient to calculate prices of options[25] and futures' options on FX. As in Black-Scholes, the drift-related parameters may be discarded. However,

[24] E.g. (1) can be described as an AR(1) process with a highly modified first-order moving-average component. Note that the difference in successive noises (dZ(t)) is multiplied not by a constant coefficient as in ARIMA models but by a function of the level of the process itself.

[25] See Dothan, Ramamurtie, and Ulman [12] for the appropriate pricing equations. In cases where one or more parameter estimates is not statistically significant, users may wish to try estimation for some alternative stochastic process.

they may prove useful in forecasting and in reassuring users that the hypothesized density indeed provides a good fit for the data.

Unfortunately, the ML estimators for the interest-rate in Panels 1-2 are *not* sufficient to price bonds or their derivative assets since the estimated set does not include the preference parameter λ. It is convenient to describe the type of analysis we have completed in this paper as Marginal MLE (MMLE) since the loglikelihood function (3) contains information from a single time-series and uses a *marginal* transition density function. To jointly estimate the preference parameter λ and the interest-rate parameters $\{\kappa,\theta,\sigma\}$, we must include information from multiple time series (the overnight Repo rate and the 200+ coupon Treasury prices (or Strips)). In a companion paper [19] we develop a general theory to conduct Joint MLE (JMLE) using a joint transition density function across the overnight RP rate and coupon bond prices; this approach provides the ability to simultaneously measure asymptotic standard errors for each parameter estimate (including the preference parameter). We make the general theory operational by hypothesizing a joint uniform distribution for bond pricing errors; this seems reasonable since bond prices are generallyy rounded to the nearest 1/32. Other error specifications (e.g., truncated normal) could also be used. We are currently working on computer algorithms to implement JMLE.[26]

6) Summary

In this paper, we have demonstrated that applying realistic economic restrictions to the dynamics for foreign-exchange and interest-rate processes leads to many relatively esoteric computational problems in marginal maximum-likelihood estimation (MMLE). Our studies indicate that many of these problems will never be uncovered if researchers use *canned* computer algorithms relying on uniform step sizes, numerical derivatives, and standard asymptotic expansions. Instead, the researchers will invariably report statistically insignificant results or nonconvergence because the canned algorithms end up by using noise to determine proper Newton steps.

The combination of daily sampling and complicated mathematical functions necessary for mean-reversion forced us to utilize analytical derivatives for the gradient and hessian, to obtain series solutions for many of the intermediary functions, and to guard religiously against underflow and overflow. By paying meticulous attention to such computational detail, we were able to produce algorithms which provided statistically significant parameter estimates for interest-rate series using a norm-preserving MRSR transition density and for FX rates using both norm-preserving MRS and MRL transition functions. These MML estimators are sufficient to make simple forecasts and to price FX options and futures.

However, in order to price bonds and their derivative assets, researchers must identify a preference parameter in addition to the estimators for the interest-rate process. To elegantly resolve this problem, we cite our contemporaneous work on joint maximum-likelihood estimation (JMLE). We believe that sequential quadratic programming can be successfully employed to solve the bond-pricing problem if the same meticulous care is used to

[26] An alternative procedure might rely on calculating the MMLE for the interest-rate process and then in a second stage using least-squares to find the preference parameter which minimizes the sum of the squared pricing errors across the set of coupon bonds conditional on the interest-rate estimators. In earlier work, Ulman [21] used this procedure to obtain a phenomenal fit (with pricing errors of between $0.05 and $0.50 on $100 face-value bonds) in the short and intermediate

insure computational integrity as was required for MMLE. Hence we rather blithely conclude that maximum-likelihood estimation is indeed alive and well in the financial markets!

term Treasury market (maturities less than 11 years). Unfortunately, this method provides no information about the joint standard errors of the estimators.

Appendix A: Analytic Gradient

Let $\nabla L = [\frac{\partial L}{\partial \kappa}, \frac{\partial L}{\partial \theta}, \frac{\partial L}{\partial \sigma}]$ be the gradient of the log likelihood function $L(\bullet, \text{n})$ in (3). Further, let $c(\bullet)$, $q(\bullet)$, $z_j(\bullet)$, n and Δt be defined as in (3). Let $\psi(\bullet)$ represent the psi function and $\Gamma(\bullet)$ the gamma function. Then the elements of the gradient may be expressed as follows.

A1) $$\nabla L_2 = \left(n \log c + \sum_{j=1}^{n} \log r_j - \sum_{j=1}^{n} \frac{(z_j/2)^q}{I_q(z_j)} \sum_{p=0}^{\infty} \frac{(z_j/2)^p \psi(q+p+1)}{p!\Gamma(q+p+1)} \right) \left(\frac{2\kappa}{\sigma^2} \right)$$

A2) $$\nabla L_1 = \frac{\theta}{\kappa} \nabla L_2 + \left(\frac{1}{c} \frac{\partial c}{\partial \kappa} \right) n(1+q) - \frac{\partial c}{\partial \kappa} \sum_{j=1}^{n} r_j - c e^{-\kappa \Delta t} \sum_{j=1}^{n} r_{j-1} \left(\frac{1}{c} \frac{\partial c}{\partial \kappa} - \Delta t \right) - \left(\frac{1}{c} \frac{\partial c}{\partial \kappa} - \frac{\Delta t}{2} \right) \sum_{j=1}^{n} z_j \frac{I_{q+1}(z_j)}{I_q(z_j)}$$

A3) $$\nabla L_3 = -\frac{2}{\sigma} \left[\sum_{j=1}^{n} z_j \frac{I_{q+1}(z_j)}{I_q(z_j)} - c \left(\sum_{j=1}^{n} r_j + r_{j-1} e^{-\kappa \Delta t} \right) + n(1+q) + \theta \nabla L_2 \right].$$

Appendix B: The analytic Hessian matrix

Let $\nabla^2 L$ represent the Hessian matrix corresponding to the loglikelihood function $L(\bullet, n)$ in (3) and the gradient ∇L defined in (A1)-(A3). Further, let $c(\bullet)$, $q(\bullet)$, $z_j(\bullet)$, n and Δt be defined as in (3). Let $\psi(\bullet)$ represent the psi function and $\Gamma(\bullet)$ the gamma function. Then the elements of the Hessian may be expressed much more simply in terms of a set of auxiliary functions and constants.

The auxiliary functions:

B1) $$T1(z,q,n) = \sum_{j=1}^{n} \frac{z_j I_{q+1}(z_j)}{I_q(z_j)}$$

B2) $$T2(z,q,n) = \sum_{j=1}^{n} z_j^2 \left[\left(\frac{I_{q+1}(z_j)}{I_q(z_j)} \right)^2 - 1 \right]$$

B3) $$T4(z,q,n) \equiv \sum_{j=1}^{n} \sum_{p=0}^{\infty} \frac{(z_j/2)^{2p+q} \left[\psi^2(p+q+1) - \psi'(p+q+1) \right]}{p!\Gamma(p+q+1) I_q(z_j)}$$

B4) $$T5(z,q,n) \equiv \sum_{j=1}^{n} \left[\left(\sum_{p=0}^{\infty} \frac{(z_j/2)^{2p+q} \psi(p+q+1)}{p!\Gamma(p+q+1) I_q(z_j)} \right)^2 \right]$$

B5) $$T6(z,q,n) \equiv 2 \sum_{j=1}^{n} \sum_{p=0}^{\infty} \frac{p\,(z_j/2)^{2p+q} \psi(p+q+1)}{p!\Gamma(p+q+1) I_q(z_j)}$$

B6) $$T7(z,q,n) \equiv \sum_{j=1}^{n} \frac{z_j I_{q+1}(z_j)}{I_q(z_j)} \sum_{p=0}^{\infty} \frac{(z_j/2)^{2p+q} \psi(p+q+1)}{p!\Gamma(p+q+1) I_q(z_j)}$$

B7) $$T8(z,q,n) \equiv \sum_{j=1}^{n} z_j \sum_{p=0}^{\infty} \frac{(z_j/2)^{2p+q+1} \psi(p+q+2)}{p!\Gamma(p+q+2) I_q(z_j)}$$

The auxiliary constants:

B8) $$C1 = \frac{1}{c} \frac{\partial c}{\partial \kappa}$$

B9) $$C2 = \frac{c\sigma^2 \Delta t e^{-\kappa \Delta t}}{2}$$

B10) $$C3 = C1 - \Delta t$$

B11) $C4 = C1 - \frac{\Delta t}{2}$

B12) $C5 = \sum_{j=1}^{n} r_j + r_{j-1} e^{-\kappa \Delta t}$

The Hessian:

B13) $\nabla^2 L_{12} = \frac{\nabla L_2}{\kappa} + \frac{2n}{\sigma^2}(1 - C2) - \frac{\partial q}{\partial \theta}\frac{\partial q}{\partial \kappa}[T5(\bullet) - T4(\bullet)] - C4\frac{\partial q}{\partial \theta}[T6(\bullet) - T7(\bullet)]$

B14) $\nabla^2 L_{22} = -\left(\frac{\partial q}{\partial \theta}\right)^2 [T5(\bullet) - T4(\bullet)]$

B15) $\nabla^2 L_{32} = -\frac{\nabla L_2}{\sigma} - \frac{4\kappa n}{\sigma^3} - \frac{\partial q}{\partial \theta}\frac{\partial q}{\partial \sigma}[T5(\bullet) - T4(\bullet)] + \frac{2}{\sigma}\frac{\partial q}{\partial \theta}[T6(\bullet) - T7(\bullet)]$

B16) $\nabla^2 L_{33} = \frac{1}{\sigma}\left[2C5\frac{\partial c}{\partial \sigma} - \nabla L_3 - 2\theta \nabla^2 L_{32} - 2n\frac{\partial q}{\partial \sigma} - \frac{4}{\sigma}[2qT1(\bullet) + T2(\bullet)] - 2\frac{\partial q}{\partial \sigma}[T8(\bullet) - T7(\bullet)]\right]$

B17) $\nabla^2 L_{31} = \frac{2}{\sigma}\left\{ce^{-\kappa \Delta t}C3\sum_{j=1}^{n} r_{j-1} + \frac{\partial c}{\partial \kappa}\sum_{j=1}^{n} r_j - \frac{2n\theta}{\sigma^2} - \theta \nabla^2 L_{12} + C4[2qT1(\bullet) + T2(\bullet)] + \frac{\partial q}{\partial \kappa}[T8(\bullet) - T7(\bullet)]\right\}$

B18)
$$\begin{aligned}\nabla^2 L_{11} &= n\theta \Delta t e^{-\kappa \Delta t}\frac{\partial c}{\partial \kappa} + \frac{\partial^2 c}{\partial \kappa^2}\sum_{j=1}^{n} r_j + e^{-\kappa \Delta t}\left[2\Delta t\frac{\partial c}{\partial \kappa} - c(\Delta t)^2 - \frac{\partial^2 c}{\partial \kappa^2}\right]\sum_{j=1}^{n} r_{j-1} + \frac{\partial q}{\partial \kappa}[T7(\bullet) - T8(\bullet)] \\ &\quad - C4[2qT1(\bullet) + T2(\bullet)] + \frac{\theta}{\kappa}\left[\nabla^2 L_{12} - \frac{1}{\kappa}\nabla L_2\right]\end{aligned}$$

References

[1] Abramowitz, M., and I. Stegun: *Handbook of Mathematical Functions*, New York: Dover Publications, 1972.

[2] Bard, Y.: *Nonlinear Parameter Estimation*, New York: Academic Press, 1974.

[3] Basawa, I.V. and B.L.S. Prakasa Rao: *Statistical Inference for Stochastic Processes*, London: Academic Press, 1980.

[4] Brennan, M.J., and E.S. Schwartz: "An Equilibrium Model of Bond Pricing and a Test of Market Efficiency", *Journal of Financial and Quantitative Analysis*, 17(1982), 301-29.

[5] Brennan, M.J., and E.S. Schwartz: "Alternative Methods for Valuing Debt Options", *Finance* 4(1983) 119-37.

[6] Brown, S.J., and P. Dybvig: "The Empirical Implications of the Cox, Ingersoll, and Ross Theory of the Term Structure of Interest Rates", *Journal of Finance*, 41(1986), 617-30.

[7] Cox, J.C., J.E. Ingersoll, and S. Ross: "A Theory of the Term Structure of Interest Rates", *Econometrica*, 53(1985), 385-408.

[8] Cox, J.C., and S. Ross: "The Valuation of Options for Alternative Stochastic Processes", *Journal of Financial Economics* 3 (1976), 145-166.

[9] Cox, J.C., J.E. Ingersoll, and S. Ross: "An Intertemporal General Equilibrium Model of Asset Prices," *Econometrica*, 53(1985), 363-84.

[10] Dennis, J.E., and R. Schnabel: *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*, Englewood Cliffs: Prentice-Hall, 1983.

[11] Dietrich-Campbell, B., and E.S. Schwartz, "Valuing Debt Options", *Journal of Financial Economics*, 16(1986), 321-43.

[12] Dothan, M., S. Ramamurtie, and S. Ulman, "Applying Economic Restrictions to Foreign Exchange-Rate Dynamics: Spot Rates, Futures, and Options", Federal Reserve Bank of Atlanta Working Paper Series, March 1996.

[13] Feller, W.: "Two Singular Diffusion Problems", *Annals of Mathematics*, 54(1951), 173-82.

[14] Gibbons, M., and K. Ramaswamy: "The Term Structure of Interest Rates: Empirical Evidence", unpublished, 1987.

[15] Heston, S.: "Testing Continuous Time Models of the Term Structure of Interest Rates", November 1988, Carnegie Mellon University working paper.

[16] Karlin, S., and H. Taylor: *A Second Course in Stochastic Processes*, New York: Academic Press, 1981.

[17] Marsh, T., and E. Rosenfeld: "Stochastic Processes for Interest Rates and Equilibrium Bond Prices", *Journal of Finance*, 38 (1983), 635-46.

[18] Prezas, A. and S. Ulman: "Forecasting the Overnight RP Rate: An Application of Nonlinear Maximum-Likelihood Methods to a Class of Diffusion Processes", University of Minnesota Supercomputer Institute Research Report UMSI 89/73, April 1989.

[19] Ramamurtie, S. and S. Ulman, "Specifying a Consistent Joint Maximum-likelihood Approach to Testing Bond Mode working paper, June, 1996.

[20] Richard, S.: "An Arbitrage Model of the Term Structure of Interest Rates", *Journal of Financial Economics*,

6(1978), 33-57.

[21] Ulman, S.: "The Overnight Repo Rate, Bond Prices and Yields: Implications of an Equilibrium Term Structure Model", University of Minnesota Supercomputer Institute Research Report UMSI 88/26, March 1988.

[22] Vasicek, O.: "An Equilibrium Characterization of the Term Structure", *Journal of Financial Economics*, 5(1977), 177-88.

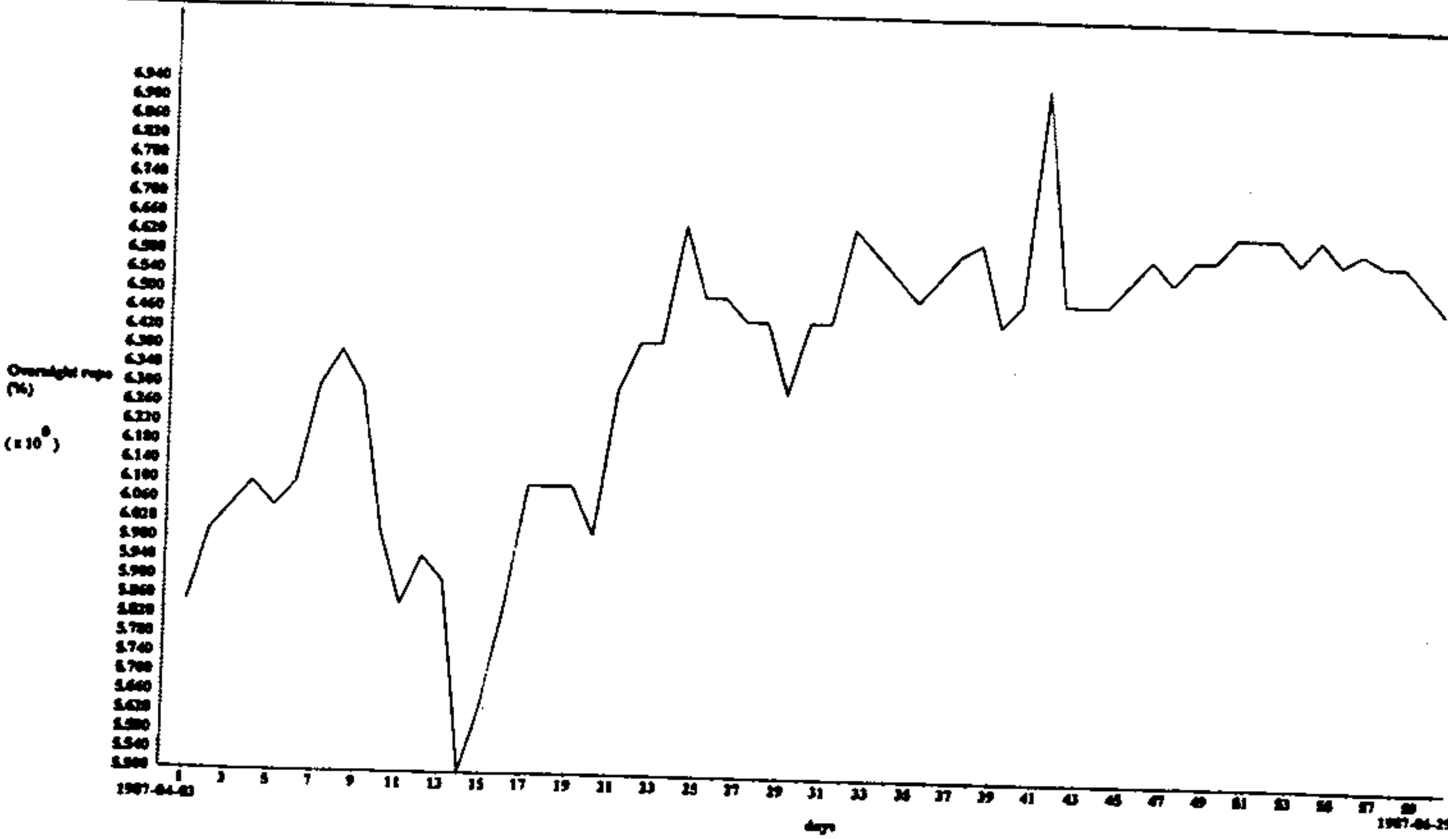
RateTool 2.1 Analytics Package
U.S. Overnight Repo 4/3/87-6/29/87

Mean-Reverting, Square-Root Process

Speed of adjustment: κ = 37.749
$s(\kappa)$ = 17.582
Loglikelihood function value = 303.98634
Sample size: 60

Long-run mean: θ = 0.06444
$s(\theta)$ = 0.00137

Rate volatility: σ = 0.09478
$s(\sigma)$ = 0.00927



Panel 1

Overnight RP rates in the U.S. during mid-1987 using daily sampling and the MRSR norm-preserving density.

Note that parameter estimates are more than double standard errors for all three parameters.

7-day LIBOR (%)
6/22/92 - 9/15/92

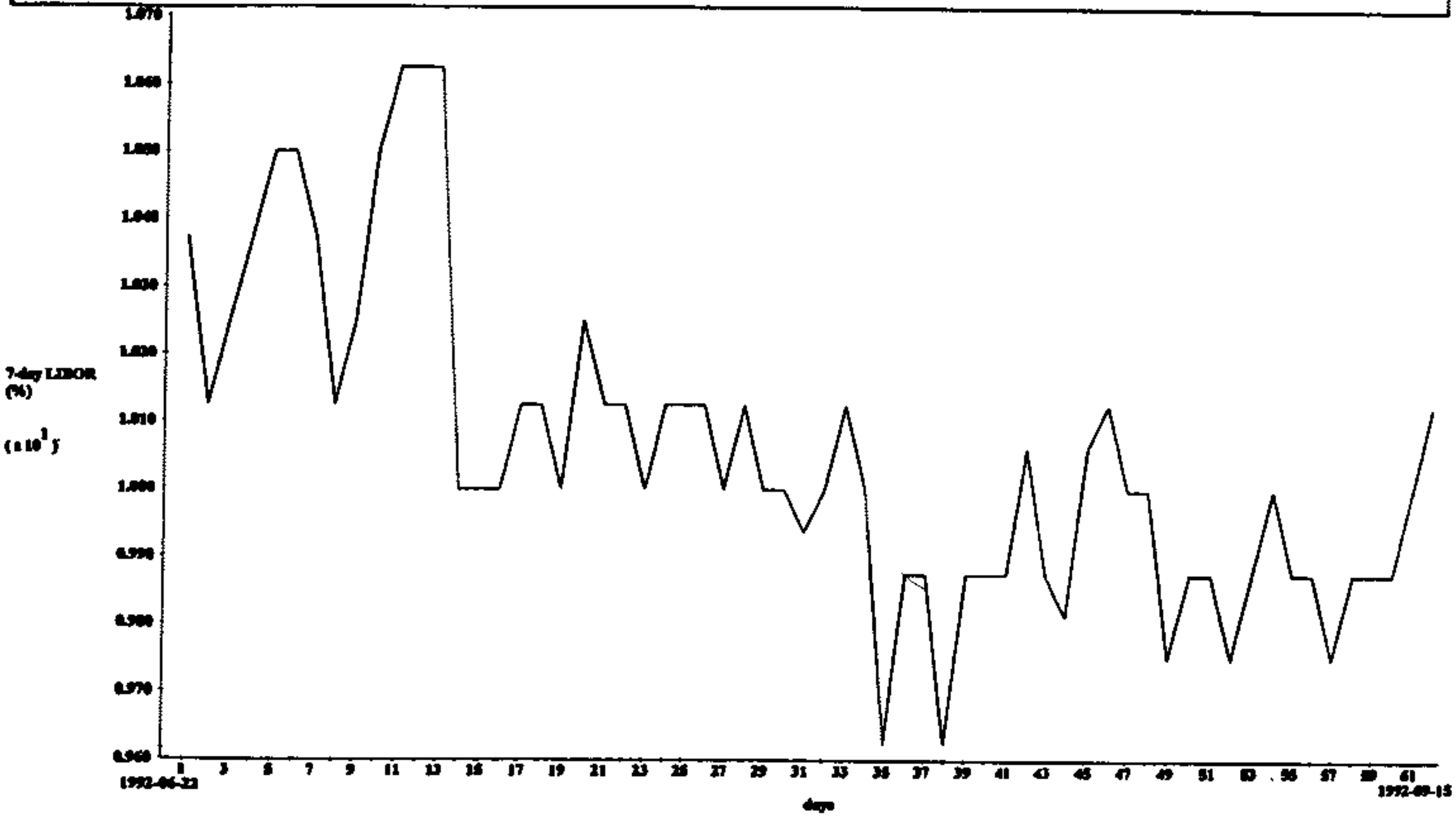
Mean-Reverting, Square-Root Process

Speed of adjustment: κ = 76.416
$s(\kappa)$ = 28.442

Long-run mean: θ = 0.10040
$s(\theta)$ = 0.00073

Rate volatility: σ = 0.08640
$s(\sigma)$ = 0.00897

Loglikelihood function value = 310.32866
Sample size: 62



Panel 2
London Interbank Offer Rates (LIBOR) during summer 1992 using daily sampling and the MRSR norm-preserving density.

Note that parameter estimates are more than double standard errors for all three parameters.

The speed of adjustment estimator κ is considerably larger than its counterpart for Overnight RP in Panel 1 due to the toothy appearance of the time series.

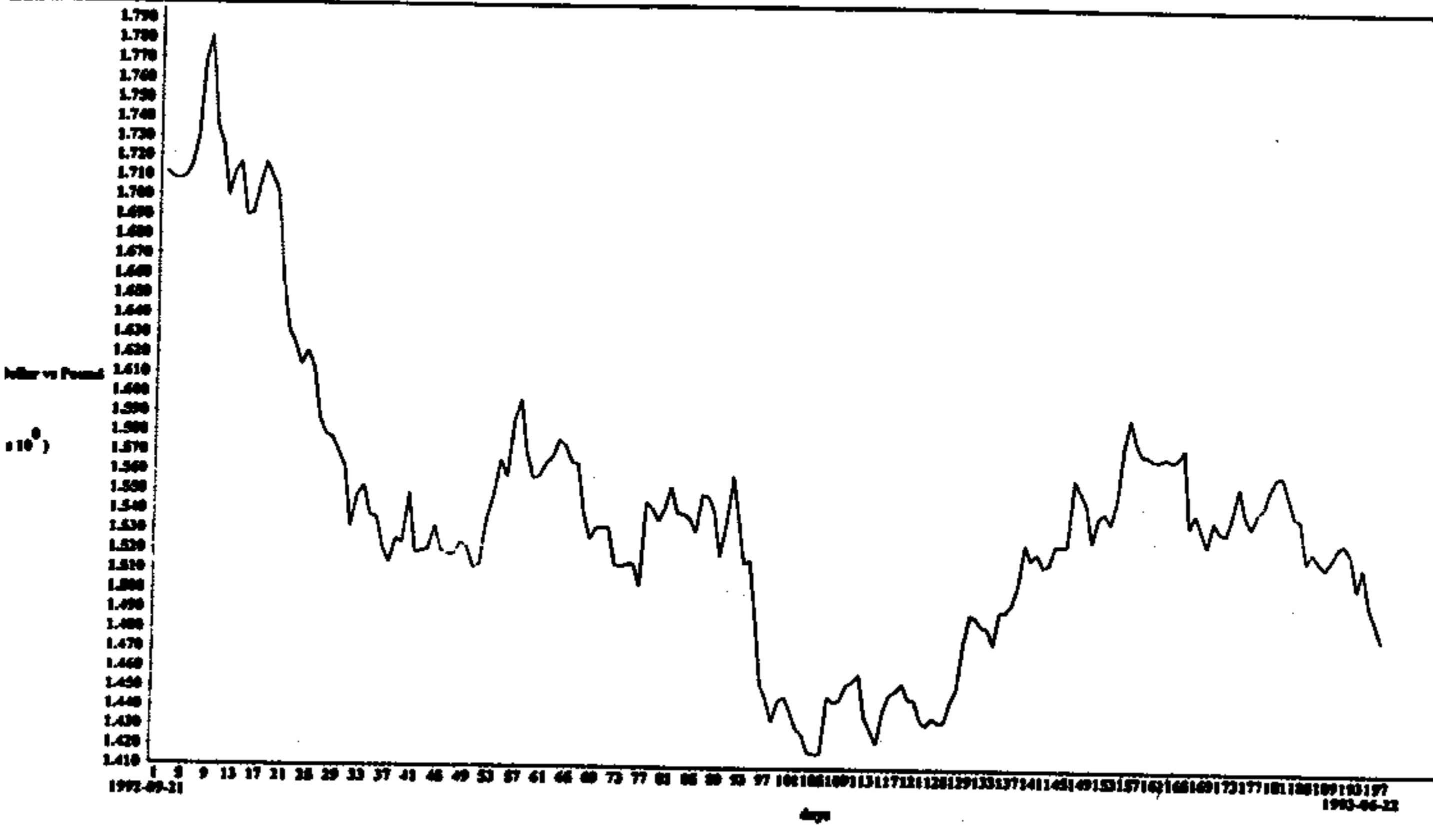
Dollar Versus British Pound
9/21/92-6/22/93 (MRSR)

Mean-Reverting, Square-Root Process

Speed of adjustment: κ = 7.653
$s(\kappa)$ = 3.458
Loglikelihood function value = 564.70549
Sample size: 197

Long-run mean: θ = 1.50056
$s(\theta)$ = 0.03623

Rate volatility: σ = 0.17630
$s(\sigma)$ = 0.00898



Panel 3
FX: Dollar vs. British Pound
(Sept. 1992 - June 1993) using daily sampling and the MRSR norm-preserving density.

Note that parameter estimates are more than double standard errors for all three parameters.

However, with an MRSR specification for this time series, the inverse rate (Pound vs. Dollar) would not have an MRSR distribution. This is counterintuitive at best.

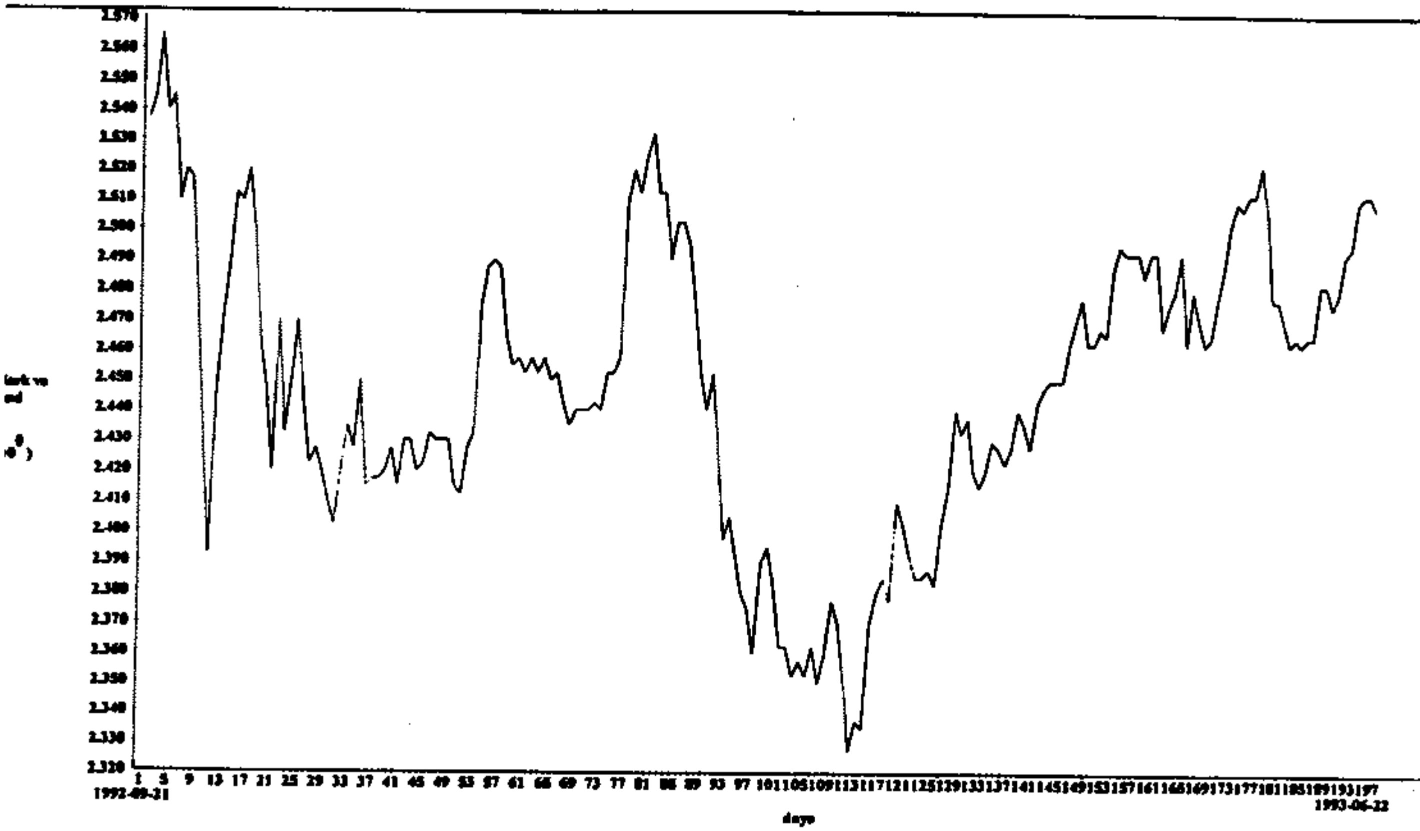
FX: Deutschmark versus British Pound
9/21/92-6/22/93 (MRSR)

Mean-Reverting, Square-Root Process

peed of adjustment: κ = 15.595
$s(\kappa)$ = 6.256
oglikelihood function value = 536.48687
ample size: 197

Long-run mean: θ = 2.44600
$s(\theta)$ = 0.01866

Rate volatility: σ = 0.16391
$s(\sigma)$ = 0.00852



Panel 4
FX: Deutschmark vs. British Pound (Sept. 1992 - June 1993) using daily sampling and the MRSR norm-preserving density.

Note that parameter estimates are more than double standard errors for all three parameters.

However, with an MRSR specification for this time series, the inverse rate (Pound vs. Deutschmark) would not have an MRSR distribution. This is counterintuitive at best.

Dollar Versus British Pound
9/21/92-6/22/93 (critical floor=1.1)

Mean-Reverting, Logarithmic Process

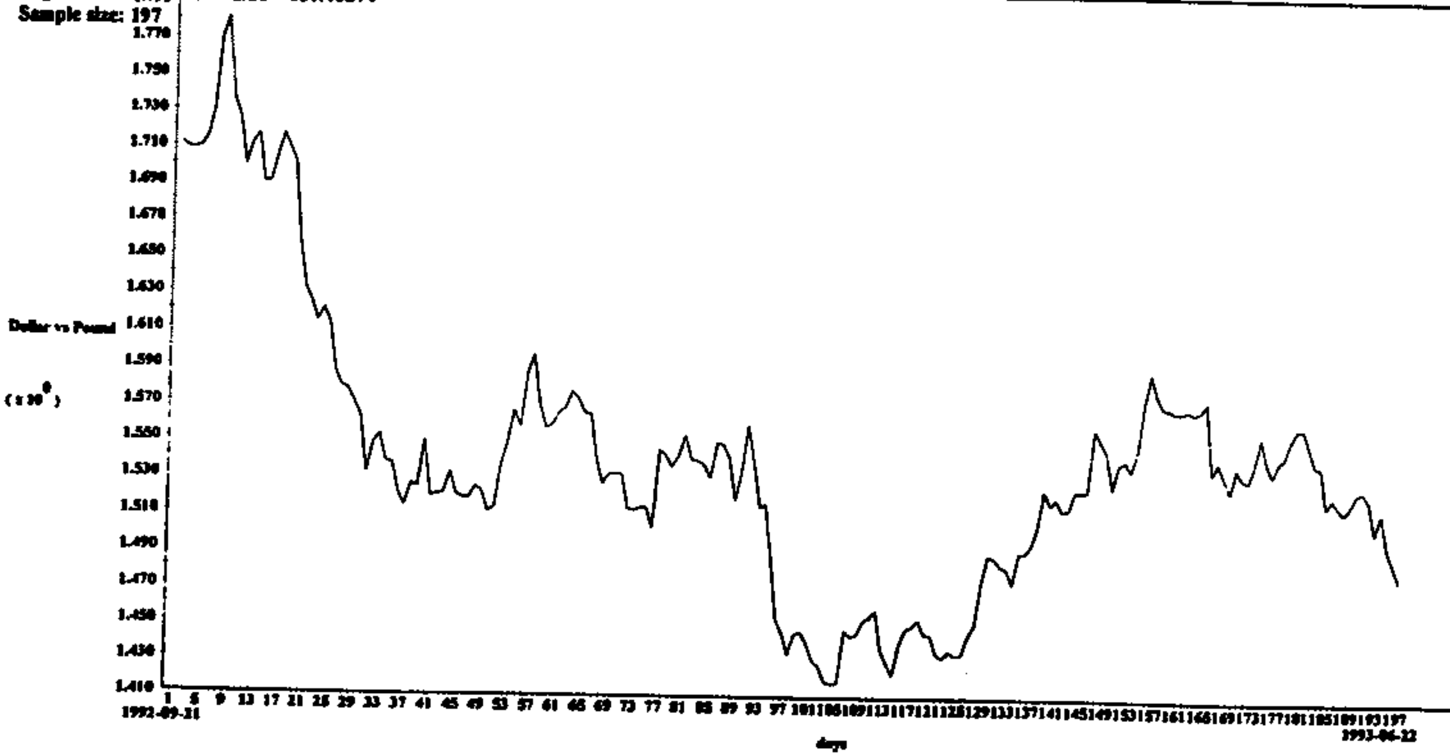
Speed of adjustment: N3 = 0.474
s(N3) = 1.242
Critical floor: N2 = 1.100

Log long-run mean: N4 = 1.1631
s(N4) = 0.50104

Rate volatility: N1 = 0.22998
s(N1) = 0.01085

Slow convergence! Desired tolerance NOT reached! Restart optimizer from different initial guesses!

Loglikelihood function value = 650.40574
Sample size: 197



Panel 6
Slow Convergence.
Occasionally, the optimizer becomes stuck at a local minimum with the default guesses. It then prints a message like this. The user must alter initial guesses in a dialog box and try again.

The convergent series is shown in Panel 6 (merely started from a different set of initial guesses).

Dollar Versus British Pound
9/21/92-6/22/93 (critical floor=1.1)

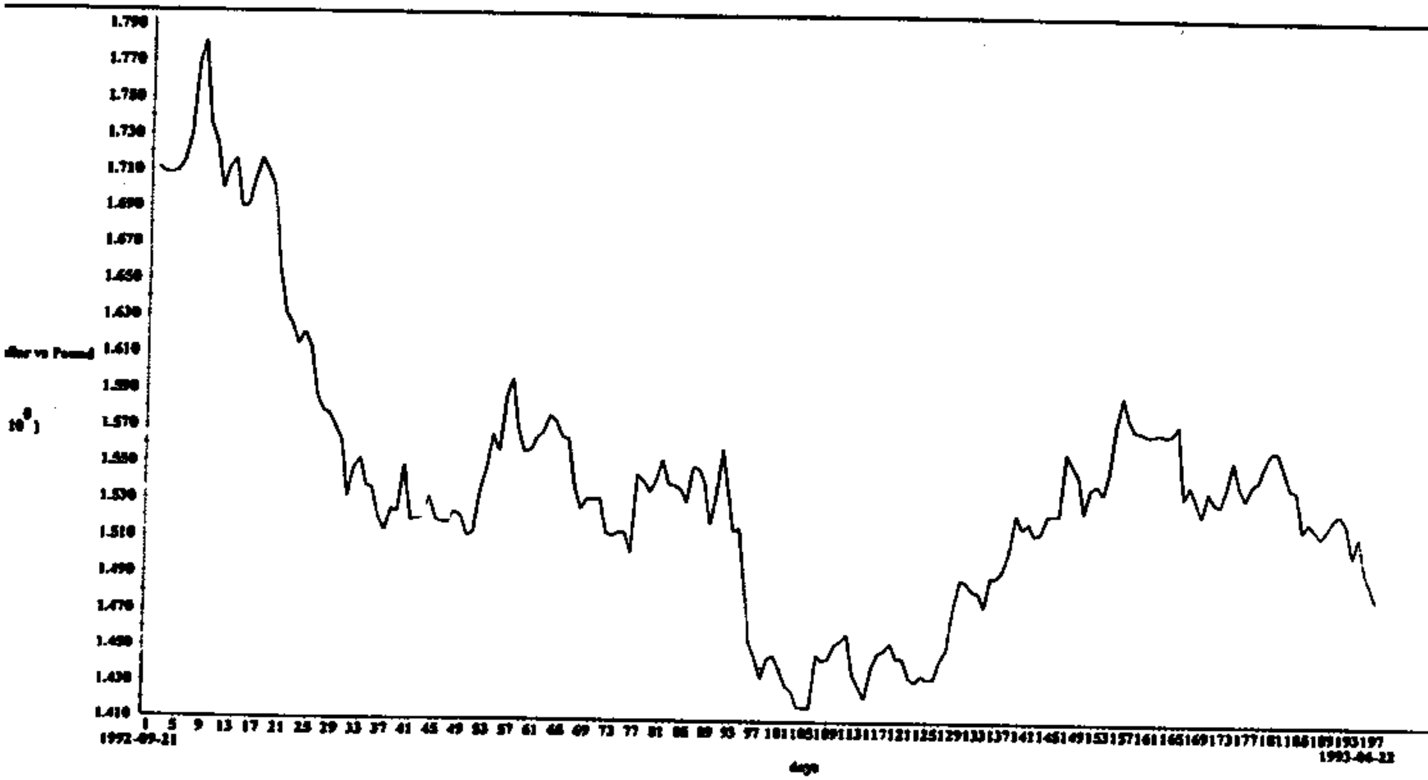
Mean-Reverting, Logarithmic Process

Speed of adjustment: N3 = 7.712
s(N3) = 3.535
Critical floor: N2 = 1.100

Log long-run mean: N4 = 1.5009
s(N4) = 0.00049

Rate volatility: N1 = 0.24185
s(N1) = 0.01233

Loglikelihood function value = 652.99032
Sample size: 197



Panel 6
FX: Dollar vs British Pound
(Sept. 1992 - June 1993) using daily sampling and the MRL norm-preserving density with a critical floor exogenously specified as 1.1.

All parameter estimates appear statistically significant (this is the same time series as in Panels 3 and 5).

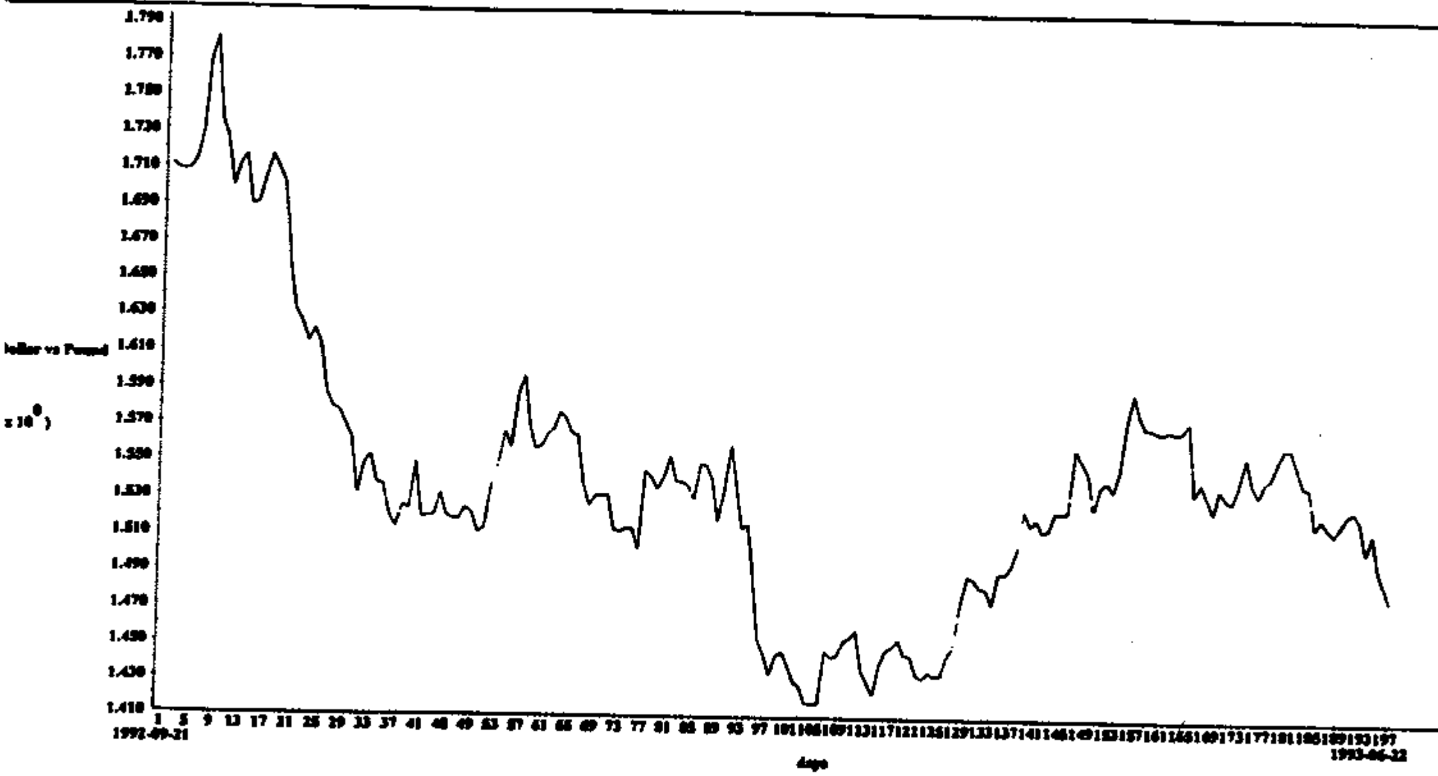
Compare results against the MRSR process from Panel 3.

Dollar Versus British Pound
9/21/92-6/22/93 (critical floor=1.2)

Mean-Reverting, Logarithmic Process

Speed of adjustment: N3 = 7.760	**Log long-run mean:** N4 = 1.5012	**Rate volatility:**	N1 = 0.28164
s(N3) = 3.554	s(N4) = 0.00046		s(N1) = 0.01436
Critical floor: N2 = 1.200			

Loglikelihood function value = 653.36631
Sample size: 197



Panel 7
FX: Dollar vs British Pound (Sept. 1992 - June 1993) using daily sampling and the MRL norm-preserving density with a critical floor exogenously specified as 1.2.

The major impact of raising the critical floor from 1.1 (Panel 6) is to increase the estimator for the volatility parameter.

With the MRL process, the inverse rate (Pounds vs Dollars) will also have an MRU distribution so users are indifferent about the direction of the analysis.

Dollar Versus British Pound
9/21/92-6/22/93 (critical floor=1.3)

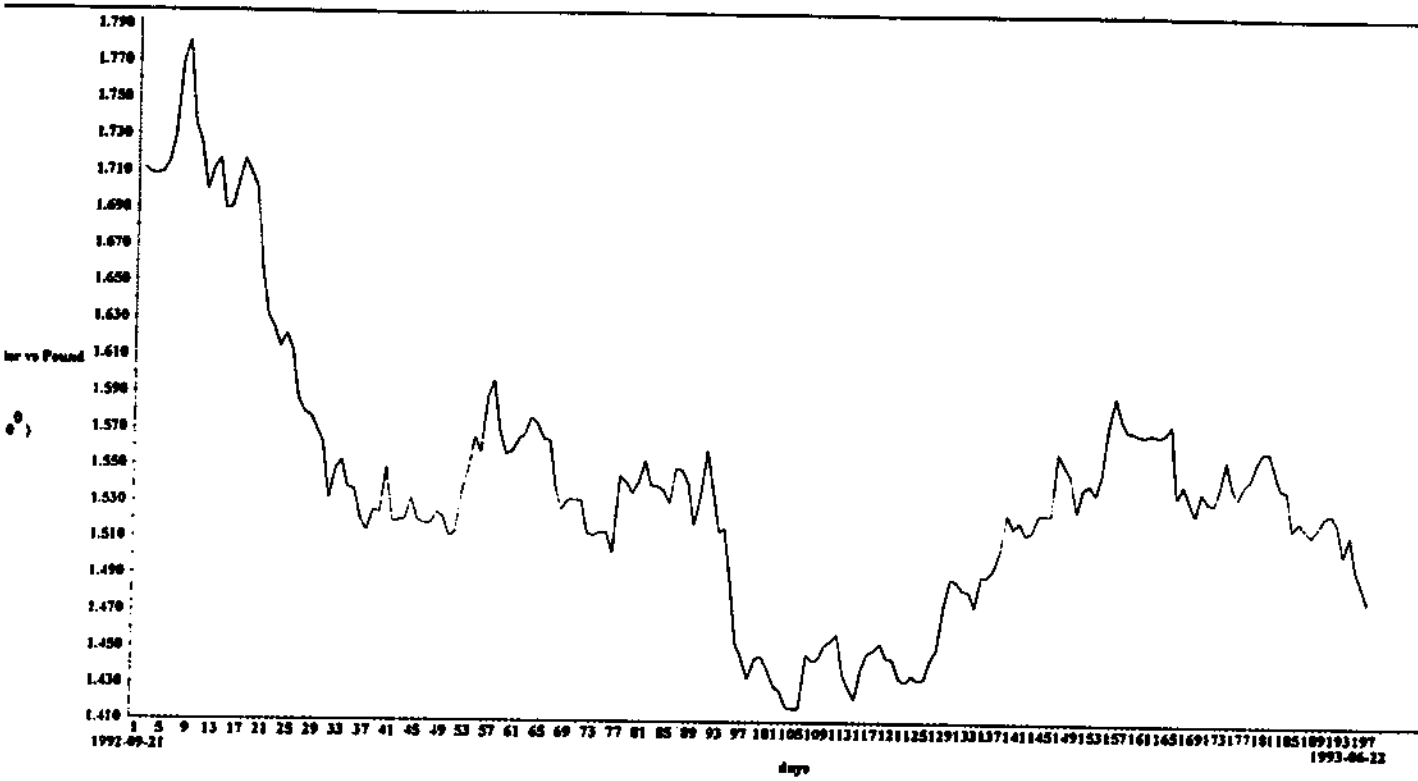
Mean-Reverting, Logarithmic Process

peed of adjustment: N3 = 7.769
s(N3) = 3.566
ritical floor: N2 = 1.300

Log long-run mean: N4 = 1.5012
s(N4) = 0.00043

Rate volatility: N1 = 0.34570
s(N1) = 0.01763

oglikelihood function value = 653.39520
imple size: 197



Panel 8
FX: Dollar vs British Pound (Sept. 1992 - June 1993) using daily sampling and the MRL norm-preserving density with a critical floor exogenously specified as 1.3.

The estimator for the volatility parameter appears to rise at an accelerating rate as the critical floor rises.

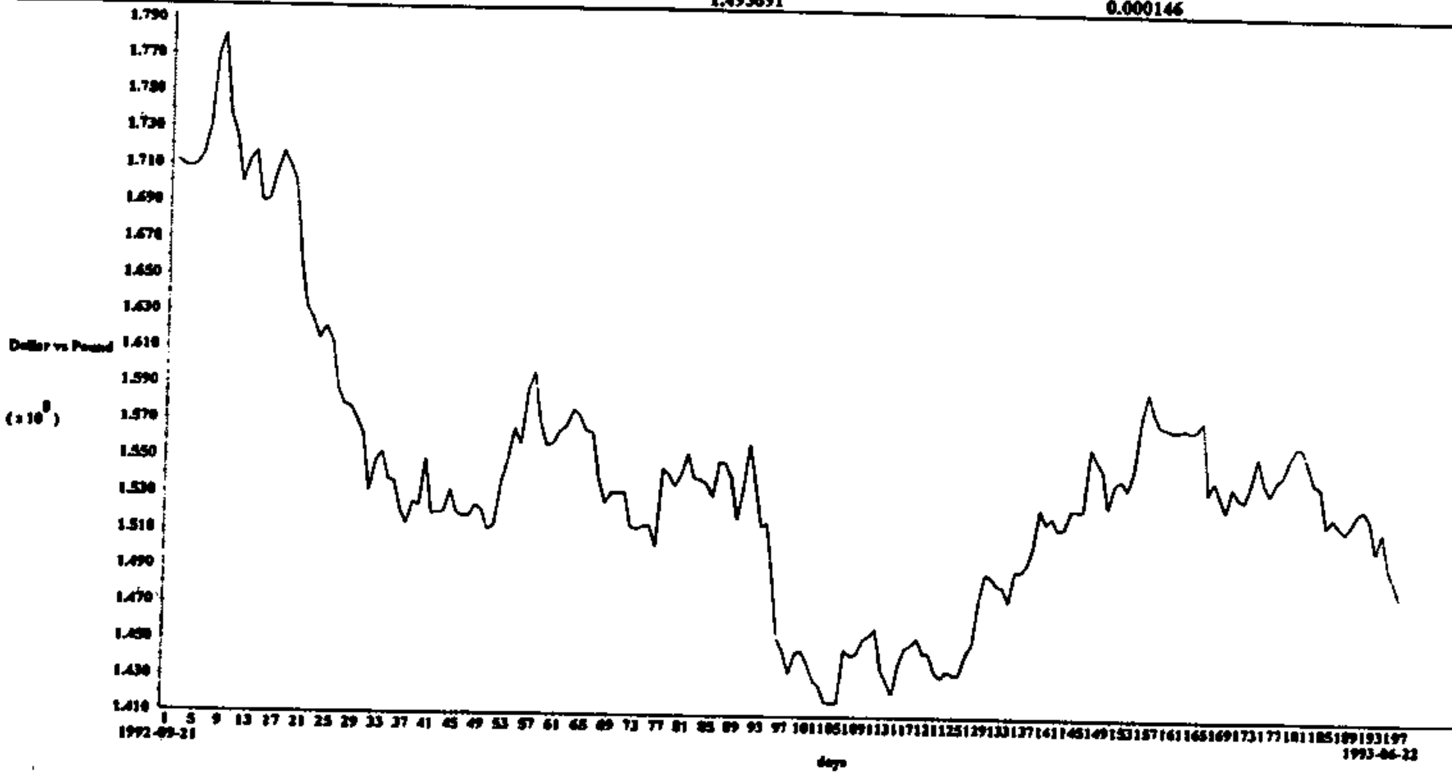
With the MRL process, the inverse rate (Pounds vs Dollars) will also have an MRL distribution so users are indifferent about the direction of the analysis.

Dollar Versus British Pound
9/21/92-6/22/93 (critical floor=1.3)

Mean-Reverting, Logarithmic Process

Forecasts using estimators of the parameter set: N3 = 7.769 N4 = 1.5012 N1 = 0.34570 N2 = 1.300
and a current rate r(s) = 1.47800 on

Weeks Ahead	Expected Rate	Variance
3	1.486785	0.000095
4	1.488923	0.000113
5	1.490759	0.000127
6	1.492336	0.000138
7	1.493691	0.000146



Panel 9
Forecasting FX: Dollar vs British Pound
(Sept. 1992 - June 1993) using daily sampling and the MRL norm-preserving density with a critical floor exogenously specified as 1.3.

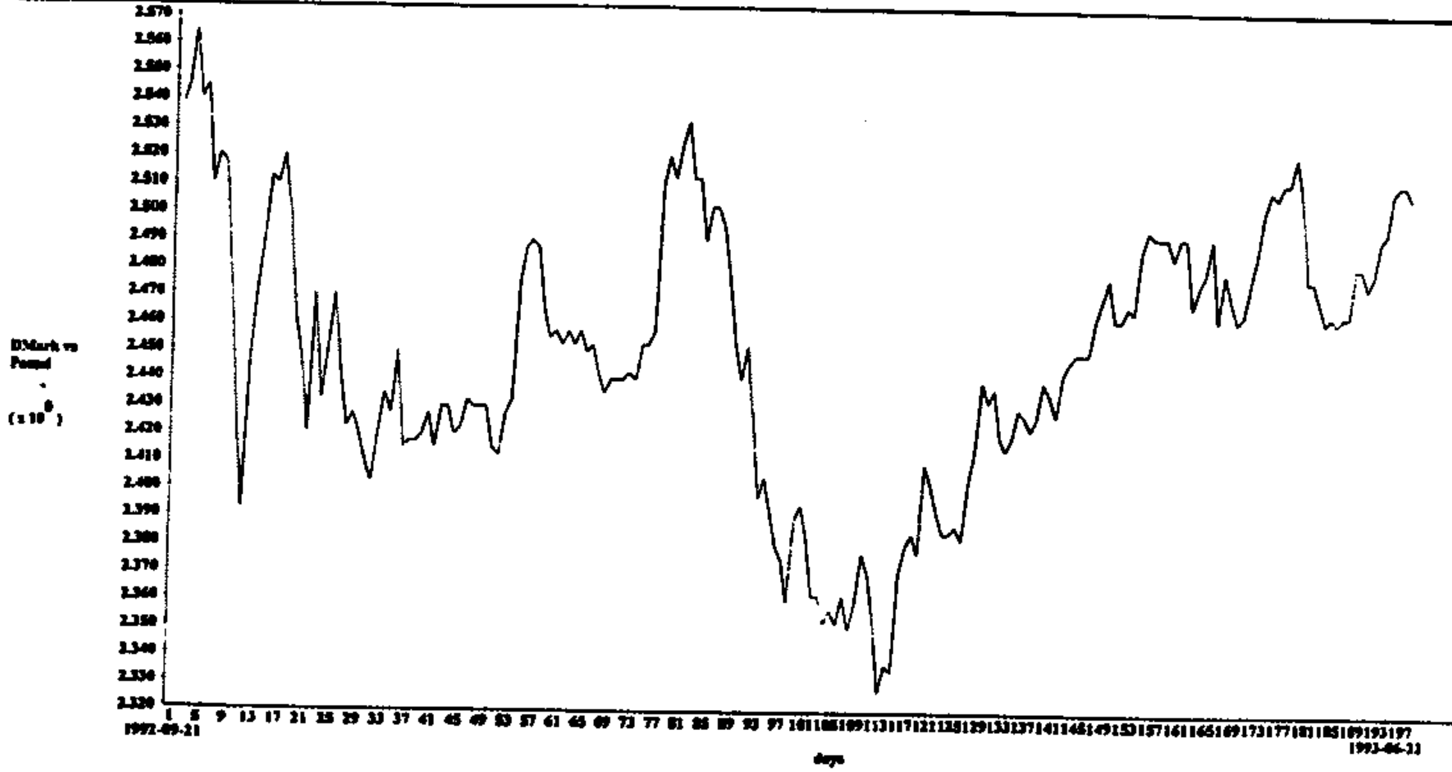
Note the forecast is for N weeks ahead from the last observation in the time series.

FX: Deutschmark versus British Pound
9/21/92-6/22/93 (critical floor = 1.96)

Mean-Reverting, Logarithmic Process

Speed of adjustment: N3 = 15.047	**Log long-run mean:** N4 = 2.4455	**Rate volatility:**	N1 = 0.22321
s(N3) = 6.163	s(N4) = 0.00006		s(N1) = 0.01159
Critical floor: N2 = 1.960			

Loglikelihood function value = 711.24755
Sample size: 197



Panel 10
FX: Deutschmark vs British Pound
(Sept. 1992 - June 1993) using daily sampling and the MRL norm-preserving density with a critical floor exogenously specified as 1.96.

Compare results with the MRSR process in Panel 4.

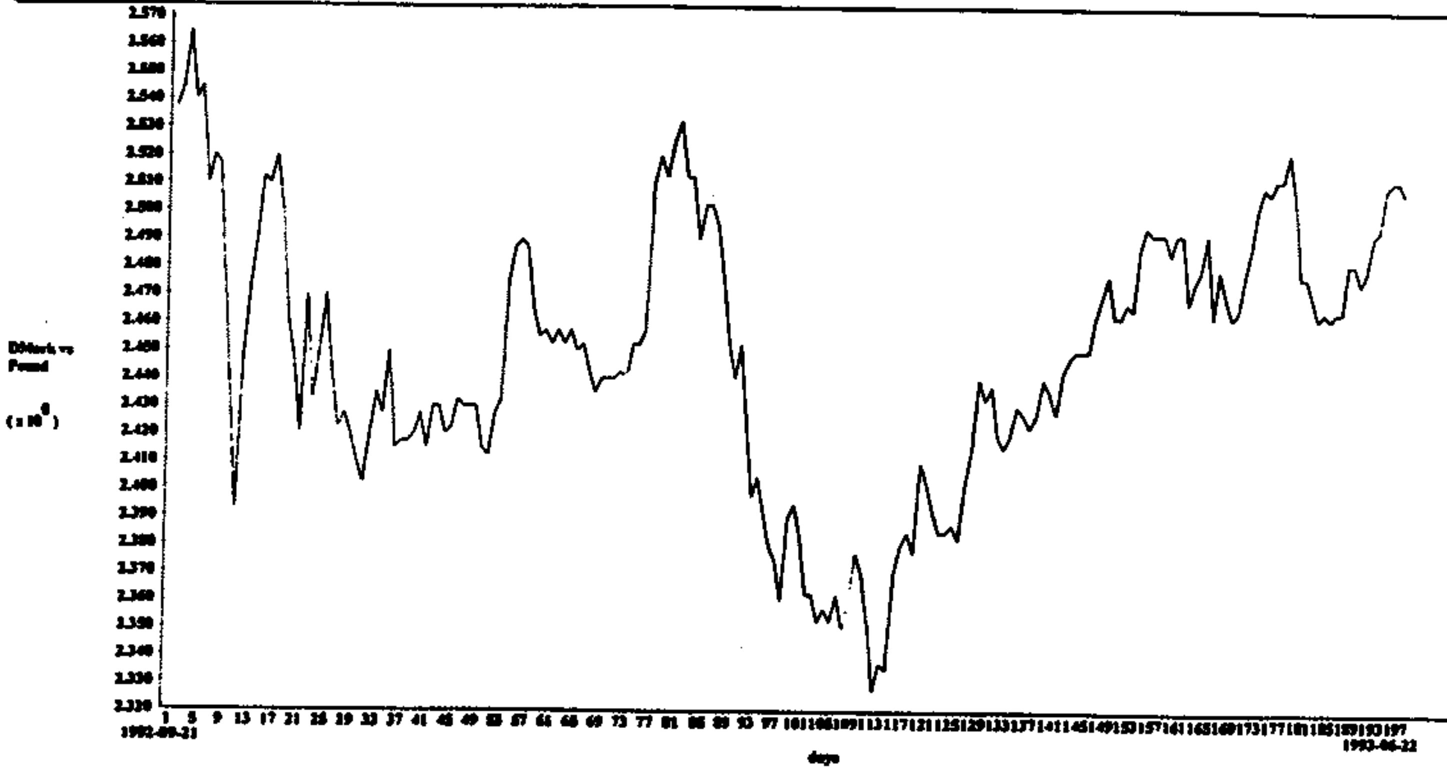
With the MRL process, the inverse rate (Pounds vs Dollars) will also have an MRL distribution so users are indifferent about the direction of the analysis.

FX: Deutschmark versus British Pound
9/21/92-6/22/93 (critical floor = 2.0)

Mean-Reverting, Logarithmic Process

Speed of adjustment: N3 = 14.985	Log long-run mean: N4 = 2.4455	Rate volatility:	N1 = 0.23431
s(N3) = 6.152	s(N4) = 0.00006		s(N1) = 0.01216
Critical floor: N2 = 2.000			

Loglikelihood function value = 711.13511
Sample size: 197



Panel 11
FX: Deutschmark vs British Pound
(Sept. 1992 - June 1993) using daily sampling and the MRL norm-preserving density with a critical floor exogenously specified as 2.0.

As in Panel 7, the major impact of the higher critical floor is an increase in the estimator for the volatility parameter.

FX: Deutschmark versus British Pound
1/7/93 - 4/21/93 (75 observations)

Mean-Reverting, Logarithmic Process

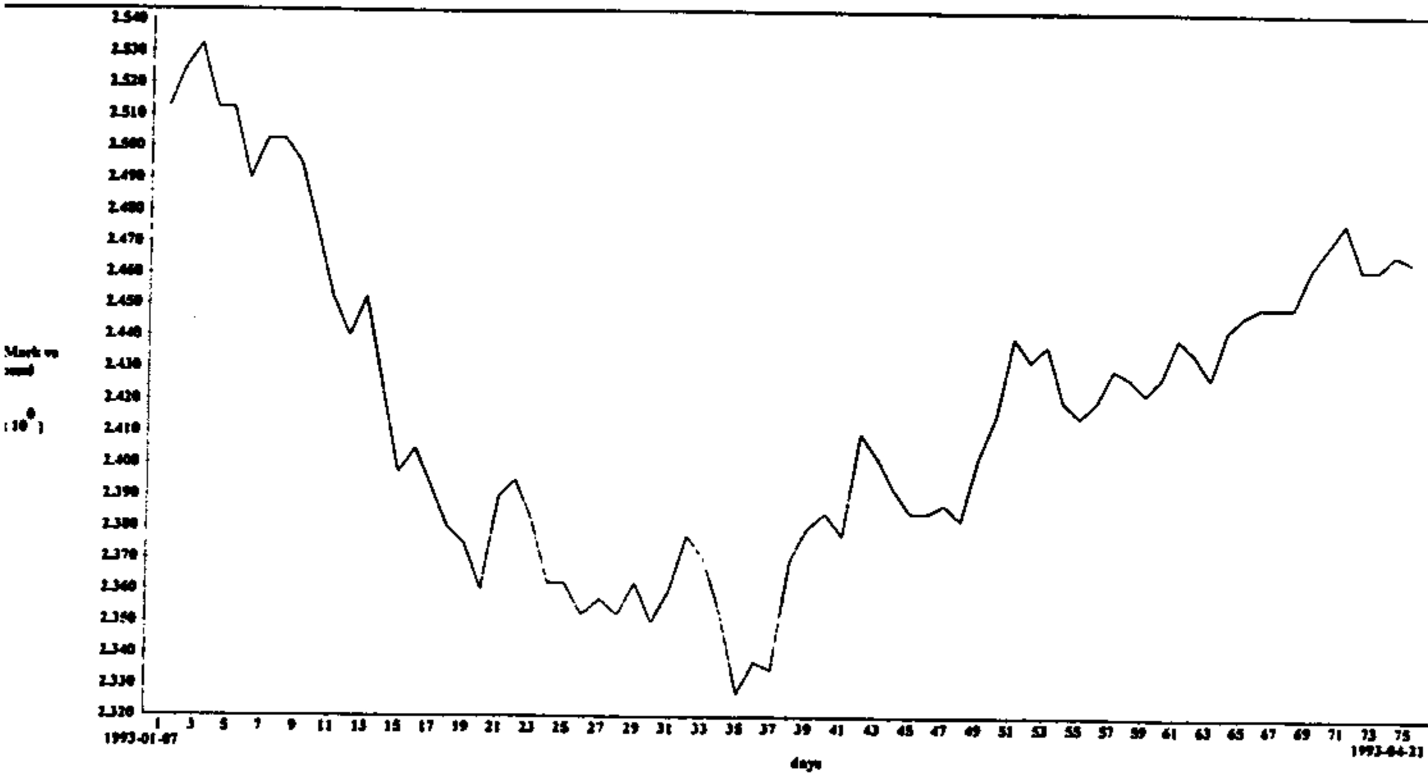
Speed of adjustment: N3 = 14.368
s(N3) = 8.460

Critical floor: N2 = 1.960

Log long-run mean: N4 = 2.4066
s(N4) = 0.00014

Rate volatility: N1 = 0.19905
s(N1) = 0.01669

Loglikelihood function value = 279.08475
Sample size: 75



Panel 12
FX: Deutschmark vs British Pound (data subset)
(Jan. 1993 - April 1993) using daily sampling and the MRL norm-preserving density with a critical floor exogenously specified as 1.96

This panel shows the results of picking a subset of the data from Panel 10. Data was selected from the area near the time series' low point.

File: ratetest.csv
Field 2: Japanese Yen to Pound Sterling

Mean-Reverting, Square-Root Process

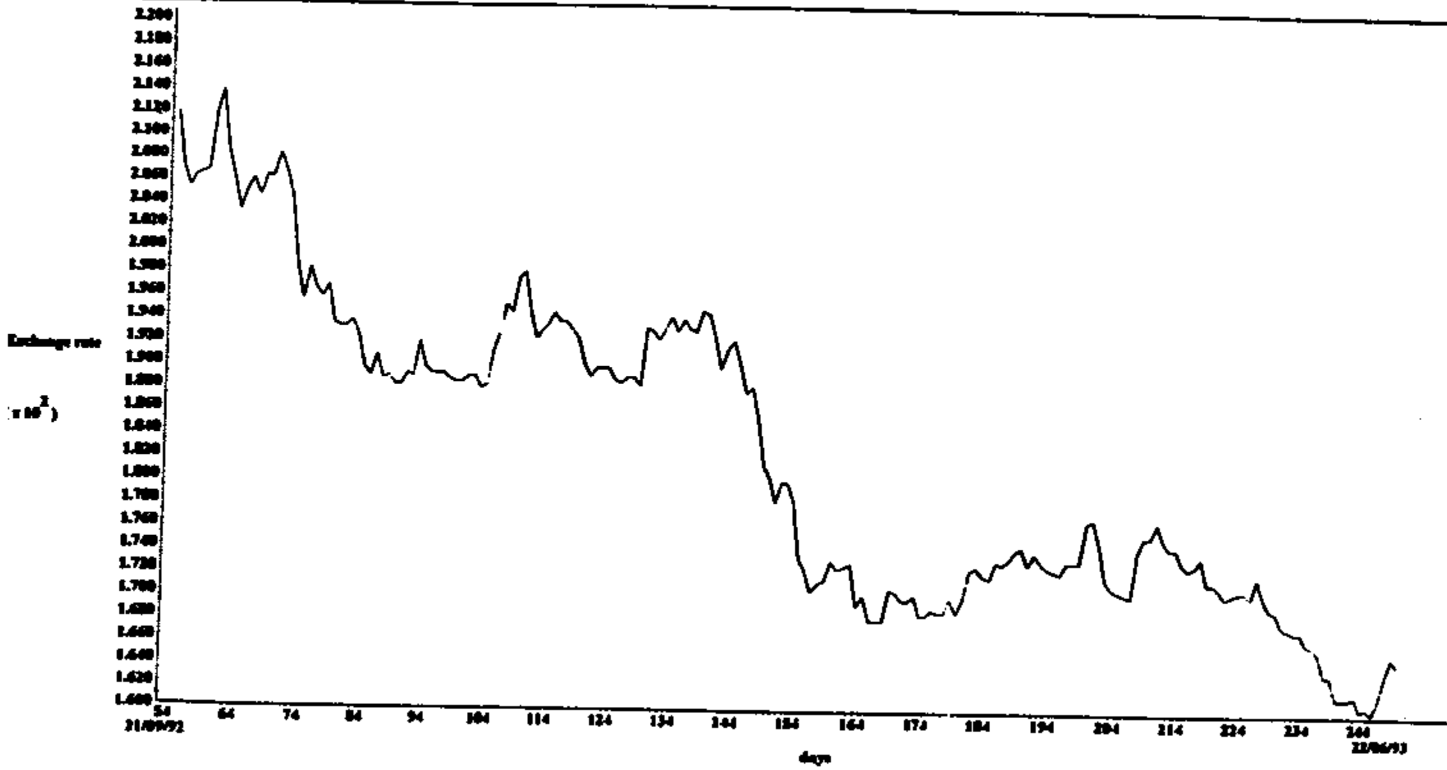
Speed of adjustment: $\kappa =$ 3.939
$s(\kappa) =$ 2.243
Loglikelihood function value = -371.39645
Sample size: 197

Long-run mean: $\theta =$ 166.87992
$s(\theta) =$ 11.07547

Starting: 21/09/92

Rate volatility: $\sigma =$ 1.90823
$s(\sigma) =$ 0.09664

Ending: 22/06/93



Panel 13
FX: Yen vs British Pound (Poor Fit)
(Sept. 1992 - June 1993) using daily sampling and the MRSR norm-preserving density.

This panel shows that the optimizer can converge to a result with high standard errors and a very low value for the log likelihood function.

Note that the time series appears to be drifting inexorably downward and displays no mean-reverting traits. Hence the selected time series is not consistent with MRSR. Better practical results might be obtained by hypothesizing a lognormal process with negative drift.