Linear or Non-Linear: This is the Dilemma!

Paolo Pasquariello*
Stern School of Business
New York University

March, 12 2000

Abstract

Estimation of the unknown drift and diffusion terms of a stochastic process representing the spot interest rate has received increased attention in the financial literature. In this brief paper, we analyze the performance of the two most popular non-parametric approaches to the problem, Sahalia’ s FGLS and Stanton’ s Taylor Expansion. An outright estimation of each of the proposed models to a sample of 22 years of daily observations of the 7-days eurodollar spot interest rate seems to suggest the existence of non-linearity in both the reversion and the volatility of the rate process. However, the results of a series of simulations from a known process reveal that none of the methods here examined has enough power to recognize a “true” linear drift, i.e. that the linearity of the interest rate drift component is rejected “too often”. More accuracy is however found for the estimated diffusion. We are then inclined to confirm Sahalia’ s claim that the volatility of the spot rate appears to be increasing with the level of the rate.

* Ph.D. candidate at the Stern School of Business. Please address comments to the author at the Leonard N. Stern School of Business, New York University, Kaufman Management Education Center, Suite 9-190, 44 West 4th Street, New York, NY 10012-1126, or through email: ppasquar@stern.nyu.edu.
1. Introduction

The correct specification of the spot interest rate as a stochastic process and a proper and efficient estimation of the resulting drift and diffusion terms have always been puzzling practitioners and traders pricing complex derivative securities, before finally capturing the attention and the efforts of the financial literature in the past few years. It has in fact been argued that an incorrect specification of the potential fluctuations of the underlying random interest rate would induce incorrect valuation of contingent claims written on bonds or currencies.

Three main approaches to the estimation of the drift and diffusion term of a general stochastic process of the form:

\[ dr_t = \mu(r_t, t)dt + \sigma(r_t, t)dw_t \]  \[ 1 \]

can be by now identified. In one case, given a specific structure for \( \mu(r) \) and \( \sigma(r) \), maximum-likelihood estimators for the resulting unknown parameters have been obtained by applying the Forward Kolmogorov Equation to derive an expression for the transition density for the data. This approach, usually known as Parametric and pioneered by Lo(1988), Pearson and Sun(1994) and more recently Sahalia (1999), provides the clear advantage of always leading to efficient estimates for the assumed structure, but is unable to distinguish between alternative structural specifications for the drift and diffusion of the rate process. In the attempt to reduce the number and significance of the structural and distributional restrictions needed to identify the Likelihood function for \( r_t \), Chan et al. (1992), Duffie and Singleton (1993) and Hansen and Scheinkman (1995) propose to estimate the given \( \mu(r) \) and \( \sigma(r) \) by matching a set of moment conditions with the corresponding estimates from the observed data. This Method of Moments attenuates, but does not eliminate the risk of mis-specification that affects any heavily structural estimation procedure. Hence, the Non-Parametric approach of Sahalia (1996 a, b) and Stanton (1997) appeared to be particularly promising. In one case, Sahalia specifies an apparently very general form for \( \mu(r) \) and \( \sigma(r) \) and shows that none of the more specific structures that are typically referred to in the literature, from the CIR process to the CEV diffusions, seems to really fit the data. Stanton uses stochastic Taylor series to generate even more general and structure-less point-wise estimates for \( \mu(r) \) and \( \sigma(r) \). Both methods are not affected by the limitations resulting from the estimation of a continuous process via discretely sampled data, but rely heavily on non-parametric estimation of the marginal and/or transitional density of the underlying and unknown process for \( r_t \) using gaussian Kernels. Unfortunately, not many of the initial promises were fulfilled. Non-parametric Kernel estimation of densities is highly imprecise, and the strong persistence in the interest rate data limits the amounts of “real” information that any of those methods can extract from the available observations.
In this paper, we estimate the drift and the diffusion of the general stochastic process of equation [1] applying the FGLS method of Sahalia and the Stanton’s approach to a data set of 5505 observations for the 7-day spot Eurodollar rate from June 1st 1973 to February 25th 1995. We then simulate series of interest rates for the same time-frame using a CIR diffusion and show that both methods fail to recognize the “true” drift and diffusion terms embedded in the constructed series unless the estimation is repeated several times, i.e. that both approaches show low power. An out-of-sample analysis of the estimates generated with Sahalia’s and Stanton’s procedures is provided. The simulated interest rate series in both cases show more variability than what was effectively observed in the four years following the original sample. In the next section, we describe the Sahalia’ s FGLS estimation. Then, in section 3, we describe and implement the Stanton’s estimation. Section 4 evaluates the power of the two strategies with repeated simulations. Section 5 concludes.

2. The FGLS Estimation

Sahalia (1996 b) proposes the following general structure for the drift and the diffusion terms in equation [1]:

$$dr_t = \left[ \alpha_0 + \alpha_1 r_t + \alpha_2 r_t^2 + \frac{\alpha_3}{r_t} \right] dt + \left[ \beta_0 + \beta_1 r_t + \beta_2 r_t^2 \right] dW_t$$  \[2\]

The original non-parametric estimation of the vector of parameters ($\alpha$, $\beta$) is carried over by observing that any specification for the drift and the diffusion, like in equation [2], implies a one-to-one mapping with the marginal and transitional densities for $r_t$. For the purposes of this paper, this estimation is achieved through the following FGLS multi-step procedure:

- **Regress:**

  $$E[r_{t+1} - r_t / r_t] = \alpha_0 + \alpha_1 r_t + \alpha_2 r_t^2 + \frac{\alpha_3}{r_t} + \epsilon_{t+1}$$  \[3\]

- **Estimate:**

  $$\hat{\epsilon}_{t+1} = E[r_{t+1} - r_t / r_t] - \hat{\alpha}_0 + \hat{\alpha}_1 r_t + \hat{\alpha}_2 r_t^2 + \frac{\hat{\alpha}_3}{r_t}$$  \[4\]

- Use Non-Linear Least Squares to estimate $\hat{\alpha}$.

---

1 The starting values necessary to generate NLSQ estimates for the betas are obtained by assuming that initially $\beta_3 = 2$, and then by estimating the following functional form $f(.) = \beta_0 + \beta_1(t^3) + \beta_2(t^2)$. We then assume that the initial $\beta_3 = 1$ for the non-linear regression described in equation [5].
\[ E[\varepsilon^2_t / r_t] = \beta_0 + \beta_1 r_t + \beta_2 r_t^2 + u_{t+1} \quad [5] \]

- Calculate:
\[ \hat{\varepsilon}^2_{t+1} = \hat{\beta}_0 + \hat{\beta}_1 r_t + \hat{\beta}_2 r_t^2 \quad [6] \]

- Use the estimated squared residuals of equation [6] to estimate \( \Sigma \), the covariance matrix for \( \varepsilon \). We assume that the residuals are heteroscedastic but not autocorrelated, i.e. that \( \Sigma \) has the form:
\[
\Sigma = \begin{bmatrix}
\sigma_1^2 & 0 & \cdots & 0 \\
0 & \sigma_2^2 & \cdots & \vdots \\
\vdots & \vdots & \ddots & \vdots \\
0 & \cdots & \cdots & \sigma_T^2
\end{bmatrix}
\quad [7]
\]
\[ \hat{\sigma}^2_{t+1} = \hat{\varepsilon}^2_{t+1} \]

- Estimate the alphas in equation [3] by Feasible GLS:
\[
(\alpha_0, \alpha_1, \alpha_2, \alpha_3) = \left( r^\top \hat{\Sigma}^{-1} r \right)^{-1} r^\top \hat{\Sigma}^{-1} (dr) \quad [8]
\]

The estimates for betas obtained from equation [6] and for alphas from equation [8] are reported in Table 1 below.

Table 1: FGLS Estimation for Equation [2]

<table>
<thead>
<tr>
<th>Parameters</th>
<th>FGLS Estimates</th>
<th>Standard Error</th>
<th>t-Stat</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \alpha_0 )</td>
<td>-0.0004</td>
<td>0.00128</td>
<td>-0.344</td>
</tr>
<tr>
<td>( \alpha_1 )</td>
<td>0.0116</td>
<td>0.01848</td>
<td>0.626</td>
</tr>
<tr>
<td>( \alpha_2 )</td>
<td>-0.0784</td>
<td>0.08138</td>
<td>-0.963</td>
</tr>
<tr>
<td>( \alpha_3 )</td>
<td>0.000008</td>
<td>0.00003</td>
<td>0.298</td>
</tr>
<tr>
<td>( \beta_0 )</td>
<td>0.000035</td>
<td>0.00003</td>
<td>1.131</td>
</tr>
<tr>
<td>( \beta_1 )</td>
<td>-0.0024</td>
<td>0.00517</td>
<td>-0.461</td>
</tr>
<tr>
<td>( \beta_2 )</td>
<td>0.0051</td>
<td>0.00368</td>
<td>1.396</td>
</tr>
<tr>
<td>( \beta_3 )</td>
<td>1.3668</td>
<td>0.64946</td>
<td>2.105</td>
</tr>
</tbody>
</table>
As evident from the t-statistic values reported in column 3 of table 1, few of the parameters’ estimates are statistically significant.

The comparison of these results with Sahalia’s suggests that the matching-density approach he designs is more effective in reducing the degree of uncertainty around the resulting estimated parameters. Figure 1 plots the drift and diffusion functions corresponding to the FGLS values of Table 1.

Figure 1: Drift and Diffusion 1973-1995 – FGLS Estimation

![Graph of drift and diffusion functions](image)

When the spot rate is far away from its mean, the estimated drift appears to reverse strongly, but is essentially zero for values of r between 4 and 13%, thus suggesting that in such an interval the interest rate process is a random walk. This finding, if confirmed, would explain why over shorter time-samples the interest rate process appears not to reject the unit root hypothesis. From the diffusion picture, the spot rate appears to be more volatile outside the same middle region we previously identified for the drift. In particular, the volatility of the spot rate appears to be increasing for higher levels of r, but smaller for levels of r below its long-term mean.

---

2 Note that the left-hand axis’ time-scale of the diffusion chart depends on the sampling interval for the data, while the proper non-parametric estimates resulting from Sahalia’s matching density method are unaffected by the frequency of the data.

3 In the Appendix, we attach a sample Limdep code for the FGLS estimation of equation [2].
3. The Stanton’s Estimation

Stanton’s strategy relies on the observation that, given an arbitrary function \( f(r) \) of the process \([1]\), it is possible to expand its conditional expectation:

\[
E_t[f(r_{t+D}, t + D)]
\]

as a stochastic Taylor expansion around \( f(r_t, t) \):

\[
E_t[f(r_{t+D}, t + D)] = f(r_t, t) + Lf(r_t, t)D + \frac{1}{2} L^2 f(r_t, t)D^2 + \ldots + \frac{1}{n!} L^n f(r_t, t)D^n + O(D^{n+1})
\]

\[
Lf(r_t, t) = \frac{\partial f}{\partial t} + \frac{\partial f}{\partial r} \mu(r) + \frac{1}{2} \sigma^2(r) \frac{\partial^2 f}{\partial r^2}
\]

From equation [10], a first-order proxy for the conditional expectation of \( f(.) \) is then:

\[
Lf(r_t, t) = \frac{1}{D} \{E_t[f(r_{t+D}, t + D) - f(r_t, t)]\} + O(D)
\]

Given a suitable choice for \( f(.) \), Stanton uses the result of equation [11] to express \( Lf(.) \) as a function of \( \mu(r) \) and \( \sigma^2(r) \). Finally, estimates for the moments of \( f(.) \) can be used to calculate point-wise estimates for \( \mu(r) \) and \( \sigma^2(r) \). Using a second-order approximation, as suggested in Stanton’s paper, the following expressions for \( \mu(r) \) and \( \sigma^2(r) \) are obtained:

\[
\mu(r_t) = \frac{1}{2D} \{4E_t[r_{t+D} - r_t] - E_t[r_{t+2D} - r_t]\} + O(D^2)
\]

\[
\sigma^2(r_t) = \frac{1}{2D} \{4E_t[(r_{t+D} - r_t)^2] - E_t[(r_{t+2D} - r_t)^2]\} + O(D^2)
\]

It is clear from equation [12] that the only remaining task left for the econometrician is to estimate the conditional moments in the square brackets. Stanton suggests to employ a non-parametric Gaussian-Kernel estimator for the marginal density of the underlying interest rate process. More specifically, we adopt the following estimate for the density of the interest rate process at \( r \):
\[
\hat{f}(r) = \frac{1}{\sigma^*} \prod_{t=1}^{T} K\left( \frac{r - r_{t}}{h(\cdot)} \right) = \frac{1}{\sigma^*} \prod_{t=1}^{T} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2} \left( \frac{r - r_{t}}{h(\cdot)} \right)^2}
\]

\[
h(\cdot) = \hat{\sigma} \left( T \frac{1}{h} \right)
\]

where \(\sigma^*\) is the estimate for the volatility of the underlying rate process over the sampling interval, and \(h(\cdot)\) is a bandwidth, based on the dispersion of the observations, that minimizes the asymptotic mean integrated squared error of the estimated density function.

Given the density resulting from equation [13], we can now estimate any moment we desire from the distribution. We will then use the following algorithm:

\[
E\left[ (r_{t+j} - r_i) / r_i \right] = \int_0^\infty (r_{t+j} - r_i) f\left( r_{t+j} / r_i \right) dr_{t+j} = \int_0^\infty \left( r_{t+j} - r_i \right) f\left( r_{t+j} / r_i \right) \frac{f\left( r_{t+j} / r_i \right)}{f_i(r)} dr_{t+j} = \int_{t=1}^{T-j} \left( r_{t+j} - r_i \right) K\left( \frac{r - r_{t+j}}{h(\cdot)} \right) \frac{r}{h(\cdot)}
\]

\[\forall i, j \]

The estimated marginal density over the interest-rate sample 1973-1995 is reported in Figure 2.

**Figure 2: Non-Parametric Estimated Density for \(r\) - 1973-1995**

---

4 Over our sample, we estimated this value to be around 3.59%.

As evident from the figure above, although most of the probability mass is allocated for rates between 4 and 13 %, levels in the 15 to 17 % range have some mass as well. This observation is going to be relevant soon for the analysis.

The next figure shows the point-wise estimates for the drift and the diffusion of the 7-days spot eurodollar interest rate resulting from equation [12] and the set of moments of equation [14].

The drift term $\mu(r)$ is substantially zero for most of the values of $r$ between 0 and 13 %. No strong reversion appears for small values of $r$, contrary to the findings of Sahalia in section 2. Nonetheless, there still appears to be a significant negative reversion push from $\mu(r)$ for values of $r$ above 15 %. It is also worth commenting on the short positive spike in the drift for $r$ between 17 and 19 %. This spike “happens” because of the fact that such values for $r$ have some “mass” in the marginal density of figure 1. In other terms, as in our sample there are more than few observations for $r$ at those levels, a consequence of the floating-interest-rate policy adopted by Paul Volcker at the Fed between 1979 and 1982, too a negative drift for that range would prevent the estimated process $r_t$ of equation [1] to fit the data. In practice, the spike is a consequence of the Stanton’ s technique over-fitting the available data.

The diffusion term appears to be convex, as suggested by Sahalia, but not decreasing at low levels of the interest rate. Again, a downward spike appears for $r_t$ between 17 and 19 %: as long as the drift term is pushing upward to reach the high levels of $r$ in the sample, the volatility drops, in order not to “obstacle” the fit.

Both Sahalia and Stanton’ s results seem to suggest that the spot interest rate behaves substantially as a random walk for most of the levels of the eurodollar rate observed in the past twelve years, that the instantaneous volatility is increasing with the level of the rate and that the drift and the diffusion in

---

6 In the Appendix, we report a sample code in C++ for the estimation of drift and diffusion through the Stanton’ s approach.
equation [1] are highly non-linear, although they disagree on the nature and entity of these non-linearities. How sensitive are these conclusions to the particular estimation strategy selected? In other terms, how statistically powerful are these conclusions? The next section tackles this compelling issue.

4. A CIR Simulation

The main difficulty we face in evaluating qualitatively the results of both Sahalia’s and Stanton’s empirical analysis lies in that we do not actually know how the true underlying interest rate process truly behaves. Hence, any of their conclusions are ex-ante acceptable, and difficult to disprove. Nonetheless, we would be able to say more about the power of their statistical techniques, i.e. more specifically about their capacity to identify the “true” drift and diffusion terms of equation [1], when we actually know ex-ante what these “true” elements really are. Fortunately, such an evaluation is possible: we can generate simulated time series of interest rates from a given interest rate process, where we specify exactly the form of $\mu(r)$ and $\sigma(r)$. Then, we simply apply the FGLS approach and the Taylor’s Expansion Method to these simulated series and observe whether any of those techniques is able to capture the main characteristics of the true drift and diffusion from which the data were actually generated.

We select a popular model for interest rate dynamics, the CIR process, of the form:

$$dr_t = k(\theta - r_t)dt + \sigma \sqrt{r_t} dw_t \quad [15]$$

We assume that at the beginning of the sample period, of the same length as the original sample, $r_{\text{CIR}}(0) = r(0)$, and then generate a sample path for $r_t$ from equation [15] with the following (asymptotically exact) daily Forward-Euler approximation:

$$r_{t+1} = r_t + k(\theta - r_t)dt + \sigma \sqrt{r_t} G_{t+1}, \quad [16]$$

where $G$ is a random variable distributed as a Gaussian Normal with mean zero and unit variance. We generate a time series of independently drawn values of $G$ with the Box-Muller algorithm. Values for $k$, $\phi$ and $\sigma$ come from Pritsker (1998)\footnote{It is worthwhile observing that Pritsker values are annualized, hence for the simulation of the CIR process we use $dt = 0.004 \equiv (1/252)$.}, and are respectively, 0.89218, 0.090495 and $(0.032742/252)^{0.5}$.

We finally repeat the two procedures described in sections 2 and 3 on this simulated interest rate series. Let’s start with the Feasible GLS of Sahalia. In figure 3 we report the estimated drift resulting from the
estimation of equation [2] on the simulated CIR data and what is supposed to be the “true” drift term (the dotted line).

**Figure 4: FGLS Estimates for $\mu(r)$ of the CIR Process**

As clearly pictured above, Sahalia’s representation appears to impose non-linearity over a truly linear drift, for very low and very high levels of the interest rate, while slightly overestimating the reversion component of the drift for values of $r$ between 4% and 15%. We argue that this happens because of the “polynomial” character of the functional form for $\mu(r)$ in equation [2]. It is in fact well known that polynomial estimates tend to “explode” at the tails of the data sample over which the estimation is run. And Sahalia identifies strong reversion exactly at the tails of “sampled” values for $r$.

**Figure 5: FGLS Estimates for $\sigma(r)$ and $\sigma^2(r)$ of the CIR Process**

The estimation for the diffusion term, reported in figure 5, appears to be more successful for most of the values of $r$ in the range 5 to 20%.

Still, Sahalia’s estimation overstates the instantaneous volatility of the CIR process for low values of $r$. Again this is a result of the over-shooting in the drift term and the need for the diffusion component to pull the process away from the unattainable level of zero.
Using the same simulated time series, we apply Stanton’s procedure and report the results in figure 6, where the “true” drift and diffusion are shown in dotted line.

Figure 6: Stanton’s Estimates for $\mu(r)$ and $\sigma(r)$ of the CIR Process

Stanton’s estimate for the drift seems to suggest that, although free of overshooting in the short end, his procedure is biased toward negative reversion for high levels of interest rates. Even in this case, the diffusion estimate appears to be quite accurate for most of the values of $r$ in the range 2 to 18%. Again, as in the case of FGLS, the accuracy drops for values at the extremities of the range.

The evidence presented so far leads us to conclude that the two methods appear to be biased by construction in the measurement of drift and diffusion for values of $r$ in the tails of the range of possible values observed in the given time series. Obviously, this conclusion is not by itself surprising, as it is exactly for very high and very low values of $r$ that we have fewer observations available, hence any inference becomes more problematic.

In order to properly evaluate the power of these strategies, we repeat the experiment described above $N$ times, then calculate the mean point-wise estimates for $\mu(r)$ and $\sigma(r)$ and 95% confidence intervals.

Table 2 reports Sahalia’s FGLS mean estimated parameters’ values, while table 3 compares the mean estimated values resulting from 30 simulations of the CIR process (and, correspondingly, 30 repetitions of the FGLS procedure) for selected values of $r$. Figure 3 reports the estimates for $\mu(r)$ and $\sigma(r)$ over the range 0 - 20% versus the “true” values for the drift and diffusion (dotted line). We also report 95% point-wise confidence intervals.

The accuracy of $\mu'(r)$ and $\sigma'(r)$ appears to be enormously increased, especially as far as the diffusion term is concerned. Nonetheless, the FGLS approach still fails to properly account for the behavior of the drift at the tails of the range of possible values for $r$. This means that even with 30 additional 22 years-long time series simulated from CIR, the general form of equation [2] keeps imposing non-linear reversion to the data.
Table 2: FGLS Estimation for N = 30 Simulated CIR Process

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Sahalia’s FGLS Estimation of the CIR Diffusion</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>FGLS Estimates</td>
</tr>
<tr>
<td>α₀</td>
<td>-0.00086</td>
</tr>
<tr>
<td>α₁</td>
<td>0.01239</td>
</tr>
<tr>
<td>α₂</td>
<td>-0.0676</td>
</tr>
<tr>
<td>α₃</td>
<td>0.000028</td>
</tr>
<tr>
<td>β₀</td>
<td>0.00000063</td>
</tr>
<tr>
<td>β₁</td>
<td>0.000114</td>
</tr>
<tr>
<td>β₂</td>
<td>0.000051</td>
</tr>
<tr>
<td>β₃</td>
<td>1.6974</td>
</tr>
</tbody>
</table>

Table 3: FGLS Estimation for N = 30 Simulated CIR Processes versus the “true” µ(r) and σ(r)

<table>
<thead>
<tr>
<th>Rates</th>
<th>FGLS Estimation of the CIR Diffusion: Analysis of Drift and Diffusion at selected values</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>µ(x)</td>
</tr>
<tr>
<td>0 %</td>
<td>n.a.</td>
</tr>
<tr>
<td>5 %</td>
<td>0.000084</td>
</tr>
<tr>
<td>10 %</td>
<td>0.000058</td>
</tr>
<tr>
<td>15 %</td>
<td>0.000283</td>
</tr>
<tr>
<td>20 %</td>
<td>0.000875</td>
</tr>
</tbody>
</table>

Figure 6: FGLS Estimates for µ(r) and σ(r) of the CIR Process: N = 30 Simulations

The same analysis is then repeated for the Stanton’ s procedure with N = 100 simulations. Table 4 reports the estimated values for µ*(r) and σ*(r) for selected r.
Table 4: Stanton’ s Estimation for $N = 100$ Simulated CIR Processes versus the “true” $\mu(r)$ and $\sigma(r)$

<table>
<thead>
<tr>
<th>Rates</th>
<th>Stanton Estimation of the CIR Diffusion: Analysis of Drift and Diffusion at selected values</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\mu(x)$</td>
</tr>
<tr>
<td>0 %</td>
<td>0.000752</td>
</tr>
<tr>
<td>5 %</td>
<td>0.000158</td>
</tr>
<tr>
<td>10 %</td>
<td>-0.000021</td>
</tr>
<tr>
<td>15 %</td>
<td>-0.000323</td>
</tr>
<tr>
<td>20 %</td>
<td>-0.001945</td>
</tr>
</tbody>
</table>

As in the FGLS case, again the technique fails to improve significantly its precision at the tails of the range. In figure 7 we report the estimates for $\mu(r)$ and $\sigma(r)$ over the range 0 - 20 % versus the “true” values for the drift and diffusion (dotted line). The confidence intervals widen significantly for small and high values of $r$, i.e. exactly where we are most interested in understanding the behavior of the underlying, and in this case known, interest rate process. Finally, figure 8 reports the estimated marginal density for $r$, and its confidence intervals. This figure, and in particular the width of the upper and lower bands, suggests the high degree of uncertainty surrounding the non-parametric estimation of the “true” density at the tails of the range.

Figure 7: Stanton’ s Estimates for $\mu(r)$ and $\sigma(r)$ of the CIR Process: $N = 100$ Simulations

![Figure 7](image)

Few points deserve additional clarification. The poor results for both methodologies even after more than $N = 30$ repetitions of the same simulation experiment confirm earlier findings by Pritsker that FGLS and Taylor’ s approach grossly overestimate the amount of information that is actually in the data. In fact, a

---

8 We remind the reader that the “true” conditional distribution of $r$, i.e. conditional on the initial value $r(0)$, is a Non-Central Chi-Squared.
failure for FGLS to recognize the linearity of the drift even after 30 simulations of newly drawn CIR time series over a span of 22 years means that the method would have failed to identify the linear reversion even with more than 660 (22 times 30) years of daily observations. The failure appears even more “embarrassing” for the Stanton’ s procedure, as more than 2200 years of daily interest rate data would have not helped the strategy in finally removing the negative drift from the fitted $\mu(r)$. The degree of persistence in interest rate data, whether real or simulated through a process (CIR) that embeds it, limits the accuracy of these approaches, hence their eventual efficacy in answering the questions for which they were originally designed.

Figure 8: Stanton’ s Estimates for the marginal density of the CIR Process: $N = 100$ Simulations

Slightly more comforting are the results for the estimation of the “true” diffusion: with just a single simulation, i.e. “just” 22 years of data, the point-wise estimates for $\sigma(r)$ are fairly accurate.

In the appendix, figure 9, we report the results of an out-of-sample analysis of the estimated process [1] over the sample period 1995-2000. Out-of-sample series for $r_t$ have been generated with a single time-series for the Gaussian Wiener Process $dw_t$ for the original Sahalia’ s Matching-Density RFS estimated process [2], the FGLS process [2], the Stanton process, and the CIR process (with Pritsker’ s coefficients given at the beginning of this section). We also report (in the darkest line) the observed process for 3-month constant-maturity interest rate, as calculated daily by the Federal Reserve Bank of St. Louis until March 8th 2000, to extend Sahalia’ s sample and compare the performance of each of the proposed models. Sahalia’ s RFS process shows the biggest fluctuations, but all of them predict a sudden increase in the interest rate between 1997 and 1998 that did not happen in the real data. The CIR process is the one that gets closer to the observed rate on March 8th 2000. All estimated processes fluctuate more than the “true” interest rate.

This analysis suggests that the estimated interest rate process [1], especially over a single time-series drawing of $dw_t$, should not be used to predict “the” future behavior of interest rates, as more economically-sound models are designed to, but simply to describe the potential future dynamics of $r$. 

13
These dynamics are then fundamental in pricing contingent claims written on assets whose price fluctuations may depend heavily on the potential fluctuations of the spot rate itself.

5. Conclusions
Estimation of the unknown drift and diffusion terms of a stochastic process representing the spot interest rate has received increased attention in the financial literature. In this brief paper, we analyze the performance of the two most popular non-parametric approaches to the problem. An outright estimation of each of the proposed models to a sample of 22 years of daily observations of the 7-days eurodollar spot interest rate seems to suggest the existence of non-linearity in both the reversion and the volatility of the rate process. However, the results of a series of simulations from a known process reveal that none of the methods here examined has enough power to recognize a “true” linear drift, i.e. that the linearity of the interest rate drift component is rejected “too often”. More accuracy is however found for the estimated diffusion. We are then inclined to confirm Sahalia’s claim that the volatility of the spot rate appears to be increasing with the level of the rate.
References
Ait-Sahalia, Yacine, 1996 a, Non-Parametric Pricing of Interest Rate Derivative Securities, Econometrica, Vol. 64, No. 3.
Appendix

Figure 9: Out-of Sample Estimation of the interest rate behavior - "true" interest rate in black