FINANCIAL ECONOMETRIC MODELS
Some Contributions to the Field

João Nicolau
Instituto Superior de Economia e Gestão, Universidade Técnica de Lisboa, Rua do Quelhas 6, 1200-781 Lisboa, Portugal, nicolau@iseg.utl.pt

Abstract: Four recent financial econometric models are discussed. The first aims to capture the volatility created by "chartists"; the second intends to model bounded random walks; the third involves a mechanism where the stationarity is volatility-induced, and the last one accommodates nonstationary diffusion integrated stochastic processes that can be made stationary by differencing.

Key words: ARCH models, diffusion processes, bounded random walk, volatility-induced stationarity, second order stochastic differential equations.

1. INTRODUCTION

1.1 The objective and scope of this work

This paper reflects some of our recent contributions to the state-of-the-art on our financial econometrics. We have selected four main contributions in the this field. Also, we briefly refer to some contributions to the estimation of stochastic differential equations, although the emphasis of this chapter is on specification of financial econometric models. We give the motivation behind the models, and the more technical details will be referred to the original papers. The structure of this chapter is as follows. In section 1.2 we refer some general properties of returns and prices. In section 2 we mention a model that aims to capture the volatility created by "chartists". This is done in a discrete-time setting in the context of ARCH models; also a continuous-time version is provided. In section 3 we present three diffusion processes, with different purposes. The first one intends to model bounded random walks; the idea is to model stationarity processes with random walk

M. Seabra Pereira (ed.), A Portrait of State-of-the-Art Research at the Technical University of Lisbon, 23–41.

behaviour. In the second one we discuss processes where the stationarity is volatility-induced. This is applicable to every time series where reversion effects occur mainly in periods of high volatility. In the last one, we focus on a second order stochastic differential equation. This process accommodates nonstationary integrated stochastic processes that can be made stationary by differencing. Also, the model suggests directly modelling the (instantaneous) returns, contrary to usual continuous-time models in finance, which model the prices directly.

1.2 Prices, returns and stylized facts

An important step in forming an econometric model consists in studying the main features of the data. In financial econometrics two of the most important variables are prices and returns (volatility is also fundamental and we shall go back to it later). Prices include, for example, stock prices, stock indices, exchange rates and interest rates. If we collect daily data, the price is usually some type of closing price. It may be a bid price, an ask price or an average. It may be either the final transaction price of the day or the final quotation. In discrete time analysis, researchers usually prefer working with returns, which can be defined by changes in the logarithms of prices (with appropriate adjustments for any dividend payments). Let P_t be a representative price for a stock (or stock indices, exchange rate, etc.). The return r_t at time t is defined as $r_t = log(P_t) - log(P_{t-1})$.

General properties (stylized facts) are well known for daily returns observed over a few years of prices. The most significant are:

- The (unconditional) distribution of r_t is leptokurtic and in some cases (for stock prices and indices) asymmetric;
- The correlation between returns is absent or very weak;
- The correlations between the magnitudes of returns on nearby days are positive and statistically significant.

These features can be explained by changes through time in volatility. Volatility clustering is a typical phenomenon in financial time series. As noted by Mandelbrot [**19**], “large changes tend to be followed by large changes, of either sign, and small changes tend to be followed by small changes.” A measurement of this fact is that, while returns themselves are uncorrelated, absolute returns $|r_t|$ or their squares display a positive, significant and slowly decaying autocorrelation function: $Corr(|r_t|, |r_{t+\tau}|) > 0$ for τ ranging from a few minutes to several weeks. Periods of high volatility lead to extreme values (and thus to a leptokurtic distribution). Figure 1 shows a typical time series of returns. Any econometric model for returns should capture these general features of financial time series data.

The statistical features of prices are not so obvious. In general, most of the series contain a clear trend (e.g. stock prices when observed over several years), others shows no particular tendency to increase or decrease (e.g. exchange rates). Shocks to a series tend to display a high degree of persistence. For example, the Federal Funds Rate experienced a strong upwards surge in 1973 and remained at the high level for nearly two years. Also, the volatility of interest rates seems to be persistent. We will resume some of these features in section 3.

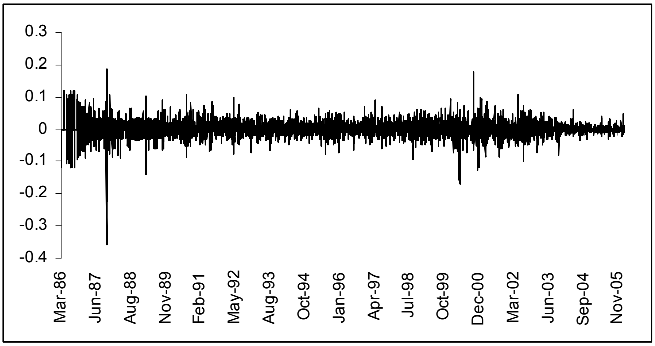

Figure 1. Microsoft daily returns from 1986 to 2006

2. DISCRETE-TIME MODELS

2.1 The ARCH family

In a seminal paper Engle [**13**] introduced the so called *autoregressive conditional heteroskedasticity* model. These models have proven to be extremely useful in modelling financial time series. Also, they have been used in several applications (forecasting volatility, CAPM, VaR, etc.). The ARCH(1) is the simplest example of an ARCH process. One assumes that the distribution of the return for period t, given past information, is

$$r_t | F_{t-1} \sim \mathrm{D}\left(\mu_t, \sigma_t^2\right) \tag{1}$$

where D is the conditional distribution, μ_t is the conditional mean and

$$\sigma_t^2 = \omega + \alpha\left(r_{t-1} - \mu_{t-1}\right)^2, \; (\omega > 0, \alpha \geq 0) \tag{2}$$

is the conditional variance. A large error in period $t-1$ (that is a high value for $(r_{t-1}-\mu_{t-1})^2$) implies a high value for the conditional variance in the next period. Generally, μ_{t-1} is a weak component of the model since it is difficult to predict the return r_{t-1} based on a F_{t-2}-mensurable stochastic process μ_{t-1}. In many cases it is a positive constant. Thus, either a large positive or a large negative return in period $t-1$ implies higher than average volatility in the next period; conversely, returns close to the mean imply lower than average volatility. The term *autoregressive* (from ARCH) comes from the fact that the squared errors follow an autoregressive process. In fact, from $\sigma_t^2=\omega+\alpha u_{t-1}^2$ where $u_{t-1}=r_{t-1}-\mu_{t-1}$ we have

$$\begin{aligned}&\sigma_t^2+u_t^2=\omega+\alpha u_{t-1}^2+u_t^2\\&u_t^2=\omega+\alpha u_{t-1}^2+v_t,\;\; v_t=u_t^2-\sigma_t^2\end{aligned} \tag{3}$$

and since v_t is a martingale difference (by construction, assuming $E[|v_t|]<+\infty$) one concludes that u_t^2 is an autoregressive process of order one. There are a great number of ARCH specifications and many of them have their own acronyms, such GARCH, EGARCH, MARCH, AARCH, etc.

2.2 One more ARCH model – the Trend-GARCH

2.2.1 Motivation

In recent literature a number of heterogeneous agent models have been developed based on the new paradigm of behavioural economics, behavioural finance and bounded rationality (see [**17**] for a survey on this subject). Basically, most models in finance distinguish between sophisticated traders and technical traders or chartists. Sophisticated traders, such as fundamentalists or rational arbitrageurs tend to push prices in the directions of the rational expectation fundamental value and thus act as a stabilising force. Chartists base their decisions mainly on statistics generated by market activity such as past prices and volume. Technical analysts do not attempt to measure the intrinsic value of a security; instead they look for patterns and indicators on stock charts that will determine a stock's future performance. Thus, there is the belief that securities move in very predictable trends and patterns.

As De Long et al. [**11**] recognise, this activity can limit the willingness of fundamentalists to take positions against *noise traders* (chartists). In fact, if noise traders today are pessimists and the price is low, a fundamentalist with a short time horizon buying this asset can suffer a loss if noise traders become even more pessimistic. Conversely, a fundamentalist selling an asset short when the price is high can lose money if noise traders become more bullish in the near future. "Noise traders thus create their own space. [...] Arbitrage does not eliminate the effect of noise because noise itself creates risk" (De Long et al., [**11**]). As a consequence, technical traders or chartists, such as feedback traders and trend extrapolators tend to push prices away from the fundamental and thus act as a destabilising force, creating volatility.

Based on these ideas, Nicolau [**26**] proposed an econometric model, in a discrete and continuous-time setting, based on a technical trading rule to measure and capture the increase of volatility created by chartists.

2.2.2 The Trend-GARCH

In order to derive the model we now focus more closely on a buy-sell rule used by chartists. One of the most widely used technical rules is based on the *moving average* rule. According to this rule, buy and sell signals are generated by two moving averages of the price level: a long-period average and a short-period average. A typical moving average trading rule prescribes a buy (sell) when the short-period moving average crosses the long-period moving average from below (above) (i.e. when the original time series is rising (falling) relatively fast). As can be seen, the moving average rule is essentially a trend following system because when prices are rising (falling), the short-period average tends to have larger (lower) values than the long-period average, signalling a long (short) position.

Hence, the higher the difference between these two moving averages, the stronger the signal to buy or sell would be and, at the same time, the more chartists detect the buy or sell signals. As a consequence, a movement in the price and in the volatility must, in principle, be expected, whenever a trend is supposed to be initiated. How to incorporate this information in the specification of the conditional variance is explained below. To simplify, we assume (as others) that the short-period moving average is just the current (or latest) market price and the long-period one is an exponentially weighted moving average (EWMA), which is also an adaptive expectation of the market price. In this formulation, the excess demand function of noise traders can be given as a function of $\log S_t - m_t$

$$q_t = f(\log S_t - m_t), \quad f'(x) > 0 \tag{4}$$

where S_t denotes the market price and m_t is the long-period moving average, represented here as an EWMA,

$$m_t = \lambda m_{t-1} + (1-\lambda) \log S_{t-1}, \quad 0 \le \lambda < 1. \tag{5}$$

The derivative of f (see equation (1)) is positive as, the higher the quantity $\log S_t - m_t > 0$, the stronger the signal to buy would be. Conversely, the lower the quantity $\log S_t - m_t < 0$, the stronger the signal to sell would be.

Based on these ideas and in Bauer [**7**], Nicolau [**26**] proposes the following model, inspired by the GARCH(1,1) specification:

$$\begin{aligned} r_t &= \mu_t + u_t, \\ u_t &= \sigma_t \varepsilon_t \\ \sigma_t^2 &= \omega + \alpha u_{t-1}^2 + \beta \sigma_{t-1}^2 + \gamma (\log S_{t-1} - m_{t-1})^2, \quad \alpha \ge 0, \beta \ge 0, \gamma \ge 0 \\ m_t &= \lambda m_{t-1} + (1-\lambda) \log S_{t-1}, \quad 0 \le \lambda < 1 \end{aligned} \tag{6}$$

where, r_t is the log return, μ_t is the conditional mean, $\{\varepsilon_t\}$ is assumed to be a sequence of i.i.d. random variables with $E[\varepsilon_t] = 0$ and $Var[\varepsilon_t] = 1$. The conditional variance σ_t^2 incorporates a measure of chartists trading activity, through the term $(\log S_{t-1} - m_{t-1})^2$.

We present some properties of this model. Suppose that $S_0 = 1$. Thus,

$$\log S_t = \log S_t - \log S_0 = \sum_{i=1}^{t} r_i \tag{7}$$

On the other hand, the EWMA process has the following solution

$$m_t = m_0 \lambda^t + (1-\lambda) \sum_{k=1}^{t} \log S_{k-1} \tag{8}$$

Combining equations (7) and (8), and assuming $m_0 = 0,$ we have, after some simplifications,

$$\log S_t - m_t = \sum_{i=1}^{t} r_i - (1-\lambda) \sum_{k=1}^{t} \lambda^{t-k} \log S_{k-1} = \sum_{i=1}^{t} \lambda^{t-i} r_i \tag{9}$$

If the sequence $\{r_i\}$ displays very weak dependence, one can assume $\mu_t = 0,$ that is $r_t = u_t$. In this case, we have

$$\sigma_t^2 = \omega + \alpha u_{t-1}^2 + \beta\sigma_{t-1}^2 + \gamma\left(\log S_{t-1} - m_{t-1}\right)^2$$
$$= \omega + \alpha u_{t-1}^2 + \beta\sigma_{t-1}^2 + \gamma\left(\sum_{i=1}^{t}\lambda^{t-i}u_i\right)^2. \qquad (10)$$

The model involves explicitly the idea of the moving average rule, which we incorporate using equation $m_t = \lambda m_{t-1} + (1-\lambda)\log S_{t-1}$. This moving average representation greatly facilitates the estimation of the model and the study of the stationary conditions. The expression $\sum_{i=1}^{t}\lambda^{t-i}r_i$ can be understood as a trend component, which approximately measures trend estimates in technical trading models. When the most recent returns have the same signal, that is, when $\left(\log S_{t-1} - m_{t-1}\right)^2 = \left(\sum_{i=1}^{t}\lambda^{t-i}r_i\right)^2$ is high, chartists see a general direction of the price (that is, a trend) which is generally classified as an *uptrend* or *downtrend*. In these cases, chartists increase their activity in the market, buying and selling and thus increasing volatility. On the other hand, when the trend is classified as *rangebound*, price swings back and forth for some periods, and as consequence, the quantity $\left(\sum_{i=1}^{t}\lambda^{t-i}r_i\right)^2$ is low (the positive returns tend to compensate the negative ones). In this case, there is much less trade activity by chartists, and the volatility associated with them is low.

It can be proved under the conditions, $\alpha \geq 0$, $\beta \geq 0$, $\gamma \geq 0$, $0 \leq \lambda < 1$ and $\{\varepsilon_t\}$ is a sequence of i.i.d. random variables with $E[\varepsilon_t] = 0$ and and $Var[\varepsilon_t] = 1$ that the process $\{u_t\}$ is covariance-stationary if and only if $\left(1-\lambda^2\right)(1-\alpha-\beta) > \gamma$. Conditions for the existence of a unique strict stationarity solution are also studied in Nicolau [**26**]. The stationarity makes sense because uptrends or downtrends cannot persist over time.

To assess the mean duration of a trend component, it could be interesting to calculate the speed of adjustment around zero. The higher the parameter $\lambda < 1$ the lower the speed of reversion. A useful indicator of the speed of adjustment is the so-called half-live indicator, which, in our case, is given by the expression $\log(1/2)/\log\lambda$.

Estimation of model (5) is straightforward. One can use the pseudo maximum likelihood based on the normal distribution (for example). A null hypothesis of interest is whether the term $\left(\log S_{t-1} - m_{t-1}\right)^2$ enters in the specification of the conditional variance, that is, H0: $\gamma = 0$. Under this hypothesis, λ is not identified, that is, the likelihood function does not depend on λ and the asymptotic information matrix is singular. One simple approach consists of considering Davies's bound when q parameters are identified only under alternative hypothesis (see Nicolau, [**26**]). An empirical illustration is provided in Nicolau [**26**]. Also, when the length of the discrete-time intervals between observations goes to zero, it is shown that, in some conditions, the discrete-time process converges in distribution to the solution of the diffusion process

$$\begin{aligned} dX_t &= \left(\bar{c} + \bar{\phi} X_t\right)dt + \sqrt{\bar{\omega} + \bar{\gamma}\left(X_t - \mu_t\right)^2}\, dW_{1,t}, \quad \bar{\omega} > 0, \bar{\gamma} > 0 \\ d\mu_t &= \bar{\theta}\left(X_t - \mu_t\right)dt + \bar{\sigma} dW_{2,t}, \quad \bar{\theta} \geq 0, \bar{\sigma} > 0. \end{aligned} \tag{11}$$

3. CONTINUOUS-TIME MODELS

3.1 A bounded random walk process

3.1.1 Motivation

Some economic and financial time series can behave just like a random walk (RW) (with some volatility patterns) but due to economic reasons they are bounded processes (in probability, for instance) and even stationary processes. As discussed in Nicolau [**21**] (and references therein) this can be the case, for example, of interest rates, real exchange rates, some nominal exchange rates and unemployment rates among others series. To build a model with such features it is necessary to allow RW behaviour during most of the time but force mean reversions whenever the processes try to escape from some interval. The aim is to design a model that can generate paths with the following features: as long as the process is in the interval of moderate values, the process basically looks like a RW but there are reversion effects towards the interval of moderate values whenever the process reaches some high or low values. As we will see, these processes can admit - relying on the parameters - stationary distributions, so we come to an interesting conclusion: processes that are almost indistinguishable from the RW process can be, in effect, stationary with stationary distributions.

3.1.2 The model

If a process is a random walk, the function $E\left[\Delta X_t \middle| X_{t-1} = x\right]$ (where $\Delta X_t = X_t - X_{t-1}$) must be zero (for all x). On the other hand, if a process is bounded (in probability) and mean-reverting to τ (say), the function $E\left[\Delta X_t \middle| X_{t-1} = x\right]$ must be positive if x is below τ and negative if x is above τ.

Now consider a process that is bounded but behaves like a RW. What kind of function should $E[\Delta X_t | X_{t-1} = x]$ be? As the process behaves like a RW, (i) it must be zero in some interval and, since the process is bounded, (ii) it must be positive (negative) when x is "low" ("high"). Moreover we expect that: (iii) $E[\Delta X_t | X_{t-1} = x]$ is a monotonic function which, associated with (ii), means that the reversion effect should be strong if x is far from the interval of reversion and should be weak in the opposite case; (iv) $E[\Delta X_t | X_{t-1} = x]$ is differentiable (on the state space of X) in order to assure a smooth effect of reversion. To satisfy (i)-(iv) we assume $E[\Delta X_t | X_{t-1} = x] = e^k \left(e^{-\alpha_1(x-\tau)} - e^{\alpha_2(x-\tau)}\right)$ with $\alpha_1 \geq 0$, $\alpha_2 \geq 0,$ $k < 0$. Let us fix $a(x) = e^k \left(e^{-\alpha_1(x-\tau)} - e^{\alpha_2(x-\tau)}\right)$. With our assumption about $E[\Delta X_t | X_{t-1} = x]$ we have the bounded random walk process (BRW) in discrete-time:

$$X_{t_i} = X_{t_{i-1}} + e^{k_\Delta}\left(e^{-\alpha_1\left(X_{t_{i-1}} - \tau\right)} - e^{\alpha_2\left(X_{t_{i-1}} - \tau\right)}\right)\Delta + \sigma_\Delta \varepsilon_{t_i}, \quad X_0 = c \qquad (12)$$

where t_i are the instances at which the process is observed, ($0 \leq t_0 \leq t_1 \leq ... \leq T$), Δ is the interval between observations, $\Delta = t_i - t_{i-1,}$ k_Δ and σ_Δ are parameters depending on Δ and $\{\varepsilon_{t_i}, i = 1,2,...\}$ is a sequence of i.i.d. random variables with $E[\varepsilon_{t_i}] = 0$ and $Var[\varepsilon_{t_i}] = 1$. It can be proved (see [**21**]) that the sequence $\{X_t^\Delta\}$ formed as a step function from X_{t_i}, that is $X_t^\Delta = X_{t_i}$ if $t_i \leq t < t_{i+1}$, converges weakly (i.e. in distribution) as $\Delta \downarrow 0$ to the solution to the stochastic differential equations (SDE) :

$$dX_t = e^k \left(e^{-\alpha_1(X_t - \tau)} - e^{\alpha_2(X_t - \tau)}\right)dt + \sigma dW_t, \quad X_{t_0} = c \qquad (13)$$

where c is a constant and W is a standard Wiener process ($t \geq t_0$). The case $a(x) = 0$ (for all x) leads to the Wiener process (which can be understood as the random walk process in continuous-time). It is still obvious that $a(\tau) = 0$, so X_t must behave just like a Wiener process when X_t crosses τ. However, it is possible, by selecting adequate values for k, α_1 and α_2 to have a Wiener process behaviour over a large interval centred on τ (that is, such that $a(x) \approx 0$ over a large interval centred on τ). Nevertheless, whenever X_t escapes from some levels there will always be reversion effects towards the τ. A possible drawback of model (12) is that the diffusion coefficient is constant. In the exchange rate framework and under a target zone regime, we should observe a volatility of shape "$\bigcap$" with respect to x (maximum volatility at the central rate) (see [**18**]**)**. On the other hand, under a free floating regime, it is common to observe a "smile" volatility (see [**18**]). For both possibilities, we allow the volatility to be of shape "$\bigcap$" or "$\bigcup$' " by assuming a specification like $exp\{\sigma + \beta(x - \mu)\}$.

Depending on the β we will have volatility of "$\cap$" or "$\cup$" form. Naturally, $\beta = 0$ leads to constant volatility. This specification, with $\beta > 0$, can also be appropriate for interest rates. We propose, therefore,

$$dX_t = e^k \left(e^{-\alpha_1(X_t-\tau)} - e^{\alpha_2(X_t-\tau)}\right)dt + e^{\sigma/2+\beta/2(X_t-\mu)^2} dW_t, \quad X_{t_0} = c \qquad (14)$$

Some properties are studied in [**21**]. Under some conditions both solutions are stationary (with known stationary densities). To appreciate the differences between the Wiener process (the unbounded RW) and the bounded RW, we simulate one trajectory for both processes in the period $t \in [0,20]$ with $X_0 = 100$. We considered $k = -2$, $\alpha_1 = \alpha_2 = 2$, $\tau = 100$ and $\sigma = 4$. The paths are presented in figure 2. In the neighbourhood of $\tau = 100$ the function $a(x)$ is (approximately) zero, so X behaves as a Wiener process (or a random walk in continuous-time). In effect, if $a(x) = 0$, we have $dX_t = \sigma dW_t$ (or $X_t = X_0 + \sigma W_t$). We draw two arbitrary lines to show that the bounded random walk after crossing these lines tends to move toward the interval of moderate values.

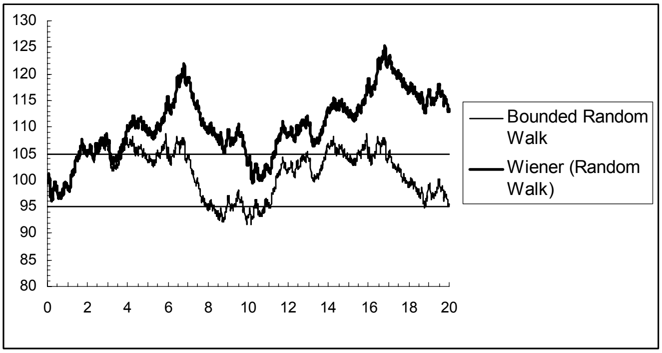

Figure 2. Bounded Randow Walk vs. Wiener Process

3.2 Processes with volatility-induced stationarity

3.2.1 Motivation

Short-term interest rate processes have shown at least two main facts. Firstly, the mean-reverting effect is very weak (see, for example, Chan et al. [**9**] or Bandi [**5**]). In fact, the stationarity of short-term interest rate processes is quite dubious. The usual unit root tests do not clearly either reject or accept the hypothesis of stationarity. Since interest rate processes are bounded by a lower (zero) and upper (finite) value a pure unit root hypothesis seems impossible since a unit root process goes to ∞ or $-\infty$ with probability one as time goes to ∞. Some authors have addressed this question. The issue is how to reconcile an apparent absence of mean-reverting effects with the fact that the interest rate is a bounded (and possibly stationary) process. While Aït-Sahalia [**1**] and Nicolau [**21**] suggests that stationarity can be drift-induced, Conley et al. [**10**] (CHLS, henceforth) suggest that stationarity is primarily volatility-induced. In fact, it has been observed that higher volatility periods are associated with mean reversion effects. Thus, the CHLS hypothesis is that higher volatility injects stationarity in the data.

The second (well known) fact is that the volatility of interest rates is mainly level dependent and highly persistent. The higher (lower) the interest rate is the higher (lower) the volatility. The volatility persistence can thus be partially attributed to the level of persistence of the interest rate.
The hypothesis of CHLS is interesting since volatility-induced stationarity can explain martingale behaviour (fact one), level volatility persistence (fact two), and mean-reversion. To illustrate these ideas and show how volatility can inject stationarity we present in figure 3 a simulated path from the SDE:

$$dX_t = \left(1 + X_t^2\right) dW_t \tag{15}$$

It is worth mentioning that the Euler scheme

$$Y_{t_i} = Y_{t_{i-1}} + \left(1 + Y_{t_{i-1}}^2\right)\sqrt{t_i - t_{i-1}}\,\varepsilon_{t_i}, \quad \varepsilon_{t_i} \sim \text{i.i.d.} N(0,1) \tag{16}$$

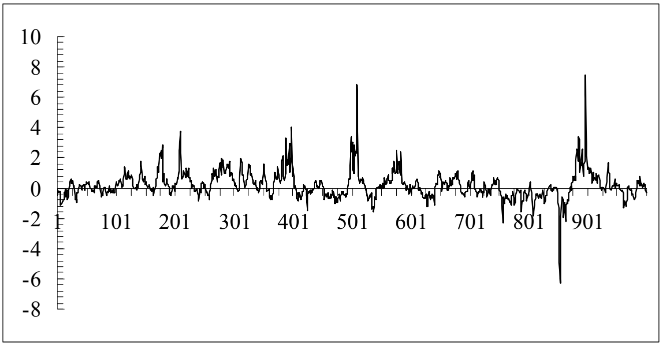

Figure 3. Simulated path from the SDE (14)

cannot be used since Y explodes as $t_i \to \infty$ (see [**24, 27**]). For a method to simulate X, see Nicolau [**24**]. Since the SDE (14) has zero drift, we could expect random walk behaviour. Nevertheless, figure 3 shows that the simulated trajectory of X exhibits reversion effects towards zero, which is assured solely by the structure of the diffusion coefficient. It is the volatility that induces stationarity. In the neighbourhood of zero the volatility is low so the process tends to spend more time in this interval. If there is a shock, the process moves away from zero and the volatility increases (since the diffusion coefficient is $1+x^2$) which, in turn, increases the probability that X crosses zero again. The process can reach extreme peaks in a very short time but quickly returns to the neighbourhood of zero. It can be proved, in fact, that X is a stationary process. Thus, X is a stationary local martingale but not a martingale since $E[X_t | X_0]$ converges to the stationary mean as $t \to \infty$ and is not equal to X_0 as would be required if X was a martingale.

3.2.2 A definition of volatility-induced stationarity

To our knowledge, CHLS were the first to discuss volatility-induced stationarity (VIS) ideas. Richter [**27**] generalizes the definition of CHLS. Basically, their definition states that the stationary process X (solution of the stochastic differential equation (SDE) $dX_t = a(X_t)dt + b(X_t)dW_t$) has VIS at boundaries $l = -\infty$ and $r = \infty$ if $\lim_{x \to l} s_X(x) < \infty$ and $\lim_{x \to r} s_X(x) < \infty$ where s_X is the scale density,

$$s_X(x) = \exp\left\{ -\int_{z_0}^{x} 2a(u)/b^2(u)du \right\} \quad (z_0 \text{ is an arbitrary value}) \tag{17}$$

There is one disadvantage in using this definition. As shown in [**25**], the VIS definition of CHLS and Richter does not clearly identify the source of stationarity. It can be proved that their definition does not exclude mean-reversion effects and thus stationarity can also be drift-induced.

A simple and a more precise definition is given in Nicolau [**25**]. Consider the following SDEs

$$\begin{aligned} dX_t &= a(X_t)dt + b(X_t)dW_t \\ dY_t &= a(Y_t)dt + \sigma dW_t. \end{aligned} \tag{18}$$

We say that a stationary process X has VIS if the associated process Y does not possess a stationary distribution (actually, this corresponds to what it defined in Nicolau [**25**] as VIS of type 2). The intuition is simple: although the process Y has the same drift as that of the process X, Y is nonstationary (by definition) whereas X is stationary. The substitution of σ for $b(x)$ transforms a nonstationary process Y into a stationary process. Thus, the stationarity of X can only be attributed to the role of the diffusion coefficient (volatility) and in this case we have in fact a pure VIS process.

The following is a simple criterion to identify VIS, in the case of state space $(-\infty,+\infty)$. We say that a stationary X process with boundaries $l=-\infty$ and $r=\infty$ has VIS if $\lim_{x\to\infty} xa(x)\geq 0$ or $\lim_{x\to-\infty} xa(x)\geq 0.$

3.2.2.1 An example: Modelling the Fed funds Rate with VIS

Processes with VIS are potentially applicable to interest rate time-series since, as has been acknowledged, reversion effects (towards a central measure of the distribution) occur mainly in periods of high volatility. To exemplify a VIS process monthly sampling of the Fed funds rate between January 1962 and December 2002 was considered. As discussed in Nicolau [**25**]**,** there is empirical evidence that supports the specification

$$dX_t = exp\{\alpha/2+\beta/2(X_t-\mu)^2\}dW_t \tag{19}$$

where $X_t = log\, r_t$ and r represents the Fed funds rate. The state space of r is $(0,\infty)$ and X is $(-\infty,\infty)$. That is, X can assume any value in R . This transformation preserves the state space of r, since $r_t = exp(X_t) > 0.$ By Itô's formula, equation (18) implies a VIS specification for interest rates

$$dr_t = r_t\frac{1}{2}e^{\alpha+\beta(log\, r_t-\mu)^2}dt + r_t e^{\alpha/2+\beta/2(log\, r_t-\mu)^2}dW_t \tag{20}$$

It can be proved that X is an ergodic process with stationary density

$$\overline{p}(x) = \frac{m_X(x)}{\int m_X(x)dx} = \frac{\sqrt{\beta}}{\sqrt{\pi}}e^{-\beta(x-\mu)^2} \tag{21}$$

i.e. $X = log\, r \sim N(\mu, 1/(2\beta))$. By the continuous mapping theorem, $r = exp(X)$ is an ergodic process. Furthermore, it has a log-normal stationary density. There is some empirical evidence that supports the above models. It is based on four facts:

1. The empirical marginal distribution of $X_t = log\, r_t$ matches the (marginal) distribution that is implicit in model (18).
2. The results of Dickey-Fuller tests are compatible with a zero drift function for $X,$ as specified in model (18).
3. Nonparametric estimates of $a(x)$ and $b^2(x)$ do not reject specification (18).

4. Parametric estimation of model (18) outperforms common one-factor models in terms of accuracy and parsimony.

The estimation of SDE (18) is difficult since the transition (or conditional) densities of X required to construct the exact likelihood function are unknown. Several estimation approaches have been proposed under these circumstances (see Nicolau [**20**] for a brief survey). To estimate the parameters of equation (18) we considered the simulated maximum likelihood estimator suggested in Nicolau [**20**] (with $N = 20$ and $S = 20$). The method proposed by Aït-Sahalia [**3**] with $J = 1$ (Aït-Sahalia's notation for the order of expansion of the density approximation) gives similar results. The approximation of the density based on $J = 2$ is too complicated to implement (it involves dozens of intricate expressions that are difficult to evaluate).

The proposed model compares extremely favourably with other proposed one-factor continuous-time models. In table 1 we compare the proposed model with other relevant models for interest rates. Only the proposed method was estimated by us. Remaining information was obtained from table VI of Aït-Sahalia [**2**]. For comparison purposes the proposed model was estimated using the same method applied to the other models (we considered the density approximation proposed by Aït-Sahalia [**3**] with $J = 1,$ in the period January-63 to December-98). Table 5 indicates that the proposed model outperforms the others in terms of accuracy and parsimony.

Table 1. Log-Likelihood of some Parametric Models, 1963-1998

Models	Log-likelihooh	Nº Parameters
$dr_t = \kappa(\tau - r_t)dt + \sigma dW_t$	1569.9	3
$dr_t = \kappa(\tau - r_t)dt + \sigma\sqrt{r_t}\,dW_t$	1692.6	3
$dr_t = r_t\left(\kappa - \left(\sigma^2 - \kappa\tau\right)r_t\right)dt + \sigma r_t^{3/2} dW_t$	1801.9	3
$dr_t = \kappa(\tau - r_t)dt + \sigma r_t^{\rho} dW_t$	1802.3	4
$dr_t = \left(\beta_1 r_t^{-1} + \beta_2 + \beta_3 r_t + \beta_4 r_t^2\right)dt + \sigma r_t^{3/2} dW_t$	1802.7	5
$dr_t = r_t \frac{1}{2} e^{\alpha+\beta(\log r_t - \mu)^2} dt + r_t e^{\alpha/2+\beta/2(\log r_t - \mu)^2} dW_t$	1805.1	3

3.3 A second order stochastic differential equation

In economics and finance many stochastic processes can be seen as integrated stochastic processes in the sense that the current observation behaves as the cumulation of all past perturbations. In a discrete-time framework the concept of integration and differentiation of a stochastic process plays an essential role in modern econometrics analysis. For instance, the stochastic process $\{y_t; t = 0,1,2,...\}$ where $y_t = \alpha + y_{t-1} + \varepsilon_t$ ($\varepsilon_t \sim i.i.d.N(0,1)$) is an example of an integrated process. Notice that y can be written as $y_t = y_0 + t\alpha + \sum_{k=1}^{t} \varepsilon_k$, or

$$y_t = y_0 + \sum_{k=1}^{t} x_k \tag{22}$$

where $x_t = \alpha + \varepsilon_t$. One way to deal with such processes is to use a differenced-data model (for example, $\Delta y_t = \alpha + \varepsilon_t$, in the previous example). Differencing has been used mostly to solve non-stationary problems viewed as unit roots although, historically, differenced-data models arose early in econometrics as a procedure to remove common trends between dependent and independent variables.

In empirical finance, most work on integrated diffusion processes is related to stochastic volatility models (see for example, Genon-Catalot and Laredo [**14**]) and realized volatility (see for example, Andersen et al. [**4**] and Barndorff-Nielsen and Sheppard [**6**]). However, integrated and differentiated diffusion processes in the same sense as integrated and differentiated discrete-time processes are almost absent in applied econometrics analysis. One of the reasons why continuous-time differentiated processes have not been considered in applied econometrics is, perhaps, related to the difficulties in interpreting the 'differentiated' process. In fact, if Z is a diffusion process driven by a Brownian motion, then all sample functions are of unbounded variation and nowhere differentiable, i.e. dZ_t / dt does not exist with probability one (unless some smoothing effect of the measurement instrument is introduced). One way to model integrated and differentiated diffusion processes and overcome the difficulties associated with the nondifferentiability of the Brownian motion is through the representation

$$\begin{cases} dY_t = X_t dt \\ dX_t = a(X_t)dt + b(X_t)dW_t \end{cases} \tag{23}$$

where a and b are the infinitesimal coefficients (respectively, the drift and the diffusion coefficient), W is a (standard) Wiener process (or Brownian motion) and X is (by hypothesis) a stationary process. In this model, Y is a differentiable process, by construction. It represents the integrated process,

$$Y_t = Y_0 + \int_0^t X_u du \tag{24}$$

(note the analogy with the corresponding expression in a discrete-time setting, $y_t = y_0 + \sum_{k=1}^{t} x_k$, equation (20)) and $X_t = dY_t / dt$ is the stationary differentiated process (which can be considered the equivalent concept to the first differences sequence in discrete-time analysis). If X represents the continuously compounded return or log return of an asset, the first equation in system (22) should be rewritten as $d \log Y_t = X_t dt.$

Nicolau [**23**] argues that (22) can be a useful model in empirical finance for at least two reasons. First, the model accommodates nonstationary integrated stochastic processes (Y) that can be made stationary by differencing. Such transformation cannot be done in common univariate diffusion processes used in finance (because all sample paths from univariate diffusion processes are nowhere differentiable with probability one). Yet, many processes in economics and finance (e.g. stock prices and nominal exchange rates) behave as the cumulation of all past perturbations (basically in the same sense as unit root processes in a discrete framework). Second, in the context of stock prices or exchange rates, the model suggests directly modelling the (instantaneous) returns, contrary to usual continuous-time models in finance, which directly model the prices. General properties for returns (stylized facts) are well known and documented (for example, returns are generally stationary in mean, the distribution is not normal, the autocorrelations are weak and the correlations between the magnitude of returns are positive and statistically significant, etc.). One advantage of directly modelling the returns (X) is that these general properties are easier to specify in a model like (22) than in a diffusion univariate process for the prices. In fact, several interesting models can be obtained by selecting $a(x)$ and $b^2(x)$ appropriately. For example, the choice $a(x) = \beta(\tau - x)$ and $b^2(x) = \sigma^2 + \lambda(X_t - \mu)^2$ leads to an integrated process Y whose returns, X, have an asymmetric leptokurtic stationary distribution (see the example below). This specification can be appropriated in financial time series data. Bibby and Sørensen [**8**] had already noticed that a similar process to (22) could be a good model for stock prices.

We observe that the model defined in equation (22) can be written as a second order SDE, $d(dY_t / dt) = a(X_t)dt + b(X_t)dW_t$. These kinds of equations are common in engineering. For instance, it is usual for engineers to model mechanical vibrations or charge on a capacitor or condenser submitted to white noise excitation through a second order stochastic differential equation. Integrated diffusions like Y in equation (23) arise naturally when only observations of a running integral of the process are available. For instance, this can occur when a realization of the process is observed after passage through an electronic filter. Another example is provided by ice-core data on oxygen isotopes used to investigate paleo-temperatures (see Ditlevsen and Sørensen [**12**]).

To illustrate continuous-time integrated processes we present in figure 4 two simulated independent paths of $Y_t = Y_0 + \int_0^t X_u du$ where X is governed by the stochastic differential equation

$$dX_t = 20(0.01 - X_t)dt + \sqrt{0.1 + 10(X_t - 0.05)^2}\, dW_t \quad (25)$$

(X is also represented in figure 4). All paths are composed of 1000 observations defined in the interval $t \in [0,10]$. It is interesting to observe that Y displays all the features of an integrated process (with a positive drift, since $E[X_t] = 0.01$): absence of mean reversion, shocks are persistent, mean and variance depend on time, etc. On the other hand, the unconditional distribution of X (return) is asymmetric and leptokurtic.

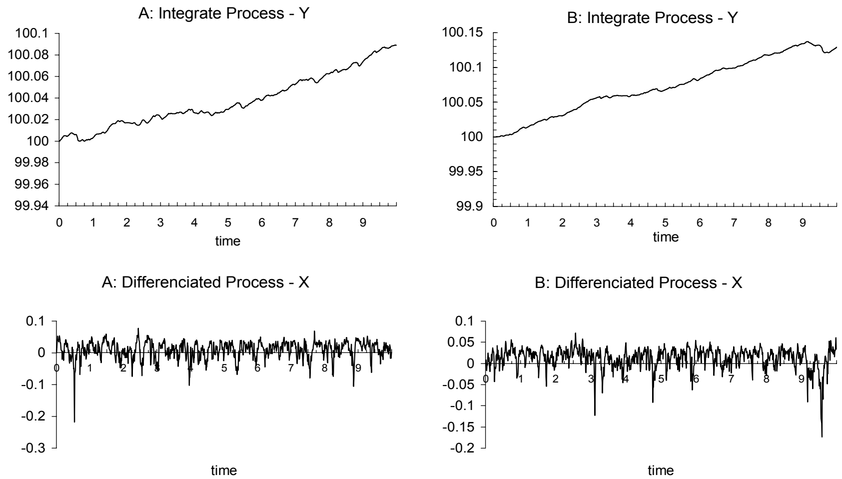

Figure 4 Simulation of two independent paths from a second order SDE

Estimation of second order stochastic differential equations raises new challenges for two main reasons. On the one hand, only the integrated process Y is observable at instants $\{t_i, i = 1,2,...\}$ and thus X in model (22) is a latent non-observable process. In fact, for a fixed sampling interval, it is impossible to obtain the value of X at time t_i from the observation Y_{t_i} which represents the integral $Y_0 + \int_0^{t_i} X_u du$. On the other hand, the estimation of model (22) cannot in principle be based on the observations $\{Y_{t_i}, i = 1,2,...\}$ since the conditional distribution of Y is generally unknown, even if that of X is known. An exception is the case where X follows an Orstein-Uhlenbeck process, which is analyzed in Gloter [**16**].

However, with discrete-time observations $\{Y_{i\Delta}, i = 1, 2, ...\}$ (to simplify we use the notation $t_i = i\Delta,$ where $\Delta = t_i - t_{i-1}$), and given that

$$Y_{i\Delta} - Y_{(i-1)\Delta} = \int_0^{i\Delta} X_u du - \int_0^{(i-1)\Delta} X_u du = \int_{(i-1)\Delta}^{i\Delta} X_u du, \quad (26)$$

we can obtain a measure of X at instant $t_i = i\Delta$ using the formula:

$$\widetilde{X}_{i\Delta} = \frac{Y_{i\Delta} - Y_{(i-1)\Delta}}{\Delta} \quad (27)$$

Naturally, the accuracy of (27) as a proxy for $X_{i\Delta}$ depends on the magnitude of Δ. Regardless of the magnitude of Δ we have in our sample, we should base our estimation procedures on the sample $\{\widetilde{X}_{i\Delta}, i = 1, 2, ...\}$ since X is not observable. Parametric and semi-parametric estimation of integrated diffusions is analyzed in Gloter [**15, 16**] and Ditlevsen and Sørensen [**12**]. In Nicolau [**23**] it is supposed that both infinitesimal coefficients a and b, are unknown. Non-parametric estimators for the infinitesimal coefficients a and b are proposed. The analysis reveals that the standard estimators based on the sample $\{\widetilde{X}_{i\Delta}, i = 1, 2, ...\}$ are inconsistent even if we allow the step of discretization Δ to go to zero asymptotically. Introducing slight modifications to these estimators we provide consistent estimators. See also [**22**].

ACKNOWLEDGEMENTS

I would like to thank Tom Kundert for helpful comments. This research was supported by the Fundação para a Ciência e a Tecnologia (FCT) and by POCTI.

REFERENCES

1. Aït-Sahalia, Y. (1996), Testing Continuous-Time Models of the Spot Interest Rate, The Review of Financial Studies 9, 385-426.
2. Aït-Sahalia, Y. (1999), Transition Densities for Interest Rate and Other Nonlinear Diffusions, The Journal of Finance LIV, 1361-1395.
3. Aït-Sahalia, Y. (2002), Maximum Likelihood Estimation of Discretely Sampled Diffusions: a Closed-Form Approximation Approach, Econometrica 70, 223-262.
4. Andersen T. & T. Bollerslev & F. Diebold & P. Labys (2001) The Distribution of Exchange Rate Volatility. Journal of the American Statistical Association 96, 42-55.

5. Bandi, F. (2002), Short-Term Interest Rate Dynamics: A Spatial Approach, Journal of Financial Economics 65, 73-110.
6. Barndorff-Nielsen, O. & N. Sheppard (2002) Econometric Analysis of Realized Volatility and its use in Estimating Stochastic Volatility Models. Journal of the Royal Statistical Society B 64, 253-280.
7. Bauer, C. "A Better Asymmetric Model of Changing Volatility in Stock Returns: Trend-GARCH", Working Paper 03-05, University Bayreuth, 2005.
8. Bibby, B. & M. Sørensen (1997) A Hyperbolic Diffusion Model for Stock Prices. Finance and Stochastics 1, 25-41.
9. Chan, K., G. Karolyi, F. Longstaff and A. Sanders (1992), An Empirical Comparison of Alternative Models of the Short-Term Interest Rate, The Journal of Finance XLVII, 1210-1227.
10. Conley, T., L. Hansen, E. Luttmer and J. Scheinkman (1997), Short-term interest rates as subordinated diffusions, The Review of Financial Studies 10, 525-577.
11. De Long, J.B., Shleifer, A., Summers,L.H. and Waldmann, R.J., (1990) Noise trader risk in financial markets, Journal of Political Economy 98, 703-738.
12. Ditlevsen S. & M. Sørensen (2004) Inference for Observations of Integrated Diffusion Processes. Scandinavian Journal of Statistics 31(3), 417-429.
13. Engle, R. "Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of the United Kingdom Inflation", *Econometrica*, **vol. 50-4**, pp. 987-1007, 1982.
14. Genon-Catalot, V. & J. Laredo (1998) Limit Theorems for Discretely Observed Stochastic Volatility Models. Bernoulli 4, 283-303.
15. Gloter A. (1999) Parameter Estimation for a Discretely Observed Integrated Diffusion Process. Preprint, Univ. of Marne-la-Vallée, 13/99.
16. Gloter A. (2001) Parameter Estimation for a Discrete Sampling of an Integrated Ornstein-Uhlenbeck Process. Statistics 35, 225-243.
17. Hommes, C. "Heterogeneous agent models in economics and finance", in Kenneth L. Judd, and Leigh Tesfatsion, ed.: Handbook of Computational Economics (North-Holland) Vol. 2: Agent-Based Computational Economics, 2005.
18. Krugman, P. & Miller, M. (1992) Exchange Rate Targets and Currency Bands. Centre for Economic Policy Research, Cambridge University Press.
19. Mandelbrot, B. "The variation of certain speculative prices", *Journal of Business* **vol. 36**, pp. 394-419, 1963.
20. Nicolau, J. "New Technique for Simulating the Likelihood of Stochastic Differential Equations" *The Econometrics Journal*, 5(1). pp. 91-103, 2002.
21. Nicolau, J. "Stationary Processes that Look Like Random Walks -- the Bounded Random Walk Process in Discrete and Continuous Time", *Econometric Theory*, **vol. 18 (1)**, pp. 99-118, 2002.
22. Nicolau, J. "Bias Reduction in Nonparametric Diffusion Coefficient Estimation", *Econometric Theory*, **vol. 19(5)**, pp. 754-777, 2003.
23. Nicolau, J. "Non-Parametric Estimation of Second Order Stochastic Differential Equations", Working Paper 3-04, CEMAPRE, 2004.
24. Nicolau, J. "A Method for Simulating Non-Linear Stochastic Differential Equations in R1", *Journal of Statistical Computation and Simulation*, **vol. 75(8)**, pp. 595-609, 2005.
25. Nicolau, J. "Processes with Volatility-Induced Stationarity. An Application for Interest Rates", *Statistica Neerlandica*, 59(4), pp. 376-396, 2005.
26. Nicolau, J. "Modelling the Volatility Created by Chartists: A Discrete and a Continuous-Time Approach", Working Paper 4-05, CEMAPRE, 2005.
27. Richter M. (2002), A study of stochastic differential equations with volatility induced stationarity (unpublished).