



HAL
open science

Modeling financial time series through second order stochastic differential equations

João Nicolau

► **To cite this version:**

João Nicolau. Modeling financial time series through second order stochastic differential equations. Statistics and Probability Letters, 2009, 78 (16), pp.2700. 10.1016/j.spl.2008.03.024 . hal-00595950

HAL Id: hal-00595950

<https://hal.science/hal-00595950>

Submitted on 26 May 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Accepted Manuscript

Modeling financial time series through second order stochastic differential equations

João Nicolau

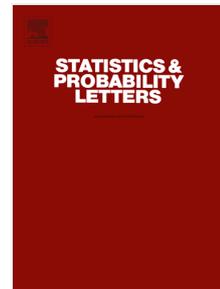
PII: S0167-7152(08)00187-9
DOI: [10.1016/j.spl.2008.03.024](https://doi.org/10.1016/j.spl.2008.03.024)
Reference: STAPRO 5026

To appear in: *Statistics and Probability Letters*

Received date: 9 October 2006
Revised date: 15 November 2007
Accepted date: 14 March 2008

Please cite this article as: Nicolau, J., Modeling financial time series through second order stochastic differential equations. *Statistics and Probability Letters* (2008), doi:10.1016/j.spl.2008.03.024

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.



Modeling Financial Time Series Through Second Order Stochastic Differential Equations

João Nicolau^{*†}

School of Economics and Management (ISEG) - Technical University of Lisbon (UTL)

November, 2007

Keywords: diffusion processes, second order stochastic differential equations.

Abstract

In this paper we motivate the use of second order stochastic differential equations in economics and finance. We provide an empirical illustration and discuss a parametric second order stochastic differential equation for stock prices and exchange rates.

^{*}Postal address: ISEG, Rua do Quelhas 6, 1200-781 Lisboa, Portugal. Telephone: +351-21 3925876. Fax: +351-21 3922782. e-mail: nicolau@iseg.utl.pt

[†]This research was supported by the Fundação para a Ciência e a Tecnologia (FCT) and by FEDER/POCI 2010.

1 Introduction

In this paper we motivate the use of second order stochastic differential equations in economics and finance. We show that second order stochastic differential equations are the right model in continuous-time to account for integrated processes that can be made stationary by differencing. We provide an empirical illustration based on non-parametric estimators discussed in Nicolau (2007). The estimation results suggest a parametric second order stochastic differential equation for stock prices and exchange rates, which we briefly discussed.

If y_t is an integrated discrete time process ($y_t \sim I(1)$), then Δy_t is a stationary process. But, what if Y is an integrated continuous-time diffusion process? Can we make Y stationary by differencing? Seemingly the answer is negative since if Y is a diffusion process driven by a Brownian motion then the differentiable process dY_t/dt does not exist with probability one, because all samples are of unbounded variation and nowhere differentiable. The difficulties in interpreting the ‘differentiated’ process is perhaps one reason why continuous-time differentiated processes are almost absent in applied econometrics. However, there is a simple way to explore integrated and differentiated diffusion processes. It implies the use of second order stochastic differential equations (SDEs) $d(dY_t/dt) = a(X_t)dt + b(X_t)dW_t$, or equivalently

$$\begin{cases} dY_t = X_t dt \\ dX_t = a(X_t)dt + b(X_t)dW_t \end{cases} \quad (1)$$

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,

$$Y_t = Y_0 + \int_0^t X_u du. \quad (2)$$

Note that if y is a discrete-time integrated process, for example, $y_t = \alpha + y_{t-1} + \varepsilon_t$ ($\varepsilon_t \sim$

i.i.d. $N(0, 1)$) then y_t can be written as $y_t = y_0 + t\alpha + \sum_{k=1}^t \varepsilon_k$, or

$$y_t = y_0 + \sum_{k=1}^t x_k, \quad (3)$$

where $x_t = \alpha + \varepsilon_t$. The analogy between the discrete and the continuous-time case is obvious if we compare equations (2) and (3). On the other hand, the process $X_t = dY_t/dt$ can be considered the equivalent concept to the first differences sequence in discrete-time analysis ($x_t = y_t - y_{t-1}$).

Model (1) can be useful in empirical finance for at least two reasons. First, the model accommodates nonstationary integrated stochastic processes (Y) that can be made stationary by differencing. As we mentioned previously, such transformation cannot be undertaken 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 modeling the (instantaneous) returns, in contrast to usual continuous-time models in finance, which directly model the prices (consider for example the elementary geometric Brownian motion and subsequent generalizations). By directly modeling the returns, the proposed model suggests following the same strategy as in a discrete time approach (first one obtains the returns from log prices and then a specific model is proposed, for example an AR(1) with GARCH innovations). 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 modeling the returns (X) is that these general properties are easier to specify in a model like (1) 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. We later suggest the choice $a(x) = \beta(\tau - x)$ and $b(x) = \sqrt{\alpha_0 + \alpha_1(X_t - \mu)^2}$ as this leads to an integrated process for the (log) prices Y whose

returns, X , have an asymmetric leptokurtic stationary distribution (note that if X represents the continuously compounded return or log return of an asset, the first equation in system (1) should be interpreted as $d \log Y_t = X_t dt$).

We observe that model (1), although not used in empirical finance and in applied economics in general, is 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 SDE. Integrated diffusions like Y in equation (2) 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, 2004).

Estimation of second order SDEs raises new challenges for two main reasons. On the one hand, only the integrated process Y is observable at instants $\{t_i, i = 1, 2, \dots\}$ and thus X in model (1) 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 (1) cannot in principle be based on the observations $\{Y_{t_i}, i = 1, 2, \dots\}$ 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 Ornstein-Uhlenbeck process, which is analyzed in Gloter (2001). Nevertheless, with discrete-time observations $\{Y_{i\Delta}, i = 1, 2, \dots\}$ (to simplify we use the notation $t_i = i\Delta$, where $\Delta = t_i - t_{i-1}$), and given that

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

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

$$\tilde{X}_{i\Delta} = \frac{Y_{i\Delta} - Y_{(i-1)\Delta}}{\Delta}. \quad (4)$$

Naturally, the accuracy of (4) as a proxy for $X_{i\Delta}$ depends on the magnitude of Δ .

Non-parametric estimation of integrated diffusions based on the observations $\{\tilde{X}_{i\Delta}\}$ is analyzed in Nicolau (2007). Gloter (1999, 2006) and Ditlevsen and Sørensen (2004) analyze parametric and semi-parametric estimation.

2 Examples

2.1 A Simulated Process

Figure 1 presents a numerical simulation of Y and X where $Y_t = 100 + \int_0^t X_u du$ and X is governed by the SDE

$$dX_t = 20(0.03 - X_t)dt + \sqrt{10 + 20(X_t - 0.07)^2}dW_t$$

defined in the interval $t \in [0, 50]$. The numerical simulations were based on the Euler-Maruyama approximation. It is interesting to observe that Y displays all the features of an integrated process: 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.

2.2 An Empirical Illustration

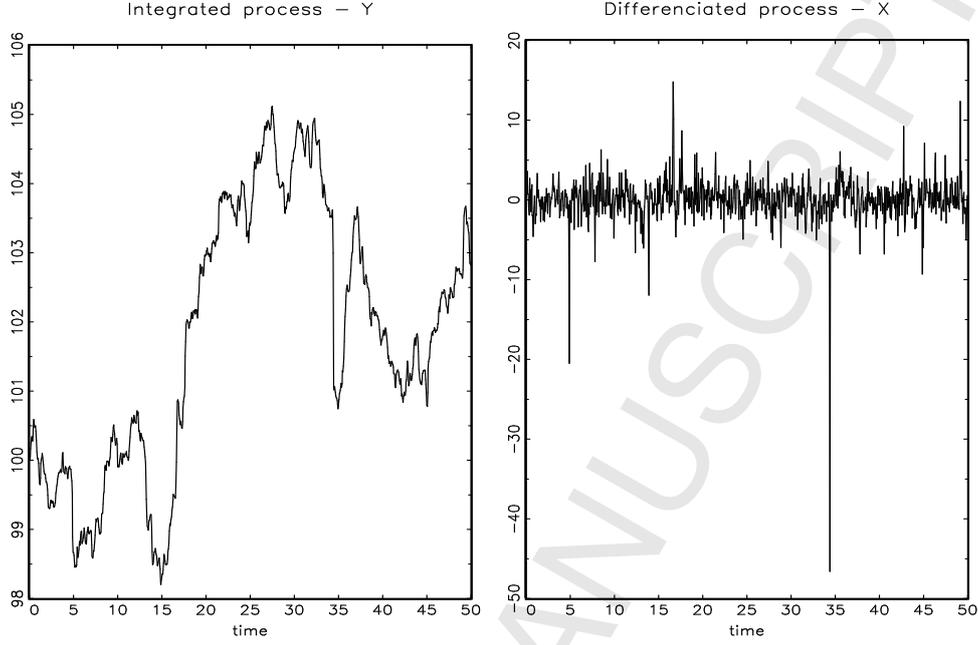
To illustrate second order SDE, we analyze daily closing stock prices of three major US companies (Applied Materials, starting from September 84, Oracle and Microsoft, both starting from March 88), three stock indices (DAX, FTSE and NASDAQ, respectively, starting from November 90, April 84 and February 71) and three US Dollar exchange rates (Euro, starting from January 99, Pound Sterling and Yen, both starting January 71). In all cases samples end in March 2006.

We assume

$$\text{(equation for log prices)} \quad d \log Y_t = X_t dt$$

$$\text{(equation for returns)} \quad dX_t = a(X_t)dt + b(X_t)dW_t.$$

In this application, we suppose that both infinitesimal coefficients, a and b , are unknown and our aim is their non-parametric functional estimation. Under some conditions (including stationarity

Figure 1: Numerical Simulation of Y and X .

of X , and that the step of discretization goes to zero at an appropriate rate) Nicolau (2007) has shown that

$$\hat{a}_n(x) = \frac{\frac{1}{nh_n} \sum_{i=1}^n K\left(\frac{x - \tilde{X}_{(i-1)\Delta_n}}{h_n}\right) (\tilde{X}_{i\Delta_n} - \tilde{X}_{(i-1)\Delta_n})}{\frac{1}{nh_n} \sum_{i=1}^n K\left(\frac{x - \tilde{X}_{(i-1)\Delta_n}}{h_n}\right)} \xrightarrow{p} a(x)$$

$$\sqrt{h_n n \Delta_n} (\hat{a}_n(x) - a(x)) \xrightarrow{d} N\left(0, K_2 \frac{b^2(x)}{\bar{p}(x)}\right)$$

where $K_2 = \int_{\mathbb{R}} K^2(u) du$ and K is the kernel function, and

$$\hat{b}_n^2(x) = \frac{\frac{1}{nh_n} \sum_{i=1}^n K\left(\frac{x - \tilde{X}_{(i-1)\Delta_n}}{h_n}\right) \frac{3}{2} \left(\frac{\tilde{X}_{i\Delta_n} - \tilde{X}_{(i-1)\Delta_n}}{\Delta_n}\right)^2}{\frac{1}{nh_n} \sum_{i=1}^n K\left(\frac{x - \tilde{X}_{(i-1)\Delta_n}}{h_n}\right)} \xrightarrow{p} b^2(x)$$

$$\sqrt{nh_n} (\hat{b}_n^2(x) - b^2(x)) \xrightarrow{d} N\left(0, \frac{4K_2 b^4(x)}{\bar{p}(x)}\right).$$

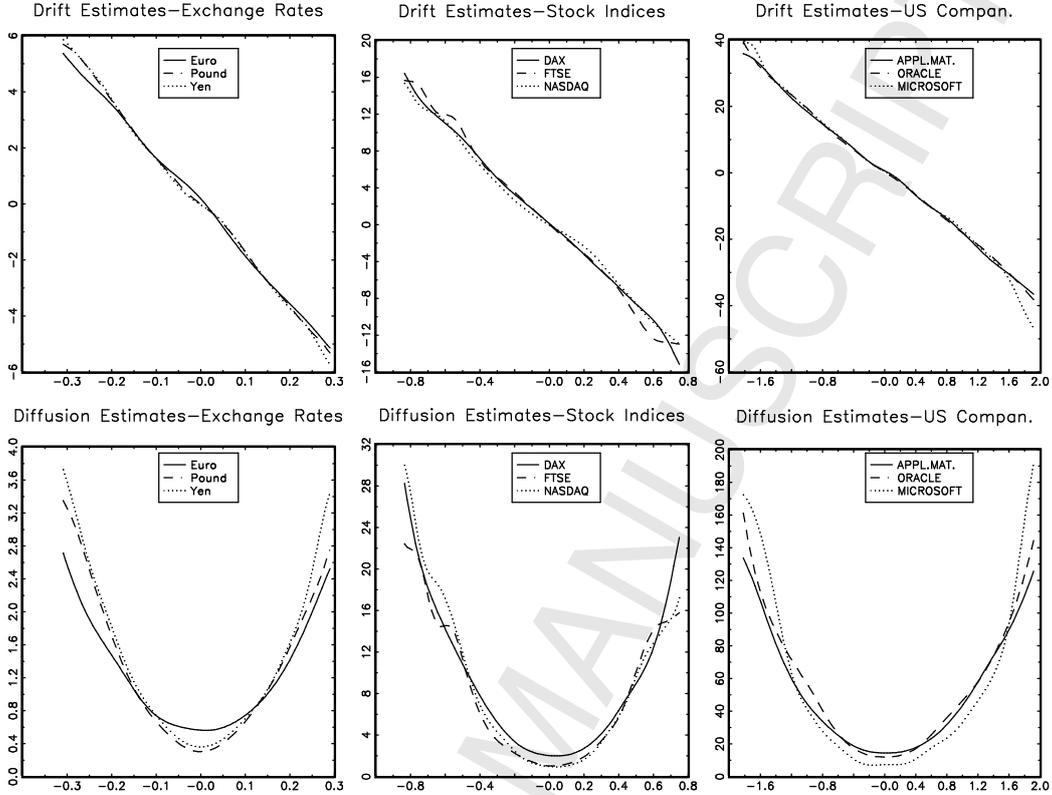


Figure 2: Non-parametric estimates of $a(x)$ and $b^2(x)$.

However, Monte Carlo simulation experiments suggest that

$$\check{a}_n(x) = \frac{\sum_{i=1}^n K\left(\frac{x - \tilde{X}_{(i-1)\Delta_n}}{h_n}\right) (\tilde{X}_{i\Delta_n} - \tilde{X}_{(i-1)\Delta_n})}{\sum_{i=1}^n K\left(\frac{x - \tilde{X}_{(i-1)\Delta_n}}{h_n}\right)}$$

performs better than $\hat{a}_n(x)$ for moderate/high values of Δ (see Nicolau, 2007). We use $\check{a}_n(x)$ and $\hat{b}_n^2(x)$ to estimate the infinitesimal coefficients $a(x)$ and $b^2(x)$ associated with these nine financial time series ($\tilde{X}_{i\Delta}$ is given by equation (4) and Y is the log of price). We fixed $\Delta = 1/20$ so the results have a monthly interpretation ($t = 1$ means one month). In figure 2 we present the results.

As we expected they show that returns from exchange rates have the least variability and individual stocks have the highest variability. It is interesting to observe a regularity pattern in all estimates. In all cases the drift is clearly linear and the diffusion is a convex function

with a minimum in the neighborhood of zero. Several other stock prices and exchange rate time series showed the same pattern. The specification $b^2(x) = \alpha_0 + \alpha_1(x - \mu)^2$ seems to fit the non-parametric estimates very well. The quadratic hypothesis for the diffusion is interesting as we show in next section.

3 A Second Order SDE

The previous section suggests that the second order SDE

$$\text{(equation for log prices)} \quad d \log Y_t = X_t dt$$

$$\text{(equation for returns)} \quad dX_t = \beta(\tau - X_t) dt + \sqrt{\alpha_0 + \alpha_1(X_t - \mu)^2} dW_t, \quad \beta \geq 0, \alpha_0 > 0, \alpha_1 \geq 0$$

is a plausible model for prices and returns of exchange rates, stocks and stocks indices. This model presents very interesting properties. It can be shown that X is ergodic and the invariant distribution P^0 has density $\bar{p}(x)$ with respect to the Lebesgue measure. The density \bar{p} is given by $\bar{p}(x) = m(x) / \int_{\mathbb{R}} m(u) du$ where $m(u) = (b^2(u) s(u))^{-1}$ and $s(z) = \exp \left\{ - \int_{z_0}^z 2a(u) / b^2(u) du \right\}$. Thus, (after some calculations) we have

$$\bar{p}(x) = \frac{1}{c} \exp \left\{ \frac{2\beta(\tau - \mu) \arctan \left(\sqrt{\frac{\alpha_1}{\alpha_0}} (x - \mu) \right)}{\sqrt{\alpha_0 \beta}} \right\} \left(\alpha_0 + \alpha_1 (x - \mu)^2 \right)^{-\frac{\beta}{\alpha_1} - 1}$$

where $c = \int_{\mathbb{R}} m(u) du$ is the normalizing constant,

$$c = \frac{4^{-\frac{\beta}{\alpha_1}} \pi \alpha_0^{-1/2 - \beta/\alpha_1} \sqrt{\alpha_1} \Gamma \left(1 + \frac{2\beta}{\alpha_1} \right)}{\Gamma \left(1 + \frac{\beta}{\alpha_1} + \frac{i\beta(\mu - \tau)}{\sqrt{\alpha_0 \beta}} \right) \Gamma \left(1 + \frac{\beta}{\alpha_1} + \frac{i\beta(\tau - \mu)}{\sqrt{\alpha_0 \beta}} \right)} \in \mathbb{R}, \quad (i = \sqrt{-1}).$$

To the best of our knowledge, \bar{p} is a new probability density function (pdf). Since the function $\exp \{ \arctan(x) \} > 0$ is bounded away from zero, the tails of \bar{p} are dominated by the expression $\left(\alpha_0 + \alpha_1 (x - \mu)^2 \right)^{-\frac{\beta}{\alpha_1} - 1}$. This means that \bar{p} has polynomial tails and thus heavy tails. Moreover, \bar{p} generalizes the t-Student distribution since if $\tau = \mu$, \bar{p} is the t-Student distribution (in this case centered in μ). The component $\exp \{ \arctan(\cdot) \}$ can be seen as a weight function that gives more weight to the right or left tails of \bar{p} according to the sign of $\tau - \mu$. If, for example, $\mu > \tau$ the

distribution is left-skewed. This tends to occur when $\mu > 0$ in the case of returns of stock prices because positive returns normally present lower variability than negative returns (we notice that the volatility is minimum at μ). It can be shown that

$$\begin{aligned}
 E[X] &= \tau, \\
 Var[X] &= \frac{\alpha_1(\mu - \tau)^2 + \alpha_0}{2\beta - \alpha_1}, \\
 E[(X - \tau)^3] &= \frac{2\alpha_1(\tau - \mu)(\alpha_0 + \alpha_1(\mu - \tau)^2)}{(\beta - \alpha_1)(2\beta - \alpha_1)}, \\
 E[(X - \tau)^4] &= \frac{3(\alpha_1(\tau - \mu)^2 + \alpha_0)(\alpha_1(3\tau^2\alpha_1 - 6\alpha_1\mu\tau - 2\mu\beta\tau + 3\alpha_1\mu^2 + \mu^2\beta - \alpha_0 + \beta\tau^2) + \alpha_0\beta)}{(2\beta - \alpha_1)(\beta - \alpha_1)(2\beta - 3\alpha_1)}.
 \end{aligned}$$

(the proofs are available from the author upon request). The pdf \bar{p} can even be explored in discrete-time modeling. One of the difficulties in modeling autoregressive conditional skewness and kurtosis is related with the choice of the conditional density. This pdf \bar{p} , after being properly normalized, can permit that the first four conditional moments evolve separately.

References

- Ditlevsen, S. and M. Sørensen, (2004), Inference for Observations of Integrated Diffusion Processes, *Scandinavian Journal of Statistics* **31**, 417-429.
- Gloter, A., (2001), Parameter Estimation for a Discrete Sampling of an Integrated Ornstein-Uhlenbeck Process, *Statistics* **35**, 225-243.
- Gloter, A., (2006), Parameter Estimation for a Discretely Observed Integrated Diffusion Process, *Scandinavian Journal of Statistics* **33**, 83-104.
- Nicolau, J., (2007), Non-Parametric Estimation of Second Order Stochastic Differential Equations, *Econometric Theory*, **23**, 880-898.

Figure Legends

Figure 1: Numerical Simulation of Y and X .

Figure 2: Non-parametric estimates of $a(x)$ and $b^2(x)$.

ACCEPTED MANUSCRIPT