A stochastic processes toolkit for risk management: Mean reverting processes and jumps

Received (in revised form): 20th August, 2009

Damiano Brigo*

is Managing Director of Fitch Solutions' Quantitative Innovation team and Visiting Professor at Imperial College London's Department of Mathematics. His work has been published in top journals for mathematical finance, systems theory, probability and statistics, and he has also written a field reference book on stochastic interest rate modelling. Damiano is Managing Editor of the *International Journal of Theoretical and Applied Finance*, a member of the Fitch Academic Advisory Board, and has been a charter member of Risk Who's Who since 2007. Damiano obtained a PhD in stochastic filtering with differential geometry in 1996 from the Free University of Amsterdam, following a BSc in mathematics with honours from the University of Padua.

Antonio Dalessandro

holds a PhD in mathematics from Imperial College London and is research associate at the University of Geneva's Department of Econometrics. From 2001 to 2004 he worked as a researcher at CERN and INFN in Geneva. He has a first-class honours MSc in electrical engineering from the Politecnico di Bari, and his research interests include probability, functional analysis, optimisation and high frequencies trading.

Matthias Neugebauer

is Senior Director of Fitch Ratings' structured credit group. His responsibilities include criteria development and credit market research, including design and development of the portfolio credit model. Prior to this, he worked with the company's European quantitative research team. He holds a master's degree in finance and economics from Durham University and an engineering degree from Technical University Cottbus. He also holds a master's degree in mathematical finance from Oxford University.

Fares Triki

obtained an engineering diploma from a French grande école, an MSc in finance from CNAM Paris, and a DEA diploma in economics from Paris Sorbonne University with summa cum laude distinction. After working for multinational consulting firms, he joined the Fitch Ratings Quantitative Financial Research team with the role of risk assessment and valuation of various financial products. Fares is now a macro finance expert at the financial stability directorate of Banque de France and is also a PhD student at the Paris School of Economics. His research areas include risk modelling and international economics and finance.

*Fitch Solutions, 101 Finsbury Pavement, London EC2A 1RS, UK
E-mail: damiano.brigo@fitchsolutions.com

Abstract In risk management it is desirable to grasp the essential statistical features of a time series representing a risk factor. This paper aims to introduce a number of different stochastic processes that can help in grasping the essential features of risk factors, describing different asset classes or behaviours. The paper does not aim to be exhaustive, but gives examples and a feeling for practically implementable models, allowing for stylised features in the data. These models can also be used as building blocks to build more complex models, although, for a number of risk management applications, the models developed here suffice for the first step in the quantitative analysis. The broad qualitative features addressed here are fat tails and mean reversion.

The introduction, the general framework and fat tails have been addressed in the authors' paper published in Vol. 2, No. 4 of the *Journal*. This paper deals with mean reversion both on its own and jointly with fat tails.

Keywords: *risk management, stochastic processes, risk factors, fat tails, mean reversion, jump diffusion processes, variance gamma processes*

INTRODUCTION

This paper analyses the following:

- *Mean reverting processes:* Vasicek, CIR (levels of interest rates or spreads, or returns in general), exponential Vasicek (levels).
- *Mean reverting processes with fat tails:* Vasicek with jumps (levels of interest rates or spreads, or returns in general), exponential Vasicek with jumps (levels).

Different financial time series are better described by different processes. In general, when first presented with a historical time series of data for a risk factor, one should decide what kind of general properties the series features. The financial nature of the time series can give some hints at whether or not mean reversion and fat tails are expected to be present. Interest rate processes, for example, are often taken to be mean-reverting, whereas foreign exchange series are often supposed to be fat-tailed and credit spreads feature both characteristics. In general one should check for mean reversion or stationarity as well as for fat tails.

THE MEAN-REVERTING BEHAVIOUR

One can define mean reversion as the property to always revert to a certain constant or time-varying level with limited variance around it. This property is true for an AR(1) process (Figure 1) if the absolute value of the autoregression coefficient is less than one ($|\alpha| < 1$). To be precise, the formulation of the first order autoregressive process AR(1) is:

$$x_{t+1} = \mu + \alpha x_t + \sigma \varepsilon_{t+1} \Rightarrow \Delta x_{t+1} = (1-\alpha)\left(\frac{\mu}{1-\alpha} - x_t\right) + \sigma \varepsilon_{t+1} \quad (1)$$

All the mean-reverting behaviour in the processes that are introduced in this section are due to an AR(1) feature in the discretised version of the relevant SDE.

In general, before thinking about fitting a specific mean-reverting process to available data, it is important to check the validity of the mean reversion assumption. A simple way to do this is to test for stationarity. As for the AR(1) process, $|\alpha| < 1$ is also a necessary and sufficient condition for stationarity, testing for mean reversion is equivalent to testing for stationarity. Note that, in the special case where $\alpha = 1$, the process behaves like a pure random walk with constant drift, $\Delta x_{t+1} = \mu + \sigma \epsilon_{t+1}$.

A growing number of stationarity statistical tests are available and ready for use in many econometric packages (eg Stata, SAS, R, E-Views, etc). The most popular are:

- the Dickey and Fuller (DF) test;[1]
- Said and Dickey's augmented DF (ADF) test;[2]

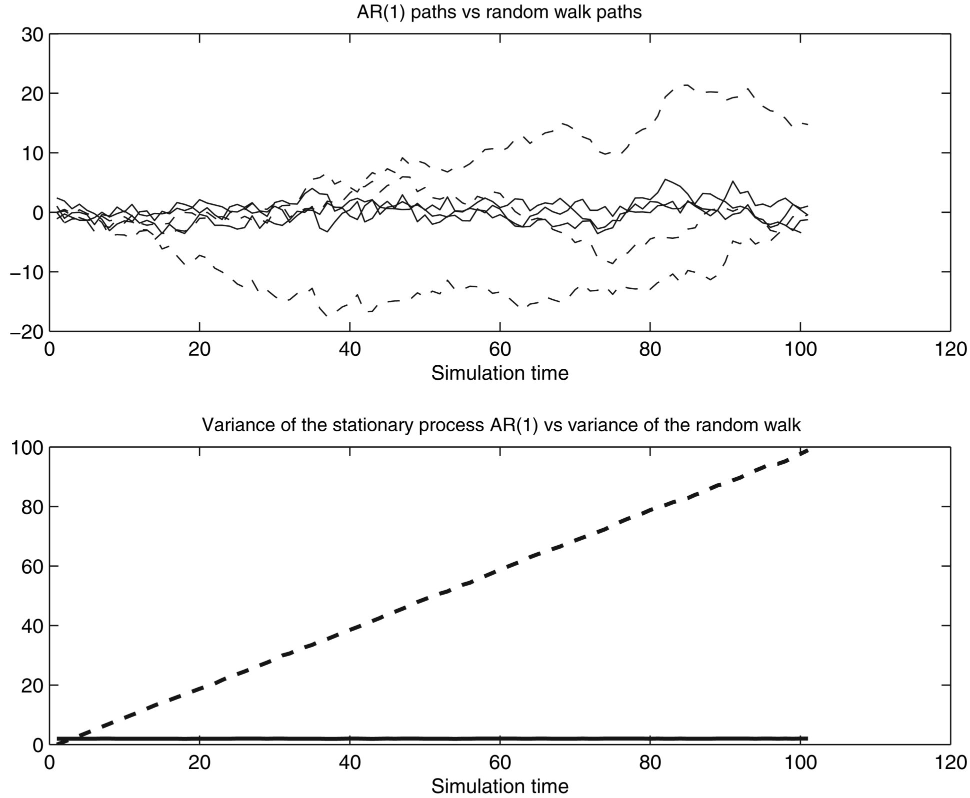

Figure 1: The random walk vs the AR(1) stationary process AR(1): $\mu = 0$, $\alpha = 0.7$, $\sigma = 1$ and random walk: $\mu = 0$, $\sigma = 1$

- the Phillips and Perron (PP) test;[3]
- the variance ratio (VR) test of Poterba and Summers[4] and Cochrane;[5] and
- the Kwiatkowski, Phillips, Schmidt and Shin (KPSS) test.[6]

All tests give results that are asymptotically valid, and are constructed under specific assumptions.

Once the stationarity has been established for the time series data, the only remaining test is for the statistical significance of the autoregressive lag coefficient. A simple OLS of the time series on its one-time lagged version (Equation (1)) allows one to assess whether α is effectively significant.

As an example, suppose that one is interested in analysing the GLOBOXX index history. Due to the lack of credit default swap (CDS) data for the time prior to 2004, the underlying GLOBOXX spread history has to be proxied by an index based on the Merrill Lynch corporate indices (asset swap spreads data of a portfolio of equally weighted Merrill Lynch indices that match both the average credit quality of GLOBOXX, its duration, and its geographical composition). This proxy index has appeared to be mean-reverting. Then the risk of having the CDS spreads rising in the mean term to revert to the index historical 'long-term mean' may impact the risk analysis of a product linked to the GLOBOXX index.

The mean reversion of the spread indices is supported by the intuition that the credit spread is linked to the economic cycle. This assumption also has some support from the econometric analysis. To illustrate this, the mean reversion of a Merrill Lynch index, namely the EMU Corporate A-rated 5-7Y index (Bloomberg Tickers ER33 — Asset Swap Spreads data from January 1998 to August 2007), is tested using the ADF test.

The ADF test for mean reversion is sensitive to outliers (for an outlier due to a bad quote, the up and down movement can be interpreted as a strong mean reversion in that particular data point and is able to influence the final result). Before testing for mean reversion, the index data have been cleaned by removing the innovations that exceed three times the volatility, as they are considered outliers. Figure 2 shows the EMU Corporate A-rated 5-7Y index cleaned from outliers and shows also the autocorrelations and partial autocorrelation plots of its log series. From this chart, one can deduce that the weekly log spread of the EMU Corporate A-rated 5-7Y index shows a clear AR(1) pattern. The AR(1) process is characterised by exponentially decreasing autocorrelations and by a partial autocorrelation that reaches zero after the first lag:

$$ACF(k) = \alpha^k, k = 0, 1, 2, \ldots;$$
$$PACF(k) = \alpha,\ k = 1;$$
$$PACF(k) = 0, k = 2, 3, 4, \ldots \quad (2)$$

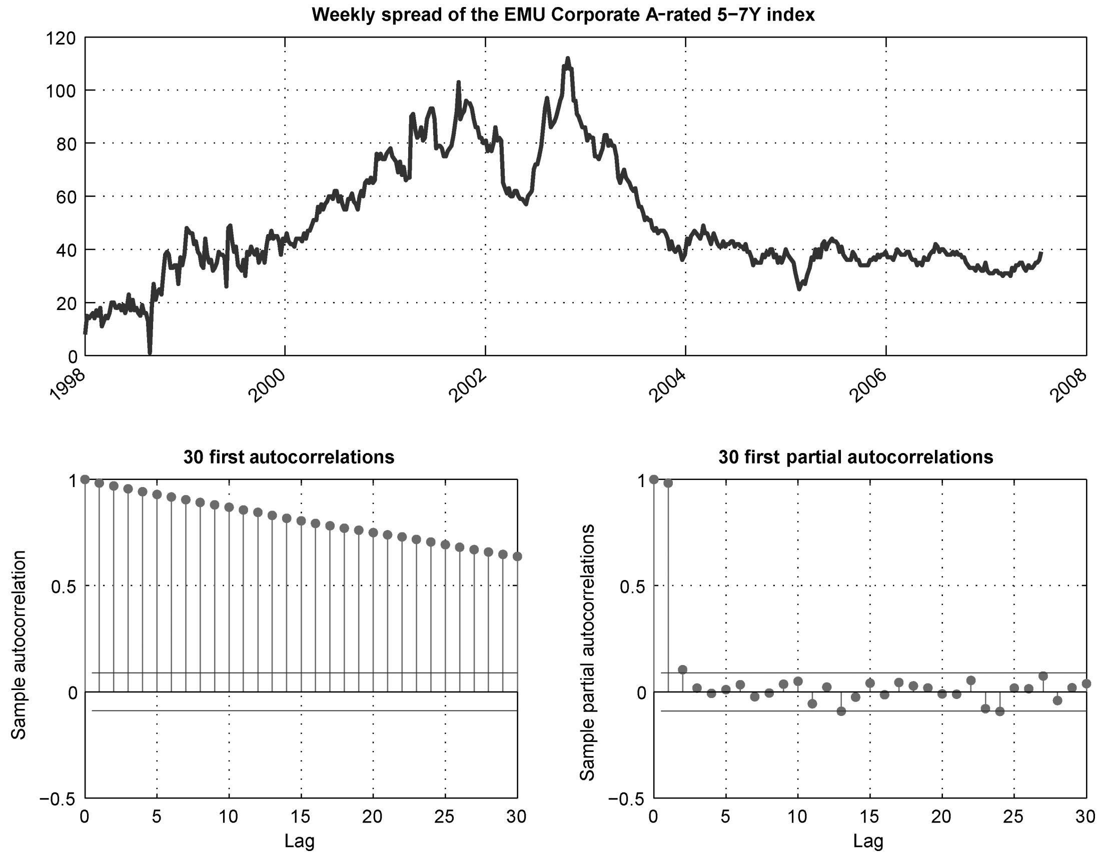

Figure 2: On the top, the EMU Corporate A-rated 5–7Y weekly index from 01/98 to 08/07; on the bottom, the ACF and PACF plot of this time series

The mean reversion of the time series has been tested using the ADF test. The figure obtained by this time series is -3.7373. The more negative the ADF statistics, the stronger the rejection of the unit root hypothesis (before cleaning the data from the outliers, the ADF statistic was -5.0226; this confirms the outliers' effect on amplifying the mean reversion detected level). The first and fifth percentiles are -3.44 and -2.87. This means that the test rejects the null hypothesis of unit root at the 1 per cent significance level.

Having discussed AR(1) stylised features in general, this paper now embarks on describing specific models featuring this behaviour in continuous time.

MEAN REVERSION: THE VASICEK MODEL

The Vasicek model,[7] is one of the earliest stochastic models of the short-term interest rate. It assumes that the instantaneous spot rate (or 'short rate') follows an Ornstein-Uhlenbeck process with constant coefficients under the statistical or objective measure used for historical estimation:

$$dx_t = \alpha(\theta - x_t)dt + \sigma dW_t \tag{3}$$

with α, θ and σ positive and dW_t a standard Wiener process. Under the risk-neutral measure used for valuation and pricing, one may assume a similar parametrisation for the same process but with one more parameter in the drift modelling the market price of risk; see for example Brigo and Mercurio.[8]

This model assumes a mean-reverting stochastic behaviour of interest rates. The Vasicek model has a major shortcoming that is the non null probability of negative rates. This is an unrealistic property for the modelling of positive entities like interest rates or credit spreads. The explicit solution to the SDE (3) between any two instants s and t, with $0 \leq s < t$, can be easily obtained from the solution to the Ornstein-Uhlenbeck SDE, namely:

$$x_t = \theta(1 - e^{-\alpha(t-s)}) + x_s e^{-\alpha(t-s)} + \sigma e^{-\alpha t}\int_s^t e^{\alpha u} dW_u \tag{4}$$

The discrete time version of this equation, on a time grid $0 = t_0, t_1, t_2, \ldots$ with (assume constant for simplicity) time step $\Delta t = t_i - t_{i-1}$ is:

$$x(t_i) = c + bx(t_{i-1}) + \delta\epsilon(t_i) \tag{5}$$

where the coefficients are:

$$c = \theta(1 - e^{-\alpha\Delta t}) \tag{6}$$

$$b = e^{-\alpha\Delta t} \tag{7}$$

and $\epsilon(t)$ is a Gaussian white noise ($\epsilon \sim N(0, 1)$). The volatility of the innovations can be deduced by the Itō isometry:

$$\delta = \sigma\sqrt{(1 - e^{-2\alpha\Delta t})/2\alpha} \tag{8}$$

whereas if one used the Euler scheme to discretise the equation, this would simply be $\sigma\sqrt{\Delta t}$, which is the same (first order in Δt) as the exact one above.

Equation (5) is exactly the formulation of an AR(1) process (see previous section). Having $0 < b < 1$ when $0 < \alpha$ implies that this AR(1) process is stationary and mean-reverting to a long-term mean given by θ. One can check this also directly by computing mean and variance of the process. As the distribution of $x(t)$ is Gaussian, it is characterised by its first two moments.

The conditional mean and the conditional variance of $x(t)$ given $x(s)$ can be derived from Equation (4):

$$\mathbb{E}_s[x_t] = \theta + (x_s - \theta)e^{-\alpha(t-s)} \quad (9)$$

$$\mathrm{Var}_s[x_t] = \frac{\sigma^2}{2\alpha}(1 - e^{-2\alpha(t-s)}) \quad (10)$$

One can easily check that as time increases, the mean tends to the long-term value θ and the variance remains bounded, implying mean reversion. More precisely, the long-term distribution of the Ornstein-Uhlenbeck process is stationary and is Gaussian with θ as mean and $\sqrt{\sigma^2/2\alpha}$ as standard deviation.

To calibrate the Vasicek model, the discrete form of the process is used. The coefficients c, b and δ are calibrated using Equation (5). The calibration process is simply an OLS regression of the time series $x(t_i)$ on its lagged form $x(t_{i-1})$. The OLS regression provides the maximum likelihood estimator for the parameters c, b and δ. By resolving the three equations system one gets the following α, θ and σ parameters:

$$\alpha = -\ln(b)/\Delta t \quad (11)$$

$$\theta = c/(1-b) \quad (12)$$

$$\sigma = \delta/\sqrt{(b^2-1)\Delta t/2\ln(b)} \quad (13)$$

One may compare the above formulas with the estimators derived directly via maximum likelihood in Brigo and Mercurio,[8] where for brevity, $x(t_i)$ is denoted by x_i and n is assumed to be the number of observations.

$$\widehat{b} = \frac{n\sum_{i=1}^{n} x_i x_{i-1} - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} x_{i-1}}{n\sum_{i=1}^{n} x_{i-1}^2 - \left(\sum_{i=1}^{n} x_{i-1}\right)^2}, \quad (14)$$

$$\hat{\theta} = \frac{\sum_{i=1}^{n}[x_i - \hat{b}x_{i-1}]}{n(1-\hat{b})}, \quad (15)$$

$$\widehat{\delta^2} = \frac{1}{n}\sum_{i=1}^{n}[x_i - \hat{b}x_{i-1} - \hat{\theta}(1-\hat{b})]^2. \quad (16)$$

From these estimators, using Equations (11) and (13), one obtains the estimators for the parameters α and σ.

Code 1: MATLAB® calibration of the Vasicek process

```
function ML_VasicekParams = Vasicek_calibration(V_data,dt)
  N=length(V_data);
  x = [ones(N-1,1) V_data(1:N-1)];
  ols = (x'*x)^(-1)*(x'*V_data(2:N));
  resid = V_data(2:N) - x*ols;

  c=ols (1) ;
  b=ols (2) ;
  delta=std(resid);

  alpha=-log(b)/dt;
  theta=c/(1-b);
  sigma=delta/sqrt((b^2-1)*dt/(2*log(b)));

  ML_VasicekParams = [alpha theta sigma];
```

For the Monte Carlo simulation, the discrete time expression of the Vasicek model is used. The simulation is straightforward using the calibrated coefficients α, θ and σ. One only needs to draw a random normal process ϵ_t and then generate recursively the new x_{t_i} process, where now t_i are future instants, using Equation (5), starting the recursion with the current value x_0.

MEAN REVERSION: THE EXPONENTIAL VASICEK MODEL

To address the positivity of interest rates and ensure their distribution has fatter tails, it is possible to transform a basic Vasicek process $y(t)$ in a new positive process $x(t)$.

The most common transformations to obtain a positive number x from a possibly negative number y are the square and the exponential functions.

By squaring or exponentiating a normal random variable y one obtains a random variable x with fat tails. As the exponential function increases faster than the square and is larger ($\exp(y) \gg y^2$ for positive and large y), the tails in the exponential Vasicek model below will be fatter than the tails in the squared Vasicek (and CIR as well). In other terms, a lognormal distribution (exponential Vasicek) has fatter tails than a chi-squared (squared Vasicek) or even noncentral chi-squared (CIR) distribution.

Therefore, if one assumes y to follow a Vasicek process:

$$dy_t = \alpha(\theta - y_t)dt + \sigma dW_t \qquad (17)$$

then one may transform y. If one takes the square to ensure positivity and define $x(t) = y(t)^2$, then by Itō's formula x satisfies:

$$dx(t) = (B + A\sqrt{x(t)} - Cx(t))dt + \nu\sqrt{x(t)}dW_t$$

for suitable constants A, B, C and ν. This 'squared Vasicek' model is similar to the CIR model that the subsequent section will address, so it will not be addressed further here.

If, on the other hand, to ensure positivity one takes the exponential, $x(t) = \exp(y(t))$, then one obtains a model that at times has been termed exponential Vasicek (for example in Brigo and Mercurio[8]).

Itō's formula this time reads, for x under the objective measure:

$$dx_t = \alpha x_t(m - \log x_t)dt + \sigma x_t dW_t \qquad (18)$$

where α and σ are positive. (A similar parametrisation can be used under the risk-neutral measure; also, $m = \theta + \frac{\sigma^2}{2\alpha}$, in other words, this model for x assumes that the logarithm of the short rate (rather than the short rate itself) follows an Ornstein-Uhlenbeck process.)

The explicit solution of the exponential Vasicek equation is then between any two instants $0 < s < t$:

$$x_t = \exp\left\{\theta(1 - e^{-\alpha(t-s)}) + \log(x_s)e^{-\alpha(t-s)} + \sigma e^{-\alpha t}\int_s^t e^{\alpha u} dW_u\right\} \qquad (19)$$

The advantage of this model is that the underlying process is mean-reverting, stationary in the long run, strictly positive and with a distribution skewed to the right and characterised by a fat tail for large positive values. These properties make the exponential Vasicek model a possible choice for modelling of the credit spreads. (As shortcomings for the exponential Vasicek model, one can identify the explosion problem that appears when this model is used to describe the interest rate dynamic in a pricing model involving the bank account evaluation (eg for the Eurodollar future pricing[8]). One should also point out that if this model is assumed for the short rate or stochastic default intensity, there is no closed-form bond pricing or survival probability formula, and all curve construction and valuation according to this model must be done numerically through finite difference schemes, trinomial trees or Monte Carlo simulation.)

The historical calibration and simulation of the exponential Vasicek model can be done by calibrating and simulating its log-process y as described in the previous section. To illustrate the calibration, the EMU Corporate A-rated 5-7Y index described previously is used. The histogram of the index (Figure 3) shows a clear fat-tail pattern. In addition, the log spread process is tested to be mean-reverting (again, see previously). This leads one to think about using the exponential Vasicek model with a log normal distribution and an AR(1) disturbance in the log prices.

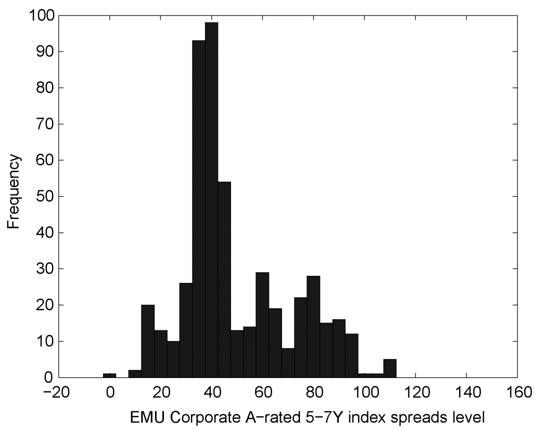

Figure 3: The histogram of EMU Corporate A-rated 5–7Y index

The OLS regression of the log-spread time series on its first lags (Equation (5)) gives the coefficients c, b and δ. Then the α, θ and σ parameters in Equation (18) are deduced using Equations (11), (12) and (13):

$$\begin{aligned} c &= 0.3625 & \alpha &= 4.9701 \\ b &= 0.9054 & \theta &= 3.8307 \\ \delta &= 0.1894 & \sigma &= 1.4061 \end{aligned}$$

Figure 5 shows how the EMU Corporate A-rated 5-7Y index time series fits into the simulation of 50,000 paths using the equation for the exponential of Equation (5) starting with the first observation of the index (first week of 1998). The simulated path illustrates that the exponential Vasicek process can occasionally reach values that are much higher than its 95 percentile level.

MEAN REVERSION: THE CIR MODEL

The model originally proposed by Cox, Ingersoll and Ross[9] was introduced in the context of the general equilibrium model of the same authors.[10] By choosing an appropriate market price of risk, the x_t process dynamics can be written with the same structure under both the objective measure and risk neutral measure. Under the objective measure, one can write:

$$dx_t = \alpha(\theta - x_t)dt + \sigma\sqrt{x_t}dW_t \qquad (20)$$

with α, θ, $\sigma > 0$ and dW_t a standard Wiener process. To ensure that x_t is

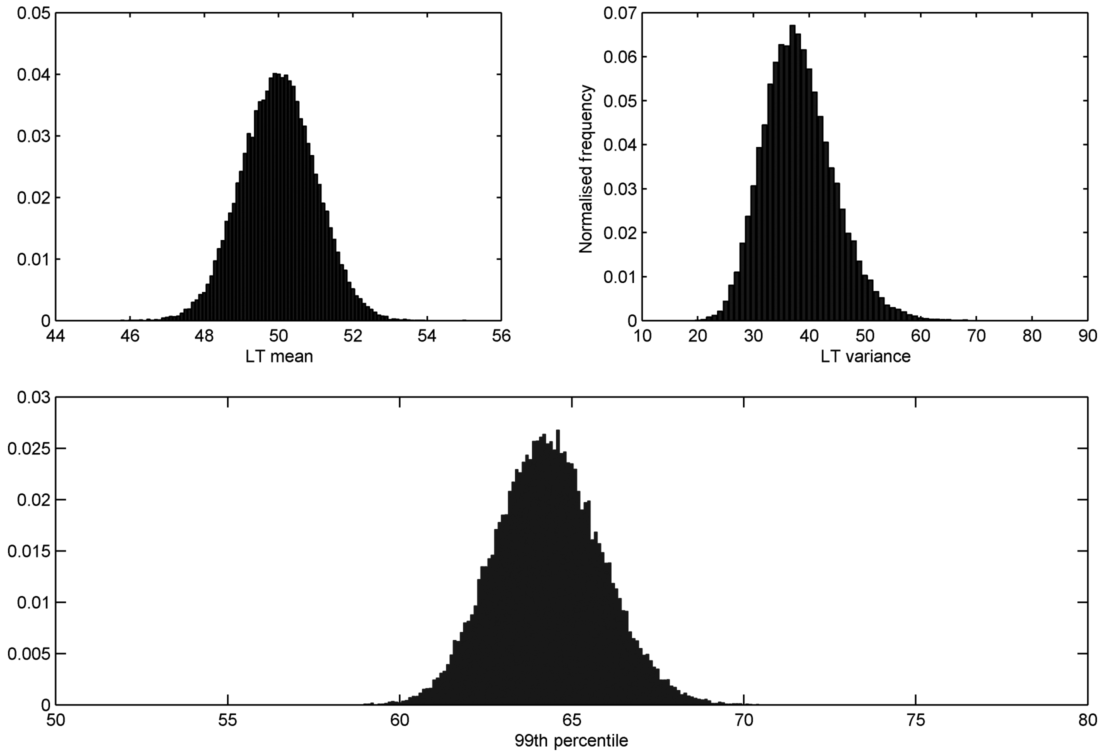

Figure 4: The distribution of the long-term mean and long-term variance and the 99th percentile estimated by several simulations of a Vasicek process with $\alpha = 8$, $\theta = 50$ and $\sigma = 25$

always strictly positive, one needs to impose $\sigma^2 \leq 2\ \alpha\theta$. By doing so, the upward drift is sufficiently large with respect to the volatility to make the origin inaccessible to x_t.

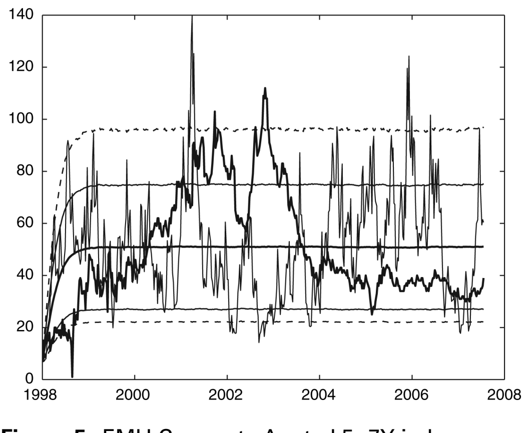

Figure 5: EMU Corporate A-rated 5–7Y index (in black) vs one simulated exponential Vasicek path (in grey). The dashed lines describe the 95th percentile and the 5th percentile of the exponential Vasicek distribution. The middle line is the distribution mean. The two remaining lines describe one standard deviation above and below the mean

The simulation of the CIR process can be done using a recursive discrete version of Equation (20) at discretisation times t_i:

$$x(t_i) = \alpha\theta\Delta t + (1 - \alpha\Delta t)x(t_{i-1}) + \sigma\sqrt{x(t_{i-1})\Delta t}\epsilon_{t_i} \qquad (21)$$

with $\epsilon \sim \mathcal{N}(0, 1)$. If the time step is too large, however, there is risk that the process approaches zero in some scenarios, which could lead to numerical problems and complex numbers surfacing in the scheme. Alternative positivity preserving schemes are discussed, for example in Brigo and Mercurio.[8] For simulation purposes, it is more precise but also more computationally expensive to draw the simulation directly from the following exact conditional distribution of the CIR process:

$$f_{CIR}(x_{t_i}|x_{t_{i-1}}) = ce^{-u-v}\left(\frac{v}{u}\right)^{q/2} I_q\left(2\sqrt{uv}\right) \quad (22)$$

$$c = \frac{2\alpha}{\sigma^2(1-e^{-\alpha\Delta t})} \quad (23)$$

$$u = cx_{t_i}e^{-\alpha\Delta t} \quad (24)$$

$$v = cx_{t_{i-1}} \quad (25)$$

$$q = \frac{2\alpha\theta}{\sigma^2} - 1 \quad (26)$$

with I_q the modified Bessel function of the first kind of order q and $\Delta t = t_i - t_{i-1}$. The process x follows a noncentral chi-squared distribution $x(t_i)|x(t_{i-1}) \sim \chi^2(2cx_{t_{i-1}}, 2q+2, 2u)$.

To calibrate this model, the maximum likelihood method is applied:

$$\underset{\alpha>0,\theta>0,\sigma>0}{\arg\max} \quad \log(\mathcal{L}) \quad (27)$$

The likelihood can be decomposed, as for general Markov processes, as a product of transition likelihoods (Equation (21) shows that $x(t_i)$ depends only on $x(t_{i-1})$ and not on previous values $x(t_{i-2})$; $x(t_{i-3})$ ... so CIR is Markov). Thus:

$$\mathcal{L} = \prod_{i=1}^{n} f_{CIR}(x(t_i)|x(t_{i-1}); \alpha, \theta, \sigma) \quad (28)$$

As this model is also a fat-tail mean-reverting one, it can be calibrated to the EMU Corporate A-rated 5-7Y index described previously. The calibration is done by maximum likelihood. To speed up the calibration, one uses the AR(1) coefficient of the Vasicek regression (Equation (5)) as an initial guess for α and then infers θ and σ using the first two moments of the process. The long-term mean of the CIR process is θ and its long-term variance is $\theta\sigma^2/(2\alpha)$. Then:

$$\alpha_0 = -\log(b)/\Delta t \quad (29)$$

$$\theta_0 = \mathbb{E}(x_t) \quad (30)$$

$$\sigma_0 = \sqrt{2\alpha_0 Var(x_t)/\theta_0} \quad (31)$$

The initial guess and fitted parameters for the EMU Corporate A-rated 5-7Y index are the following:

$$\alpha_0 = 0.8662 \quad \alpha = 1.2902$$
$$\theta_0 = 49.330 \quad \theta = 51.7894$$
$$\sigma_0 = 4.3027 \quad \sigma = 4.4966$$

Figure 6 shows how the EMU Corporate A-rated 5-7Y index time series fits into the simulation of 40,000 paths using Equation (21) starting with the first observation of the index (first week of 1998).

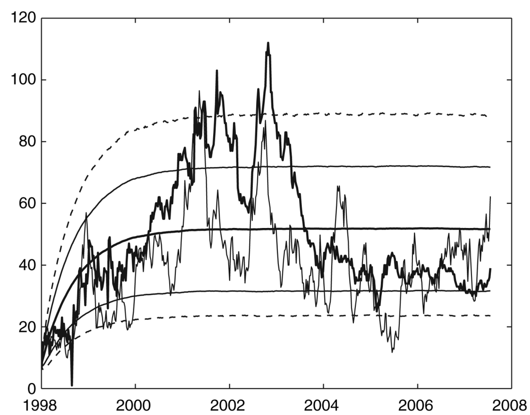

Figure 6: EMU Corporate A-rated 5–7Y index (in black) vs one simulated CIR path (in grey). The dashed lines describe the 95th percentile and the 5th percentile of the distribution. The middle line is the distribution mean. The two remaining lines describe one standard deviation above and below the mean

Code 2: MATLAB® MLE calibration of the CIR process

```
function ML_CIRparams = CIR_calibration(V_data,dt,params)
  % ML_CIRparams = [alpha theta sigma]
  N=length(V_data);
  if nargin <3
    x = [ones(N-1,1) V_data(1:N-1)];
    ols = (x'*x)^(-1)*(x'*V_data(2:N));
    m=mean(V_data); v=var(V_data);
    params = [-log(ols(2))/dt,m,sqrt(2*ols(2)*v/m)];
  end
  options = optimset('MaxFunEvals', 100000, 'MaxIter',
    100000);
  ML_CIRparams = fminsearch(@FT_CIR_LL_ExactFull, params,
    options);
  function mll = FT_CIR_LL_ExactFull(params)
    alpha = params (1) ;teta = params (2) ;sigma = params (3) ;
    c = (2*alpha)/((sigma^2)*(1-exp(-alpha*dt)));
    q = ((2*alpha*teta)/(sigma^2))-1;
    u = c*exp(-alpha*dt)*V_data(1:N-1);
    v = c*V_data(2:N);
    mll = -(N-1)*log(c)+sum(u+v-log(v./u)*q/2-...
      log(besseli(q,2*sqrt(u.*v),1))-abs(real(2*
        sqrt(u.*v))));
  end
end
```

As observed before, the CIR model has less fat tails than the exponential Vasicek, but avoids the above-mentioned explosion problem of the log-normal processes bank account. Furthermore, when used as an instantaneous short rate or default intensity model under the pricing measure, it comes with closed-form solutions for bonds and survival probabilities. Figure 7 compares the tail behaviour of the calibrated CIR process with the calibrated Vasicek and exponential Vasicek processes. Overall, the exponential Vasicek process is easier to estimate and to simulate, and features fatter tails, so it is preferable to CIR.

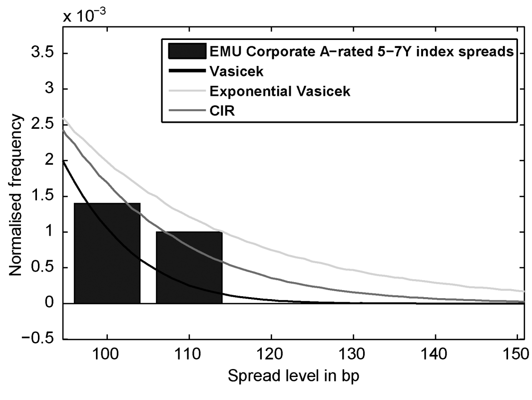

Figure 7: The tail of the EMU Corporate A-rated 5–7Y index histogram compared with the fitted Vasicek process probability density (in black), the fitted exponential Vasicek process probability density (in light grey) and the fitted CIR process probability density (in dark grey)

However, if used in conjunction with pricing applications, the tractability of CIR and non-explosion of its associated bank account can make CIR preferable in some cases.

Remark 1 (Confidence levels for parameter estimates and distribution percentiles): *None of the three mean-reverting models provides analytical solutions for the parameter distribution, which has to be derived using the bootstrapping technique described in a Remark in the authors' previous paper. In general, the bootstrapping approach would be to simulate a sample path with the estimated best-fit parameters and re-estimate the parameters for each sample path, again as described in a Remark in the authors' previous paper. However, the three models described above do provide a closed-form solution for the marginal density function for the underlying risk factor. This allows one to re-sample the risk factor for the best-fit parameters and then re-estimate the parameters without re-simulating the whole path of the risk factor in time.*

MEAN REVERSION AND FAT TAILS TOGETHER

In the previous paper and sections, two possible features of financial time series were studied, namely fat tails and mean reversion. The first feature accounts for extreme movements in a short period, whereas the second one accounts for a process that over long periods approaches some reference mean value with a bounded variance. The question in this section is: can one combine the two features and have a process that can feature extreme movements in short periods, beyond the Gaussian statistics, and in the long run feature mean reversion?

In this section, an example of one such model is formulated: a mean-reverting diffusion process with jumps. In other terms, rather than adding jumps to a GBM, one may decide to add jumps to a mean-reverting process such as Vasicek. Consider:

$$dx(t) = \alpha(\theta - x(t))dt + \sigma dW(t) + dJ_t \tag{32}$$

where the jumps J are defined as:

$$J(t) = \sum_{j=1}^{N(t)} Y_j, \quad dJ(t) = Y_{N(t)} dN(t)$$

where $N(t)$ is a Poisson process with intensity λ and Y_j are iid random variables modelling the size of the j-th jump, independent of N and W.

The way the previous equation is written is slightly misleading, as the mean in the shocks is nonzero. To have zero-mean shocks, so as to be able to read the long-term mean level from the equation, one needs to rewrite it as:

$$\begin{aligned} dx(t) = {} & \alpha(\theta_Y - x(t))dt + \sigma dW(t) \\ & + [dJ_t - \lambda \mathbb{E}[Y]dt] \end{aligned} \tag{33}$$

where the actual long-term mean level is:

$$\theta_Y := \theta + \lambda \mathbb{E}[Y]/\alpha,$$

and can be much larger than θ, depending on Y and λ. For the distribution of the jump sizes, Gaussian jump sizes are considered, but calculations are also possible with exponential or other jump sizes, particularly with stable or infinitely divisible distributions.

Otherwise, for a general jump size distribution, the process integrates into:

$$x_t = m_x(x_s, t-s) + \sigma e^{-\alpha t}\int_s^t e^{\alpha z} dW_z + e^{-\alpha t}\int_s^t e^{\alpha z} dJ_z$$

$$m_x(x_s, t-s) := \theta(1-e^{-\alpha(t-s)}) + x_s e^{-\alpha(t-s)},$$

$$v_x(t-s) := \frac{\sigma^2}{2\alpha}(1-e^{-2\alpha(t-s)}) \tag{34}$$

where m_x and v_x are, respectively, mean and variance of the Vasicek model without jumps.

Notice that taking out the last term in Equation (34) one still gets Equation (4), the earlier equation for Vasicek. Furthermore, the last term is independent of all others. As a sum of independent random variables leads to convolution in the density space and to product in the moment generating function space, the moment generating function of the transition density between any two instants is computed:

$$M_{t|s}(u) = \mathbb{E}_{x(s)}[\exp(ux(t))] \tag{35}$$

which in the model is given by

$$M_{t|s}(u) = \exp\left\{m_x(x_s, t-s)u + \frac{v_x(t-s)}{2}u^2 + \lambda\int_s^t (M_Y(ue^{-\alpha(t-z)})-1)dz\right\} \tag{36}$$

where the first exponential is due to the diffusion part (Vasicek) and the second one comes from the jumps.

The characteristic function is a better notion than the moment generating function. For the characteristic function one would obtain:

$$\begin{aligned}\phi_{t|s}(u) &= \mathbb{E}_{x(s)}[\exp(iux(t))] \\ &= \exp\left\{m_x(x_s, t-s)iu - \frac{v_x(t-s)}{2}u^2 + \lambda\int_s^t (\phi_Y(ue^{-\alpha(t-z)})-1)dz\right\}\end{aligned} \tag{37}$$

where i is the imaginary unit from complex numbers theory. The characteristic function always exists, whereas the moment generating function may not exist in some cases.

Now, one knows that to write the likelihood of the Markov process it is enough to know the transition density between subsequent instants. In this case, such density between any two instants $s = t_{i-1}$ and $t = t_i$ can be computed by the above moment generating function through inverse Laplace transform (or better from the above characteristic function through inverse Fourier transform).

There are however some special jump size distributions where one does not need to use transform methods. These include the Gaussian jump size case, which is covered in detail below.

Suppose then that the jump size Y is normally distributed, with mean μ_Y and variance σ_Y^2. For small Δt, one knows that there will be possibly just one jump. More precisely, there will be one jump in $(t - \Delta t, t)$ with probability $\lambda\Delta t$, and no jumps with probability $1 - \lambda\Delta t$.

Code 3: MATLAB® calibration of the Vasicek process with jumps

```
function
[alpha_,theta_,sigma_,lambda_,mu_y,sigma_y]=VasicekJumps_calibration
(V_data,dt,initparams)
  options=optimset('MaxFunEvals',100000,'MaxIter',100000);
  parameters = fminsearch(@FT_LL, parameters, options);
  function mll = FT_LL(x)
    alpha_=abs(x (1)) ;theta_=abs(x (2)) ;sigma_=abs(x (3));
    lambda_=abs(x (4)) ;mu_y=abs(x (5)) ;sigma_y=abs(x (6)) ;
    epsil = V_data(2:end)-(1-alpha_*dt)*V_data(1:end-1)-theta_*alpha_*dt;
    LL = 0; BP = lambda_*dt;
    for time=1:length(epsil)
      condV = dt*sigma_^2; condM= -mu_y*lambda_*dt;
      PD = (1-BP)*exp(-(epsil(time)-condM)^2/condV/2)/sqrt(pi*2*condV);
      condV = condV+sigma_y^2; condM = condM+mu_y;
      PD = PD+BP*exp(-(epsil(time)-condM)^2/condV/2)/sqrt(pi*2*condV);
      LL = LL + log(PD);
    end
    mll=-LL;
    end
  end
```

Furthermore, with the approximation:

$$\int_{t-\Delta t}^{t} e^{\alpha z} dJ_z \approx e^{\alpha(t-\Delta t)} \Delta J(t),$$
$$\Delta J(t) := J(t) - J(t-\Delta t). \qquad (38)$$

one can write the transition density between $s = t - \Delta t$ and t as:

$$f_{x(t)|x(t-\Delta t)}(x) = \lambda \Delta t f_N(x; m_x(\Delta t, x(t-\Delta t)) + \mu_Y, v_x(\Delta t) + \sigma_Y^2) + (1-\lambda\Delta t) f_N(x; m_x(\Delta t, x(t-\Delta t)), v_x(\Delta t)) \qquad (39)$$

This is a mixture of Gaussians, weighted by the arrival intensity λ. It is known that mixtures of Gaussians feature fat tails, so this model is capable of reproducing more extreme movements than the basic Vasicek model.

To check that the process is indeed mean-reverting, in the Gaussian jump size case, through straightforward calculations, one can prove that the exact solution of Equation (33) or equivalently Equation (34) satisfies:

$$\mathbb{E}[x_t] = \theta_Y + (x_0 - \theta_Y)e^{-\alpha t} \qquad (40)$$

$$Var[x_t] = \frac{\sigma^2 + \lambda(\mu_Y^2 + \sigma_Y^2)}{2\alpha}(1 - e^{-2\alpha t}) \qquad (41)$$

This way one sees that the process is indeed mean-reverting, since as time increases, the mean tends to θ_Y while the variance remains bounded, tending to $(\sigma^2 + \lambda(\mu_Y^2 + \sigma_Y^2))/(2\alpha)$. Notice that, either letting λ or μ_Y and σ_Y go to zero (no jumps) gives back the mean and variance for the basic Vasicek process without jumps.

It is important for the jump mean μ_Y not to be too close to zero, or one again obtains almost symmetric Gaussian shocks that are already there in Brownian motion and the model does not differ much from a pure diffusion Vasicek model. On the contrary, it is best to assume μ_Y to be significantly positive or negative with a relatively small variance, so the jumps will be markedly positive or negative. If one aims at modelling both markedly positive and negative jumps, one may add a second jump process $J_-(t)$ defined as:

$$dJ_-(t) = Z_{M(t)} dM(t)$$

where now M is an arrival Poisson process with intensity λ_- counting the jumps, independent of J, and Z are again iid normal random variables with mean μ_Z and variance σ_Z^2, independent of M, J, W. The overall process reads:

$$\begin{aligned} dx(t) &= \alpha(\theta - x(t))dt + \sigma dW(t) \\ &\quad + dJ(t) - dJ_-(t) \text{ or} \\ dx(t) &= \alpha(\theta_J - x(t))dt + \sigma dW(t) \\ &\quad + [dJ(t) - \lambda \mu_Y dt] \\ &\quad - [dJ_-(t) - \lambda_- \mu_Z dt], \\ \theta_J &:= \theta + \frac{\lambda \mu_Y - \lambda_- \mu_Z}{\alpha} \end{aligned} \tag{42}$$

This model would feature as transition density, over small intervals $[s, t]$:

$$\begin{aligned} & f_{x(t)|x(t-\Delta t)}(x) \\ &\quad = \lambda \Delta t f_N(x; m_x(\Delta t, x(t-\Delta t)) \\ &\quad + \mu_Y, v_x(\Delta t) + \sigma_Y^2) \\ &\quad + \lambda_- \Delta t f_N(x; m_x(\Delta t, x(t-\Delta t)) \\ &\quad - \mu_Z, v_x(\Delta t) + \sigma_Z^2) \\ &\quad + (1 - (\lambda + \lambda_-)\Delta t) \\ &\quad f_N(x; m_x(\Delta t, x(t-\Delta t)), v_x(\Delta t)) \end{aligned} \tag{43}$$

This is now a mixture of three Gaussian distributions, which in principle allows for more flexibility in the shapes the final distribution can assume.

Again, one can build on the Vasicek model with jumps by squaring or exponentiating it. Indeed, even with positive jumps, the basic model still features positive probability for negative values. One can solve this problem by exponentiating the model: define a stochastic process y following the same dynamics as x in Equation (42) or Equation (32), but then define the basic process x as $x(t) = \exp(y(t))$. Then historical estimation can be applied working on the log-series, by estimating y on the log-series rather than on the series itself.

From the examples in Figures 8 and 9, one sees the goodness of fit of Vasicek with jumps, and note that the Vasicek model with positive-mean Gaussian size jumps can be an alternative to the exponential Vasicek. Indeed, Figure 8 shows that the tail of exponential Vasicek almost exceeds the series, whereas Vasicek with jumps can be a weaker tail alternative to exponential Vasicek, somehow similar to CIR, as one sees from Figure 9. However, by playing with the jump parameters or better by

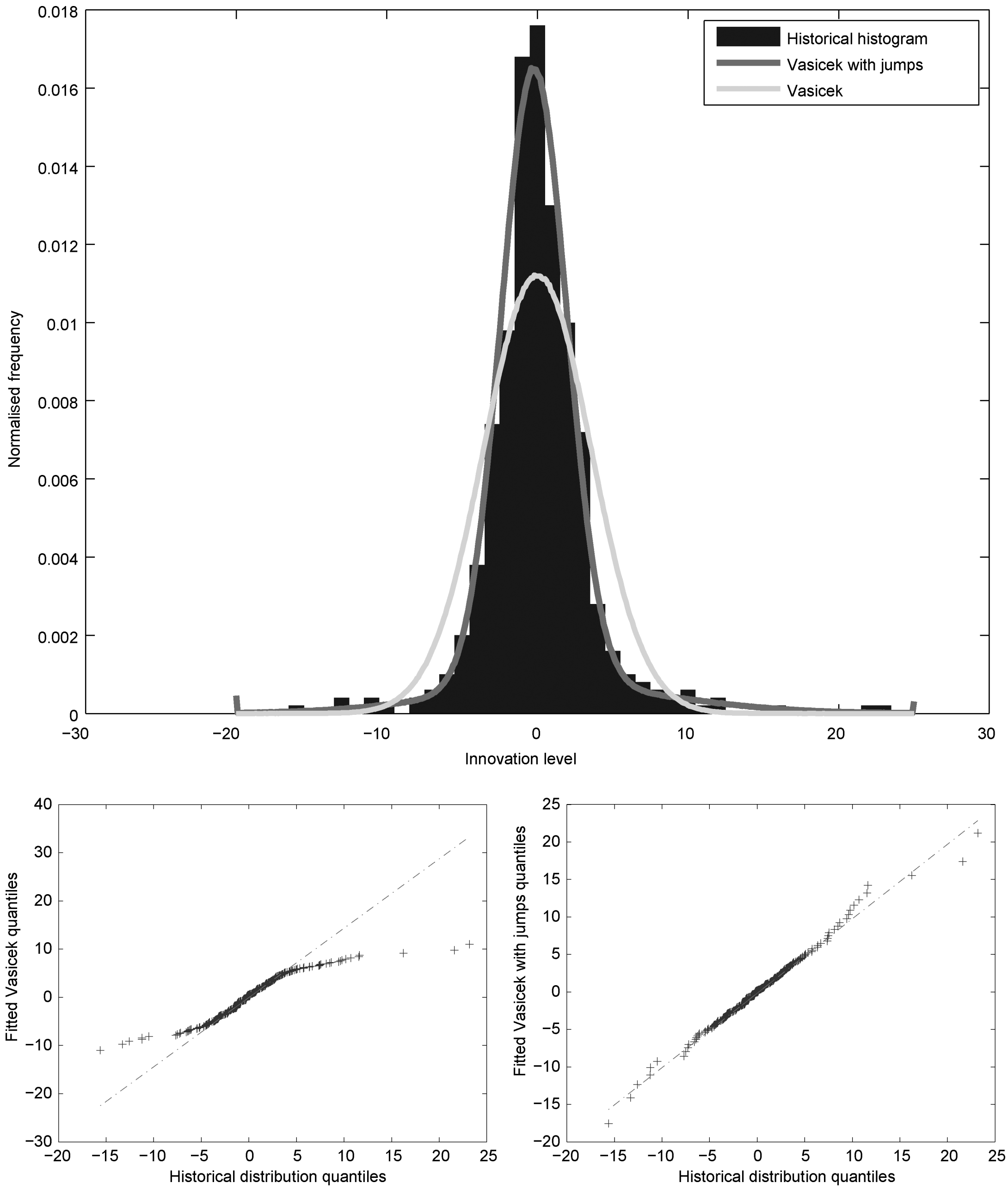

Figure 8: EMU Corporate A-rated 5–7Y index charts. The top chart compares the innovation fit of the Vasicek process (in light grey) to the fit of the Vasicek process with normal jumps (in dark grey). The bottom left figure is the QQ-plot of the Vasicek process innovations against the historical data innovations. The bottom right figure is the QQ-plot of the Vasicek process with normal jumps innovations against the historical data innovations

exponentiating Vasicek with jumps, one can easily obtain a process with tails that are fatter than the exponential Vasicek when needed. While, given Figure 7, it may not be necessary to do so in this example, in general the need may occur in different applications.

More generally, given Vasicek with positive and in case negative jumps, by playing with the jump parameters and in

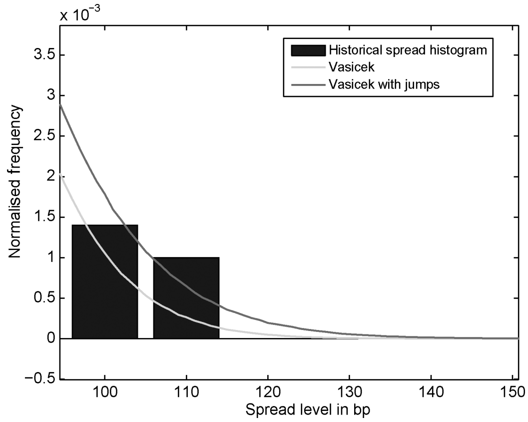

Figure 9: The tail of the EMU Corporate A-rated 5–7Y index histogram compared with the fitted Vasicek process probability density (in light grey) and the fitted Vasicek with jumps process probability density (in dark grey)

case by squaring or exponentiating, one can obtain a large variety of tail behaviours for the simulated process. This may be sufficient for several applications. Furthermore, using a more powerful numerical apparatus, especially inverse Fourier transform methods, one may generalise the jump distribution beyond the Gaussian case and attain an even larger flexibility.

References

1 Dickey, D. and Fuller, W. (1979) 'Distribution of the estimators for autoregressive time series with a unit root', *Journal of the American Statistical Association*, Vol. 74, pp. 427–431.

2 Said, S. and Dickey, D. (1984) 'Testing for unit roots in autoregressive moving average models of unknown order', *Biometrika*, Vol. 71, No. 3, pp. 599–607.

3 Phillips, P. and Perron, P. (1988) 'Testing for unit root in time series regression', *Biometrica*, Vol. 75, No. 2, pp. 335–346.

4 Poterba, J. and Summers, L. J. (1988) 'Mean reversion in stock prices: Evidence and implications', *Journal of Financial Economics*, Vol. 22, No. 1, pp. 27–59.

5 Cochrane, J. H. (2001) 'Asset Pricing', Princeton University Press, Princeton, NJ.

6 Kwiatkowski, D., Phillips, P. C. B., Schmidt, P. and Shin, Y. (1992) 'Testing the null hypothesis of stationarity against the alternative of a unit root', *Journal of Econometrics*, Vol. 54, Nos. 1–3, pp. 159–178.

7 Vasicek, O. A. (1977) 'An equilibrium characterization of the term structure', *Journal of Financial Economics*, Vol. 5, No. 2, pp. 177–188.

8 Brigo, D. and Mercurio, F. (2006) 'Interest Rate Models: Theory and Practice — With Smile, Inflation and Credit', Springer Verlag, New York.

9 Cox, J., Ingersoll, J. and Ross, S. A. (1985) 'A theory of the term structure of interest rates', *Econometrica*, Vol. 53, No. 2, pp. 385–408.

10 Cox, J., Ingersoll, J. and Ross, S. A. (1985) 'An intertemporal general equilibrium model of asset prices', *Econometrica*, Vol. 53, No. 2, pp. 363–384.

APPENDIX: NOTATION

General

$\mathbb{E}$	expectation
Var	variance
SDE	stochastic differential equation
$P(A)$	probability of event A
$M_X(u)$	moment generating function of the random variable X computed at u
$\varphi_X(u)$	characteristic function of X computed at u
q	quantiles

$f_X(x)$	density of the random variable X computed at x
$F_X(x)$	cumulative distribution function of the random variable X computed at x
$f_X\|_{Y=y}(x)$ or $f_X\|_Y(x;y)$:	density of the random variable X computed at x conditional on the random variable Y being y
$N(m,\ V)$	normal (Gaussian) random variable with mean m and variance V
$f_N(x;m,\ V)$	normal density with mean m and variance V computed at x $= (2\pi V)^{-\frac{1}{2}} \exp\left(\frac{-(x-m)^2}{2V}\right)$
Φ	standard normal cumulative distribution function
$f_P(x;L)$	Poisson probability mass at $x = 0,1,2,\ldots$ with parameter $L=$ $\exp(-L)(L)^x/(x!)$
$f_\Gamma(x;\kappa,\ \zeta)$	gamma probability density at $x \geq 0$ with parameters κ and $\zeta := x^{\kappa-1}\exp(-x/\zeta)/(\zeta^\kappa\Gamma(\kappa))$
$\sim$	distributed as

Statistical estimation

CI	confidence interval
ACF	autocorrelation function
PACF	partial ACF
MLE	maximum likelihood estimation
OLS	ordinary least squares
Θ	parameters for MLE
$\mathcal{L}(\Theta)(X_1,\ldots,X_n)$	$:= f x_1, X_2, \ldots, X_n\|\Theta(X_1,\ldots,X_n)$ or $f x_1,\ldots,x_n\ (x_1,\ldots,x_n;\Theta)$: likelihood distribution function with parameter Θ in the assumed distribution, for random variables $X_1, \ldots, X_n$, computed at $x_1, \ldots, x_n$
$\mathcal{L}^*(\Theta)(X_1,\ldots,X_n)$	$:= \log(\mathcal{L}(\Theta)(X_1,\ldots,X_n))$, log-likelihood
$\hat{\Theta}$	estimator for Θ

Geometric Brownian motion (GBM)

ABM	arithmetic Brownian motion
GBM	geometric Brownian motion
μ, $\hat{\mu}$	drift parameter and sample drift estimator of GBM
σ, $\hat{\sigma}$	volatility and sample volatility
W_t or $W(t)$	standard Wiener process

Fat tails: GBM with jumps, GARCH models

J_t, $J(t)$	jump process at time t
N_t, $N(t)$	Poisson process counting the number of jumps, jumping by one whenever a new jump occurs, with intensity λ
$Y_1, Y_2, \ldots, Y_i, \ldots$	jump amplitudes random variables

Time, discretisation, maturities, time steps

Δt	time step in discretisation

s, t, u	three time instants, $0 \leq s \leq t \leq u$
$0 = t_0, t_i, \ldots, t_i, \ldots$	a sequence of equally spaced time instants, $t_i - t_{i-1} = \Delta t$
T	maturity
n	sample size

Variance gamma models

VG	variance gamma
$\bar{\mu}$	log-return drift in calendar-time
$g(t) \sim \Gamma(t/\nu, \nu)$	transform of calendar time t into market activity time
$\bar{\theta}$	log-return drift in market-activity time
$\bar{\sigma}$	volatility

Vasicek and CIR mean-reverting models

θ, α, σ	long-term mean reversion level, speed and volatility