Simulating Term Structure of Interest Rates with Arbitrary Marginals

First draft: September, 2009

Abstract

Decision models under uncertainty need to be fed with scenarios of the interest rate curve. Such scenarios have to comply, as close as possible, with the empirical distribution of each rate. Simulation models of the term structure usually assume that the conjugate distribution of the interest rates is lognormal. Dynamic models, like vector auto-regression, implicitly postulate that the logarithm of the interest rates is normally distributed.

Statistical analyses have, however, shown that stationary transformations (yield changes) of the interest rates are substantially leptokurtic, thus posing serious doubts on the reliability of the available models.

We propose in this paper a VARTA model to simulate term structures of the interest rates with arbitrary marginals. We will show that such an approach is able to simulate paths of the entire term structure with distributional properties very close to those found in the empirical data.

Keywords: Simulation; Term structure of interest rates; Vector Autoregressive models; Fat tails.

1 Introduction

Simulation models of economic and financial factors are nowadays widely used to support decisions or to assess risk exposures. For example, scenario generators are essential parts of the stochastic programming methodology aimed at planning under uncertainty. Value-at-Risk or stress testing of financial positions are routinely performed and based on simulation models of the major factors affecting the company portfolios.

A key economic factor is the term structure of the interest rates. The need for a consistent description of the interest rate curve is justified by the rapid development of decision models for insurance policies, pension funds and optimal debt allocation.

Interest rate curves are particularly complex to simulate. Linear correlation between the short term and long term rate, and correlations at different lags are very high. Broadly speaking, it is not just the random motion of a point to be modeled, but rather the evolution of a whole curve (corresponding to different maturities) which needs to be simulated.

Another important issue, related to the assessment of risk and its evaluation, is that scenarios have to be simulated for medium-long time horizons. For example, minimum guarantee options embedded in insurance contracts expire, on average, after 10 years (Consiglio and De Giovanni, 2008; Consiglio et al., 2007, 2008); stochastic programming models for pension funding (Cariño et al., 1998; Mulvey and Thorlacius, 1998) have to comply with 20 or 30 years simulation span. For such applications, it is fundamental that the tails of the distribution are more realistic as possible since they will determine the risk structure of the selected portfolios.

Vector auto-regression models (VAR), which takes into account of cross and lagged correlations, are quite suitable to describe the term structure evolution (see Ang and Piazzesi, 2003, for recent developments). VAR models can be used for simulation by adding to the estimated *deterministic* part a sample draw from the *innovation* term. However, the assumption that the white noise is multivariate Gaussian implies that the unconditional marginal distributions are Gaussian too.

Empirical observations show that stationary transformations of the interest rates (yield changes) significantly deviates from the Gaussian distribution, and that the degree of such a deviation depends on the maturity of the yield. A possible resort would be that of adding a noise vector with a multivariate fat-tailed distribution (i.e., a multivariate t-Student). Such an approach is not a viable solution because *i)* marginal distributions of the rates show different degree of fatness, and *ii)* estimates of the VAR parameters do not coincides with those obtained assuming a normal distribution of the error term[1].

An alternative approach for scenario generation is the Høyland and Wallace (2001)'s model who builds multivariate trees by matching moments, extreme events and autocorrelations. Such an model has the merit of being distribution-free, and, thanks to the note by Klaassen (2002), of being able

[1]Recently, Ni and Sun (2005) proposes a MCMC procedure to estimate VAR parameters with a Student's t distribution of the error term, and unknown ν.

to generate arbitrage free scenarios. The last feature is fundamental to price derivative securities in incomplete markets (King, 2002). However, the higher is the number of the parameters to fit, the higher must be the number of arcs springing from each node. When multiple periods are needed to represent the time flow, the size of the tree grows exponentially.

Other simulation models are based on the copula approach (see Cherubini et al., 2004, for applications to finance). The celebrated Sklar's theorem shows that any n-dimensional joint distribution function may be decomposed into its n marginal distributions, and a copula, which completely describes the dependence between the n variates. Copula–based models for autoregressive processes are studied in Patton (2006) and Chen and Fan (2006). However, their approach is mainly focused on univariate Markov model and the extension to the multivariate case is not given.

A paper dealing with copulas in the context of time dependence is by Fermanian and Scaillet (2003). They propose a kernel approach to determine nonparametric estimates of the copula. However, their method is mainly focused on estimating copula when data are dependent, and no explicit VAR dynamics is supplied.

The objective of this paper is to build a model with the following features:

- fit the unconditional distribution of some representative rates;
- being able to capture the dynamic properties of the interest rate curves and the relations across maturities;
- use correlations to describe the relations at different lags and across each rate.

The first requirement is motivated by the need to simulate scenarios that, on average, mirror the empirical characteristics of the interest rates. The second requirement is needed to describe smooth interest rate curves whose dynamics resemble that observed in reality. Finally, we choose correlations to account for dependency because it is a measure well-understood by practitioners, and because through correlations we can fully characterize VAR models. We are aware that correlation is a weak measure of dependence. As pointed out in Embrechts et al. (2002), knowledge of the marginals and correlation matrix is not sufficient to determine the underneath joint distribution.

The *VAR to Anything* (VARTA) model, developed by Biller and Nelson (2003), fits the above cited features. The contribution of this paper is to show that a VARTA approach is able to generate scenarios of the interest rates whose marginal distributions replicate empirical stylized facts.

The paper is organized as follows: in Section 2 we analyzes a VAR model with arbitrary marginals. Section 3 describes the data used and provides implementation notes. Section 4 reports tests of the model and validation. Section 5 contains conclusions as well as some suggestions for future research.

2 A simulation model with arbitrary marginals

We describe here a simulation model to generate time-series of interest rate curves with arbitrary marginal unconditional distributions. Details about modeling multivariate time-series using a vector auto-regression technique and arbitrary marginals can be found in Biller and Nelson (2003).

We assumed that the term structure is described by k representative interest rates, and denoted by $\mathbf{X}_t$ the stationary time-series process of such variates, hence, $\mathbf{X}_t = (X_{1,t}, X_{2,t}, \ldots, X_{k,t})'$.

To characterize a wide variety of distributional shapes, we let each $X_{i,t}$ to have a marginal distribution from the Johnson translation system (Johnson, 1949). The cumulative distribution function (cdf) of a Johnson translation system has the following form,

$$F_X(x) = \Phi\{\gamma + \delta f[(x - \xi)/\lambda]\}, \tag{1}$$

where, γ and δ are shape parameters, ξ is a location parameter, and λ is a scale parameter. The function $f(\cdot)$ is defined by one of the following transformation:

$$f(y) = \begin{cases} \log(y) & \text{for the } S_L \text{ (lognormal) family} \\ \log\left(y + \sqrt{y^2+1}\right) & \text{for the } S_U \text{ (unbounded) family} \\ \log\left(\dfrac{y}{1+y}\right) & \text{for the } S_B \text{ (bounded) family} \\ y & \text{for the } S_N \text{ (normal) family} \end{cases} \tag{2}$$

There is a unique family (choice of f) for each feasible combination of the skewness and the kurtosis that determine the parameters γ and δ. Any mean and variance can be attained by any of the families by the manipulation of the parameters λ and ξ.

Dependency across rates, at lags $h = 0, 1, 2, \ldots, p$, is described by the linear correlation coefficient,

$$\rho_{\mathbf{X}}(i, j, h) = \frac{E[X_{i,t}X_{j,t-h}] - E[X_{i,t}]E[X_{j,t-h}]}{\sqrt{\mathrm{Var}[X_{i,t}]\mathrm{Var}[X_{j,t-h}]}}, \tag{3}$$

for all $i, j = 1, 2, \ldots, k$, and $i \neq j$ when $h = 0$.

Following Biller and Nelson, we denoted by $\mathbf{Z}_t$ a k-variate standard gaussian VAR process of order p. The conjugate distribution $\mathbf{Z}_{t+h}$, $h = 0, 1, 2, \ldots, p$, plays the same role of a gaussian copula (see Biller, 2006, for further details about the relation between copulas and VARTA). It provides the associative structure for the process with arbitrary marginals $\mathbf{X}_t$. Recall that, to simulate a gaussian copula, you need first to generate a multivariate gaussian random sample, say $\mathbf{Z}$, and use the probability-integral transformation $\mathbf{U} = \Phi(Z)$ to obtain the uniform dependent variates. A further nonlinear transformation will determine the sample $\mathbf{X}$, with arbitrary marginals. As pointed out in Embrechts et al. (2002), linear correlation has the serious deficiency that it is not invariant under nonlinear strictly increasing transformations. This implies that the empirical correlations will be different from those of the copula. Even rank correlations would not deliver the observed correlations. As we will see, the autoregressive structure of the process $\mathbf{X}_t$ will be inferred from the observed correlations, therefore, any copula based approach is deemed to fail.

The procedure devised by Biller and Nelson consist of adjusting the correlations of the base process $\mathbf{Z}_t$ in order to match the empirical correlations, and therefore the autoregressive structure, of the VARTA process $\mathbf{X}_t$.

2.1 Fitting a VARTA process

A k-variate p–th order VAR$_k(p)$ model is defined as follows:

$$\mathbf{Z}_t = \alpha_1 \mathbf{Z}_{t-1} + \alpha_2 \mathbf{Z}_{t-2} + \ldots + \alpha_p \mathbf{Z}_{t-p} + \mathbf{u}_t, \tag{4}$$

where, $\mathbf{Z}_t$ is $(k \times 1)$ random vector, α_i are $(k \times k)$ fixed autoregressive coefficient matrices, and $\mathbf{u}_t$ is a k-dimensional white noise vector such that,

$$E[\mathbf{u}_t] = \mathbf{0} \text{ and } E[\mathbf{u}_t \mathbf{u}_{t-h}] = \begin{cases} \mathbf{\Sigma}_u & \text{if } h = 0 \\ \mathbf{0}_{(k\times k)} & \text{otherwise} \end{cases} \tag{5}$$

A VAR$_k(p)$ model is fully described by its autocovariance structure,

$$\mathbf{\Sigma_Z} = \begin{pmatrix} \mathbf{\Sigma}_Z(0) & \mathbf{\Sigma}_Z(1) & \ldots & \mathbf{\Sigma}_Z(p-2) & \mathbf{\Sigma}_Z(p-1) \\ \mathbf{\Sigma}'_Z(1) & \mathbf{\Sigma}_Z(0) & \ldots & \mathbf{\Sigma}_Z(p-3) & \mathbf{\Sigma}_Z(p-2) \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{\Sigma}'_Z(p-1) & \mathbf{\Sigma}'_Z(p-2) & \ldots & \mathbf{\Sigma}'_Z(1) & \mathbf{\Sigma}_Z(0) \end{pmatrix} \tag{6}$$

To determine α_i, for $i = 1, 2, \ldots, p$ and $\mathbf{\Sigma}_u$, we simply solved the multivariate Yule–Walker equations given by

$$\alpha = \mathbf{\Sigma}\mathbf{\Sigma}_{\mathbf{Z}}^{-1} \tag{7}$$

$$\mathbf{\Sigma}_u = \mathbf{\Sigma}_Z(0) - \alpha_1 \mathbf{\Sigma}'_Z(1) - \ldots - \alpha_p \mathbf{\Sigma}'_Z(p) \tag{8}$$

where, $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_p)_{(k\times kp)}$, and $\mathbf{\Sigma} = (\mathbf{\Sigma}_Z(1), \mathbf{\Sigma}_Z(2), \ldots, \mathbf{\Sigma}_Z(p))_{(k\times kp)}$ (see Lütkepohl, 2005, pag. 85).

Let $\mathbf{U}_t = (U_{1,t}, U_{2,t}, \ldots, U_{k,t})'$ be a series of autocorrelated uniform random variables. We obtain such a series via the probability-integral transformation $U_{i,t} = \Phi(Z_{i,t})$, where $\Phi(\cdot)$ is the cdf of the standard normal distribution. We then determine the i-th time series via the transformation $X_{i,t} = F_{X_i}^{-1}[U_{i,t}]$, where F_{X_i} is the Johnson-type cdf. The latter transformation ensures that the i-th marginal X_i has the required distribution.

As highlighted in Section 1, such transformations distort the correlations, in sense that $\mathbf{\Sigma}_Z$ does not match $\mathbf{\Sigma}_X$. The fitting of a VARTA process consists in finding $pk^2 + k(k-1)/2$ correlations of $\mathbf{\Sigma}_Z$ such that, $\mathbf{Z}_t$ is a stationary $\mathrm{VAR}_k(p)$ process and $\mathbf{\Sigma}_X$ are the empirical correlations[2].

The problem of finding the correlation $\rho_Z(i,j,h)$ that matches $\rho_X(i,j,h)$ is formulated by observing that *(i)* each pair $(Z_{i,t}, Z_{j,t-h})'$ has a nonsingular standard bivariate normal distribution, and that *(ii)* the correlation between each pair $(X_{i,t}, Xj, t-h)'$ can be parameterized as a function of ρ_Z. Biller and Nelson show that the parametric function $c_{ijh}(\cdot)$, linking ρ_Z to ρ_X, is given by,

$$c_{ijh}(\rho_Z) = \frac{\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} F_{X_i}^{-1}[\Phi(z_{i,t})]F_{X_j}^{-1}[\Phi(z_{j,t-h})]f(z_{i,t}, z_{j,t-h}, \rho_Z)dz_{i,t}dz_{j,t-h} - \mu_i\mu_j}{\sigma_i\sigma_j}, \tag{9}$$

where $f(z_{i,t}, z_{j,t-h}, \rho_Z)$ is the density function of a standard bivariate normal distribution, μ_i, μ_j and σ_i, σ_j are, respectively, the expected values and standard deviations of X_i and X_j.

We summarize here the necessary steps to fit a VARTA process:

1. Fit to each variate X_i, $i = 1, 2, \ldots, k$ a Johnson distribution (e.g., by a moment matching method).

2. Solve for the $pk^2 + k(k-1)/2$ correlations of $\mathbf{\Sigma}_Z$ the matching problem $c_{ijh}[\rho_Z(i,j,h)] \approx \rho_X(i,j,h)$.

Note that, the double integral in (9) has to be solved numerically. The choice of a Jonhnson translation system to model the marginal distributions

[2]Biller and Nelson prove that the stationarity condition of the $\mathbf{Z}_t$ process implies stationarity of the $\mathbf{X}_t$ process.

facilitate such a task. In fact, $F_X^{-1}[\Phi(z)] = \xi + \lambda f^{-1}[(z-\gamma)/\delta]$, where

$$f^{-1}(a) = \begin{cases} e^a & \text{for the } S_L \text{ (lognormal) family} \\ \dfrac{e^a - e^{-a}}{2} & \text{for the } S_U \text{ (unbounded) family} \\ \dfrac{1}{1+e^{-a}} & \text{for the } S_B \text{ (bounded) family} \\ a & \text{for the } S_N \text{ (normal) family} \end{cases} \quad (10)$$

thus avoiding the explicit evaluation of $\Phi(z)$.

The function $c_{ijh}(\rho_Z)$ holds some properties ensuring that the matching problem converges smoothly. In particular:

1. For any distributions F_{X_i} and F_{X_j}, $c_{ijh}(0) = 0$ and $\rho_{\mathbf{Z}}(i,j,h) \geq 0 (\leq 0)$ implies that $c_{ijh}[\rho_{\mathbf{Z}}(i,j,h)] \geq 0 (\leq 0)$.

2. For any distributions F_{X_i} and F_{X_j}, $c_{ijh}(-1) = \underline{\rho}_{ij}$ and $c_{ijh}(-1) = \overline{\rho}_{ij}$, where $\underline{\rho}_{ij}$ and $\overline{\rho}_{ij}$ are the minimum and maximum bivariate correlation attainable.

3. The function $c_{ijh}(\rho_{\mathbf{Z}}(i,j,h))$ is nondecreasing for $-1 \leq \rho_{\mathbf{Z}}(i,j,h) \leq 1$.

Therefore, being $c_{ijh}(\rho_Z)$ a bounded, continuous, nondecreasing function, any line search procedure will be able to find ρ_Z such that $c_{ijh}(\rho_Z) \approx \rho_X$.

However, some limitations need to be highlighted. First, the Höffding-Frechét's theorem (see Embrechts et al., 2002, Theorem 4) shows that given two random variables with fixed marginals and unspecified dependence structure, the set of all possible correlations is a closed interval $[\rho_{\min}, \rho_{\max}]$. This implies that the matching problem could not be feasible for some combinations of F_{X_i}, $i = 1, 2, \ldots, k$, and $\mathbf{\Sigma_Z}(h)$, $h = 0, 1, \ldots, p$.

Second, the $\mathbf{\Sigma_Z}$ matrix obtained by the matching problem must be nonnegative definite. Unfortunately, there is no guarantee that such a property holds, and it is not sufficient that $\mathbf{\Sigma_X}$ is nonnegative definite (Ghosh and Henderson, 2003a,b).

A solutions for the first problem is to use rank correlation, whereas nonnegative matrixes can be handled by computing the nearest correlation matrix (see, for example, Higham, 2002; Lurie and Goldberg, 1998; Qi and Sun, 2006).

3 Data and implementation notes

The data set consists of monthly yields on actively traded non-inflation-indexed issues adjusted to constant maturities. The series are made available

	$\Delta 3m$	$\Delta 1y$	$\Delta 7y$	$\Delta 20y$
Mean	-1.7	-3.1	-3.4	-3.3
St. dev.	0.27	0.32	0.3	0.27
Skewness.	-0.63	-1.09	-0.29	-0.29
Kurtosis	7.42	7.51	4.72	4.76
ADF	-6.8	-5.6	-5.7	-5.6
PP	-211.0	-191.2	-173.0	-172.1
KPSS	0.28	0.30	0.19	0.18

Table 1: Summary statistics of Δi for the sample period February, 1982 to December, 2006. All the series are level stationary and show a high level of the kurtosis.

by the Treasury Department of the U.S. Federal Reserve Bank[3], and the time window goes from February, 1982 to December, 2006. We selected a set of rates to represent the term structure, in particular, we fitted the model on $k = 4$ variates corresponding to 3m, 1y, 7y and 20y.

The most basic requirement of any statistical analysis of market data is the existence of some statistical properties of the data under study which remain stable over time. The invariance of statistical properties of the return process in time corresponds to the stationarity hypothesis. To this purpose, we focused our attention on the series of interest rate changes ($\Delta i_t = i_t - i_{t-1}$). In Table 1, we displayed the summary statistics of Δi_t for the four selected rates. Tests to asses the stationary of the series are also reported. In particular, we computed the Augmented Dickey-Fuller test (ADF) and the Phillips-Perron test (PP) whose null hypotheses are that the series has a unit root. A test for level or trend stationarity is provided by the Kwiatkowski-Phillips-Schmidt-Shin test (KPSS) whose null hypothesis is that the series is level or trend stationary. The ADF and PP test both reject the null hypothesis with a p-value less than 1%, thus assessing stationarity of the levels. The KPSS accept that the series are level stationary with a p-value higher than 10%.

The key issue that inspired the present paper is the high level of the kurtosis found for each rate, especially for the short rates. As explained in Section 1, term structures must be estimated as an whole entity, therefore, it would be meaningless to use different models for the short and long rates. Moreover, when term structures are used for the simulation of medium-long period projects (like, pension funds or debt allocation), not just for one-step-

[3]http://www.federalreserve.gov/Releases/H15/data.htm

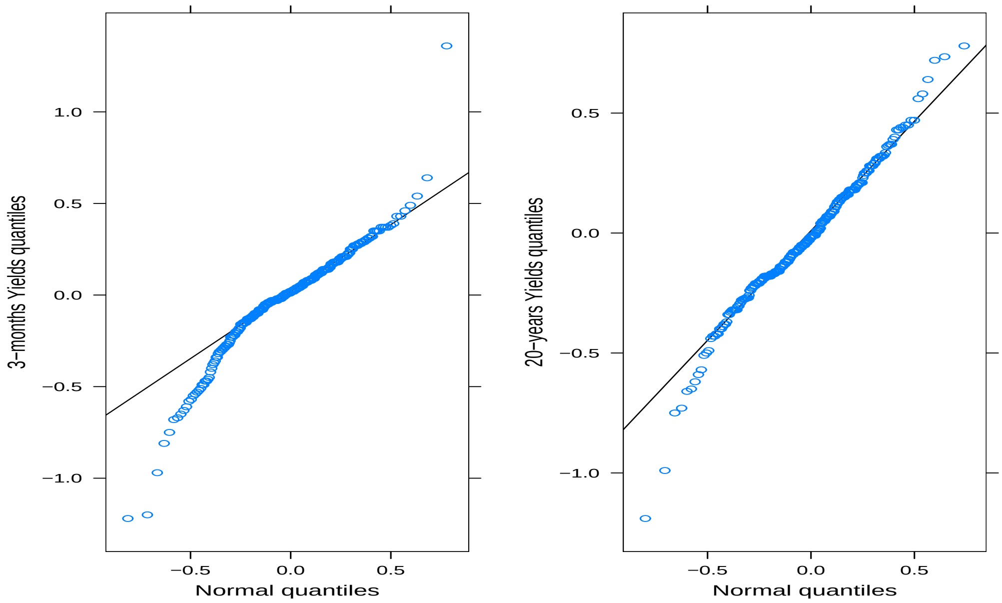


Figure 1: Quantile-quantile plots for $\Delta 3m$ and $\Delta 20y$. If data were gaussian, all the markers would line up on the reference line. The substantial deviation from the straight line reveals a fat tailed distribution.

ahead forecasts (like, short-term economic analysis), fat tails play a relevant role on the decision models. In Figure 1, one can observe the deviation from normality of the tails of the distribution of $\Delta 3m$ and $\Delta 20y$.

3.1 Fitting a VARTA model to interest rate series

The first step in fitting a VARTA model consists in determining the Johnson's translation system parameters for each variable. To this purpose, we use the moment matching approach by Hill et al. (1976)[4]. In Table 2, we show the results obtained for each interest rate series selected. Note that, all the series are found to belong to the S_U family, and that the level of the parameter δ is lower than 2. The latter is a characteristics of leptokurtic distributions[5]. In Figure 2, we compare the empirical distribution function of $\Delta 3m$ (left panel) and $\Delta 20y$ (right panel) with the best fit obtained by the Hill et al.'s moment matching approach.

The core of the VARTA procedure is the solution of the double integral in (9). Following Biller and Nelson (2003), we first operated a variable transfor-

[4] An R implementation of this algorithm can be found in the package `SuppDists`.

[5] As explained in Johnson, the skewness also affects the level of the shape parameter δ.

	$\Delta 3m$	$\Delta 1y$	$\Delta 7y$	$\Delta 20y$
ξ	0.023	0.065	0.0033	-0.0017
λ	0.314	0.469	0.488	0.432
γ	0.313	0.636	0.235	0.218
δ	1.507	1.919	1.929	1.915
Family	S_U	S_U	S_U	S_U

Table 2: Estimates of the Johnson parameters for each series Δi. Distributions with δ lower than 2 are characterized by an high level of the kurtosis.

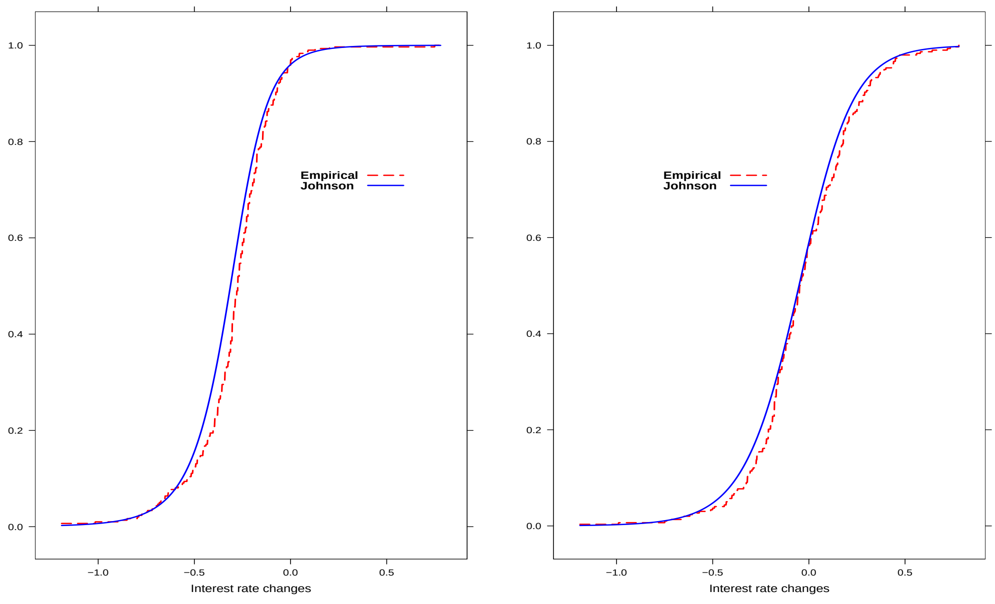


Figure 2: Empirical cdf vs fitted Johnson's translation system cdf. We show, respectively, the $\Delta 3m$ series (left panel) and the $\Delta 20y$ series (right panel).

mation from the infinite region $[-\infty, \infty]^2$ to the finite region $[-1, 1]^2$. This is done by setting $z_i = \tan(\pi z_i^*/2)$ and $dz_i = (\pi/2)[1 + \tan^2(\pi z_i^*/2)]dz_i^*$, where $-1 < z_i^* < 1$ and $i = 1, 2$. We used an adaptive integration routine and tried different cubature rules. That was done by interfacing our C++ code with the HIntlib library by Schürer (2006).

At an outer level, we solved the matching problem $c_{ijh}[\rho_Z(i, j, h)] \approx \rho_X(i, j, h)$ through a one-dimensional root finding algorithm. Since we know the size of the bounded region containing the root, we adopted the root bracketing scheme given in the Brent's algorithm and implemented in the GSL library (GSL Team, 2007).

4 Model testing and validation

We recall here that the main objective of this paper is to provide scenarios of the interest rate curve for medium–long horizon. We are not interested in one-step ahead forecasts.

To this purpose, we fitted a VARTA(2) to the set of rates selected. The choice of lag $p = 2$ is motivated by the search for a parsimonious model, and from the observation that after two lags the autocorrelations of the interest rate changes are negligible. Moreover, when fitting a VAR model to the set of available data, a VAR(2) is suggested by the various information criteria[6].

We generated $S = 1000$ scenarios of length $T = 360$ months. The scenario generation is accomplished by recursively applying a series of Gaussian white noise vectors, $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_T$, to the estimated process $\mathbf{z}_t = \alpha_1 \mathbf{z}_{t-1} + \alpha_2 \mathbf{z}_{t-2} + \ldots + \alpha_p \mathbf{z}_{t-p} + \mathbf{u}_t$, for $t = 1, 2, \ldots, T$. To obtain the starting values $\mathbf{z}_{-p+1}, \mathbf{z}_{-p+2}, \ldots, \mathbf{z}_0$ and the series of independent Gaussian white noise vectors, $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_T$, we sampled, respectively, from the kp–dimensional multivariate normal distribution, with variance-covariance matrix $\mathbf{\Sigma}_\mathbf{Z}$, and from the k-dimensional multivariate normal distribution with variance-covariance matrix $\mathbf{\Sigma}_u$[7]. Finally, the realization $\mathbf{x}_t$ of the process $\mathbf{X}_t$ is obtained by the probability-integral transformation $\mathbf{x}_t = F_X^{-1}[\Phi(\mathbf{z}_t)]$.

The performance of the two models is validated by setting the quantiles of the empirical distribution against the quantiles obtained by averaging over the set of scenario S. In particular, we denoted by q_α^l the quantile, under scenario l, at which α% of the data fall below. The average quantile $\bar{q}_\alpha$ is

[6]We fitted the VAR model using the `vars` package available in R.

[7]The matrixes $\mathbf{\Sigma}_\mathbf{Z}$ and $\mathbf{\Sigma}_u$ are definite positive, therefore, a standard algorithm based on the Cholesky factorization of the two matrixes can be used to draw samples from the two multivariate normal distributions.

then given by,

$$\bar{q}_\alpha = \frac{1}{S} \sum_l q_\alpha^l. \tag{11}$$

We also added a band to delimit the area under which 90% of the scenarios fall. To this purpose, we ordered the quantiles q_α^l, for all $l = 1, 2, \ldots, S$, from the highest to the lowest, and denoted by q_α^{UP} the quantile under which 95% of the quantiles fall, and by q_α^{LO} the quantile under which 5% of the quantiles fall.

In Figure 3 and 4, we show the quantile–quantile plots for, respectively, the $\Delta 3m$ and $\Delta 20y$ variates. In the upper panel, we display the results obtained when scenarios are generated by a VAR(2) process, in the bottom panel, when scenarios are generated by an VARTA(2) model. We highlight here the most significative findings:

1. The VARTA process is able to reproduce more faithfully the empirical distribution of the data. In particular, we observe that the average quantile curve line up on the straight line. This behavior is also confirmed for the majority of the scenarios, since the band, delimited by the two dashed lines, closely aligns with the reference line.

2. The VAR process plainly deviates from the theoretical quantile line. This bias also applies to the 90% band. We could not expect a different behavior, since the VAR processes implicitly assume that the marginal distributions of the variables are Gaussian.

3. The VAR process slightly improves for $\Delta 20y$ (see Figure 4). This occurs because the $\Delta 20y$ variate shows less deviation from normality (see Figure 1, right panel). The VARTA process, consistently, delivers scenarios that match the empirical distribution. This is a valuable feature for the VARTA process. In fact, due to the nature of the interest rate curve, we cannot estimate separately variables whose marginals are approximately Gaussian, and those whom deviate from normality. Broadly speaking, the VARTA process is flexible enough to allow for the fitting of distributions with different degree of "Gaussianity". This is also possible when using copula based models, but, as explained in Section 1, applications to this problem are somewhat troublesome.

4. A more accurate estimate of the VAR model (i.e., structural models, cointegration analysis, etc.) will not lead to better results, since for medium-long period simulation what matters are the higher moments of the unconditional marginal distributions.

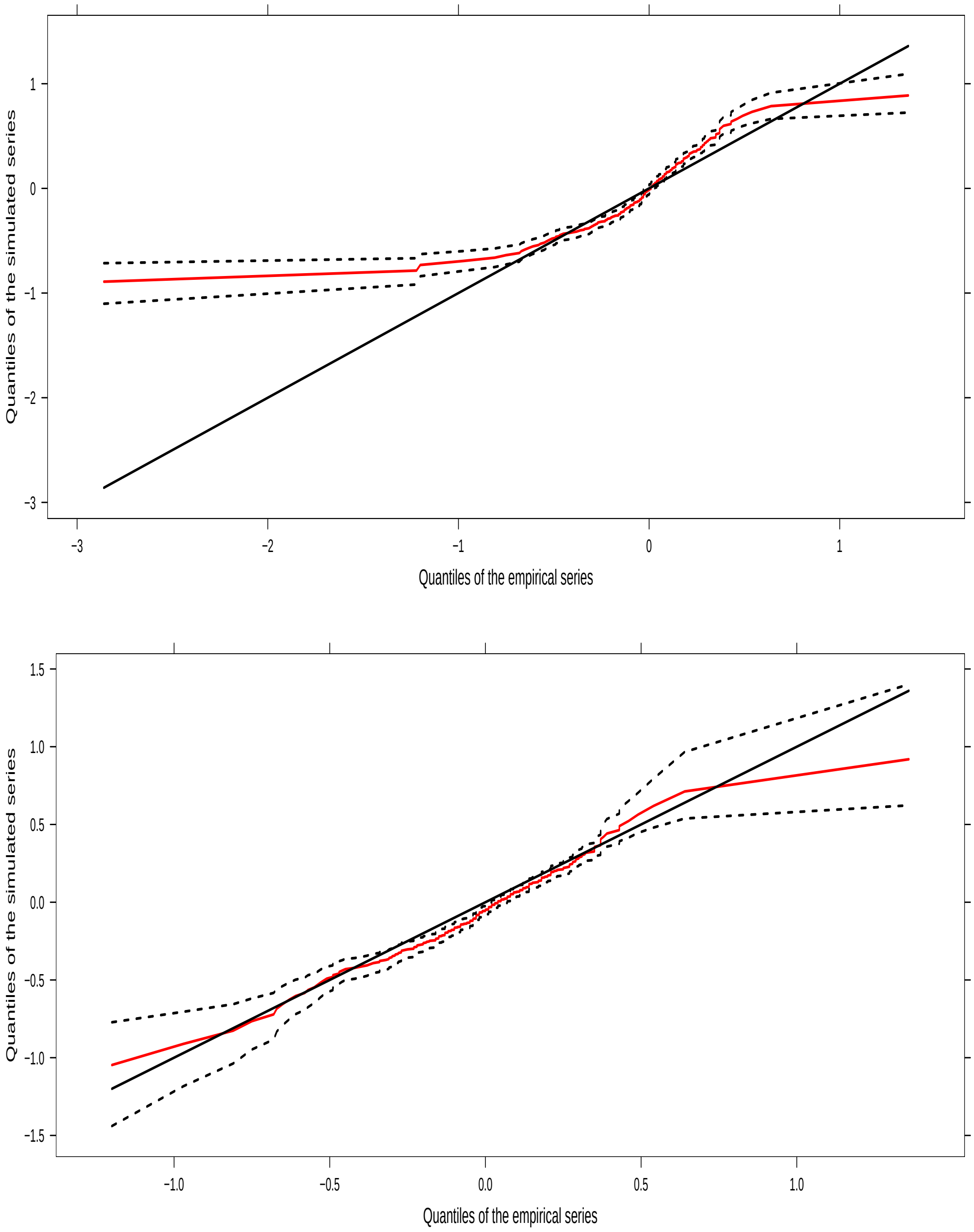
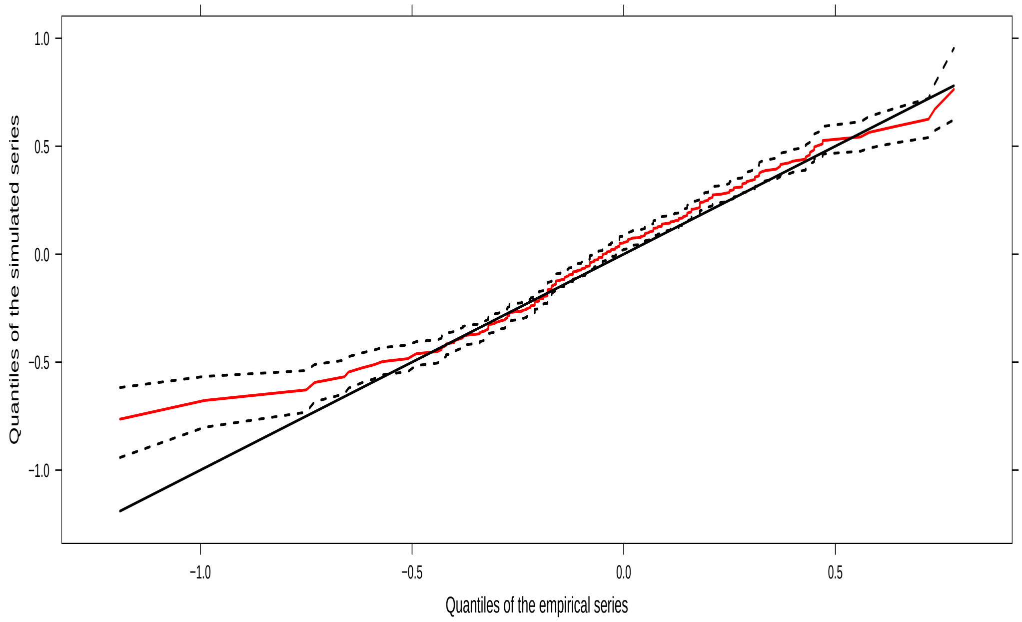


Figure 3: Quantile–quantile plot for the $\Delta 3m$ variate. The upper panel displays the results obtained by the VAR(2) process. The bottom panel displays the results obtained by the VARTA(2). The reference line is represented by the straight line. For each level of α, the dashed piecewise lines join the points q_α^{UP} and q_α^{LO}, the solid piecewise line joins the average quantiles $\bar{q}_\alpha$.



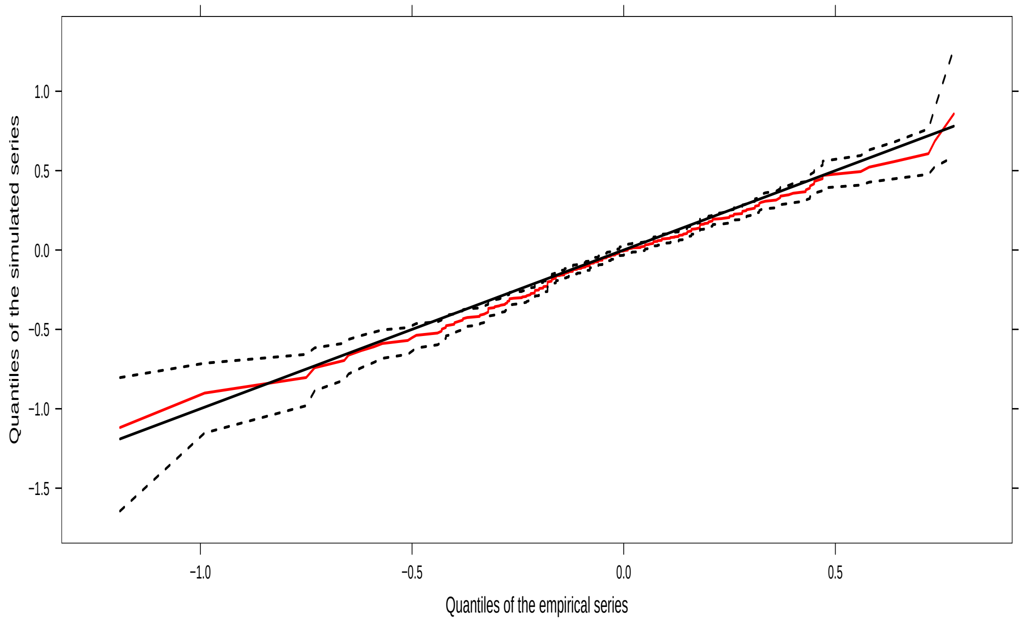


Figure 4: Quantile–quantile plot for the $\Delta 20y$ variate. The upper panel displays the results obtained by the VAR(2) process. The bottom panel displays the results obtained by the VARTA(2). The reference line is represented by the straight line. For each level of α, the dashed piecewise lines join the points q_α^{UP} and q_α^{LO}, the solid piecewise line joins the average quantiles $\bar{q}_\alpha$.

5 Conclusions

Empirical analyses have shown that interest rate changes have distributions that substantially deviate from the normal assumption. The need of simulation models for medium-long time horizons have led us to build a vector auto-regression model with arbitrary marginals. Such simulation models are fundamental for risk management and optimal debt planning. Our model provides more reliable scenarios in term of fit to the empirical distribution. I also improves upon copula–based models, since it provides a framework to reproduce the dynamics of the random–vector process underneath the term structure of interest rates. A possible extension of the model consists in describing the interest rate curve more parsimoniously. To this purpose, we are adapting the VARTA process to a representation of the term structure due to Diebold and Li (2006).

References

A. Ang and M. Piazzesi. A no-arbitrage vector autoregression of term structure dynamics with macroeconomic and latent variables. *Journal of Monetary Economics*, 50(4):745–287, 2003.

B. Biller. Copula-based multivariate time-series models for stochastic simulations. Working paper, Tepper School of Business, Carnegie Mellon University, 2006.

B. Biller and B.L. Nelson. Modeling and generating multivariate time-series input processes using a vector autoregressive technique. *ACM Transactions on Modeling and Computer Simulation*, 13(3):211–237, July 2003.

D.R. Cariño, D.H. Myers, and W.T. Ziemba. Concepts, technical issues, and uses of the russel–yasuda kasai financial planning model. *Operations Research*, 46(4):450–462, 1998.

X. Chen and Y. Fan. Estimation of copula-based semiparametric time series models. *Journal of Econometrics*, 130:307–335, 2006.

U. Cherubini, E. Luciano, and W. Vecchiato. *Copula methods in finance.* Wiley, 2004.

A. Consiglio and D. De Giovanni. Evaluation of insurance products with guarantee in incomplete markets. *Insurance: Mathematics & Economics*, 42(1):332–342, 2008.

A. Consiglio, F. Cocco, and S.A. Zenios. The Prometeia model for managing insurance policies with guarantees. In W.T. Ziemba and S.A. Zenios, editors, *Handbook of Asset and Liability Management 2. Applications and Case Studies*, Handbooks of Finance, chapter 15, pages 664–705. Elsevier, The Netherlands, 2007.

A. Consiglio, F. Cocco, and S.A. Zenios. Asset and liability modelling for participating policies with guarantees. *European Journal of Operational Research*, 186(1):380–404, 2008.

F.X. Diebold and C. Li. Forecasting the term structure of government bond yields. *Journal of Forecasting*, 130:337–364, 2006.

P. Embrechts, A. McNeil, and D. Straumann. Correlation and dependence in risk management: properties and pitfalls. In M. Dempster, editor, *Risk management: value at risk and beyond*. Cambridge University Press, Cambridge, 2002.

J. Fermanian and O. Scaillet. Nonparametric estimation of copulas for time series. *The Journal Of Risk*, 5(4):25–54, 2003.

S. Ghosh and S.G. Henderson. Behaviour of norta method for correlated random vector generation as dimension increases. *ACM Transactions on Modeling and Computer Simulation*, 13:276–294, 2003a.

S. Ghosh and S.G. Henderson. Behaviour of norta method for correlated random vector generation as dimension increases. *ACM Transactions on Modeling and Computer Simulation*, 13:276–294, 2003b.

GSL Team. *GNU Scientific Library*, 2007. Available at http://www.gnu.org/software/gsl.

N.J. Higham. Computing the nearest correlation matrix—a problem from finance. *IMA Journal of Numerical Analysis*, 22(3):329–343, 2002.

I.D. Hill, R. Hill, and R.L. Holder. Fitting Johnson curves by moments. *Applied Statistics*, 1976.

K. Høyland and S.W. Wallace. Generation scenario trees for multistage decision problems. *Management Science*, 47(2):295–307, 2001.

N.L. Johnson. Systems of frequency curves generated by methods of translation. *Biometrika*, 36:149–176, 1949.

A.J King. Duality and martingales: A stochastic programming perspective on contingent claims. *Mathematical Programming, Series B*, 91:543–562, 2002.

P. Klaassen. Comment on “Generating scenario trees for multistage decision problems”. *Management Science*, 48(11):1512–1516, 2002.

P. M. Lurie and M. S. Goldberg. An approximate method for sampling correlated random variables from partially-specified distributions. *Management Science*, 44(2):203–218, 1998.

H. Lütkepohl. *New introduction to multiple time series analysis*. Springer-Verlag, Heidelberg, 2005.

J.M. Mulvey and A.E. Thorlacius. The Towers Perrin global capital market scenario generation system. In W. Ziemba and J.M. Mulvey, editors, *Worldwide asset and liability modeling*, pages 286–312. Cambridge University Press, Cambridge, UK, 1998.

S. Ni and D. Sun. Bayesian estimates for vector autoregressive models. *Journal of Business & Economic Statistics*, 23(1):105–117, 2005.

A. Patton. Modelling asymmetric exchange rate dependence. *International Economic Review*, 47(2):527–556, 2006.

H. Qi and D. Sun. A quadratically convergent newton method for computing the nearest correlation matrix. *Journal on Matrix Analysis and Applications*, 28(2):360–385, 2006.

R. Schürer. *High-dimensional integration library*, 2006. Available at http://www.cosy.sbg.ac.at/rschuer/hintlib.