Scenario-Generation Methods for an Optimal Public Debt Strategy

Massimo Bernaschi, Maya Briani, Marco Papi and Davide Vergni

Istituto per le Applicazioni del Calcolo “Mauro Picone” – CNR
Viale del Policlinico, 137, 00161 Roma, Italy

We describe the methods we employ for the generation of possible scenarios of the term structure evolution. The problem is originated by the request of the Italian Ministry of Economy and Finance of finding an optimal strategy for the issuance of Public Debt securities. The basic idea is to split the evolution of each rate in two parts. The first component is driven by the evolution of the official rate (the European Central Bank official rate in the present case). The second component of each rate is represented by the fluctuations having null correlation with the ECB rate.

Submitted to: Quantitative Finance
1. Introduction

The management of the public debt is of paramount importance for any country. Mathematically speaking, this is a stochastic optimal control problem with a number of constraints imposed by national and supranational regulations and by market practices.

The Public Debt Management Division of the Italian Ministry of Economy decided to establish a partnership with the Institute for Applied Computing in order to face this problem from a quantitative viewpoint. The goal is to determine the composition of the portfolio issued every month which minimizes a predefined objective function that can be described as an optimal combination between cost and risk of the public debt service.

Since the main stochastic component of the problem is represented by the evolution of interest rates, a key point is to determine how various issuance strategies perform under different scenarios of interest rates evolution. In other words, an optimal strategy for the management of the public debt requires a suitable modelling of the stochastic nature of the term structure of interest rates.

Note that hereafter we are not going to present how to forecast the actual future term structure but how to generate a set of realistic possible scenarios. For our purposes, the scenarios should cover a wide range of possible outcomes of future term-structures in order to provide a reliable estimate of the possible distribution of future debt charges. This distribution is useful in a risk-management setting to estimate, for instance, a sort of Value at Risk (VaR) of the selected issuance policy.

The paper is organized as follows. Section 2 describes the problem. Section 3 presents a first, simple, solution based on the well-known Vasicek model. Section 4 describes other, more sophisticated, models we employ for generating future interest rates scenarios. Section 5 introduces some criteria to validate a scenario. Section 6 concludes with the future perspectives of this work.

2. Problem description and basic guidelines

It is widely accepted that stock prices, exchange rates and most other interesting observables in Finance and Economics cannot be forecast with certainty.

Interest rates are an even more complex issue because it is necessary to consider the term structure, that is a multi-value observable. Despite this difficulty, there are a number of studies that, from both the academic and the practitioner viewpoint, deal with interest rate modeling (for a comprehensive survey of interest rate modelling see).

The most common application of existing term structure models is the evaluation of interest-rate-contingent claims. However our purpose is pretty different since we aim to find an optimal strategy for the issuance of the Italian public debt securities.

In a very simplified form the problem can be described as follows. The Italian Treasury Department issues (about) ten different types of securities including one with floating rate. Securities differ in the maturity (or expiration date) and in the rules
for the payment of interests. Short term securities (those having maturity up to two years) do not have coupons. Medium and long term securities (up to thirty years) pay cash dividends, every 6 months, by means of coupons. The problem is to find a strategy for the selection of public debt securities that minimizes the expenditure for interest payment (according to the ESA95 criteria) and satisfies, at the same time, the constraints on debt management. The cost of future interest payments depends on the future value of the term structure (roughly speaking when a security expires or a coupon is paid, there is the need to issue a new security whose cost depends on the term structure at expiration time). That is the reason why we need to generate scenarios of future interest rates. In we show that for a set of term structure evolutions and Primary Budget Surplus (PBS) realizations, such optimization problem can be formulated as a linear programming problem with linear constraints.

A first issue to consider in the model selection process is the time frame. For the purposes of the Ministry, a reasonable planning period is 3-5 years. Within such a long period, the term structure may change significantly as shown in figure 1.

In figure 2 we report the monthly evolution of the swap interest rates for the following maturities: 3, 6, 12, 24, 36, 60, 120, 180, 360 months along with the value of the European Central Bank (ECB) official rate and a simple interpolation of such rate in the same period (January 1999-December 2002). The interpolation is obtained by joining two successive jumps of the ECB official rate by means of a line (i.e., it is not a linear interpolation based on the values of the ECB rate in the whole time period). Such interpolation mimics in a pretty good way the evolution of the interest rates, especially for short maturities, and we use it as a simple approximation of the ECB trend.

This is our basic dataset for the analysis and generation of new scenarios of the term structure evolution.
3. A simple model to generate interest rates scenarios

A very simple way to simulate the time evolution of an interest rate relies on the Vasicek model\(^\text{18}\) whose main characteristic is the presence of a mean-reverting term and a constant volatility term. From a mathematical viewpoint it is a stationary Gaussian Markovian model:

\[
dr_t = k(\hat{r} - r_t)dt + \sigma dB_t, \tag{3.1}
\]

where \(\{B_t : t \geq 0\}\) is a Brownian motion with unit variance and \(k, \hat{r}, \sigma\) are constant parameters.

Usually, model parameters must be estimated looking at market data. There are many methods to estimate the value of the parameters and some of these are very sophisticated. Just to mention two of them, both the Method of Moments and the Maximum Likelihood\(^\text{12}\) are widely used to this purpose.

However, for our purpose, that is the generation of long time series of realistic interest rates, we can resort to a more simple technique based on the assumption that the mean reverting level \(\hat{r}\) corresponds to the current interest rate value \(r_0\). Starting from this assumption, it is easy to compute the statistical properties of the \(r_t\) given by Eq. (3.1):

\[
\mathbb{E}(r_t) = \hat{r}, \quad \text{Cov}(r_t, r_s) = \frac{\sigma^2}{2k} e^{-k|t-s|}. \tag{3.2}
\]

Therefore we can evaluate \(\sigma\) and \(k\) by means of the second relation and a simple linear fit, in semi-logarithmic scale, of the historical covariance.

It is possible to repeat such calibration procedure for each maturity (from 3 to 360 months) obtaining a set of parameters \((\hat{r}^h, k^h, \sigma^h)\), where \(h = 1, \ldots, M\) is an index that identifies the \(M\) different maturities \(m^h\). Therefore, in principle, starting from the
stochastic equation of evolution with suitable parameters it is possible to generate the whole term structure of the interest rates.

However, to obtain a realistic term structure we need to consider the correlation among different securities. To this purpose, we first generate independent rates, then we correlate them by means of the addition of an exponentially weighted moving average term. This solution guarantees that short term rates are less correlated whereas long term rates are strongly correlated, in accordance to the observed behaviour of the interest rates.

The whole procedure can be summarized as follows: from the historical time series of the interest rates, we determine the parameters \((\hat{r}^h, \alpha^h, \sigma^h)\). Starting from the last value of each rate, \(r^h_{t-\Delta t}\), and using a discretization of Eq. (3.1), we generate a set of new rates \(r^h_t\). Then, we construct a complete yield curve defining \(\tilde{y}(t, T) = r^h_t\) if \(T = m^h\). For the lacking maturities we resort to a linear interpolation. The exponentially weighted moving average is applied to the maturity variable, \(T\), of the piecewise linear function \(\tilde{y}(t, T)\):

\[
y(t, T) = \frac{1}{2e^{cT} + 1} \sum_{S=T-e^{cT}}^{T+e^{cT}} \tilde{y}(t, S).
\] (3.3)

The moving average window has a width that increases exponentially with respect to maturity. Obviously, since the domain of \(y(t, T)\) is bounded: \(\min_h m^h \leq T \leq \max_h m^h\), the sum has to be limited to this region.

Since the Vasicek model does not guarantee that interest rates remain positive, it is necessary to check the evolution of the different rates and set a barrier to prevent any rate from becoming negative.

Figure 4 shows the results of a simulation, based on the Vasicek model, of the term structure evolution for a planning period of 48 months.

4. More advanced models for the generation of interest rates scenarios
Even if the approach based on the Vasicek model is valuable for its simplicity, it cannot be fully satisfactory from a theoretical viewpoint. In this section we present other, more sophisticated, ways to simulate the evolution of interest rates.

If we look at Figure 2, it is apparent that any rate (regardless of its maturity) has a strong correlation with the ECB rate. This is not surprising and we use this observation to develop a new approach to the generation of future term-structure scenarios that can be described as a multi-step process:

(i) generation of a scenario for the future ECB official rate;

(ii) generation of the fluctuations of each rate with respect to the ECB official rate;

(iii) validation of the resulting scenario to determine whether it is acceptable or no.

The basic idea is that the ECB official rate represents a “reference rate” and all the other rates are generated by a composition of this “reference rate” plus a characteristic, maturity dependent, fluctuation.

In other words, each rate is determined by the sum of two components: the first component is proportional to the (linearly interpolated) ECB official rate; the second component is the orthogonal fluctuation of that rate with respect to the ECB. In mathematical terms, each rate $r^h_t$ is decomposed as

$$r^h_t = \alpha_h r^{\text{ecb}}_t + p^{h,\perp}_t$$

where $r^{\text{ecb}}_t$ is the linear interpolation of the ECB official rate and $\alpha_h$ is the correlation coefficient between $r^h$ and $r^{\text{ecb}}$, given by

$$\alpha_h = \frac{\mathbb{E}((r^{\text{ecb}} - \mathbb{E}(r^{\text{ecb}})) \cdot (r^h - \mathbb{E}(r^h)))}{\sqrt{\text{Var}(r^{\text{ecb}}) \cdot \text{Var}(r^h)}}.$$  \hfill (4.5)

By construction, the time series of $p^{h,\perp}_t$ has null correlation with $r^{\text{ecb}}$. The factors $\alpha^h$ are different for each maturity and a larger value of $\alpha^h$ means a larger correlation with...
the ECB official rate. Table 1 reports the values of the factors $\alpha^h$ for each maturity we consider. As expected, longer maturities are less correlated with the ECB official rate.

<table>
<thead>
<tr>
<th>Maturity</th>
<th>$\alpha^\lambda$</th>
<th>Maturity</th>
<th>$\alpha^\lambda$</th>
<th>Maturity</th>
<th>$\alpha^\lambda$</th>
</tr>
</thead>
<tbody>
<tr>
<td>3 months</td>
<td>0.9833</td>
<td>2 years</td>
<td>0.7693</td>
<td>10 years</td>
<td>0.3949</td>
</tr>
<tr>
<td>6 months</td>
<td>0.9514</td>
<td>3 years</td>
<td>0.6695</td>
<td>15 years</td>
<td>0.3545</td>
</tr>
<tr>
<td>12 months</td>
<td>0.9094</td>
<td>5 years</td>
<td>0.5425</td>
<td>30 years</td>
<td>0.3237</td>
</tr>
</tbody>
</table>

Table 1. The values of the factors that measure the correlation with the ECB rate

Figure 5 shows the result of the decomposition applied to the data of figure 2 (only the component having null correlation with the ECB rate is shown in the figure). In Section 4.1 and 4.2 we describe two approaches for the simulation of the future fluctuations, based on the analysis of these data.

However, before that, we describe how we generate scenarios of the future ECB official rate.

There is a long tradition of studies that try to explain the links between prices, interest rates, monetary policy, output and inflation. Babbs and Webber $^{19}$ were among the first authors to propose yield-curve models able to capture some aspects of monetary policies. In $^{16}$ the author introduces a class of linear-quadratic jump-diffusion processes in order to develop an arbitrage-free time-series model of yields in continuous time that incorporates the central bank policy and estimates the model with U.S. interest rates and the Federal Reserve’s target rate. A general framework for these models has been recently presented in $^8$, where the authors introduce factors which influence the marginal productivity of capital, and thus the interest rates, in the economy. Their technique
is general, since it applies to any multi-factor, exponential-affine term structure model with multiple Wiener and jump processes.

From an empirical point of view, there is some evidence that important central banks, like the U.S. Federal Reserve, conduct a monetary policy \( i.e., \) set the official rate that is well described by the so called Taylor’s rule \(^{17}\). Basically the rule states that the “real” short-term interest rate (that is, the interest rate adjusted for inflation) should be determined according to three factors: (1) where actual inflation is relative to the targeted level that the central bank wishes to achieve, (2) how far economic activity is above or below its “full employment” level, and (3) what the level of the short-term interest rate is that would be consistent with full employment. Although the Taylor’s rule appears more robust than more-complex rules with many other factors, it requires the knowledge of inflation and real Gross Domestic Product (GDP). The simulation of the future inflation and GDP is far from being easy so, in some sense, the application of the Taylor’s rule changes but does not solve the problem of generating meaningful scenarios for the evolution of the ECB official rate.

Currently, instead of resorting to an existing macroeconomic model, we simulate the future ECB official rate as a stochastic jump process.

The starting point is the observation of some features that are readily apparent by looking at the evolution of the ECB official rate since January 1999 (see Figure 2) when the ECB official rate replaced the national official rate for the Euro “zone” countries:

- there are, on average, three interventions of the ECB per year;
- the ECB rate jumps (with, approximately, the same probability) by either 25 or 50 basis-point;
- there is a strong “persistence” in the direction of the jumps. That is, there is a high probability that a rise is followed by another rise and that a cut is followed by another cut.

We model the ECB interventions as a combination of two processes: \( i) \) a Poisson process that describes when the ECB changes the official rate and \( ii) \) a Markov process that describes the sign of the change. Then we resort to an exponential distribution to simulate the waiting times between two changes of the ECB official rate, and to a Markov Chain for the simulation of the direction (positive or negative) of the change.

The parameter of the exponential distribution can be easily estimated from available data (that is from the waiting times, in months, between the jumps occurred in the past) by means of the Maximum Likelihood Estimation (MLE) method. It turns out to be approximately equal to 0.25.

Since there are two possible states (positive and negative) in the Markov chain, the corresponding transition matrix has four entries (positive-positive, positive-negative, negative-negative, negative-positive). We estimate the values of each entry by looking at the historical data and in particular at the probability that a change is in the same, or in the opposite, direction of the previous one. It is interesting to note that the
probability of having a change in the same direction of the previous one is pretty high, approximately 85%.

Finally, the width of the jump is selected between two possible values (25 or 50 basis point) with the same probability.

In conclusion, the ECB official rate at time $t$ is defined as:

$$\text{ECB}_t = \text{ECB}_0 + \sum_{s=1}^{N_t} a_s C_{s-1,s}$$ (4.6)

where $N_t$ is a realization at time $t$ of the Poisson process that represents the total number of jumps up to time $t$; $a_s$ is the random width of the jump $s$ and $C_{s-1,s}$ represents the sign of the jump $s$ given the sign of the jump $s - 1$.

The algorithm to simulate the future ECB rate can be described as follows:

Future ECB rate = current ECB rate;
Set the direction (up/down) of the next jump;
Foreach month in the planning period {
    Determine, according to the exponential distribution, if there is a jump;
    if (jump) {
        Determine if the jump is in the same direction of the previous one;
        Determine if the jump is of 25 or 50 basis point;
        Future ECB rate = Future ECB rate + jump; (note that jump can be < 0)
    }
}

Figure 6 shows a few simulations of the future ECB official rate produced by this algorithm.

In the rest of this section, we describe the approaches we use for the generation of the fluctuations of each rate with respect to the simulated ECB official rate.
4.1. Principal Component Analysis

Principal Component Analysis (PCA) is a well-known technique in time series analysis and it has been in use for a number of years in the study of fixed income markets. In general PCA assumes that the underlying process is a diffusion. The data we employ do not have the jump components produced by the ECB interventions thanks to the decomposition procedure (4.4) we described previously. From this viewpoint the data appear suitable to a PCA. The procedure we follow is pretty much the standard one:

- For each de-trended rate $p^{h,\perp}_t$ we calculate the differences $\delta^h_t = p^{h,\perp}_t - p^{h,\perp}_{t-1}$ ($t$ is the time index).

- We apply the PCA to the $m$ time series of the differences that we indicate with $\delta$.

This means:
- to calculate the empirical covariance matrix ($\Lambda_{ij}^{\text{hist}}$) of the $\delta$;
- to find a diagonal matrix of eigenvalues $\Lambda^d$ and an orthogonal matrix of eigenvectors $E = [e_1, ..., e_m]$ such that the covariance matrix is factored as: $\Lambda^d = E\Lambda^{\text{hist}}E^T$.

The eigenvectors with the largest eigenvalues correspond to the component that have the strongest correlation in the dataset.

For the dataset of the monthly yield changes from January 1999 to December 2002, the first three components represent 98.3% of the total variance of the data. The projections of the original data on the first three principal components do not show autocorrelation phenomena (the covariance matrix is, as expected in the PCA, very close to the $\Lambda^d$ matrix). Usually the first three components are interpreted respectively as i) a level shift; ii) a slope change; iii) a curvature change. Since the PCA is applied, in our case, to de-trended data (with respect to the ECB official rate) the meaning of the components could be different. However the plot of the first three components shown in figure 7 does not seem too different from similar studies that consider directly the yield changes.

To create a new scenario for the fluctuations it is necessary to take a linear combination of the principal components. If $N$ is the number of principal components ($N = 3$ in the present case), an easy way is to compute $\bar{\delta} = F\bar{\mu}$ where $F = [e_1, ..., e_N]$ is a $m \times N$ matrix composed of the eigenvectors corresponding to the first $N$ eigenvalues and $\bar{\mu}$ a vector $\mu$ with $N$ elements. In principle, the value of each element of this vector could be selected, at will. However, two are the most common choices: the first one is to assign a value taking into account the meaning of the corresponding principal component. For instance, if the purpose is to test the effect of a level shift with no slope or curvature change, it is possible to assign a value only to the first element leaving the other two elements equal to zero. The other choice is to compute $\bar{\delta} = F\sqrt{\Lambda^d}\bar{Z}$ where $\bar{Z}$ is a vector of $N$ independent, normally distributed, variables. Therefore each element of the vector $\bar{\mu}$ is drawn from a normal distribution with variance equal to the eigenvalue corresponding to the principal component. In the present work, we followed the second
approach since we did not want to make *a-priori* assumptions. Hence, the evolution of \( p_{t}^{h,\perp} \) is described by the following equation:

\[
p_{t}^{h,\perp} = p_{t-1}^{h,\perp} + \sum_{j=1}^{N} F_{hj} \sqrt{\lambda_j} Z_j ,
\]

(4.7)

where \( \lambda_j \) is the \( j \)-th eigenvalue. Note that a linear combination of the principal components provides a vector of fluctuations for one time period only. Actually, since the planning period is three-five years and the time step is one month, we need, for a single simulation, a minimum of 36 up to a maximum of 60 vectors of fluctuations. Obviously, it is possible to repeat the procedure for the generation of the fluctuations, but there is no guarantee that the covariance matrix of the simulated fluctuations will be close to the covariance matrix of the historical fluctuations. Although a different behaviour does not necessarily imply that the simulated fluctuations are meaningless, it is clear that we must keep this difference under control. In Section 5 we describe how we faced this issue.

Figure 8 shows the results of a simulation, based on the PCA, of the term structure evolution for a planning period of 48 months.

### 4.2. Cox-Ingersoll-Ross model

#### 4.2.1. The Basic Stochastic Process

The widely used Cox-Ingersoll-Ross (CIR) model \(^6\) proposes the following equation for the short term interest rate:

\[
dr_t = k(\mu - r_t)dt + \sigma \sqrt{r_t}dB_t
\]

(4.8)

where \( k, \mu \) and \( \sigma \) are positive constants. This model is similar to the Vasicek model
but it has better properties. Foremost, the interest rates are continuous and remain positive if the condition $2k\mu > \sigma^2$ is fulfilled. Secondly, the volatility is proportional to the interest rate. As the Vasicek model, it is a mean reverting model (i.e., the short rate is elastically pulled to the constant long-term value $\mu$). The CIR model belongs to the class of exponential-affine models. In these models the yield to maturity on all bonds can be written as a linear function of the state variables (the affine models have been thoroughly analyzed in $^{10}$). A review of some of the estimation methods for these models is reported in the Appendix.

Unfortunately, if we used the simple one factor model (4.8), we would neglect a fundamental point for the generation of scenarios of the future term structure, that is the correlation among interest rates corresponding to different maturities. In order to capture this correlation element, in the following, we propose a simple multi-dimensional extension of model (4.8) for the generation of the orthogonal fluctuations with respect to the ECB official rate. We consider the following model:

$$dp_{h,t}^{h,\perp} = k_h(\mu_h - p_{h,t}^{h,\perp})dt + \sqrt{p_{h,t}^{h,\perp}}\sum_{j=1}^{M}\sigma_{hj}dB_t^j \quad \text{for} \quad h = 1, \ldots, M$$  \hspace{1cm} (4.9)

on the interval $[0,T]$, $T$ being the time horizon for the generation of scenarios. Here $B_t = (B_t^1, \ldots, B_t^M)$ is an $M$-dimensional Brownian motion representing $M$ sources of randomness in the market, and $k_h > 0, \mu_h > 0, \sigma_{hj}$ are constants satisfying $\text{det}(\sigma\sigma^\top) \neq 0$ with

$$k_h\mu_h > \frac{1}{2}(\sigma\sigma^\top)_{hh}$$ \hspace{1cm} (4.10)

for any $h = 1, \ldots, M$. It is easy to show that the inequality (4.10) is a sufficient condition for the existence of a unique solution with positive trajectories to the stochastic differential equation (4.9) starting at $p_0 \in (0, \infty)^M$, see in particular $^{9}$. Unfortunately, model (4.9) does not belong to the usual affine yield class $^{10}$, since the var-cov matrix

---

**Figure 8.** Simulation of the term structure evolution based on the PCA
\( \Sigma(p) \Sigma(p)^\top \), with \( \Sigma_{i,j}(p) = \sqrt{p_i p_j} \sigma_{i,j} \), \( i, j = 1, \ldots, M \) is not an affine function of \( p \) unless \( \sigma \) is diagonal, which is the case of \( M \) uncorrelated CIR univariate models mentioned previously. Moreover, there is not a linear change of variable that allows to reduce the model to an affine one. However, model (4.9) preserves some features of the class studied in \(^9\) and \(^{10}\), such as analytical tractability and convenient econometric implementation, while the volatility of each component is proportional only to the component itself. In the sequel we focus our attention on the problem of the estimation of (4.9).

4.2.2. The Estimation of Model (4.9)

There is a growing literature on estimation methods for term structure models, see in particular \(^2\). A vast literature is specifically devoted to the estimation of affine models, see \(^3\) and references therein. Here we shall discuss an ad-hoc method to estimate model (4.9), which is based on the maximization of maximum likelihood function for the Euler discretization. The discrete version of model (4.9) is

\[
p^h_{ti+1} = p^h_{ti} + k_h(\mu_h - p^h_{ti}) \Delta + \sqrt{p^h_{ti} \Delta} \sum_{j=1}^{M} \sigma_{hj} Z^j_i, \quad i = 1, \ldots, n - 1, \tag{4.11}
\]

for \( h = 1, \ldots, M \), where \( t_i = i \Delta \) and \( \Delta = T/n \). Here \( Z_i = (Z^1_i, \ldots, Z^M_i) \), \( i = 1, \ldots, n - 1 \), are independent multivariate normal random variables with zero mean and covariance matrix \( I_M \), \( I_M \) being the identity matrix of order \( M \). We observe that the distributional properties of the process (4.11) depend only on \( k, \mu \) and \( \sigma \sigma^\top \). Therefore, in order to generate scenarios from model (4.11), it suffices to know an estimate of this matrix. Hence we shall introduce the matrix \( \Gamma^{-1} = \sigma \sigma^\top \). The conditional log-likelihood function associated, for a sequence of observations \( \hat{p}^h_{ti} > 0 \), is:

\[
\ell_n = \zeta + \frac{n - 1}{2} \log(\det \Gamma) - \frac{1}{2} \sum_{i=1}^{n-1} e_i^\top \Gamma e_i \tag{4.12}
\]

where \( \zeta = -[(n - 1)M \log(2\pi)]/2 \) and

\[
e_i^h := e_i^h + \eta_i^h k_h \mu_h + \psi_i^h k_h, \tag{4.13}
\]

\( \epsilon_i^h, \eta_i^h \) and \( \psi_i^h \) being constants independent of the parameters of the model, given by:

\[
\begin{align*}
\epsilon_i^h &= \frac{\hat{p}^h_{ti+1} - \hat{p}^h_{ti}}{\sqrt{\hat{p}^h_{ti} \Delta}} \\
\eta_i^h &= -\sqrt{\frac{\Delta}{\hat{p}^h_{ti}}} \\
\psi_i^h &= \sqrt{\hat{p}^h_{ti} \Delta}
\end{align*} \tag{4.14}
\]

for every \( i = 1, \ldots, n - 1 \) and \( h = 1, \ldots, M \). We reduce the number of parameters from \( 2M + M^2 \) to \( 2M + [M(M + 1)/2] \), by introducing the following new variables:

\[
\Gamma = C^\top C, \quad \alpha_h = k_h \mu_h, \quad h = 1, \ldots, M, \tag{4.15}
\]

where \( C \) is a lower triangular matrix with strictly positive entries on the main diagonal\(^\dagger\). We associate to this matrix a vector \( c \in \mathbb{R}^{M(M + 1)/2} \) according to the relation

\[
C_{i,j} = c_{(i-1)^2+j} \quad \forall \quad M \geq i \geq j \geq 1. \tag{4.16}
\]

\(^\dagger\)Using the Cholesky factorization of \( \Gamma^{-1} = A^\top A \), with \( A \) upper triangular, hence \( C = [A^{-1}]^\top \).
Using these variables, we have \( \det \Gamma = [\det C]^2 = \prod_{h=1}^M C_{hh}^2 \), while (4.12) can be rewritten as follows

\[
\ell_n(c, \alpha, k) = \zeta + (n - 1) \sum_{h=1}^M \log(C_{hh}) - \frac{1}{2} \sum_{i=1}^{n-1} |C_{ii}|^2.
\] (4.17)

Therefore the calibration of model (4.9) can be obtained by computing the maximum of \( \ell_n \) for \( \alpha, k \in (0, \infty)^M \) and \( c \in \mathbb{R}^{M(M+1)/2} \), satisfying \( c_{h(M+1)/2} > 0 \), for any \( h = 1, \ldots, M \). We observe that the computation can be reduced to the maximum of \( U(\alpha, k) = \sup_c \ell_n(c, \alpha, k) \), (4.18)
on \((0, \infty)^M \times (0, \infty)^M \). Since \( \ell_n \) is concave compared to \( c \), it is easy to show that \( U \) is well defined and, for every \( (\alpha, k) \in (0, \infty)^{2M} \), there is a \( c^* = c^*(\alpha, k) \) such that

\[
U(\alpha, k) = \ell_n(c^*, \alpha, k)
\] (4.19)
holds true. The optimizer \( c^* \) is related, via relation (4.16), to the lower triangular matrix \( (R^*_{ij}/\sqrt{R^*_{ii}})_{i \geq j} \) defined by

\[
\begin{pmatrix}
R^*_{i1} \\
R^*_{i2} \\
\vdots \\
R^*_{ii}
\end{pmatrix} = V_i^{-1}
\begin{pmatrix}
0 \\
0 \\
\vdots \\
n - 1
\end{pmatrix} \quad i = 1, \ldots, M,
\] (4.20)
where

\[
(V_i)_{h_1 h_2} = (V_i(\alpha, k))_{h_1 h_2} := [e^{h_1} \, e^{h_2}] \quad h_1, h_2 = 1, \ldots, i.
\] (4.21)

For the details of this result, we refer the reader to 15.

Summarizing, it is possible to pick up an optimal value of \( C = C(\alpha, k) \) and then calculate the maximum of the function \( U \) which depends only on \( 2M \) variables, thus reducing the computational burden.

Paths generation using model (4.9) is carried out by means of the discrete version (4.11). The generation of a scenario with time frequency \( \Delta > 0 \), assigns the last observed data to \( p^h_i \), for every \( h = 1, \ldots, M \). Then at each time step \( i \), where \( 1 \leq i \leq n - 1 \), we generate independent random vectors \( z_i \) from the multivariate normal distribution \( N(0, I_M) \), and we set

\[
p^h_{i+1} = p^h_i + k_h(\mu_h - p^h_i)\Delta + \sqrt{p^h_i \Delta (C^{-1} \cdot z_i)_h}
\] (4.22)
for \( i = 1, \ldots, n - 1 \), where \( k, \mu \) and \( C \) are the MLE estimators.

In Table 2, we show the estimation of model (4.9) using the reduction method described above. Since we do not know whether the function \( U \) is concave or not, gradient methods are no capable of capturing a global maximum. For such reason, we resorted to a stochastic algorithm based on Adaptive Simulated Annealing (ASA) combined with Newton’s method.
Table 2. MLE of the Multivariate Model: The parameters of the multivariate model (4.9) of the orthogonal components (the $p^{h,\perp}_t$ of Eq. (4.4)) of Italian Treasury rates. The estimate is obtained from daily data of the 1999-2003 period by means of the techniques described in section 4.2 using Adaptive Simulated Annealing.

\[
\sigma\sigma^\top = \begin{pmatrix}
0.5927 & 0.3092 & 0.1475 & 0.1022 & 0.0557 & 0.0244 & 0.0130 & 0.0041 \\
0.3092 & 0.5651 & 0.3887 & 0.3120 & 0.2398 & 0.1575 & 0.1342 & 0.1149 \\
0.1475 & 0.3887 & 0.3363 & 0.2824 & 0.2283 & 0.1583 & 0.1383 & 0.1221 \\
0.1022 & 0.3120 & 0.2824 & 0.2552 & 0.2107 & 0.1493 & 0.1319 & 0.1181 \\
0.0557 & 0.2398 & 0.2283 & 0.2107 & 0.1864 & 0.1373 & 0.1235 & 0.1123 \\
0.0244 & 0.1575 & 0.1583 & 0.1493 & 0.1373 & 0.1158 & 0.1066 & 0.0997 \\
0.0130 & 0.1342 & 0.1383 & 0.1319 & 0.1235 & 0.1066 & 0.1014 & 0.0959 \\
0.0041 & 0.1149 & 0.1221 & 0.1181 & 0.1123 & 0.0997 & 0.0959 & 0.0952 \\
\end{pmatrix}
\]

Figure 9. Simulation of the term structure evolution based on the multivariate CIR model (4.9).

Figure 9 shows the results of a simulation, based on the extension of the CIR model described in this section, of the term structure evolution for a planning period of 48 months.

5. The validation of the simulated scenarios

The stochastic models presented above allow the generation of a “realistic” future...
term structure. However, we need to generate a (pretty long) temporal sequence of term structures starting from the present interest rates curve. As we mentioned in Section 4, this requires to control the evolution in time of the simulated term structure in such a way that, for instance, its behaviour is not too different from the behaviour observed in the past.

First of all we impose lower and upper bounds to the simulated ECB official rate to prevent it from reaching unrealistic values. Besides that, we resort to two techniques to assess the reliability of the sequence of simulated term structures.

The first method, that we classify as “local”, assures that, at each time step of the planning period, the simulated term structure is “compatible” with the historical term structure. The second one (that we classify as “global”), considers the whole term structure evolution, and controls that the correlation among the increments of the $p^{h,1}_t$ be close to the correlation of the increments of the historical fluctuations.

The local test is based upon the observation that the shape of the term structure does not vary wildly over time. Obviously, there is some degree of variability. For instance, it is known that the term structure is usually an increasing function of the maturity, but, at times, it can be inverted for a few maturities. This is the fundamental assumption of the Nelson and Siegel parsimonious model of the term structure. In the present case, since we are not interested in a functional representation of the term structure but in the interest rates at fixed maturities, we take as indicator of the term structure shape the relative increment and the local convexity of the interest rates:

$$d^h_t = r^h_t - r^{h-1}_t, \quad c^h_t = \tilde{r}^h_t - r^h_t$$

where $\tilde{r}^h_t$ is the value at $m^h$ of the linear interpolation between $r^{h+1}_t$ and $r^{h-1}_t$:

$$\tilde{r}^h_t = r^{h-1}_t + \frac{r^{h+1}_t - r^{h-1}_t}{m^{h+1} - m^{h-1}}(m^h - m^{h-1}).$$

A positive convexity ($c^h_t > 0$) means that the rate $r^h_t$ at maturity $m^h$ is below the line joining the rates $r^{h-1}_t$ and $r^{h+1}_t$ at maturities $m^{h-1}$ and $m^h$, respectively, therefore the curvature opens upward. The case of $c^h_t < 0$ can be interpreted along the same line but the curvature opens downward.

At any time step, we accept the simulated interest rates if the corresponding values of $d^h_t$ and $c^h_t$ are not too different from the historical values. Briefly, we compute the historical mean and standard deviation ($\mu^h_d, \sigma^h_d$) and ($\mu^h_c, \sigma^h_c$) of $d^h_t$ and $c^h_t$, respectively, then we check that

$$\sum_{h=1}^{M-1} \frac{1}{(\sigma^h_d)^2} (d^h_t - \mu^h_d)^2 \leq (M - 1)$$

$$\sum_{h=2}^{M-1} \frac{1}{(\sigma^h_c)^2} (c^h_t - \mu^h_c)^2 \leq (M - 2).$$

The meaning of such tests is straightforward. If the simulated interest rates pass the tests, we are quite confident that their increments and the local convexity have a statistical behaviour similar to the historical rates.
The global test controls the correlation among the increments of the $p_t^{h,\perp}$, since it can be quite different from the correlation of the historical data. In order to avoid “pathological” situations like anti-correlated increments or increments too much correlated each other, we compute the “one-norm” of the matrix difference between the correlation matrix of the increments of the historical fluctuations and the correlation matrix of the simulated increments. The “one-norm” is defined as follows: for each column, the sum of the absolute values of the elements in the different rows of that column is calculated. The “one-norm” is the maximum of these sums:

$$||\Sigma_{i,j}^{\text{hist}} - \Sigma_{i,j}||_1 = \sup \sum_{j=1}^{M} |\Sigma_{i,j}^{\text{hist}} - \Sigma_{i,j}|.$$ (5.25)

We compare the result of this calculation with a predefined acceptance threshold. If the “one-norm” is below the threshold, the simulated scenario is accepted, otherwise it is rejected and a new set of fluctuations is generated. Currently the threshold is set equal to 0.05. Since the correlation is a number within the range $[-1,1]$, this simple mechanism guarantees that the covariance among the increments of the interest rates of different maturities in the simulated scenarios is pretty close to the historical covariance.

Note how the local and global tests are complementary since the first test involves the shape of the term structure, whereas the second one concerns how the rates’ increments are correlated in time, that is, how the term structure evolves from a time step of the simulation to the next.

6. Conclusions and future perspectives

The management of Public Debt is of paramount importance for any country. Such issue is specially important for European countries after the definition of compulsory rules by the Maastricht Treaty. Together with the Italian Ministry of Economy and Finance, we study the problem of finding an optimal strategy for the issuance of public debt securities. This turned out to be a stochastic optimal control problem where the evolution of the interest rates plays a crucial role.

We presented the techniques we employ to simulate the future behaviour of the interest rates for a wide range of maturities (from 3 months up to 30 years). Since the planning period that we consider is pretty long (up to five years), most existing models need to be modified in order to provide realistic scenarios. In particular the evolution of the “leading” rate (that we assume to be the European Central Bank’s official rate) is simulated as if it were a stochastic jump process.

The models proposed in the present paper are fully integrated in the software prototype in use at the Ministry. The time required by the simulations (few seconds on a personal computer) is such that we can afford on-line generation of the simulated scenarios even if the validation procedures described in Section 5 may require multiple executions before a scenario is accepted.

Open problems and future analysis directions include
• Modelling of Primary Budget Surplus, Gross Domestic Product and Inflation in
order to implement the Taylor’s rule and possibly other models of the European
Central Bank’s official rate evolution.

• The overcoming of the assumption that interest rates are independent of the
portfolio of existing securities and independent of the new securities issued every
month. To limit the impact on the optimization problem, we should devise a
description of these interactions that is compatible with the linear formulation of
the problem.

7. Acknowledgments

We thank A. Amadori, D. Guerri and B. Piccoli for useful discussions.


We deal with the case of a diffusion process \( X_t \) (i.e., the interest rate \( r_t \)) that is
observed at discrete times \( 0 = t_0 < t_1 < \cdots < t_n \), not necessarily equally spaced.
If the transition density of \( X_t \) from \( y \) at time \( s \) to \( x \) at time \( t \) is \( p(x,t,y,s;\theta) \),
where \( \theta \) is a collection of parameters to be estimated, we can resort to the Maximum
Likelihood Estimator (MLE) \( \hat{\theta}_n \) which maximizes the likelihood function

\[
L_n(\theta) = \prod_{i=1}^{n} p(X_{t_i}, t_i, X_{t_{i-1}}, t_{i-1}; \theta) \tag{8.26}
\]

or equivalently the log-likelihood function

\[
\ell_n(\theta) = \log(L_n(\theta)) = \sum_{i=1}^{n} \log(p(X_{t_i}, t_i, X_{t_{i-1}}, t_{i-1}; \theta)) \tag{8.27}
\]

In case of observations equally spaced in time, the consistency and asymptotic normality
of \( \hat{\theta}_n \) as \( n \to \infty \) it can be proved (see 7). In general, the transition density of \( X_t \) is
not available. In this case the classical alternative estimator is obtained by means of
an approximation of the log-likelihood function for \( \theta \) based on continuous observations
of \( X_t \). Unfortunately, this approach has the undesirable property that the estimators
are biased, unless \( \pi = \max_i |t_i - t_{i-1}| \) is small. To overcome the difficulties due to the
dependence of the parameters on \( \pi \), different solutions have been proposed. One of the
most efficient methods resorts to martingale estimating functions 5. This method is
based on the construction of some estimating equations having the following form:

\[
G_n(\theta) = 0 \tag{8.28}
\]

where

\[
G_n(\theta) = \sum_{i=1}^{n} g_{i-1}(X_{t_{i-1}}; \theta)(X_{t_i} - \mathbb{E}(X_{t_i}|X_{t_{i-1}})), \tag{8.29}
\]

\( g_{i-1} \) being continuously differentiable in \( \theta \), for \( i = 1, \ldots, n \). In 5 it is proved, under
technical conditions on \( g_{i-1} \), that an estimator that solves the equation (8.28) exists
with probability tending to one as \( n \to \infty \), and this estimator is consistent and asymptotically normal.

A more simple approach, which can be used when \( \pi \) is sufficiently small, is based on the Euler discretization of the diffusion equation associated to \( X_t \). In this case, one can use the log-likelihood approach since the transition density can be easily computed. We discuss the application of this method to (4.8). Let

\[
    r_{t_{i+1}} = r_{t_i} + k(\mu - r_{t_i})\Delta + \sigma \sqrt{r_{t_i}\Delta} \ Z_i, \quad i = 1, \ldots, n - 1,
\]

be the first order discretization of the equation (4.8) on a time interval \([0, T]\), where \( \Delta = T/n \) and with \( Z_i \) the increment \( \Delta B_i \) of the Brownian motion between \( t_i = i\Delta \) and \( t_{i+1} = (i+1)\Delta \). Since these increments are independent \( N(0, \Delta) \) distributed random variables, the transition density from \( y > 0 \) to \( x \) during an interval of length \( \Delta \) is

\[
    p(x|y; \theta) = \frac{1}{\sqrt{2\pi y\sigma^2\Delta}} \exp \left( -\frac{1}{2y\sigma^2\Delta} (x - k(\mu - y)\Delta)^2 \right),
\]

where \( \theta = (k, \mu, \sigma) \). Therefore, given a sequence of observations \( \{r_i\}_i \), an estimator of (8.30) is obtained by maximizing the log-likelihood function:

\[
    \max_\theta \sum_{i=1}^{n-1} \log(p(r_{t_{i+1}}|r_i; \theta)).
\]

If one expands the function in (8.32) and sets equal to zero its partial derivatives with respect to \( (k, \mu, \sigma^2) \), it becomes easy to show that the corresponding equations admit a unique solution that is a maximum point of the log-likelihood function. The following relations represent this ML estimator:

\[
    \mu = \frac{EF - 2CD}{2EB - FD}, \quad k = \frac{D}{2\mu B - F},
\]

\[
    \sigma^2 = \frac{1}{n-1} \left( A + k^2 \mu^2 B + k^2 C - k\mu D + k E - k^2 \mu F \right)
\]

where

\[
    A = \sum_{i=0}^{n-1} \frac{(r_{t_{i+1}} - r_i)^2}{r_i \Delta}, \quad B = \Delta \sum_{i=0}^{n-1} \frac{1}{r_i}, \quad C = \Delta \sum_{i=0}^{n-1} r_i
\]

\[
    D = 2 \sum_{i=0}^{n-1} \frac{(r_{t_{i+1}} - r_i)^2}{r_i}, \quad E = 2(r_n - r_0), \quad F = 2\Delta(n - 1).
\]

Here \( r_i \) stands for the interest rate at time \( t_i \). We report the results of this method in Table 3. In Table 4, the method is applied to each orthogonal component \( p_t^{h,\perp} \).
Scenario-Generation Methods for an Optimal Public Debt Strategy

<table>
<thead>
<tr>
<th>Statistics (N=500)</th>
<th>Parameter</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$k$</td>
</tr>
<tr>
<td>True parameter</td>
<td>0.8</td>
</tr>
<tr>
<td>Estimated parameter</td>
<td>0.79928</td>
</tr>
<tr>
<td>Standard error</td>
<td>0.018222</td>
</tr>
<tr>
<td>$T$ Statistics</td>
<td>-0.88811</td>
</tr>
<tr>
<td>$H_0$</td>
<td>0</td>
</tr>
</tbody>
</table>

**Table 3. Assessment of the Discrete Estimator.** Results for a simulation evaluation of the log-likelihood estimator to the univariate CIR model. Using the true values of the parameters we simulated 500 sample paths of length 2610 daily observations each. For each sample path we undertook discrete MLE estimation via Euler discretization. The table presents summary statistics of the simulated estimations. We computed the mean and standard error for the estimator ($\hat{k}$, $\hat{\mu}$, $\hat{\sigma}$), computed t-statistics for the difference in the simulated parameter estimate versus the true parameter and t-test ($H = 0$ do not reject the null hypothesis at significance level 95%).

<table>
<thead>
<tr>
<th>Parameter</th>
<th>BOT 3</th>
<th>BOT 6</th>
<th>BOT 12</th>
<th>CTZ</th>
</tr>
</thead>
<tbody>
<tr>
<td>$k$</td>
<td>5.4611</td>
<td>4.0217</td>
<td>2.1634</td>
<td>1.1265</td>
</tr>
<tr>
<td>$\mu$</td>
<td>0.32712</td>
<td>0.43815</td>
<td>0.72852</td>
<td>1.457</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>0.78513</td>
<td>0.76817</td>
<td>0.74801</td>
<td>0.57573</td>
</tr>
<tr>
<td>$\log(L)$</td>
<td>1866.541</td>
<td>2174.4863</td>
<td>1920.0013</td>
<td>1791.6797</td>
</tr>
<tr>
<td>$</td>
<td>\nabla \log(L)</td>
<td>$</td>
<td>$1.2268 \times 10^{-13}$</td>
<td>$3.0103 \times 10^{-15}$</td>
</tr>
<tr>
<td>$\lambda_L$</td>
<td>-0.54044</td>
<td>-1.1819</td>
<td>-1.5844</td>
<td>-1.7826</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Parameter</th>
<th>BTP 36</th>
<th>BTP 60</th>
<th>BTP 120</th>
<th>BTP 180</th>
<th>BTP 360</th>
</tr>
</thead>
<tbody>
<tr>
<td>$k$</td>
<td>1.1488</td>
<td>1.4039</td>
<td>1.7635</td>
<td>1.9973</td>
<td>2.4133</td>
</tr>
<tr>
<td>$\mu$</td>
<td>2.0531</td>
<td>2.8686</td>
<td>3.9144</td>
<td>4.2957</td>
<td>4.5355</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>0.50163</td>
<td>0.4293</td>
<td>0.33892</td>
<td>0.31716</td>
<td>0.30738</td>
</tr>
<tr>
<td>$\log(L)$</td>
<td>1756.6086</td>
<td>1743.7829</td>
<td>1822.044</td>
<td>1839.9547</td>
<td>1840.362</td>
</tr>
<tr>
<td>$</td>
<td>\nabla \log(L)</td>
<td>$</td>
<td>$1.4118 \times 10^{-14}$</td>
<td>0</td>
<td>$9.0949 \times 10^{-13}$</td>
</tr>
<tr>
<td>$\lambda_L$</td>
<td>-1.8906</td>
<td>-1.8071</td>
<td>-1.5539</td>
<td>-1.3116</td>
<td>-0.93247</td>
</tr>
</tbody>
</table>

**Table 4. MLE of the Univariate CIR Model.** Results for the estimation of the univariate CIR model of the $p_t^{h-L}$ of Eq. 4.4. Estimation is carried out using discrete maximum-likelihood (8.32). We report the value of the log-likelihood ($\log(L)$), the norm of its gradient ($|\nabla \log(L)|$) and the maximum eigenvalue ($\lambda_L$) of its Hessian matrix at the maximum point.


